

Learning and Teaching Vowels

Piers Messum

Introduction

Learning unfamiliar sounds and sequences of sounds has not come easily to me. Over the years I have spent many hours struggling with the pronunciation of French and Japanese, bewildered by my inability to sound authentic in either language and often by my inability even to know whether I sounded authentic or not. In itself this is not much of a qualification for writing on a subject, but puzzlement has spawned curiosity and with a great deal of help from other people I think that I have made progress both as a learner and teacher of sounds, and that I now have some understanding of the process involved.

The ideas below are connected as follows. The first section is about how we learn vowel sounds, both as infants and as students. If the account I give of this is correct, then the most common approach to teaching vowels is misconceived and therefore likely to fail with many learners. As an alternative I describe an approach which many teachers have found to be more successful.

This alternative approach needs particular tools and techniques to support it, including a model of the vowel space suited to the task of teaching. At present language teachers use the IPA quadrilateral from phonetics to meet this need, but in the second section of this article I explain why I think we should abandon it. Again, I suggest an alternative, the TTL (or TJL) model.

In the third section I describe a vowel chart based on this new model which incorporates other features of what we could call 'pedagogical' rather than 'academic' phonetics. Happily, developing this distinction does not require us to adopt any fictions; just a different view of the phonetic facts that is adapted to language teaching rather than academic research and cross-linguistic comparison.

My own teaching has been hugely influenced by Caleb Gattegno and his Silent Way (SW) approach. This can arouse strong feelings. It is sometimes misidentified with particular tools and techniques that he developed, and criticised when these don't, on their own, turn out to be a panacea for the challenges they address. Alternatively, when it is understood that the SW should actually be seen as one particular implementation of a general principle, that of the subordination of teaching to learning, people can resent the way it seems to hijack what could otherwise just be called 'good teaching'.

Personally I feel that what Gattegno did and argued for was distinctive and rigorous enough to merit its own name, but I sympathise with those who are not able to agree. I rely on his ideas and quote him at length below, so perhaps I should add here that when reading Gattegno it can help to keep in mind that he was a French speaker before he was an English one. His writing often seems to me to be thoughts in French that have been expressed using English words, and the same may be true for the technical terms he developed and which I have adopted.

1 Learning Vowels

Before we can decide how to teach vowels we should, of course, investigate how vowels are learnt. This section starts by unpacking the superficial explanation that they are learnt ‘by imitation’ into a more detailed account of what we do both as infants learning a first language and as students learning subsequent ones.

To do this, I examine some of the very different meanings we have for the word ‘imitation’ and the consequent need for us to distinguish between them when involved in a pedagogical discussion. I then present Gattegno’s account of the process of vowel (and other speech element) learning. This seems to me to describe both my own experiences as a learner and what I believe I have observed in others when teaching.

Gattegno’s theory explains why common approaches to pronunciation teaching are unsuccessful, particularly when judged by the collective achievement of all the students around the world who are subject to them, rather than just by the achievement of the relatively few who end up with competence in English pronunciation. His ideas can justify us abandoning conventional wisdom in this area. They also point the way to a practical alternative approach, that many teachers and students have already found to be successful.

Learning by imitation: mimicry, emulation and ‘PLI’

Imitation turns out to be a complex, problematic and controversial field (particularly when imitation in animals is included). However as a first definition we could say that imitation occurs when someone purposely does something similar to what he perceives another (or others) to do. This allows us to say, for example, that children learn to speak ‘by imitation’. After all, it is the language of the speakers around them that they (eventually) learn as opposed to any other, so they must have made use of this model in some way.

At this superficial level ‘imitation’ is functioning as a cover term, in the same way that ‘red’ can sometimes be a useful cover term for what lies within a wide range of shades of colour. However, such terms lead to confusion when they obscure differences that are significant, and for our purposes we need to examine the shades of ‘imitation’ more closely.

One sense of the word is that of mimicry or impersonation. Here the aim is to reproduce salient aspects of how an act is performed (ie the process) so that the overall impression given is recognisably similar to the original. The way we do this is first by abstraction – stressing some aspects of the model performance and ignoring others – and then by making use of our existing capabilities to reproduce the features we have selected as significant, or by installing new capabilities to enable us to do this if necessary.

Notice that the final result of mimicry doesn’t have to be authentic (although this possibility is not excluded). The Swedish cook in the Muppets could mimic speaking Swedish by adopting a sing-song pattern of intonation without any real language being produced at all.

A second sense of imitation is emulation. This is a less common word, and its use among different speakers and in different dictionary definitions is quite variable. In its technical sense, emulation (or ‘goal emulation’ (Heyes 2001)) is imitation that attends to the final result rather than the process involved. Thus we can say that George W. Bush has emulated his father’s achievement in becoming president without suggesting that he did this using any of the

same methods or by taking the same route to the top. Similarly, manufacturers of electronic equipment produce ‘emulation boards’, which perform the same final function as the original from another manufacturer but do not use the same underlying processes or programming code (to avoid infringing intellectual property protection).

For the learning of speech sounds I would suggest that emulation is the sense of imitation that we need to take. Our concern both as children and students is, after all, to achieve a final result which matches what speakers around us achieve when they use vowels and other speech elements. In contrast, while mimicry is a field that we all explore at certain times in childhood and that some of us find compelling enough to develop into an entertaining skill, we do not end up mimicking our parents in our own speech, even though we clearly could have chosen to do so¹.

Mimicry and emulation complement each other by being concerned either with the ‘imitation’ of the process, or the ‘imitation’ of the end result of a model’s actions. A third type of imitation is different again. On occasions we take advantage of another’s expertise by breaking a complex process down into the stages or sub-goals we observe in the model execution to guide us in our own learning. Ie, we imitate the structure while furnishing the details of the actions ourselves. This has been called ‘program level imitation’ (PLI) by Byrne and Russon (1998), and suggested as a description of how a child learns to tie a shoelace: observing the stages involved in the model execution of the skill and then learning for himself how to execute each stage, perhaps through trial and error.

When a child’s productive gestures for phonemes and syllables are in place and a new, multisyllabic word is perceived in the environment PLI may be the best description of how its pronunciation is acquired (Byrne and Russon 1998:716). If the word is sounded out slowly by an adult the ‘sub-goals’ involved in its production are made clear and the mastery of its production can then be relatively quick and effortless.

Learning speech sounds

I have already suggested one reason why we should not view normal sound acquisition as involving mimicry. Gattegno (1973) makes a more general case for this,

No one can learn to speak by imitation simply because we hear with our ears and speak with our vocal system. The one system is submitted to the given impacts from the environment and the other is a totally voluntary system.

In the context of this discussion we might say that, “No one can learn to speak by imitation *alone*”, because at some stages in the speech acquisition process we may consider certain forms of imitation to play a role as I described above. Gattegno’s point, though, is that there must be more to the process than mimicry. Imagine a child born dumb who at age 10 regains use of a paralysed tongue. He has heard and understood speech for some time so would he be able to imitate it immediately? Of course not. There would have to be a prior process of education of his tongue that would be personal and unrelated to anything we would describe as mimicry.

¹ Other instances of emulation in language acquisition will occur when we are inspired by the persuasive, coercive, consoling, humorous and other results that speakers we observe achieve with language and attempt to achieve these ourselves.

When it appears that someone can immediately mimic something then the gestures involved must have been installed at an earlier time, quite possibly for a very different purpose, and just happen to be adaptable to the present circumstance. Young (2000) makes this point for adults,

When I was a child, I learnt to walk a tightrope. So if I installed a rope between this rooftop and that one, and told you, "Now off we go. Just do like me." would you try? Of course not, because you know as well as I do that you can't imitate me in this. You have to develop the sensitivity to your centre of gravity and all the other technical skills and the muscular power in your feet and in your abdomen which will allow you to do it.

This is an extreme example, but a little thought shows that in all circumstances, without exception, it is only possible to imitate what one can already do. If I don't already possess the gesture I can't imitate someone else doing it.

If imitation were part of the learning process we could all be champions in any discipline we wanted to. Just watch and do.

Imitation exists of course, but when someone is imitating they are not faced with the unknown. They are using skills which they already possess.

Deciding that emulation is a better description of sound acquisition than mimicry is a step forward which will help us to understand what we need to avoid in our teaching. However, it does not tell us how children (and adults) do install the gestures of the tongue and other articulators that they use to speak. We now need to examine this part of the process.

Gattegno's proposals on speech sound acquisition

In 'The Universe of Babies' (1973), Gattegno describes how babies and infants develop the ability first to talk and then to speak in terms of their developing awarenesses of themselves and the environment².

"When a baby makes sounds in his crib he will be able to hear them if he wills it, and if his ear and acoustic nerve function properly. As soon as he acknowledges that it is he who is making the sounds that he hears, he has at his disposal two sources of information, his throat and his ear, to investigate the world of his own utterances." p.64

"From his *in utero* experience [a baby] knows how to affect the muscle tone of every muscle involved in phonation. This knowledge is conscious, and therefore phonation is a field he can reach in as much detail as he needs. Hence he can study separately or in conjunction the parts played by an air flow impinging upon the various components of the chamber that a closed mouth represents." p.64

"Every baby a few weeks old does this during those hours of crying that no adult any longer understands. This necessary knowledge in terms of muscle tone gives the muscles the quality of being instruments for knowing because awareness of a set of impacts informs the self what is happening in the chamber." p.65

² I quote Gattegno at length here partly because I find his description of babies at work so vivid and exciting and partly because the extracts illustrate how reexamining questions of learning in terms of awareness and awarenesses can give such rich rewards. In the early 1970's, when he was writing, the mainstream view of the learning baby was very different from the current one, which is increasingly in line with Gattegno's overall vision of babies as alert, pro-active and intelligent learners from the start.

“For the non-deaf baby the two systems of conscious experiences, concerning utterances and impacts on the ear, are put into correspondence so that the ear and the mouth are functionally connected. The time needed to create this systematic correspondence is found in the crib when explicit experimentation takes place.” p.67

“Babies have no difficulty in emitting a sound and noting its impact upon their ear. There is no danger that the baby will confuse this sound with the sounds of the environment since the utterance can take place at any time of the day or night – at those times that the baby knows himself to be left alone and can be sure that *he* and no one else has uttered *this* sound. He at once gives himself a test by doing it again, and, if need be, again and again.

Concentrating on both ends, the production of the sound and its reception, he knows what he does to make the utterance and almost at once he also perceives the particular alterations in his ear and the part of the brain connected with it. He can therefore keep the two awarenesses connected, or else voluntarily emit a sound and switch his attention to knowing it as an entity reaching his ear. When satisfied that his ear is as good an informant of what his throat does as his own direct contact with the throat (through consciousness), he can relax one of the watchers and leave one system to keep vigil over the functions of the phonation system. From then on the ear holds the feedback mechanism to inform consciousness of what the throat is doing when uttering sounds.” p.67

“The ear now holds the clues that make the mouth capable of interpreting the sounds that it hears in terms of utterance. Since a baby can evoke all he knows, he can play at the virtual level the game of giving himself a symbolic sound, evoked in the ear or the brain, and then discover whether the organs of phonation take the position and shape necessary to produce the sound if they were actualized by an additional amount of energy.

The key to the entry into a language is in this apprenticeship: the study of the parts of the mouth which correspond to the sounds that are heard.” p.71

For understanding Gattegno’s account it may be helpful to add one more quote, which does not seem as challenging now as it must have done to readers at the time, but may still be ahead of conventional wisdom.

“... my not knowing *now* what I did to achieve such a thorough knowledge of my skin that it now permits me to coordinate the actions needed to chase or crush an insect creating a sensation on my neck or back ... will not preclude that I did work consciously to become aware of how to interpret messages before entering into an action. The development of the somatic self – that is, *in utero* – is accompanied by a process of awareness.

However reluctant readers may be to adopt the assumption that somatic consciousness has been as vivid and as well known to the self as social or intellectual consciousness is today, I suggest that they entertain it for a while in order to find out what help it provides in understanding babies, or for that matter any age and any condition.”

Young (1995) summarises how Gattegno’s ideas apply to foreign language learners learning completely new sounds,

“First of all, learning a new sound requires that the student realize that there is in fact a new sound to learn. He can then try to create the sound. In this case he is dealing with

two independent but closely related systems, the mouth and the ear. Only one of these systems, the mouth, can be controlled voluntarily. All the muscles of the ear are involuntary muscles. The student can only modify the voluntary system. With his mouth he produces a sound which he guesses might be as close as possible to the sound he is aiming for. He hears the sound with his ears. Since he produced it with his own mouth, he knows that, muscularly speaking, his mouth was used in a new or special way and consequently he knows he should listen for a sound which is different from what he usually hears. He can probably predict at least to some extent in what ways the sound will be different from what he usually produces. He speaks here with the deliberate intention of hearing something unusual and he listens to the result with the specific intention of hearing this unusual sound he has produced, creating a double feedback loop. He has feedback from his mouth telling him what it is doing and his ears give him feedback about what changes they detect as a result. Gattegno proposes that this is the process we all use to learn to produce new sounds.

Once the student has managed to produce the sound to his satisfaction, he must practise it in a wide variety of different situations and contexts until he is completely at ease with the sound. He then reaches a stage where the sound has become completely automatised and the learning process for that particular sound is over.”

I do not know if Gattegno’s account has been tested explicitly in the SLA literature, but with respect to babies and infants I have paid attention to the published research on language development and I have come across a great deal of evidence that would appear to be consistent with it and none that contradicts it. Indeed his overall vision of babies and how they use their time, which seemed so unconventional and ‘unscientific’ when he proposed it, could probably now be described as the mainstream view. As I mentioned earlier, through observing myself and others I have found that we are engaged in exactly the activities he describes.

A summary of where we are now may be helpful. I have argued that the process of learning new vowel sounds is not one of mimicry but one of emulation. This, however, does not tell us about what is happening at a detailed level of learning. So I have presented Gattegno’s idea that learning a new vowel involves a process of interaction between the learner’s vocal apparatus and the learner’s ear, and one which requires his presence (or attention) to what is going on in both of these systems. This applies whether the learner is an infant or an adult. We can now ask about the implications of these ideas in the language classroom.

What this means for the teaching of vowels

A common approach when teaching a new vowel sound is for the teacher or a recording to provide a model, and for the student to be asked to repeat it. A variant of this is when multiple instances of a vowel are presented first, in order, it is hoped, to sensitise the learner’s hearing. The student is then invited to try it for himself. (The rationale being that until the student can ‘hear’ the sound he won’t be able to repeat it, and that being exposed to multiple tokens is going to improve his ‘hearing’.)

In most cases this doesn’t work. The failure of this approach is even more evident if we consider all the students throughout the world who undergo this procedure and not just the very few who retain enough enthusiasm to persevere with learning English and end up in our classes. (I will explain shortly why I think it does ‘work’ in some cases.)

The reason for this may now be clear. When a teacher asks a student to ‘repeat after me’ he is effectively, if unwittingly, asking the student to engage his powers of mimicry. After all, as a

learner asked to ‘repeat after me’ I am immediately aware of two things: firstly that the criteria I have for discriminating sounds are inadequate (I cannot hear the difference that is so clear to the teacher), and secondly that I have no idea what to do to produce this thing I cannot ‘hear’. I do, however, know that a performance is expected of me, so I do what I can. I remain present to the (inadequate) internal representation of the teacher’s model that I have created and I produce what I can by mimicry, using articulatory gestures I installed in myself in the past.

What is clear is that the task proposed has led me to take my presence away from where Gattegno suggests it actually needs to be – mainly in my throat and only at a secondary, feedback level in my ear - if I am to learn to produce and discriminate this new sound. Even if my rendition was acceptable to the teacher at the time I will not have developed any criteria for production on future occasions in the absence of a model to mimic³.

How is it, then, that anyone ever makes any progress in pronunciation when the ‘repeat after me’ approach to teaching is used? My assumption (drawing, again, on my own experience) is that successful students subvert the process, either by not following the instructions the teacher has given or by supplementing them. Aware at some level that they need to be present to their vocal apparatus and ear to learn a new sound (and not to be using their powers of mimicry) this is where their attention goes, either during the time the teacher is working with them or afterwards.

Many students, though, will scrupulously follow the instructions they have been given and assume (not unreasonably) that doing this should lead to progress. How much better it would be for us to propose activities which encourage all our students to be present to what they really need to do to learn a new sound⁴.

Gattegno and other SW teachers have developed many techniques for resisting mimicry on the part of the student, although a refusal to provide models is the most important one. At the same time the so-called ‘articulatory approach’ to teaching new sounds has a long tradition of support within phonetics, and has certain parallels with Gattegno’s ideas in its pedagogical justification. In recent times it has been most consistently championed by Catford (1987, 2001), who made experimental attempts to prove its superiority (Catford and Pisoni 1970) although no one to my knowledge has consolidated (or contradicted) the positive initial results he obtained.

When a teacher has decided not to provide model utterances, there is an even more pressing need for ways to help students to become more aware of what they must do with their articulators to produce new sounds. It is relatively easy to suggest manoeuvres for most

³ A minor caveat to this argument is that some aspects of vowel production are amenable to direct perception, and then the use of pre-existing gestures to reproduce them. Lip rounding is an obvious example. However our primary channel for vowel perception is a sense of hearing that has not been educated to perceive the sound at issue, and which is best educated through a dialogue with our own vocal apparatus.

⁴ One of Gattegno’s techniques for doing this, silence on the part of the teacher, was so striking to observers at the time that he called his whole approach to foreign language teaching the ‘Silent Way’. In retrospect this was a bit unfortunate, since it makes the approach sound both like an exotic, spiritual path and like something that is only really concerned with pronunciation. Neither of these things are true.

consonants, where landmarks within the mouth are guides, but rather difficult for vowels and vowel-like consonants.

Part of the solution is for the teacher to encourage and to be content with trials and many errors on the part of students as they resensitise themselves to their articulatory apparatus and its potential, and to accept progress that comes step by step rather than in a revelatory instant. In addition, though, we can look to phonetics for a model of prototypical (or perhaps hyperarticulated) production that we can use to guide our students.

At present, the vowel model that teachers are most familiar with is the IPA quadrilateral. My next section argues that this is actually not suitable for our purposes, and suggests something we might put in its place.

2 Pedagogical models of vowel production

While there are teachers of English who have successfully developed their own ‘pedagogical phonetics’ or who have incorporated standard phonetic ideas in their work, many teachers I have met admit to feeling inadequate in this area despite, in many cases, having been exposed to a phonetic element in their training courses. Not unnaturally they are reluctant to instruct in an area where they feel they lack expertise.

(As an aside I might add that this lack of confidence is also common among British speech and language therapists, who have a significant component of speech science and phonetics in their extensive professional training. A common complaint I hear from them is that, “I never understood phonetics.”)

We will help teachers to overcome a lack of confidence if we can present a coherent model of English pronunciation that will enable them to give practical assistance to their students. We should be clear that we are not aiming to produce phoneticians. Instead we need to develop a pedagogical phonetics of English to complement the pedagogical grammars of the language that exist and serve us well.

With respect to vowels there are two principal tools on offer to teachers at present. The first is a set of phonemic symbols which help to disentangle the sounds of the language from its spellings. We certainly need labels for speech sounds, and the main alternative to the IPA symbol approach is the device of using combinations of letters whose ‘natural’ reading expresses a particular sound (“ee” for /i:/, for example). One significant problem with this approach is its idiosyncrasy from author to author and in general I think that the use of IPA symbols is preferable to this⁵.

The second tool is the IPA quadrilateral (IPAq), which aims to illustrate the relationships between vowels by their position within it. How it does this and how most of us think it does it, turn out to be two rather different matters, which I start by discussing below. This leads me

⁵ For many classes, however, the use of colour coding as seen in the SW charts is better. Misleading L1 associations between written forms and sounds are avoided, fossilised mispronunciations can be bypassed, and the system can be extended to word and spelling charts which then combine orthography with phonemic colour coding.

to the conclusion that the IPAq is unsuited to any role within EFL. (What phoneticians do with it, of course, is a completely separate matter, although it has its critics within this community as well, e.g. Butcher (1982) and Lindblad (2001).)

We do however need something to perform the role it purports to play, and what I think we need stems from the understanding of the process of learning sounds that I described earlier. Later in this section I describe what I think our needs are, and how an alternative approach to the description and classification of vowels might serve us better.

What we think the IPA quadrilateral represents

At two recent IATEFL events, the 2001 conference in Brighton and the subsequent PronSig day in London, I asked the audience what the IPAq labels ‘Front-Back’ and ‘Open-Close’ referred to. Without a single dissenting voice I was told that they describe the position of the tongue in the mouth (by reference to its highest point in a given vowel articulation).

The EFL books on teaching pronunciation that I have at home (Ginésy 1995, Kenworthy 1987, Underhill 1994) tell a similar story. The books by phoneticians written to be accessible to language teachers (Gimson 1989, Roach 1991) and the new editions of the main pronunciation dictionaries *appear* to be saying the same; although careful reading shows that the authors are usually scrupulous about not asserting any such thing.

With good reason. They do not assert it because it is not the case.

We can trace the roots of the IPAq back to the Bells, father and son, who believed that their classificatory system based on tongue height was both physiologically and phonetically correct. Daniel Jones incorporated some of their ideas into his Cardinal Vowel system, and at this point there starts to be some controversy among phoneticians as to what aspects of articulatory reality were and what were not being represented. Russell’s x-ray analysis of vowels (Russell 1928) were a warning to anyone who believed in what he labelled the ‘tongue arching theory’, and the development of the modern theory of speech production in the 1950’s finally made it untenable.

It is now understood that the most significant gesture that the tongue makes when we produce a vowel is the one that has the greatest effect in dividing the vocal tract (the space between the larynx and the lips) into two resonant chambers. This ‘point of maximum constriction’ is sometimes the same as the highest point of the tongue below the hard palate, in line with the Bell model (for example when producing [i] type vowels); but often it is not.

Over the course of recent human evolution our larynx has descended in the throat to the point where, while we have exposed ourselves to the danger of food ‘going down the wrong way’, we have gained the opportunity of making new divisions of the vocal tract through rearward gestures of the tongue towards the back of the throat wall. In this way we can make so-called pharyngeal speech sounds that include /ɑ:/ and /ɔ:/ type vowels, which are not properly represented by reference to the highest point of the tongue⁶.

⁶ To be strict about this, we should note that the tongue is a ‘muscular hydrostat’, meaning that it conserves its overall volume whatever shape it adopts. So a pharyngeal gesture, translating some of its volume rearwards, will result in a more or less reliable deformation of the rest of the tongue, defining, of course, the position of a highest point below the palate. Hence Whalen, Min Kang, et al. (1999), who

Phoneticians have been aware of this problem for some time, and of the accumulation of X-ray evidence which seems to show, for example, some speakers articulating /e/ with a higher tongue position than the one they use for /i/, and so on (Wood 1979). They now view the IPAq as defining an auditory/acoustic space, and the labels ‘front-back’, ‘open-close’ and ‘high-low’ are given no formal articulatory significance. Instead they are understood to describe the first and second formant frequencies of the sounds. This is rarely made explicit in books written by phoneticians for non-specialists, but an honourable exception is the unequivocal statement in Ladefoged (2001).

The usefulness of the IPAq to language teachers

An auditory/acoustic space may be useful to phoneticians for cross linguistic comparisons and other purposes. However I cannot see it being of any benefit whatsoever to language teachers faced with students who want to learn just one language, particularly since the training necessary to use the cardinal vowel system is highly specialist, not widely available and very time consuming. It is impractical for any but the most pronunciation obsessed language teacher to master it.

Instead, I would suggest that we need an articulatory model that allows us to distinguish all the vowels of English in a ‘distinctive features’ fashion. This will not resolve all the difficulties students face nor will it mean that lessons proceed dramatically differently from now. Feedback on how close students are getting to the auditory targets of vowels will continue to be given by the teacher (without, of course, reference to formant values). However a genuine articulatory model will give us some reliable pointers as to how our students might work in order to improve their attempts, and give the students some ‘triggers’ for reproducing success.

Obviously such a model should be as phonetically authentic as possible. But it will be a pedagogical model, and it can therefore be based on prototypical gestures and hyperarticulations rather than the reality of skilled, economical adult speech if this will give us better means to help our students in practice.

The throat-tongue-lip model of vowel articulation

As an alternative to the IPAq, Wood (1993) describes the throat-tongue-lip (TTL) model of vowel articulation. This recognises

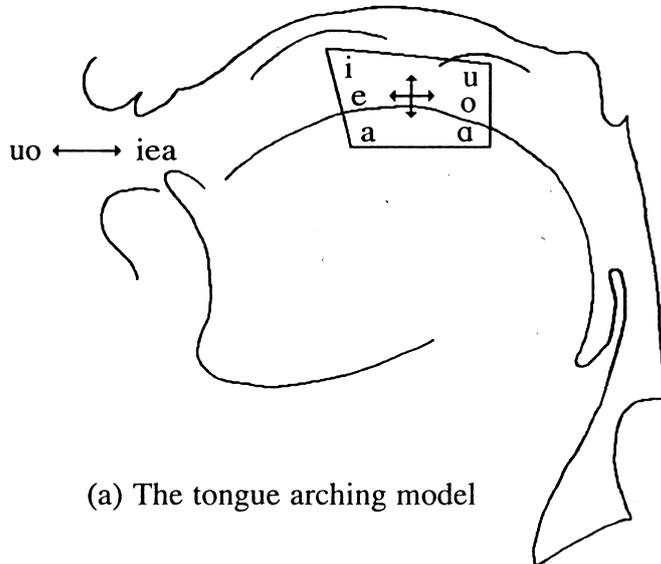
- three tongue manoeuvres (towards the hard palate, the velum and the back wall of the pharynx)
- jaw position
- lip configuration

and describes vowels using these parameters. (See figure 1.)

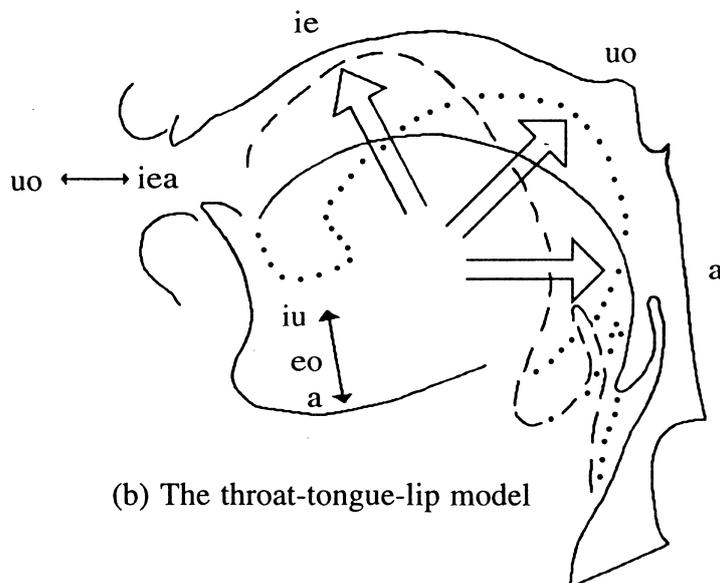
He traces its long history prior to the introduction of Alexander Melville Bell’s model in 1867, and points out its

investigated tongue positions with magnetic resonance imaging, were able to conclude that (p.600), “The IPA reliance on tongue position in the oral cavity for defining the vowel symbols ... can be considered roughly accurate [for English].” This does not mean that the IPA approach is of any help to teachers and learners, for whom the actual articulatory gesture is significant - not the unintended side-effect.

correspondence [with] modern speech production theory inasmuch as the ancient model was based on manoeuvres that cause local narrowing or expansion of the vocal tract at what can be recognized today as spectrally significant locations.



(a) The tongue arching model



(b) The throat-tongue-lip model

1. The tongue arching model assigns to each vowel a unique position of the top of the tongue arch in a coordinate system. Raising, lowering, advancing and retracting the tongue arch anywhere in the quadrilateral yields a new vowel defined by the new coordinates. The throat-tongue-lip model recognized three tongue manoeuvres (palatal, velar, pharyngeal), jaw position and lip position.

Wood (1987) describes his preferred variant of this historical approach, creating four classes of tongue gestures: palatal, velar, upper pharyngeal and lower pharyngeal. Within each class vowels are differentiated by variations of lip position, tongue blade, posture, larynx depression and jaw opening. How the RP English vowels map onto this is explained in the next section.

The TTL model may not be ‘true’ in the sense that all speakers produce vowels in the way it specifies all of the time, but it describes articulatory prototypes based on gestures that should be understandable and usable by teachers and students alike. Suggestions for experimentation based on this model should enable us to get students into the cycle of creating new articulations with promising characteristics and listening for the differences these make, as described in section 1. The one modification I would like to make would be to re-christen it, for EFL purposes, as the tongue-jaw-lip (TJL) model, thus making the primary parameters we would be working with a more prominent part of its name.

The IPAq is enshrined in many language teaching materials, but not used as it should be as an auditory model because it is impractical for teachers to use in this way. Instead it is appealed to, inappropriately, for the articulatory guidance which teachers find they do require. My proposal is simply that we replace the IPAq by a model that actually represents vowel articulation. To illustrate this proposal I describe the implementation of such a model in teaching materials in the next section.

3 A Silent Way sound/colour chart for British English

For some years I have been working with Roslyn Young of the Centre de Linguistique Appliquée (CLA) in Besançon on new ways of arranging the sounds of English on a phonemic chart for best pedagogical effect. Our overall aim has been the design of a complete set of Silent Way materials consistent with an RP/Southern British English (SBE)/Australian etc model, to complement the General American set published by and available from Educational Solutions Inc., New York and Education Solutions Ltd, Reading. We have been greatly assisted by Roslyn’s colleagues at the CLA.

One of Gattegno’s innovations in his original SW and Words in Colour charts (SW for foreign languages and WiC for native speaker literacy) was his colour coding of phonemes and his subsequent use of these colours in sets of word and spelling charts (which he called fidels). When used this way, with sound/colour correspondence, I believe that our design would fall within the copyright owned by Educational Solutions. However, the idea of arranging phonemic symbols on a chart predates Gattegno, and it is in this form that I am presenting the design here. In that sense, then, the discussion that follows is independent of the Silent Way.

Having a phonemic chart available is a boon in any class where pronunciation is going to be taken seriously. If nothing else, at least one aspect of the challenge every student faces is immediately apparent: “To speak this language well I will have to be able to produce and discriminate between all these elements – but no more.”

Perhaps the first part of this thought might daunt a Japanese or Spanish speaker moving from a 5 vowel system to one of 19 in English, but we should not underestimate the confidence and inspiration that is engendered by the second part, as the chart gradually becomes familiar and the student can feel his mastery of all its elements growing. Having a task well defined from the outset is an empowering device.

Beyond this, a chart enables sounds to be referred to without a spoken model being produced by the teacher (in SW teaching usually by using extensible pointers). One of the reasons why this is desirable – taking mimicry out of the classroom - was explained in section 1. A further

benefit is that the particular sound (or sequence of sounds) that a student is working on is unambiguously signalled to the rest of the class in this way. This is often not the case when no external label for a sound is available, and an attempted discussion of a mispronunciation gives rise to general confusion.

The purpose of this section is to explain the arrangement of one version of the chart that we have developed (figure 2.), and to point out some of its benefits in practical teaching. To some extent this will involve drawing on principles developed by Gattegno within the SW. The extent of his influence on us is, I am sure, already clear.

The overall design of the chart

The particular chart reproduced here is the version of the basic design that depicts the greatest amount of phonetic detail. It reflects a degree of sensitivity to pronunciation required of the undergraduate students of English that I have taught in France. A simpler design, omitting much of the detail in the bottom row and with some other rectangles arranged differently, is currently in use at the CLA, and is probably more appropriate for a general public. For readers of *Speak Out!*, though, it seems best to describe the more detailed version which can always be simplified to what may be appropriate in a particular teaching context.

The chart is divided into three sections: strong vowels at the top, consonants in the middle and weak vowels at the bottom. The semi-vowels /j/ /w/ and /r/ (produced as vowels in RP but used as consonants) lie on the main dividing line immediately below the particular family of vowels with which they share a basic articulation.

The innovation in the chart lies in its treatment of vowels. In the previous section I argued that a major analytical tool in phonetics, the IPA quadrilateral, is of no relevance to EFL teaching. In its place I proposed that we adopt the 'Tongue-Jaw-Lip' (TJL) model, and the arrangement of the full vowels is based on this.

IATEFL 2001 Chart devised by Piers Messum and Roslyn Young, 1993-1999 Please do not reproduce without acknowledgment							
— iə —	— eə —	— uə —					
	— eɪ —	— ɔɪ — — aɪ — — aʊ — — əʊ —					
ɪ	e	ʊ ɒ ʌ æ					
i:	u:	ɔ:	ɑ:	ɜ:			
— j — — w — — r —							
	l	h					
	m	n	ŋ				
v	ð	b	d	g	z	ʒ	dʒ
f	θ	p	t	k	s	ʃ	tʃ
ɪ	ʊ	ə	i	u	ə		

The full vowels

The four ‘families’ of vowels which are produced with palatal, velar, upper pharyngeal and lower pharyngeal tongue gestures are arranged from left to right across the page. The odd man out, /ɜ:/ occupies the rightmost position in the line of tense vowels. Each family contains a single ‘pure’, tense vowel on the row above the main dividing line, but sometimes one and sometimes two lax ones immediately above this. The most basic strategy, then, for helping a student who is struggling with any of these sounds is to encourage him to experiment with tongue gestures in the appropriate direction, and to attend to the auditory result.

The lax vowels are reliably distinguished in prototypical articulations (and hyperarticulations) by a difference in the opening of the jaw, and /ʊ/ and /ɒ/ are also distinguishable in this way⁷. The relationship is reflected on the chart by the more open jaw element of each pair being positioned slightly lower. The relation only holds within the particular pairs; no comparison, for example, between the jaw heights of /ɪ/ and /ʌ/ is being depicted.

This illustrates an important point about our priorities. The requirement is for a pedagogical tool not a precise and exhaustive representation of phonetic reality. There is often no conflict between these two possible objectives, but where there is we have been guided by our beliefs about what will assist and what might impede or distract our students from their desire to speak English. (We don’t always agree about this, of course.) It is not our aim to teach them English phonetics, and we will only ‘explain’ the layout of the charts to the extent that it assists their learning of the language. Mostly that means that we explain little or nothing; we just use the chart as a tool to help us to work on the awarenesses the students must have in order to improve their pronunciation.

In this context, the relative displacement of various lax vowels described above is probably right on the borderline of what is appropriate. It may help teachers and some students with an analytical turn of mind, but the real insight, of course, needs to be implemented by other techniques – in this case undoubtedly including the simple suggestion to students that they experiment with more or less open jaws when making these sounds if they find their pronunciation problematic.

For this reason there is also no formal indication of lip rounding on the charts. It could certainly be done; by, for example, changing the background colour around the /u:/ and /ɔ:/ families to grey to mark them as distinct from the others in this aspect of their articulation. But we have never found a problem with students’ failing to round their lips adequately for these phonemes of English, so it seems to us that the device is unnecessary and would just be a distraction.

The diphthongs are arranged based on their starting elements and whether they glide towards schwa, or towards schwi and schwu (names that are probably self evident, but are explained below). In the coloured - Silent Way - version of the chart the top half of the rectangle is filled with the same colour as the lax vowel positioned immediately beneath the diphthong. This is not a precise definition of the starting point any more than the first element of the standard IPA symbol for a diphthong is. However it is close enough for teaching purposes and actually, we think, has two distinct advantages over the use of standard symbols.

⁷ Try hyperarticulating pairs like “hid” and “head” with your forefinger on your nose and your thumb under your chin to illustrate this.

Firstly, in RP/SBE the starting elements of /aɪ/ and /aʊ/ seem closer to us to /ʌ/ and /æ/ than they do to each other. In practical teaching, at any rate, we have had good results with this approach. So filling them with /ʌ/ and /æ/ colours is helpful, and avoids both the need to add another colour to the chart and the confusion sometimes created by the standard use of IPA symbols here.

Secondly, it seems an advantage to start /əʊ/ with the colour of a strong vowel (/ɜ:/) rather than with that of the weak schwa. There can be enough problems working on low energy schwa with learners from certain linguistic backgrounds without suddenly implying by our notation that they need to produce a high energy version for the start of this diphthong. (Of course, neither /ə/ nor /ɜ:/ is ideal from either an auditory or articulatory point of view⁸).

The reduced vowels

Our second innovation is to place the reduced vowels at the bottom of the chart, in an area that is quite distinct from the full vowels. This has been very successful in practice. We distinguish the two areas on the basis of (articulatory) energy and students are clearly informed by the arrangement that something very different is going to be required for sounds in each area.

In the particular chart reproduced here, I have distinguished two groups of reduced vowels at the bottom of the chart. (As I mentioned before, the team at the CLA, working with a general public, have no need to work at this level of phonetic precision/pedantry.) On the right, within half-height rectangles, are the schwa which forms the final element of “butter”, and the /i/ and /u/ which are found word finally in “happy” and “to Aberdeen”, and pre-vocalically in “radiator” and “influence”. This /i/ is sometimes called the ‘happy’ vowel, but we call it ‘schwi’ and call the /u/ ‘schwu’. (Ashby (1997) is a helpful discussion of the distribution of schwi in English.)

It seems to me that the way these sounds are realised by native speakers justifies calling them vowels, but that they should be qualified as being ‘weak’ or ‘reduced’ since the articulatory energy with which they are produced is certainly less than that given to full vowels. I am happy therefore with the now conventional categorisation of these sounds.

However, the sounds which appear when /ə/, weak /ɪ/ and weak /ʊ/ either start words or appear between consonants seem to me to be better characterised differently, at least for pedagogical purposes. When teachers have produced texts along the following lines, “John c.n go t. th. party b.t Jane can’t”, they have been reflecting the essential nothingness of schwas in these contexts. Certainly there are syllables created in each case, but for production purposes, to help students whose first languages do not reduce vowels, we can suggest that they simply produce one consonant and then follow it by another. The result is usually very satisfactory and the dot filled with a schwa symbol on the chart is what I would use to separate the consonants when pointing out the sequences of phonemes involved.

⁸ This illustrates another advantage of working with colours rather than symbols, in that we are liberated somewhat from the weight of history. Wells (1990) mentions that he would have liked to change /əʊ/ to /ɜʊ/, presumably for similar reasons to ours, but that he rejected the idea to maintain conformity with earlier transcription standards.

To accommodate the cases where this transition has the flavour of /ɪ/ or /ʊ/ – in words like “basic” or “educate” - I have given the option of pointing to a ‘dot’ filled with an appropriate colour (or here, symbol). In practical terms, though, and outside university courses on English phonetics and pronunciation, any transition between the flanking consonants is probably acceptable, and the two extra dots on the left hand side would be unnecessary.

The source of this analysis is Catford (1985). He points out that the typical duration of a schwa, schwi or schwu clearly marks them as ‘real’ vowels. But he argues that the nucleus of the first syllable of “terrain”, for example, is very different in its duration and in a number of other ways. He labels it an ‘open transition’ between the two consonants involved, meaning that the first consonant is produced and that when this is complete we follow it with the second. This contrasts with a ‘closed transition’ when the articulations of the two consonants overlap. This is more conventionally called a consonant cluster, of course, with the first two sounds of “train” being a contrastive example.

Other pairs of words he gives to illustrate this include “polite/plight”, “collapse/claps” and “tussock/tusk” (and he is more precise and careful in his description of the phenomenon than I report here). He suggests that this analysis will have pedagogical value and my own teaching has borne this out.

As an aside, it may be that more phoneticians will reanalyse English this way in the future. Recent studies of interconsonantal schwa, both acoustic and articulatory, have suggested that it is ‘targetless’ (Bates 1995, van Bergem 1995). That is, that we cannot say either that a speaker is trying to produce a particular acoustic image in these cases, or that he is aiming at a particular articulation. Instead the trajectory of the tongue leads directly from the first consonant to the second, and the exact nature of the sound which appears between them as a result is of no particular significance.

If there was a ‘real’ vowel present we would expect a deviation in the trajectory towards the articulation required to produce it, and some sort of target area for the sound to fall within. But if interconsonantal schwa is indeed ‘targetless’ then it seems very respectable for us, as teachers, to adopt an ‘open transition’ phonetic analysis rather than a schwa/vowel one, quite apart from the pedagogical justification we may feel we have for doing so.

The consonants

Readers of *Speak Out!* will grasp the overall scheme of the arrangement of the consonants without difficulty. It is only one of many similar possibilities.

In a Silent Way chart these consonants are coloured rather than filled with symbols, and the left to right arrangement mainly reflects a convenient grouping of the sounds so that their position is easily distinguished and quickly retained by the students.

The vertical arrangement reflects the sonority of the elements. There is no pedagogical advantage in this *per se*, but it means that when pointing out words which include consonant clusters the pointer moves more consistently upwards or downwards than would otherwise be the case. (Try pointing out a word like “strand” to illustrate this. The pointer moves sideways or up for the onset of the syllable, and down or sideways from the vowel during the rhyme. Of

course, a word with a reduced vowel or open transition will require a gesture in the opposite direction, to the low energy part of the chart.)

There is no particular rationale for the position of /h/, beyond a desire that it should not appear to be in a relationship to any other sound.

Using the chart

A full description of all the ways that a phonemic chart, whether symbol or colour based, can support pronunciation work is the subject of a book not an article. (Underhill (1994) is such a book although it will be clear from what I have said here that I disagree with some of what he proposes.) To finish, though, I would like to make a few miscellaneous observations, prompted by recent conversations with teachers who have never used a phonemic chart and clearly have some misapprehensions about how they might do so.

A chart plays an important role throughout a course, and certainly not just in classes labelled as ‘pronunciation’. At the start it enables teachers to devise activities where sounds and sequences of sounds can be worked on in the absence of meaning. For many difficult areas in pronunciation (meant in a wide sense to include the whole sound system and melody of a language) meaning is a distraction, and more precise work can be done when meaning is temporarily dispensed with. But throughout a course pronunciation problems will arise, and a chart is a good tool for teacher and students to use to ‘discuss’ sounds and sequences of sounds in a precise way.

Although it will be necessary to spend a small amount of time on tense vowels and diphthongs spoken in isolation (since these form legitimate English words and interjections) the overwhelming proportion of work when using a chart will be with syllables and multisyllable phrases; i.e. with vowels embedded in words and utterances. Very long sentences can be produced by both teachers and students on a chart, and lead to profitable work on suprasegmental phenomena.

Incidentally, I feel it is a great mistake ever to ask students to produce lax vowels in isolation. This breaks a fundamental regularity of English phonology (in legitimate syllables they are always ‘checked’ by immediately subsequent consonants) and feels unnatural to native speakers. Fortunately, having a chart to hand means that a teacher need never say /ɪ/ or /æ/ out loud. To refer to them, he need only point.

Finally, teachers who work on these things will notice that the position of /j w r/ supports an understanding of their patterning when they take a linking role between vowels. Pointing out “A bee, a shoe, a four, a car and the purr of a cat” on the chart will illustrate this.

Conclusion

This article has covered quite a lot of ground. I would summarise the main points I have made as follows.

1. Students need to emulate the vowels in the language they are learning, not to mimic them. If our teaching techniques encourage or demand mimicry, then they will distract students’ attention from where it needs to be for the students to acquire a new sound.

2. Learning a new sound involves a process of interaction between the learner's vocal apparatus and the learner's ear, and one which requires the learner's presence (or attention) to what is going on in both of these systems. One technique that can encourage this is for a teacher to refuse (generally) to give models.
3. Articulatory guidance from the teacher may help a student. The IPA quadrilateral, cannot be the basis for this. Instead, we should adopt a model that is based in articulatory reality, such as the Throat/Tongue/Lip (or Tongue/Jaw/Lip) model. The elements of this support a 'distinctive features' approach that differentiates English vowels on the basis of the direction of the basic tongue gesture, its extent (tenseness/laxness), jaw opening and lip rounding.
4. A 'pedagogical phonetics' incorporating this could also include clearly separating out reduced from full vowels, and reanalysing many instances of reduced vowels as open transitions.

References

- Ashby M (1997), Weak vowel /i/ in the fifth edition of the OALDCE, *Journal of English Phonetics Society of Japan* 1: 51-59
- Bates SAR (1995), Acoustics characteristics of schwa: a comparative study, in *ICPhS 95, Volume 3*, Stockholm p 230-233
- Butcher A (1982), Cardinal vowels and other problems, in *Linguistic Controversies*, ed. Crystal D, Arnold, London p 50-72
- Byrne RW, Russon AE (1998), Learning by imitation: a hierarchical approach, *Behavioral and brain sciences* 21: 667-721
- Catford J (1987), Phonetics and the teaching of pronunciation: a systematic description of English phonology, in *Current perspectives on pronunciation*, ed. Morley J, TESOL, Washington DC
- Catford J (2001), *A practical introduction to phonetics*, OUP,
- Catford J, Pisoni D (1970), Auditory vs. articulatory training in exotic sounds, *Modern Language Journal* 54: 477-481
- Catford JC (1985), 'Rest' and 'open transition' in a systematic phonology of English, in *Systemic Perspectives on Discourse Vol 1*, ed. Greaves W and Benson J, Ablex, Normal NJ
- Gattegno C (1973), *The Universe of Babies*, Educational Solutions, New York
- Gimson AC (1989), *An Introduction to the Pronunciation of English*, Edward Arnold, London
- Ginésy M (1995), *Mémento de phonétique anglaise*, Éditions Nathan, Paris
- Heyes C (2001), Causes and consequences of imitation, *Trends in Cognitive Sciences* 5: 253-261
- Kenworthy J (1987), *Teaching English Pronunciation*, Longman, London
- Lindblad P (2001), Critique of the IPA vowel quadrilateral, especially the use of point vowels in it, *Working Papers, Department of Linguistics, Lund University* 49: 104-107
- Roach P (1991), *English Phonetics and Phonology*, CUP, Cambridge
- Russell OG (1928), Tongue-arching vowel triangle - a fallacy, in *The Vowel*, Ohio State U Press
- Underhill A (1994), *Sound Foundations*, Heinemann ELT, Oxford
- van Bergem D (1995), *Acoustic and Lexical Vowel Reduction*, IFOTT (Institute for Functional Research into Language and Language Use), Amsterdam
- Wells JC (1990), *Longman Pronunciation Dictionary*, Longman, Harlow

- Whalen DH, Min Kang A, Magen HS, Fulbright RK, Gore JC (1999), Predicting midsagittal pharynx shape from tongue position during vowel production, *Journal of Speech, Language and Hearing Research* 42: 592-603
- Wood S (1987), The history of the classical vowel articulation model - a reply to Catford and Fischer-Jørgensen, in *Proceedings of the XIth ICPHS*, Academy of Sciences of the Estonian SSR
- Wood SAJ (1979), A radiographic analysis of constriction locations for vowels, *Journal of Phonetics* 7: 25-43
- Wood SAJ (1993), The throat-tongue-lip model of vowel articulation, *Phonum (Umea)* 2: 139-149
- Young R (1995), Caleb Gattegno's "Silent Way": some of the reasons why, *Methoden der Fremdsprachenvermittlung, University of Mainz* 55-74
- Young R (2000), Round Table Discussion, *Prism, Paris*

Piers Messum <p.messum@ucl.ac.uk> has taught in Japan, France and the UK. At present he is trying to finish a Ph.D on the significance of speech breathing within English phonetics, while working part-time at Laryngograph Ltd, a voice instrumentation company.