

**COMMON EUROPEAN FRAMEWORK
OF REFERENCE FOR LANGUAGES:
LEARNING, TEACHING, ASSESSMENT**

**PHONOLOGICAL SCALE REVISION
PROCESS REPORT**

Language Policy

Enrica Piccardo

COMMON EUROPEAN FRAMEWORK
OF REFERENCE FOR LANGUAGES:
LEARNING, TEACHING, ASSESSMENT

**PHONOLOGICAL SCALE REVISION
PROCESS REPORT**

Enrica Piccardo

2016

Language Policy Programme
Education Policy Division
Education Department
Council of Europe

*The opinions expressed in this work
are the responsibility of the authors
and do not necessarily reflect the
official policy of the Council of
Europe.*

All requests concerning the
reproduction or translation of all or
part of this document should be
addressed to the Directorate of
Communication (F-67075
Strasbourg Cedex or
publishing@coe.int). All other
correspondence concerning this
document should be addressed to
DG II – Education Department
(language.policy@coe.int)

The reproduction of extracts is
authorised, except for commercial
purposes, on condition that the
source is quoted [© Council of
Europe – Language Policy]

This document is available for
download on the following website:
www.coe.int/lang-cefr

CONTENTS

INTRODUCTION	6
PHASE ONE	7
<u>Analysis of existing concept and scales (CEFR 2001, provisional version of revised descriptors)</u>	<u>7</u>
<u>Identification of weaknesses (also based on existing criticism)</u>	<u>8</u>
<u>Rationale for revising the existing phonology scale/creation of new scales</u>	<u>9</u>
PHASE TWO	10
<u>Literature review</u>	<u>10</u>
<u>Analysis of resources (scales) and identification of key concepts</u>	<u>15</u>
<u>Identification of key concepts</u>	<u>16</u>
<u>Operationalisation of the concepts in order to create descriptors</u>	<u>16</u>
PHASE THREE	17
<u>Creation of three scales (first draft)</u>	<u>17</u>
<u>Revision of drafts and preparation of the workshop</u>	<u>17</u>
<u>Workshop (preliminary internal validation)</u>	<u>17</u>
<u>Revision of the scales based on the feedback and analysis of results</u>	<u>18</u>
<u>Consultation with experts on the revised scales for feedback</u>	<u>18</u>
<u>Preparation for validation</u>	<u>18</u>
PHASE FOUR	19
<u>Qualitative validation</u>	<u>19</u>
<u>Quantitative validation</u>	<u>19</u>
REFERENCES	24
APPENDICES	28
<u>Appendix 1 – Phase 1: Rating descriptors for quality – and judging category and level</u>	<u>28</u>
<u>Appendix 3 – Video performances for assessment in the Phase 3 assessment task</u>	<u>36</u>
<u>Appendix 4 – Phase 3: Using descriptors for assessment: Bookmark method with core group</u>	<u>37</u>
<u>Appendix 5 – Item analysis: All information on items</u>	<u>39</u>
<u>Appendix 6 – Descriptor scales – after analysis and discussion</u>	<u>42</u>

Introduction

The paper reports on a sub-project of a Council of Europe initiative to provide an extended version of the CEFR illustrative descriptors, one of the main aims of which was to provide descriptors for communicative activities and strategies related to mediation. Once underway, however, the main Council of Europe project quickly moved beyond its initial scope. Not only did the conceptualization of mediation prove to be a powerful endeavour capable of encapsulating a new vision of the role and function of the language learning, teaching and assessment process, but, beyond this, it revealed grey areas in the CEFR that needed to be addressed in order to respond to the developments in pedagogical and societal visions occurring since the CEFR publication. This work allowed us to address underdeveloped or underconceptualized aspects of language teaching pedagogy and to take into consideration new perspectives that have been emerging in language education. Some of the areas covered are computer-mediated communication, plurilingual and inter/pluricultural interaction, a more agency-oriented relationship to literature and the strategic dimension of the mediation process itself.

Needless to say, one of these grey areas is phonology, an aspect of language pedagogy that has been under-researched by scholars working in second/foreign language education. In fact, while the articulatory phenomena and the difficulties related to phonological features of speakers belonging to different linguistic communities have been investigated extensively by phoneticians and linguists, still little has been published by applied linguists when it comes to core principles of the teaching of pronunciation in different languages, and consequently its assessment. At the same time, a new sensibility has been emerging in the applied linguists' scholarly community when it comes to re-evaluating the traditional idea of the 'native speaker' as a model or perception of the norm in pronunciation. This is especially visible in English considering the movement towards 'global Englishes' or 'English as a Lingua Franca', but similar considerations have been applied to all languages.

The CEFR presents a solid conceptualization of phonological competence, albeit necessarily adapted to an audience of language educators not of phoneticians, yet this conceptualization is arguably not translated into appropriate descriptors in the single existing scale for Phonological control. Needless to say, this scale does not incorporate the new vision of pronunciation and phonological models referred to above. The present work on the phonological scale of the CEFR aims to address some of these issues and to provide realistic scales and concrete descriptors to support practitioners and learners alike, in the delicate and crucial process of acquiring an appropriate and effective pronunciation of the target language.

The work conducted on producing descriptors for phonological competence has proceeded in a similar way as that focusing on the other areas of the CEFR where new descriptors needed to be produced. It has included alternate phases of documentary research, conceptualization, drafting, sharing and revision of descriptors. A detailed description of each phase follows.

Phase One

Analysis of existing concept and scales (CEFR 2001, provisional version of revised descriptors)

Naturally the initial phase of the work included a thorough analysis of what exists in the CEFR related to phonological competence, and an evaluation of what had already been addressed in the first draft of the revision of the existing set of CEFR illustrative descriptors.

The analysis of the CEFR produced the following conclusions:

- 1) Phonological competence takes an important role in the descriptive scheme of the CEFR even though this does not translate into an extended and accurate series of scales and descriptors. In particular, it is included in
 - a. The description of **communicative competence** (Section 2.1.2, p. 22):

“Linguistic competences include lexical, phonological, syntactical knowledge and skills and other dimensions of language as system, independently of the sociolinguistic value of its variations and the pragmatic functions of its realisations.”
 - b. The description of the **communicative language process** (Section 4.5) when it comes to execution (Section 4.5.2) and especially Production (4.5.2.1):

“The production process involves two components: The formulation component takes the output from the planning component and assembles it into linguistic form. This involves lexical, grammatical, phonological (and in the case of writing, orthographic) processes which are distinguishable and appear (e.g. in cases of dysphasia) to have some degree of independence but whose exact interrelation is not fully understood. The articulation component organises the motor innervation of the vocal apparatus to convert the output of the phonological processes into co-ordinated movements of the speech organs to produce a train of speech waves constituting the spoken utterance, or alternatively the motor innervation of the musculature of the hand to produce hand- written or typewritten text.”
 - c. The description of **ability to learn** (*savoir apprendre*) (Section 4.1.4), in particular the General phonetic awareness and skills (Section 5.1.4.2) which are described as follows:

“Many learners, particularly mature students, will find their ability to pronounce new languages facilitated by:

 - ▶ an ability to distinguish and produce unfamiliar sounds and prosodic patterns;
 - ▶ an ability to perceive and catenate unfamiliar sound sequences;
 - ▶ an ability, as a listener, to resolve (i.e. divide into distinct and significant parts) a continuous stream of sound into a meaningful structured string of phonological elements;
 - ▶ an understanding/mastery of the processes of sound perception and production applicable to new language learning. These general phonetic skills are distinct from the ability to pronounce a particular language.”
 - d. The description of **communicative language competences** (Section 5.2), under linguistic competences (Section 5.2.1) where it is detailed as follows:

5.2.1.4 Phonological competence involves a knowledge of, and skill in the perception and production of:

- ▶ the sound-units (*phonemes*) of the language and their realisation in particular contexts (*allophones*);
- ▶ the phonetic features which distinguish phonemes (*distinctive features*, e.g. voicing, rounding, nasality, plosion);
- ▶ the phonetic composition of words (*syllable structure*, the sequence of phonemes, word stress, word tones);
- ▶ sentence phonetics (*prosody*)
- ▶ sentence stress and rhythm
- ▶ intonation;
- ▶ phonetic reduction
- ▶ vowel reduction
- ▶ strong and weak forms
- ▶ assimilation
- ▶ elision.

On top of this it is also mentioned when paralinguistic features are discussed (Section 4.4.5.2) to mark how this area is distinguished from the proper phonological system, and it is listed among the 12 qualitative categories relevant to oral assessment (Section 9.4, p. 193).

Finally, it is mentioned in the appendixes in the form of 'phonological control' as a component of the illustrative scales of different examinations/assessment (Cambridge Certificate in Advanced English, p. 203, Eurocentres – Small Group Interaction Assessment, p. 204, and in the CEFR itself under linguistic – control, p. 232)

At the level of scales, only the phonological control scale exists in the CEFR 2001 (p.117), which limits itself to providing one descriptor per level from A1 to C1.

In the draft of the update of the 2001 scales¹, this scale had been revised to include a new descriptor at C2, two new descriptors at B2, a descriptor at B2+ (previously at B2), a revised descriptor at B1.

Identification of weaknesses (also based on existing criticism)

From the analysis, strengths and weaknesses clearly appear:

Strengths:

- ▶ The construct of the CEFR in relation to phonology is thorough and sufficiently broad to allow a revision and extension of the scales/descriptors in order to capture the new developments of reflection in second/foreign language education.
- ▶ The link with the descriptive scheme is clear and extensive.
- ▶ The pedagogical aspect is present, as learnability (and consequently “teachability”) of phonology is envisaged.
- ▶ The integration of phonology among relevant categories of assessment is clearly mentioned.

¹ This draft, which included all the modifications/revisions done to the main set of descriptors included in the CEFR (2001) was presented in June 2014 at a consultant meeting in Strasbourg.

Weaknesses:

- ▶ The existing scale *Phonological Control* does not capture this conceptual apparatus, and appears fully unrealistic when it comes to issues such as accent, or progression (particularly in moving from B1 to B2).
- ▶ Furthermore, it is not consistent as it mixes such diverse factors as stress/intonation, pronunciation, accent and intelligibility without providing clear indication of progression in any of these factors specifically.
- ▶ It is not complete which results in jeopardizing its applicability and usefulness.

The analysis of the proposed revision proved that none of these weaknesses had been addressed. Even though this initial revision represented an improvement in relation to the previous scale, it still did not really integrate the conceptual apparatus of the CEFR or part thereof.

Rationale for revising the existing phonology scale/creation of new scales

From this analysis a rationale emerges which can guide the work on the phonological aspects of the CEFR in view of producing useful and realistic descriptors.

The rationale proposed here includes the need to:

- ▶ identify explicit categories that inform the scales based (i) on the CEFR construct and (ii) on recent development in relation to pronunciation use and teaching;
- ▶ think through the internal progression of phonological competence as well as its delicate relation to progression in language proficiency;
- ▶ provide both a general scale and specific ones in order for teachers/learners to both have a snapshot of their phonological competence and to identify areas for improvement.

Phase Two

Literature review

Systematic documentary research was conducted with the following aims:

- ▶ to collect further sources of information beyond the CEFR-related discussion of phonological competence;
- ▶ to explore if and how phonology was dealt with, from a pedagogical point of view, in relation to second/foreign language education; and
- ▶ to acquire possible resources to inform our future work.

When we started this work, very little had been published, which focused specifically on the CEFR phonology scale (Cauvin, 2012; Horner, 2010, 2013, 2014), although that scale was critically mentioned in some studies (Galaczi, Post, Li, & Graham, 2011; Isaacs & Trofimovich, 2012; Harding, 2013). Among the very few studies, particularly enlightening for the present work were the contributions of Horner which discuss the relationship between pronunciation and proficiency and point to some fundamental issues. By examining which pronunciation features have an impact on communicative efficiency, Horner questions the possibility of establishing hierarchies and analyses the aspects of pronunciation that are more or less important, “irritating” (and for whom), difficult to acquire (and in what order), and having an impact on intelligibility. Particularly on this last point –intelligibility - he identifies factors, such as word stress, accurate reproduction of phonemes and sentence stress, as core features, while he considers intonation, rhythm, and phonetic reduction as more peripheral. Although he refers to English language phonology in his articles, these considerations can apply across languages, albeit with individual variations. Even more of a cross-linguistic nature are the considerations of the role and attitude of the listener/interlocutor, and his/her familiarity with the speech community when it comes to phonology.

Even though Horner’s proposed solutions do not really solve the problem or effectively address his concerns, principally because he adopts a modulation based on descriptive adverbs, which has been proved ineffective and vague by the literature (Alderson 1991, North, 2000), his analysis is nevertheless insightful and points at precise issues that can be addressed effectively. Another important study strengthening the need for a reconsideration of the CEFR phonological scale is Trofimovich & Isaacs, where the authors underline how the CEFR combination of “descriptions of easily understandable speech and a noticeable foreign accent in the same band descriptor” (2012: 914) can be problematic and stress the need to “disentangle accent from different aspects of communicative effectiveness, including comprehensibility.” (2012: 914). Finally, the TOEFL iBT Research Report (Jamieson & Poonpon, 2013) also contributed to support our view that we need to consider the different dimensions of phonology in view of developing targeted effective descriptors, both global and analytic. More recently, we were able to access a publication by Harding, still in press (expected for 2016), which criticises directly the CEFR scale for phonological control. In particular the author addresses the issue of the usability of the scale. He used mixed-methods research to investigate the construct underlying the scale, particularly its orientation towards a nativeness principle or a comprehensibility principle respectively. Also another article by Frost and O’ Donnell (in press) concerns the empirical development of an analytic grid for assessing the phonological features of French learners of English inspired by the CEFR existing scale.

In addition to this work related specifically to the CEFR, a more general literature search was conducted using different databases.²

² Three databases were consulted: The Teacher Reference Centre (teacherreference.com), a resource of peer-reviewed journals covering resources on language teaching related issues (including assessment, best practices, continuing education, curriculum development and instructional media), the MLA International Bibliography – Modern Language Association (www.mla.org) and Education Resources Information Center (ERIC) (https://eric.ed.gov/). On top of this a research using Google scholar was also conducted.

Of the articles found, the vast majority did not focus on the teaching/learning component of phonology but rather investigated such topics as patterns and issues related to phonology typical of specific linguistic communities, possible avenues for understanding specific phonological problems, speech disorders in children and adults. Much less numerous were the articles focusing specifically on the teaching/learning/assessing of pronunciation. However, a growing body of research in the last few years was observed which investigates pedagogical aspects related to pronunciation (Saito, 2012; Lee, Jang & Plonsky, 2015; Thomson & Derwing, 2014). **Specific attention to the pedagogical dimension of phonology is a recent phenomenon and so is a specific focus on assessment, which appears to be even less frequent.** Furthermore, the articles dealing with the pedagogical dimension of phonology covered a wide range of studies in terms of literature reviews which related of trends in the field, pointed at lacks and limitations, and strongly advocated the need for more empirically targeted research. Above all, they pointed at the need for a synergy between researchers and practitioners' work in the field. It became clear that even if there has been a growing body of research on the pedagogic aspects over the past few years, research and reflection on the role of phonology in second/foreign language education is still limited, and so are resources supporting teachers who intend to teach pronunciation in their classes. Particularly scarce are studies investigating the criteria that can be employed to realistically assess phonological competence, which could potentially constitute the backbone of effective phonological assessment scales.

In a seminal article published in *TESOL Quarterly*, Derwing and Munro described the situation as follows: "The study of pronunciation has been marginalized within the field of applied linguistics [and that] as a result; teachers are often left to rely on their own intuitions with little direction." (2005: 379). Even though "[a]n extensive, growing literature on L2 speech has been published in journals that focus on speech production and perception, for example, *Journal of the Acoustical Society of America*, *Journal of Phonetics*, and *Language and Speech* [...] this work is rarely cited or interpreted in teacher-oriented publications." (2005: 382). The research tradition in phonology has often been associated with the activity of specialized laboratories published in highly technical journals and regarded as only peripherally relevant in applied linguistics (Derwing & Munro, 2015). On the whole, as Murphy and Baker (2015) put it, it was only during the 90s that pronunciation instruction started to be seriously researched in an empirical manner.

In their recent publication, Thomson and Derwing acknowledged a shift in the field since 2005: "Many more L2 pronunciation studies have appeared in peer-reviewed venues, conference proceedings, and graduate theses, laying the foundation for increasingly rigorous research. Further cementing L2 pronunciation as an area worthy of investigation is the emergence of annual conferences (e.g. *Pronunciation in Second Language Learning and Teaching*) and the new *Journal of Second Language Pronunciation*³ (2014: 1-2). Moreover, the research of Cauldwell (2011) and Levis (2005) calls for teaching pronunciation based on spontaneous speech and not isolated words and sentences. In fact, learners need "first to hear and understand authentic, fast, spontaneous speech and, second, to speak rapidly with both fluency and accuracy" (Levis 2005: 559). This points to more attention to the connections between listening comprehension, chunking down, and sound recognition as preparation for pronunciation teaching.

However, the marginalization of pronunciation teaching, which Derwing and Munro pointed out in their 2005 article, as having potentially serious consequences, continues in spite of the increased interest in pronunciation among educators. In a very recent publication (2015) the same authors relate of a wide series of studies that revealed that "teachers are hesitant about systematically teaching pronunciation" (p. 78), that they feel a "need for access to more professional development" (p.80) and that "[t]he curricula in the various programs in which the teachers worked did not focus on pronunciation" (ibid.) with vague and unhelpful indications if any. Teachers do not teach pronunciation because they "lack confidence, skills and knowledge" (MacDonald, 2002: 3), they have received little or no specific training (Burgess & Spencer, 2000⁴; Breitzkreutz, Derwing & Rossiter, 2002; Baker, 2011) and therefore they "have no formal preparation to teach pronunciation" (Derwing & Munro, 2005: 389).

³ The first issue of the *Journal of Second Language Pronunciation* was published in 2015.

⁴ Burgess & Spencer, 2000, for instance, speak about the fact that in teacher training programs teachers are being taught "phonology" instead of being offered courses on "how to teach pronunciation" p. 192-194 of their article.

Also, methodology books dedicate minimal or no attention to pronunciation and phonology and/or do not base their indication on research. All of this unfortunately confirms what had already been written at the turn of the century (Pennington, 1998) about the minimal space in programs for language teachers for phonology, which, when offered – if at all – focus on general rather than applied linguistics. In fact, the trend of increased interest in pronunciation only started in the last decades of the 20th century, especially as far as intelligibility is concerned (Pennington & Richards, 1986; Catford, 1987; Crawford, 1987; Morley, 1992; Pennington, 1998). This increased interest has continued ever since (Levis, 2005; Isaacs, 2008; Munro & Derwing, 2011), but unfortunately has not had a serious impact at the level of teacher education, material development and language policy documents.

And yet pronunciation plays *de facto* a primary role in the language class, as accurate pronunciation often ends up by being used as measure of students' progress and indicator of students' proficiency. This reveals a "mismatch between theory-driven research and practice" (Munro & Derwing, 2011: 317). Derwing and Munro (2005) underlie that while some teachers who rely on their intuition in teaching pronunciation may develop some practical and effective skills, this is at best haphazard. Inconsistencies in the teaching of pronunciation (in terms of dedicated time, methodology, use of textbooks and support) have been confirmed by several studies recently conducted in various countries (as reported in Derwing & Munro, 2015). In reality, the two researchers point at a series of risks linked to the lack of teachers' preparation when it comes to teaching pronunciation. This can include issues like students not receiving any instruction at all to being directed to focus on pronunciation-related aspects which do not have a particular influence on intelligibility, like for example the pronunciation(s) of the segment *th* (Derwing & Rossiter, 2002) in English. It can imply that teachers, feeling unprepared in this domain, "rely too heavily on pronunciation textbooks and software without regard for their own students' problems" (Derwing & Munro, 2005: 389). This they see as problematic: both paper and electronic resources often do not have a solid scientific basis, nor they provide a rationale that would help teachers make informed choices and customize the activities to suit their students' needs, besides, the 'one size fits all' is prevailing due to market pressure. In particular, the worry that researchers expressed at the perceived idea that "technology is a panacea for correcting pronunciation" (Derwing & Munro, 2005: 390) has been confirmed after analysing recent research and dedicated software. Although advances are being made in computer-assisted pronunciation training (CAPT) with some promising automatic speech recognition (ASR) software (Thomson, 2012a, 2012b; Cucchiarini, Neri & Strik, 2009; Chun, Hardison & Pennington, 2008), 'off-the-shelf ASR technology is of little use in identifying production problems because it is unable to distinguish accent features that do and do not affect intelligibility" (Derwing & Munro, 2015: 130). With very few exceptions (Thomson, 2012b; Germain & Martin, 2000) software does not provide individualized feedback, nor does it adjust to students' speech as human beings would do, this results in artificial interactions that can reveal themselves useless or even counterproductive in real-life situations. Once again, the problem is the human being behind the machine, and teachers should treat technology as one of the various tools they can use to improve their students' pronunciation. However, for doing this effectively they must have the necessary knowledge of the foundations of pronunciation research on top of the pedagogical skills to be able to use the resources. Unfortunately, as Derwing & Munro (2015 pp. 80–81) attest, this is still far from being the case.

All in all, based on the increased interest in teaching of phonology and pronunciation among professionals, Derwing & Munro (2005; 2015) continue to strongly advocate solid programmatic research in applied linguistics. In doing so they provide some important orientation, such **as attention to intelligibility "to establish the most effective ways of assessing it and to identify the factors that contribute to it"** (2005 p.391). This is reinforced in their recent book in which they indicate two avenues for research: "greater focus on the factors that contribute to intelligibility and comprehensibility", and secondly "more instructional research on pronunciation to help us identify efficient and effective teaching strategies and techniques" (2015: 168). Also, they list a series of pronunciation-related phenomena that have been identified as crucial to the success of any communicative activity by various researchers (Gass & Varonis, 1984; MacIntyre, Baker, Clément, & Donovan, 2003; Rubin, 1992), namely "listener factors, such as **familiarity with foreign-accented speech, willingness to communicate, and attitudes toward L2 speakers** [which] require extensive further research" (Derwing & Munro, 2005: 392). Finally, "little research has explored which linguistic features of speech are most crucial for intelligibility and which – while noticeable or irritating – merely contribute to the perception of an accent" (Trofimovich & Isaacs, 2012: 906), which reinforces the need for moving away from the native speaker model when it comes to pronunciation assessment, "given the well-established fact that **accent is partially independent of comprehensibility and intelligibility**

and that the latter two are more important to successful communication” (Derwing & Munro 2015: 168).

Some very recent studies are starting to address the issue of pronunciation assessment. In particular some studies (Isaacs, 2008; Trofimovich & Isaacs, 2012) focus on the definition of assessment criteria for pronunciation proficiency, which highlight the important notion of intelligibility, and to the need to better investigate listeners’ perceptions. They attempt to provide rating scale specifications able to move beyond the native speaker standard and to bring light into what are often just intuitive impressions (Isaacs & Thomson, 2013).

Derwing and Munro’s work, which aims to study phonology from a pedagogical perspective, aptly builds on the work of other scholars and reinforces the case made for the teachability of second language phonology, in particular to adult learners, something already advocated by Pennington (1998). This last study arrived at the conclusion that teaching pronunciation to adult learners is not only possible but also desirable, thus dispelling the misconception that excluded any causal link between pronunciation instruction and the acquisition of phonology in a second language. Historically, such misconceptions had originated from studies conducted in the 1960s and 70s, especially Suter (1976) and Purcell and Suter (1980). The latter in particular, widely cited as showing the futility of pronunciation instruction had had a strong influence in the field of language teaching and of ESL in particular, thus further hindering the development of the field of study both at the level of teaching and research⁵. Pennington questioned the validity of this study on three grounds: types of measures used, response of human raters, nature and quality of pronunciation training. Interestingly, **Pennington showed how human raters have a strong tendency to conflate all measures of oral proficiency even after several lengthy training sessions, thus phonological features get mixed with linguistic features like grammatical or lexical control**. This confirms what has been said about holistic scoring of speaking performance where the first impression can blur criteria-informed rating (North, 2000) while underlining the nature of speaking itself, which is a complex event where different elements act simultaneously (Bygate, 2002; Luoma, 2004) and “[t]he challenge for raters ... to reconcile their possibly idiosyncratic, intuitive, or nonlinear impressions of an L2 performance with rating scale specifications, including discretizing what might be a “grey area” performance” (Isaacs & Thomson, 2013: 135).

Furthermore, proficient speakers can be unfairly penalised due to the mismatch between their level of functional proficiency and their level of phonological competence. “[C]ertain stereotyped or frequent errors can give an overriding impression of poor pronunciation, even when, objectively, the speaker is otherwise accurate (or fluent) in the second language” (Pennington, 1998: 327). These holistic overgeneralisations and the deep-seated attitudes they lead to become critical in everyday interactions. The fact that stigmatisation of a foreign accent can lead to an experience of accent discrimination and harassment is increasingly recognised in the literature. Intolerance towards non-standard speakers, no matter how intelligible they are, is well documented (Coupland & Bishop 2007; Hansen, Rakic’ & Steffens 2014; Rakic’, Steffens & Mummendey 2011). Indeed, many adult immigrants and people working in foreign languages seek out specialised courses to neutralise their accents as a result.⁶ This emphasises the need for a shift in descriptions that inform the creation of standards away from native speaker norms towards an acceptance of accent, with a focus on intelligibility. This is particularly crucial in relation to the assessment criteria used in high stakes examinations with a gatekeeping function.

The situation is exacerbated further when tests are based on mechanical voice recognition. As already Pennington argued, resorting to “mechanical instruments that make quantitative measurements of human articulation (e.g., by electropalatography) or of the sound waves produced when speaking (by various types of acoustic analysis)” (1998: 327) raises the questionable issue of defining a native speaker standard against which the mechanical device should compare the production that needs to be judged. Even where speech recognition software averages out the phonological features of

⁵ More recently, Derwing and Munro extensively re-analyze the studies by Suter (1976) and Purcell and Suter (1980) on the origin of the misconceptions about teachability of pronunciation and confirm Pennington’s remarks also providing extra considerations not only directly on these studies but also mentioning other empirical studies done in the 70s and 80s that precisely provide evidence of pronunciation improvement as a result of different instructional methods and that oddly enough had almost never been cited.

⁶ For example, an entire industry has emerged to service call centre employees (Derwing & Munro, 2015: 136).

thousands of sampled speakers, it cannot identify from this process alone the most essential features for intelligibility. As to the third ground, Pennington mentioned several studies that counterbalance Suter (1976) and Purcell and Suter's (1980) claim that pronunciation training has no effect. All these studies agree on measurable effectiveness of pronunciation training when this follows a carefully designed program of instruction. According to the author, the two major factors that seem to have a negative impact on pronunciation training are the quality of training itself and difficulty in evaluating progress in pronunciation, **this last one benefitting greatly from knowledge of phonology by the raters, use of clear assessment criteria and consistency between the nature of tasks and the nature of assessment.**

Besides teachers' lack of knowledge and/or targeted training when it comes to pronunciation, there are two main obstacles to introduction of explicit, analytic-linguistic approaches that aim to raise learner awareness as opposed to purely intuitive-imitative approaches (Celce Murcia 2010). These are the 'learner age' question and socio-political arguments respectively.

While the "politically correct" arguments at the basis of the latter, which question the acceptability of any corrective intervention on learners' accents, seen as a threat to individuals' identity do not appear very strong and even may result in social and professional disadvantage of the learners, the former argument, which stresses the physiological limitations of adults when it comes to replicating the pronunciation of languages learned after childhood, remains powerful, and gives an excuse to ignore pronunciation teaching. This is despite the fact that, as already Pennington said several years ago "various types of structured learning experiences [...] may help the adult language learner to re-contextualize pronunciation at the perceptual, motor, cognitive, psychological, and sociocultural levels" (1998: 337) thus fostering a new attitude "which makes learning a second language a less threatening and a more ego-enhancing activity" (1998: 337). More recent research calls into question the whole idea of a critical period for learning second language pronunciation (Abu-rabia & Kehat 2004).

Derwing and Munro point out how research on the powerful influence of the learner's age on accent "has proved misleading because it emphasises what learners *cannot* usually achieve (perfect pronunciation) rather than what they are capable of achieving (intelligible speech)" (2015: 53). They underline that the view of classroom pronunciation instruction as ineffective has been refuted by systematic research and that "providing in-class instruction on aspects of pronunciation that influence intelligibility" (ibid.) is indeed useful.

Another example of successful, specialised pronunciation teaching is the training of foreign language diction to opera singers (explicit articulatory instruction: EAI)⁷ around the world on a daily basis.

Even though a more accurate and extended literature review is necessary, scholars who are investigating the pedagogical dimension of phonology seem to agree in pointing at **a generalized lack of consideration both at the level of targeted research and at the practical level.** They identify several reasons for this including a stereotyped view of the possibility of teaching and assessing pronunciation, the difficulty of making often highly technical research accessible to educators, very limited or absent formal teacher preparation and training, and lack of specific, research-informed support and resources, including assessment tools. The result is that **teachers are often left alone and very often neglect explicit teaching of pronunciation, thus disadvantaging precisely those learners that would mostly benefit from such instruction.**

This literature review aimed to capture the main tendencies and developments in research, and in particular to focus on aspects that could inform conceptualization and development of assessment tools for phonology as well as supporting/enhancing reflection on the underlying construct.

We consider that for the sake of the intended work, we have compiled sufficient documentary evidence to:

- ▶ Justify the need for increased work on phonological aspects in language education;

⁷ Leigh, 2015, unpublished MA thesis, University of Toronto. The research showed that EAI, when properly conducted, can have a crucial impact on intelligibility and accent among adult learners, and revolutionise their attitude towards pronunciation and precipitated a sense of membership in the new culture's identity. The results of this study also support the idea that sequences organized around a realistic pronunciation hierarchy, from the small phonemic units to the large prosodic units, would make pronunciation learning easier and more attainable (and confidence building) for language learners and easier to teach for instructors.

- ▶ Identify the key elements, which should inform work on phonology (in terms of goal setting, explicit teaching and assessment);
- ▶ Support the creation of phonological competence-related scales.

Analysis of resources (scales) and identification of key concepts

At the level of availability of specific tools for assessing phonological competence, beyond the CEFR phonology scale, the result of the documentary research showed that **pronunciation is rarely explicitly included among the features to be assessed, being generally subsumed under ‘speaking’ or ‘fluency’**. In the TOEFL iBT Test, for instance, references to pronunciation are to be found within the Speaking rubric under the category ‘Delivery’, which includes in the same descriptor four different, albeit overlapping features (fluency, pronunciation, intonation, and intelligibility). In the Hong Kong Diploma of Secondary Education (HKDSE), specific descriptors are provided organized around 5 levels of “pronunciation and delivery.” Only the IELTS test defines 9 bands of “pronunciation” (Cambridge IELTS).

Some of the issues we mentioned earlier as being problematic (modulation of levels through adverbs or expressions of quantity) are visible (e.g. ‘full range of pronunciation features’, ‘wide range’, ‘range’, ‘limited range’, ‘some features’ in the Cambridge scale) and in general the scales concentrate only on some pronunciation-related aspects (e.g. stress and pronunciation of words in HKDSE).

Clearly, we do not mean to be exhaustive by presenting such a limited number of existing scales, we only intend to show that recurring limitations can be observed when it comes to phonology assessment scales even those used by highly renowned institutions.

This review of the debate around the pedagogy of phonology (including the CEFR sections/scale and criticism of them, targeted literature and existing resources), albeit limited, was useful to highlight some central issues that need to be considered when dealing with pronunciation teaching and assessment. In particular it was crucial in helping to clear the ground on the effectiveness of pronunciation instruction and identify core concepts and terms that would inform future practically oriented work aimed at producing targeted descriptors.

As Munro & Derwing aptly point out in their research timeline on accent and intelligibility in pronunciation (2011) there is incongruence between research interests and pedagogical ones. In fact, the focus on accent and on accuracy instead of on intelligibility has been detrimental to the development of pedagogically-oriented work on pronunciation. **“Identifying native-like production as the central goal in pronunciation teaching inevitably leads to the conclusion that pronunciation is probably not worth teaching because of the limited likelihood of achieving that end.”** (Munro & Derwing, 2011: 317) and even though intelligibility has been seen as a priority in language instruction for a long time (the authors mention Sweet’s handbook dating back to 1900!), conceptualization of the notion for assisting teachers in their practice is a very recent process (Munro & Derwing, 1995; Jenkins, 2000, 2002; Derwing, Munro & Thomson, 2008; Zielinski, 2008, Trofimovich, Lightbown, Halter & Song, 2009). Nowadays, they say, ‘intelligibility principle’ has been clearly distinguished from ‘nativeness principle’ (Levis, 2005) and **intelligibility “is generally identified by pedagogical specialists as the most important outcome of pronunciation instruction”** (Derwing & Munro, 2005: 384) **in communicative ESL teaching, as “it is intelligibility – rather than native-like pronunciation – that is most critical for successful communication in an L2”** (Munro & Derwing, 2011: 316-317).

Furthermore, by analysing the evolution of research in the field, we can clearly notice a movement towards stressing (and consequently studying) the importance of explicit pronunciation instruction. This tendency increasingly manifested itself in expansion and diversification of the more general term of pronunciation to include different features like prosody, accent, intelligibility as well as the relation between them. This translates in the clear need for a paradigm shift where the context and learners’ needs are taken in greater consideration than idealized models even at the level of textbooks and curricula (Levis, 2005). Above all it translates into the need for explicit pronunciation instruction, where concepts such as intelligibility, comprehensibility and willingness to communicate play a crucial role (Couper, 2006; Derwing, Munro & Thomson, 2008; Trofimovich, Lightbown, Halter & Song, 2009).

Identification of key concepts

We therefore proceeded to identify the following core areas (and concepts) that would inform our work on scales/descriptors production:

- ▶ **Articulation** (including pronunciation of sounds/phonemes);
- ▶ **Prosody** (including intonation, rhythm and stress –word stress/sentence stress – and speech rate/chunking);
- ▶ **Accentedness** (accent and deviation from a ‘norm’);
- ▶ **Intelligibility** (i.e. actual understanding of an utterance by a listener) and
- ▶ **Comprehensibility** (i.e. listener’s perceived difficulty in understanding an utterance).

Operationalisation of the concepts in order to create descriptors

As Munro & Derwing clearly remind us “it makes little sense to assess pronunciation on scales of the type that range from *not accented, perfectly comprehensible* at one endpoint to *accented and difficult to understand* at the other.” (2000: 305). Not only do scales on the different phonological aspects need to be considered separately from one another but also assessment needs to be modulated according to different factors like the goals of the instructor, the learner, the curriculum and the context.

We proceeded afterwards to a possible operationalization of these concepts into targeted scales and descriptors.

It is evident that a clear distinction of each of these aspects appears unrealistic when it comes to practical application. Not only does a certain overlap exist between them but also, above all, they all influence the final judgment of raters in an often-inextricable way. We are well aware therefore that a reasonable compromise needs to be found to ensure feasibility and practicality of phonological assessment. In particular, we decided not to apply the academic distinction between intelligibility and comprehensibility in the scales, since this might confuse teachers.

Considering the common need to have a scale that provides a snapshot of a certain competence and other scales that allow for more precise and targeted work, we decided to operationalize the above-mentioned concepts into three scales:

1. General phonology scale
2. Pronunciation (sound articulation)
3. Prosody (intonation, stress and rhythm)

Accentedness as well as intelligibility and perceived comprehensibility are meant to be transversal across different levels. Therefore, they have informed the scales in a different way: **intelligibility has been key for the discrimination of levels and explicit mention of accentedness has been used to raise awareness of the non-causal relationship between accent and phonological competence.**

Production of descriptors, and the subsequent fine-tuning and categorization of them has therefore proceeded bottom up as we will explain in the following section.

Phase Three

Creation of three scales (first draft)

Following the operationalization of the concepts and the macro-categorization into three scales we developed descriptors at various levels with a careful consideration of ways of discriminating progress and of expressing it in a pedagogically sound and practically usable form. We took special care to seek for coherence with the other CEFR descriptors as far as positive expression of competence is concerned and clarity in indicators of progression. When possible we drew upon existing descriptors, but as I mentioned this did not contribute extensively to the work due to the nature of the descriptors we were able to find.

Revision of drafts and preparation of the workshop

The first draft was revised in a restricted team, with a particular focus on clarity, and the descriptors were color-coded (blue for overall phonology descriptors, red for the sound articulation descriptors and green intonation, stress and rhythm descriptors). This helped moving sections around and/or adding or eliminating sections to assure better identification of descriptors belonging to the different scales.

Workshop (preliminary internal validation)

This second draft was used as a basis for the workshop conducted with consultant experts participating at the Second consultation meeting concerning the extended version of the CEFR illustrative descriptors of language competences (30 June - 2 July 2015) within the above-mentioned Council of Europe project. The Workshop took place in Strasbourg on July 1st, 2015. Eleven experts participated, coming from different European countries working in the domain of second/foreign language education, mainly in English as a second/foreign language, but also in Français language seconde/étrangère. The workshop lasted one hour plus a follow up discussion. Experts were asked in pairs to work on two worksheets: one containing the descriptors for both the prosody/intonation (I) scale and the pronunciation (P) scale mixed and in random order; one scale (O) containing the descriptors of the overall phonological competence in random order. Participants were asked to:

1. Identify descriptors belonging to each of the I and P scales;
2. Rank the descriptors of each identified scale according to the 6 main levels of the CEFR;
3. Provide qualitative feedback on the scales and/or the single descriptors.

The rationale for the procedure followed during this workshop was to test the clarity and appropriateness of the descriptors, the coherence of what the specific aspects they intended to assess and the coverage of all relevant issues in phonological competence. All three tasks were completed by all the participants within the indicated time frame. After the workshop, the completed worksheets were collected and analysed. We received 7 worksheets as some experts worked in pairs and some individually. Some of those who worked in pairs still decided to fill in the worksheets individually to record individual differences.

Worksheet 1 (scale O):

The recognition of the levels was generally very good. Few participants ranked one or two descriptors one level higher or lower. Only one single occurrence of a movement of two levels in ranking was observed for one descriptor. There were almost no doubts at the C levels (100% recognition of the C1 level, 60% recognition of the C2 level, the remaining 40% ranked the descriptor at the C1+ level). In the A levels recognition was 100% for the A1 level, 60% recognition for the A2 level, the remaining 40% ranked the descriptor at the B1 level. Generally, a good recognition of the B levels, with some more hesitations especially in the B2 level which tended to be ranked lower (at B1 or B1/B2 level). In general, all participants completed this task without difficulty and without visible disagreement.

Worksheet 2 (scales P and I)

Task one: discrimination between the scales: The participants accomplished the first of the two tasks concerning these scales without noticeable difficulties. Out of the 26 descriptors included, only a couple were misplaced per worksheet and/or there were some hesitations marked in the margin. One descriptor which was focusing exclusively on accent was perceived as difficult to place on either scale. On one worksheet there were also some blanks that were probably due to lack of time as the discussion on the ranking of the descriptors took over. Generally speaking, clarity and consistency of the scales appeared confirmed by the workshop results.

Task two: ranking descriptors according to levels: This second task generated much more discussion than the other two, and participants would have benefitted from a longer time to accomplish the task, especially considering the specific nature of descriptors of phonological competence. This resulted in a few participants being obliged to complete the final part of the task in a rush or even to leave some of the descriptors at the end of the worksheet without ranking them. On the whole though, this task produced satisfactory results with a percent of almost 70% of correct ranking. The total of descriptors of the P and I scales together was 174 (26x7 worksheets). Out of these descriptors 55 were not ranked at the expected level. Of these, 46 were ranked either one level higher or lower, and only 9 were ranked two levels higher or lower. On top of this, some hesitations were recorded which were expressed either by indicating only the macrolevel (A, B or C) or inserting two levels separated by a slash.

Revision of the scales based on the feedback and analysis of results

Qualitative feedback: After completion of the task, an extensive discussion followed where participants provided qualitative feedback both on the descriptors and scales and on their experience in sorting out scales and ranking the descriptors. The synthesis of these insights was very helpful to inform the revision of the three scales, which is still in progress.

Provisional conclusion: The work conducted thus far clearly proved that work on phonological competence, its teachability and evaluation is timely. Even though it is a complex domain, which requires accurate reflection and has no straightforward solution, it appears that comprehensive scales can be developed and specific descriptors produced which can help identify specific issues related to phonology, and thus inform both curricula and pedagogical practice.

Consultation with experts on the revised scales for feedback

The revised scales were shared with a larger group of experts working on phonology in relation to the English and French languages.⁸ Reactions to the descriptors were very positive, some consultants gave very detailed feedback. The specific feedback on each descriptor that was received was collated and compared in order to inform revision.

Preparation for validation

On the basis of this feedback, and through wider consultation within the authoring group of the project, the descriptors were again revised and translated into French, in order to make them available for the three-phase validation process defined for the project.

⁸ Sophie de Abreu, Aline Germain-Rutherford, David Horner, Thalia Isaacs, Murray Munro

Phase Four

Qualitative validation

The study replicated the three phases of the mediation project and took place in two stages between 02.01.16 and 07.02.16:

- ▶ Phase 1: Rating descriptors for quality – and judging category and level (English only);
- ▶ Phase 2: Judging CEFR level;
- ▶ Phase 3: Using descriptors for assessment of performances on video (English and French).

The Respondents

People who had participated in the Phase 3 survey of the Mediation Project who had indicated an interest in a follow up survey for phonology were contacted. The vast majority had participated in all three phases of the Mediation Project and so were accustomed to the tasks involved. Some 250 respondents completed Phase 1 survey and 272 the combined survey for Phases 2 & 3. There was a very considerable overlap between the two groups.

Phase 1

Before Phase 1 of the validation, the set of descriptors was revised in accordance with feedback received from the experts mentioned above. The task in Phase 1 combined the tasks from Phases 1 & 2 of the Mediation Project. Respondents were asked to:

- ▶ Identify the category of the descriptors;
- ▶ Rate the descriptors for (a) clarity and (b) pedagogic usefulness;
- ▶ Judge the CEFR level (A1, A2, B1, B2, C1 or C2) of the descriptors.

The first task had a trick element, in that descriptors were included (e.g. on accent) that did not fit under the two categories given 'Sound articulation' and 'Prosodic features.' They were intended to land in the 'Can't decide' column. Analysis of the third task (judgements of level) was purely qualitative. The results are given in Appendix 1. Only one descriptor was not assigned to the correct category (D26: highlighted red), though another three were not as clearly identified with their category as was desirable. Only two descriptors were rated at less than 80% for the two quality criteria: clarity and usefulness (D8 and D22). 21 descriptors were assigned to the level intended by 40% of the respondents, with two more assigned by 40%, but with a noticeably wider spread of level, and another eight assigned to a related 'plus' level. Only three descriptors were assigned to a wide range of level. Three were also assigned evenly to C1 and C2. As a result of the exercise, five descriptors were dropped and several others reformulated in preparation for the main survey. This was to follow in English and French, combining the methodologies of Phase 2 and Phase 3 of the Mediation Project.

Quantitative validation⁹

Phase 2 (Main survey)

One task in the main survey was to rate the CEFR level of 34 descriptors grouped into the four categories:

- ▶ Sound articulation;
- ▶ Prosodic features;

⁹ This section is provided by Brian North, author of the original 2001 scales and coordinator of the 2014 – 2017 project to update and extend them.

- ▶ Other (these were more global descriptors, including those on accent);
- ▶ Sound recognition (This category was put last because the previous three had been used with the Phase 3 methodology, reported below, which actually preceded this task in the survey).

The instruction for the descriptor rating task was identical to the Phase 2 task, to answer the following question for each item:

At what CEFR level do you think a person can do what is defined in the descriptor?

However, participants were given only the 6 CEFR criterion levels for which descriptors had been prepared. There had been no attempt to make finer distinctions, because this is notoriously difficult in phonology. As for Phase 2 in the Mediation Project and as in the follow up survey for plurilingual and pluricultural competences, the data was prepared, by Cambridge Assessment, for two separate analyses:

- ▶ Classical analysis showing, for each descriptor: the percentage that rated the intended level;
- ▶ Rasch analysis (Linacre, 2014): to link the ratings to the mathematical (logit) scale created in the 1994 Swiss project analysis.

Classic analysis: The spread of levels was colour-coded as shown in Image 1:

Image 1

SOUND ARTICULATION					
	1. Can articulate a high proportion of the sounds in the target language clearly in extended stretches of production; is intelligible throughout, despite a few systematic mispronunciations.	2. Can articulate a limited number of sounds, so that speech is only intelligible if the interlocutor provides support (e.g. by repeating correctly and by eliciting repetition of new sounds).	3. Can articulate the majority of the sounds of the target language reasonably clearly in extended speech.	4. Can articulate virtually all of the sounds of the target language with a high degree of control. He/she can usually self-correct if he/she noticeably mispronounces a sound.	5. Can articulate virtually all the sounds of the target language with clarity and precision.
A1	0,00%	77,94%	0,00%	0,00%	0,00%
A2	5,88%	19,49%	5,51%	1,10%	1,47%
B1	32,72%	1,10%	41,91%	2,94%	2,21%
B2	40,07%	1,47%	42,65%	18,75%	14,71%
C1	20,22%	0,00%	9,56%	55,88%	26,84%
C2	1,10%	0,00%	0,37%	21,32%	54,78%
Intend	B2	A1	B1	C1	C2
Result	B2	A1	B1+	C1	C2

The second, fourth and fifth items in Image 1 are clearly interpreted as being the same level as that intended. The first item, *Can articulate a high proportion of the sounds in the target language clearly in extended stretches of production; is intelligible throughout, despite a few systematic mispronunciations*, is also squarely targeted at the intended level, but with a wider spread and a suggestion that it may be at the lower end of the B2 band. The third item, *Can articulate the majority of the sounds of the target language reasonably clearly in extended speech*, is coming out at B1+, straddling B1 and B2.

The images for the full set of descriptors are given in Appendix 2. The following item, on Sound recognition, (No 29) *Can recognise and reproduce in the target language sounds that are found in other language(s) he/she speaks, or that have been explicitly practised* was spread across all levels. This item also 'misfitted' in the Rasch analysis (see below) and was therefore dropped.

Rasch analysis: The quality of the data was very good, better than that in Phases 2 and 3 of the Mediation Project. 20 respondents (about 10%) were excluded from the data because of 'misfit'.

What is misfit? To simplify, a person is “misfitting” if they behave in a way that is inconsistent and improbable (e.g. rating easy items as difficult and difficult items as easy). In the current context, a ‘misfitting’ item is a descriptor that people find confusing, or which describes something completely different to the construct formed by the main body of descriptors. In our context a ‘misfitting’ person is someone who doesn’t understand the concepts involved, cannot read descriptors, or cannot make consistent judgements instance. ‘Misfit’ is balanced by ‘overfit:’ people or items that are improbably good. If a teacher did their own test they would ‘overfit.’ If an item ‘overfits,’ it is probably one that people have come across before in training, or, in our case, that uses an expression that *everyone* associates with a specific level. The reason a Rasch methodology is so effective is because by identifying sources of ‘misfit’ and, where possible, removing such persons or items from the analysis, one increases the accuracy of the measurement. Item 29, remarked on above for having a wide spread of level, was also excluded because of misfit.

The scale produced in the analysis was very long, from 4.20 logits (a mid C2) to -4.60 logits (just below the A1 cut-off). Because of the scale length and slope, straightforward anchoring to the cut-offs for the CEFR levels produced in the Swiss project seemed appropriate. However, only CEFR item available as an anchor. Therefore three more ‘global’ items adapted from Cambridge English Language Assessment criteria for A2, B1 and B2 were used as anchor items, anchored at the value of the midpoint of the band of proficiency for the criterion level. This approach worked well. However, for technical reasons, as in Phase 2 of the Mediation Project, positives and negatives were reversed in the reporting scales and therefore had to be also reversed in the anchor items.

Phase 3 (Main survey)

The main task in the survey, however, had been to rate three video performances with 27 descriptors (i.e. excluding the seven for Sound recognition). The videos used were those developed by the CIEP in 2008 and are available on line. Respondents were able to choose whether to rate a performance in English, French, German or Spanish. The links are given in Appendix 3. There were 872 responses from the 272 participants, but some 200 were excluded for misfit, leaving 667 in the unanchored analysis and 664 in an anchored one.

The unanchored analysis produced a scale from 4.81 logits (very high C2) to -4.96 logits (Pre-A1). The three Cambridge anchor items mentioned in the previous section were used to anchor the item values to the scale underlying the CEFR levels. The scale from the anchored analysis was similar to the unanchored scale, from 4.96 (very high C2) to -4.52 (just below the A1 cut-off). However, because the anchoring was to some extent artificial, a second standard-setting method was used to help define cut-points. The method used was a simplified form of the Bookmark Method described in the Council of Europe’s Manual for relating tests to the CEFR. Each member of the Authoring Group independently selected the point for the cut-off between levels on the unanchored scale. The result is shown in Appendix 4, with the different coloured horizontal lines representing the 5 judges. As can be seen from Appendix 4, the cut-points selected for A2/B1, B1/B2 and B2/C1 were identical. Four of the five chose the same cut-point for A1/A2, leaving the C1/C2 cut-point as the only controversial one.

This is not surprising considering that:

- ▶ people often have difficulty defining this C1/C2 cut-point, especially when assessing video performances;
- ▶ as can be seen from Appendix 2 (percentages of judgements at different levels), whilst respondents could distinguish clearly between C1 and C2 descriptors for *Sound articulation*, this was not the case for *Prosodic features*.
- ▶ the C1 CEFR descriptor, *Can vary intonation and place sentence stress correctly in order to express finer shades of meaning*, calibrated in 1994 as a high C1 (at 3.32) was assigned to C2 by approximately half the participants in the Phase 2 judging task and came out as by far the most difficult item in the Phase 3 assessment task (4.81 unanchored; 4.96 anchored analysis). Indeed for this reason, this items could not itself be used as an anchor item.

Treatment of accent

The other issue that emerged was treatment of accent. The implicit native speaker model and existing B2 descriptor *Has acquired a clear, natural, pronunciation and intonation* appeared to have instilled the unrealistic expectation that user/learners at the C levels would not have any accent, This B2 CEFR descriptor had actually been calibrated in 1994 at 2.53, a very high B2+, within the margin of error to the C1 cut-off. Yet research has demonstrated that that accent remains a feature of the speech of many people with even a very high level of language proficiency. It is not the ‘naturalness’ of native

speakers that is essential: it is intelligibility, which is not necessarily the same thing. Not surprisingly, the statements about accent, intended to be part of descriptors at the B and C levels on an overall scale, tended to come out lower than the level intended, as Table 1 shows. The second (C1) descriptor, however, which was retained intact as a 'double-barrelled' descriptor, demonstrates that when statements about accent are added as detail in this way, the effect of rating at a lower level is reduced.

Table 1 – Accent

Descriptor	Intended	Phase 2 Judging	Phase 3 Assessment
Any features of accent retained from other language(s) do not affect intelligibility or effective conveyance and enhancement of meaning.	C2	B2	B2/C1
Can articulate virtually all the sounds of the target language; some features of accent retained from other language(s) may be noticeable, but they do not affect intelligibility at all.	C1	B2+/C1	B2+/C1
Accent tends to be influenced by other language(s) he/she speaks, but has little or no effect on intelligibility.	B2	B1+	B2
Accent is generally influenced by other language(s) he/she speaks, and this may sometimes affect intelligibility.	B1	A2	Misfit

The intention was to include a statement about accent at all levels, including C2. Yet, as Table 1 shows, it is the intended C2 descriptor that is the most problematic. In order to retain a mention of accent at all levels, this descriptor was later rewritten as follows:

Intelligibility and effective conveyance and enhancement of meaning are not affected in any way by features of accent that may be retained from other language(s).

The B1 descriptor was dropped.

Results

Table 1 also shows that descriptors were interpreted in different ways in the two tasks. Deciding which level to assign to descriptors was therefore a question of judgement, based on all the evidence available. This was particularly the case because the anchoring to the CEFR scale was artificial, since the very reason for the project was dissatisfaction with the existing CEFR scale. Anchoring was based on three rather global items adapted from Cambridge descriptors, as stated earlier. Appendix 5 collates all the information about each of the 34 items.

The items are presented in approximate rank order of difficulty, not by entry number. The columns are the following:

- ▶ Intended level
- ▶ Entry number
- ▶ Serial number
- ▶ Phase 2 Judging task: Classic analysis (%) information 1
- ▶ Phase 2 Judging task: Rasch anchored analysis information 2
- ▶ Phase 3 Assessing task: Bookmark Method information 3
- ▶ Phase 3 Assessing task: Rasch anchored analysis information 4
- ▶ DECISION: Level decided on
- ▶ Descriptor: Wording in survey, English version

There were four pieces of information for most items to guide the decision, although the Rasch calibrations for the Phase 2 judging task appear suspect for the C levels. However, the seven items for Sound recognition did not have an entry for the Phase 3 assessment task, since they could not be used for it. Therefore these seven items have only two information points. In Appendix 5, there are 15 cases in which all information points agree with the intended level and another five in which three of the

four agree with it, making 20 items for which the decision was automatic. With a further seven items, the information was mixed and for five items it is definitely different to what was intended. One item, for Sound recognition, was dropped without discussion since it was distributed across all levels (*Can recognise and reproduce in the target language sounds that are found in other language(s) he/she speaks, or that have been explicitly practise*). Of the seven items with mixed information, three had the decision on level varied slightly from what had been intended, two items which were amended slightly in order to make clearer the intended level, and two items on dealing with accent were retained where originally intended. The five remaining items were completely rewritten once the rest of the scales had been completed.

Conclusion

The survey worked well, producing a good, long scale. The 12 items that had been interpreted differently than intended were reformulated on the basis of the experience gained rather than being eliminated in order to provide a coherent assessment scale with entries for all aspects at each proficiency level. The resulting descriptor grid is shown in Appendix 6. It is presented in three columns, the overall phonological control on the left and the two sub-scales on the right. The two descriptors for B1+ are not used in the final, published version. The grid is available in its final form in the *CEFR Companion Volume with extended illustrative descriptors* (downloadable from the Council of Europe's website).

In replacing the original CEFR scale for Phonological Control, this analytic grid can provide the basis for teachers to include appropriate objectives for phonology in their teaching and to develop assessment criteria appropriate to the level(s) concerned. Those involved in the project hope that the provision of this more realistic, analytic scale will foster attention to phonological aspects in language teaching, encourage a more analytical and comparative approach and thus contribute to the plurilingual awareness and effective communication of our learners.

References

- Abu-rabia, S. and Kehat, S. (2004). The Critical Period for Second Language Pronunciation: Is there such thing? *Educational Psychology* 24(1), 77-97.
- Alderson, J. C. A. (1991). Bands and scores. In J. C. A. Alderson and B. North (eds.) *Language testing in the 1990s*. (pp.71-86). London: Macmillan.
- Baker, A. A. (2011). ESL teachers and pronunciation pedagogy: Exploring the development of teachers' cognitions and classroom practices. In J. Levis & K. LeVelle (Eds.). Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference, Sept. 2010. (pp. 82-94), Ames, IA: Iowa State University.
- Breitkreutz, J., Derwing, T. M., & Rossiter, M. J. (2002). Pronunciation teaching practices in Canada. *TESL Canada Journal*, 19, 51–61.
- Burgess, J., & Spencer, S. (2000). Phonology and pronunciation in integrated language teaching and teacher education. *System*, 28, 191–215.
- Bygate, M. (2002). Speaking. In R. B. Kaplan (ed.). *The Oxford Handbook of Applied Linguistics* (pp. 22-38). Oxford: Oxford University Press.
- Catford, J. C. (1987). Phonetics and the teaching of pronunciation: A systemic description of English phonology. In J. Morley (Ed.), *Current perspectives on pronunciation* (pp. 87–100). Alexandria, VA: TESOL.
- Caudwell, R. (2011). Pronunciation in spontaneous speech. In P. Roach, J. Setter, & J. Esling. [Eds]. *Cambridge Pronouncing Dictionary*. [18th Edition]. pages xx-xxi. Cambridge: Cambridge University Press.
- Cauvin, E. (2012). Le CECRL et l'évaluation de la compétence phonologique: point de vue à partir d'un corpus sur l'anglais : Evaluer avec le CECRL. *Les Langues modernes* 106(1), 33-40
- Celce-Murcia, M., Brinton, D., & Goodwin, J. (2010). *Teaching pronunciation: A course book and reference guide* (2nd ed.). Cambridge: Cambridge University Press.
- Chun, D. M., Hardison, D. M., & Pennington, M. C. (2008). Technologies for prosody in context: Past and future L2 research and practice. In J. Hansen & M. Zampini (Eds.), *Phonology and second language acquisition* (pp. 323–346). Amsterdam: John Benjamins.
- Couper, G. (2006). The short and long-term effects of pronunciation instruction. *Prospect* 21.1, 46–66.
- Coupland N., Bishop, H. (2007). Ideologised values for British accents. *Journal of Sociolinguistics* 11(1), 74–93.
- Hansen, K.; Rakic´, T., & Steffens, M.C. (2014). When Actions Speak Louder Than Words: Preventing Discrimination of Nonstandard Speakers. *Journal of Language and Social Psychology* 33(1). 68–77.
- Rakic´, T., Steffens, M.C., Mummendey, A. (2011). Blinded by the Accent! The Minor Role of Looks in Ethnic Categorization. *Journal of Personality and Social Psychology* 100(1), 16–29.
- Crawford, W. W. (1987). The pronunciation monitor: L2 acquisition considerations and pedagogical priorities. In J. Morley (Ed.), *Current perspectives on pronunciation: Practices anchored in theory* (pp. 101–121). Alexandria, VA: TESOL.
- Chucchiarini, C., Neri, A. & Strik, H. (2009). Oral proficiency training in Dutch L2: The contribution of ASR-based corrective feedback. *Speech Communication*, 51(10), 853–863.
- Derwing & Munro (2015). *Pronunciation Fundamentals. Evidence-based perspectives for L2 Teaching and Research*. Amsterdam/Philadelphia: John Benjamins.
- Derwin & Rossiter (2002).

- Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, 39(3), 379-398.
- Derwing, T. M., M. J. Munro & R. I. Thomson (2008). A longitudinal study of ESL learners' fluency and comprehensibility development. *Applied Linguistics*, 29.3, 359–380.
- Frost, D. & O' Donnell, J. (in press). Evaluating the essentials, the place of prosody in oral production. In J. Volin (ed). *Pronunciation of English as a Foreign or Second Language*.
- Galaczi, E., Post, B., Li, A. and Graham, C. (2011) Measuring L2 English phonological proficiency: Implications for language assessment. In J. Angouri, M. Daller and J. Treffers-Daller (eds) *The Impact of Applied Linguistics: Proceedings of the 44th Annual Meeting*
- Gass, S. & E. M. Varonis (1984). The effect of familiarity on the comprehensibility of nonnative speech. *Language Learning* 34.1, 65–89.
- Germain, A. & and Martin, P. (2000). Présentation d'un logiciel de visualisation pour l'apprentissage de l'oral en langue seconde . ALSIC. Vol. 3, n° 1. Pp. 61-76. Retrieved from <http://alsic.revues.org/1796>
- Harding, L. (2013). Pronunciation assessment. In C.A. Chapelle (ed.) *The Encyclopedia of Applied Linguistics*. Oxford: Blackwell.
- Harding, L. (2016, in press). 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*. Bristol, UK: Multilingual Matters.
- Homer, D. (2010). *Le CECRL et l'évaluation de l'oral: Anglais*. Paris: Belin.
- Homer, D. (2013) Towards a new phonological control grid. In E. D. Galaczi & C. J. Weir (Eds.), *Studies in Language Testing: Vol. 36 Exploring Language Frameworks. Proceedings of the ALTE Kraków Conference* (pp.187-204). Cambridge: UCLES/Cambridge University Press.
- Homer, D. (2014) And What about Testing Pronunciation? A Critical Look at the CEFR Pronunciation Grid and a Proposal for Improvement. in L.M Rupp & R. van den Doel (eds.), *Pronunciation Matters*, University of Utrecht.
- Isaacs, T & Thomson, R. (2013). Rater Experience, Rating Scale Length, and Judgments of L2 Pronunciation: Revisiting Research Conventions. *Language Assessment Quarterly*, 10, 135–159.
- Isaacs, T. (2008). Towards Defining a Valid Assessment Criterion of Pronunciation Proficiency in Non-Native English-Speaking Graduate Students. *The Canadian Modern Language Review/La Revue canadienne des langues vivantes*, 64(4), 555–580.
- Isaacs, T. and Trofimovich, P. (2012) Deconstructing comprehensibility: identifying the linguistic influences on listener 'L2 comprehensibility ratings. *Studies in Second Language Acquisition* 34 (3), 475-505.
- Isaacs, T., Trofimovich, P., Yu, G., & Chereau, B. M. (2015). Examining the linguistic aspects of speech that most efficiently discriminate between upper levels of the revised IELTS pronunciation scale. *IELTS research reports online series*, 4.
- Jamieson, J. & Poonpon, K. (2013). *Developing Analytic Rating Guides for TOEFL iBT® Integrated Speaking Tasks*. TOEFL iBT® Research Report TOEFL iBT–20. Available online: <https://www.ets.org/Media/Research/pdf/RR-13-13.pdf>
- Jenkins, J. (2000). *The phonology of English as an international language*. Oxford: Oxford University Press.
- Jenkins, J. (2002). A sociolinguistically-based, empirically-researched pronunciation syllabus for English as an international language. *Applied Linguistics*, 23, 83–103.
- Lee, J., Jang, J., & Plonsky, L. (2015). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics* 36, 345-366.
- Leigh, S. (2015). Testing an approach to teaching Italian lyric diction to opera singers: An action research study. Unpublished MA thesis. University of Toronto.
- Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly* 39(3), 369–377.

- Linacre, John M. (2014) *Winsteps: Rasch-model Computer Program*, Chicago: MESA Press.
- Luoma, S. (2004). *Assessing speaking*, Cambridge: Cambridge University Press.
- MacDonald, S. (2002). Pronunciation—views and practices of reluctant teachers. *Prospect*, 17(3), 3–18.
- MacIntyre, P. D., Baker, S. C., Clément, R., & Donovan, L. A. (2003). Talking in order to learn: Willingness to communicate and intensive language programs. *The Canadian Modern Language Review*, 59, 587–605.
- Morley, J. (1992). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly*, 25, 481-520.
- Munro, M. J., & Derwing, T. M. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech*, 38, 289–306.
- Munro, M. J., & Derwing, T. M. (1999). Foreign Accent, Comprehensibility, and Intelligibility in the Speech of Second Language Learners. *Language Learning* 49(1), 285-310.
- Munro, M. J., Derwing, T. M. (2011). The foundations of accent and intelligibility in pronunciation research. *Language Teaching* 44(3), 316–327
- Murphy, J. M. & Baker, A. A. (2015). History of ESL pronunciation teaching. In M. Reed & J. M. Levis (Eds.), *The handbook of English pronunciation* (pp. 36–65). Hoboken, NJ: Wiley Blackwell.
- North, B. (2000) *The Development of a Common Framework Scale of Language Proficiency*, New York: Peter Lang.
- Pennington, M. C. (1998). The teachability of phonology in adulthood: a re-examination. *International Review of Applied Linguistics in Language Teaching* 36(4): 323-341.
- Pennington, M. C., & Richards, J. C. (1986). Pronunciation revisited. *TESOL Quarterly*, 20, 207–225.
- Purcell, E. T., Suter, R. W. (1980). Predictors of pronunciation accuracy: A re-examination. *Language Learning*, 30(2), 271-88.
- Rubin, D. L. (1992). Nonlanguage factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants. *Research in Higher Education*, 33, 511–531.
- Saito, K. (2012). Effects of instruction on L2 pronunciation development: A synthesis of 15 quasi-experimental intervention studies. *TESOL Quarterly*, 46(4), 842-854.
- Suter, R. W. (1976). Predictors of pronunciation accuracy in second language learning. *Language Learning*, 26, 233-53.
- Thomson, R. I. (2012a). English Accent Coach: Not quite a fairy godmother for pronunciation instruction, but a step in the right direction. *Contact*, 38(1), 18-24.
- Thomson, R.I. (2012b). English accent coach [Online game]. Retrieved from www.englishaccentcoach.com.
- Thomson, R. I. & Derwing, T. M. (2014). The effectiveness of L2 pronunciation instruction: a narrative review. *Applied Linguistics* 36, 326–344
- Trofimovich, P & Isaacs, T. (2012). Disentangling accent from comprehensibility, *Bilingualism: Language and Cognition* 15 (4), 905–916.
- Trofimovich, P., P. Lightbown, R. H. Halter & H. Song (2009). Comprehension based practice: The development of L2 pronunciation in a listening and reading program. *Studies in Second Language Acquisition* 31.4, 609–639.
- Zielinski, B. W. (2008). The listener: No longer the silent partner in reduced intelligibility. *System* 36.1, 69–84.

Exams-related resources:

Hong Kong Examinations and Assessment Authority (HKEAA). Hong Kong Diploma of Secondary Education Examination. Available online: <http://www.hkeaa.edu.hk/en/hkdse/>

HKEAA. Hong Kong Diploma of Secondary Education Examination. English Language Level Descriptors. Speaking Descriptors. Available online: http://www.hkeaa.edu.hk/DocLibrary/HKDSE/Subject_Information/eng_lang/LevelDescriptors-ENG-Speaking.pdf

IELTS SPEAKING: Band Descriptors (public version). Available online: <https://www.ielts.org/pdf/SpeakingBanddescriptors.pdf>

Educational Testing Service (2014). Independent speaking rubrics/Integrated speaking rubrics. (TOEFL iBT Test). Available online: http://www.ets.org/s/toefl/pdf/toefl_speaking_rubrics.pdf, and scoring guides- Speaking. Available online:

http://www.ets.org/s/toefl/pdf/toefl_student_test_prep_planner.pdf

Trinity College London (2015) Integrated Skills in English (ISE) Specifications – Speaking & Listening ISE Foundation to ISE III. Available online: www.trinitycollege.com/resource/?id=6298

Appendices

Appendix 1 – Phase 1: Rating descriptors for quality – and judging category and level

Descriptor	Intended	Sound	Prosody	Can't decide	Clearly Formulated (% yes)	Useful (% yes)	A1	A2	B1	B2	C1	C2				
D1	Overall	13%	79%	8%	82%	90%	28%	57%	13%	2%	0%	0%	A2	A2	√	A very strong influence from other language(s) he/she speaks on stress, rhythm and intonation may affect intelligibility, requiring collaboration from interlocutors. Nevertheless, pronunciation of familiar words is clear.
D2	Overall	52%	27%	21%	81%	81%	4%	34%	53%	8%	1%	0%	B1	B1	√	Accent is generally influenced by other language(s) he/she speaks, and this may sometimes affect intelligibility.
D3	Overall	47%	27%	25%	85%	81%	0%	2%	28%	56%	11%	2%	B2	B2	√	Accent tends to be influenced by other language(s) he/she speaks, but has little or no effect on intelligibility.
D4	Overall	38%	40%	22%	84%	83%	21%	64%	14%	0%	0%	0%	A1	A2	(√)	Articulation of longer utterances requires (significant) effort, so that collaboration and support from interlocutors may (well) be needed to assure intelligibility.
D5	Prosody	3%	96%	1%	93%	92%	1%	7%	50%	28%	12%	2%	B1	B1	√	Can approximate common prosodic features of the target language in longer utterances (e.g. rising intonation for open questions), in order to convey the appropriate meaning.
D6	Sounds	97%	1%	1%	97%	96%	94%	6%	0%	0%	0%	0%	A1	A1	√	Can articulate a limited number of sounds, so that speech is only intelligible if the interlocutor provides support (e.g. by repeating correctly and by eliciting repetition of new sounds).
D7	Overall (Sounds)	97%	2%	1%	92%	87%	7%	31%	26%	28%	7%	1%	A2	-	X	Can articulate a range of sounds in the target language correctly and can recognize new sounds and generally pronounce them intelligibly.

Descriptor	Intended	Sound	Prosody	Can't decide	Clearly Formulated (% yes)	Useful (% yes)	A1	A2	B1	B2	C1	C2				
D8	Sounds	94%	4%	2%	56%	59%	3%	18%	49%	21%	7%	1%	B1	(B1)	X	Can articulate sounds that compensate for his/her mispronunciations in a way that enhances intelligibility.
D9	Overall (Sounds)	96%	3%	1%	96%	96%	1%	2%	11%	44%	32%	10%	B2	(B2)	√	Can articulate the majority of the sounds in the target language with clarity and precision clearly and intelligibly.
D10	Overall (Sounds)	97%	2%	1%	92%	91%	1%	11%	39%	38%	10%	0%	B1	B1+	(√)	Can articulate the majority of the sounds of the target language reasonably clearly precision.
D11	Overall (Sounds)	96%	3%	0%	98%	96%	0%	2%	16%	35%	37%	9%	C1	B2+	(√)	Can articulate virtually all of the sounds of the target language intelligibly with precision, though occasional phonological lapses still occur. He/she can usually self-correct if he/she noticeably mispronounces a sound.
D12	Sounds	96%	3%	1%	93%	89%	0%	1%	6%	19%	30%	44%	C1	(C2)	C1/C2	Can articulate virtually all the sounds of the target language.
D13	Sounds	82%	7%	11%	92%	91%	0%	8%	43%	34%	12%	1%	B2	B1+	(√)	Can be understood (throughout), despite habitual mispronunciation of some individual words or sounds.
D14	Sounds	88%	5%	6%	96%	92%	0%	0%	1%	26%	35%	38%	C2	-	C1/C2	Can clearly articulate virtually all the sounds of the target language and is fully intelligible throughout with precision.
D15	Prosody	2%	97%	1%	99%	94%	0%	0%	0%	1%	51%	48%	C2	C1	(√)	Can convey (fine) shades of meaning by manipulating prosodic features of spoken discourse (e.g. stress, rhythm and intonation).
D16	Prosody	7%	87%	5%	87%	86%	0%	2%	9%	49%	22%	17%	C1	(B2)	(√)	Can convey his/her message in a (fully-) intelligible way, regardless of any prosodic features (e.g. stress, intonation, rhythm) retained from the language(s) he/she speaks
D17	Prosody	3%	94%	3%	94%	93%	0%	16%	67%	16%	1%	0%	B1	B1	√	Can convey the main point of his/her message in an intelligible way in spite of a strong influence on stress, intonation and rhythm from other language(s) he/she speaks.
D18	Overall	29%	51%	20%	86%	86%	0%	0%	5%	37%	39%	18%	B2	B2+	(√)	Can draw on his/her broad repertoire to predict the phonological features of

Descriptor	Intended	Sound	Prosody	Can't decide	Clearly Formulated (% yes)	Useful (% yes)	A1	A2	B1	B2	C1	C2					
																	most unfamiliar words (e.g. word stress).
D19	Overall	2%	76%	23%	91%	92%	0%	0%	0%	2%	28%	71%	C2	C2	√		Can employ the full range of phonological features in the target language with a high level of control – including prosodic features such as word and sentence stress, rhythm and intonation – so that the finer points of his/her message are clear and precise.
D20	Sounds	88%	4%	9%	96%	95%	0%	11%	64%	25%	1%	0%	B1	B1	√		Can generally be understood throughout, despite regular mispronunciation of individual sounds and words he/she is less familiar with.
D21	Prosody	0%	99%	1%	97%	92%	0%	1%	36%	57%	6%	1%	B2	B2	√		Can generally employ prosodic features (e.g. stress, intonation, rhythm) to support the message he/she intends to convey, though with some noticeable influence from other languages he/she speaks.
D22	Sounds	93%	2%	5%	76%	67%	9%	12%	31%	31%	14%	4%	A2	B1-B2	X		Can predict and approximate the pronunciation of phonemes and unfamiliar words on the basis of the other language(s) he/she speaks.
D23	Sounds	86%	6%	8%	85%	83%	2%	13%	44%	29%	10%	2%	B2	(B1)	?		Can predict the probable articulatory features of unfamiliar words by generalising from words he/she does know.
D24	Prosody	22%	47%	31%	91%	95%	72%	27%	1%	0%	0%	0%	A1	A1	√		Can pronounce a limited repertoire of simple words and phrases intelligibly, in spite of a very strong influence on prosody (e.g. stress, rhythm, intonation) from other language(s) he/she speaks; his/her interlocutor needs to be collaborative.
D25	Prosody	22%	48%	31%	94%	93%	17%	66%	15%	2%	0%	0%	A2	A2	√		Can pronounce everyday words and phrases intelligibly, in spite of a very strong influence on prosody (e.g. stress, intonation, rhythm) from other language(s) he/she speaks.
D26	Sounds	34%	28%	38%	89%	82%	1%	1%	4%	11%	36%	48%	C1	(C2)	(√)		Can recognise and adjust to features of regional and socio-linguistic varieties of

Descriptor	Intended	Sound	Prosody	Can't decide	Clearly Formulated (% yes)	Useful (% yes)	A1	A2	B1	B2	C1	C2					
																	pronunciation, so that they are not a significant obstacle to comprehension.
D27	Sounds	96%	3%	1%	90%	85%	61%	24%	10%	6%	0%	0%	A1	A1	√		Can recognise and reproduce sounds in the target language correctly if carefully guided.
D28	Sounds	95%	3%	3%	86%	86%	33%	43%	19%	5%	0%	1%	A2	(A2)	(√)		Can recognise and reproduce the sounds he/she is familiar with (e.g. from teaching, noticing them him/herself, similarity to other language(s) he/she speaks).
D29	Sounds	71%	11%	18%	91%	88%	1%	4%	33%	34%	22%	6%	B2	B1-C1	B1 to C1		Can recognise common words when pronounced in a different regional variety from the one(s) he/she is accustomed to.
D30	Sounds	46%	19%	35%	90%	82%	1%	1%	2%	28%	35%	35%	C2	-	B2+ to C2		Can recognise features of regional and socio-linguistic varieties of pronunciation well enough to incorporate them (e.g. through adjusting, describing or imitating articulation).
D31	Sounds	86%	4%	10%	87%	85%	21%	38%	28%	10%	2%	1%	A1	A2-B1	X		Can recognise sounds and words if they are clearly articulated in the variety of the language with which he/she is most familiar.
D32	Sounds	53%	14%	33%	88%	84%	3%	9%	44%	35%	8%	2%	B1	(B1)	(√)		Can recognise the presence of regional varieties of pronunciation, though this may hinder comprehension.
D33	Overall	41%	17%	41%	94%	95%	63%	34%	3%	1%	0%	0%	A1	A1	√		Can reproduce correctly a limited range of sounds as well as the stress on simple, familiar words and phrases.
D34	Prosody	1%	99%	0%	96%	94%	0%	1%	5%	44%	40%	10%	C1	B2+	(√)		Can sustain flexible use of prosodic features of spoken discourse (e.g. stress, rhythm and intonation) with only occasional lapses of control.
D35	Overall	32%	30%	38%	85%	85%	1%	0%	19%	48%	22%	10%	C2	(B2)	X		Features of accent retained from another language do not affect intelligibility or effectiveness.
D36	Overall	31%	32%	36%	82%	81%	1%	1%	38%	48%	9%	3%	C1	(B2)	(√)		Features of accent retained from other language(s) he/she speaks may well be noticeable but do not affect intelligibility.

Descriptor	Intended	Sound	Prosody	Can't decide	Clearly Formulated (% yes)	Useful (% yes)	A1	A2	B1	B2	C1	C2				
D37	Overall	74%	7%	19%	95%	94%	25%	65%	10%	1%	0%	0%	A2	A2	√	Pronunciation is generally intelligible when communicating in simple everyday situations, provided the interlocutor makes an effort to understand specific sounds.
D38	Overall (Prosody)	2%	98%	0%	94%	89%	1%	1%	34%	52%	11%	2%	C1	B2	(√)	Prosodic features may retain influence from other language(s) that he/she speaks, but this does not impede clarity intelligibility????
D39	Overall (Prosody)	1%	98%	1%	98%	95%	21%	60%	18%	1%	0%	0%	B1	A2	(√)	Prosodic features (e.g. word stress) are appropriate for familiar words and simple utterances.
D40	Sounds	92%	5%	4%	87%	87%	26%	49%	22%	3%	0%	0%	A2	(A2)	√	Regular mispronunciation of phonemes does not hinder intelligibility, provided the interlocutor makes an effort to recognise and adjust to the influence of the speaker's language background on pronunciation.
D41	Prosody	3%	97%	1%	97%	96%	0%	0%	2%	16%	44%	38%	C2	(C1)	(√)	Stress, rhythm and intonation are fully adequate to produce consistently smooth, intelligible discourse. Any lapses arising from the influence of other language(s) he/she speaks do not affect intelligibility or effectiveness.
														√	17	
														(√)	15	
														C1/C2	3	
														DROP	5	
														?	1	

Appendix 2 – Phase 2: Judging CEFR level: Percentages for different levels

SOUND ARTICULATION									
	1. Can articulate a high proportion of the sounds in the target language clearly in extended stretches of production; is intelligible throughout, despite a few systematic mispronunciations.	2. Can articulate a limited number of sounds, so that speech is only intelligible if the interlocutor provides support (e.g. by repeating correctly and by eliciting repetition of new sounds).	3. Can articulate the majority of the sounds of the target language reasonably clearly in extended speech.	4. Can articulate virtually all of the sounds of the target language with a high degree of control. He/she can usually self-correct if he/she noticeably mispronounces a sound.	5. Can articulate virtually all the sounds of the target language with clarity and precision.	6. Can generalise from his/her repertoire to predict the phonological features of most unfamiliar words (e.g. word stress) with reasonable accuracy (e.g. whilst reading).	7. Is generally intelligible throughout, despite regular mispronunciation of individual sounds and words he/she is less familiar with.	8. Pronunciation is generally intelligible when communicating in simple everyday situations, provided the interlocutor makes an effort to understand specific sounds.	9. Systematic mispronunciation of phonemes does not hinder intelligibility, provided the interlocutor makes an effort to recognise and adjust to the influence of the speaker's language background on pronunciation.
1	0	212	0	0	0	3	2	40	82
2	16	53	15	3	4	13	52	192	130
3	89	3	114	8	6	60	159	39	54
4	109	4	116	51	40	117	55	1	4
5	55	0	26	152	73	57	4	0	0
6	3	0	1	58	149	22	0	0	2
checker	272	272	272	272	272	272	272	272	272
A1	0,00%	77,94%	0,00%	0,00%	0,00%	1,10%	0,74%	14,71%	30,15%
A2	5,88%	19,49%	5,51%	1,10%	1,47%	4,78%	19,12%	70,59%	47,79%
B1	32,72%	1,10%	41,91%	2,94%	2,21%	22,06%	58,46%	14,34%	19,85%
B2	40,07%	1,47%	42,65%	18,75%	14,71%	43,01%	20,22%	0,37%	1,47%
C1	20,22%	0,00%	9,56%	55,88%	26,84%	20,96%	1,47%	0,00%	0,00%
C2	1,10%	0,00%	0,37%	21,32%	54,78%	8,09%	0,00%	0,00%	0,74%
		1		1	1		1	1	
			1			1			1
	1								
Intend	B2	A1	B1	C1	C2	B2	B1	A2	A2
Result	B2	A1	B1+	C1	C2	B2	B1	A2	A2

PROSODIC FEATURES

	10. Can approximate common prosodic features of the target language in longer utterances (e.g. rising intonation for open questions), in order to convey the appropriate meaning.	11. Can convey the main point of his/her message in an intelligible way in spite of a strong influence on stress, intonation and/or rhythm from other language(s) he/she speaks.	12. Can exploit prosodic features (e.g. stress, rhythm and intonation) appropriately and effectively in order to convey finer shades of meaning (e.g. to differentiate and emphasise).	13. Can generally employ prosodic features (e.g. stress, intonation, rhythm) to support the message he/she intends to convey, though with some noticeable influence from other languages he/she speaks.	14. Can produce smooth, intelligible spoken discourse with only occasional lapses in control of stress, rhythm and/or intonation, which do not affect intelligibility or effectiveness.	15. Can use the prosodic features of a limited repertoire of simple words & phrases intelligibly, in spite of a very strong influence on stress, rhythm, and/or intonation from other language(s) he/she speaks; needs a collaborative interlocutor.	16. Can use the prosodic features of everyday words and phrases intelligibly, in spite of a very strong influence on stress, intonation and/or rhythm from other language(s) he/she speaks.	17. Can vary intonation and place sentence stress correctly in order to express finer shades of meaning.	18. Prosodic features (e.g. word stress) are appropriate for familiar, everyday words and simple utterances.
1	2	10	0	1	0	143	29	0	50
2	16	77	1	16	1	91	163	2	126
3	91	146	7	110	15	30	64	3	78
4	113	31	31	121	87	6	16	28	14
5	39	6	110	22	126	2	0	111	3
6	11	2	123	2	43	0	0	128	1
checker	272	272	272	272	272	272	272	272	272
A1	0,74%	3,68%	0,00%	0,37%	0,00%	52,57%	10,66%	0,00%	18,38%
A2	5,88%	28,31%	0,37%	5,88%	0,37%	33,46%	59,93%	0,74%	46,32%
B1	33,46%	53,68%	2,57%	40,44%	5,51%	11,03%	23,53%	1,10%	28,68%
B2	41,54%	11,40%	11,40%	44,49%	31,99%	2,21%	5,88%	10,29%	5,15%
C1	14,34%	2,21%	40,44%	8,09%	46,32%	0,74%	0,00%	40,81%	1,10%
C2	4,04%	0,74%	45,22%	0,74%	15,81%	0,00%	0,00%	47,06%	0,37%
		1				1	1		
					1				1
	1		1	1				1	
Intend	B1	B1	C2	B2	C1	A1	A2	C1	A2
Result	B1+	B1	C1/C2	B1+	C1	A1	A2	C1/C2	A2

OTHER									
	19. A very strong influence from other language(s) he/she speaks on stress, rhythm and intonation may affect intelligibility, requiring collaboration from interlocutors. Nevertheless, pronunciation of familiar words is clear.	20. Accent is generally influenced by other language(s) he/she speaks, and this may sometimes affect intelligibility.	21. Accent tends to be influenced by other language(s) he/she speaks, but has little or no effect on intelligibility.	22. Any features of accent retained from other language(s) do not affect intelligibility or effective conveyance and enhancement of meaning.	23. Can articulate virtually all the sounds of the target language; some features of accent retained from other language(s) may be noticeable, but they do not affect intelligibility at all.	24. Can employ the full range of phonological features in the target language with sufficient control to ensure intelligibility throughout.	25. Can generally use appropriate intonation, place stress correctly and articulate individual sounds clearly.	26. Can reproduce correctly a limited range of sounds as well as the stress on simple, familiar words and phrases.	27. Pronunciation is generally intelligible; can approximate intonation and stress at both utterance and word levels.
1	123	42	0	0	0	0	3	118	8
2	125	122	11	5	4	1	15	111	56
3	18	101	123	24	35	18	56	35	159
4	4	6	114	124	77	58	132	4	38
5	1	0	21	73	114	113	44	2	9
6	1	1	3	46	42	82	22	2	2
checker	272	272	272	272	272	272	272	272	272
A1	45,22%	15,44%	0,00%	0,00%	0,00%	0,00%	1,10%	43,38%	2,94%
A2	45,96%	44,85%	4,04%	1,84%	1,47%	0,37%	5,51%	40,81%	20,59%
B1	6,62%	37,13%	45,22%	8,82%	12,87%	6,62%	20,59%	12,87%	58,46%
B2	1,47%	2,21%	41,91%	45,59%	28,31%	21,32%	48,53%	1,47%	13,97%
C1	0,37%	0,00%	7,72%	26,84%	41,91%	41,54%	16,18%	0,74%	3,31%
C2	0,37%	0,37%	1,10%	16,91%	15,44%	30,15%	8,09%	0,74%	0,74%
									1
		1		1		1	1		
	1		1					1	
					1				
Intend	A2	A2	B2	! C2 !	C1	C1	B2	A1	B1
Result	A1/A2	A2	B1+	B2	C1	C1	B2	A1/A2	B1

Appendix 3 – Video performances for assessment in the Phase 3 assessment task

English	Duration	Link
Amélie	3.13	http://www.ciep.fr/en/books-and-cd-roms-dealing-with-assessment-and-certifications/dvd-spoken-performances-illustrating-the-6-levels-of-the-common-european-framework-of-reference-for/english-amelie-a-book
Xavier	3.25	http://www.ciep.fr/en/books-and-cd-roms-dealing-with-assessment-and-certifications/dvd-spoken-performances-illustrating-the-6-levels-of-the-common-european-framework-of-reference-for/english-xavier-role-of-teachers-and-parents
Tifaine	2.05	http://www.ciep.fr/en/books-and-cd-roms-dealing-with-assessment-and-certifications/dvd-spoken-performances-illustrating-the-6-levels-of-the-common-european-framework-of-reference-for/english-tifaine-likes-and-dislikes
German		
Caroline	2.06	http://www.ciep.fr/de/bucher-und-cd-roms-zum-the-ma-evaluierung-und-zertifizierungen/dvd-mundliche-leistungen-beispiele-fur-die-6-niveaustufen-gemeinsamen-europaischen-referenzrahmens/deutsch-caroline-rolle-der-lehrer-eltern
Mathias	3.15	http://www.ciep.fr/de/bucher-und-cd-roms-zum-the-ma-evaluierung-und-zertifizierungen/dvd-mundliche-leistungen-beispiele-fur-die-6-niveaustufen-gemeinsamen-europaischen-referenzrahmens/deutsch-mathias-ein-film
Matheiu	1.49	http://www.ciep.fr/de/bucher-und-cd-roms-zum-the-ma-evaluierung-und-zertifizierungen/dvd-mundliche-leistungen-beispiele-fur-die-6-niveaustufen-gemeinsamen-europaischen-referenzrahmens/deutsch-mathieu-hobbys-und-interessen
French		
Tobias	3.35	http://www.ciep.fr/ressources/ouvrages-cedero ms-en-evaluation-certification/s/dvd-productions-orales-illustrent-les-6-niveaux-cadre-europeen-commun-reference-les-langues/francais/francais-tobias-livre
Inge	2.31	http://www.ciep.fr/ressources/ouvrages-cedero ms-en-evaluation-certification/s/dvd-productions-orales-illustrent-les-6-niveaux-cadre-europeen-commun-reference-les-langues/francais/francais-inge-role-professeurs
Julietta	1.09	http://www.ciep.fr/ressources/ouvrages-cedero ms-en-evaluation-certification/s/dvd-productions-orales-illustrent-les-6-niveaux-cadre-europeen-commun-reference-les-langues/francais/francais-julietta-journee-habituelle
Spanish		
Vincent	1.28	http://www.ciep.fr/es/publicaciones-y-cd-roms-dedicados-a-evaluacion-y-a-certificacion/dvd-producciones-orales-ilustran-los-6-niveles-del-marco-comun-europeo-referencia-para-las-lenguas/espanol-vicente-las-nuevas-tecnologias
Alexandros	2.07	http://www.ciep.fr/es/publicaciones-y-cd-roms-dedicados-a-evaluacion-y-a-certificacion/dvd-producciones-orales-ilustran-los-6-niveles-del-marco-comun-europeo-referencia-para-las-lenguas/espanol-alexandros-libro
Katherina	2.05	http://www.ciep.fr/es/publicaciones-y-cd-roms-dedicados-a-evaluacion-y-a-certificacion/dvd-producciones-orales-ilustran-los-6-niveles-del-marco-comun-europeo-referencia-para-las-lenguas/espanol-katharina-tiempo-libre

Appendix 4 – Phase 3: Using descriptors for assessment: Bookmark method with core group

ORDER OF DIFFICULTY – VIDEO ANALYSIS					
C1	4.81	17	688	Can vary intonation and place sentence stress correctly in order to express finer shades of meaning.	
C2	4.76	12	686	Can exploit prosodic features (e.g. stress, rhythm and intonation) appropriately and effectively in order to convey finer shades of meaning (e.g. to differentiate and emphasise).	
C2	3.64	5	677	Can articulate virtually all the sounds of the target language with clarity and precision.	C2
C1	3.55	4	678	Can articulate virtually all of the sounds of the target language with a high degree of control. He/she can usually self-correct if he/she noticeably mispronounces a sound.	
C1	2.64	14	687	Can produce smooth, intelligible spoken discourse with only occasional lapses in control of stress, rhythm and/or intonation, which do not affect intelligibility or effectiveness.	
C1	2.64	24	697	Can employ the full range of phonological features in the target language with sufficient control to ensure intelligibility throughout.	
C1	1.79	23	696	Can articulate virtually all the sounds of the target language; some features of accent retained from other language(s) may be noticeable, but they do not affect intelligibility at all.	C1
C2	1.76	22	695	Any features of accent retained from other language(s) do not affect intelligibility or effective conveyance and enhancement of meaning.	
B2	1.43	6	680	Can generalise from his/her repertoire to predict the phonological features of most unfamiliar words (e.g. word stress) with reasonable accuracy (e.g. whilst reading).	
B1	1.33	10	690	Can approximate common prosodic features of the target language in longer utterances (e.g. rising intonation for open questions), in order to convey the appropriate meaning.	
B2	1.18	25	703	Can generally use appropriate intonation, place stress correctly and articulate individual sounds clearly.	B2
B2	.55	21	698	Accent tends to be influenced by other language(s) he/she speaks, but has little or no effect on intelligibility.	
B2	.45	1	679	Can articulate a high proportion of the sounds in the target language clearly in extended stretches of production; is intelligible throughout, despite a few systematic mispronunciations.	
B1	-.07	3	681	Can articulate the majority of the sounds of the target language reasonably clearly in extended speech.	
B2	-.24	13	689	Can generally employ prosodic features (e.g. stress, intonation, rhythm) to support the message he/she intends to convey, though with some noticeable influence from other languages he/she speaks.	B1

B1	-1.05	27	704	Pronunciation is generally intelligible; can approximate intonation and stress at both utterance and word levels.
B1	-1.53	7	682	Is generally intelligible throughout, despite regular mispronunciation of individual sounds and words he/she is less familiar with.
A1	-2.21	18	694	Prosodic features (e.g. word stress) are appropriate for familiar, everyday words and simple utterances.
A2	-2.79	19	700	A very strong influence from other language(s) he/she speaks on stress, rhythm and intonation may affect intelligibility, requiring collaboration from interlocutors. Nevertheless, pronunciation of familiar words is clear.
A2	-2.82	9	683	Systematic mispronunciation of phonemes does not hinder intelligibility, provided the interlocutor makes an effort to recognise and adjust to the influence of the speaker's language background on pronunciation.
B1	-3.02	11	691	Can convey the main point of his/her message in an intelligible way in spite of a strong influence on stress, intonation and/or rhythm from other language(s) he/she speaks.
A2	-3.02	16	692	Can use the prosodic features of everyday words and phrases intelligibly, in spite of a very strong influence on stress, intonation and/or rhythm from other language(s) he/she speaks.
A1	-4.23	26	701	Can reproduce correctly a limited range of sounds as well as the stress on simple, familiar words and phrases.
A2	-4.59	8	684	Pronunciation is generally intelligible when communicating in simple everyday situations, provided the interlocutor makes an effort to understand specific sounds.
A1	-4.96	15	693	Can use the prosodic features of a limited repertoire of simple words and phrases intelligibly, in spite of a very strong influence on stress, rhythm, and/or intonation from other language(s) he/she speaks; his/her interlocutor needs to be collaborative.

A2

A1

Dropped (for now) – the A1 item was rated 78& A1 in the other task – but No20 definitely doesn't work.

A1	2	685	Can articulate a limited number of sounds, so that speech is only intelligible if the interlocutor provides support (e.g. by repeating correctly and by eliciting repetition of new sounds).
B1	20	699	Accent is generally influenced by other language(s) he/she speaks, and this may sometimes affect intelligibility.

Appendix 5 – Item analysis: All information on items

Intended			Phase 2 Judging Classic %	Phase 2 Judging Rasch Anchored	Phase 3 Assessing Bookmark Method	Phase 3 Assessing Rasch anchored	DECISION	
C2	28	705	C2	C2	-	-	C2	Can consciously incorporate relevant features of regional and socio-linguistic varieties of pronunciation appropriately.
C1	17	688	C1/C2	C1	C2	C2	C1	Can vary intonation and place sentence stress correctly in order to express finer shades of meaning precisely what he/she means to say.
C2	5	677	C2	C1	C2 (CD:C1)	C2	C2	Can articulate virtually all the sounds of the target language with clarity and precision.
C2	12	686	C1/C2	C1	C2	C2	C2	Can exploit prosodic features (e.g. stress, rhythm and intonation) appropriately and effectively in order to convey finer shades of meaning (e.g. to differentiate and emphasise).
C1	32	709	C1	C1	-	-	C1	Can recognise features of regional and socio-linguistic varieties of pronunciation and consciously incorporate the most prominent of them in his/her speech.
C1	24	697	C1	B2+	C1	C1	C1	Can employ the full range of phonological features in the target language with sufficient control to ensure intelligibility throughout.
C1	4	678	C1	B2+	C1 (TG:C2)	C1	C1	Can articulate virtually all of the sounds of the target language with a high degree of control. He/she can usually self-correct if he/she noticeably mispronounces a sound.
C1	14	687	C1	B2+	C1	C1	C1	Can produce smooth, intelligible spoken discourse with only occasional lapses in control of stress, rhythm and/or intonation, which do not affect intelligibility or effectiveness.
C1	23	696	B2+/C1 B1-C2	B2+	C1	B2+	C1	Can articulate virtually all the sounds of the target language; some features of accent retained from other language(s) may be noticeable, but they do not affect intelligibility at all.
C2	22	695	B2	B2+	C1	B2+	C2	Any features of accent retained from other language(s) do not affect intelligibility or effective conveyance and enhancement of meaning. Intelligibility and effective conveyance and enhancement of meaning are not affected in any way by features of accent that may be retained from other language(s).
B2	6	680	B2	B2	B2	B2	B2	Can generalise from his/her repertoire to predict the phonological features of most unfamiliar words (e.g. word stress) with reasonable accuracy (e.g. whilst reading).
B2	25	703	B2	B2	B2	B2	B2	Can generally use appropriate intonation, place stress correctly and articulate

Intended			Phase 2 Judging Classic %	Phase 2 Judging Rasch Anchored	Phase 3 Assessing Bookmark Method	Phase 3 Assessing Rasch anchored	DECISION	
								individual sounds clearly.
B2	31	708	B1-C1 = B2	B2	-	-	B2	Can recognise common words when pronounced in a different regional variety from the one(s) he/she is accustomed to.
B1	33	710	B2	B2	-	-	B2	Can recognise the presence of regional varieties of pronunciation, though this may hinder comprehension. Can recognise when his/her comprehension difficulty is caused by a regional variety of pronunciation.
B2	1	679	B2	B2	B2	B1+	B2	Can articulate a high proportion of the sounds in the target language clearly in extended stretches of production; is intelligible throughout, despite a few systematic mispronunciations.
B1	10	690	B1+	B1+/B2	(B2)	(B2)	B1+	Can approximate common prosodic features of the target language in longer utterances (e.g. rising intonation for open questions), in order to convey the appropriate meaning.
B2	13	689	B1+	B1+	B1	B1+	B2	Can generally employ prosodic features (e.g. stress, intonation, rhythm) to support the message he/she intends to convey, though with some noticeable influence from other languages he/she speaks.
B2	21	698	B1+	B1+	B2	B1+/B2	B2	Accent tends to be influenced by other language(s) he/she speaks, but has little or no effect on intelligibility.
B1	3	681	B1+	B1+	B1	B1+	B1+	Can articulate the majority of the sounds of the target language reasonably clearly in extended speech.
B1	7	682	B1	B1	B1	A2+/B1	B1	Is generally intelligible throughout, despite regular mispronunciation of individual sounds and words he/she is less familiar with.
B1	27	704	B1	B1	B1	B1	B1	Pronunciation is generally intelligible; can approximate intonation and stress at both utterance and word levels.
B1	11	691	B1	B1	A2	A2	B1	Can convey the main point of his/her message in an intelligible way in spite of a strong influence on stress, intonation and/or rhythm from other language(s) he/she speaks.
B1	20	699	A2/A2+	A2+	-	-	B1	Accent is generally influenced by other language(s) he/she speaks, and this may sometimes occasionally affect intelligibility.
A2	16	692	A2	A2+	A2 (MS: A1)	A2	A2	Can use the prosodic features of everyday words and phrases intelligibly, in spite of

Intended			Phase 2 Judging Classic %	Phase 2 Judging Rasch Anchored	Phase 3 Assessing Bookmark Method	Phase 3 Assessing Rasch anchored	DECISION	
								a very strong influence on stress, intonation and/or rhythm from other language(s) he/she speaks.
A2	18	694	A2	A2+	A2	A2+	A2	Prosodic features (e.g. word stress) are appropriate adequate for familiar, everyday words and simple utterances.
A2	8	684	A2	A2	(A1)	(A1)	A2	Pronunciation is generally intelligible when communicating in simple everyday situations, provided the interlocutor makes an effort to understand specific sounds.
A2	9	683	A2	A2	A2	A2	A2	Systematic mispronunciation of phonemes does not hinder intelligibility, provided the interlocutor makes an effort to recognise and adjust to the influence of the speaker's language background on pronunciation.
A1	26	701	A1/A2	A2	A1	A1	A1	Can reproduce correctly a limited range of sounds as well as the stress on simple, familiar words and phrases.
A1	30	707	A1-B1	A2/A2+	-	-	A1	Can recognise and reproduce sounds in the target language correctly if carefully guided. Can reproduce sounds in the target language if carefully guided.
A2	19	700	A1/A2	A2	A2	A2	A2	A very strong influence from other language(s) he/she speaks on stress, rhythm and intonation may affect intelligibility, requiring collaboration from interlocutors. Nevertheless, pronunciation of familiar words is clear.
A1	34	701	A1/A2	A1/A2	-	-	A1	Can reproduce correctly a limited range of sounds as well as the stress on simple, familiar words and phrases.
A1	15	693	A1	A1	A1	A1	A1	Can use the prosodic features of a limited repertoire of simple words and phrases intelligibly, in spite of a very strong influence on stress, rhythm, and/or intonation from other language(s) he/she speaks; his/her interlocutor needs to be collaborative.
A1	2	685	A1	A1	-	-	A1	Can articulate a limited number of sounds, so that speech is only intelligible if the interlocutor provides support (e.g. by repeating correctly and by eliciting repetition of new sounds).
A2	29	706	A1-C1	dropped	-	-	DROP	Can recognise and reproduce in the target language sounds that are found in other language(s) he/she speaks, or that have been explicitly practised.

Appendix 6 – Descriptor scales – after analysis and discussion

	OVERALL PHONOLOGICAL CONTROL	SOUND RECOGNITION AND ARTICULATION	PROSODIC FEATURES
C2	Can employ the full range of phonological features in the target language with a high level of control – including prosodic features such as word and sentence stress, rhythm and intonation – so that the finer points of his/her message are clear and precise. Intelligibility and effective conveyance and enhancement of meaning are not affected in any way by features of accent that may be retained from other language(s).	Can consciously incorporate relevant features of regional and socio-linguistic varieties of pronunciation appropriately. Can articulate virtually all the sounds of the target language with clarity and precision.	Can exploit prosodic features (e.g. stress, rhythm and intonation) appropriately and effectively in order to convey finer shades of meaning (e.g. to differentiate and emphasise).
C1	Can employ the full range of phonological features in the target language with sufficient control to ensure intelligibility throughout. Can articulate virtually all the sounds of the target language; some features of accent retained from other language(s) may be noticeable, but they do not affect intelligibility at all.	Can recognise features of regional and socio-linguistic varieties of pronunciation and consciously incorporate the most prominent of them in his/her speech. Can articulate virtually all of the sounds of the target language with a high degree of control. He/she can usually self-correct if he/she noticeably mispronounces a sound.	Can produce smooth, intelligible spoken discourse with only occasional lapses in control of stress, rhythm and/or intonation, which do not affect intelligibility or effectiveness. Can vary intonation and place sentence stress correctly in order to express precisely what he/she means to say.
B2	Can generally use appropriate intonation, place stress correctly and articulate individual sounds clearly. Accent tends to be influenced by other language(s) he/she speaks, but has little or no effect on intelligibility.	Can recognise common words when pronounced in a different regional variety from the one(s) he/she is accustomed to. Can articulate a high proportion of the sounds in the target language clearly in extended stretches of production; is intelligible throughout, despite a few systematic mispronunciations. Can generalise from his/her repertoire to predict the phonological features of most unfamiliar words (e.g. word stress) with reasonable accuracy (e.g. whilst reading).	Can employ prosodic features (e.g. stress, intonation, rhythm) to support the message he/she intends to convey, though with some influence from other languages he/she speaks.
B1	Pronunciation is generally intelligible; can approximate intonation and stress at both utterance and word levels. Accent is generally influenced by other language(s) he/she speaks, and this may occasionally affect intelligibility.	<i>Can articulate the majority of the sounds of the target language reasonably clearly in extended speech. (B1+) NOT USED</i>	<i>Can approximate common prosodic features of the target language in longer utterances (e.g. rising intonation for open questions), in order to convey the appropriate meaning. (B1+) NOT USED</i>
		Can recognise when his/her comprehension difficulty is caused by a regional variety of pronunciation. Is generally intelligible throughout, despite regular mispronunciation of individual sounds and words he/she is less familiar with.	Can convey his/her message in an intelligible way in spite of a strong influence on stress, intonation and/or rhythm from other language(s) he/she speaks.
A2	Pronunciation is generally clear enough to be understood, but conversational partners will need to ask for repetition from time to time. A strong influence from other language(s) he/she speaks on stress, rhythm and intonation may affect intelligibility, requiring collaboration from interlocutors. Nevertheless, pronunciation of familiar words is clear.	Pronunciation is generally intelligible when communicating in simple everyday situations, provided the interlocutor makes an effort to understand specific sounds. Systematic mispronunciation of phonemes does not hinder intelligibility, provided the interlocutor makes an effort to recognise and adjust to the influence of the speaker's language background on pronunciation.	Can use the prosodic features of everyday words and phrases intelligibly, in spite of a strong influence on stress, intonation and/or rhythm from other language(s) he/she speaks. Prosodic features (e.g. word stress) are adequate for familiar, everyday words and simple utterances.
A1	Pronunciation of a very limited repertoire of learnt words and phrases can be understood with some effort by interlocutors used to dealing with speakers of his/her language group. Can reproduce correctly a limited range of sounds as well as the stress on simple, familiar words and phrases.	Can reproduce sounds in the target language if carefully guided. Can articulate a limited number of sounds, so that speech is only intelligible if the interlocutor provides support (e.g. by repeating correctly and by eliciting repetition of new sounds).	Can use the prosodic features of a limited repertoire of simple words and phrases intelligibly, in spite of a very strong influence on stress, rhythm, and/or intonation from other language(s) he/she speaks; his/her interlocutor needs to be collaborative.

www.coe.int

The Council of Europe is the continent's leading human rights organisation. It comprises 47 member states, 28 of which are members of the European Union. All Council of Europe member states have signed up to the European Convention on Human Rights, a treaty designed to protect human rights, democracy and the rule of law. The European Court of Human Rights oversees the implementation of the Convention in the member states.