Haptic (Movement and Touch for Better) Pronunciation

Brian D. Teaman
Osaka Jogakuin University

William R. Acton
Trinity Western University

In this paper we describe a series of new techniques for the teaching of pronunciation using movement and touch. The "haptic approach" described here assumes that speaking is essentially a physical act that engages the entire body and not just the speech organs. This paper reviews the theoretical foundations of a haptic system, describes 9 haptic-based techniques, and explores the specific application of these techniques with Japanese learners of English.

この論文は、現在開発中の身体の動きと接触を利用した発音指導のための新しいテクニックについて書かれたものである。「触覚アプローチ」とは動作と接触を用いるという意味で、話すこととは、本質的に身体全体を使った身体的行動であり、単なる「発話器官」ではないという考え方に基づいている。この論文では、触覚アプローチのシステムの論理的根拠を考察し、その触覚に基づく9つの教授テクニックを紹介し、なぜ日本人の学習者にそれらのテクニックが効果的かを述べる。

WHERE IS the spoken language spoken? Using articulatory descriptions of spoken language, one might conclude that the spoken language is located only in the mouth. Detailed diagrams, such as a cross section of the human head, include terms that should be familiar to language teachers including *upper lip, upper teeth, alveolar ridge, hard palate,* and *velum*. There are more. How many more depends on how precisely you want to portray the vocal mechanism. From that perspective, it is almost as if the human head were but a laboratory for producing different sounds, not unlike the ingenious do-it-yourself vowel resonators created by Huckvale (2013). On the website, Huckvale shows how to add tubing of different shapes to a duck call in order to produce fairly convincing reproductions of a few English vowels. In the same way, we might be able to get our students to produce a few beautiful vowels in isolation, but it turns out language use is much more complicated than reproducing individual sounds.

In contrast to "mouth-centered" learning, we present a holistic or, to be more specific, a *haptic* (movement plus touch) approach that attempts to engage more of the rest of the body. Indeed, pronunciation is considerably more than a handful of vowels and consonants. If a student learns these 38 sounds of American English, as described in Ladefoged (1999, pp. 41-42), is the student now ready to speak a new second language? Unfortunately, those sounds are not spoken in isolation but may morph substantially when they appear in connected speech.

Making a **Difference**

In addition, there are complicated interactions with the body itself and with other levels of the language, such as the suprasegmentals that include other linguistic areas of intonation, rhythm, pitch register, and word and sentence stress. The haptic approach addresses the body first. So the answer to our question as to where spoken language resides is: in the body—the whole body.

The term haptic in its most basic form refers to touch plus movement. Haptics has recently gained importance in modern technology in varying applications such as haptic interfaces, which allow users to interact with devices such as smart phones or tablet computers using touch. Haptic gaming in applications such as the now familiar Wii connect the user with the gaming environment using touch and movement. Haptic video (or cinema) engulfs the watcher in a sensory shell that engages senses beyond the aural and visual aspects familiar to movie watching for the last 100 years or so (Marks, 2009.) Theatregoers are now able to "feel" the movie with vibrations, movement, wind, and other sensations, along with seeing and hearing it.

The Roots of Haptic Pronunciation Teaching: Acting and ESL

The work of Arthur Lessac (1997), well-known voice and acting teacher, was influential in the early development of the Essential Haptic-Integrated English Pronunciation (EHIEP) framework (Acton, 1994, 1997). His dictum, "Train the body first," serves as one of the basic principles of the approach outlined here. Lessac is recognized as one of the first to successfully reconcile fundamental mind-body "antagonism" in vocal training. Of course, in reality, virtually all speaking or pronunciation teachers use movement and gesture in class, whether for emphasis or depicting the "shapes" of sounds such as intonation contours "in the air." Beyond pronunciation alone, Asher (e.g., 1972) developed

the widely used Total Physical Response (TPR) approach for language teaching, which involves the use of movement in the form of commands and student nonverbal responses (McCafferty, 2004).

Examination of any student pronunciation textbook will reveal numerous recommendations for use of gesture or physical gimmicks to reinforce pronunciation learning, such as clapping hands or tapping feet during music, poetry, or jazz chants (e.g., Celce-Murcia, Brinton, Goodwin, & Briner, 2010; Morley, 1991). The important distinction of the haptic approach, however, is the extent to which movement and gesture are used *systematically* in classroom teaching.

Using the Visual Field to Teach Pronunciation: Observed Experiential Integration

The concept of the visual field in this approach is derived to some degree from the *Observed Experiential Integration* approach to psychotherapy developed by Cook and Bradshaw (2013) and others. In that system, eye movement is exploited in a number of ways, along with other sensory modalities such as focused touch (massage) or aroma therapy, in enhancing the efficacy and efficiency of therapy. EHIEP essentially establishes positions in the visual field in front of the learner and then directs learners to move their hands to designated positions as they say words or sounds to help them learn and recall new pronunciation. It was in that context about 6 years ago (Acton, 2010), that a possible solution to the problem of ineffective or inconsistent kinesthetic pronunciation teaching techniques came into focus.

Having made extensive use of kinesthetic procedures such as gesture and body movement for decades, Acton (e.g., 1984) had been looking for ways to use directed movement so that learner and instructor actions were performed *with sufficient consistency* and in prescribed patterns so that results could be measured

and methods replicated. Haptic engagement provided the answer (Acton, Baker, & Burri, 2009). Given almost any gesture in the visual field, if it could be anchored (or terminated) with touch (either hands touching each other or some spot on the upper body), the effect of the technique became much better defined and regularized—so that the impact could be explored in multiple contexts (Acton, Baker, Burri, & Teaman, 2012).

English Haptic-Integrated English Pronunciation

EHIEP represents the culmination of decades of work in pronunciation teaching by Acton (2013). It should be noted that the EHIEP model is quite experiential in nature. As such, this text-representation of it can only be a partial introduction to it, at best. One needs to at least see, if not experience it firsthand, to really understand its power in shaping the new phonology of an L2 English speaker. (See the *Haptic-integrated clinical pronunciation* blog, HICPR, 2013, for links to video demonstrations of EHIEP techniques.)

The core of the EHIEP system is the concept of *haptic-integrated*, defined as "the systematic engagement of hand movement through the visual field with a touch termination on a stressed syllable as the word is spoken" (HIPCR, 2013). In the EHIEP system, hand movement through the visual field is a crucial link in the haptic experience. Although the center of haptic anchoring is vocal resonance tied to movement and touch, that process has a visual complement, in which the haptic event is marked by the learner in a very multisensory process, making for a richer and deeper learning experience. There are three crucial, nearly simultaneous events that characterize our use of haptic: hand movement through the visual field, touch on a stressed syllable, and simultaneously spoken language. For a representation of one of the haptic movements, see Figure 1. For a video version please see https://vimeo.com/61198065. The nature of those elements will become clearer as the protocols and techniques are introduced in

the following section. This haptic "trinity" come together and are embodied in *pedagogical movement patterns* (PMPs).
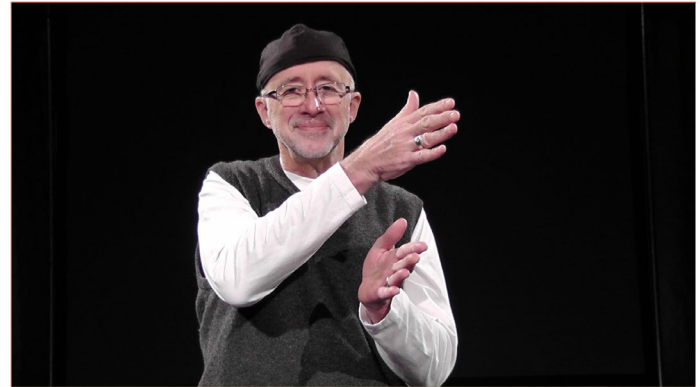

Figure 1. Demonstrates the EHIEP Rise-Fall Intonation Pattern. The right hand moves at chest level from the right to left (left to right here for the viewer) and strikes the stationary left hand on the nucleus of the intonational phrase, moves upwards with the rise of the intonation (this peak position between the rise and fall is pictured here), before falling once again to chest level along with a fall in pitch at the end of the phrase.

The EHIEP Protocols

There are nine basic protocols, which are ordered sets of procedures that train a learner in how to work with one PMP. Those PMPs are designed to be later used in classroom instruction or independent study. The protocols generally target one particular aspect of the L2 phonology. In briefly characterizing the nine protocols, note the use of the terms *left* hand and *right* hand or direction

across the visual field. The left/right distinction could, however, easily be reversed in some instances and often is performed both left/right and right/left within one protocol. There are both important pedagogical and neurophysiological reasons for choosing the specific handedness of a PMP for particular phrases (HICPR, 2013; Acton, in press). Essentially it has to do with exploiting brain hemispheric specificity. Some PMPs may be more effectively learned or used by designating more right hand (left brain) or left hand (right brain) engagement (Minogue & Jones, 2006).

The Warm-Up Protocol

The Warm-Up Protocol (WUP) functions to enliven the student's body and begin to accustom the student to the coming EHIEP tasks by having students move and produce several nonsense (English) syllables such as /i/, /wi/, /yi/, /hi/, or other pure vowels or diphthongized vowels. As in this protocol and the other protocols, the movement, touch, and vocal productions are modeled by a person recorded on video but could also be modeled by a trained instructor. The WUP is intended as a gentle introduction to the visual, tactile, and expressive or auditory anchoring experienced throughout the EHIEP protocols. The PMPs employed include moving from a central position in front of the body outward, movements of the hands from a low position below the waist upwards to a head level and snapping the fingers while speaking the syllables. The PMPs are then mirrored by the learner, as is the case for most of the protocols. Performing this warm-up should feel relaxing and rhythmic. It serves to connect the learner's body, visual space, and vocal tract in the L2.

Visual Field Anchoring Protocol

The Visual Field Anchoring Protocol (VAP) is a PMP that maps the vowel space to the positions on the clock. It serves as a

critical stepping-stone to the entire system. It is as if a clock is superimposed on the human body with 6 at the lateral center of the body at the level of the waistline and 12 centered just above the head. Three is at mouth height to the right and 9 at the same level but to the left. The corresponding vowels for these positions (going around the clock in a clockwise fashion) the tense /e/ at 3, open /a/ at 6, /o/ at 9, and finally the y-offglide at 12. There are vowels corresponding to each of the other numbers of the clock. It is not necessary to go into all of that detail for this brief description of the protocols, in part because that can vary, depending on the dialect of English that is targeted.

Vowel and Word Stress Protocol

The Vowel and Word Stress Protocol (VWSP) is based on the VAP clock described above. There are really three sub-protocols that follow the same basic form, but cover different parts of the vowel space. One VWSP is for the lax vowels—*rough vowels* in our terminology. Phonetically (using symbols from Celce-Murcia et al. (2010), they are written as [ɪ], [ʊ], [ɛ], [ɔ], [æ], [ʌ], and [ɑ]. The tense vowels and diphthongs, represented phonetically as [iy], [uw], [ey], [ow], [ay], [ɔy], and [aw] (again the symbols are taken from Celce-Murcia et al., 2010)—are called *smooth vowels* in our terminology. This protocol involves learning the vowels as isolated syllables and then practicing them in poems that cover the vowels from "top to bottom" on the vowel clock in words. For example the VWSP for rough (lax) vowels uses the poem: "if it fits the f<u>oo</u>t, I b<u>e</u>t you b<u>ou</u>ght, the h<u>a</u>t and the st<u>u</u>ff from the sh<u>o</u>p of my p<u>o</u>p!" The poem is recited rhythmically with precise movements that correspond to the key underlined vowels.

Syllable Protocol

The Syllable Protocol focuses on syllables and their relative

prominence. For this protocol, one hand rests on the opposite shoulder with the other hand's palm cupping the opposite elbow. In utterances containing from one to seven syllables, of which one and only one syllable is stressed, fingers gently tap near the elbow for unstressed syllables and the other fingers tap near the shoulder for the one stressed syllable. To demonstrate the tapping pattern here we will use "S" for the shoulder (focal) tap and "e" for the elbow tap marking the nonfocal syllables. The tapping pattern for the two-syllable phrase *that's nice* the pattern is e-S**.** The seven-syllable phrase *that's very interesting* would have the corresponding tap pattern of e-e-e-S-e-e-e.

Intonation Protocol

The Intonation Protocol focuses on embodying intonation contours (Acton, Baker, & Burri, 2009). For this protocol one stationary hand is raised to shoulder level while the opposite hand moves at the same level, crossing the center of the body and touching the stationary hand on the stressed syllable. After meeting the stationary hand, one of several possible intonation contours is traced by the moving hand, depending on the intended phrase. For a level monotone contour, the voice and the hand continue in a flat trajectory without varying pitch. For a rise or fall, the voice and hand either rise together or fall together. At the end of the sentence, there can be an additional final fall after the tonic contour.

Fluency Protocol

The Fluency Protocol targets the speed and fluidity of speech by accompanying speech with quick fluid movements with simultaneous tapping. In the starting position, the learner's left hand is open, positioned adjacent to the left quadricep. The right arm is in the same position, on the right side. For example, as the phrase *Tricky?* is spoken with a rising intonation, the right hand moves upwards and taps the left hand on the stressed syllable and continues upward. For *That's tricky?* the exact same movement is performed with the addition of a hip tap with the right hand on *that*. Other phrases are repeated as a reply using the reverse hand positions. This protocol creates a soothing, fluid, and rhythmic effect using Tai Chi-like moves.

Rhythm Protocol

The Rhythm Protocol (RP) is performed with hands in the position of a jab used in boxing. The hand alternates syllables with a forward (f) punch and return (r) to the original position. On the tonic syllable of the phrase, a large forward jab (F) accompanies it. So the phrase *That's easy* would be r-F-r. *That's very easy* would be r-f-r-F-r. These two phrases begin with the return because you need to end up with a forward movement on the stressed syllable jab. So the phrase *That's amazing!* would start with a quick forward jab and therefore be f-r-F-r. The function of the rhythm protocol is to compact the syllables of speech, especially the unstressed syllables, creating a much more conversational "felt sense" for the learner. The RP is generally the one with the most immediate impact on conversational speaking style.

Expressiveness Protocol

Building on the Intonation Protocol, the Expressiveness Protocol takes the intonational contours and then situates them in conversational discourse by assigning them (a) pitch, (b) volume, and (c) pace, along with explicit reference to "discourse orientation," that is the relation of that discourse turn to the previous one of the other person in the conversation. For example, if the conversational turn of one speaker was FAST, HIGH PITCH, and LOUD, the response might well match those three parameters or intentionally change the intensity somewhat, depending on the emotional and textual qualities of what is to follow.

For high pitch, the hands would be moving near the top of the visual field; for low pitch, near the bottom. For differences in volume, the hands may either move more quickly or move further away from the body. For differences in pace, the silence after a rhythm group is either shortened or lengthened by changing the time between gestures.

Integration Protocol

The Integration Protocol has been developed, as its name suggests, to assist learners in integrating key elements of the EHIEP system: vowel quality, stress assignment, rhythm grouping, intonation contour/tone groups, expressiveness, fluency, conversational speed—and new or changed consonant sounds. Just as in the case of the conductor of an orchestra, the baton (or pencil or rod or stick of some kind) serves to set up the basic rhythmic beat or tempo and expressive intensity or volume of the phrase or sentence being repeated or produced for the first time. The effect is to drive the speech mechanism to more rapid and contrastive production. The baton simply takes over direction and the voice and body follow. It is also, in some contexts, a very reliable tool in initial diagnostic work. If a learner can move his or her baton with their speech at the outset, the prognosis for their rapid improvement is good.

These nine protocols can be used in varying ways. Using a rather selective approach, certain protocols could be introduced as stand-alone activities used to reach specific teaching goals. For example, the WUP can be done in a few minutes at the beginning of class and serves as a great way to invigorate students and get them focused and ready to use their bodies for speaking. The Syllable Protocol can be quickly taught and used to reinforce the learning of syllabification and stress for new vocabulary. Many other stand-alone scenarios are possible; however, the most effective use of these protocols would involve working with the complete set over a couple of months. In the

approach that we have developed and that we commonly use in our classes, one or two protocols are introduced every week in class with a video and practice for about 20 to 30 minutes. Students then practice the protocol by themselves a few times in the following week. As new protocols are added, old protocols can be reviewed and practiced periodically to keep the varying haptic skills alive for the learner. Within each unit are chances to practice short dialogues that allow learners to focus on producing language in context. With this kind of extended training and practice from individual sounds to words, phrases, and conversations, learners can expect gains in their pronunciation accuracy even in spontaneous speaking.

Conclusion

Several aspects of the EHIEP system were developed in Japan when we were teaching together in the 1990s at Nagoya University of Commerce. In working in large conversation classes, it was essential to be able to assist individual students with correcting pronunciation. An earlier kinesthetic version of the EHIEP method that we developed there provided important insights that later evolved into the current system. The EHIEP system is especially effective with Japanese learners for two reasons. First, the body-based rhythm group focus is very helpful in assisting student in moving away from their more syllable-by-syllable way of speaking of English. Second, the location in the visual field of several key vowels of English that are not part of the Japanese phonological system has proven to be a feature of the system that learners immediately identify as most helpful, especially in being able to produce the distinctions between word pairs such as: *sit/seat, let/late, kook/cook,* and *coat/caught.*

One major advantage of the EHIEP system is that for most instructors, regardless of background in pronunciation teaching, it is reasonably simple for them to learn the techniques and teach them to students so that they can later be used in

integrated classroom instruction. In other words, once students have experienced the haptic PMPs, those PMPs are ready to be used whenever the pronunciation of a new word is targeted or a correction is executed. Another key advantage of the EHIEP approach should be the ability to use improved pronunciation in spontaneous speech. The haptic anchoring of sounds enables integration of changed pronunciation, just as the use of haptics in various kinds of physical training has been shown to be exceedingly effective in integrating knowledge in many disciplines. Finally, it is guaranteed to be a moving experience for both you and your students.

Bio Data

Brian D. Teaman does research in CALL and speaking/phonetics/pronunciation. In his recent sabbatical he spent a lot of time learning and researching haptic pronunciation techniques in British Columbia with Bill Acton. He is also an avid cyclist and vocalist. <teaman@wilmina.ac.jp>

William R. Acton's research and publications center on the role of movement in pronunciation teaching. He also plays harmonica and is a distance runner. His blog address is: www.hipoeces.blogspot.com. <william.acton@twu.ca>

References

Acton, W. (1984). Changing fossilized pronunciation. *TESOL Quarterly, 18*, 71-85.

Acton, W. (1994). Directed movement in language instruction, *Journal of Communication and International Studies (Nagoya University of Commerce)*, *1*, 43-52.

Acton, W. (1997). Direct speaking instruction (and the mora-bound, focal-stress blues). *The Language Teacher, 21*(9), 7-12. Retrieved from http://jalt-publications.org/tlt/files/97/sep/acton.html

Acton, W. (2010, March). *Full-bodied, multiple modality pronunciation teaching*. Paper presented at the 2010 TESOL Convention, Boston, MA.

Acton, W. (2012, March). *Post pronunciation pronunciation*. Paper presented at the 2012 TESOL Convention, Philadelphia.

Acton, W. (2013). *Essential haptic-integrated English pronunciation teaching system* (haptic video instructional modules, student workbook and audio recordings, instructor manual). Manuscript in preparation.

Acton, W. (in press). Accent on OEI. In R. A. Bradshaw & A. Cook (Eds.), *Toward integration: One eye at a time: Clinical applications*. Vancouver, BC, Canada: SightPsych Seminars.

Acton, W., Baker, A., & Burri, M. (2009, March). *Haptic approaches to grammar and intonation instruction.* Paper presented at the 2009 TESOL Convention, Denver, CO.

Acton, W., Burri, M., Baker, A., & Teaman, B. (2012, August). *Preliminaries to haptic-integrated pronunciation instruction.* Paper presented at the 2012 Pronunciation in Second Language Learning and Teaching Conference, Vancouver, BC, Canada.

Asher, J. (1972). Children's first language as a model for second language learning. *Modern Language Journal, 56*, 133-139.

Celce-Murcia, M., Brinton, D. M., Goodwin, J. M., & Griner, B. (2010). *Teaching pronunciation: A reference for teachers of English to speakers of other languages* (2nd ed.). Cambridge: Cambridge University Press.

Cook, A., & Bradshaw, R. (2013). Sight-psych seminars. Retrieved from http://www.sightpsychology.com/

HICPR. (2013). Haptic-integrated clinical pronunciation. Retrieved from http://hipoeces.blogspot.ca/

Hucktvale, M. (2013). Make your own vowel resonators. Retrieved from http://www.phon.ucl.ac.uk/home/mark/vowels/

Ladefoged, P. (1999). American English. In *Handbook of the International Phonetic Association* (pp. 41-44). Cambridge: Cambridge University Press.

Lessac, A. (1997). *The use and training of the human voice: A bio-dynamic approach to vocal life.* Boston, MA: McGraw-Hill.

Marks, L. (2009). *Touch: Sensuous theory and multisensory media*. Minneapolis: University of Minnesota Press.

McCafferty, S. G. (2004). Space for cognition: gesture and second language learning. *International Journal of Applied Linguistics, 14*, 148-165. doi: 10.1111/j.1473-4192.2004.0057m.x

Minogue, J., & Jones, M. (2006). Haptics in education: Exploring an untapped sensory modality. *Review of Educational Research*, 76, 317-348.

Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly, 25*, 481-520.