WHAT KIND OF ENGLISH PRONUNCIATION SHOULD WE TEACH?

Laurie Bauer

Victoria University of Wellington

Introduction

When I was first involved in the teaching of English as a foreign language, the teaching of pronunciation was considered a vitally important part of what the teacher had to do. It was seen as an area in which a relatively small amount of intervention could have a huge effect on comprehensibility and, what is more, that gain could be continued throughout the student's career as a learner once good habits were instituted. This prioritisation of pronunciation resonated with me particularly because of my experience in learning Italian. I have never learnt much Italian, but I did go to an introductory course in Italian, and then visited Italy. Hitch-hiking through Italy (in the days before hitch-hiking was viewed as quasi-suicidal) and talking to a large number of Italians, I was regularly complemented on the quality of my Italian. I knew this was a joke (even if not intended as one). I knew a few fixed phrases and could guess at much of what was being said to me but had little ability to formulate sentences. But the warm welcome I was given was because my pronunciation was good enough not to sound like the broken Italian produced by most foreigners. A good pronunciation brings good will in its wake. Conversely, students whose English pronunciation is not good enough for them to be understood, feel that they lose respect from their peers and teachers, become frustrated and demotivated (Tananuraksakul, 2012).

In the years since then, pronunciation has often been downplayed. It is viewed as being too technical for beginning students, and other things, particularly vocabulary, have been given priority. There has, unfortunately, been little communication between the phoneticians (like me) and the language teachers (which I ceased to be when I moved to New Zealand) to translate the jargon of phonetics into usable pedagogical practice. It may not be easy to tell an Anglophone learner of French to make dental plosives or to use syllable-timed rhythm, but it should be relatively straightforward to tell them to push the tongue forward in the mouth or to try to sound nastier. I see both sides as having a part in this lack of intercommunication, and I still think that it is worth working at a good pronunciation. But things have got rather more complicated now.

At the time, in the early 1970s, it was assumed that the target forms of English were either British (what we then called R[eceived] P[ronunciation], though the

label has become less popular) or American (a broadcast standard General American). Each teacher modelled, as best they possibly could, one of these varieties, usually the one closest to their own native variety if they were Anglophones, even though few people actually spoke natively either of the varieties that were being targeted. Sometimes things were even worse than that: stories still abound of teachers from New Zealand being forced to teach 'American English' in classes in East Asia, although the model they were presenting to their students must be stereotypical, inconsistent and not fully understood.

In what can now perhaps be seen as an early objection to colonialism, the dominance of native models was challenged (e.g., by Jenkins, 2000) and a more neutral model was suggested. But any normative model has its problems, and later approaches have looked for more learner-oriented models, stressing the need for a variety of different input styles and an acknowledgement of the phonological constraints of first languages on learner varieties. This leaves the whole question still open to discussion and can be seen as a recipe for chaos.

Why we need models and why we can't use them

An argument against the dual-model image of learner English is that not only is it an unrealistic picture of the reality of the English language, but also that the two models (British and American) are phonetically and phonologically quite difficult. On many analyses, RP has 21 or more vowels and 24 consonants. The consonants can be difficult enough for speakers of languages which have just 10 or fewer, but 21 vowels is overwhelming for speakers of languages which have only 3 (e.g., Classical Arabic) or 5 (e.g., Russian, Spanish, Swahili). The models are also overly prescriptive in the sense that speakers can be well understood even if they make fewer distinctions.

Hans Hartvigsson, my professor when I worked in Denmark, used to talk of the 'spy school' of language learning, where the aim was to speak the foreign language so well that you could pass as a native. While we probably all know a few people who achieve this, it is clearly an elitist aim, and one which is impractical for most learners. Not only are most of us too entwined in the sound systems of our first languages to be able to produce sounds from other languages accurately, or even to hear them properly, many do not want to pass as a native – for various reasons. French learners of English frequently find that having a 'sexy' French accent is a social benefit when speaking English, for instance. Many speakers find that being audibly foreign can be valuable in that listeners are more likely to make allowances for linguistic errors in foreigners. A Spanish friend of mine, whose command of English is in most ways admirable, has difficulty with words ending in N. At first sight, this is odd, since standard Spanish has many

words which end in N (e.g., *comen* 'they eat'). But my friend comes from the south of Spain where final N is replaced with a velar nasal. My friend uses this in English and says *theng* for *then* and so on, quite consistently. His linguistic partisanship (and linguistic identity) in his first language affects the way in which he speaks English, even though that identity is not communicated to English-speaking listeners. More generally, many people feel that they need to mark linguistically their ingrained identity (national, social) as part of a different linguistic community, and do not want to achieve inconspicuousness in the new community. Some even consider that a good English pronunciation will mean that they are perceived as mocking natives (Tananuraksakul, 2012).

Even if we look away from these social factors, important as they are, and think of someone who may be attempting to match the model, we know that, as with so many physical skills from archery to playing the piano, you cannot hit the target all the time. The natural result is a splatter effect round the target, even for professionals: you can get much better, but you cannot achieve perfection all the time. The more different targets you have to aim at, the greater the splatter effect. If you argue that more communication takes place between speakers of English as an additional language than between people who have English as a primary language, the model keeps people focused on something that allows mutual comprehension. If you allow multiple models (one for Northern Europe, one for the Middle East, one for China, one for the Indian subcontinent, and so on) the splatters can take you so far apart that mutual comprehensibility breaks down, and the whole value of having a 'common language' in English fails. Ironically, teachers of English are the people least able to judge this accurately. If you spend a lot of time teaching native speakers of Vietnamese or Mandarin, you come to be able to understand far more of their output than another speaker of English as a primary language would be able to, and a huge amount more than a speaker from the Indian sub-continent would be able to. Despite some arguments in the literature, multiplying models can easily have the effect of limiting communication between speakers of English as an additional language rather than supporting it.

Can we, then, have a different kind of model – either an artificial one as presented by Jenkins (2000) or another real variety of English? A good candidate for the latter, it seems to me, is Scottish English. It is widely recognised, socially accepted and phonologically far simpler than RP (it has only 12 vowels). It is, of course, no more homogeneous than American English is, but we have that problem to deal with under any set of circumstances. Unfortunately, trying to impose Scottish English, or an artificial variety of English, is no less normative and coercive than trying to impose the dual-model system we already have. It is also open to the same set of problems.

The only option is to teach from a model, but to allow for divergence from the model. The model is there to prevent too much divergence. The divergence is natural because of the splatter effect and the primary language of the learners. The job of the teacher, in this view, is to determine how much divergence can be permitted and where it can be permitted. This means that the teacher must think in terms of what might impede communication (which may depend on the target audience of the learner) and what is relatively innocuous. This is where we turn next.

What is important and what is not

In this section I look at some of the facets of pronunciation that teachers might want to deal with and try to sort out some of the factors that might help them prioritise what to cover in their teaching. The views are my own, though many of them can be found discussed in the literature. Contrasting priorities can be found in many places, including, for example, Jenkins (2000) and Collins and Mees (2003).

Intonation

Intonation is relevant only when the learners can string words together fluently enough to gain a certain amount of fluency. This means that intonation tends to get added after some syntax is learnt, just the opposite of what happens in first language acquisition where children are born already familiar with the intonation of their mothers (Cruttenden, 1994, p. 249). The traditional patterns described for RP statements are not necessarily widespread in other varieties, and a relatively flat intonational contour is usually sufficient for comprehension and is unlikely to be intrusive. On a world scale, in modern English, it is neither true that statements always fall, nor that questions always rise. Learners who use English in the community will be able to copy surrounding norms, but their primary language norms may be perfectly acceptable.

Three points of intonation are worth considering. The first is contrast, which some teachers think of as sentence stress, though it involves intonational patterns. It seems that learners are not good at using this or perceiving it, but English speakers use it to structure information where other languages may use word order. There is an important difference between *It wasn't my BROTHER who saw the flying saucer* and *It wasn't MY brother who saw the flying saucer* that is worth teaching time. Also, worth teaching is the fall-rise intonation pattern which signals a reservation and implies that there is a *but* still to come. The third is less obviously worth teaching time and is of use only to advanced learners. Many English speakers use intonation to mark sarcasm and being able to recognise this can be very valuable, even if students are not trained to use it.

Rhythm

The distinction between stress-timed (or morse-code) languages and syllable-timed (or machine-gun) languages (often including mora-times languages like Japanese and Māori) is well established in the literature, though the simple dichotomy does not do justice to the complexities of the issue (Laver, 1994, Ch. 16). Far more languages are syllable-timed than are stress-timed, and this leads to time spent on teaching weak forms in English and teaching vowel reduction. There are several things which make this seem of lesser importance.

The first is that there are many varieties of English which are syllable-timed, including Indian English, Nigerian English, Singaporean English, and so on. These varieties typically are perfectly comprehensible to speakers of the stress-timed version of English, or at least, when they are not, it is not usually the rhythm that is the problem. We cannot simply dismiss all the excellent speakers of syllable-timed English as in some way deficient on that basis alone (this point was made to me by David Deterding, personal communication). The second is that not all varieties of stress-timed English are equally stress-timed: New Zealand English is rather more syllable-timed than RP is, for example (Bauer & Warren, 2008, p. 61; Bauer, 2015). Thirdly, one of the effects of syllable timing is to make the precise form of prepositions and articles (for example) clearer, and learners may find this helpful. Fourthly, speakers who speak slowly, whether because of hesitancy or because of insufficient experience to be fully fluent, typically use a more syllable-timed version of the additional language anyway, and this cannot be overcome until fluency is achieved.

The real difficulty with syllable-timed English is not its comprehensibility, but the fact that speakers who use it may have problems understanding fluent spoken English from native syllable-timed speakers. Cauldwell (2018) proposes that for students likely to be in this situation, lessons in decoding stress-timed English may be necessary.

Stress

The first thing to say about stress is that it is not incompatible with syllable-timing. Spanish is a language which is at the syllable-timed end of the spectrum, but which has contrastive stress. Having said that, syllable-timing does reduce vowel reduction.

The literature has plenty of examples of incorrect stress causing misunderstanding or lack of understanding. My favourite example (reported many years ago in the British *Daily Telegraph*) is from a vote of thanks provided

to a stress-timed speaker by a syllable-timed student, who referred to the speaker as the 'most important' man in his field but stressed *important* on the first syllable. So genuine problems do occur. Nevertheless, there are thousands of cases where the precise placement of the stress is not crucial to understanding. Even apparently contrastive examples like *export* (noun with initial stress and verb with final stress) are not necessarily used consistently by stress-timed native speakers, and the distinction in stress between *driving licence* and *driving rain* is not always maintained by people reading aloud in, for instance, news broadcasts. Teachers need to be aware of words which might not be understood with wrong stress but need not necessarily spend time in ensuring that every single word is properly stressed. This is an area where more research is required.

Consonants

Consonant sounds often take up more than their fair share of pronunciation teaching time, but the effort needs to be weighed against the achievement of comprehensibility. The pronunciation of [θ] and [ð] (in *thigh* and *thy* respectively) is not as important as it often seems; despite minimal pairs (*breathe* vs *breed* vs *breeze* vs *breve*, *fie* vs *thigh* vs *sigh* vs *thigh*, *thing* vs *sing*, *thought* vs *fought*) there are many native speakers of English who never use either and genuine misunderstanding is likely to be rare despite the famous advertisement from Berlitz (n.d.). The distinction between [r] and [l], on the other hand, is surprisingly important and disruptive, which is problematic for learners who are primary language speakers of Cantonese and Japanese, for example, although the precise phonetic realisation of the [r] may be less important, and the distinction between clear and dark [l], at the beginning and the end of words like *little* and *lull*, may not matter at all. The use of vocalised [l] (that is some kind of [ʊ] sound) in suitable environments (e.g., in *salt*, *spill*, *smile*) works very well for many Chinese learners of English. The difference between [p] and [b] etc. can be seen as a difference in aspiration as much as (or more than) a difference in voicing. Voicing in fricatives is often less important, and many learners get away with no [s]::[z] distinction.

One of the most contentious problems in English pronunciation is what to do with non-prevocalic [r], in words like *farm*, *bear* (this problem is frequently hidden under a discussion of which model to use but see Wells, 2016,p. 151–2 for a brief summary). There is value to keeping a rhotic pronunciation because it reduces the number of vowel sounds that have to be distinguished, it reflects the spelling, and it avoids the problems that arise in expressions like *idea of*, *better off*, where the use of [r] is variable and hard to grasp. There is a half-way house in New Zealand: many New Zealanders (especially, but not only, from Southland, Māori and Pasifika communities) pronounce the [r] only in words like *bird*, *nurse*, *word*, *work* with that specific vowel sound (Marsden, 2013; Bauer, 2015). I suspect this

is not helpful for learners except in as far as it is supported by the community around them.

Vowels

Vowel sounds are the main source of problems for learners. This is not only because there are so many of them (though that doesn't help), but also that the spelling does not accurately reflect the vowel quality, and that knowledge of etymology is in many cases required in order to sort out what phonetic value to assign to vowel letters – something which most learners do not have. Using the labels for Wells's (1982) lexical sets to name the vowels, we get situations like those in Table 1 to Table 4.

Table 1. Different spellings of the same vowel sound

DRESS	ate, bed, friend, head, heifer, leopard, said
FACE	gaol, lay, paid, rate, steak, straight
FLEECE	aegis, bleat, field, Oedipus, scene, seen, visa, weir
FOOT	good, push, should
GOOSE	blue, boot, canoe, flute, grew, guru, rouge, sleuth, tomb
KIT	build, busy, crystal, English, sieve, sit, wanted
LOT	cauliflower, cot, cough, encore, knowledge, wand, yacht
STRUT	cup, does, flood, one, rough
TRAP	bade, cat, plaid

Table 2. Words pronounced with different vowels by different speakers

amen	PALM	FACE	
because	LOT	STRUT	THOUGHT
boutique	GOOSE	GOAT	
covert	STRUT	GOAT	
dance	PALM	TRAP	
economic	DRESS	FLEECE	
leisure	DRESS	FLEECE	
lever	DRESS	FLEECE	
mass	TRAP	PALM	
oestrus	FLEECE	DRESS	
patent	FACE	TRAP	
shaman	TRAP	FACE	PALM
zebra	DRESS	FLEECE	

Table 3. Parallel spellings with variant pronunciations

done	STRUT	lone	GOAT
donkey	LOT	monkey	STRUT
flower	MOUTH	rower	GOAT
mould	GOAT	would	FOOT
road	GOAT	broad	THOUGHT

Table 4. Homophones with different vowel spellings

air	heir	
bare	bear	
ewe	you	yew
I	aye	eye
key	quay	
mew	mu	
nun	none	
peer	pier	
pray	prey	
sale	sail	
son	sun	
steak	stake	
tea	tee	ti
toe	tow	
vale	veil	

The standard notion is that learners must be able to make all phonemic distinctions (see e.g., Cruttenden, 1994, p. 273). This is too much to ask of most of them. The truth is that lack of contrast between pairs which are phonemically distinct will not cause too many problems, as long as there are not too many such cases. Lack of contrasts between the pairs in Table 5 are widely found (sometimes even in native varieties) and can be tolerated.

Table 5. Contrasts which can be lost

KIT	FLEECE	Although these belong to separate systems in English, and although there are multiple minimal pairs, some of them potentially embarrassing (like *shit* and *sheet* or *piss* and *piece*), most speakers are used to hearing English spoken by people who fail to make this distinction.
DRESS	TRAP	Singapore English loses this contrast, and some Australian and New Zealand speakers lose the distinction before [l]. The loss often passes unremarked.
STRUT	PALM	These differ mainly in length in some varieties of English so that a common quality is not unexpected.
LOT	THOUGHT	The distinction is lost in many varieties of North American (including Canadian) English. Despite this, the loss may occasionally cause problems, even in America.
FOOT	GOOSE	Despite much fronting of GOOSE in the last half-century, and the more recent fronting and unrounding of FOOT in many

		places, loss of the distinction does not usually lead to incomprehensibility.
MOUTH	PALM	These merge in varieties of standard British English and in Southern hemisphere varieties, but not consistently.
PRICE	PALM	These can be merged in some varieties, though not usually consistently.

Even if the contrasts mentioned in Table 5 can be lost without causing too much disruption, too many mergers can lead to difficulties. The merger of series like *cot*, *caught*, *coat*, for example, by Spanish learners of English can make individual words incomprehensible, though it is not clear to me whether this is because monophthongs and diphthongs are merged or just because too many vowels are merged.

Voice quality

Voice quality covers a host of matters from typical articulatory settings to laryngeal states to degrees of nasalisation, typical volume, precision of articulation and so on. It is voice quality which gives rise to stereotypes such as 'Italians are very excitable people' and 'The Chinese always sound angry' (for a technical discussion of voice quality and articulatory settings see Laver 1994). Voice quality may be particularly resistant to change, but in some cases, particularly in instances where speakers of tone languages are learning non-tone languages like English, may be worth spending some time on.

Phonotactics

English allows for extraordinarily complex sequences of consonants (although such sequences are often simplified in rapid speech). Consonants which provide difficulty for learners will cause extra difficulties in consonant clusters, and teachers need to think about how to deal with this.

The major phonotactic problems for many learners of English is that you cannot lose word-final consonants. There are many languages which allow no word-final consonants (Sāmoan) or only very limited word-final consonants (Italian, Spanish, Japanese, Mandarin), and that means that final [t], [d], [s] and [z] are often difficult, but these sounds are grammatically important in English so that failure to produce them makes the language sound less grammatical. Adding a vowel at the end (as in a stereotypical Italian accent of English, for instance) sounds very intrusive, even if it is comprehensible. Breaking up word-internal consonant clusters with vowels (as is often done by Chinese learners) sounds equally odd and may be less comprehensible. Because of the general difficulty with clusters (many languages allow no or very limited consonant clusters), clusters at the ends of words are doubly difficult. Extra care therefore must be taken with the clusters

at the ends of words like *dogs*, *life's*, *breathes* and even *pills* and *prince*, where the first consonant is a sonorant. While many textbooks comment on the need to retain clusters – especially in initial position – and give rules for reducing clusters, few point out that for a student having trouble with a [θ] sound, for example, a word like *sixths* causes even greater problems of pronunciation.

Speed of utterance

Although there are common perceptions that some languages (or even dialects) are spoken faster than others, we do not need to worry about that here because it will not affect learners' output, and input from speakers talking directly to learners is likely to be slower than normal to ensure comprehensibility. More to the point is that beginning learners speak extremely slowly, as they struggle with getting the right word and (where relevant) the right inflection. Where production is slow, we expect the pronunciation to be complete and explicit. You may never pronounce the first R in *February*, but if you do it will be when speaking slowly; you may rarely pronounce the A, but if you do, it will be when speaking slowly. Elisions and assimilations arise only once a certain degree of fluency has been acquired. *Handbag* becomes [hambag] only when you are familiar enough with the word to speak it confidently. Elision and assimilation are natural processes that affect most languages (perhaps all, in some form), but they do not affect all languages in precisely the same way. Speakers of languages with no consonant clusters may have no experience of moulding one consonant to the next one. Some of this can be taught as the pronunciation of individual items (like *February*); some more general principles may be needed to teach advanced fluency.

Implications beyond English-teaching

The questions here have been framed in terms of the teaching of English, but the same problems arise in the teaching of other languages, and the same set of questions need to be asked. In terms of the general model that should be used, we find questions with the teaching of Spanish, where Iberian and Latin American models abound, but also in terms of teaching Māori since many Māori people, not wanting to prioritise one dialect over any other, find themselves opposed to the notion of a standard variety of Māori (Keegan, 2017). In this latter case, the varieties are largely mutually comprehensible, but can give rise to social distinctions being drawn.

Where intonation is concerned, English speakers may have a problem with learning languages which have more dynamic intonation patterns, because they tend to sound too emotional. Using intonation as the only mark of questions may be problematic (as in Italian on occasions), or using intonation to mark

grammatical structure (e.g., 'going up at the commas' in French) can be equally difficult. Some facets of German intonation sound very rude to English speakers, who avoid them for that reason.

Where rhythm is concerned, stress-timed speakers of English have great difficulty in not reducing vowels, that is, in using full vowels in every syllable. This is problematic for learners of Spanish and Māori, for example.

Where stress is concerned, English speakers who are used to the notion of words being stressed have problems with a language like French, where stress is a function of a larger unit than the word. They also have problems with putting equal weight on all syllables.

Where consonants are concerned, English speakers have problems with some places of articulation, such as uvular [ʁ] in French or German, palatals in Italian, back consonants in Arabic and Hebrew, but also with bilabial and velar approximants in Spanish. These things are largely predictable, but some evaluation of how important they are is required, and some thought needs to be given to strategies for getting students to produce them (or suitable approximations to them). Geminate consonants in Italian (and also in Japanese) require some training but are usually not too difficult to teach – the problem is knowing just when to use them. Glottal stops in languages like Sāmoan are difficult, not because English speakers do not use them, but because English speakers are not aware of them and cannot hear them.

Vowel qualities in other languages are not often problems for English speakers, except for languages with rounded front vowels (French, German, Mandarin). As I was told by colleagues in the French Department when I first arrived in New Zealand in 1979, New Zealand learners, who half a century ago could pronounce *vous* [vu] in French but not *tu* [ty], now find themselves unable to pronounce the *vous*-vowel, which needs to be equated with New Zealand English *school* rather than New Zealand English *goose*. Japanese unrounded-U is also difficult. Languages which have distinctive vowel length do create problems, though note that many modern Māori speakers, under the influence of English, are losing contrastive vowel length (marked with a macron) in their language. Languages which have a lot of diphthongs (Cantonese, Dutch) may cause difficulties in keeping the various distinctions separate.

Phonotactics do provide problems for English speakers who have difficulties at first with word-initial [ŋ] (e.g., in words like *ngaio*) and lasting difficulties with word final [ɛ] (in Māori *wheke* 'octopus', French *donnait* 'he/she/it used to give', Spanish *doce* 'twelve', Italian *come* 'how'). Surprisingly, given that most New Zealand speakers have word-final [ɔ] in words like *fore, maw*, final [ɔ] is often

replaced by a diphthong (the GOAT vowel), perhaps because the [ɔ] is so often associated with a written R-letter.

Voice quality can, to some extent, be approximated by ensuring an appropriate hesitation vowel (the equivalent of what is written as *er* in English). The French use of [ø], for instance, indicates the importance of lip position in French, and moves the tongue forward from where it would be in English.

Just as spelling causes problems for learners in predicting the appropriate vowel to use in English (and also in failing to distinguish between [θ] and [ð]), there can be problems caused by spelling in other languages. Although Spanish is often cited as a language whose orthography is helpful, having C and Z representing the same sound is potentially awkward. French has large numbers of homophones with different spellings (e.g., *cent* 'a hundred', *sans* 'without', *sent* 'smells') where the writing can be more difficult for those who can pronounce the language than for those who cannot.

Conclusion

I remember being asked, many years ago, by a pianist, how to pronounce a certain sound. I mimicked it for him, and said, 'Like that.' He got very cross with me. He said if he wanted to teach someone how to play something on the piano, he told them where to put their fingers, what position to have their hand in, and a number of other details, he didn't just play it for them and say 'Like that'. I was right in the sense that we do not all have a built-in capacity to mimic playing the piano, although we do all have a built-in capacity to copy speech – it is how we learn in the first place. He was right in the sense that our ability to copy speech, if not kept alive by constant practice, tends to fade, and we might need attention drawn to things we cannot notice, might need to be told explicitly about lip position or about glottal stops or aspiration. Copying is a very good strategy as far as it works, but when it doesn't work (and that is at different points for different learners), the teacher needs to be able to support further learning. Presenting a model for students may be a good first step, but it is not all that is needed to teach pronunciation.

My overall message in this article is that just what is taught in terms of additional language pronunciation cannot have a single solution for all instances. It cannot be found in a single textbook. It needs to be determined in the light of the language background of the students, the goals of the students, the personalities of the students, the expertise of the teacher, the goals of the teacher, and so on. Aspiring actors and spies will need far more support than aspiring holiday-makers or aspiring business people. Students aiming for credentials in international examinations will have to consider the requirements of the tests they wish to pass.

What I see as important about this message is that it puts the teacher at the heart of the teaching. The teacher has to decide what to model, how close to the model the students need to be able to get, what aspects of the model can be ignored without losing too much comprehensibility, and so on, as well as how to turn phonetic knowledge (either inbuilt because they are native speakers of the relevant variety or learnt) into useful support for the students.

But if the teacher is key, that does not mean that the teacher with no background in phonetics can just waive responsibility and hope that learners will copy enough of the presented model to be successful learners at some level. The pronunciation teacher has to know enough about the target language and about the primary language of the students to be able to predict the problem areas and to predict where intervention will be required, and then to formulate a plan for teaching those parts. The knowledge of the learners' language may come from books or may come from experience, just like the knowledge about the target language. Pronunciation teaching need not be terribly technical – teaching a student phonetics before teaching them another language might be an ideal, but is impractical – but it needs a solid knowledge base, and it requires consideration of the relevant factors and prioritising of goals.

References

Bauer, L. (2015). Australian and New Zealand English. In M. Reed & J. M. Lewis (Eds.), *The handbook of English pronunciation* (pp. 269-285). Wiley Blackwell.

Bauer, L & Warren, P. (2008). New Zealand English: Phonology. In K. Burridge & B. Kortmann (Eds.), *Varieties of English Vol 3: The Pacific and Australasia* (pp. 39-63). Mouton de Gruyter.

Berlitz n.d. German coastguard. https://www.youtube.com/watch?v=KLSdOY-6R_U

Cauldwell, R. (2018). *A syllabus for listening: Decoding*. Speech in Action.

Collins, B. & Mees, I. M. (2003). *Practical phonetics and phonology*. Routledge.

Cruttenden, A. (1994). *Gimson's pronunciation of English*. Arnold.

Jenkins, J. (2000). *The phonology of English as an international language*. Oxford University Press.

Keegan, P. J. (2017). Māori dialect issues and Māori language ideologies in the revitalisation era. *MAI Journal: A New Zealand Journal of Indigenous Scholarship, 6*(2), 129-142.

Laver, J. (1994). *Principles of phonetics*. Cambridge University Press.

Marsden, S. (2013). *Phonological variation and the construction of regional identities in New Zealand English*. [Unpublished PhD thesis]. Victoria University of Wellington.

Tananuraksakul, N. (2012). Non-native English students' linguistic and cultural challenges in Australia. *Journal of International Students*, *2*(1), 107–115.

Wells, J.C. (1982). *Accents of English*. Cambridge University Press.

Wells, J.C. (2016). *Sounds fascinating*. Cambridge University Press.