Visualizing speech in a classroom setting using interactive ultrasound imaging

Sonya Bird\textsuperscript{1} and Heather Bliss\textsuperscript{1,2}

\textit{University of Victoria\textsuperscript{1} and UBC\textsuperscript{2}}
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 /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
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
   • ~ $25,000 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)

Figure 8. Ultrasound images of one participant’s improved /r/ word

(from Tsui 2012, p. 58)
US data: quantitative analysis

• Static measurements: “magic moment” studies

• Dynamic measurements: movement between sounds
  • Can be automated

(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

<table>
<thead>
<tr>
<th>French</th>
<th>IPA</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>bout</td>
<td>[bu]</td>
<td>end</td>
</tr>
<tr>
<td>but</td>
<td>[by]</td>
<td>goal</td>
</tr>
<tr>
<td>pou</td>
<td>[pu]</td>
<td>(head) louse</td>
</tr>
<tr>
<td>pu</td>
<td>[py]</td>
<td>was able to</td>
</tr>
<tr>
<td>French</td>
<td>IPA</td>
<td>English</td>
</tr>
<tr>
<td>----------</td>
<td>------</td>
<td>------------</td>
</tr>
<tr>
<td>bout</td>
<td>[bu]</td>
<td>end</td>
</tr>
<tr>
<td>but</td>
<td>[by]</td>
<td>goal</td>
</tr>
<tr>
<td>vous</td>
<td>[vu]</td>
<td>you (pl)</td>
</tr>
<tr>
<td>vu</td>
<td>[vy]</td>
<td>seen</td>
</tr>
<tr>
<td>nous</td>
<td>[nu]</td>
<td>we</td>
</tr>
<tr>
<td>nu</td>
<td>[ny]</td>
<td>naked</td>
</tr>
<tr>
<td>loup</td>
<td>[lu]</td>
<td>wolf</td>
</tr>
<tr>
<td>lu</td>
<td>[ly]</td>
<td>read (past)</td>
</tr>
<tr>
<td>roue</td>
<td>[ru]</td>
<td>wheel</td>
</tr>
<tr>
<td>rue</td>
<td>[ry]</td>
<td>street</td>
</tr>
<tr>
<td>voulu</td>
<td>[vuly]</td>
<td>wanted</td>
</tr>
</tbody>
</table>
**Post-test**

<table>
<thead>
<tr>
<th>French</th>
<th>IPA</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>bout</td>
<td>[bu]</td>
<td>end</td>
</tr>
<tr>
<td>but</td>
<td>[by]</td>
<td>goal</td>
</tr>
<tr>
<td>pou</td>
<td>[pu]</td>
<td>(head) louse</td>
</tr>
<tr>
<td>pu</td>
<td>[py]</td>
<td>was able to</td>
</tr>
</tbody>
</table>

- 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


