Ultrasound-Integrated Pronunciation Teaching and Learning

Noriko Yamane, Jennifer Abel, Blake Allen, Strang Burton, Misuzu Kazama, Masaki Noguchi, Asami Tsuda, and Bryan Gick

University of British Columbia

1. Introduction

Pronunciation is an integral part of communication, as it directly affects speakers' communicative competence and performance, and ultimately their self-confidence and social interaction. Second language (L2) pronunciation is one of the most challenging skills to master for adult learners. Explicit pronunciation instruction from language instructors is often unavailable due to limited class time; even when time is available, instructors often lack knowledge of effective pronunciation teaching and learning methods. Imitating native speakers' utterances can be done independently from classroom learning, but the absence of feedback makes it difficult for learners to improve their skills (e.g., de Bot, 1980; Neri et al., 2002). As well, learning to articulate difficult or unusual sounds can be made more challenging when learners have only auditory input, as the mapping from sound to articulation is not always straightforward (e.g., Wilson & Gick, 2006; Gick et al., 2008).

In an effort to improve pronunciation instruction, the Department of Linguistics and the Japanese language program in the Department of Asian Studies at the University of British Columbia began a collaboration in 2014 designed to develop new multimodal approaches to pronunciation teaching and learning. The Japanese language program is the largest language program at UBC, with more than 1,500 students enrolled every year, and is also known to be the most diverse in terms of learners' language backgrounds. The project is developing online resources to allow learners of Japanese to improve their pronunciation, as well as to allow Linguistics students to better understand sound production. The key technological innovation of this project is the use of ultrasound overlay videos, which combine mid-sagittal ultrasound images of tongue movement in speech with external profile views of a speaker's head to allow learners to visualize speech production. This technology is currently being extended to create an interactive tongue visualizer, which will allow learners to see their lingual articulations overlaid on video of their head in real time.

2. Methods

Ultrasound of native speakers of Japanese and of English was recorded using an Aloka ProSound SSD-5000 system, and the exterior video was recorded using a JVC camcorder (GZ-E300AU). Both recordings were made at 30 frames per second. The exterior video showed the left profile of the speaker's head. A clapper was used to generate an audio alignment point.

The ultrasound overlay videos were created from raw footage using a four-step process. First, the ultrasound and exterior video were trimmed using Adobe Premiere to ensure alignment. Next, all elements of the ultrasound image aside from the tongue were manually erased using Adobe After Effects. The brightness of the tongue was increased, and the colour was changed from white to a shade of pink (colour #DE8887 in Adobe After Effects) to more closely resemble the human tongue. Then, the erased ultrasound image was overlaid on the exterior face video using Adobe After Effects. Scaling of the two sources was achieved by ensuring that the shadow of the probe in the ultrasound image is the same width as the top of the probe in the exterior video. The results of this process are exemplified in Figure 1.
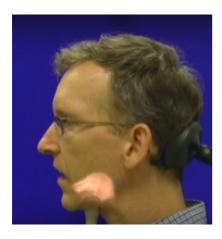
Figure 1. Ultrasound overlay video frame of [χ]

3. Results

The videos are available to the public through the eNunciate website (http://enunciate.arts.ubc.ca/), and are licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. The videos are categorized into 'Linguistics' and 'Japanese' content, although all pages are open to all students and instructors.

3.1 Linguistics content

The Linguistics pages feature ultrasound overlay videos of canonical examples of the sounds of the world's languages produced in basic contexts: in the [_a] and [a_a] positions for consonants, and in isolation or in [C_C] contexts for vowels. Freeze frames are inserted in these videos to capture key moments in the articulation (e.g., the stop closure in a stop articulation), and beginning and end titles are inserted. These videos can be accessed through interactive IPA consonant and vowel charts. In addition to the ultrasound overlay videos, videos introducing the use of ultrasound in linguistics and the basics of vowel and consonant articulation are available.

In the Fall 2015 term, four UBC Linguistics courses used the resources: two general introductory courses (Linguistics 100 and 101), one introduction to phonetics and phonology course (Linguistics 200), and one upper-year acoustic and instrumental phonetics course (Linguistics 313). In Linguistics 200, of the 26 students who responded to a voluntary survey, 23 (88%) indicated that the resources were easy to use and that they helped them understand how sounds are articulated, 21 (81%) indicated that the resources helped them understand the differences between sounds, and 24 (92%) indicated that they would recommend the resources to other students. Data collection on student use of and satisfaction with the resources from these courses is ongoing.

3.2 Japanese content

The Japanese pages include instructional and exercise videos for Japanese pronunciation teaching and learning. These videos incorporate narration, cartoons, and animations in addition to ultrasound overlay elements, and are augmented with quizzes to allow students to reinforce what they have learned using the videos. The videos are grouped into three categories: introductory, which includes introductions to Japanese sounds and to basic phonetic concepts; 'challenging sounds', which features videos focusing on problems that L2 learners from different language backgrounds may encounter; and intonation.

In the Fall 2015 term, the eNunciate video resources are being used in two sections of the beginner-level Japanese 102 course, which are taught by the same instructor. In one of these sections, the student will also receive a half-hour ultrasound feedback session with the first author to help improve their pronunciation. These sections are being compared with a third section, also taught by the same instructor, in which neither eNunciate resources nor ultrasound feedback are being used, to determine if use of these resources will lead to a greater improvement in students' Japanese pronunciation than 'traditional' pronunciation practice.

Table 1. Implementation of eNunciate resources and ultrasound feedback session in three sections of Japanese 102 at the University of British Columbia.

	'Traditional' Pronunciation Practice Section	Pronunciation Practice with eNunciate Section	Pronunciation Practice with eNunciate and Ultrasound Section
Activities	• Shadowing • Listening to Audio	• Watching eNunciate 'Challenging Sounds' Videos • Listening to Audio	• Watching eNunciate 'Challenging Sounds' Videos • Ultrasound Feedback Session
Assessment by Students	• Survey • Reflection Paragraph	• Survey • Reflection Paragraph	• Survey • Report on Ultrasound Feedback Session
Assessment of Students	• Perception Test • Recording Assignment	• Perception Test • Recording Assignment	• Perception Test • Recording Assignment

Data collection on student use of and satisfaction with the resources from these courses is ongoing.

4. Discussion: developments in progress

4.1 Interactive tongue visualizer

As part of our plan to use biofeedback to facilitate L2 pronunciation learning, we are developing an interactive tongue visualizer, which will automate creation of the type of ultrasound overlay videos described in section 2 based on ultrasound and video feeds of a speaker producing sounds in real time. Development of this tool is still in the early stages. The visualizer will be implemented at a physical location ("Pronunciation Station") at UBC, and will be equipped with a CHISON ECO 1 portable ultrasound with a 6.0MHz D6C12L Transvaginal probe.

4.2 Ultrasound training

To overcome the lack of a standardized procedure for the teaching of L2 pronunciation with ultrasound imaging, we are developing guidelines based on the procedures previously used in the settings of L2 learning (Gick et al., 2008) and speech language pathology (Bernhardt et al., 2005). The guidelines target three consecutive days of teaching to allow teachers to use the Pronunciation Station: (1) initial evaluation of students' pronunciation, (2) training with ultrasound images as biovisual feedback, and (3) post-training evaluation of students' pronunciation. As a case study, we implemented the protocols in teaching Japanese pronunciation to four native speakers of Korean, particularly focusing on the acquisition of the contrast between alveolar and alveo-palatal sibilants (e.g. [za] vs. [ʑa]), which is known to be especially difficult for Korean speakers. The results suggest that the protocols are effective:

the two beginner learners, one advanced learner, and one heritage speaker, none of whom had any significant contrast between those sounds in pre-training recording, showed a significant contrast in post-training recording.

4.3 Expansion to additional languages

In 2016, we intend to begin development for materials for additional languages being taught at UBC: Chinese, French, Spanish, German, and English as a second/additional language.

Acknowledgements

References

Bernhardt, B., et al. (2005). Ultrasound in speech therapy with adolescents and adults. *Clinical Linguistics & Phonetics* 19.6-7: 605-617.

de Bot, C. L. J. (1980). The role of feedback and feedforward in the teaching of pronunciation. *System*, 8, 35-45.

Gick, B., et al. (2008). Ultrasound imaging applications in second language acquisition. In J. G. Hansen Edwards and M. L. Zampini (eds.), *Phonology and Second Language Acquisition* (pp. 309-322). Amsterdam: John Benjamins.

Neri, A., et al. (2002). The pedagogy-technology interface in computer assisted pronunciation training. *Computer Assisted Language Learning*, 15(5), 441-467.

Pillot-Loiseau, C., et al. (2015). French /y/-/u/ contrast in Japanese learners with/without ultrasound feedback: vowels, non-words and words. Paper presented at ICPhS 2015. Retrieved August 12, 2015 from http://www.icphs2015.info/pdfs/Papers/ICPHS0485.pdf.

Wilson, I., & Gick, B. (2006). Ultrasound technology and second language acquisition research. In M. Grantham O'Brien, C. Shea, and J. Archibald (eds.), *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006)* (pp. 148-152). Somerville, MA: Cascadilla Proceedings Project.