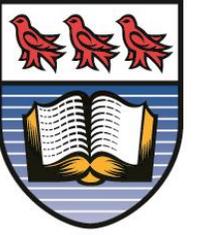


Visualizing speech in a classroom setting using interactive ultrasound imaging

Sonya Bird¹ and Heather Bliss^{1,2}
University of Victoria¹ and UBC²



**University
of Victoria**



Acknowledgments

- Ute, Paiute, Goshute, Shoshone and Navajo Nations
- Bryan Gick, Murray Schellenberg, Strang Burton, and other members of UBC's Interdisciplinary Speech Research Lab
- Workshop organizers and volunteers
- Banting post-doctoral fellowship to Heather
- SSHRC/UVic grant to Sonya
- NIH grant to Haskins Laboratories (Bryan)

Introductions

- Sonya and Heather
- Ultrasound imaging as a visualization tool for pronunciation learning and teaching - Two questions:
 - a) How can we use US to help with pronunciation teaching and learning?
 - b) How can we assess improvement, if we want to?



Benefits of ultrasound for pronunciation



Video credit: UBC Communications

Outline

1. Introduction to ultrasound imaging (20 minutes)

- Ultrasound imaging in (L2) speech research
- Affordability and accessibility
- Methodology
 - Different contexts: one-on-one, classrooms, in community
- Data processing and analysis
 - Qualitative vs. quantitative; static vs. dynamic; articulatory vs. other measurements

2. Interactive ultrasound session (30 minutes)

- French /u/ vs. /y/

3. Discussion: questions and comments (10 minutes)



Ultrasound imaging in
speech research

Ultrasound in L2 research

- Long history of US in phonetic research and speech therapy, but US has been used in L2 research only for the past decade or so
- No studies comparing US with other methods (to our knowledge)
- Studies tend to be focused on a small number of learners (10 or fewer) acquiring a small number of segments (e.g., English /ɹ l/; vowels)
- For a review of US in pronunciation, see Bliss, Abel, & Gick, to appear
- For a list of L2 studies using US, visit <http://enunciate.arts.ubc.ca/research-and-case-studies/other-research/>

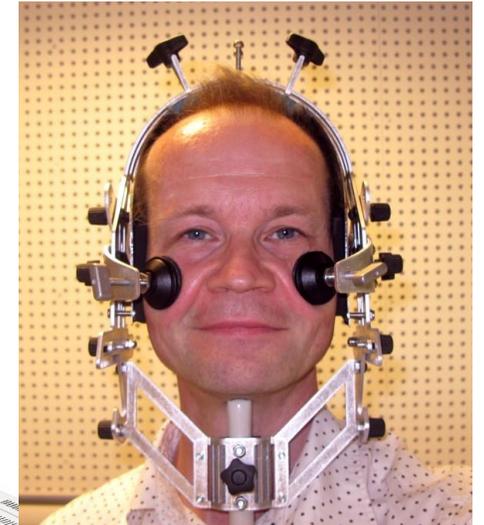
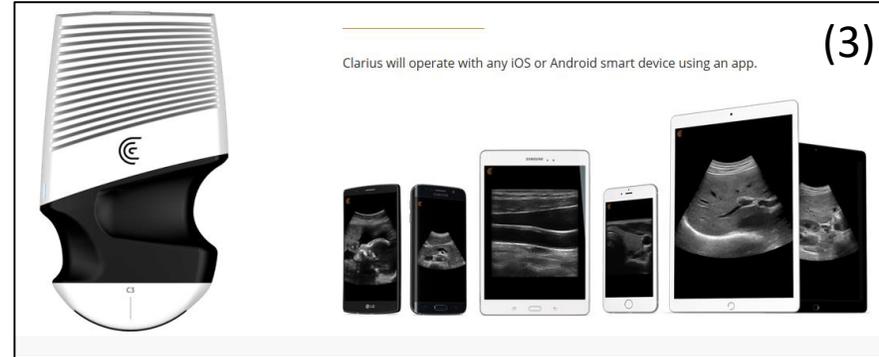
Affordability and accessibility



(1)



(2)



(5)



(4)



Technical notes on affordability/accessibility

(1) Aloka SSD-5500

- Not portable
- ~ \$150,000 CAD

(2) CHISON ECO 1 with EV probe (also with us!)

- Portable
- ~ \$25000 CAD

(3) Clarius handheld ultrasound probe

- Released in 2016
- For use with smartphone app
- ~ \$7000-10000 USD

(4) Seemore USB probe

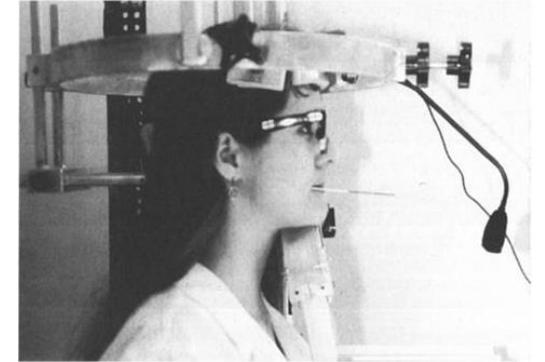
- \$4900 CAD
- includes software compatible with most windows-based machines
- downloadable to multiple computers;
- “the ultimate in portability and ideal for shared use”

(5) Articulate Instruments (Info from Scott Moisik, May 2017)

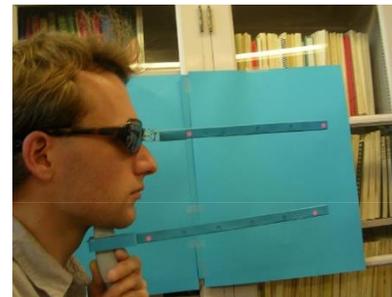
- 3 probes, a probe stabilization helmet for UTI, and Alan Wrench’s software (among other goodies for audio, etc.) for about GBP9500

Methodology: Invasiveness

- Head and the probe stabilization may be important for quantitative US analysis
- ... Stabilization need not be particularly invasive (e.g., bottom right) (Gick et al., 2005)
- ... And there are “work-arounds” (e.g., bottom left)
- ... And stabilization is not crucial for other kinds of analysis or if US is primarily for pedagogic purposes



Stone & Davis 1995



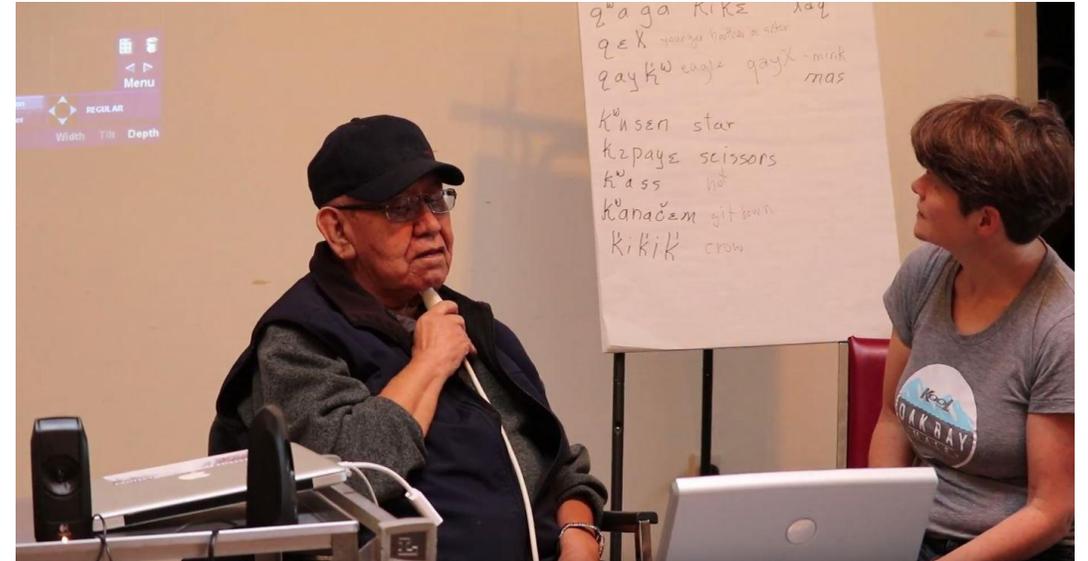
Mielke et al., 2005



Gick et al., 2012

Methodology: Different contexts

- One-on-one
- Classroom settings
- In community

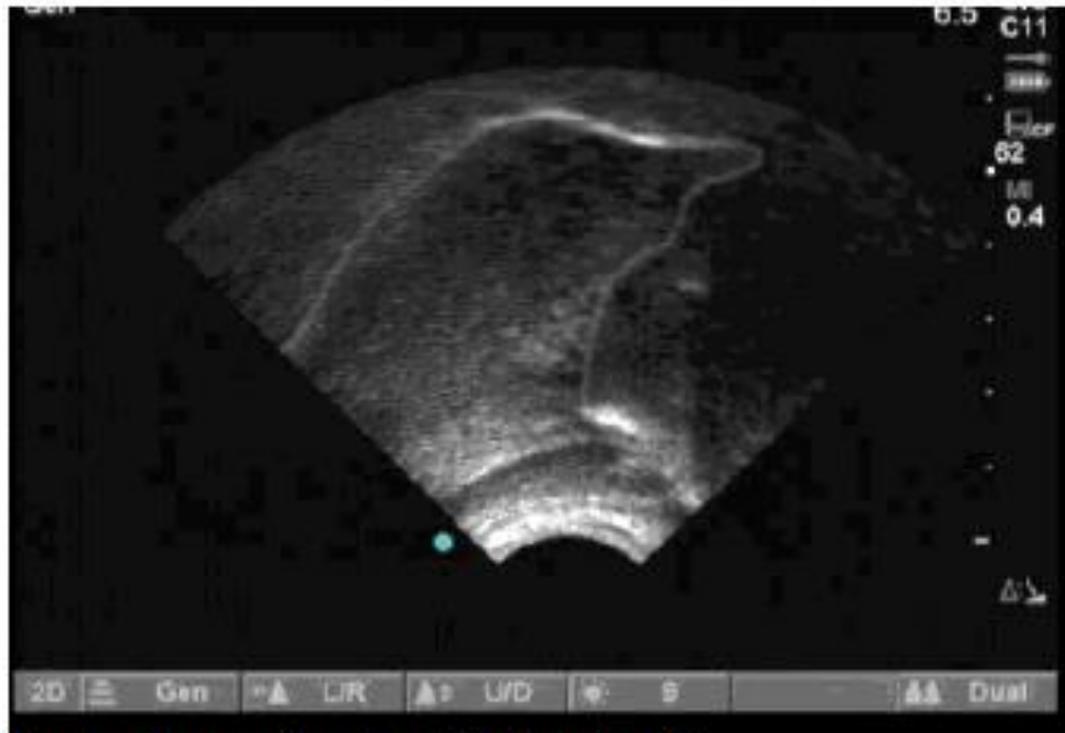


Data analysis

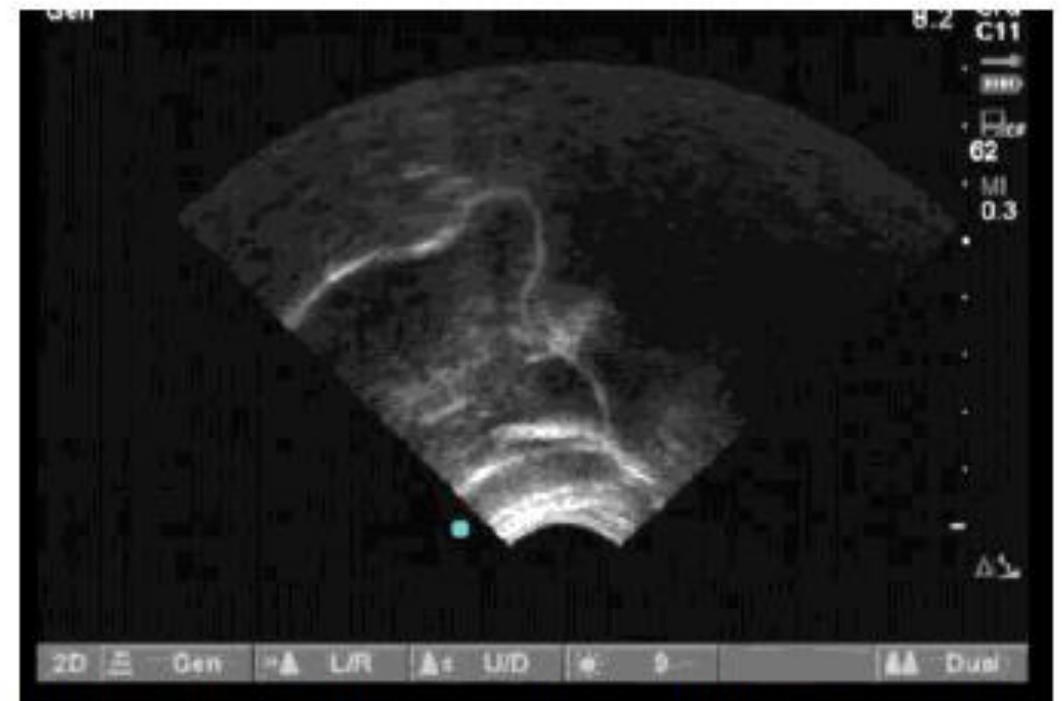
- Non US data
 - Listener judgements (auditory analysis of acoustic recording)
 - Acoustic analysis
 - Self reflections (benefit of tool to learning)
- US data
 - Generally requires capturing US video + audio onto an external machine
 - Variable process; not necessarily complex!
 - Qualitative vs. quantitative

US data: qualitative analysis

- Eyeballing tongue contours (visual analysis of US image/video)



Pre-training production of /ɪ/ in "Read"



Post-training Production of /ɪ/ in "Read" by the same speaker

Figure 8. Ultrasound images of one participant's improved /ɪ/ word

(from Tsui 2012, p. 58)

US data: quantitative analysis

- Static measurements: “magic moment” studies
- Dynamic measurements: movement between sounds
 - Can be automated

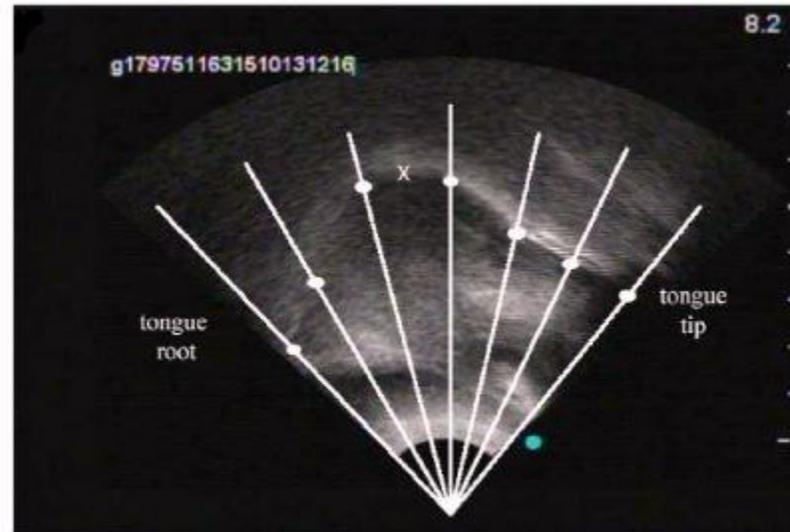
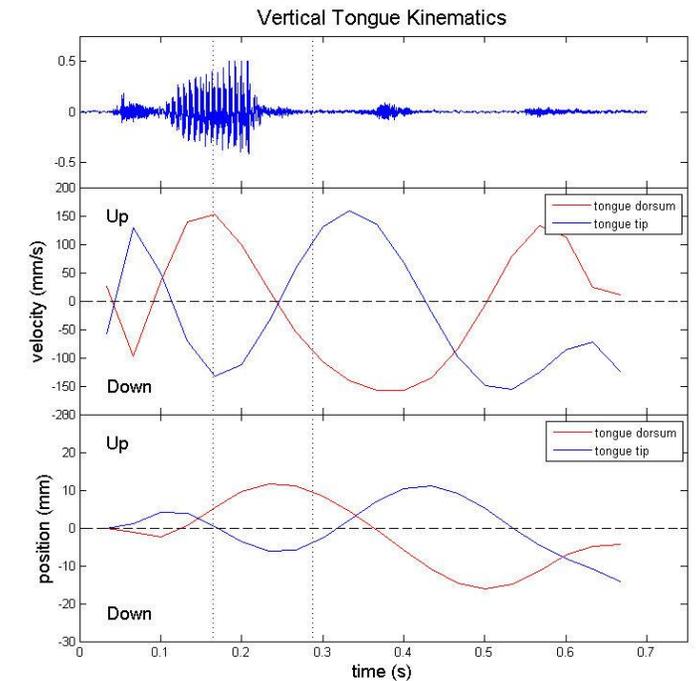


FIG. 2. (Color online) Midsagittal ultrasound image of the maximally raised position of the tongue dorsum for /q/ in “Baghdad” with a radial grid overlay. The white ellipses indicate where the radii intersect the tongue curve. The “X” indicates the location of maximum constriction along the tongue curve.

(from Davidson, 2006)



Interactive US session: French /u/ vs. /y/

- Mock mini-experiment
- Structure of the study
 - Pre-test: baseline for potential improvement + description of L2 pronunciation
 - Training
 - One-on-one US session with a fluent speaker (alternative: pre-recorded videos)
 - (Possible other conditions: auditory-only training; no training)
 - Post-test: assessment of improvement
- Methods = qualitative: self-reflection, listener judgments (alternative: quantitative: acoustics, articulation)

Pre-test

French	IPA	English
bout	[bu]	end
but	[by]	goal
pou	[pu]	(head) louse
pu	[py]	was able to

Training

French	IPA	English
bout	[bu]	end
but	[by]	goal
vous	[vu]	you (pl)
vu	[vy]	seen
nous	[nu]	we
nu	[ny]	naked
loup	[lu]	wolf
lu	[ly]	read (past)
roue	[ru]	wheel
rue	[ry]	street
voulu	[vuly]	wanted

Post-test

French	IPA	English
bout	[bu]	end
but	[by]	goal
pou	[pu]	(head) louse
pu	[py]	was able to

- Listener judgments
- Self-reflection: engagement, motivation, confidence

Discussion

- Appropriate target sounds/sequences
- Ways of adding complexity
 - Different segmental and prosodic contexts; sound sequences
 - Different ways of assessing value of speech visualization
- Other considerations

Summing up

- Ultrasound imaging has intuitive appeal to learners
- It is entirely feasible
 - Doesn't need to be invasive
 - Can be done in a variety of settings
 - Doesn't necessarily require complex articulatory data analysis
- More studies are needed to show what its benefits are
 - Comparative (US vs. other visualization methods)
 - Classroom studies (e.g., beyond single participant)
 - Local (particular articulations) vs. global (awareness and motivation)

Thank you

- sbird@uvic.ca
- hbliss@uvic.ca

- Slides:

<http://enunciate.arts.ubc.ca/research-and-case-studies/team-research/>

References

- Bird, S. (2010). An ultrasound investigation of possible covert contrasts in first language acquisition: the case of /sw ~ sp ~ sm/ > [f] mergers. *Canadian Acoustics: Proceedings of Acoustics Week in Canada*.
- Bliss, H., J. Abel, B. Gick (to appear). Computer-Assisted Visual Articulation Feedback in L2 Pronunciation Instruction: A Review. To appear in the *Journal of Second Language Pronunciation*.
- Davidson, L. (2006). [Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance](#). *Journal of the Acoustical Society of America* 120:1, 407-415.
- Gick, B., S. Bird and I. Wilson (2005) [Techniques for field application of lingual ultrasound imaging](#). *Clinical Linguistics and Phonetics* 19:6/7, 503-514.
- Gick, B., H. Bliss, K. Michelson, B. Radanov. Articulation without acoustics: "Soundless" vowels in Oneida and Blackfoot. *Journal of Phonetics* 40: 46-53.
- Mielke, J., A. Baker, D. Archangeli, S. Racy (2005). Palatron: a technique for aligning ultrasound images of the tongue and palate. *Coyote Papers: Working Papers in Linguistics, Linguist Theory at the University of Arizona* http://arizona.openrepository.com/arizona/bitstream/10150/126629/1/azu_alc_cp_n14_96_107_w.pdf
- Stone, M., and E. Davis. (1995). A head and transducer support (HATS) system for use in ultrasound imaging of the tongue during speech. *Journal of the Acoustical Society of America*, 98, 3107-3112.
- Tsui, H.M. (2012). Ultrasound speech training for Japanese adults learning English as a second language. Unpublished MSc thesis, University of British Columbia.