Richards, M., & Cotos, E. (2019). Teaching Tip: Segmental accuracy: A recommended training sequence for moving learners from accurate perception to accurate and automatic production in the stream of speech. In J. Levis, C. Nagle, & E. Todey (Eds.), *Proceedings of the 10th Pronunciation in Second Language Learning and Teaching Conference*, ISSN 2380-9566, Ames, IA, September 2018 (pp. 413-422). Ames, IA: Iowa State University.

TEACHING TIP

SEGMENTAL ACCURACY: A RECOMMENDED TRAINING SEQUENCE FOR MOVING LEARNERS FROM ACCURATE PERCEPTION TO ACCURATE AND AUTOMATIC PRODUCTION IN THE STREAM OF SPEECH

Monica Richards, Iowa State University
Elena Cotos, Iowa State University

A key reason L2 learners struggle to pronounce new segmentals is because their L1 has trained them to hear L2 phonemes as allophonic. When learners cannot accurately *hear* a word's phonemic structure, they are able to self-assess their L2 pronunciation only by comparing their conscious knowledge of how the word should be pronounced with the physiological "feel" of their vocal organs. But is it possible to perform this task consciously on a regular basis? After all, L2 speakers must simultaneously engage in several additional, higher-level cognitive processes also harder in L2 than L1, for example, comprehending what others are saying, drawing connections between what is said to what is already known, planning what to say next and figuring out how best to say it. L2 learners must therefore develop the ability to self-assess *sub*consciously whether the phones they pronounce are categorized by the L2 as the phonemes they intend. Not only that, but their physical production of accurate L2 phoneme distinctions must become *habitual*. This paper therefore introduces a recommended training sequence for moving learners from accurate perception to accurate and automatic production of challenging L2 segmentals in the stream of speech.

INTRODUCTION

L2 segmental training frequently focuses on the mouth. However, pronunciation is an automatic skill including receptive as well as productive components. Just as the mind originally trained the mouth to produce reliably accurate segmental pronunciations in the L1, the mind is vitally important for re-training the mouth to produce reliably accurate segmental pronunciations in an L2. We therefore overview three key realities underpinning maximally effective segmental pronunciation pedagogy: *1) perception and production are linked*, *2) segmental pronunciation is a physical activity* and *3) segmental pronunciation is a habit*. On this foundation, we present a recommended training sequence grounded in these three key realities in order to maximize L2 learners' success in developing and automatizing new, more accurate pronunciation of segmentals.

UNDERSTANDING THE PERCEPTION/PRODUCTION LINK

The ear and mind can potentially learn to distinguish hundreds of sounds (MRT [magnetic resonance tomography] YouTube video). However, babies acquiring their L1 are exposed to the smaller subset of sounds found in that L1, and their minds begin developing sound categories accordingly (usually less than a hundred – Moran, McCloy, & Wright, 2014). Each phoneme, or mental label for a sound category, may include a range of phones (i.e., distinct physical pronunciations). For example, the English /l/ phoneme is not only produced as light/clear [l] but

also syllable-finally as dark [ɬ]). In part, this is because a sound's environment, its surrounding sounds, may facilitate slightly varying realizations. Additionally, people's vocal tracts differ slightly in size, shape, etc., also producing minute variation in a given phoneme's pronunciation. The mind learns to disregard this "noise" in the input to a large extent. Nevertheless, unfortunately for L2 learners, sounds that pertain to the same phoneme label are sometimes not phonetically close at all. For example, in General American English, the phoneme /t/ has 5 distinct allophones (phones categorized as belonging to the same phoneme even though they are in fact articulated differently). Some allophones of General American /t/ are quite different from the prototypical voiceless alveolar stop we commonly think of as /t/, as demonstrated in Figure 1.

In spite of how some L1-defined allophones are quite different from each other phonetically, L1 child learners nevertheless develop mental sound categories enabling their minds to mirror for any allophone they hear their L1's phoneme categorization as efficiently as possible. This "tuning" of the mind to notice only L1-relevant aspects of any spoken language input maximizes the processing efficiency of L1 speech perception, but it complicates learning to categorize L1 allophones as phonemic in the L2, i.e., as phonetically different sounds belonging to different L2 phonemes. As a result, learners frequently hear L2 phonemes as allophonic (e.g., L1 Japanese speakers characteristically struggle to hear the difference between the English /ɹ/ and /l/ phonemes and L1 Korean speakers the difference between English /p/ and /b/ because their respective languages categorize these sound pairs as allophones) (Best & Tyler, 2007; Broersma & Cutler, 2008; Flege, 1995; Jenkins, 2000; Qian, Chukharev-Hudalainen, & Levis, 2018; Richards, 2012).

General American allophones of /t/	Environment	Example
aspirated [t^h]	the beginning of words and stressed syllables	"time", "return"
flap [ɾ]	usually between a vowel (± /ɹ/) and an unstressed vowel or syllabic [ɬ], [m̩], or [ɻ] – also, between any vowel (± /ɹ/) and a word-initial vowel (Vaux, 2000)	"water", "liberty", "bottom", "turtle", "thought it over", "report it immediately"
nasalized flap [ɾ̃]	substituting for "nt"	"winter", "front of. . ." *(i.e., the same phone used to pronounce /n/ in "winner" and "tunnel")*
glottal stop [ʔ]	between a vowel (± /l-ɹ/) and syllabic [n̩]	"button", "bought an. . . .", "carton", "Hilton" *(cf. the medial glottal stop in "Uh-oh")*
unaspirated [t]	everywhere else	"wait", "doctor", "multiply", "distance", "stop"

Figure 1. The five General American allophones of /t/.

It is important teachers realize that when learners fail to recognize phonetically different L2 allophones as belonging to the same phoneme, it is not only their L2 *listening*, but also their L2 *pronunciation* that is affected. After all, how much impact can a teacher's occasional pronunciation correction be expected to make if learners' continued misclassification of L2 phonemes as allophones prevents them from recognizing the *vast majority* of their mispronunciations of a particular problem segmental? Indeed, when learners cannot hear a word's phonemic structure accurately, they can self-assess their L2 pronunciation only by comparing their conscious knowledge of how the word *should* be pronounced with the physiological "feel" of their vocal organs. The impracticability of this for the stream of speech should be evident in light of how the mind must simultaneously accomplish several higher-level cognitive processes also much harder in the L2 than L1, namely:

1) Comprehending what others are saying;
2) Identifying connections between what others are saying and what one already knows;
3) Figuring out what ideas one wants to say next and how to express them (in terms of information structure, politeness, grammar, etc.).

It is therefore vitally important that teachers prioritize developing learners' ability to passively hear whether or not the L2 categorizes their various phone pronunciations as distinct phonemes.

One final factor impacting L2 sound perception and L2 segmental production is orthography. Many L2 learners have stronger reading than listening skills, resulting in their L2 sound perception being impacted by spelling. That is, the pronunciation L2 listeners sometimes perceive is that which the spelling of a word leads them to expect (Detey & Nespoulous, 2008; Erdener & Burnham, 2005). Understandably, they may also speak this spelling-derived pronunciation (Young-Scholten & Archibald, 2000). It is therefore vital teachers 1) recognize which words learners misperceive and mispronounce due to orthographic ambiguity and 2) facilitate learners' perceptive and productive acquisition of these words' standard spoken form.

Nevertheless, maximally effective segmental pronunciation pedagogy is grounded not only in teachers' understanding of the perception/production link, but also in their recognition that segmental pronunciation *is* a physical activity.

UNDERSTANDING SEGMENTAL PRONUNCIATION AS A PHYSICAL ACTIVITY

Many L2 speakers exhibit near-native proficiency in L2 skills other than pronunciation (Scovel, 1969). However, advanced L2 speakers frequently struggle to acquire standard L2 pronunciation in part because segmental pronunciation has an unmistakably physical component.

That is, even if L2 speakers successfully learn to hear new L2 phonemes as phonemes (versus allophones), they often experience great difficulty learning *how* to move their vocal organs differently than how they have always moved them before. In part, this is because L2 learners are accustomed to regularly adjusting their understanding of new vocabulary and even grammar, but many have had very little pressure to adjust their pronunciation since first learning to speak their L1 as a child. Teachers' efforts to develop students' conscious understanding – or metaphonological awareness – of *how* a target segmental is physically articulated can therefore be

helpful (Ingels, 2010). Nevertheless, teachers must not only provide L2 learners metaphonological information about *how* a sound is physically articulated; they must also help learners automate newly learned L2 segmental articulations via the development of new pronunciation *habits*.

UNDERSTANDING SEGMENTAL PRONUNCIATION AUTOMATICITY

The process of learning an L1 involves not only "tuning" the mind to hear some phonetic distinctions as phonemic and others as not, but also tuning the vocal organs to default to L1 configurations. As a result, students beginning to hear the difference between members of an L2 minimal pair and to be able to articulate new L2 phoneme(s) nevertheless often fail at pronouncing "learned" L2 phonemes accurately except in the rare instances when they consciously attend to pronunciation. As explained earlier, this is because pronunciation more than other language skills is necessarily nearly always produced automatically (i.e., subconsciously) since the mind does not have the processing capacity necessary to support the multiple higher-level cognitive processes required for fluent speech production as well as to simultaneously also control pronunciation consciously.

L1-trained pronunciations are so deeply entrenched that successfully replacing them in the stream of speech with more accurate L2 pronunciations can be the hardest step of all in L2 pronunciation acquisition. The almost magnetic nature of L1-based articulatory habits makes them very hard for L2 learners to break. Additionally, unless L2 learners are true beginners, their battle to entrench new L2 segmental pronunciation habits will include to a large extent working to *un*learn previously acquired (i.e., now fossilized) mispronunciations of particular L2 words containing their problem phonemes. Students may additionally have habits of mispronouncing certain words due to interference from L1 sound-spelling correspondences (Young-Scholten & Archibald, 2000).

It is therefore vitally important teachers realize segmental training on *how* a sound is physically articulated is not enough to accomplish successful L2 segmental acquisition. Just as activities promoting the development of accurate L2 sound *perception* are vitally important to accurate L2 sound *production*, activities promoting accurate L2 sound production as a *habit* are mandatory. L2 learners' pronunciation will only be fluently accurate in the stream of speech if they can 1) *hear* L2 phonemes accurately, 2) *articulate* them accurately and 3) accomplish both tasks subconsciously. We therefore recommend the training sequence described below to maximize L2 learners' success in developing and automatizing new, more accurate segmental pronunciations.

A RECOMMENDED TRAINING SEQUENCE

In light of the three key realities impacting L2 segmental pronunciation acquisition described above, to be maximally effective, segmental pronunciation teaching must involve the following:

1) Developing the accuracy of students' L2 sound perception categorization (including the ability to discriminate L2 contrasts not existing in their L1)
2) Developing students' awareness of *how* L2 sounds are articulated and building their ability to physically articulate these sounds
3) Helping students undo old pronunciation habits for problem sounds and automatize new ones

Recommended resources and assignments for accomplishing each step of this training sequence are described below.

Developing accurate L2 sound perception

To build students' ability to hear the difference between various L2 English segmentals, we recommend two tools designed to build learners' proficiency in distinguishing minimal pairs their L1 may have trained them to hear as allophones. Linguatorium Auris (Qian, Chukharev-Hudalainen, & Levis, 2018) diagnoses each student's L2 English perception difficulties, adapting its recommended 10 minutes of daily High-Variability Phonetic Training (HVPT) homework to continually focus on the student's current problem segmentals. Linguatorium Auris' online version as well as app interface are transparent and easy-to-use for both students and teachers and its pricing is adaptable to different lengths of school term (with all proceeds dedicated to continued maintenance and research development by its parent nonprofit, the Andrey A. Hudyakov Center for Linguistic Research). Another valuable HVPT tool available for free online and as an iOS (but not Android) app is English Accent Coach (Thomson, 2012), though it requires students to learn the International Phonetic Alphabet (IPA) as well as to manually adapt its games to reflect their step-by-step segmental training progress (Sheppard, 2016).

Ideally, only when these tools indicate learners have receptively mastered a particular phoneme distinction should they begin being taught how to configure their vocal organs in order to *produce* that distinction. However, since the mind's process of breaking L1 allophones into separate L2 phonemes is not instantaneous but instead takes place apparently incrementally over time, we have found that if learners reach an apparent impasse at some intermediate point in the process of mastering a particular phoneme distinction receptively, taking several days off from perception training (presumably providing their minds time to consolidate what they have already learned) or even moving ahead to metaphonological training as well as physical articulation practice can help clarify to students what in their perception training they should be listening *for* and often provide the breakthrough needed. In other words, according to our experience, the old chicken-and-egg conundrum applied to pronunciation pedagogy – "Which comes first – perception or production?" – is misleading because complete mastery of new L2 phonemes in terms of either perception or production appears not to occur at a single point in time, but rather to develop incrementally.

A useful consciousness-raising tool for helping students notice their nonstandard pronunciation of particular words is YouGlish. Students asked to identify how their pronunciation differs from five or more YouGlish pronunciations of the U.S. state of "Illinois," for example, are often able to identify the point of difference accurately (e.g., YouGlish speakers do not usually pronounce "Illinois" with a final [z]). Not only that, but because of this exposure to many YouGlish speakers whose shared pronunciation for a given word differs from their own, students are often highly motivated to adjust individual word pronunciations toward a more standard form.

Developing accurate L2 sound production

To develop students' conscious understanding – or metaphonological awareness – of *how* a target segmental is physically articulated, a useful website for *assessing* whether students' inaccurate pronunciation results from lack of explicit knowledge about *how* to articulate a particular

segmental is Daniel Currie Hall's Interactive Sagittal Section. A useful website for *teaching* students how the various English (as well as Spanish and German) segmentals are physically articulated is the University of Iowa's Sounds of Speech.

#10	#9	#8	#7	#6	#5	#4	#3	#2	#1
/b, p/	**/f, h/**	/v, w/	/b, v/	/ð, v/	*/θ, ð/*	*/t, θ/*	*/tʃ, dʒ/*	*/ʃ, tʃ/*	*/f, θ/*
/f, p/	**/d, t/**	/s, z/	/f, v/	/s, ʒ/	*/s, θ/*			*/ʃ, ʒ/*	*/j, dʒ/*
/m, n/	**/g, k/**		/ð, z/		*/d, ð/*			*/j, ʒ/*	
/l, n/			/s, ʃ/		*/z, dʒ/*				
/l, r/					*/n, ŋ/*				

Highest functional load = **bold**
Lower functional load = regular
Lowest functional load = *italic*

Figure 2. Relative importance of English consonant pairs based on functional load (Brown, 1988; see also Catford, 1987; Jenkins, 2000, 2002; Munro & Derwing, 2006).

Because of the difficulty of breaking one's L1-trained pronunciation defaults, it is important students focus on only a few problem segmentals at any one time. The high-functional-load segmentals most responsible for distinguishing utterances should, of course, be prioritized over those that are low functional load, as exemplified in Figure 2 (Brown, 1988, 1991; Catford, 1987; Jenkins, 2000, 2002; King, 1967; Munro & Derwing, 2006).

Developing accurate L2 sound production automaticity

To develop students' segmental pronunciation automaticity, students can be provided adequate practice with new L2 segmentals via a series of activities that are at first relatively controlled but increasingly become less controlled and more authentic. Students who work just 10-15 minutes per day on building new L2 segmental pronunciation habits are likely soon to find their old L1-based habits no longer sound quite right to them and their new, more standard L2 pronunciations increasingly do.

A good initial tool for facilitating students' repeated practice of newly learned L2 segmentals and/or minimal pair contrasts toward the goal of entrenching more standard vocal configuration habits is Nilsen & Nilsen's (2010) *Pronunciation Contrasts in English.* A useful consciousness-raising tool for facilitating learners' noticing of *which* English words contain a newly learned or habitually mispronounced phoneme is *RelateWorldwide's North American English Pronunciation Highlighter*. Students simply:

1) Copy-and-paste into the Pronunciation Highlighter from a PowerPoint or Prezi presentation (or any other source text)
2) Indicate which problem segmentals they want highlighted (again, no more than 2-3 at a time is best since learners are limited in how many new segmental articulations they can focus on at any one time)
3) Optionally specify the HTML color code they prefer for highlighting each segmental
4) Click "Submit."

The *Pronunciation Highlighter* then outputs the text, providing lists of potential problem words broken down according to problem phoneme as well as displaying the entire text with all words containing one or more of the student's problem phonemes appropriately color-coded. Only when students can read through each phoneme's wordlist fluently as well as accurately should they practice the entire highlighted text, paying attention to pronouncing the problem phonemes in each highlighted word accurately.

For students in the habit of mispronouncing technical terms in their fields, either segmentally or in terms of lexical stress, a useful technique for identifying their nonstandard technical term pronunciations is having them systematically read aloud through the glossary of an undergraduate textbook introducing their field. For each listed term, students should:

1) pronounce the term,
2) embed the term in a sentence and
3) pronounce the term again.

Their teacher, meanwhile, should take notes on any pronunciation issues to be able to provide students individual instruction and tailored homework later. Students can similarly read aloud through conference or seminar presentation slides/posters to obtain their teacher's help in identifying general and technical vocabulary mispronunciations they characteristically use when discussing their specific research niche. This can be followed up by students repeatedly talking through one slide at a time of a PowerPoint/Prezi presentation, paying careful attention to their pronunciation of words (and particularly technical terms) they have historically mispronounced, as identified by RelateWorldwide's *North American English Pronunciation Highlighter*.

Once students are capable of articulating the standard pronunciation of a word but struggle to do so reliably, a useful method for "resetting" learners' pronunciation of a particular problem word is having them repeat the word in a thought group context for 10-15 new YouGlish examples every day for a week, paying careful attention to pronouncing each accurately (e.g., https://youglish.com/search/live vs. https://youglish.com/search/leave). By the end of this period, learners are likely immediately to recognize if they revert to their former mispronunciation and be able to correct it. Depending on how frequently students use a given problem word, in a very short time this can lead to their relatively completely recalibrating historically problematic word and ultimately phoneme pronunciations. YouGlish is thus a powerful tool for building L2 learners' fluency in *consistently* applying their new L2 word and segmental pronunciations to the stream of speech in real-life communication contexts.

CONCLUSION

To be maximally effective, L2 pronunciation instructors must work *with* rather than *against* the cognitive processes every human uses to understand receptively and speak productively any languages they know. A key reason L2 learners struggle to pronounce new segmentals is because their L1 has trained them to hear L2 phonemes as allophonic (Best & Tyler, 2007; Broersma & Cutler, 2008; Flege, 1995; Jenkins, 2000; Qian, Chukharev-Hudalainen, & Levis, 2018; Richards, 2012). Only if learners can reliably accurately *hear* L2 phoneme differences as phonemic are they

likely to be able reliably accurately to self-assess their pronunciation of these phonemes. Only if learners can reliably accurately self-assess their L2 pronunciation without consciously attending to the task are they likely to make progress toward *reliably* accurately configuring their vocal organs to distinguish challenging L2 phoneme distinctions in their real-life spoken communication. Certainly, it is important that L2 pronunciation instructors put focus on learners' mouths in their segmental pronunciation teaching. However, it is equally important that instructors take into account learners' minds along with the habitual nature of segmental pronunciation. Only then can optimal L2 segmental pronunciation training take place.

ABOUT THE AUTHORS

Monica Richards is a Postdoctoral Fellow in the Center for Communication Excellence (CCE) at Iowa State University. In this role, she has set up the CCE's English Speaking Consultation program as well as developed extensive materials on best practice in the teaching of L2 English pragmatics, listening, speaking, pronunciation and lexicogrammar for training new English Speaking Consultants. Monica is passionate about research-grounded L2 pedagogy, materials development and teacher training. Monica also cares deeply about building L2 English researchers' self-editing skills so their submissions to international English-medium journals are judged solely in terms of research quality rather than (lack of) adherence to English research writing norms.

Elena Cotos is an Assistant Professor in the Applied Linguistics Program at Iowa State University. She is also the Director of the Center for Communication Excellence of the Graduate College at Iowa State University. She investigates computer-assisted language learning and assessment, English for specific and academic purposes, functional language in instructional discourse, corpus and genre-based pedagogy, genre writing in the disciplines and automated writing evaluation.

REFERENCES

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning* (pp. 13-34). Amsterdam: John Benjamins.

Broersma, M., & Cutler, A. (2008). Phantom word activation in L2. *System, 36*(1), 22-34. https://doi.org/10.1016/j.system.2007.11.003

Brown, A. (1988). Functional load and the teaching of pronunciation. *TESOL Quarterly, 22*, 593-606.

Brown, A. (1991). Functional load and the teaching of pronunciation. In A. Brown (Ed.), *Teaching English pronunciation: A book of readings* (pp. 211-224). London, UK: Routledge.

Catford, J. C. (1987). Phonetics and the teaching of pronunciation: A systemic description of English phonology. In J. Morley (Ed.), *Current perspectives on pronunciation: Practice anchored in theory* (pp. 86-99). Washington, DC: TESOL.

Detey, S., & Nespoulous, J. L. (2008). Can orthography influence second language syllabic segmentation? Japanese epenthetic vowels and French consonantal clusters. *Lingua, 118*(1), 66-81.

Erdener, V. D., & Burnham, D. K. (2005). The role of audiovisual speech and orthographic information in nonnative speech production. *Language Learning, 55*(2), 191-228.

Flege, J. E. (1995). Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). Timonium, MD: York Press.

Ingels, S. (2010). The effects of self-monitoring strategy use on the pronunciation of learners of English. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 1st Pronunciation in Second Language Learning and Teaching Conference*, Sept. 2009. (pp. 67-89). Ames, IA: Iowa State University.

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

Jenkins, J. (2002). A sociolinguistically based, empirically researched pronunciation syllabus for English as an International Language. *Applied Linguistics, 23*(1), 83-103.

King, R.D. (1967). Functional load and sound change. *Language, 43,* 831–852.

Munro, M. J., & Derwing, T. M. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System, 34*, 520-531.

Moran, S., McCloy, D., & Wright, R. (Eds.). (2014). PHOIBLE Online. Retrieved November 28, 2018, from http://phoible.org

Nilsen, D. L. F., & Nilsen, A. P. (2010). *Pronunciation contrasts in English* (2nd ed.). Long Grove, IL: Waveland Press, Inc.

Qian, M., Chukharev-Hudilainen, E., & Levis, J. (2018). A system for adaptive high-variability segmental perceptual training: Implementation, effectiveness, transfer. *Language Learning & Technology, 22*(1), 69-96. https://doi.org/10125/44582

Richards, M. (2012). Helping Chinese learners distinguish English /l/ and /n/. In. J. Levis & K. LeVelle (Eds.), *Proceedings of the 3rd Pronunciation in Second Language Learning and Teaching Conference*, Sept. 2011. (pp. 161-167). Ames, IA: Iowa State University.

Scovel, T. (1969). Foreign accents, language acquisition, and cerebral dominance. *Language Learning, 19*(3–4), 245–253. https://doi.org/10.1111/j.1467-1770.1969.tb00466.x

Sheppard, B. (2016). English Accent Coach. *TESL-EJ, 20*(1), Available online: http://www.tesl-ej.org/wordpress/issues/volume20/ej77/ej77m2/

Thomson, R. I. (2012). Improving L2 listeners' perception of English vowels: A computer-mediated approach. *Language Learning, 62*(4), 1231-1258.

Vaux, B. (2000, January). *Flapping in English*. Paper presented at the meeting of the Linguistic Society of America, Chicago, IL.

Young-Scholten, M., & Archibald, J. (2000). Second language syllable structure. In J. Archibald (Ed.), *Second language acquisition and linguistic theory* (pp. 64-97). Oxford, England: Blackwell.