

Article

<https://doi.org/10.52696/FXWR8735>

Reprints and permission:

The Malaysian English Language Teaching Association

Corresponding Author:

Zuraidah Mohd Don zuraidah.mohddon@utm.myORCID ID: <https://orcid.org/0000-0002-3627-9889>**Content Knowledge and the Teaching of English Stress, Rhythm and Intonation to Malaysian Learners of English**

Zuraidah Mohd Don

*Universiti Teknologi Malaysia & Universiti Malaysia Kelantan***ABSTRACT**

This paper makes a positive response to the CEFR Companion, which requires learners to speak the target language with appropriate prosodic features defined as stress, rhythm and intonation. When, as in Malaysia, the teaching of English is aligned with the CEFR, teachers are expected to teach prosody to enable their students to progress, despite the serious practical problem that they may not have the necessary content knowledge. The paper aims to provide teachers with the content knowledge, while leaving to teachers the prerogative to decide what and how to teach in their own classrooms. It outlines some basic prosodic concepts and covers what learners have to do when speaking English other than producing acceptable speech sounds. Although teachers who read up on English prosody may develop greater knowledge and increased awareness of prosodic features, just knowing about these things is not enough, and it is also essential to know how to recognize and produce them. In order to enable learners to acquire appropriate prosodic features, teachers require support at the point of need in the classroom, and the discussion includes suggestions how this can be done with modern technology.

KEYWORDS: word stress, speech rhythm, intonation, prosody, prosodic features**Introduction**

This paper addresses the problems of content knowledge which makes it difficult for teachers to meet the requirements of teaching prosody, usually taken to include stress, rhythm and intonation. The Companion volume to the CEFR (Council of Europe, 2020, pp. 133–135) supersedes the phonological scale included in the original version of the CEFR (Council of Europe, 2001), which has received considerable criticism (Kremmel et al, 2023, pp. 69-71). The Companion volume to the CEFR (Council of Europe, 2020, pp. 133–135) presents a new

scale relevant to the teaching and learning of spoken English, and this new scale includes a column headed “Prosodic features”, which specifies what learners can do from A1 to C2. Inspection of the “can do” statements indicates that the features of which learners are expected to develop increasing command include stress, rhythm and intonation. For example, for L2/FL students at school or at university, pronunciation “is generally intelligible; intonation and stress at both utterance and word levels do not prevent understanding of the message” at B1, and at B2 they “Can generally use appropriate intonation; place stress correctly and articulate individual sounds clearly [...]” (Council of Europe, 2020, p. 134). The rather undeveloped nature of the construct of prosodic features and the lack of definition of prosodic terms will pose a challenge for the teaching of spoken English to ESL/EFL learners.

Most teachers would surely agree that one of the important goals of ELT is to enable their students to take part successfully in activities that require them to speak in English and to understand what other people say in English. This is of course the idea behind the new CEFR scale (Piccardo, 2016). The problem is that by no means all teachers will be familiar with the terms *prosodic features* and the related *prosody*, and with *stress*, *rhythm* and *intonation* (see e.g. Alameen, 2012), and this reflects the lack of training in pronunciation expertise (Galaczi, Post, Li, Barker & Schmidt, 2017). Someone unfamiliar with these terms will find it difficult to judge whether intonation is appropriate or whether stress is placed correctly. The first objective of this paper is to explain the meaning and use of these terms in the context of ESL/EFL teaching.

The situation in Malaysia is presumably much like that in other countries, in that although many students, particularly from less privileged areas, leave education with poor English, there are also many who leave education with excellent English. The factors that distinguish poor and excellent spoken English are difficult to identify, and are sometimes assigned vaguely to “intonation”. The practical problem to be solved, and the second objective of this paper, is therefore to identify these factors with a view to improving the existing teaching and learning of spoken English (Gregersen, 2019).

While understanding terms such as *stress*, *rhythm* and *intonation* may help to raise awareness of English prosody, it is by no means sufficient. If students are required to improve their stress, rhythm and intonation in order to progress up the CEFR scale, then there must be clearly defined targets that teachers and their students can realistically work towards. In addition, if teachers are expected to teach so that their students can do the things required, then teachers themselves have not only to understand but also to be able to produce appropriate stress, rhythm and intonation when speaking English (Gilakjani & Sabouri, 2016). Since teachers cannot be expected to improve their English prosody spontaneously and on their own, the third objective is to outline how this can be achieved in the real world.

A major problem for teachers and students is how to find out about stress, rhythm and intonation. In the middle decades of the last century, there were many books for learners, notably including O’Connor and Arnold’s *Intonation of Colloquial English* (1973). More recently the interest has been in theoretical work, which has advanced the theory but which is not of much practical value in the classroom, notably including the work of Couper-Kuhlen and Selting (2018), which deals with intonation in the context of social interaction. The practical study of intonation has been greatly enhanced by the development of speech waveform editors, including PRAAT (<https://www.fon.hum.uva.nl/praat/>), which make it

possible to see and measure prosodic features. Much of the information in this paper is not taken from conventional books and articles but based on the study of speech waveforms, and the extensive close study of intonation in naturally produced speech.

The next section outlines some prosodic features which it is hoped will be of use in the classroom. This is followed by the discussion and a separate conclusion.

Prosody

The most important prosodic feature is the pitch of the voice, which goes continuously up and down as we speak. Although we may think of words as units of speech made up of speech sounds, we also have to consider the pitch, and this is true whatever language we are speaking. Perhaps the clearest illustration of the role of pitch is found not in English but in Chinese. Mandarin Chinese is like many languages around the world in that different words can be distinguished by their pitch patterns, known as *lexical tones*, or just *tones*. The best known and most cited example involves the syllable /ma/, which can have four tones numbered 1 to 4. With tone 1 (high level pitch) it means ‘mother’, with tone 2 (pitch rising to high) ‘hemp’, with tone 3 (pitch falling to low and rising again) ‘horse’, and with tone 4 (pitch falling to low) ‘scold’. These are very simple patterns, and similar patterns are found in many other languages, although their role in speech is likely to be different. English also sometimes uses pitch patterns to distinguish words, but as part of the system of word stress.

Word Stress

A characteristic of English words is that they can have several vowel sounds, and the speaker puts different degrees of emphasis on these vowel sounds. Every English word has a property known as *word stress* or just *stress*, which reflects particular emphasis (Sadat-Tehrani, 2017). We can mark the stressed vowel informally by putting the corresponding letter or letters in capitals. For example, *lOrry* is stressed on the first vowel and *beLOng* on the second. What we actually do to stress a vowel is to give it a fall in pitch. In this way, English stress is related to Mandarin tone 4, although it is not exactly the same. In order to fall, the pitch first has to rise to a higher level, and there is indeed a preliminary rise in most cases in both languages; but people generally report hearing just the fall. Unlike Mandarin, in which the fall contrasts with other tones, every English word taken out of the dictionary or read aloud on its own has this fall on the stressed syllable.

If you use a dictionary which represents the pronunciation of words, you are likely to see the stress marked by a raised vertical dash ('). This symbol can be used with either ordinary spelling or phonetic transcription, and it is by convention placed at the beginning of the stressed syllable, thus 'lorry, be'long or /'lɒrɪ, bɪ'lɒŋ/. Since the number of syllables in a word is the same as the number of vowel sounds, this convention indicates which vowel is stressed, and in view of its widespread use, this convention is followed here. However, if learners were trained to complete the fall in pitch on the stressed vowel itself, their English would sound rather strange. More precisely, the fall begins on the stressed vowel and is completed on the rest of the word. For example, in *lorry* the fall begins on the first vowel and is typically completed on the second vowel.

The effect of stress is to make stressed syllables stand out more than unstressed syllables, a property known as prominence. Syllables can accordingly be divided into strong (stressed) and weak (unstressed): *'happy* is strong + weak, *per 'haps* weak + strong, and *to 'morrow* weak + strong + weak. The position of the stress can distinguish different words in English; for example, *'billow* and *be 'low* have the same speech sounds but different stress and prominence patterns. There are also several paired nouns and verbs with different stressed syllables, including *con 'vict* (verb) and *'convict* (noun). A consequence of stress in English and some other languages is that the prominence of unstressed vowels is in some cases reduced by replacing it with one of the so-called reduced vowels, [ɪ] as in *belong*, or the first vowel of *today*, which is known as schwa and for which the phonetic symbol is [ə]. In this way the first vowel sound of *con 'vict* /kən'vɪkt/ is reduced to [ə]. The [ɪ] vowel occurs in both stressed and unstressed syllables, and so in *'convict* /'kɒnvɪkt/ it is the same whether it is stressed or not.

Prominence is sometimes confused with stress, and the claim is made that there has to be a contrasting weak syllable for a syllable to be classed as stressed. Following this view, monosyllables (i.e. words with only one syllable) have no weak syllable and so cannot be stressed. More realistically, monosyllables such as *good* and *bad* are spoken with a fall in pitch and so have stress just like any other words: they happen not to have a weak syllable to contrast in prominence with that stressed syllable. The stress mark is not used for monosyllables in dictionaries because it would be redundant.

Rhythm

Rhythm is essentially concerned with the organisation of events in time. It is closely connected with music, and indeed the first attempt to describe English rhythm was made by Joshua Steele (1775), who used musical notation to represent the performance of an actor reciting Shakespeare. Events may occur at more or less equal intervals, and although the events themselves can be identical, as in the case of the ticking of a clock, a rhythm is more typically made up of different events, such as long and short events, or strong and weak events. As indicated above, stress produces patterns of strong and weak syllables, and these patterns form the basis of the rhythm of spoken English. We now have to consider how strong and weak syllables are organised in time. To do this, we need to go beyond stress in short words, and consider longer words.

Let us start with the word *uniform*, which is stressed on the first syllable, thus *'uniform*. The first syllable is strong and the second weak; but the third syllable seems to be intermediate in prominence. To explain this, we need to consider another property of rhythm. When listening to music, you might find yourself tapping on a surface in time to the beat of the music. Similarly, you can tap in time with the stressed syllables in spoken English, including the first syllable of *uniform*. If you say the word very slowly, you can tap to the first syllable, and then again to the third syllable. However, if you put the word into a phrase such as *a uniform pattern*, it would be very odd to tap to that third syllable. When you tap in this way, you are responding to perceived beats in the spoken form. The beat on the third syllable is optional, and it is clearly a rhythmical phenomenon, because it depends on the speech tempo. Reduced vowels are rarely accompanied by a beat, but the unreduced second vowel of *uniform* can take a beat at slow tempo.

This is important to understand the pronunciation of long English words. For example, *'photograph* is stressed on the first syllable, and the third syllable can take a beat if the word is spoken slowly. *Pho'tographer* /fə'tɒgrəfə/ is stressed on the second syllable, and the other syllables have reduced vowels. *Photo'graphic* /fəʊtə'græfɪk/ is stressed on the third syllable, and has another rhythmical property: the vowel of the first syllable is not reduced, and can take a beat at slow tempo. Like many languages, English has a preference for alternating strong and weak syllables, which facilitates the emergence of a beat two syllables away from the stress, but not immediately before a stress. These syllables with potential or actual rhythmical beats are commonly said to have “secondary stress”, for which the phonetic symbol is a lowered vertical dash [ˌ], thus *'uni form*, *'photo graph*, *ˌphoto graphic*. The notion of secondary stress is not used here, because it leads to the confusion of stress, prominence and rhythm.

Intonation

The term *intonation* is used in different senses, but the general idea is that it concerns pitch patterns, stress and rhythm above the level of the word including conversation (Reed, 2012). It also includes two different language systems, of which the first is known vaguely as “normal intonation”, and the second is concerned with strategies to convey the speaker’s meaning to the addressee.

Normal Intonation

In everyday speech, words that belong together are produced as a group, often consisting of two stressed syllables. There are many examples of such groups in the first paragraph of section 2.2, and they include *es'sentially con'cerned*, *e'vents in 'time*, *'Joshua 'Steele*, *'spoken 'English*, and *'organised in 'time*. When these are taken out of the text and spoken in isolation, the grouping is indicated by the pitch pattern associated with the first stress. When words are spoken in isolation, the stress is accompanied by a fall in pitch; but when a word occurs at the beginning of a group, its stressed syllable remains quite high in pitch. This makes it rather like Mandarin tone 1. Since we are now dealing with different pitch movements, we need to replace the simple stress marks with symbols that indicate the pitch movement, thus *es[↗]entially con[↘]cerned*, *e[↗]vents in [↘]time*, etc. When English words are spoken in their dictionary form, they always take a fall on the stressed syllable; when the same words are used in the context of several words, the stressed syllable can take one of a number of different pitch patterns, also known as tones.

The rhythm of these phrases can be informally described as (de-)DUM-de-DUM, and as for longer words, you can beat time to the stresses by tapping on a surface with a finger or pencil. *The[↗] ticking of a [↘]clock* is a more complex example of the same pattern. In this case, it is necessary to make a distinction between content words (nouns, verbs and adjectives, etc) and function words (articles, prepositions, some pronouns, etc), because the stresses are generally associated with content words. In their dictionary form, the function words *the*, *of* and *a* are stressed as usual; but when they are included in a phrase, they typically lose their stress, and have to fit in between the content words to keep time. Paired stresses of this kind are an important characteristic of spoken English, and they often occur at the end of a spoken sentence.

Paired stresses can be embellished, for example by alliteration, in which the stressed syllables begin with the same sound: →*safe and* ↘*sound*; →*house and* ↘*home*; →*fast and* ↘*furious*. Some English proverbs are made up of two sets of paired stresses: →*out of* ↘*sight*, →*out of* ↘*mind*; a →*rolling* ↘*stone* →*gathers no* ↘*moss*. Nursery rhymes often follow a similar pattern: →*baa baa* ↘*black sheep*, →*have you any* ↘*wool?* etc. In this last case, you might notice that we normally say →*black* ↘*sheep*, but here the rhythm of the nursery rhyme takes over, and *sheep* loses its stress.

Intonation Strategies

At B2 and above, learners are expected to use stress, rhythm and intonation “to support the message they mean to convey” (Council of Europe, 2020, p. 134, B2). Stress and rhythm are determined by the nature of the language, and leave little or no room for imaginative performance. So-called normal intonation varies according to the speed of speaking, and so gives some limited scope for more fluent learners to show what they can do. A characteristic of intonation in ordinary conversation is that it is so varied that it might seem to be constrained by no rules at all. However, there are strategies that skilled speakers use to convey their meaning to the addressee.

Although it is possible to put the stress on the wrong syllable of a word, and to group words into phrases in inappropriate ways, strategies are more effective or less effective, and cannot really be said to be right or wrong. In using intonation, the speaker has to take the needs of the listener into account. The more speakers do to make their meaning clear, the less listeners have to do to understand what the speaker intends. In the case of learners, the more they do as speakers, the more highly what they say is valued. This has important consequences for assessing intonation: anything beyond separate words cannot really be said to be wrong, but may need to be improved by the use of a more effective strategy. We consider two strategies here: intonation in lists and intonation in speech acts.

Lists

We can usually tell when someone is making a list, and in this respect, English is like many other languages. Non-final items are accompanied by a rise in pitch (↗), and the final item is given the usual fall, e.g. ↗*Monday*, ↗*Wednesday and* ↘*Friday*. This rise roughly resembles Mandarin tone 3. Although many lists are obvious lists, such lists of days or months, or items in a shopping list, the listing strategy is very flexible, and many things can be treated as lists.

Listing is a strategy, which means that its use is optional. Imagine the case of a learner we can call Soo, who is asked to say something about herself. When she is a beginner, she is given the frame “My name is <name>. I live in <place>.”, and she replies “My name is Soo I live in Ipoh.” a word at a time with pauses between the words, and without any discernible structure. When she is more advanced, she might use paired stresses: *My* →*name is* ↘*Soo*. *I* →*live in* ↘*Ipoh*. A little later she may think of bringing the two sentences together in the form of a list: *My* →*name is* ↗*Soo*, and *I* →*live in* ↘*Ipoh*. Later on, now speaking English fluently, she can go beyond listing. Imagine that she takes part in a competition in which all the contestants give their names and say where they live, and she says *My name is* →*Soo and I live in* ↘*Ipoh*. She simplifies the intonation pattern by bringing the two items together as paired stresses. She now

has so much material before the first stress and between the stresses that the rhythmical beats of slow speech have been lost. English is sometimes said to have a regular beat or “isochronous stress”, but this really applies to slow speech, and any regular beat is lost when speakers speed up. Soo’s teachers will recognise intuitively that she is making progress and becoming a confident and fluent speaker of English, even if they cannot say explicitly what new things she is learning. Of course, the assessment of intonation ideally needs to go beyond intuition, and to be based on observable characteristics of the learner’s performance.

Soo’s progress illustrates the important point that as they advance, learners are unlikely to replace incorrect intonation with correct intonation, and will more probably acquire new and more advanced strategies which are more appropriate in the context. The best known example of replacing a list with paired stresses is *fish and chips*. It would actually be very odd for someone to make a list, and say they were eating *fish, and chips*, because the dish is always known as *fish and chips*. The word *and* here is a function word, and is pronounced /ən/, which explains the spelling “fish ’n’ chips” sometimes seen in restaurant menus.

Speech Acts

Speech acts (Austin, 1962; Searle, 1969) are things that people do with words, such as making a statement, asking a question, or giving an instruction. The last stressed syllable in an utterance takes a fall by default, but other tones are also used for different kinds of speech act. Since the variety of tones and their connection with speaker meaning is complex and beyond the scope of this paper, a single illustration will perhaps suffice. When Soo says *I live in Ipoh*, she is making a statement, and the fall is appropriate in this case. However, if someone were to ask her *Do you live in Ipoh?*, their pitch would probably go up at the end to form the high rise, thus *Do you live in Ipoh?* This high rise resembles Mandarin tone 2. The high rise is in fact not obligatory, because the interrogative grammatical structure is already associated with questions. The high rise makes it clearer that the speaker intends the speech act to be understood as a question.

Reading Aloud

Reading aloud is a special skill developed by teachers and newsreaders, and teachers get beginning readers and language learners to read aloud in order to assess their prosodic competence. The learner who can only read one word at a time, and pronounces each word with a falling pitch and leaves a pause before the next word is either a complete beginner or having difficulty. Those who can read word groups are clearly more advanced, and those who can read “with expression” are becoming successful readers. Although reading aloud is not used as a criterion to determine CEFR levels, it can be used in the classroom to develop the speaking skills required for progress up the scale (Adrian, 2014). However, reading aloud has become something of a Cinderella in recent decades, because it has been regarded as not authentic in comparison with naturally produced speech. Abercrombie (1963) described reading aloud text as “spoken prose”, and this view has been echoed ever since. Nevertheless, reading aloud has an important place in the development of spoken language skills. When we speak normally, we have to carry out a sequence of complex cognitive tasks: we think of something to say; we choose words to express our meaning; we draw on our knowledge of grammar to put the words together in a meaningful way; we recall the pronunciation of the

words; we draw on our knowledge of intonation to prepare for speaking aloud; and then we say our words aloud (Levelt, 1989). When we read aloud, most of this work has already been done. We have to identify the words and which words belong together to form groups, and we have to choose appropriate intonation patterns to convey the meaning of the text to the listener. In this way, reading aloud gives a good indication of the reader's understanding of the text.

Discussion and Implications

The previous section outlines the nature of the English prosodic system, and raises a number of issues and their implications. Just knowing about prosody is insufficient in itself, and teachers need to use prosodic information to enable learners to improve their proficiency in speaking English, and possibly to improve their own performance. Prosody cannot be taught by passing on facts, and it is important to understand the role of procedural knowledge in developing spoken proficiency. Secondly, teachers need to know how to relate general ideas about prosody to the specific circumstances of the materials they are using for teaching. In this case, they need support at the point of need in the classroom, and this can now be provided in principle using modern technology. Thirdly, it is important for the time and effort needed to teach prosody to lead to measurable improvement in the performance of the learners. These issues are explored in the next three subsections.

The Development of Procedural Knowledge of Prosody

An important distinction which is relevant here is made between declarative knowledge and procedural knowledge. Declarative knowledge consists of facts of which we are consciously aware; for example, that Malaysia is located between Thailand and Singapore. Procedural knowledge involves knowing how to do things, such as swimming or riding a bicycle. Children learn to do many things without having any idea how they learn, including acquiring their first language; and in this case the knowledge remains unconscious (see Dehaene, 2014).

A traditional academic approach to teaching and learning involves presenting learners with logically ordered declarative information which they then apply to data. For example, undergraduates may be told about phonetic transcription and then required to transcribe a text. An equivalent in teaching children to ride a bicycle might be to start with words such as *pedal*, *brake* and *mudguard*. This is unlikely to be successful because most children will not see knowing these words as a problem they want to solve: what they want to do is to get on a bicycle and ride it. It is when they can ride that they will have a motivation to learn words about bicycles. Much the same applies to prosody. Teachers reading section 2 of this paper may become more aware of what is involved in teaching spoken English, but they are unlikely to be able to go into their classrooms and immediately use what they are now aware of in their teaching. The declarative information required to understand prosody is auditory in nature, and the first step towards learning is extensive listening. If students are expected to satisfy the spoken English requirements for the higher CEFR levels, they have to start by listening extensively to authentic spoken English in the classroom, and preferably more generally in daily life. For those learning British English, this means listening to varieties of English reasonably close to British Received Pronunciation (RP). This is how students sent abroad to study in an English-speaking country tend to come back able to speak much better English, but quite unable to explain what they have learned or how they learned it.

It is possible to learn the prosodic patterns of a new language, because we have brains designed to learn spoken language, and one of our abilities is to copy and repeat what someone says. This ability can be lost, perhaps after a stroke, and its loss is known as conduction aphasia (Ingram, 2007, p. 52). There is a difference between imitating what we hear and reproducing it for ourselves. An English speaker who knows no Chinese can imitate *ni hao* (both words spoken in isolation with tone 3), but reproduce it with English prosody as the nonsense phrase *~knee show*. Malay has no stress system, and although Malays can imitate English words with stress patterns, when they produce the same words for themselves, there may be no discernible stress pattern. This is a major problem for Malay learners of English. Malays probably start off using Malay prosody when speaking English, which is described as a Malay accent, and successful learners are able to modify their Malay prosodic system to bring it closer to English. The CEFR Companion (Council of Europe, 2020, pp. 134–135) regards the influence of a foreign language as something to be avoided, but it is probably the necessary initial stage in the learning process. An important finding about how we learn (Dehaene, 2020) is that the unconscious brain sets up informal hypotheses, tests them and replaces them as necessary by better hypotheses. This is how we learn things without being aware of what we are learning; it is also a property of the human brain that has long since become a fundamental principle of scientific method.

To learn English stress effectively, learners first have to listen to many, many authentic examples of appropriately stressed English words, and keep practising until they can stress words themselves. When we have already learned something, it is appropriate to ask what we have learned, and in this case think about the stress position, the fall in pitch, and perhaps the process of vowel reduction that weakens unstressed syllables. The same applies to prosodic phenomena in general. Although the usual order is to provide declarative knowledge which learners are then expected to use to develop procedural knowledge, it is probably more effective for learners to develop procedural knowledge and then to reflect on what it is that they have learned. A word of caution is required, however, because speaking is normally an unconscious activity, and people differ greatly in their ability to bring normally unconscious information into conscious awareness. What one person finds perfectly obvious may be totally incomprehensible to someone else.

In practice, there is likely to be some interaction in the development of declarative and procedural knowledge. Very young children given sufficient exposure to a new language can learn it much as they learn their first language. For older learners, including those starting English at school, appropriate declarative information may speed up the development of the procedural knowledge that lies behind proficiency. They are likely to learn new words more effectively by being given semantic information that relates words and groups of words to each other, e.g. *girl* in relation to *boy* and to *woman*. At some stage, learners can become autonomous learners, and learn new things from their exposure to the language without being explicitly taught. What is subsequently learned depends on the input to which the learners are exposed, and this is why if learners are to produce and understand international English, they have to be exposed in large amounts to authentic examples of international spoken English.

Transmission to the Classroom

Relevant information needs to be made available at the point of need. This is taken for granted in the case of word meaning, and someone who wants to know the meaning of a word looks it up in a dictionary, or now more probably using an online search engine. It would be quite extraordinary to organise an academic course on word meaning in English just in case it met someone's needs at some time. The same applies to prosody. It would be unrealistic to expect teachers to attend courses on English prosody even if they were available, and equally unrealistic to expect them to listen to large amounts of authentic material in their own time, especially if they do not know how they are going to benefit from the expenditure of time and effort. It is therefore necessary to find a completely different approach. Teachers need information in a form that they can use in teaching, and learners need it in a form that enables them to learn. Using modern technology, the needed information can be delivered in the classroom to the teacher's or a student's telephone or related device.

Teachers preparing a text for a class at a particular CEFR level need to ensure that pronunciation is included when familiarising the class with the word list for that level. In some cases, such as science texts, words recommended for a higher level may be included, e.g. B2 words in an A1 text, which raises problems with difficult words. The pronunciation itself can often be obtained from an online pronouncing dictionary, but it may need further explanation. Giving this explanation can lead to a cascade of problems, one leading on to the next, and at some point, possibly going beyond the teacher's personal knowledge. For example, a teacher might need to check the pronunciation of *tortoise* to prepare for a lesson. If the teacher expects the pronunciation to end like *voice*, the nature of the vowel spelt "oi" needs some explanation, which involves explicitly identifying the phonemes that make up the word, a process known in phonics as sounding out. This can confuse someone who is unaware of the phonemes of English, and who needs to find out more about phonemes. Someone unaware of English stress may have no idea of the stress on the first syllable of *tortoise*, and need further information on stress. When the words of the text have been covered, there will be another cascade of questions on the intonation, which will be particularly important if some of the students are expected to read parts of the text aloud.

Providing this information is of course impossible using traditional technology, even with an online dictionary. It can be done by means of a website that teachers and learners can explore by following the links from one page to another. (This is not in fact a mere imaginary possibility, because the linguistic part of the website described here already exists.) A user who looks up *tortoise* can click not only on a link to the pronunciation but also on a link to a page that divides words into their component phonemes, and so can play just the vowel sound corresponding to the spelling "oi". Repeated listening to many words and their phonemes is a good and painless way of understanding the phoneme system of the target language, and much more effective than reading theoretical explanations. Phonemes are represented by special symbols, which soon become familiar when associated with the corresponding sounds. It may (and probably will) require several attempts to grasp the information made available in this way, because one cannot expect to listen to an audio recording (of the word *tortoise* for example) just once and understand everything in the accompanying text. Understanding stress may require several attempts, listening for and noticing something different each time: noticing that part of the word is made prominent; noticing the fall in pitch starting on a particular vowel; noticing the stress mark at the beginning of the stressed syllable; noticing the symbols representing the vowels in each syllable; noticing the rhythm created by the stressed and unstressed syllables. As the user is able to hear more detail in the spoken form, written

explanations will begin to make more and more sense. Understanding the details of a spoken text increases theoretical expertise, and hearing the detail is the first step towards being able to produce it and so improve proficiency. Learning one thing in this way can lead on to another. Learning about the phoneme system is an essential part of pronouncing the words of the target language, even if learners do not recognise it as a relevant problem. However, learning about and making lexical stress makes the phoneme system something relevant to find out about, and this could include learning the symbols that conventionally represent them. Learning about the phonemes themselves before the symbols that represent them makes a lot more sense than trying to understand a table of phonetic symbols and then finding out about the entities that the symbols represent.

Measuring Progress in English prosody

The measurement of the progress made by language learners is closely associated with the CEFR and more specifically with the “can do” statements which specify what is required of learners at each CEFR level. Although what learners can do is doubtless taken in practice to refer to the written language, proficiency in the spoken language necessarily involves the oral skills specified in the CEFR. Here we encounter a problem, because it is difficult to relate the general skills referred to in the “can do” statements with prosodic patterns which the learner can be expected and even required to use. It follows that it is likewise difficult to use prosodic patterns as objective evidence for CEFR levels from A1 to C2.

It would be convenient if learners could be tested on the pronunciation of words, while intermediate learners were expected to produce well-formed phrases and advanced learners to tackle the intonation of longer expressions, and produce speech acts appropriately. It is however not so simple. Malaysians whose English is generally very good can still have problems with words like *photographer*. Medical students and others have to cope with difficult words of Greek or Latin origin for which learning the pronunciation requires some understanding of how the words are formed. Competent speakers use connected speech rules when uttering phrases, but since these rules are optional, it would be inappropriate to penalise a learner for not using them. Producing speech acts is complex and native speakers not infrequently make performance errors or say things in an inappropriate way. Learners cannot fairly be penalised for failures that can baffle native speakers. In short, the objective measurement of the prosody used by learners is at the very edge of what is currently known about prosody, and requires further research which is alas beyond the scope of this paper.

Conclusion

The CEFR Companion does not have much to say about prosodic features beyond referring to stress, rhythm and intonation. Teachers who do not know what stress, rhythm and intonation are, are unlikely to be able to do much to help their students satisfy the spoken English requirements to progress from one CEFR level to the next. This paper has sought to identify the simplest components of English prosody, starting with stress in English words and the rhythms created by stress patterns, and briefly considering intonation beyond the word (Mennen & Leeuw, 2014). Stress and rhythm are auditory experiences, and understanding them has to start with listening extensively to authentic examples. Academic text can raise awareness of what teachers and learners need to know about in spoken English, but it cannot substitute for the listening experience, and in any case the time to read and understand the text is after

listening and not before. Students who rely on their classroom experience to learn spoken English need spoken English support at the point of need, and while support of this kind was until recently quite impossible, it can now be made available thanks to technological developments such as the internet and the mobile telephone. Teachers can indeed help their students speak English with English prosody, but they do not need to rely on their own intuitive knowledge of spoken English, and in addition to textbooks and other materials they need support for prosody delivered to the classroom.

Funding statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of interest statement

The author declares that there is no conflict of interest.

References

- Abercrombie, D. (1963). Conversation and Spoken Prose. *ELT Journal*, XVIII(1), 10–16. <https://doi.org/10.1093/elt/XVIII.1.10>
- Adrián, M. M. (2014). The efficacy of a reading aloud task in the teaching of pronunciation. *Journal of English Studies*, 12, 95-112. <https://doi.org/10.18172/jes.2825>
- Alameen, G. (2012). Learner digital stories in a Web 2.0 age. *TESOL Journal*, 2(3), 355-369. <https://doi.org/10.5054/tj.2011.259954>
- Austin, J. L. (1962). *How to do things with words*. Cambridge: Harvard University Press.
- Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*. Strasbourg: Council of Europe Publishing.
- Couper-Kuhlen, E., & Selting, M. (2018). *Interactional Linguistics: Studying language in social interaction*. Cambridge: Cambridge University Press.
- Dehaene, S. (2014). *Consciousness and the Brain*. New York: Viking.
- Dehaene, S. (2020). *How we learn*. New York: Penguin.
- Galaczi, E, Post, B, Li, A., Barker, F. & Schmidt, E. (2017). Assessing second language pronunciation: distinguishing features of rhythm in learner speech at different proficiency levels. In Issacs ,T. & Trofimovich, P. (eds.) (2017) *Second Language Pronunciation Assessment*. Bristol: Multilingual Matters/Channel View Publications Ltd
- Gilakjani, A. P., & Sabouri, N. B. (2016). Why is English pronunciation ignored by EFL teachers in their classes?. *International Journal of English Linguistics*, 6(6), 195-208. <https://doi.org/10.5539/ijel.v6n6p195>
- Gregersen, T. (2018). Teaching English prosodic features. *The TESOL Encyclopedia of English Language Teaching*. Hoboken: Wiley Blackwell.
- Ingram, J. C. L. (2007). *Neurolinguistics: An introduction to spoken language processing and its disorders*. Cambridge: Cambridge University Press.
- Kremmel, B., Eberharter, K., Konrad, E., Guggenbichler, E., Moser-Frötscher, D., Ebner, V., & Spöttl, C. (2023). The CEFR Companion Volume: opportunities and challenges for Language Assessment. *DiSlaw* 3, 65-83
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge: MIT Press.

Content Knowledge and the Teaching of English Stress, Rhythm and Intonation to Malaysian Learners of English

- Mennen, I., & de Leeuw, E. (2014). Beyond segments: prosody in SLA. *Studies in Second Language Acquisition*, 36, 183-194. <https://doi.org/10.1017/S0272263114000138>
- O'Connor, J. D., & Arnold, G. F. (1973). *Intonation of Colloquial English*. London: Longman.
- Piccardo, E. (2016). *Common European Framework of reference for Languages: Learning, Teaching and Assessment. Phonological Scale Revision Process Report*. Strasbourg: Council of Europe
- Reed, B. S. (2012). Prosody in Conversation: Implications for Teaching English Pronunciation. In J. Romero-Trillo (Ed.), *Pragmatics and Prosody in English Language Teaching* (pp. 147-168). Dordrecht: Springer. <https://doi.org/10.1007/978-94-007-3883-6>
- Sadat-Tehrani, N. (2017) Teaching English Stress: A Case Study. *TESOL Quarterly*, 8(4): 943-968. <https://doi.org/10.1002/tesj.332>
- Searle, J. R. (1969). *Speech Acts: An essay in the philosophy of language*. Cambridge: Cambridge University Press.
- Steele, J. (1775). *The Melody and Measure of Speech*. Menston: The Scholar Press.