Ultrasound Research in Linguistics

Diana Archangeli and Jeff Mielke
University of Arizona

October 8, 2005

Why to image

Acoustic analysis doesn't tell us everything about speech.
Articulatory imaging is important for:

- analyzing the articulatory parameters of speech
- studying sounds which can be articulated in more than one way (e.g. American English /r/)
- studying articulatory trade-offs such as tongue backness/lip rounding or tongue height/pharynx expansion.
- real-time feedback for second-language learners and people in speech therapy.
- ...

How to image

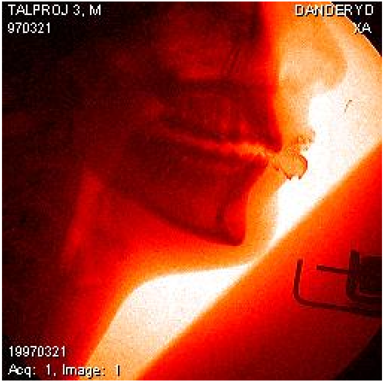
Several different technologies are available for imaging the vocal tract, such as:

- X-ray
- MRI
- EMMA
- X-ray microbeam
- ultrasound

Each has its own set of advantages and disadvantages. . .

X-ray

X-ray imaging works because x-rays are variably obstructed by different types of tissue according to their density.



X-ray pros and cons

Pros

- high temporal resolution
- locates tongue in mouth
- images entire length of vocal tract

Cons

- expensive
- not portable
- toxic

MRI (Magnetic Resonance Imaging)

MRI works because hydrogen in tissue is jiggled by a strong magnetic field.

MRI pros and cons

Pros

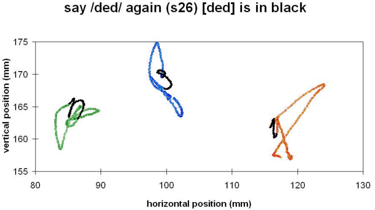
- non-toxic
- beautiful pictures
- locates tongue in mouth
- images entire length of vocal tract
- 3-dimensional imaging

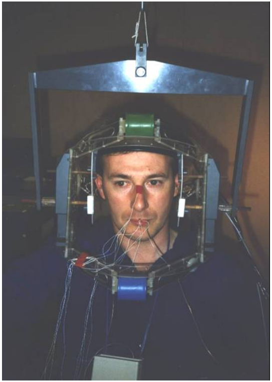
Cons

- scary for some
- subject must lie down and can't move
- very noisy
- poor temporal resolution
- expensive
- not portable

EMMA (ElectroMagnetic Midsagittal Articulometry)

EMMA works because receiver coils glued to the vocal tract report their distance from three stationary transmitter coils.





EMMA pros and cons

Pros

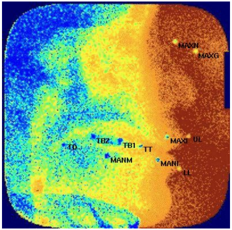
- non-toxic
- precise data points (on tongue)
- high resolution
- good temporal resolution

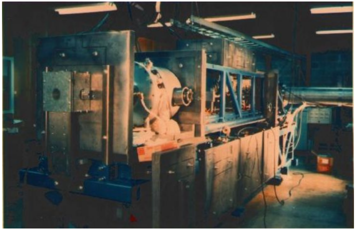
Cons

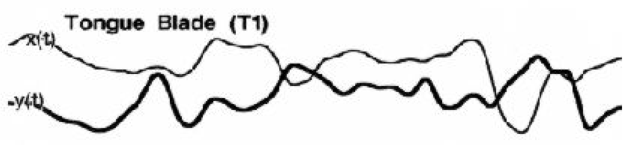
- invasive
- expensive
- not portable
- tongue tip and blade only
- must blow-dry tongue (to glue pellets)

X-ray microbeam (images from U. of Wisconsin Med. Sch.)

Uses a narrow x-ray beam to track gold pellets.







X-ray microbeam pros and cons

Pros

- less toxic than older x-ray imaging
- precise data points (on tongue)
- high resolution
- good temporal resolution

Cons

- invasive
- expensive
- not portable
- front of tongue only

Ultrasound

Ultrasound works by emitting high-frequency sound waves which are reflected back to the transducer by surfaces with sharp changes in density.

Ultrasound pros and cons

Pros

- non-toxic
- high temporal resolution (30+ fps)
- inexpensive
- portable
- subjects are comfortable

Cons

- picture can be grainy
- shows tongue only (no passive articulators)
- limited to area between thyroid cartilege and front of mandible (which cast shadows)
- difficult fixed point identification

Pros and cons highlights/summary

	safe	portable	inexpensive	high resolution	high temporal res.	passive artic.	whole vocal tract	surfaces (not points)	3D
X-ray				x	x	x	x	x	
MRI	x			x		x	x	x	x
EMA	x			x	x				
X-ray microbeam	?			x	x				
Ultrasound	x	x	x		x			x	

Pros and cons highlights/summary

	safe	portable	inexpensive	high resolution	high temporal res.	passive artic.	whole vocal tract	surfaces (not points)	3D
X-ray				x	x	x	x	x	
MRI	x			x		x	x	x	x
EMA	x			x	x				
X-ray microbeam	?			x	x				
Ultrasound	**x**	**x**	**x**		**x**			**x**	

Pros and cons highlights/summary

	safe	portable	inexpensive	high resolution	high temporal res.	passive artic.	whole vocal tract	surfaces (not points)	3D
X-ray				x	x	x	x	x	
MRI	x			x		x	x	x	x
EMA	x			x	x				
X-ray microbeam	?			x	x				
Ultrasound	**x**	**x**	**x**		**x**			**x**	

Pros and cons highlights/summary

	safe	portable	inexpensive	high resolution	high temporal res.	passive artic.	whole vocal tract	surfaces (not points)	3D
X-ray				x	x	x	x	x	
MRI	x			x		x	x	x	x
EMA	x			x	x				
X-ray microbeam	?			x	x				
Ultrasound	**x**	**x**	**x**		**x**	**?**	**?**	**x**	**?**

Ultrasound imaging

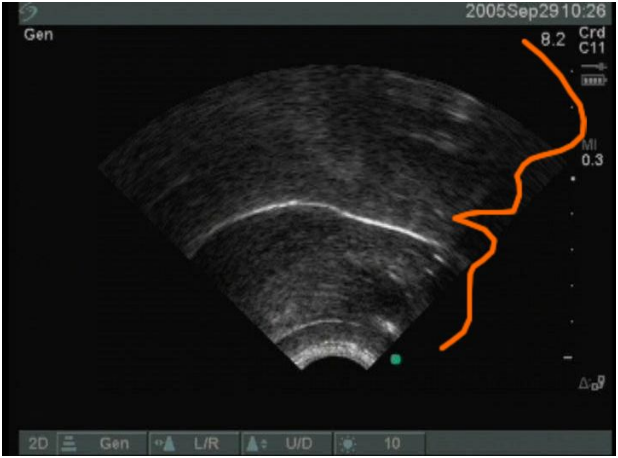
- Ultrasound imaging employs high-frequency sound waves to generate images of objects.
- Abrupt changes in density create echoes.

Ultrasound imaging of the vocal tract

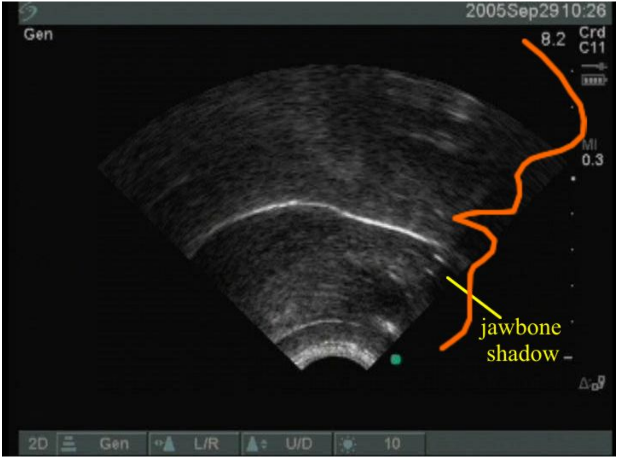
- The tongue-air interface is strongly echogenic (because of the large difference in density between air and muscle)
- An ultrasound transducer placed beneath the chin can produce a real-time movie of the full length of the tongue surface.

An ultrasound image of the tongue

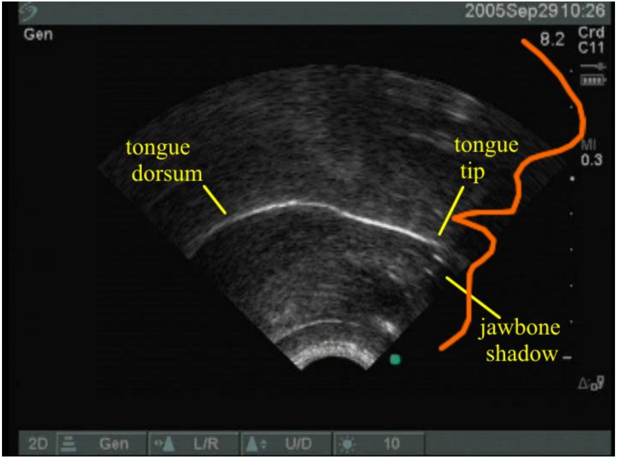
An ultrasound image of the tongue



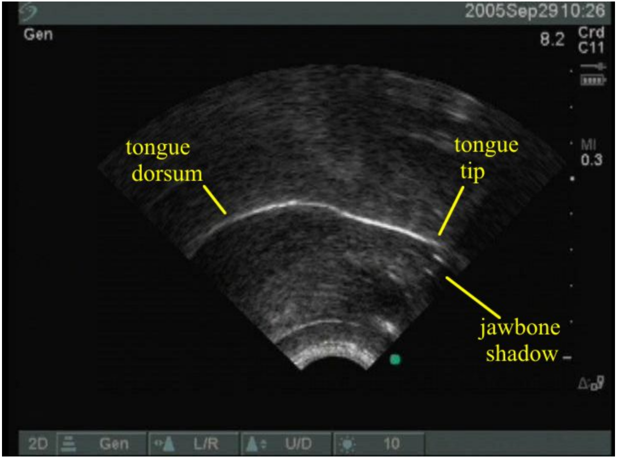
An ultrasound image of the tongue



An ultrasound image of the tongue



An ultrasound image of the tongue

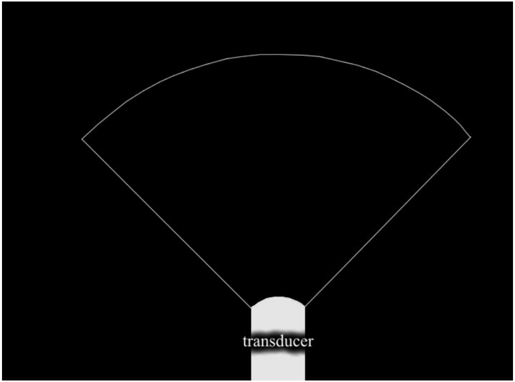


Things to look at:

- /a/ vs. /i/ vs. /u/
- tense and lax vowels
- Clear /l/ vs. dark /l/
- Retroflex /r/ vs. bunched /r/
- Grooved vs. non-grooved fricatives
- Front /k/ vs. back /k/
- /h/ before different vowels
- ...

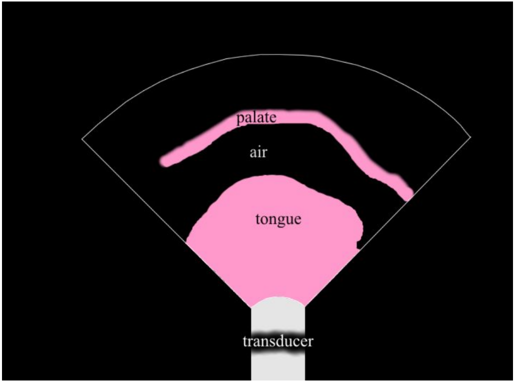
Palate obscured by air

The ultrasound transducer's field of view:



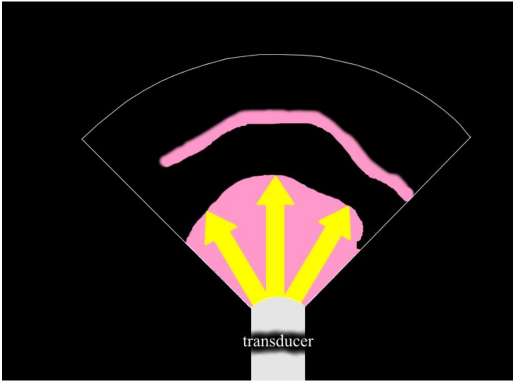
Palate obscured by air

Normally there is air between the tongue and the palate.



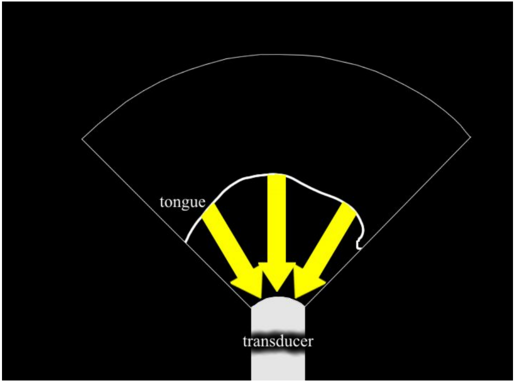
Palate obscured by air

The tongue-air interface is quite echogenic.



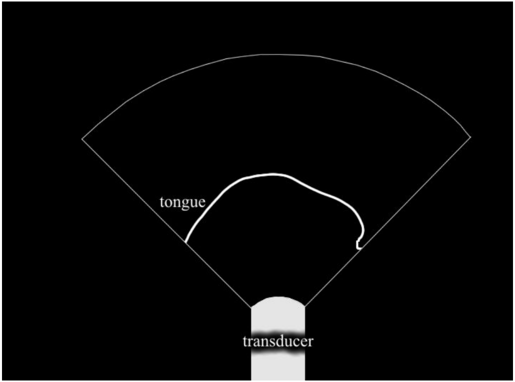
Palate obscured by air

This makes the top of the tongue image brightly.



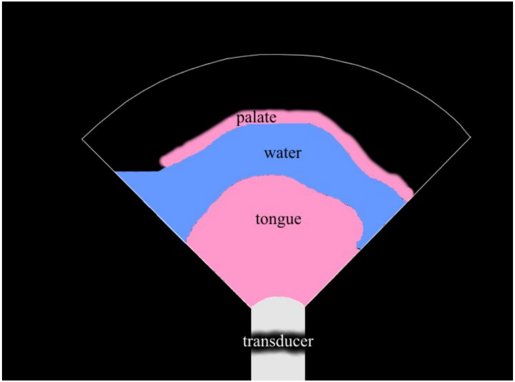
Palate obscured by air

Nothing beyond the tongue is visible.



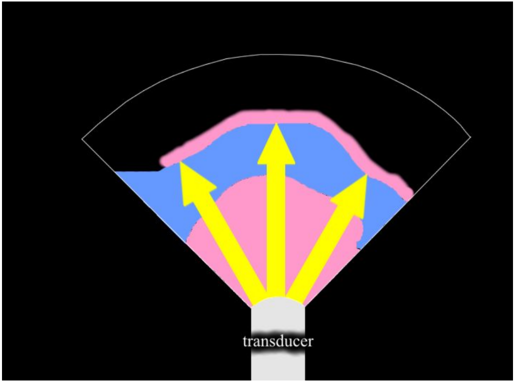
Palate visible through water

The space between the tongue and palate can be filled with water.



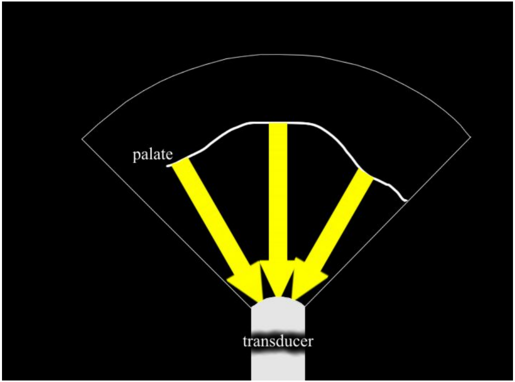
Palate visible through water

Sound waves are no longer impeded by the surface of the tongue (because tongue-water is less echogenic than tongue-air)



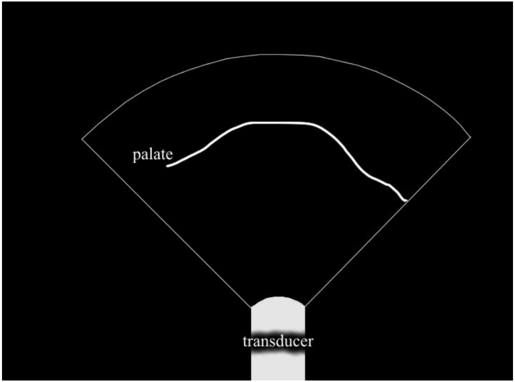
Palate visible through water

This makes the surfaces of the palate appear in the image.

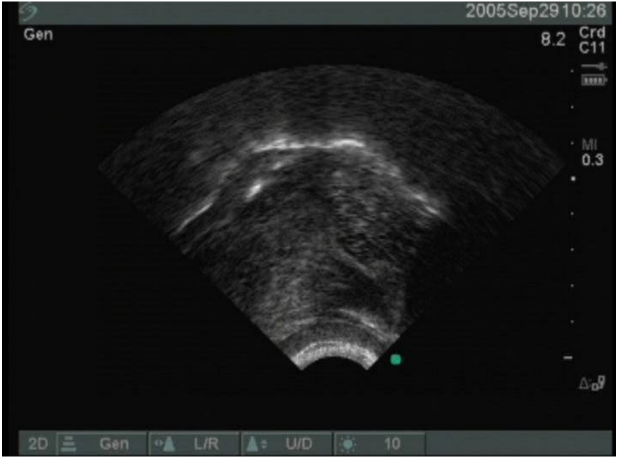


Palate visible through water

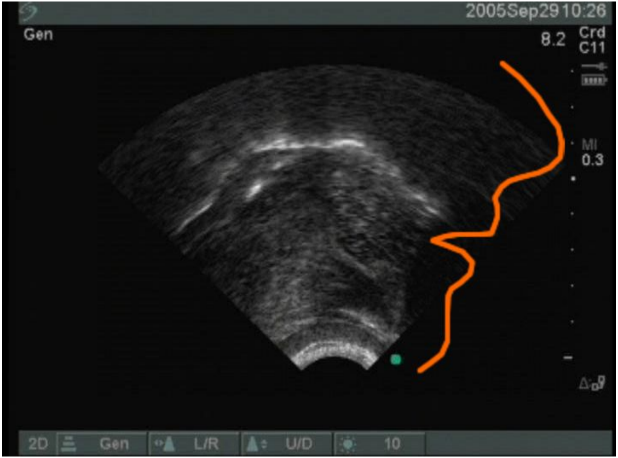
The tongue is less visible.



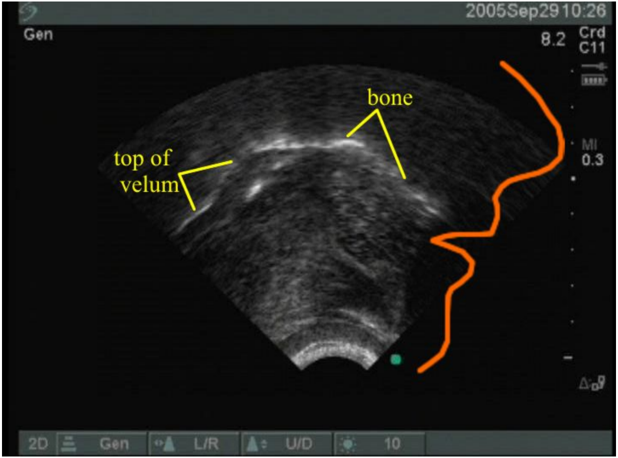
An ultrasound image of the palate



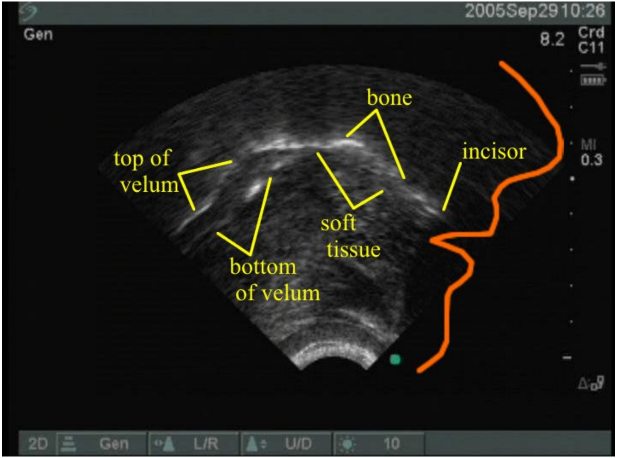
An ultrasound image of the palate



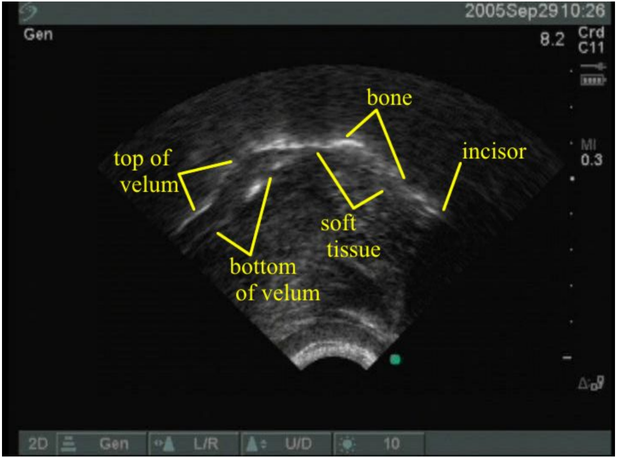
An ultrasound image of the palate



An ultrasound image of the palate

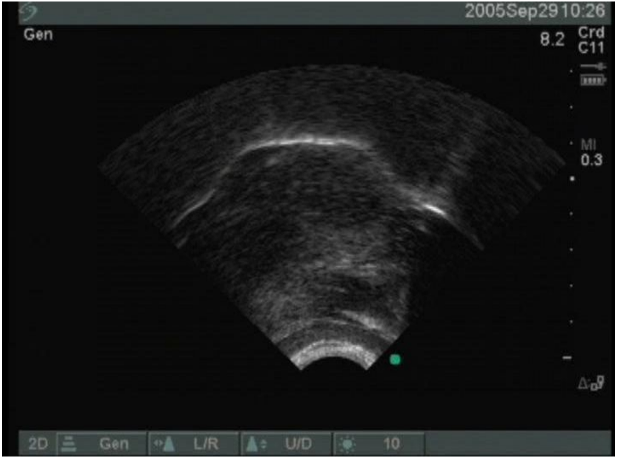


An ultrasound image of the palate

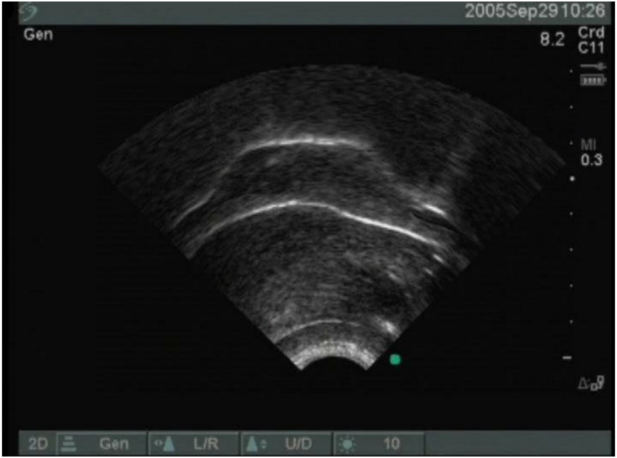


Tongue image. . .

. . . Palate image



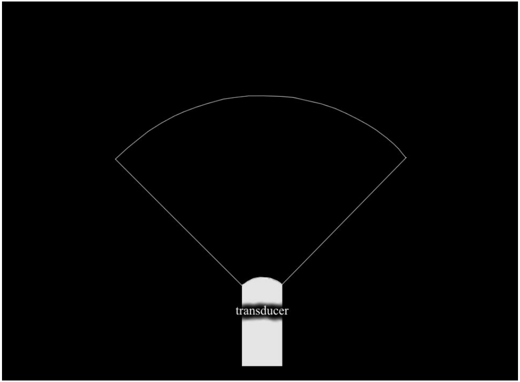
Tongue+Palate image



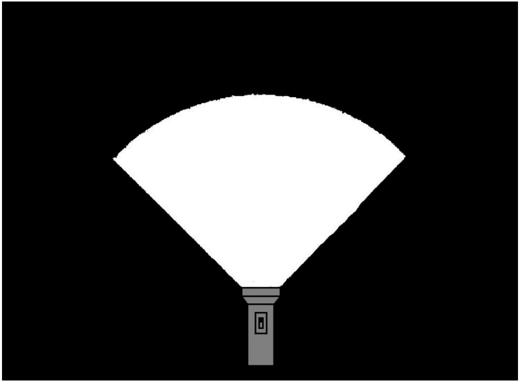
Problem...

The head and transducer can move

Flashlights

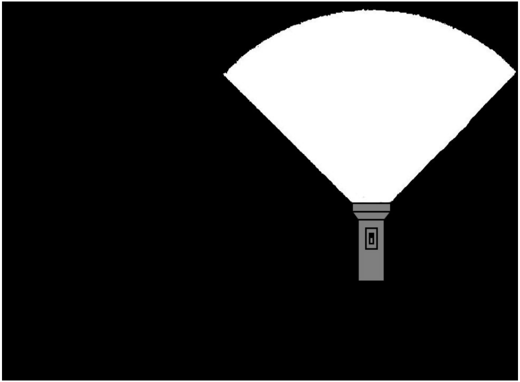


Flashlights

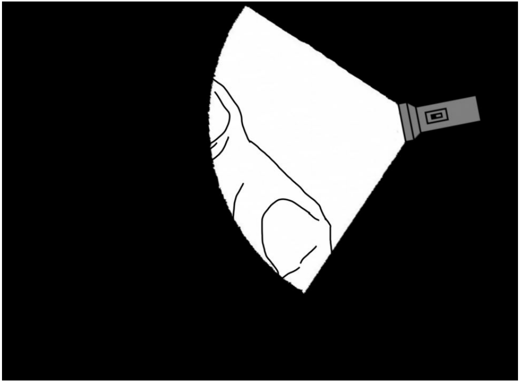


Flashlights

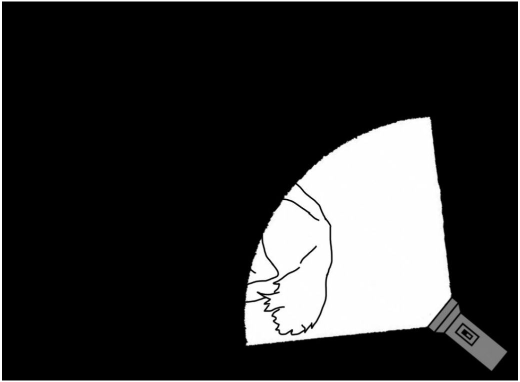
Flashlights



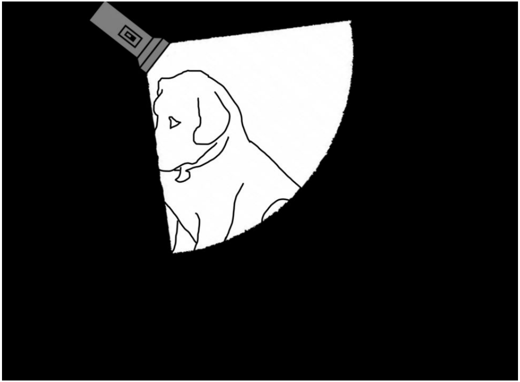
Flashlights



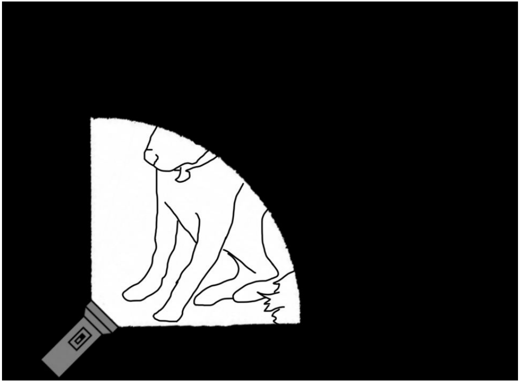
Flashlights



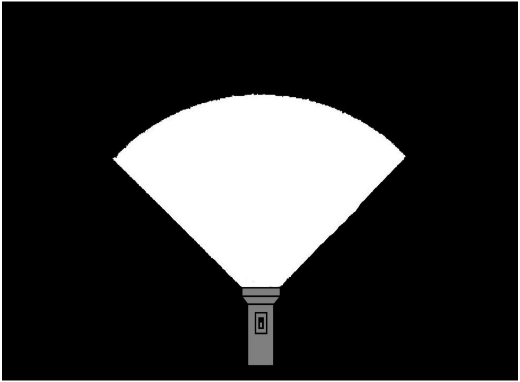
Flashlights



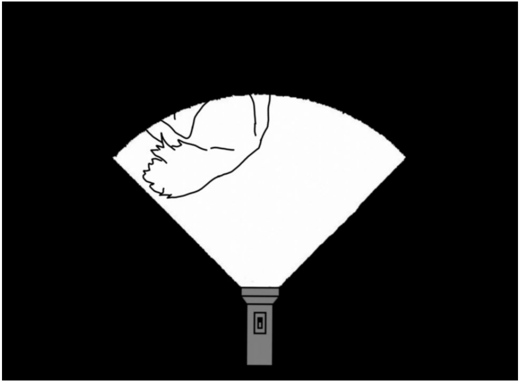
Flashlights



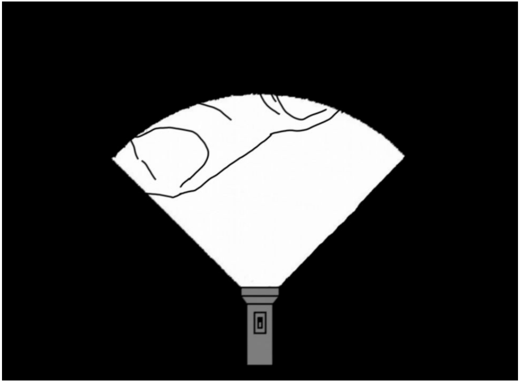
Flashlights



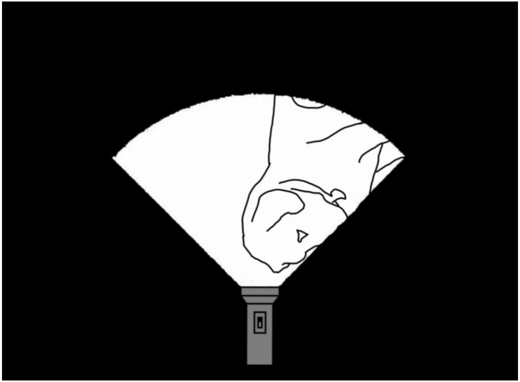
Flashlights



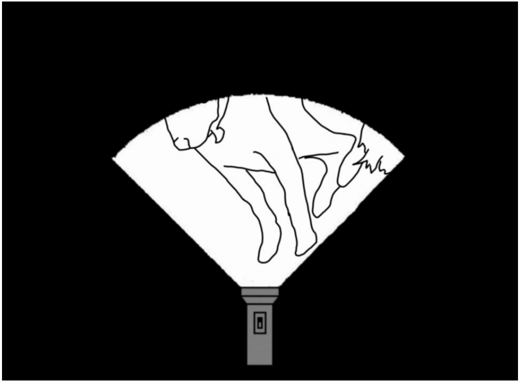
Flashlights



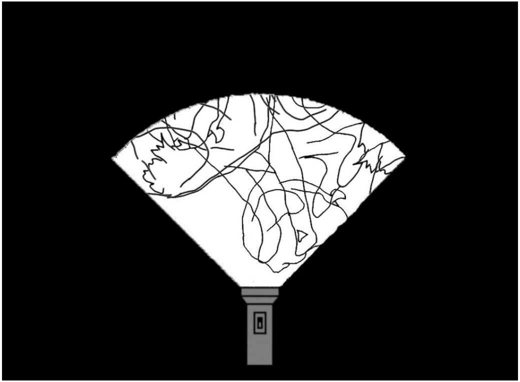
Flashlights



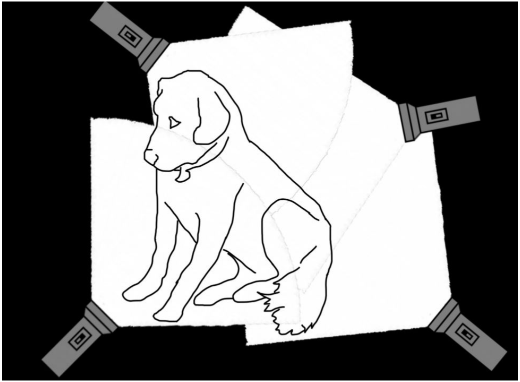
Flashlights



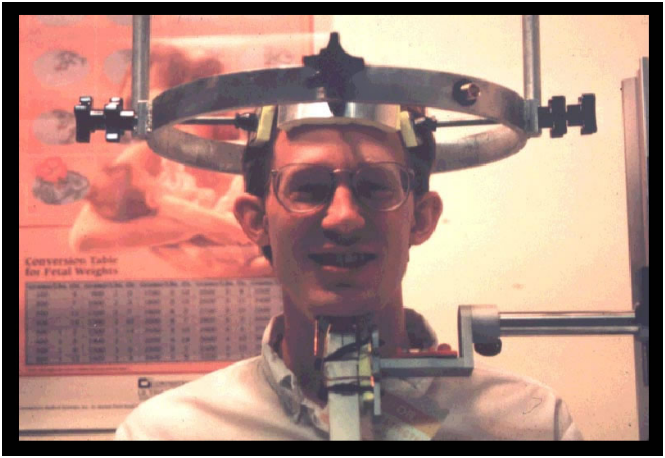
Flashlights



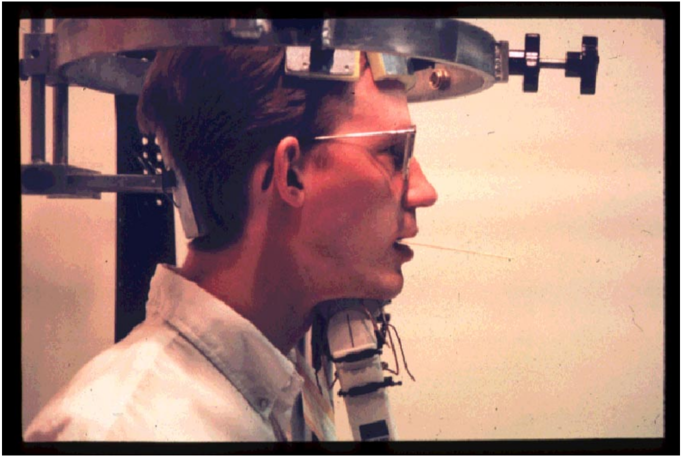
Flashlights



Head stabilization: HATS (http://speech.umaryland.edu/ahats.html)

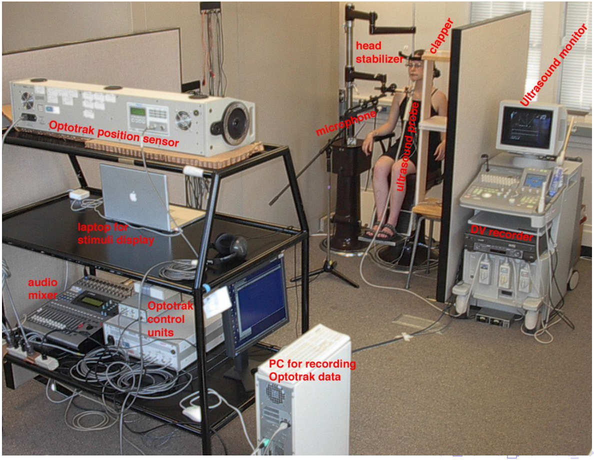


Head stabilization: HATS cont'd

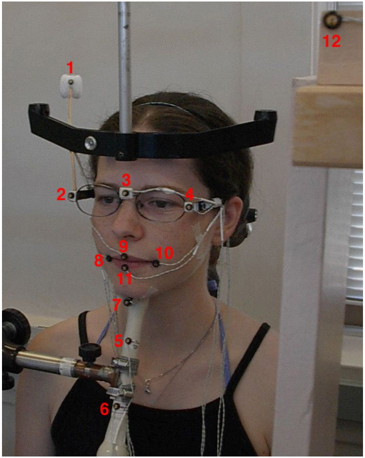


Head stabilization: UQAM (Ménard et al. 2005)

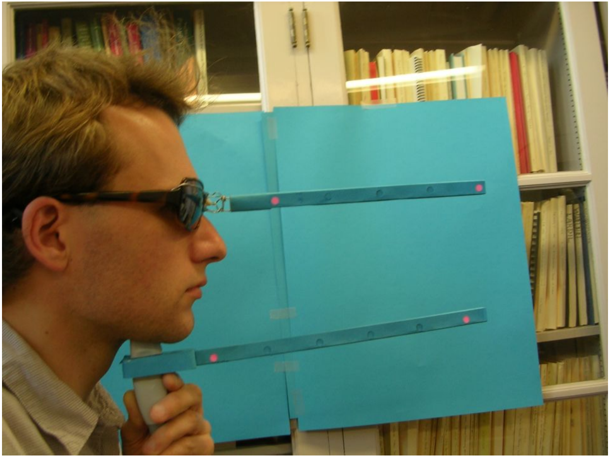
Head movement correction: UBC's optotrak (photo: Campbell et al. 2005)



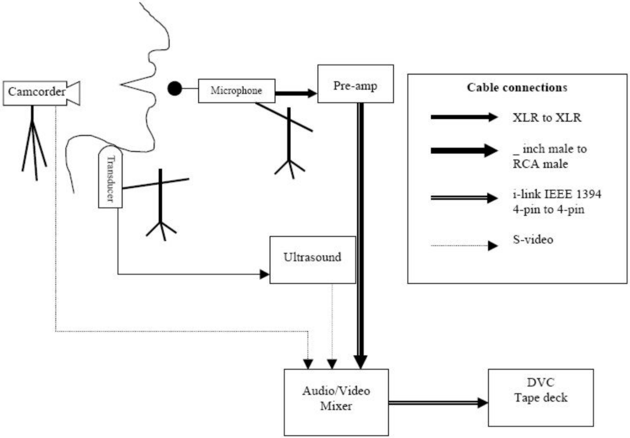
Head movement correction: UBC's optotrak cont'd



Head movement correction: Palatron



Signal paths



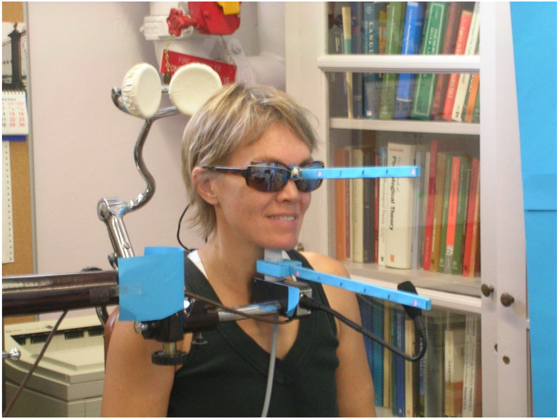
SonoSite TITAN and C-11 transducer

Camcorder and microphone preamplifier

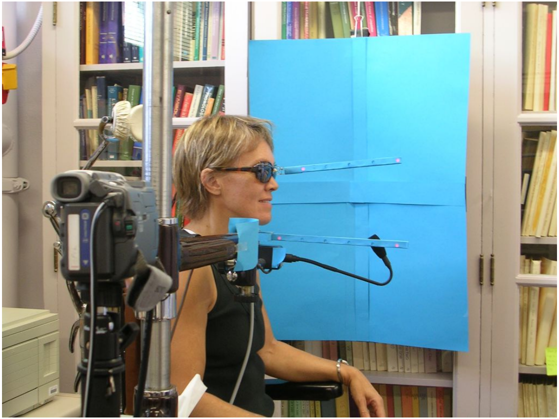
Video mixer and Digital Video recorder

APIL subject in chair

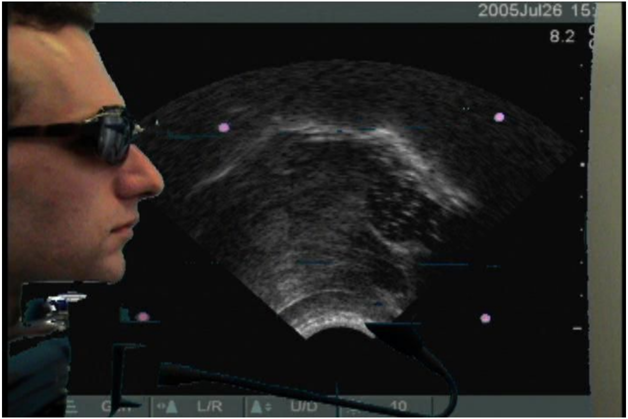
APIL subject in chair



APIL subject in chair



A palate image

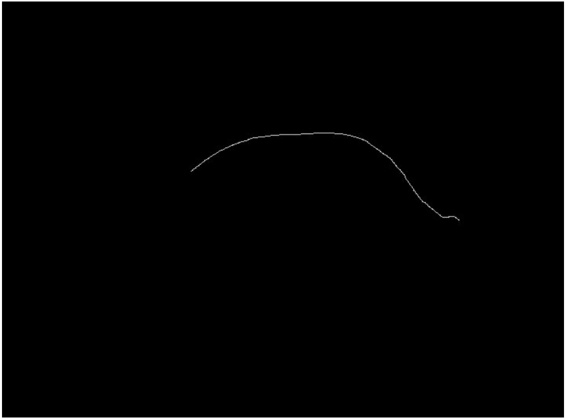


The palate image after transformation

Traced palate surfaces

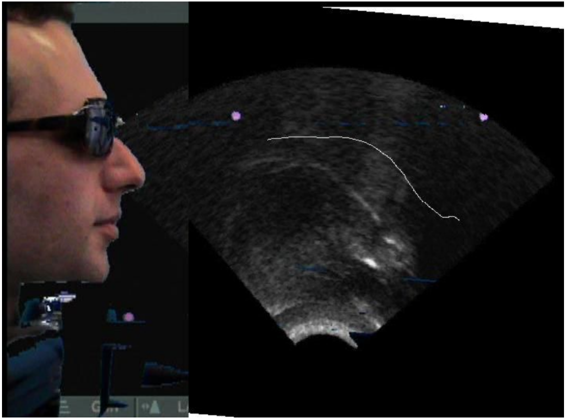
Tracings from several palate images

A composite palate tracing

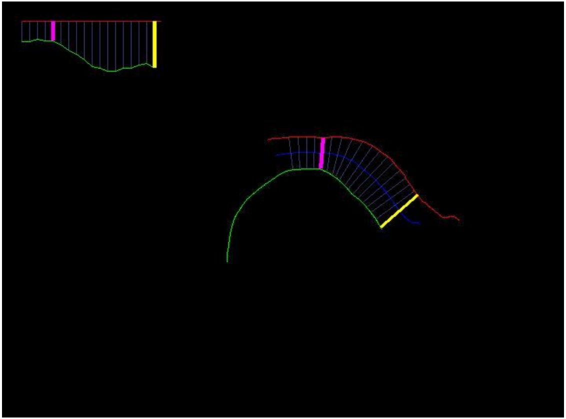


A tongue image ([u])

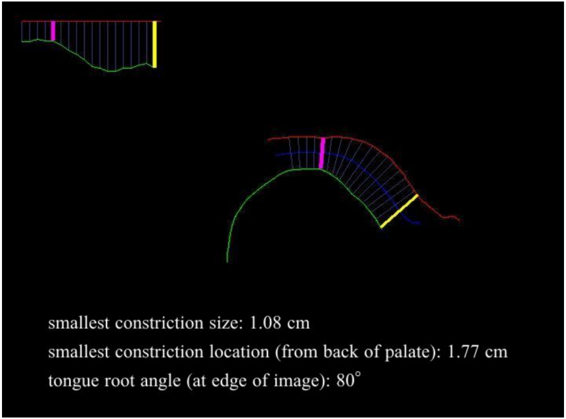
The tongue image after transformation



The tongue image after measurement



The tongue image with measurements

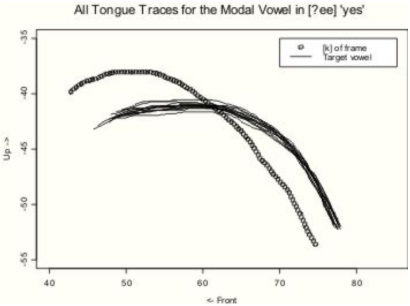


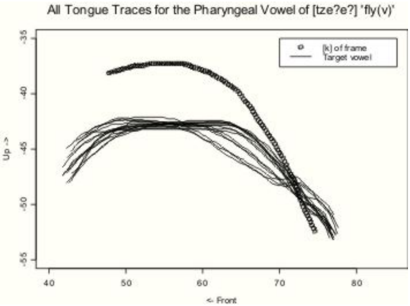
Speech applications of ultrasound

- documenting languages in the field or laboratory
- 2nd language acquisition and speech therapy
- in the clinic (e.g. for partial glossectomy patients or patients with orthodontic appliances)
- in linguistics, analyzing many aspects of speech that are not obvious from the acoustic record

Brugman (2005): Pharyngeal Vowels in N|u

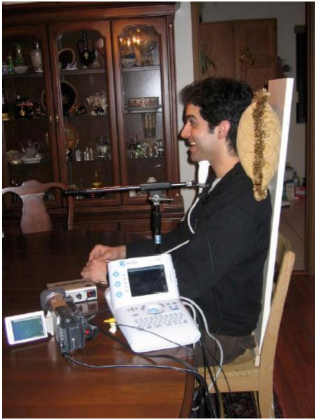
Brugman (2005): Pharyngeal Vowels in N|u



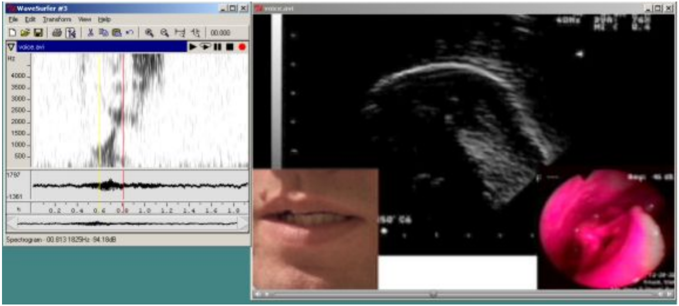


A field setup (Gick and Wilson 2005, Namdaran in prep)

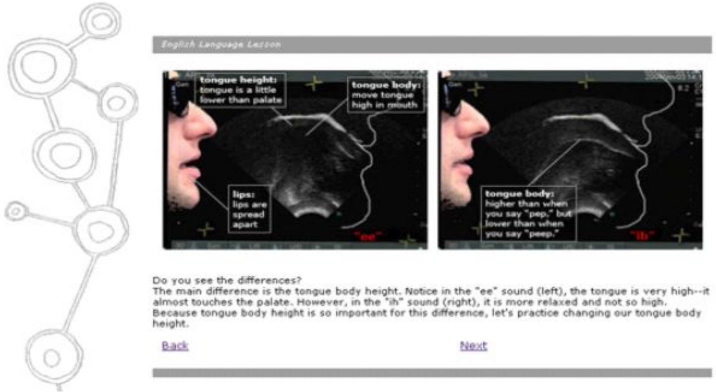
- Special chair
- SonoSite
- Desktop mic stand w/counterweighted arm
- Lapel mic
- Mini mic amp
- Sony miniDV cam



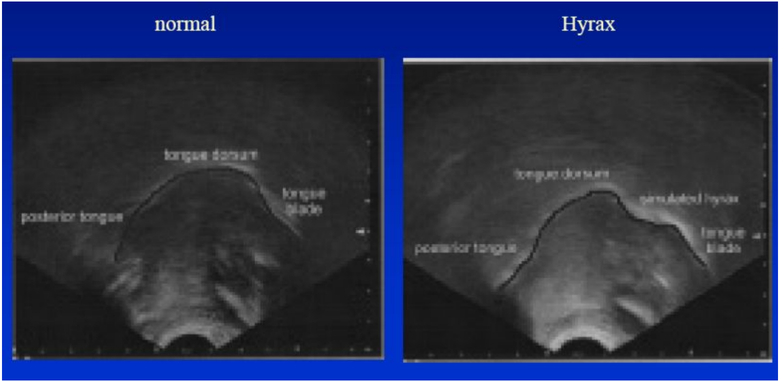
The University of South Florida audiovisual phoneme database (Frisch et al. 2005)



Ultrasound in the ESL Classroom (Meadows et al. tomorrow)

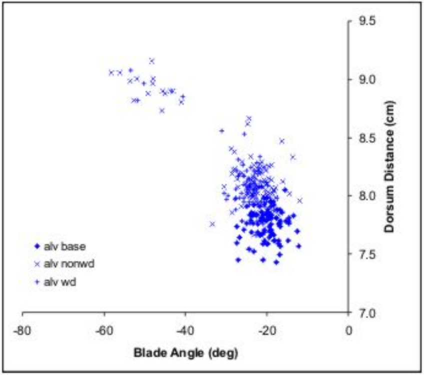


Impact of an orthodontic palatal expander appliance on tongue movement and speech (Bressmann et al. 2005)

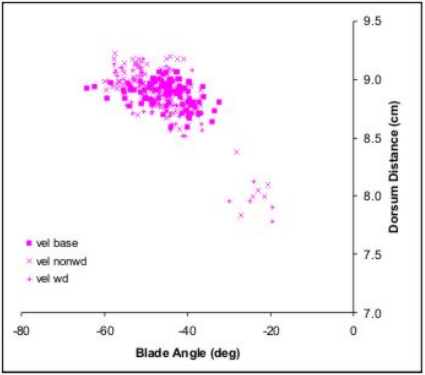


The phonetics of speech errors (Frisch et al. 2005)

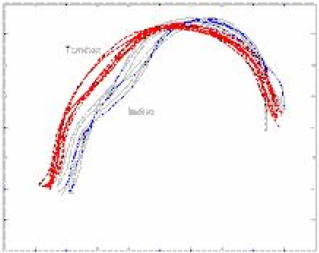
Alveolar targets

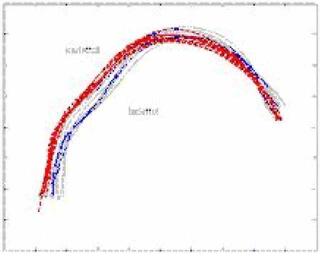


Velar targets



Transparent vowels in Hungarian (Benus 2005)

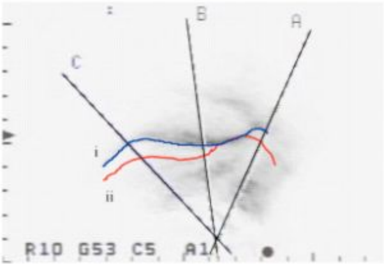




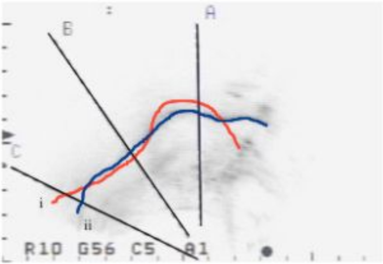
Transparent vowels in front harmony contexts are less retracted than transparent vowels in back harmony contexts.

Gestural Timing and Magnitude of English /r/ (Campbell et al. 2005)

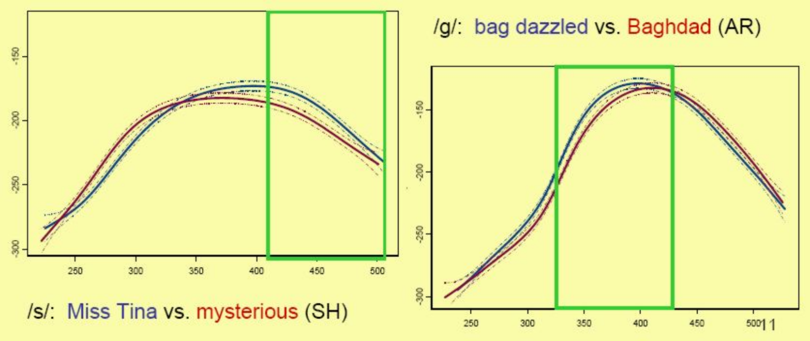
Overlaid tracings of /r/ for AGL (contrasting vowels)
(i) tip up (from /ra/), (ii) blade up (from /re/).



Overlaid tracings of /r/ for MIY (contrasting position)
(i) tip down (from /ar#ha/), (ii) tip up (from /a#ra/)

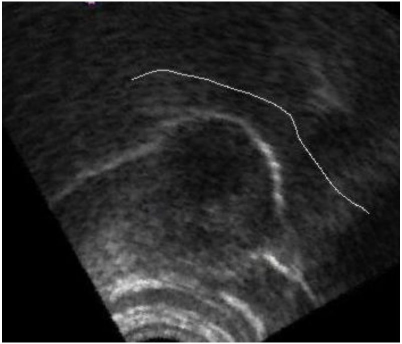


Syllabic position effects in articulation (Davidson 2005)

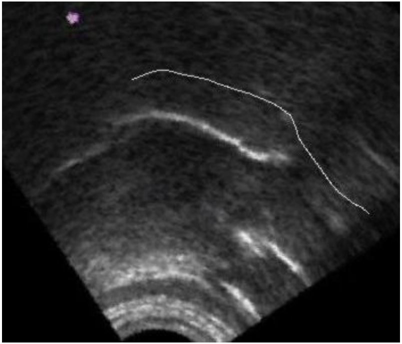


American English /r/ production strategies

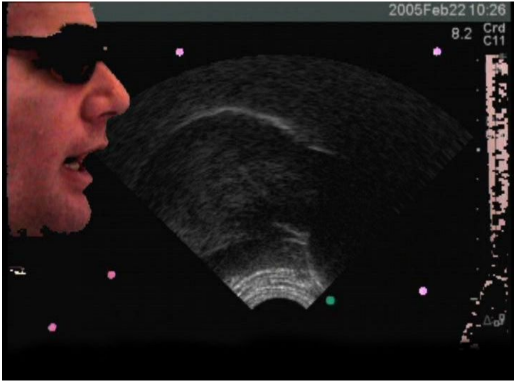
bunched: *crow*



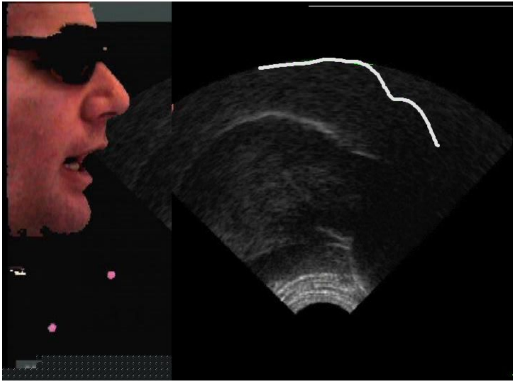
retroflex: *prose*



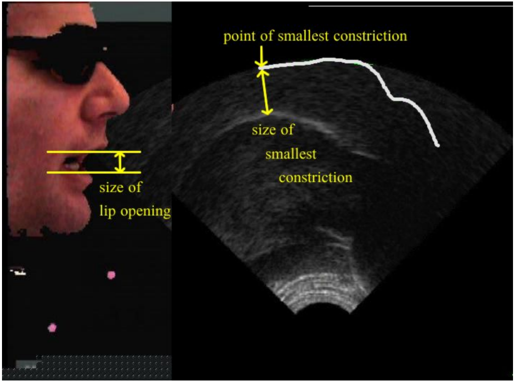
Articulatory similarity between segments: measurements



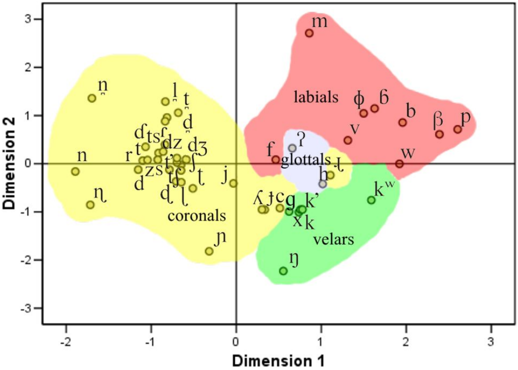
Articulatory similarity between segments: measurements



Articulatory similarity between segments: measurements



Articulatory similarity between segments: results



- University of Arizona
- University of British Columbia
- Haskins Laboratories
- University of Maryland-Baltimore
- New York University
- Université du Québec à Montréal
- Queen Margaret University College
- University of South Florida
- University of Toronto

Equipment used in an ultrasound lab

Ultrasound machine	$30,000
video camera	$1,000
video mixer	$1,000
computer	$1,000
video software	$500
imaging software	$100
video recorder	$1,000
microphone	$100
preamplifier	$100
chair and table	
incidentals such as gel, tissue, wipes	

Software used in an ultrasound lab

Applications

- Image-J (NIH)
- Adobe Photoshop or Corel PHOTO-PAINT (not open source)
- Final Cut Express (not open source)

Ultrsound-specific

- Ultra-CATS (Toronto)
- EdgeTrak (Maryland)
- Surfaces (Maryland)
- QT Extractor (Arizona)
- Palatron (Arizona)
- Glossotron (Arizona)
- Measure-o-tron (Arizona)

Time for trying out the machine. . .