Ultrasound Technology and Second Language Acquisition Research

Ian Wilson and Bryan Gick

University of Aizu and University of British Columbia

1. Introduction

Research on the acquisition of the ability to pronounce a second language has been dominated by acoustic studies as opposed to articulatory studies (e.g., Birdsong, in press; Flege et al., 2003; Morrison, 2006). Although acoustic studies are certainly a valuable means of observing one's progress in learning to produce a second language, they deal with the sound that is produced and infer the articulation that produced the sound. Articulatory studies look directly at the articulators (e.g., the tongue, the lips, the jaw, etc.) and can often give a more accurate picture of the actions performed by the pronunciation learner. In the case of pronunciation *teaching*, methods that employ acoustic data risk a lack of understanding on the part of the learner as to how to map the acoustic information onto articulatory movements. If learners are able to see directly the articulators, then they probably have an improved perception of the articulatory adjustments needed to improve their pronunciation.

Although the configuration of a teacher's lips and jaw can be observed easily enough, and the configuration of one's own lips and jaw can be seen in a mirror, the position and movements of the tongue, especially all but the most anterior part, cannot be seen. It is true that proprioceptive feedback of the position of one's own tongue is available, but the amount of feedback depends on the amount of contact the tongue makes with the teeth, gums and palate. Learners producing sounds such as vowels (other than high, front vowels) and liquids usually have less proprioceptive feedback than when producing other sounds. Thus, for these sounds direct visual biofeedback would be helpful both to a learner and to a researcher interested in observing differences in articulation of sounds within and across languages as well as interlanguages. This paper reviews ultrasound imaging, one such method of direct visual feedback.

2. Ultrasound Imaging of the Tongue

Ultrasound imaging is a non-invasive method of observing the position and movements of the tongue in real-time. As a method of observing the movements of the tongue for speech research, one-dimensional ultrasound was first used about 40 years ago (e.g., Kelsey et al., 1969) effectively allowing one point at a time on the tongue's surface to be seen. Two-dimensional ultrasound has been used in speech research for 25 years (since Sonies et al., 1981). However, only recently have higher image quality and greater affordability made ultrasound viable for research and pedagogical use in the acquisition of the pronunciation of a second language.

Figure 1 shows a setup used for pronunciation teaching and learning at the University of Aizu. The learner is seated in front of the ultrasound monitor and is holding the ultrasound probe against his neck in the submental region. On the ultrasound monitor, he can see the image of his tongue moving in real-time as he speaks. The learner's speech is recorded using a microphone, and this audio signal together with the ultrasound video signal are fed through a digital video converter and captured on a laptop computer. The captured movies are then available to be used later by the teacher and learner in post-session analysis. If the movies are required to be of sufficient quality for research purposes (i.e., to use for making measurements of tongue position), then it is recommended to limit the amount of probe movement relative to the head (Gick et al., 2005). This can be done by placing the probe on a microphone stand and having the subject lean his/her head against a stable object such as a wall or the high back of a chair.

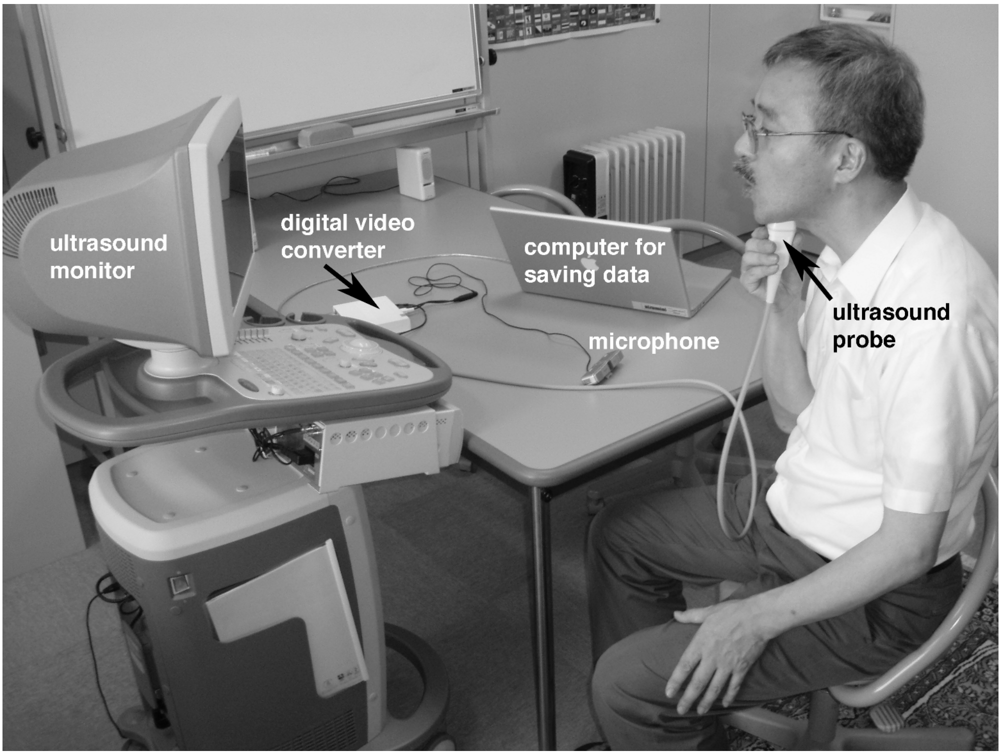

Figure 1. Ultrasound equipment setup for pronunciation session at the University of Aizu

Very high frequency sound waves from the ultrasound probe travel upward through the tongue and are reflected back off the upper surface of the tongue (see Epstein & Stone, 2005 for more details). By computing the time it takes for the sound waves to reflect back to the probe, the ultrasound machine can calculate the distance from the probe to the tongue's surface and then plot an image of the tongue with a dominant white line for the upper surface. A typical midsagittal view of the tongue observed using ultrasound can be seen in figure 2.

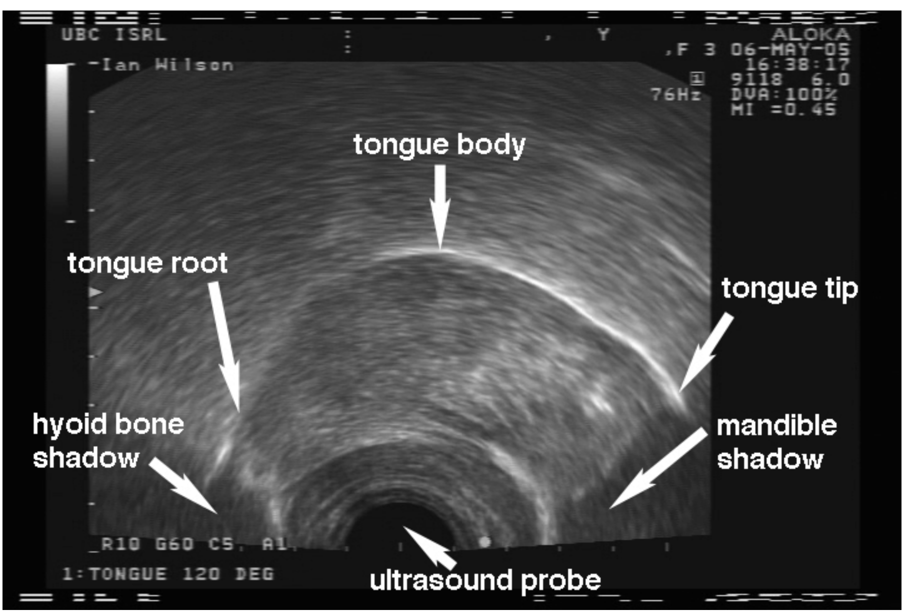

Figure 2. A midsagittal ultrasound image of the tongue (prominent white line = tongue surface)

As Epstein and Stone (2005, p. 2128) point out, an ultrasound image of the tongue is “limited anteriorly and posteriorly by ‘acoustic shadows’ (black regions) created by refraction of the ultrasound beam off the mandible and hyoid bones.” However, depending on the subject, as well as the field of view of the probe, it is often the case that the whole tongue from the tongue tip to the tongue root can be simultaneously viewed (as in figure 2). In addition, by twisting the probe 90 degrees, a coronal view of the tongue can be observed (not shown here), a view that is useful when studying midline tongue grooving in sounds such as /s/ and when studying small differences in tongue height.

3. Specific Applications of Ultrasound in SLA Research and Pedagogy

The complexity of the shape of the tongue for North American /r/ and /l/ is something that can be observed using only a method such as ultrasound, and in recent studies ultrasound has proven useful in teaching these complex sounds. In Adler-Bock (2004) and Adler-Bock et al. (forthcoming), two monolingual adolescents, who had not yet acquired the pronunciation of /r/, showed significant improvement after 13 training sessions with ultrasound. In Gick et al. (to appear), three reasonably advanced Japanese learners of English made noticeable improvements to their /r/ and /l/ pronunciation after only 1 hour of ultrasound training. The training was very interactive though, and included such activities as using tracing paper to draw the shape of one’s tongue as seen on the ultrasound. This type of activity, asking the learner to focus intensely on the articulation has been shown to be useful in other studies (e.g., Catford & Pisoni, 1970).

One of the ways that ultrasound is particularly valuable is in its ability to show the front and back of the tongue simultaneously. In the production of some consonants, e.g., /l/ and /r/, the timing of the movements of different parts of the tongue (e.g., the tongue tip, tongue dorsum, and tongue root) is important and depends on the position of the consonant in the syllable—i.e., onset versus coda (Campbell, 2004). In such cases, mistiming by an L2 learner can result in pronunciation errors and ultrasound serves as a useful diagnostic tool for determining the nature of those errors. For example, Wilson (2003) presented ultrasound data showing a Japanese native speaker attempting to say the English word “bell”, but actually pronouncing the word with a vocalized /l/. Her tongue dorsum was retracting during the pronunciation of the vocalized /l/, but then her tongue tip was raising after the sound had faded away. What the ultrasound data indicated (and the ear couldn’t hear) was that this L2 learner was aware of the tongue tip gesture for /l/. Thus, the pronunciation problem was a timing problem, not a problem of the reduced magnitude or complete absence of a gesture.

Ultrasound has also been shown to be a promising tool for hearing impaired individuals, fluent in sign language, who wish to improve their pronunciation of a spoken language (Bernhardt et al., 2003; Bernhardt et al., 2005). An example, other than the English liquids, of where ultrasound is extremely valuable is in the production of tense versus lax English vowels. MacKay (1977) showed that, except for /o/, all of the English tense vowels are produced with an advanced tongue root. With ultrasound biofeedback of the position of the tongue root, it becomes much easier to see and make the distinction between English tense and lax vowels. For L2 learners of English who do not have a tense/lax distinction in their native language, this is particularly salient.

Another area of SLA pronunciation research that has benefited from the use of ultrasound is research on articulatory setting, the underlying setting of the tongue and other articulators during speech. Although articulatory setting is something that has interested phoneticians for centuries (see Kelz, 1971, Laver, 1978, and Jenner, 2001 for historical surveys), its existence had not been shown empirically until a study by Gick et al. (2004). That study investigated articulatory setting by looking at inter-speech posture (the position of the articulators when they are motionless during inter-utterance pauses). To observe the position of the tongue, Gick et al. (2004) used existing x-ray movie films with limited spatial resolution and clarity, and they had no control over the linguistic stimuli or how they were presented to the subjects. However, they did find significant differences between, among other things, the inter-utterance tongue position for Québécois French and that for Canadian English.

Wilson (2006) partially replicated the Gick et al. (2004) study, both improving on the methodology by using ultrasound (thus allowing new stimuli to be constructed and collected) and expanding the study to include bilingual speakers of various proficiencies. He found that, on average, a group of seven monolingual English speakers held their tongue tip significantly higher during inter-utterance pauses than a group of eight monolingual French speakers. This finding mirrored that of Gick et al. for the tongue tip, thus providing strong evidence for the existence of articulatory settings (if they are indeed observed during inter-utterance pauses). Among the nine bilingual speakers tested, two of the four who were perceived as being native speakers of both English and French had a significantly higher English tongue tip during pauses (mirroring the difference across monolingual groups). None of the other seven bilinguals showed this difference. Thus, this bilingual result partially supported the monolingual group findings showing that one aspect of the articulatory setting for Canadian English is that it has a higher tongue tip than the articulatory setting for Québécois French.

4. Conclusions and Future Directions

Preliminary work described above has shown beneficial effects of the role of ultrasound biofeedback in the acquisition of the pronunciation of a second language. The importance of ultrasound in second language acquisition research has also been described, especially with respect to articulatory settings. However, there exists an immediate need for further studies of the effects of using ultrasound and other methods of direct biofeedback in the teaching and acquisition of the pronunciation of a second language. Studies are necessary to determine not only the effect ultrasound has on the acquisition of pronunciation, but more specifically the effect it has on the rate of acquisition and the ultimate level of attainment. Further studies of articulatory setting are also needed, expanding the coverage of the languages observed.

Using ultrasound as a means of visual biofeedback could shed light on the relationship between L2 speech perception and L2 speech production (see Llisterri, 1995 for a good review of the literature on this relationship). For example, it may be possible to train speakers to *produce* a consistent difference between two sounds before they are able to *perceive* the difference. Goto (1971) and Sheldon and Strange (1982) found that at least some Japanese adults had more accurate production of the difference between English /r/ and /l/ than their perception of that difference in their own speech or that of others. It is interesting to speculate how accurate one could get in producing two sounds without being able to perceive the difference. Riney and Flege (1998, p. 240) point out that "over time, greater accuracy in perception does not necessarily lead to greater accuracy in production." However, to our knowledge the question remains unanswered of whether a marked increase in one's ability to *produce* a difference between phonemes leads to an improvement in one's *perception* of those differences.

References

Adler-Bock, M. (2004). Visual feedback from ultrasound in remediation of persistent /r/ errors: Case studies of two adolescents. MS thesis, University of British Columbia.

Adler-Bock, M., Bernhardt, B., Gick, B., & Bacsfalvi, P. (forthcoming). The use of ultrasound in remediation of North American English /r/ in two adolescents.

Bernhardt, B., Bacsfalvi, P., Gick, B., Radanov, B., & Williams, R. (2005). Exploring electropalatography and ultrasound in speech habilitation. *Journal of Speech-Language Pathology and Audiology 29,* 169–182.

Bernhardt, B., Gick, B., Bacsfalvi, P., & Ashdown, J. (2003). Speech habituation of hard of hearing adolescents using electropalatography and ultrasound as evaluated by trained listeners. *Clinical Linguistics and Phonetics 17,* 199–216.

Birdsong, D. (in press). Nativelike pronunciation among late learners of French as a second language. In M. J. Munro & O.–S. Bohn (Eds.), *Second language speech learning: The role of language experience in speech perception and production*. Amsterdam: John Benjamins.

Campbell, F. M. (2004). The gestural organization of North American English /r/: A study of timing and magnitude. MA Thesis, University of British Columbia.

Catford, J. C., & Pisoni, D. B. (1970). Auditory vs. articulatory training in exotic sounds. *Modern Language Journal 54,* 477–481.

Epstein, M. A., & Stone, M. (2005). The tongue stops here: Ultrasound imaging of the palate. *Journal of the Acoustical Society of America 118,* 2128–2131.

Flege, J. E., Schirru, C., & MacKay, I. R. A. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication 40,* 467–491.

Gick, B., Bernhardt, B., Bacsfalvi, P., & Wilson, I. (to appear). Ultrasound imaging applications in second language acquisition. In J. Hansen & M. Zampini (Eds.), *State–of–the–Art Issues in Second Language Phonology*. Cambridge, U.K.: Cambridge University Press.

Gick, B., Bird, S., & Wilson, I. (2005). Techniques for field application of lingual ultrasound imaging. *Clinical Linguistics & Phonetics 19,* 503–514.

Gick, B., Wilson, I., Koch, K., & Cook, C. (2004). Language-specific articulatory settings: Evidence from inter–utterance rest position. *Phonetica 61,* 220–233.

Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds "L" and "R". *Neuropsychologia 9,* 317–323.

Jenner, B. (2001). 'Articulatory setting': Genealogies of an idea. *Historiographia Linguistica XXVIII,* 121–141.

Kelsey, C. A., Woodhouse, R. J., & Minifie, F. D. (1969). Ultrasonic observations of coarticulation in the pharynx. *Journal of the Acoustical Society of America 46,* 1016–1018.

Kelz, H. P. (1971). Articulatory basis and second language teaching. *Phonetica 24,* 193–211.

Laver, J. (1978). The concept of articulatory settings: An historical survey. *Historiographia Linguistica V,* 1–14.

Llisterri, J. (1995). Relationships between speech production and speech perception in a second language. In K. Elenius & P. Branderud (Eds.), *Proceedings of the XIII[th] International Congress of Phonetic Sciences* (Vol. 4, pp. 92–99). Stockholm: KTH / Stockholm University.

MacKay, I. R. A. (1977). Tenseness in vowels: An ultrasonic study. *Phonetic, 34,* 325–351.

Morrison, G. S. (2006). L1 & L2 production and perception of English and Spanish vowels: A statistical modelling approach. PhD dissertation, University of Alberta.

Riney, T. J., & Flege, J. E. (1998). Changes over time in global foreign accent and liquid identifiability and accuracy. *Studies in Second Language Acquisition 20,* 213–243.

Sheldon, A., & Strange, W. (1982). The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics 3,* 243–261.

Sonies, B. C., Shawker, T. H., Hall, T. E., Gerber, L. H., and Leighton, S. B. (1981). Ultrasonic visualization of tongue motion during speech. *Journal of the Acoustical Society of America 70,* 683–686.

Wilson, I. (2003). An ultrasound study of articulatory tongue gestures in bilingual Japanese-English children's speech. Paper presented at the 4[th] International Symposium on Bilingualism (ISB4), Arizona State University, Tempe, AZ.

Wilson, I. (2006). Articulatory settings of French and English monolingual and bilingual speakers. PhD dissertation, University of British Columbia.

Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006): The Banff Conference

edited by Mary Grantham O'Brien, Christine Shea, and John Archibald

Cascadilla Proceedings Project Somerville, MA 2006

Copyright information

ISBN 1-57473-416-4 library binding

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, e-mail: sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Wilson, Ian and Bryan Gick. 2006. Ultrasound Technology and Second Language Acquisition Research. In *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006)*, ed. Mary Grantham O'Brien, Christine Shea, and John Archibald, 148-152. Somerville, MA: Cascadilla Proceedings Project.

or:

Wilson, Ian and Bryan Gick. 2006. Ultrasound Technology and Second Language Acquisition Research. In *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006)*, ed. Mary Grantham O'Brien, Christine Shea, and John Archibald, 148-152. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #1497.