Teaching Students to Pronounce English: A Motor Skill Approach in the Classroom

Piers Messum (iD) and Roslyn Young

Pronunciation Science Ltd, UK

Abstract

Just as forming letters is the motor skill component of writing, pronouncing—forming sounds and other features of the spoken language—is the motor skill component of pronunciation. The motor task for L2 (second language) learners is to invent for themselves the actions needed for pronouncing L2, either from matching a model they hear (goal emulation) or in some other way, and then to automatise their use of these actions. Many students—and their teachers—are dissatisfied with the results they achieve in learning to pronounce through current teaching practices. As an alternative, we argue for taking an Articulatory Approach, based on a motor skill coaching paradigm. However, learning to pronounce an L2 differs from the learning of other socially transmitted motor skills because the learners cannot see most of the significant actions that produce the results which they and others hear. This means that the coaching paradigm for pronunciation cannot be the same as that used for those motor skills where learners can watch a performance. We explain how an appropriate paradigm is implemented in the teaching techniques we describe.

Determining the Task in Learning to Pronounce L2

Conventional Approaches to Teaching Pronunciation

Celce-Murcia et al. (2010: 2), following Kelly (1969: 61), state that pronunciation is taught by either Imitative-Intuitive or Analytic-Linguistic approaches. In the former, the archetypal exercise is 'listen and repeat': it is assumed that given an auditory target,

Corresponding author:
Piers Messum, Pronunciation Science Ltd, 112 Warner Road, London SE5 9HQ, UK.
Email: p.messum@gmail.com

tudents will readily be able to invent the articulatory gestures whose audible results will match this token. In an Analytic-Linguistic approach, a similar learning paradigm is relied upon but with additional tools employed to 'reinforce' (Kelly, 1969: 61) the supposed natural ability of students to imitate spoken models. Historically, this reinforcement involved giving students phonetic information, but recently the focus has been on improving the recognition and discrimination of unfamiliar L2 sounds that students' L1-habituated perception obscures.

'Actions' and 'Results' in Motor Learning

In pronunciation teaching, the process of matching an acoustic target is generally referred to as 'imitation'. However, the word 'imitation' in the professional literature is used to describe a process in which the observer copies both the actions employed and the results obtained by the modeller. 'Emulation' (or 'goal emulation') is used when the observer copies the results, but invents their own means, which may or may not be those used by the modeller. For completeness, 'mimicry' is used when the observer copies the actions with no concern for the results achieved by the modeller (Messum, 2007).

Most motor skills involve moving an object, some with the actions and the results visible (e.g. kicking the ball in soccer), and some with actions visible and results audible (e.g. playing musical instruments). A smaller proportion involve creating bodily postures (e.g. ballet movements, sign language gestures), where the results are the appearance of the actions themselves, and are visible, though sometimes not directly to the performer (e.g. springboard diving). Speech also involves creating bodily postures, but the actions used are largely invisible to both the speakers and their interlocutors, and the audible results are attended to instead. Matching these is, therefore, a process of emulation rather than imitation.

Furthermore, for L2 (second language) pronunciation, it is well known that learners cannot make straightforward use of these audible results (Best and Tyler, 2007; Flege, 1995). Their expert and automatic sound-categorising behaviours for their L1 obscure, for example, differences that are insignificant in L1 but significant in L2. It is less well appreciated that there are motor skill aspects of pronouncing a language that only the most gifted L2 learners are likely to invent by themselves. In English, these include the basis of articulation of the tongue, part of the specific articulatory setting of the language (Borissoff, 2012), and the open transition as a mechanism for producing schwa (see below for elaboration and teaching ideas). Indeed, in our experience even the actions that create less fundamental features of English are beyond the range of personal invention for most students if given little more than an acoustic target to work towards.

Learning to pronounce L2 is, therefore, a highly unusual activity, at least some of whose motor skill aspects must be taught. The two conventional 'listen first' approaches we described earlier share acoustic matching-to-target as their principal mechanism for learning. In the next section, we explain how such a mechanism is unsuited to learning to pronounce three underlying aspects of English, and how these can be better taught. Then we explain how learning to pronounce the new sounds of an L2 is unusual among other socially transmitted motor skills. This requires teachers to take a distinctive

approach to coaching this activity even within a motor learning paradigm. We describe how this can be done successfully with reference to Caleb Gattegno's Silent Way and our development of his techniques.

Teaching Three Underlying Systems of English Pronunciation

The Starting Point

Learners must be sensitive to, and attentive to, the actions of their articulators if they are to both inhibit the automatic gestures of L1 and exert careful, conscious control to develop new gestures for pronouncing L2. Developing sensitivity and attentiveness is, thus, a prerequisite for any work on pronunciation taught as a motor skill. Most students come to a language class having lost contact with the way that they produce speech. Re-sensitisation can start with just a few minutes' work but will deepen only if the teacher makes it a repeated focus of students' attention during early lessons. As an example, see Young and Messum (2015: 15) for a description of initial work with the tongue and inside of the mouth; such work can be seen in a video with Young (2015).

Before starting work to develop new sounds, it is particularly advantageous for learners of English to work on the underlying systems of its pronunciation so that sounds and systems can then be developed and practised together (Messum, 2015). We now describe work on three of these systems.

Stress as a Motor Skill

If one teaches using a 'listen first' approach, it is entirely natural to take the pronunciation of a language to be what one hears. With this orientation, some textbook writers simply describe English sentence stress as greater loudness, length, and vowel clarity in the stressed syllable, as these are the distinguishing acoustic qualities of stress. The authors set the students the task of directly copying these qualities in their own speech.

But pronunciation is always simultaneously something a speaker does as well as something that others hear. Examining stress as a production phenomenon, one finds an impressive consensus among the previous generations of phoneticians, for whom these matters were a primary concern. English stress is greater effort being applied by the respiratory system to the syllable in question (Catford, 1977; for further citations see Messum, 2009). Within an Articulatory Approach, students are taught a gesture for stress directly, as a pulse of compressive effort, most conveniently made by most learners as a contraction of their abdominal muscles (Messum, 2017a). The acoustic qualities of sentence stress come as an automatic result of this action.

One of the ways we teach this is to use a cycle in which students say a sentence using stage whisper (i.e. loud whisper), normal (soft) whisper, and finally normal voice; and then repeat the cycle (Messum, 2017b). Stage whisper (a 'whisper' that has to be audible throughout the theatre) requires considerable effort, most apparent in the compressive actions of the abdominal muscles on stressed syllables. The students' new sensitivity to the actions that underlie stress is then sharpened when these actions are

less apparent, as in normal whisper, and they can then be intentionally deployed in normal voice as the stress-producing mechanism. This development of a new use of physical capacities is exactly what any student does when learning a new sport.

Reduction as a Motor Skill

Similarly, schwa has to be considered as a vowel sound within 'listen first' approaches, and it then becomes notoriously difficult to teach. It is often exaggerated to make it easier for students to detect it, and this destroys its reduced nature. In contrast, viewing schwa from a production perspective makes it straightforward to teach. As Catford (1985, 2001: 111–116) pointed out, many noises analysed as schwa in perception can be described as 'open transitions' between flanking consonants in production. Compare *train* with *terrain: train* has overlapping [t] and [r] articulations while *terrain* has an open transition between the two consonants, and a reduced syllable is heard as a result.

An open transition between homorganic consonants is akin to a stutter—*p-p-park*—and can easily be taught in this way. Other open transitions are readily learnt by analogy, and students then have a successful production strategy for producing this fundamental feature of English pronunciation. We often introduce this notion using the sentence *It's a quarter to two*, building it up in a back-chaining process: /tu:/, /t.t.tu:/, /kwɔːt.t.tu:/, etc. (where a dot represents a stutter with no vowel actually articulated). It is remarkable both how natural the final result always is, and how shocked even advanced students are when they discover how native speakers actually articulate this simple sentence. We use the lyrics of *Tea for two* to show how 'stuttering' between heterorganic consonants—/ti:f.tu:/, etc.—is also used in English (Messum and Young, 2019).

The English Articulatory Setting

Neither massive exposure nor intensive listening work seem to lead to the development of a third underlying characteristic of English, its articulatory setting (Messum and Young, 2017a; Gick et al., 2017; Honikman, 1964). If L2 learners do not adopt this, they will never completely master its pronunciation (Honikman, 1964: 74). When taught within an Articulatory Approach, students find the idea of a new basic setting for their tongue to be entirely natural, even if adopting the English articulatory setting is a significant practical challenge for many learners.

To introduce the topic, we play our students two magnetic resonance imaging (MRI) videos which show characteristic tongue motions in Chinese and English. These easily convince them of the reality of articulatory settings and of the need to adjust their tongue setting when speaking English. We then show them models of the tongue setting for French (which is similar to that of many other languages) and English (Figure 1). The steps we take are described in Messum and Young (2017a).

For further discussion of these three underlying systems and teaching them, see Messum (2015) and Messum and Young (2019).
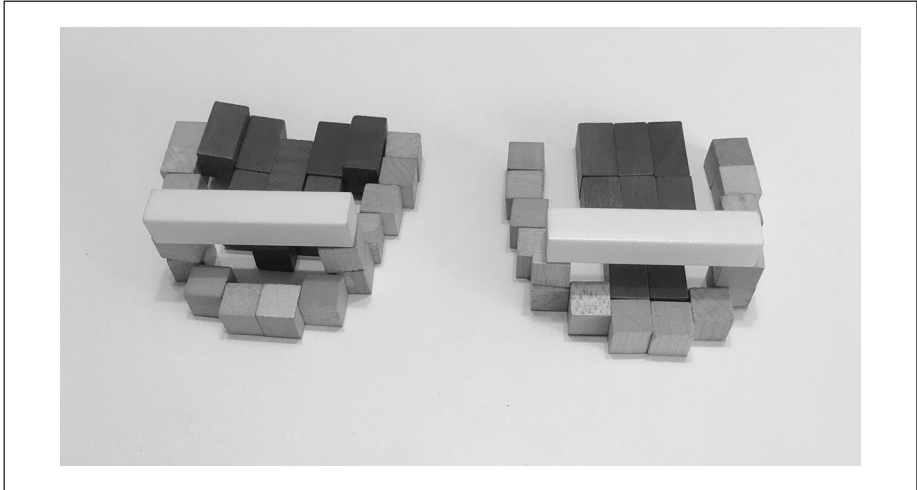
Figure 1. Cuisenaire rod model of the basis of articulation for English (left) and French (right). The pale rod bridging the front of the mouth represents the alveolar ridge. During speech, the tongue does not 'float' freely in the mouth. For French (and many other languages), the tip of the tongue is predominantly kept in contact with the lower front teeth. For English, speakers brace their tongue against the upper molars. Because the tongue is a muscular hydrostat, this widening at the back of the mouth causes a retraction of the tip.

Teaching New L2 Sounds

Learning Motor Skills

Through personal experience, everyone is familiar with how motor skills like playing a sport are learnt. The action–perception cycle is at the heart of the process; it is the principal mechanism by which learners develop sensorimotor contingencies between their actions and the results these produce. Additionally, there will usually be some external criteria of success against which the results must be evaluated. Learning to pronounce an L2 requires both the development of sensorimotor contingencies and evaluation, but the unusual nature of pronunciation means that some aspects of its teaching must differ from the norm for the coaching of other socially transmitted motor skills.

It is instructive to compare learning to produce a new L2 sound with learning to form the letters or characters of an unfamiliar script (Persian, or Amharic from Ethiopia, perhaps). For writing, developing gestures for each letter can be achieved through a copying process. This involves cycles during which the learners undertake two principal activities: (1) they look at a model, create a mental image of it, and try to recreate this on paper; and (2) they evaluate their attempt against the model. These two distinct activities can be performed repeatedly because both the model and the attempt are marks on a page: they are stable and persist in time. Mental images can be refreshed whenever necessary and don't need to be maintained in working memory. The learner's attentional resources can be switched between each activity without any conflict. In practice, the learner's eyes move repeatedly between the two images.

As previously discussed, learning L2 sounds is a process of emulation rather than imitation, because the vocal actions involved are largely invisible. In this respect, it is similar to learning to write unfamiliar L2 letters or characters from examples presented on the page, without having seen them being written. However, unlike marks on a page, sounds are ephemeral and perceptually opaque (they may be experienced differently when produced by oneself or by others).

The ephemerality of sounds creates a conflict for attentional resources within any learning-by-copying process. There are no stable, persistent images between which attention can be switched. Learners face a dilemma. Should they use working memory to maintain the fading image of the aural model for later evaluation of their own output against it? Or should they attend to the sensorimotor consequences of their attempts at novel actions, as they undertake the difficult task of re-asserting conscious control over a largely unconscious speech production system that is in the 'grip' of L1? If they do the former, the development of novel actions will be impaired since the prioritisation of storage in working memory comes at a cost in the performance of a separate processing task (Rhodes et al., 2019). If they do the latter, they will not be able to retain the image of the original model in working memory (Couffe and Michael, 2017: 168), leaving them unable to evaluate their performance.

In our experience, few L2 learners have the skill and experience to manage this conflict well, particularly if the teacher is not aware of the problem and believes that copying is a straightforward, 'natural' way to learn new sounds. The way in which a sound copying exercise is usually presented, with the expectation of an immediate response from the students, induces them to produce a token in response to the teacher's token that is usually the result of what they can already do—drawn from their inventory of routine, already learned vocal actions—rather than something willed and genuinely novel (Norman and Shallice, 1986). This makes the exercise more akin to a test than to an opportunity to learn.

Ensuring Motor Skill Development in Pronouncing New Sounds

The development of the motor schema which will produce an L2 sound that is new to the learner requires both the learning of sensorimotor contingencies involved in the production of the new sound and reliable evaluation and feedback about what is being produced.

As Gattegno demonstrated, one way to ensure that these two conditions of learning are reliably met is for the teacher to refrain from providing a model at all, either in person or recorded (Gattegno, 1977: 19–21, 1985: 49–53). A demand to produce a sound without a model to copy (but identified in some other way and cued, perhaps, by a facial gesture) throws the students into the action–perception cycle. They attend to their articulators, to the various sounds they are producing, and to the differences in these sounds that are created by changes they make in the gestures and positions of their articulators. The responsibility for evaluation lies with the teacher, who has well-developed perceptual criteria for this, and whose feedback provokes further cycles of sensorimotor experimentation on the part of the students (Figure 2).

As typically seen in motor skill learning, production educates perception in parallel processes of development. Thus, the learners' action–perception cycles evolve: the responsibility for evaluation can gradually move from teacher to student.
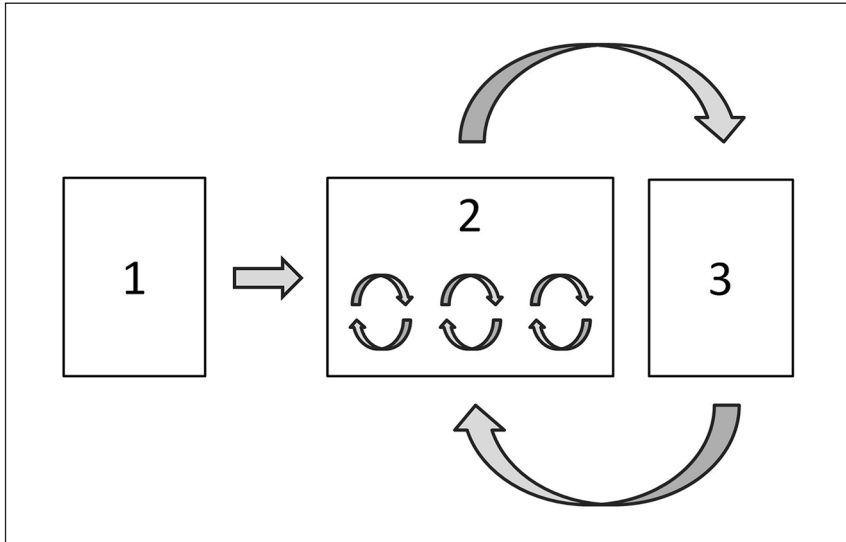
Figure 2. Ensuring motor skill development in pronouncing new sounds. (1) Teacher initiates work on a sound without giving a model. (2) Students explore sensorimotor contingencies in a series of action/perception cycles. (3) Teacher evaluates a trial and gives feedback and encouragement.

Teachers should be 'silent' but need not be mute. They should not model what the students are working on because this will draw many students into attempts to copy them, but they can speak: to coach the students in what they should be doing with themselves physically, and to encourage them to continue to explore. Within this learning paradigm, the teacher's functions of evaluation and feedback are the same as those of a coach of any perceptually opaque skill, such as springboard diving.

In the Classroom

The authors have more than 50 years of collective experience both of applying Gattegno's paradigm in class and of being students in classes taught this way. We have found that his approach is effective, efficient, and readily accepted by students. The belief that students need an aural model to work from is incorrect.

If the teacher doesn't model the sounds of the L2, some other way to refer to them will be needed. In 1978, Gattegno's solution for the Silent Way was to place the inventory of sounds, represented by coloured rectangles, on a chart to be hung on the classroom wall, with teachers and students using telescopic pointers to identify what was being worked on. Such a chart will be largely phonemic, but will depart from strict phonemic principles when pedagogy demands.

An early task for students in pronouncing an L2 is to develop distinct motor gestures for the whole inventory of sounds, and a chart is a useful tool to support this. However, its real value is in work on producing strings of sounds—chunks of language—that, for
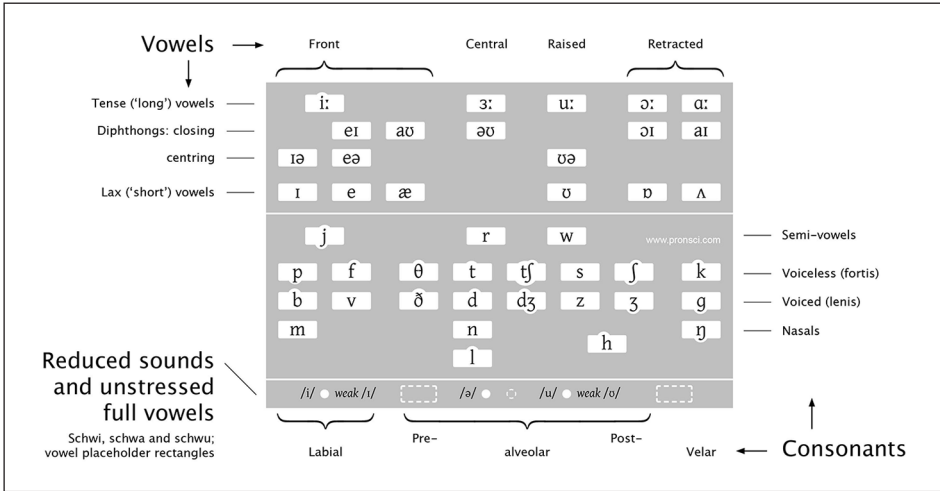
Figure 3. British English phonemic chart laid out on an articulatory basis. The arrangement of the full vowel section (top) follows Esling (2005). The unstressed vowel section (bottom) includes placeholders for full-but-unstressed vowels. See Messum and Young (2017b) for a full explanation of the layout.

English, will require the use of stress and reduction. On the charts that we ourselves have developed (Messum, 2018a), we have, therefore, also embedded the four levels of non-accentual syllabic prominence in English (Figure 3). The chart enables full stressed vowels, full-but-unstressed vowels, and two reduced vowel forms—schwa-type sounds and open transitions—to be distinguished when appropriate. Thus, the segmental and suprasegmental aspects of pronunciation can be worked on simultaneously, as they should be.

A wall chart and pointing techniques that allow whole-class work on pronunciation have many other pedagogical advantages (Messum, 2018b; for good practice, see Young, 2018). To note just two:

- The recurring use of a tool dedicated to pronunciation that is always visible in the classroom helps give the spoken language a prominent place within the language skills being learned.
- Using a chart and a pointer, the teacher can quickly and precisely direct the class's attention—for example, to areas presenting difficulties or to the changes in stress and reduction which determine different levels of formality.

In this work, the role of the teacher is to propose a step-by-step path through L2 sounds and sound combinations with the exact progression being developed as students' difficulties become apparent. The type of work required is described by Catford (2001): sensitisation, trial and error, and continual feedback from the teacher.

Taken as motor skills, new sounds and combinations of sounds first need a short period of intensive work for learners to develop a feel for how they can be produced.

Then they need practice throughout the course during other language work. The gestures needed can be fine-tuned in micro lessons. With motor skills that are developed in this way, students can give themselves worthwhile and effective practice outside the classroom as soon as they have developed some motor criteria for correctness.

Conclusion

If students are to learn to pronounce an L2, some classroom time must be spent on the motor skill aspect of pronunciation. Most current pronunciation teaching ignores this need or does not address it properly. As Catford (2001: 2) pointed out, articulatory instruction as it is generally implemented is not the way to coach pronunciation as a motor skill.

We have argued that to teach students how to pronounce, teachers need to find out what actions native speakers use to pronounce their language. Then, through a combination of teaching and coaching, they lead their students into experimental work to reinvent these gestures for themselves in action–perception learning cycles. In these, the teachers will be a source of evaluation and feedback. They should avoid modelling the results the students are expected to produce. Doing so will immediately induce attempts to match-to-target, drawing the students' attention away from their articulators and, therefore, from the action aspect of the action–perception cycle, which is the basis for motor learning, and through which perception is educated in parallel with production.

ORCID iD

Piers Messum https://orcid.org/0000-0001-5450-3189

References

Best CT, Tyler MD (2007) Nonnative and second-language speech perception: Commonalities and complementarities. In: Munro MJ, Bohn O-S (eds) *Second Language Speech Learning: The Role of Language Experience in Speech Perception and Production*. Amsterdam: John Benjamins, 13–34.

Borissoff C (2012) Basis of articulation or articulatory setting? *Speak Out! (Whitstable, IATEFL)* 46: 9–13.

Catford JC (1977) *Fundamental Problems in Phonetics*. Edinburgh: Edinburgh University Press.

Catford JC (1985) 'Rest' and 'open transition' in a systemic phonology of English. In: Greaves WS, Benson JD (eds) *Systemic Perspectives on Discourse Vol 1*. Normal, NJ: Ablex, 333–349.

Catford JC (2001) *A Practical Introduction to Phonetics*. 2nd ed. Oxford: OUP.

Celce-Murcia M, Brinton D, and Goodwin JM (2010) *Teaching Pronunciation: A Course Book and Reference Guide*. 2nd ed. New York: Cambridge University Press.

Couffe C, Michael GA (2017) Failures due to interruptions or distractions: A review and a new framework. *The American Journal of Psychology* 130(2): 163–181.

Esling JH (2005) There are no back vowels: The Laryngeal Articulator model. *The Canadian Journal of Linguistics / La revue canadienne de linguistique* 50(1): 13–44.

Flege JE (1995) Second language speech learning: Theory, findings, and problems. In: Strange W (ed) *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Timonium, MD: York Press, 233–277.

Gattegno C (1977) Further insights into learning languages: A brief history of the Silent Way. *Educational Solutions Newsletter* 7(2): 19–21.

Gattegno C (1985) *The Science of Education. Chapter 13: The Learning and Teaching of Foreign Languages*. New York: Educational Solutions.

Gick B, Allen B, Roewer-Després F, et al. (2017) Speaking tongues are actively braced. *Journal of Speech Language and Hearing Research* 60(3): 494.

Honikman B (1964) Articulatory Settings. In: Abercrombie D (ed) *In Honour of Daniel Jones*. London: Longman, 73–84.

Kelly LG (1969) *25 Centuries of Language Teaching*. Rowley, MA: Newbury House.

Messum PR (2007) *The role of imitation in learning to pronounce*. PhDThesis. University College London.

Messum PR (2009) Grounding stress in expiratory activity. *Speak Out! (Whitstable, IATEFL)* 41: 12–16.

Messum PR (2015) What to teach before you teach sounds. British Council. Available at: https://www.youtube.com/watch?v=D5DteiSDTs4 (accessed 1 July 2020).

Messum PR (2017a) Sentence stress: Learning how. Teaching Pronunciation Differently. Available at: https://youtu.be/EufYzScL2I0 (accessed 7 September 2020).

Messum PR (2017b) Sentence stress: The right actions. Teaching Pronunciation Differently. Available at: https://youtu.be/eKZsgCts3CQ (accessed 7 September 2020).

Messum PR (2018a) The PronSci phonemic charts: overall design. Available at: https://www.youtube.com/watch?v=iRCeA8bJTlg (accessed 7 September 2020).

Messum PR (2018b) Why we should use a chart and a pointer for teaching pronunciation. *Speak Out! (Whitstable, IATEFL)* 58: 53–60.

Messum PR, Young R (2017a) Bringing the English articulatory setting into the classroom: (1) the tongue. *Speak Out! (Whitstable, IATEFL)* 57: 29–39.

Messum PR, Young R (2017b) The PronSci rectangle charts: phonetic guide. Available at: https://www.pronunciationscience.com/support/ (accessed 7 September 2020).

Messum PR, Young R (2019) Teaching the underlying systems of English pronunciation. *Speak Out! (Whitstable, IATEFL)* 61: 18–31

Norman DA, Shallice T (1986) Attention to action. In: Davidson RJ, Schwartz GE, and Shapiro D (eds) *Consciousness and Self-Regulation: Advances in Research and Theory Volume 4*. Boston, MA: Springer US, 1–18.

Rhodes S, Jaroslawska AJ, Doherty JM, et al. (2019) Storage and processing in working memory: Assessing dual-task performance and task prioritization across the adult lifespan. *Journal of Experimental Psychology: General* 148(7): 1204–1227.

Young R (2015) Working on Pronunciation in English Part 1: Increasing students' sensitivity to their mouth. Besançon, France. Available at: https://www.youtube.com/watch?v=J15-i8uK-tDE (accessed 7 September 2020).

Young R (2018) How to use a chart and a pointer for teaching pronunciation. *Speak Out! (Whitstable, IATEFL)* 59: 20–26.

Young R, Messum PR (2015) Teaching English pronunciation in an intensive Silent Way course. Unpublished manuscript. Available at: https://www.pronunciationscience.com/downloads/.