



HAL
open science

EVALUATING THE ESSENTIALS: THE PLACE OF PROSODY IN ORAL PRODUCTION

Dan Frost, Jean O'Donnell

► **To cite this version:**

Dan Frost, Jean O'Donnell. EVALUATING THE ESSENTIALS: THE PLACE OF PROSODY IN ORAL PRODUCTION. Volín, Jan; Skarnitzl, Radek. The Pronunciation of English by Speakers of Other Languages, chapter 12, Cambridge Scholars Publishing, pp.228-259, 2018, 978-1-5275-0390-8. hal-02085252

HAL Id: hal-02085252

<https://hal.science/hal-02085252>

Submitted on 30 Mar 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CHAPTER 12

EVALUATING THE ESSENTIALS: THE PLACE OF PROSODY IN ORAL PRODUCTION

DAN FROST & JEAN O'DONNELL

12.1 Introduction

As in many countries, university students in France study a foreign language no matter what they major in. In France, this is compulsory for all students, most of whom study English, and conditions vary a great deal from one context to another (Taillefer, 2002). Studies such as EPTiES¹ (Henderson et al., 2012) have shown that pronunciation teaching is often neglected by both teachers and by teacher trainers. In France, as we shall see, there is a great need to work on certain aspects of pronunciation.

The French Ministry of Education and Research has specified the levels which it expects learners to achieve in their foreign languages at various stages of their education, and for LSP/LAP² learners, this means B2 according to the CEFRL³ on arrival at university (Goullier, 2005: 38). In fact, most students arrive after eight–ten years of secondary schooling with a level that is closer to A2 (Taillefer, 2007; Macré, 2015). The Innovalangues project⁴ is an attempt to address this issue: its primary mission is to help to bring the levels of LSP/LAP learners to a certified B2 level as defined by the CEFRL over the three years of their university career and the main thrust of the project is the creation of a “digital ecosystem” based around the platform Claroline Connect (Masperi & Quintin, 2014). The Innovalangues project is a six-year project which started in 2012 and is coordinated by Monica Masperi at Université Stendhal (Grenoble 3). It is composed of a team of about 60 teachers, researchers,

¹ English Pronunciation Teaching in Europe Survey.

² Languages for Specific Purposes / Languages for Academic Purposes.

³ Common European Framework for Reference in Languages (COE, 2001).

⁴ <http://innovalangues.fr/>

resource developers, IT developers and administrative staff, some of whom are employed full-time and some part-time to work on the project paid for with substantial funding from the ANR (*Agence Nationale pour la Recherche*) – the research funding body of the French Ministry of Education and Research. Initially, the target learners are students in Grenoble, but the project will also reach learners nationally and internationally and already has several national & international partners. The project contains several teams, and the subject of this chapter is part of the work of one of those teams: *THEMPPO*⁵ (Frost & Picavet, 2014). For reasons we shall see in the next part, *THEMPPO* was created to work specifically on the prosody of English on two fronts: firstly, the development of teaching resources (tools, media and activities) for use within the digital eco-system as part of a blended learning programme and secondly, a series of teacher-training seminars. The approach adopted by the team is essentially an articulatory approach (Honikman, 1964), especially towards the beginning of each course or learning path. We believe that as pronunciation is a complex set of physical gestures and as prosody is so iconic (Pennington, 1996: 137) and physical (Messum, 2009), work must begin with the body and the articulators. After “raising awareness to the articulatory settings” of English, constant reference is made to the body and articulators and various kinesthetic techniques are often used to reinforce the acquisition of stress patterns, etc. The pedagogical approach and the tools and activities which we have developed are presented elsewhere (Frost & Picavet, 2014; Frost & Guy, forthcoming), and are not the main focus of the present work. This chapter concentrates solely on the conception and calibration of one of those tools, i.e., the prosody-based descriptors for assessing oral production in English.

The descriptors are based on the CEFRL scales, which are, for various laudable political, pedagogical and linguistic reasons, plurilinguistic and based on communicative competence. We are aware, however, that our starting point is a tool which is neither language-specific nor form-based. We have developed an assessment tool which is similar in *form* to the CEFRL descriptors but is concerned only with “phonological control”, i.e., the subjects’ accent in English. In undertaking the development of this tool, we are attempting to address two main research questions: Firstly, is it feasible to peg such a tool to the CEFRL levels? And secondly, is such a tool useful, for teaching, for assessment and for research? In the next section, we will examine issues surrounding our choice to focus on

⁵ *THEMatique Prosodie et Production Orale*.

prosody, including the differences between English and French prosody, language transfer, pronunciation in language instruction and relevant assessment issues. In the following section, we will describe the conception and calibration of the descriptors and finally, we will present the results of a preliminary pilot study and discuss some of the implications.

12.2 Theoretical background

12.2.1 English and French prosody

French and English, despite sharing much of their vocabulary, are extremely different both phonetically and phonologically, particularly in the domain of prosody (see Frost, 2011 for a more complete contrastive analysis). What interests us particularly as teachers of English is how prosody is used to mark stress, both at the word level and at the level of the tone unit (TU). It is of course unrealistic to separate intonation and prominence in natural speech, but in order to identify some of the problems faced by French learners of English, let us look first at issues concerned more with intonation, then prominence, and finally segmental questions. The prosodic and segmental features which are treated in this section explain to a large extent our choices in developing the prosody-based descriptors which we will present later in the chapter.

Intonation is very closely related to prominence, especially in English, and the range between the high and low points for an average native speaker (NS) of English is greater than for an average French speaker (see Campione & Véronis, 1998 for a comparison of F0 range across five European languages). Across larger prosodic units, in particular tone units, a French learner of English typically has flatter intonation patterns than an English NS, with a step-up or a step-down at the end of a TU, as we will see later.

Regarding stress, English and French are very different. French does not have lexical stress and it exhibits relative prominence to a lesser degree than in English (Rossi, 1979). To compound difficulties for French NSs, this “relative prominence” is fixed, i.e., it is usually present only on the last syllable of prosodic units (Dahan & Bernard, 1996). These units may be shorter “stress groups” (Di Cristo, 1998) or longer “accentual phrases” (Jun & Fougeron, 1995). This final prominence, marked by a lengthening and often a fall in F0, may be explained by articulatory features, such as there simply being less acoustic energy at the end of a breath group. Wenk and Wioland (1982: 204) therefore describe French as

“trailer-timed” and English as “leader-timed”. This leads to an alternation between strong (prominent) syllables and weak (reduced) syllables, which is what characterises the rhythm of English. The concept of isochrony (Pike, 1945) has been questioned by many authors (see Bertinetto, 1989 for an overview), and is clearly too simplistic a model for describing the reality of natural speech, although many teachers and researchers use models inspired by Pike and Bolinger’s work even today (See Dickerson, 2015) and in our work, we have found that treating English as “stress-timed” and French as “syllable-timed” over shorter TUs can be useful for raising the awareness of learners and teachers to different metrical structures. There are many possible explanations for this, not least of which is the relative importance of the acoustic cues F₀, amplitude, duration and formant structure in English. F₀, perceived as pitch, is essential to producing and perceiving prominence in English, and it is this which led Bolinger (1958) to call the marking of stress in English “pitch prominence”. Prominence is key to segmenting the speech signal, and research points to the relative importance of F₀ compared to other cues in English compared to French (Frost, 2011; Coughlin & Tremblay, 2012). French, however, marks the group-final syllable mentioned above with an increased duration – this is not to say that amplitude and F₀ are not factors, but syllable-lengthening is the most salient feature (Benguerel, 1973; Di Cristo, 1998; Lacheret-Dujour & Beaugendre, 1999: 41; Jun & Fougeron, 2000; Astesano, 2001).

Finally, there are also many segmental differences between French and English. French has simple vowels, some of which are nasalised. French vowels are all tense and there are far fewer than in English – only 10–13 oral vowels and 3–4 nasal vowels (Fougeron & Smith, 1993) whereas English has a complex system comprising of lax short vowels, long tensing vowels, diphthongs and triphthongs. As for consonants, English has consonants which do not exist in French (/h/, /θ/ and /ð/) and there are many differences in the place and manner of plosives, fricatives and glides. We are, however, less concerned with these segmental differences apart from those which have a close relationship with the prosodic features of English, as we shall explain in the next part. Segmental features are important to this work for three reasons. Firstly, at the syllable level, unstressed syllables are often reduced in English. As Jenkins (2000: 147) points out, the weak/strong syllable alternation is a characteristic feature of all varieties of English. We believe this is at the heart of the production and comprehension problems that many learners, especially French NSs, have with English. Jenkins goes on to say that weak forms are “unteachable”, an assertion which is not backed up by research and which

we refute strongly. Secondly, we consider full, unreduced vowels to be important as they ‘carry’ the stress. It is therefore important to focus to some degree on the “correct” pronunciation of full vowels so as to differentiate them from reduced vowels. Thirdly, regarding consonants, we are essentially concerned only with phonotactic phenomena, particularly at word boundaries, as they are problematic for oral production and comprehension for many French NSs learning English.

The differences between French and English pronunciation for pedagogical purposes (i.e., the priorities which research has led us to establish) may be summed up in Table 12-1:

		English	French
Prosody	Rhythm & stress	<ul style="list-style-type: none"> • Lexical stress • F0 an important cue • + Stress-timed • Strongly marked nuclear stress • Final lengthening if nuclear stress is final 	<ul style="list-style-type: none"> • No lexical stress • F0 a less important cue • + Syllable-timed • Weakly marked nuclear stress • Evident final lengthening
	Intonation	<ul style="list-style-type: none"> • Large range • Smooth and varied contours throughout TUs 	<ul style="list-style-type: none"> • Narrow range • Step up / step down changes more frequent
Segmental	Syllables	<ul style="list-style-type: none"> • Very frequent reductions (mainly /ə/) 	<ul style="list-style-type: none"> • Reductions are rare
	Vowels	<ul style="list-style-type: none"> • + Lax • Complex (short and long simple vowels, diphthongs, triphthongs) 	<ul style="list-style-type: none"> • + Tense • Simple vowels only
	Consonants	<ul style="list-style-type: none"> • Deletion, assimilation, etc. very frequent 	<ul style="list-style-type: none"> • Deletion, assimilation, etc. less frequent

Table 12-1. Some differences between English and French pronunciation.

12.2.2 Language transfer: Perception, pronunciation & prosody

The differences between the phonetics and phonology of French and English which we outlined in the previous part are relevant to language teaching because of the phenomena of language transfer, or interference (Weinrich, 1953). Although it is not within the scope of this chapter to provide a full review of language transfer and L2 phonology (for a critical review, see, e.g., Major, 2008), we would like to highlight a few points which are relevant to the current study. Firstly, some of the very first work on transfer was done in an attempt to understand issues related to pronunciation and perception. The Prague Circle was already aware of “phonological deafness” (Polivanov, 1931; Trubetzkoy, 1939) well before the instruments of modern acoustic phonetics were able to measure such perceptual phenomena. Secondly, the degree to which negative transfer may be problematical depends on the languages concerned. The differences between English and French prosody are substantial, as we saw in the previous section. Lado’s “contrastive analysis hypothesis” (1957) suggests that the greater the difference between a language feature in a learner’s L1 and the target language, the harder it will be to learn and Eckmann’s “markedness differential hypothesis” (1977) supports this idea. Thirdly, following Selinker’s definition of interlanguage (1972) and Corder’s work on learner errors (1981), many teachers chose to accept the effects of language transfer on pronunciation, because fluency was the emphasis in the communicative approach which dominated English teaching for this period. The integration of interlanguage into Krashen’s “Monitor Model” (1981) was an example of this. Fourthly, many researchers have preferred to focus on segmental features. For example, Flege’s “Speech Learning Model” (Flege, 1995; Flege, Schirru & MacKay, 2003) and MacWhinney’s “Unified Model” (MacWhinney, 2008) devote considerable importance to the impact of L1 phonology on the production and perception of other languages, but as far as our purposes are concerned, do not place enough emphasis on prosody. Finally, however, recent work on transfer has focused more on prosody, as teachers and researchers are increasingly concerned with intelligibility and teaching pronunciation for comprehension. Dolbec and Santi refer to a “linguistic filter” (1995: 46) and Dupoux and Peperkamp and their colleagues in Paris identified stress in English as being problematical for French and Spanish NS learners coining the term “stress deafness” (Dupoux & Peperkamp, 1999; Dupoux, Peperkamp & Sebastien-Galles,

2001; Peperkamp & Dupoux, 2002). Furthermore, research also shows that prosody is one of the most stubborn features of a learner's L1, even with proficient language users (Bailey, Plunket & Scarpa, 1999; Flege, Schirru & MacKay, 2003; Gabriel & Kireva, 2014). As we have seen in the previous section, the differences between English and French prosody are both numerous and considerable, and in the next section, we shall examine the implications of L1 transfer for our pronunciation instruction and our pedagogical choices.

12.2.3 Pronunciation and prosody instruction

The importance given to teaching pronunciation has waxed and waned over the years, with little emphasis on oral language skills at all in formal education before the direct methods of the early 20th century. Behaviourist ideas and the development of analogue language laboratories saw an expansion of repetitive drilling and a focus on accuracy in pronunciation. The “Army method” which drew heavily on behaviourist ideas developed by Skinner led to the audio active comparative drills used in schools and universities throughout the world for decades. More recently, pronunciation was neglected throughout the nineties and early part of this century during the hegemony of the communicative approach, with authors such as Judy Gilbert referring to pronunciation as the orphan of L2 teaching (Gilbert, 2010), but things are changing now. Certainly the Web as we know it today with much user-generated media-rich content, downloading and streaming of music, films and TV series involves more active engagement with audio and video than ever before. Many authors and teachers have been pushing for the integration of pronunciation into English language teaching (ELT) (Kjellin, 1999b; Henderson, 2008; Gilbert, 2008, 2010; Derwing, 2010; Munro & Derwing, 2015). Moreover, the existence of conferences such as EPIP (English Pronunciation Issues and Practices) and the inauguration of the *Journal of Second Language Pronunciation* in 2015 are evidence of this renewed interest. There has, however, been a major shift in the goals of pronunciation instruction over the last decade: the more pragmatic targets of intelligibility and comprehension are now central to the work of many teachers and researchers (Zielinski, 2006; Munro & Derwing, 2011, 2015; Harding, 2012). This is one of the key factors in our choice to focus on prosody in our pedagogical approach. All too often, teachers in France focus on features such as /θ/ and /ð/ which do not impair intelligibility or impinge on comprehension to the same degree – after all, Irish speakers of English realise inter-dental fricatives as alveolar plosives (Hickey, 2004) and -th-

in parts of London and elsewhere is often pronounced as /f/ or /v/ (Kerswill, 2006) without causing communication problems. Research shows that prosody is important for the perception of a foreign accent (Jilke, 2000) but more importantly, it is one of the major factors contributing to the intelligibility and comprehension of L2 speakers (Hahn, 2004; Munro & Derwing, 2011; Piske, 2012; Acton, Baker, Burri & Teaman, 2013).

There are as many different ways of learning and teaching pronunciation as there are learning situations, and various studies have demonstrated the effectiveness of teaching pronunciation (see for example Scarcella & Oxford, 1994 and Saito, 2012 for overviews). As described in the introduction to this chapter, the context of this work is courses of 20 to 60 hours per year with undergraduate students or adults who are not language specialists. On such short courses with non-specialists, our work and the work of others (Derwing & Munro, 1998, 2005; Kjellin, 1999b; Celik, 2001; Murphy, 2004) has shown that focusing on prosody is an efficient use of limited time. In the French context, other authors have chosen to work on prosody (Cooke, 1993; Herry, Nishinuma & Ghio, 2003; Stenton, 2011; Horner, 2013, 2014) for the same reasons as we describe. The realisation of prosodic features, particularly stress, involves increased acoustic energy and physical effort. We therefore consider it extremely important to work on the body, breathing and the articulators. This view is shared by other teachers and researchers interested in prosody (Kjellin, 1999a, 1999b; Borrell & Salsignac, 2002; Messum, 2009; Soulaïne, 2014). As we mentioned in the introduction, our approach therefore begins with raising awareness and training of the body and the voice before progressing to freer production tasks. In the context of the *Innovalangues* project, this work is carried out not only in classrooms, but also in an online environment (mentioned in the introduction) using video recordings and playback tools and a variety of tasks. If we are to put prosody at the centre of our teaching, then we need a tool to define objectives, assess progress and allow learners to assess their own progress. As we shall see in the next part, however, there is a paucity of such tools in language teaching.

12.2.4 Pronunciation and prosody instruction

In the French university system when it comes to assessing what is referred to as either spoken language, speaking, oral production, speaking proficiency, fluency, speaking skills, oral expression, communicative language competence or speech, students are generally graded on a scale

ranging from 0 (lowest grade) to 20 (highest grade) with 10 representing a pass. Grades on oral exams are usually based on in-house scales devised either by individual teachers or small teaching teams all of whom have a great deal of freedom regarding teaching and assessment practices. Lack of a common reference frequently leads to incoherence of grades obtained by students not only within a given course but also from one course to another. The CEFRL (see Table 12-2) is rarely actually used as an in-house tool to assess students' oral competence, indeed only one of the 32 scales within the CEFRL is devoted to pronunciation and is entitled "phonological control" (Council of Europe, 2001: 117). Despite this fact, Henderson et al. (2012) found that it was the most frequently quoted set of scales by teachers when asked what, if any, tools they used to assess pronunciation.

	PHONOLOGICAL CONTROL
C2	As C1
C1	Can vary intonation and place sentence stress correctly in order to express meaning.
B2	Has acquired a clear, natural, pronunciation and intonation.
B1	Pronunciation is clearly intelligible even if a foreign accent is sometimes evident and occasional mispronunciations occur.
A2	Pronunciation is generally clear enough to be understood despite a noticeable foreign accent, but conversational partners will need to ask for repetition from time to time.
A1	Pronunciation of a very limited repertoire of learnt words and phrases can be understood with some effort by native speakers used to dealing with speakers of his/her language group.

Table 12-2. CEFRL "phonological control" descriptors (COE, 2001: 117).

These "phonological control" descriptors can be criticised on several fronts. For example the use of the terms "*clear and natural*" at level B2 tends to reinforce the perception of the model of the native speaker. This runs contrary to what is stated in the global scales with respect to C2, the highest level attainable in the CEFRL global scales, defined as "the degree of precision, appropriateness and ease with the language which typifies the speech of those who have been highly successful language learners" (Council of Europe, 2001: 36). Harding (2013) points out that the use of vague terminology can be interpreted differently by different assessors, for example "clear enough to be understood", "can be understood with some effort", "foreign accent is sometimes evident", etc.

The issue of the acquisition and therefore the assessing of language competence existing on a continuum is a much-debated topic (Fulcher, 2004). Horner (2014) questions whether or not the acquisition of phonological control can be considered as a hierarchy. When for example should intonation be taught? At what stage should word stress be introduced? Unfortunately, the lack of research in general, and more specifically regarding French learners of English, means we cannot provide definite answers to such questions. Thus scales with ascending levels are not an interpretation of SLA findings as there are as many routes to learn a language as there are learners (Luoma, 2004). This means that creating scales is a real challenge and scales will reflect the designer's beliefs, experience and understanding of the learning process. An example of a re-working of the phonological control scales is proposed by Horner (2014), who tentatively suggests a new grid for pronunciation which is more holistic than analytic. He insists on the importance of intelligibility and is strongly inspired by the CEFRL, thus he accepts to a certain degree the construct of a hierarchy.

Despite the criticism levelled at the CEFRL “phonological control” descriptors, they can be considered as a starting point in attempting to assess pronunciation. As our primary concern is assessing prosody and due to the fact there are very few methods available to do so – the PEPS-C test (Peppé & McCann, 2003) is a rare example, but designed for native speakers with autism and Asperger's – the CEFRL thus paved the way for the development of the prosody descriptors presented in this chapter.

12.3 Designing the descriptors

The prosody descriptors take account of pronunciation factors not for the mere sake of correct pronunciation according to some native model or other, but with the aim of improving intelligibility and comprehension. They were developed mainly for assessment purposes but also as a tool for raising learner and teacher awareness and for structuring learning objectives, and they came into being because no other suitable tool existed. During the ELLO project (Frost & O'Donnell, 2013), we used the CEFRL scales of descriptors for speaking to assess hundreds of subjects over a three-year period. The subjects also assessed their own performances using the same scales. We found it quite constraining that the scales didn't address the particular prosody-related language-specific issues which are, as we have seen, not only responsible for intelligibility problems, but also for problems understanding spoken English. Indeed the subjects often mentioned pronunciation in the surveys and interviews we

conducted. In devising the first set of descriptors, we were therefore thinking primarily of assessment, both by teachers and by learners themselves, but also of the pedagogical value of such a tool. The primary stated goal of the CEFRL is to improve practices – assessment is just a means to an end and the scales themselves encourage the various stakeholders “to reflect on their current practice” and, amongst other things, provide a common basis for “the elaboration of language syllabuses”, (COE, 2001: 1). As we have also seen, prosody is one of the key factors in intelligibility and comprehension problems, especially in English and especially for French learners. Although we refer to the descriptors as “prosody descriptors”, segmental features are included. It is not therefore a question of segmental features *or* prosody: segmental features are therefore dealt with in a way that makes them *secondary* to prosody. This is why there is a column for reduced syllables, one for full vowels and one for connected speech phenomena, but not a column simply for the correct realisation of consonants such as /θ/ and /ð/.

The development of the current prosody descriptors is based on the conviction that prosody can be described analytically rather than holistically. In our approach, prosody and its closely linked segmental features are broken down into components, namely rhythm and stress, intonation, syllables, vowels and consonants. In keeping with the CEFRL and on account of our experience as teachers and researchers we acknowledge a certain hierarchy, although not universal or totally rigid, in the progression of the acquisition of the above prosodic features. Whereas the CEFRL descriptors are not language-specific, the tool which we have devised is specific to English and has been calibrated for use by French learners. The implications of this choice will be discussed later in this chapter. The prosody descriptors are calibrated to the CEFRL oral production descriptors, but it is of course the case that a learner may have relatively poor phonological control, but good grammatical and lexical control or vice versa. Consequently, the “level” which the prosody-based descriptors permit a user to attribute to a given performance for a given subject may or may not correspond exactly to overall oral performance as determined by the CEFRL descriptors for speaking. In this section, we will describe the choices we made and the procedures we adopted for the design and calibration of the descriptors. The prosody descriptors and corresponding assessment sheets were initially designed intuitively based on twenty years of teaching experience, observations and assessment. The descriptors then underwent a period of revision, calibration and piloting that has lasted two years. In line with the CEFRL guidelines, the grid and assessment sheet were designed to be as positive, clear, brief and

independent as possible. We also set out to ensure that they would be “flexible”, “dynamic”, “user-friendly”, and “non-dogmatic” (COE, 2001: pp. 7–8). “Can do” statements were used to ensure the overall perspective of the grid would be perceived positively by students and teachers alike.

The starting point for the calibration process was a set of “full” prosody descriptors containing eight columns (Appendix 3) – four with quantitative criteria (QTT) and four with qualitative criteria (QLT). The original eight columns were as follows:

1. Word Stress (QTT)
2. Pause & Tone Units (QLT)
3. Focus (Nuclear & Contrastive Stress) (QLT)
4. Rhythm (number of beats) (QTT)
5. Reductions (QTT)
6. Stressed & Unstressed Vowel (QLT)
7. Connected Speech (QTT)
8. Intonation (QLT)

The full set of descriptors with eight columns was then calibrated in three stages. Firstly, there was a reading task, secondly two spontaneous speech tasks and thirdly, a preliminary pilot study. As we shall see below, to make the tool easier to use, the final version of the descriptors contains only five columns (see Appendices 1 and 2).

12.3.1 Reading task

It was decided to begin with a reading text because in this way we could control for all the features which interest us; although the prosody and the phonotactic phenomena of spontaneous speech would be more interesting, it would be hard to compare like with like. A text – “The Mallory text” – containing all the prosody features in the descriptors was selected, then read and recorded by two native North American speakers and two native British speakers. Using an orthographic transcription, the prosody features were annotated by three native speakers to obtain a harmonised annotation.

Some of the features were obligatory, such as correctly placing word stress, but some of the features were optional: the number of phonotactic phenomena realized at word boundary level, for example, depends on speaker style, speed of delivery, etc.

The calibration using the reading task was carried out using the full set of prosody descriptors, i.e., all eight columns mentioned above. The

annotated text broke down into twenty-eight individual potential tone units (TUs). For each of these TUs, the number of potential occurrences of the four QTT features was calculated (yielding between 0 and 117 possible “scores” per feature/column). For the QLT features a possible score (0, 1 or 2) was decided upon (yielding between 0 and 50 possible “scores” per feature/column). (See Appendix 3 for the calibration sheet for the reading task).

Three representative subjects’ recordings (A2, B1 and B2+) were then selected. The overall oral production level of these speakers had previously been obtained by double expert assessment and self-assessment using the CEFRL overall speaking descriptors. These three speakers then recorded the Mallory text. Two raters spent three days listening to and “scoring” each of the three recordings TU by TU. The total number of scores for each of the eight prosody features per subject was obtained independently by each of the raters, then discussed (with further listening when necessary) in order to reach a consensus for each subject’s “scores”. Based on the results of the discussions the descriptors underwent several adjustments. “Can do” statements were reworded, moved from one level to another, others were added or omitted and a simplified “basic” version of the descriptors was adopted, which evolved constantly over the next few weeks until the current version, version 15. (See Appendix 1 for the full set of descriptors, and Appendix 2 for the accompanying assessment sheet).

12.3.2 Spontaneous speech task

The reading task was a useful way of controlling for most of the features which interested us, but a prosody assessment needs to take account of spontaneous speech. As the reading text contained no interaction, the intonation was very repetitive and the number of phonotactic phenomena was quite limited. We therefore decided on two spontaneous tasks: one monologue and one interaction. Two sets of video recordings, (i.e., monologues and interactions) from the *ELLO* project were used in the second stage of the calibration process, both of which were obtained following the WebCEF project protocol (Bijnens, 2009). The monologues involved subjects describing a 30-second television advertisement (two minutes) and the interactions involved the students chatting about their mobile phones (five minutes). The two raters (watched and) listened to twenty subjects’ monologues and interactions independently (forty recordings in all) several times without knowing the subjects’ CEFRL level. They used the “basic” descriptors to give each a “prosody level”.

This was followed by a discussion and harmonisation phase. A log of issues, solutions and comments was kept. For each of the twenty monologues and interactions the “prosody level” obtained was compared to the overall CEFRL speaking level previously obtained in the *ELLO* project. This led to further vertical realignment and verification of the descriptors as well as to the current set of “basic” descriptors along with its corresponding assessment sheet.

12.3.3 Pilot study

The third stage of the calibration, still ongoing, is the piloting of the descriptors. The first part of this process, the initial pilot study, involved contacting teachers, trainee teachers and students, fourteen in all, to test out the current version of the descriptors and corresponding assessment sheet on a video recording (the same task as used in the WebCEF and *ELLO* projects). The participants were also asked to reply to a 6-point, 22-item Likert scale questionnaire on the ease of use and usefulness of the descriptors. Based on the results of the pilot study, work is underway to develop a handbook and to road-test the descriptors in a variety of learning situations. These tests will result in further modifications to the descriptors and to the handbook and will lead to a tool which will be usable by teachers, trainers and students alike.

12.4 Results and discussion

12.4.1 Stage 1 calibration – Reading

Applying the descriptors to the reading of the “Mallory text” resulted in the subjects obtaining high scores across the board for “Word stress” and “Rhythm (number of beats)”. These two features were initially considered intuitively to be discriminating features at the different levels on the grid. Fig. 12-1 shows that all three subjects produced approximately 90% of all potential occurrences of words stress irrespective of their level. Similarly, all three subjects produced between 70% and 80% of all possible beats. This indicated that subjects were sensitive to the phenomena of word stress and rhythm but marked both to varying degrees in their speech. Consequently, in order to differentiate between the three subjects’ levels, the concept of the “quality” (use of the acoustic cues to stress, i.e., amplitude, duration, F0 curve and formant structure) or how well the words were stressed thus replaced the “quantity” (i.e., stressed or not) to

differentiate the subjects' scores. The original eight columns were merged into five columns (see Appendix 1):

1. Rhythm and stress
2. Reduced syllables
3. Stressed and unreduced vowels
4. Connected speech
5. Intonation

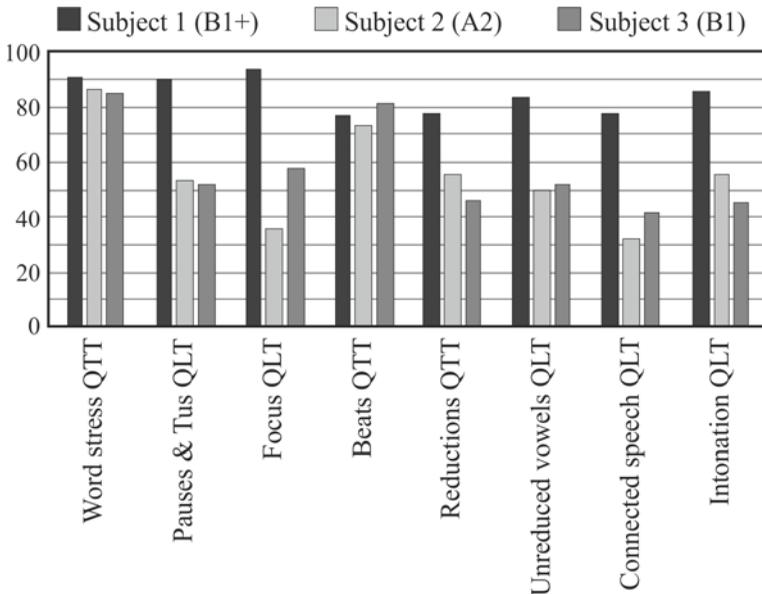


Fig. 12-1. 8 criteria for measuring prosody profiles (reading scores/100).

12.4.2 Stage 2 calibration – Spontaneous speech

This stage involved the spontaneous speech tasks (a monologue and an interaction) and actually comprised two steps: the first with all twenty subjects, and then some “fine tuning” based on the results of three selected subjects. Firstly, the average levels for the twenty subjects were considerably higher for both monologues and interactions when assessed with the prosody descriptors than when assessed using the CEFRL scales. Fig. 12-2 shows that the average level (twenty subjects) for the monologue and interaction using the CEFRL was A2+ whilst with the prosody

descriptors the average level was B1+. It was therefore clear that several of our descriptors were too “generous”. This led to further adjustments in wording and vertical realignment, i.e., certain “can do” statements being moved down one or two levels. For example, in the pre-calibrated version “can usually place word stress correctly” was in B2, whereas in the post-calibrated version, this descriptor is at A2 level.

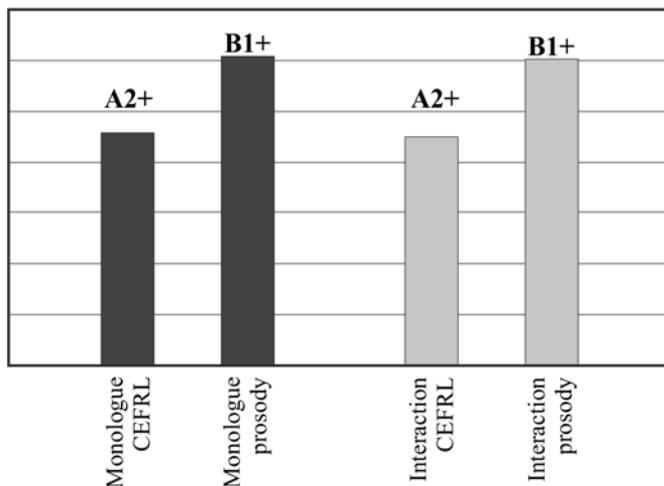


Fig. 12-2. Pre-calibration CEFRL and prosody levels for monologues and interactions (averages for all 20 subjects).

After this realignment and rewording, we looked at the individual subjects’ levels and identified three subjects from across the range of levels who had a marked discrepancy between their level using the CEFRL descriptors and their level using the prosody descriptors (subject 6 = B2/C1, subject 13 = A2/B2, subject 18 = A2/B2+ respectively). This enabled us to “fine-tune” the descriptors by making minor adjustments, especially in the “Rhythm and stress”, “Stressed and unreduced vowels” and “Phonotactics” columns.

If we compare Fig. 12-3 and Fig. 12-4 below we can see that the levels obtained after calibration were closer to the CEFRL levels. In other words, the level of a given subject’s oral performance regarding prosody using the present tool is closer to the subject’s overall oral production level obtained using the CEFRL descriptors. These final modifications resulted in the

current version (V15) of the prosody descriptors and the next stage was to let other users use them in a pilot study.

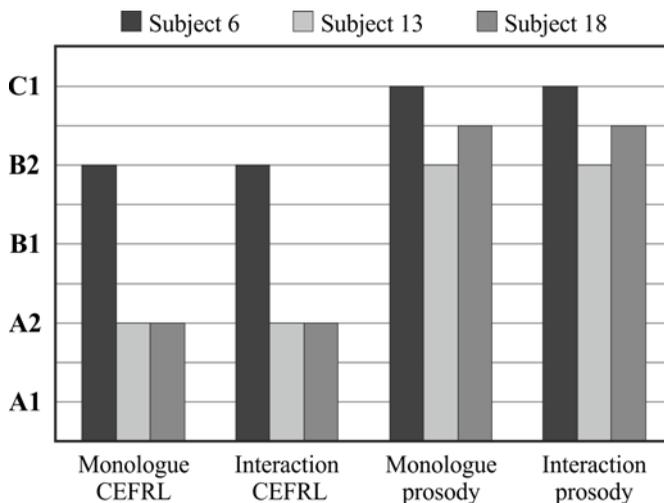


Fig. 12-3. CEFR and prosody levels for monologues and interactions before stage 2 calibration.

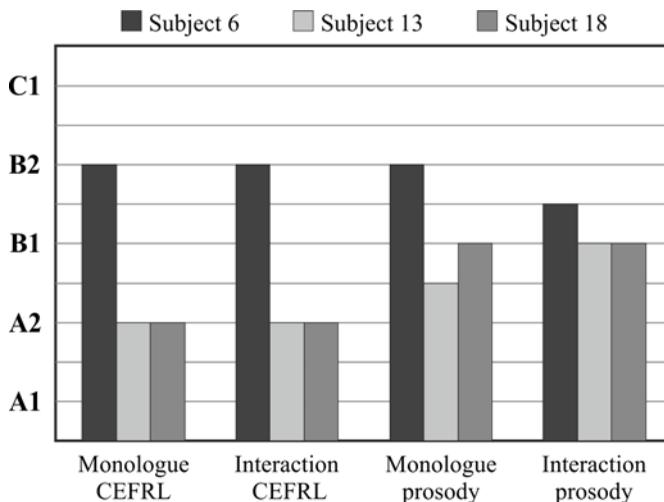


Fig. 12-4. CEFR and prosody levels for monologues and interactions after stage 2 calibration.

12.4.3 Initial pilot study

Firstly, all fourteen respondents were generally satisfied with the tool, and in the main, negative comments referred to context and training-related issues. The prosody descriptors were perceived as being both user-friendly and a good awareness-raising tool. More specifically, of the fourteen users questioned, ten of them agreed to varying degrees with the statement that the prosody descriptors were easy to use. Only four subjects found them difficult or very difficult to use. Suggestions were made that practice in using the descriptors was necessary as they referred to some concepts with which they were unfamiliar. As to the usefulness of the descriptors as an awareness-raising tool, thirteen respondents were in agreement to some extent and remarked that the descriptors enabled them to grasp the nuances between the otherwise unfamiliar elements of prosody. There was a consensus that stress was the easiest element to assess whilst connected speech was usually considered the most difficult. Finally, about half of those questioned stated that some training was required to use the tool or that a reminder of the basic technical terms and definitions related to phonology was necessary.

12.5 Conclusion

The research questions which we set ourselves were whether it was feasible to develop such a tool and to peg it to the CEFRL levels and whether such a tool would be useful for assessment, pedagogical and research purposes. We have shown that it is possible to develop a set of descriptors based on prosody and to peg them to the CEFRL levels, however the calibration process is never going to be entirely satisfactory, as it will inevitably involve subjectivity on the part of the assessor and because each individual learner has his or her own profile with different strengths and weaknesses. As the CEFRL document states: “A scale, like a test, has validity in relation to contexts in which it has been shown to work. Validation – which involves some quantitative analysis – is an ongoing and, theoretically never-ending, process.” (COE, 2001: 22). The content and calibration of these descriptors will continue based on the results of ongoing pilots in different contexts. The second research question mentions three areas where the descriptors may prove useful: assessment, setting learning objectives and research. As an assessment tool, the descriptors will be used by a number of teachers and learners both within the Innovalangues project and in other learning situations to assess

learners' needs and direct them towards sequences of remedial activities accordingly. These prosody-based descriptors are therefore, in the same way as the CEFRL descriptors, a tool for diagnostic, summative and formative assessment (*ibid* 186). The remedial activities are still being developed and will be piloted in the next year, but the descriptors have already proved valuable as they form a key part of the syllabus for developing learning objectives for oral production activities for the Innovalangues platform. Finally, on the question of research, Munro & Derwing (2015: 13) call for further defining of the construct: "Although much pronunciation-related research investigates accent, comprehensibility, and intelligibility, there has been far from perfect unanimity on how these constructs should be defined and operationalized." We believe that this tool contributes to defining some of the fundamental elements which are essential for intelligibility and comprehension in the context where this work is taking place.

References

Acton, W., Baker, A., Burri, M. & Teaman, B. (2013). Preliminaries to haptic-integrated pronunciation instruction. In: Levis, J. & LeVelle, K. (Eds.), *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference, August, 2012*, pp. 234–244. Ames, IA: Iowa State University.

Alazard, C., Astésano, C. & Billières, M. (2010). The Implicit Prosody Hypothesis applied to Foreign Language Learning: From oral abilities to reading skills. In: *Proceedings of the 5th Speech Prosody 2010*. Retrieved from <http://www.speechprosody2010.illinois.edu/papers/100648.pdf>.

Astésano, C. (2001). *Rythme et accentuation en français. Invariance et variabilité stylistique*. Paris: Editions L'Harmattan, Collection Langue et Parole.

Bailey, T., Plunket, K. & Scarpa, E. (1999). A cross-linguistic study in learning prosodic rhythms: Rules, constraints and similarities. *Language and Speech*, 42(1), pp. 1–38.

Benguerel, A-P. (1973). Corrélat physiologiques de l'accent en français. *Phonetica*, 27, pp. 21–35.

Bertinetto, P.M. (1989). Reflections on the dichotomy "stress" vs. "syllable-timing". *Revue de phonétique appliquée*, 91–93, pp. 99–130.

Bijnens, H. (2009). (Ed.) WebCEF. *Collaborative evaluation of oral language skills through the web*. Leuven: AVNet, K.U.Leuven.

Bolinger, D. (1958). A theory of pitch accent in English. *Word*, 14, pp. 109–149.

Borrell, A. & Salsignac, J. (2002). Importance de la prosodie en didactique des langues (application au FLE). In: Renard, R. (Ed.), *Apprentissage d'une langue étrangère/seconde*. Vol. 2, pp. 163–182. De Boeck Supérieur: Pédagogies en développement.

Campione, E. & Véronis, J. (1998). A statistical study of pitch target points in five languages. In: *Proceedings of the 5th International Conference on Spoken Language Processing (ICSLP'98)*, pp. 1391–1394.

Celik, M. (2001). Teaching English Intonation to EFL/ESL students. *The Internet TESL Journal*, 7/12. Retrieved on September 15, 2015, from <http://iteslj.org/Techniques/Celik-Intonation.html>

Coughlin, C. & Tremblay, A. (2012). The role of prosodic information in L2 speech segmentation. In: *Proceedings of the 35th Boston University Conference on Language Development (Supplement)*. Somerville, MA: Cascadilla Press. Retrieved from <http://www.bu.edu/buclid/files/2012/03/CoughlinTremblayBUCLDProceedingsFinal.pdf>

Cooke, R. (1993). Reducing word stress errors: Time restricted help for ESP students. *ASP*, 2, pp. 164–175.

Corder S.P. (1981). *Error Analysis and Interlanguage*. Oxford: Oxford University Press.

Council of Europe. (2001). *A Common European Framework of Reference for learning, teaching and assessment*. Cambridge: Cambridge University Press.

Cutler, A., Dahan, D. & Van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, 40(2), pp. 141–201.

Dahan, D. & Bernard, J.-M. (1996). Interspeaker variability in emphatic accent production in French. *Language and Speech*, 39/4, pp. 341–374.

Derwing, T. M. & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, 39, pp. 379–397.

Derwing, T. M., Munro, M. J. & Wiebe, G. (1998). Evidence in favour of a broad framework for pronunciation. *Language Learning*, 48/3, pp. 393–410.

Derwing, T. (2010). Utopian goals for pronunciation teaching. In: Levis, J. & LeVelle, K. (Eds.), *Proceedings of the 1st Pronunciation in Second Language Learning and Teaching Conference*, pp. 24–37. Ames, IA: Iowa State University.

Di Cristo, A. (1998). Intonation in French. In: Hirst, D. & Di Cristo, A. (Eds.), *Intonation Systems: A Survey of Twenty Languages*, pp. 3–45. Cambridge: Cambridge University Press.

Dickerson, W. (2015). A nail in the coffin of stress-timed rhythm. In: Levis, J., Mohammed, R., Qian, M. & Zhou, Z. (Eds.), *Proceedings of the 6th Pronunciation in Second Language Learning and Teaching Conference*, pp. 184–196. Ames, IA: Iowa State University.

Dolbec, J. & Santi, S. (1995). Effet du filtre linguistique sur la perception de l'accent: étude exploratoire, *Travaux de l'Institut de Phonétique d'Aix*, 16, pp. 41–60.

Dupoux, E. & Peperkamp, S. (1999). Fossil markers of language development: Phonological “deafness” in adult speech processing. In: Laks, B. & Durand, J. (Eds.), *Cognitive Phonology*, pp.168–190. Oxford: Oxford University Press.

Dupoux, E., Peperkamp, S. & Sebastian-Galles, N. (2001). A robust method to study stress-deafness. *Journal of the Acoustical Society of America*, 110(3), pp. 1606–1618.

Flege, J. E., Schirru, C. & MacKay, I. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication*, 40(4), pp. 467–491.

Fougeron, C. & Smith, C. L. (1993). Illustrations of the IPA: French. *Journal of the International Phonetic Association*, 23(2) pp. 73–76.

Fraser, H. (2006). Helping teachers help students with pronunciation: A cognitive approach. *Prospect: An Australian Journal of TESOL*, 21/1, pp. 80–95.

Frost, D. (2011). Stress and cues to relative prominence in English and French: A perceptual study. *Journal of the International Phonetic Association*, 41(1), pp. 67–84.

Frost, D. & O'Donnell, J. (2013). Combatting the “can't do mentality”: Expert, peer & self-assessment in a French university context. In: Colpaert, J., Simons, M., Aerts, A. & Oberhofer, M. (Eds.), *Proceedings of the 2nd International Conference “Language Testing in Europe: Time for a New Framework?”*, pp. 104–109. Antwerp: University of Antwerp.

Frost, D. & Picavet, F. (2014). Putting prosody first – some practical solutions to a perennial problem: The Innovalangues project. *Research in Language*, 12/2, pp. 1–11.

Frost, D. & Guy, R. (Forthcoming). L'innovation est le ton qui fait la chanson dans le secteur LANSAD: musique et prosodie dans le projet Innovalangues. *Recherche et pratiques pédagogiques en langues de spécialité – Cahiers de l'APLIUT*.

Fulcher, G. (2004). Are Europe's tests being built on an ‘unsafe’ framework? *Guardian weekly*, 18 March, 2004. Retrieved on September 15, 2015, from <http://www.theguardian.com/education/2004/mar/18/tefl2>

Gabriel, C. & Kireva, E. (2014). Prosodic transfer in learner and contrast varieties. Speech rhythm and intonation of Buenos Aires Spanish and L2 Castilian Spanish produced by Italian native speakers. *Studies in Second Language Acquisition*, 36(2), pp. 257–281.

Gervain, J. & Werker, J. (2013). Prosody cues word order in 7-month-old bilingual infants. *Nature Communications*, 4, pp. 1–6.

Gilbert, J. (2008). *Teaching Pronunciation Using the Prosody Pyramid*. Cambridge: Cambridge University Press.

Gilbert, J. (2010). Pronunciation as orphan: What can be done? *As We Speak (TESOL Newsletter)*, 7(2) pp. 1–9.

Goullier, F. (2005). *Les outils du Conseil de l'Europe en classe de langue. Cadre européen commun et Portfolios*. Paris: Didier.

Eckman, F. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning*, 27, pp. 315–330.

Hahn, L. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly*, 38/2, pp. 201–223.

Harding, L. (2012). Nateness or intelligibility: Locating the construct in pronunciation scales (roundtable presentation). In: *Proceedings of the 4th Annual Pronunciation in Second Language Learning and Teaching conference*. Vancouver.

Harding, L. (2013). Investigating the construct underlying the phonological control scale. Presented at *Language Testing in Europe: Time for a new framework?* Antwerp: University of Antwerp.

Henderson, A. (2008). Towards intelligibility: Designing short pronunciation courses for advanced field experts. *ASp*, 53/54, pp. 89–110.

Henderson, A., Frost, D., Tergujeff, E., Kautzsch, A., Murphy, D., Kirkova-Naskova, A., Waniek-Klimczak, E., Levey, D., Cunningham, U. & Curnick, L. (2012). The English Teaching in Europe Survey: Selected results. *Research in Language*, 10(1), pp. 5–28.

Herry, N., Nishinuma, Y., & Ghio A. (2003). Auto-apprentissage de la prosodie anglaise. In: Desmoulins, C., Marquet, P. & Nissen, E. (Eds.), *EIAH 2003 Environnements Informatiques pour l'Apprentissage Humain, Actes de la conférence*, pp. 113–116. Retrieved on September 15, 2015, from <http://hal.archives-ouvertes.fr/hal-00136760/en/>

Hickey, R. (2004). *A Sound Atlas of Irish English*. Berlin/New York: Mouton de Gruyter.

Hirst, D. & Ding, H. (2015). Using melody metrics to compare English speech read by native speakers and by L2 Chinese speakers from Shanghai. In: *Proceedings of Interspeech*.

Honikman, B. (1964). Articulatory settings. In: Abercrombie, D., Fry, D. B., MacCarthy, P. A. D., Scott, N. C. & Trim, J. L. M. (Eds.), *In Honour of Daniel Jones*, pp. 73–84. London: Longman.

Horner, D. (2013). Towards a new phonological grid. In: Galaczi, E. & Weir, C. (Eds.), *Exploring Language Frameworks: Proceedings of the ALTE Krakow Conference, Studies in Language Testing*, pp. 227–251. Cambridge: Cambridge University Press.

Horner, D. (2014). And what about *Testing* pronunciation? A critical look at the CEFR Pronunciation Grid and a proposal for improvement. In: van der Doel, R. & Rupp, L. (Eds.), *Pronunciation Matters*, pp. 109–124. Amsterdam: VU University Press.

Jenkins, J. (2000). *The Phonology of English as an International Language*. Oxford: Oxford University Press.

Jilka, M. (2000). *The contribution of intonation to the perception of foreign accent*. Ph.D. dissertation, Institut für Maschinelle Sprachverarbeitung, University of Stuttgart.

Jun, S.-A. & Fougeron, C. (1995). The accentual phrase and the prosodic structure of French. In: *Proceedings of the 13th International Congress of Phonetic Sciences*, vol. 2, pp. 722–725.

Jun, S.-A. & Fougeron C. (2000). A phonological model of French intonation. In: Botinis, A. (Ed.), *Intonation: Analysis, modeling and technology*, pp. 209–242. Dordrecht: Kluwer.

Kerswill, P. (2006). RP, Standard English and the standard/non-standard relationship. In: Britain, D. (Ed.), *Language in the British Isles*, pp. 34–51. Cambridge: Cambridge University Press.

Kjellin, O. (1999a). Accent addition: Prosody and perception facilitates second language learning. In: Fujimura, O., Joseph, B. D. & Palek, B. (Eds.), *Proceedings of LP'98 (Linguistics and Phonetics Conference)*, vol. 2, pp. 373–398. Prague: The Karolinum Press.

Kjellin, O. (1999b). Five cornerstones for second-language acquisition – the neurophysiological opportunist's way. Unpublished, pp. 1–9.

Krashen S. (1981). *Second Language Acquisition and Second Language Learning*. Oxford: Pergamon Press.

Lacheret-Dujour, A. & Beaugendre, F. (1999). *La prosodie du français*. Paris: CNRS.

Lado, R. (1957). *Linguistics Across Cultures*. Ann Arbor: Michigan University Press.

Luoma, S. (2004). *Assessing Speaking*. Cambridge: Cambridge University Press.

MacWhinney, B. (2008). A Unified Model. In: Robinson, P. & Ellis, N. (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*, pp. 341–371. Abingdon: Routledge.

Macré, N. (2015). Situation paradoxale ou *catch 22*: le cas d'un dispositif hybride mis en place pour des étudiants LANSAD de niveau A2/B1 en anglais en première année d'université, *Recherche et pratiques pédagogiques en langues de spécialité*, 34(1), pp. 106–126.

Major, R. (2008). Transfer in second language phonology: A review. In: Edwards, J. G. & Zampini, M. (Eds.), *Phonology and Second Language Acquisition*, pp. 63–94. Amsterdam: John Benjamins.

Maspero, M. & Quintin, J-J. (2014). L'innovation selon Innovalangues. *LEND - Lingua e Nuova Didattica* 1, pp. 6–15.

Messum, P. (2009). Grounding stress in expiratory activity. *Speak Out*, 41, pp. 12–15.

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

Munro, M. J. & Derwing, T. M. (2015). A prospectus for pronunciation research in the 21st century. A point of view. *Journal of Second Language Pronunciation*, 1, pp. 11–42.

Murphy, J. (2004). Attending to word stress while learning new vocabulary. *English for Specific Purposes*, 23, pp. 67–83.

Peperkamp, S. & Dupoux, E. (2002). A typological study of stress 'deafness'. In: Gussenhoven, C. & Warner, N. (Eds.), *Laboratory Phonology* 7, pp. 203–240. Berlin: Mouton de Gruyter.

Peppé, S. & McCann, J. (2003). Assessing intonation and prosody in children with atypical language development: The PEPS-C test and the revised version. *Clinical Linguistics and Phonetics*, 17, pp. 345–354.

Pike, K. (1945). *The Intonation of American English*. Ann Arbor: University of Michigan Press.

Piske, T. (2012). Factors affecting the perception and production of L2 prosody: Research results and their implications for the teaching of foreign languages. In: Romero-Trillo, J.(Ed.), *Pragmatics, Prosody and English Language Teaching*, pp. 41–59. New York: Springer.

Polivanov E. (1931). La perception des sons d'une langue étrangère. *Travaux du Cercle Linguistique de Prague* 4, pp. 79–96.

Rossi, M. (1979). Le français, langue sans accent? In: Fónagy, I. & Léon, P. (Eds.), *L'accent en français contemporain*, pp. 93–106. Paris: Didier.

Saito, K. (2012). Effects of instruction on L2 pronunciation development: A synthesis of 15 quasi-experimental intervention studies. *TESOL Quarterly*, 46/4, pp. 842–854.

Scarcella, R. & Oxford, R. (1994). Second language pronunciation: State of the art in instruction. *System*, 22(2), pp. 221–230.

Selinker L. (1972). Interlanguage. *IRAL*, 10, pp. 209–231.

Soulaine, S. (2014). Le défi d'une transposition didactique inédite: création d'un gestographe. *E-CRINI*, 6, pp. 1–21.

Stenton, A.J. (2011). Managing the monolingual mindset. SWANS: An authoring system for raising awareness of L2 lexical stress patterns and for inhibiting mother-tongue interference. *ReCALL*, 18, pp. 75–85.

Taillefer, G. (2002). L'anglais dans les formations spécialisées à l'Université: un cheveu sur la soupe? Peut-on rendre le plat plus appétissant? *Asp*, 37–38, pp. 155–172.

Taillefer, G. (2007). Le défi culturel de la mise en œuvre du Cadre européen commun de référence pour les langues: implications pour l'enseignement supérieur français. *Les Cahiers de l'APLIUT*, 16/2, pp. 33–49.

Trubetzkoy, N. S. (1939). Grundzüge der Phonologie. *Travaux du cercle linguistique de Prague* 7.

Weinreich U. (1953). *Languages in Contact, Findings and Problems*. New York: Linguistic Circle of New York.

Wenk, B. J. & Wioland, F. (1982). Is French really syllable-timed? *Journal of Phonetics*, 10, pp. 193–216.

Zielinski, B. (2006). The intelligibility cocktail: An interaction between speaker and listener ingredients. *Prospect: An Australian Journal of TESOL*, 21/1, pp. 22–45.

Appendix 1: Prosody descriptors (V15)

	RHYTHM & STRESS	INTONATION
Level	<p><i>Is the correct syllable stressed & marked correctly?</i> (word stress & focus, i.e. nuclear and contrastive stress)</p> <p><i>The stressed syllable should be:</i></p> <ul style="list-style-type: none"> • <i>higher (there is some pitch change)</i> • <i>louder</i> • <i>longer</i> • <i>pronounced more clearly / "correctly"</i> 	<p>Intonation</p> <ul style="list-style-type: none"> • <i>Are the intonation patterns varied and appropriate to the speaker's intentions?</i> • <i>Is the range clearly marked or is the intonation flat?</i> <p><i>NB1: Monologues may not provide opportunities for different patterns.</i> <i>NB2: Some younger native speakers use almost exclusively HRT (High rising Terminal). This is also true of learners exposed to lots of recent English, including film & TV series.</i></p>
C2	Can place and mark word stress & nuclear stress at will using all 4 cues without disrupting flow. Errors are extremely rare.	Can produce native level & natural-sounding intonation patterns appropriately, including for attitudes, emotions, humour, etc.
C1	Can almost always correctly place word stress and nuclear stress. Can clearly mark stress with all 4 cues. Errors are rare.	Can produce nearly all natural-sounding intonation patterns appropriately, even for attitudes, emotions, humour, etc.
B2	Can correctly place both word stress and nuclear stress nearly all of the time using all 4 cues to varying degrees.	Can produce a wide variety of appropriate patterns, including some attitudes, emotions, humour, etc. Can often produce a good range between high and low tones.
B1	Can correctly place word stress most of the time. Can correctly place nuclear stress most of the time, especially in shorter sentences. Can use the 4 cues but not consistently.	Can produce a variety of appropriate patterns, including the more obvious attitudes, emotions, etc.,. Can sometimes produce higher tones.
A2	Can usually place word stress correctly. Nuclear stress is placed correctly some of the time. Occasionally uses all 4 cues together.	Can produce some appropriate patterns, especially on shorter sentences. Range between low and high is minimal.
A1	Can audibly place stress on isolated words or short sentences only. Limited control of the cues which mark stress so the stressed syllable is usually difficult for the listener to identify.	Can occasionally produce appropriate intonation patterns on short learnt phrases (e.g. greetings). Range between low and high is minimal.

SOUNDS	
Level	<p>Reduced syllables <i>Reduced syllables are usually:</i></p> <ul style="list-style-type: none"> • less high • less loud • shorter • pronounced less clearly / "correctly" <p><i>E.g. schwa [ə], (doctor): /l/, short final /i/, etc. (happy): syllabic /n/ & /l/ (button, bottle). 30-35% of all vowels should be schwa!</i></p>
C2	<p>Can reduce the full range of forms, including syllabic /n/ & /l/, etc. resulting in a natural-sounding alternation of strong & weak syllables.</p>
C1	<p>Can reduce nearly the full range of forms. Strong & weak syllable alternation is evident most of the time.</p>
B2	<p>Can partially or fully reduce most possible reductions. Strong & weak alternation is often evident.</p>
B1	<p>Can partially reduce half of all reduced syllables. Some full reductions, especially schwa. Strong & weak alternation is evident on shorter sentences.</p>
A2	<p>Can produce a few partially reduced syllables, mainly schwas.</p>
A1	<p>Can rarely if ever reduce syllables; reduced syllables are usually pronounced the same as stressed syllables.</p>
	<p>Stressed & unreduced vowels <ul style="list-style-type: none"> • Are the stressed & unreduced vowels closer to English or closer to the speaker's native language? </p> <p>Can produce all / nearly all vowels in a stable and consistent accent close to e.g. US, GB.</p> <p>Can produce nearly the full range of English vowels with a stable accent often close to e.g. US, GB.</p> <p>Can produce nearly all vowels, but L1 transfer is noticeable on many vowels.</p> <p>Can produce most vowels but L1 interference is evident on most vowels, but rarely causes comprehension problems.</p> <p>Can differentiate between tense & lax vowels & diphthongs, but L1 interference very evident on all / most vowels.</p> <p>Can produce a limited number of English vowels on isolated words & expressions. L1 vowels in place of target vowels on almost all / all vowels.</p>
	<p>Connected speech (phonotactics) <ul style="list-style-type: none"> • Contractions (<i>ll, it's, gonna</i> etc.) • Linking (e.g. <i>an egg</i>) • Linking with /j/ & /w/ (e.g. <i>go-<u>ay</u>-away</i>) • Deletion of final /t/ & /d/, etc., (e.g. <i>first question</i>) • Assimilation (e.g. <i>Breaking Bad > Break<u>in</u>bad</i>) • Geminates (e.g. <i>keep <u>pl</u>aying</i>) • Etc. <p><i>NB: Higher level speakers will pause between sentences and tone units. Lower level speakers hesitate more, disrupting the flow.</i></p> <p>Can produce native speaker level connected speech phenomena making for a smooth and natural sounding flow. Can use more or fewer connected speech phenomena according to register, speed, etc.</p> <p>Can produce mainly smooth and natural-sounding connected speech. Nearly the full range of phenomena, etc., but the occasional missed opportunity, especially in longer sentences. Hesitations are rare.</p> <p>Can produce a variety of phenomena, including assimilation & deletion well over half of the time, especially in shorter sentences without many hesitations. Can produce fluent stretches of less "staccato-sounding" speech with lots of contractions.</p> <p>Can produce about half of all possible linking phenomena, including deleted /t/ and /d/, assimilations, etc., but hesitations are fairly frequent in longer sentences.</p> <p>Can occasionally produce contractions (<i>ll, d, gonna, wanna</i>, etc.) Can link between words less than half the time with word-final consonants and some /j/ & /w/ between words (e.g. <i>go out</i>). Frequent hesitations.</p> <p>Can produce basic and isolated contractions (e.g. <i>l'm, it's, gonna</i>, etc. Can produce word internal /j/ & /w/ linking (<i>going</i>). Occasional linking between words when a word-final consonant is followed by an initial vowel (e.g. <i>Are-egg</i>). Very frequent & long hesitations.</p> </p>

SOUNDS			
Level	Reduced syllables <i>Reduced syllables are usually:</i> <ul style="list-style-type: none"> • less high • less loud • shorter • pronounced less clearly / “correctly” E.g. schwa /ə/, (doctor): /ɪ, short final /ɪ/, etc. (happy): syllabic /n/ & /l/ (button, bottle), 30-35% of all vowels should be schwa!	Stressed & unreduced vowels <ul style="list-style-type: none"> • Are the stressed & unreduced vowels closer to English or closer to the speaker’s native language? 	Connected speech (phonotactics) <ul style="list-style-type: none"> • Contractions (I, it’s, gonna etc.) • Linking (e.g. an egg) • Linking with /j/ & /w/ (e.g. go-<u>u</u>l-away) • Deletion of final /t/ & /d/, etc., (e.g. fir<u>s</u>t question) • Assimilation (e.g. Breaking Bad > Break<u>i</u>mbad) • Geminates (e.g. keep <u>pl</u>aying) • Etc. NB: Higher level speakers will pause between sentences and tone units. Lower level speakers hesitate more, disrupting the flow.
C2	+	-	+
C1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
B2	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
B1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A2	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Comments:

Appendix 3: Calibration sheet (V11)

TU no	Word stress QTT			Pauses and Tone Units (TU _s) (0/1/2) - Y/N QLT			Focus: nuclear & contrastive stress (0/1/2) - Y/N QLT			Rhythm (no of beats) QTT				Reductions (schwa, /l/ & /n/) QTT			Stressed & unreduced vowels (0/1/2) QLT			Connected speech QTT			Intonation (0/1/2) QLT					
	max	notes	real	max	notes	real	max	notes	real	min	max	notes	real	max	notes	real	max	notes	real	max	notes	real	max	notes	real	max	notes	
1	1			2			2			3	4			2			2			2			2			2		
2	1			2			2			3	3			4			2			4			2			2		
3	0			2			2			1	1			0			2			0			2			2		
4	2			2			2			4	5			4			2			4			2			2		
5	1			2			2			3	4			4			2			4			2			2		
6	0			2			2			1	1			0			2			1			2			2		
7	4			2			2			4	6			4			2			4			2			2		
8	6			2			2			7	9			8			2			6			2			2		
9	6			2			2			7	9			11			2			6			2			2		
10	2			2			2			3	3			3			2			3			2			2		
11	2			2			2			5	6			5			2			4			2			2		
12	0			2			2			1	2			1			2			1			2			2		

TU no	Possible Tone Units (TUs). NB - there may be more or fewer depending on speed, stylistic factors, etc.
1	(LF)"The first question which you will ask,
2	(LR) and which I must try to answer,
3	(LF) Is this:
4	(LF) 'What is the use of climbing Mount Everest?'
5	(LR) And my answer must at once be
6	(LF) 'It is no use.
7	(LF) 'There is not the slightest prospect of any gain whatsoever.
8	(FR) Oh, we may learn a little about the behaviour of the human body at high altitudes,
9	(LR) and possibly medical men may turn our observation to some account for the purposes of aviation
10	(LF) But otherwise nothing will come of it.
11	(LR) We shall not bring back a single bit of gold or silver,
12	(LR) not a gem,
13	(LR) nor any coal or iron.
14	(LR) We shall not find a single foot of earth that can be planted with crops to raise food.
15	(RF) So it is no use
16	(LF) If you cannot understand that there is something in man which responds to the challenge of this mountain and goes out to meet it,
17	(LR) that the struggle is the struggle of life itself upward and forever upward,
18	(LF) then you won't see why we go.
19	(LF) What we get from this adventure is just sheer joy.
20	(LR) And joy,
21	(LR) after all
22	(LF) is the end of life.
23	(LR) We don't live to eat and make money.
24	(RF) We eat and make money to be able to live.
25	(RF) That is what life means and what life is for."

Student: _____ **Assessor:** _____

Institution: _____ **Class:** _____ **Date:** _____