Ultrasound Imaging in Speech Research

Pertti Palo

13 Feb 2020

Who's this guy?

- Pertti Palo
- I have couple of degrees in Engineering and a PhD in Phonetics.
- For the latter my main experimental method was ultrasound.
- I've also done a bit of method development on the ultrasound analysis side.

Outline for today

- First half: Basic properties and background
  - Introduction
  - Advantages
  - Limitations
  - How a session is run
- Second half: Data analysis - what can we get out of it?
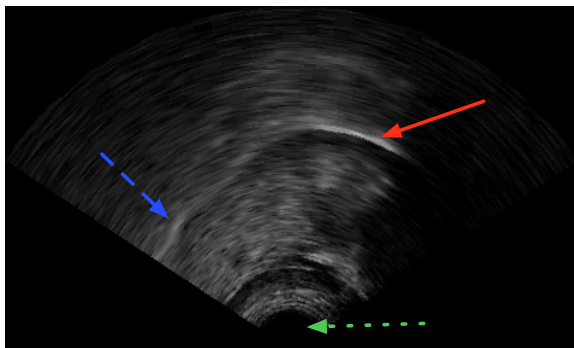 - Videos
  - Splines
  - Computational measures

First half:

Basic properties and background

Introduction: Ultrasound Tongue Imaging (UTI)

- Increasingly used for investigations of speech production.
- Clinical speech use is being studied.

Ultrasound: The physical principles

- The probe is a crucial component: Sender-receiver unit which utilises an array of piezo-electric crystals.
- Emits sound pressure waves of very high frequency (2-12 MHz, 1MHz = 1000kHz = 1 million Hz)
- The waves travel through soft tissue and reflect back when they reach a border to structures with different density (bone, air ...)

Advantages

- Very safe - no known health risks when operated at low power.
- Easy to use - after some initial setup.
- Undemanding on the participant as long as the sessions are not long.
- Relatively cheap.
- Frame rate on regular systems 20-120 Hz (and up to almost 400 Hz on exceptional systems), i.e. relatively good time resolution

Limitations I: Method

- Limited image quality
  - Time resolution is inversely related to spatial resolution.
  - Resolution also depends on probe frequency being used: Low frequencies penetrate well, but provide poorer resolution, high frequencies give better resolution but do not penetrate so deep in tissues.
  - Especially to unaccustomed eyes the pictures are far from clear.
- Limited range of structures that can be observed.
  - Generally nothing behind a bone or an airgap can be imaged.
  - Complex tissue internal features may also make it difficult to meaningfully interpret images.

Limitations II: People

- Choice of participants:
  - Very small children can not use the helmet and tend to move a lot.
  - Hand held data is possible but more difficult to analyse.
  - Big headed people are generally difficult to image.
  - A lot of fatty tissue under the chin makes imaging more difficult.
  - Age adds connective tissue inside the tongue and fuzzies the image.

How a session is run

- Clean everything. This is medical equipment and people are dirty.
- Inform your participants, do the paper work, show them how it works.
- Fit the helmet.
- Run data.
- Tell your participant they've been a great help and help them get out of the helmet.
- Also help them get the blue stuff off of themselves.
- Clean everything.

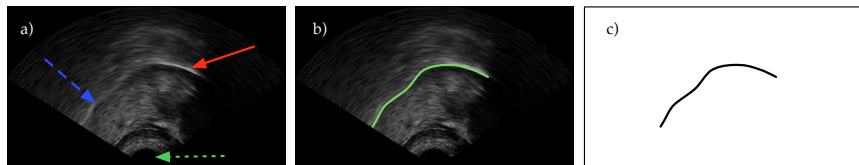Let's play!

Who wants blue stuff on their chin?

Second half:

Data analysis – what can we get out
of it?

Just the videos

- Ultrasound videos can be annotated like any other videos.
- Use either AAA or export the videos and use a video annotation program.
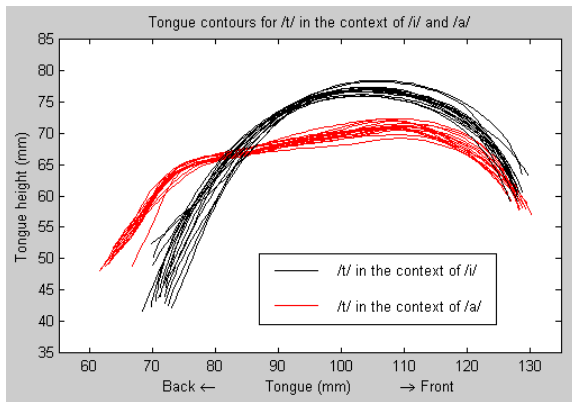- It is also possible to do direct measurements on the images.

Splines I

- Splining means we draw a contour on the image and extract it.
- Can be done either manually or automatically.
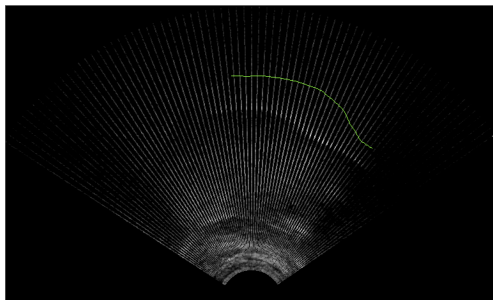- Results depend a lot on the quality of the data.

Splines II

Splines: the results

Tongue contours for /t/ in the context of /i/ and /a/
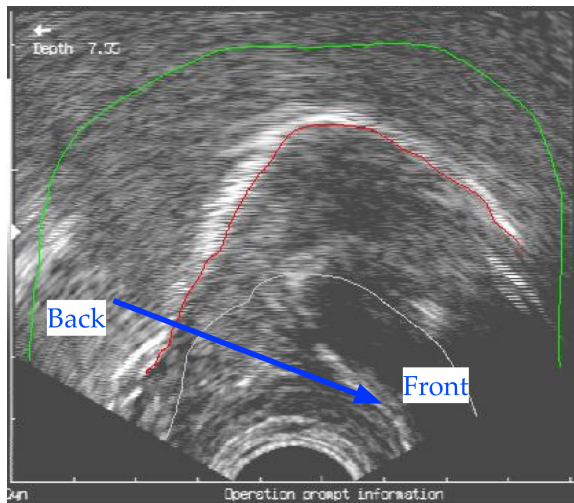
- ▶ Be traditional - get point measures in time from two or more contexts and compare them.
- ▶ Be modern - get movement sequences splined and analyse the data with functional data analysis.

Computational measures

- We can also analyse the information content of the videos without extracting features such as the tongue contour. For example:
  - Pixel difference tracks over all change and is based on using raw ultrasound data.
  - Optic flow tracks movement by statistically matching local features in the videos.
- These methods do not replace splining, but rather provide options.

Raw data

a)

First scanline

Last scanline

b)

412 px

38 px
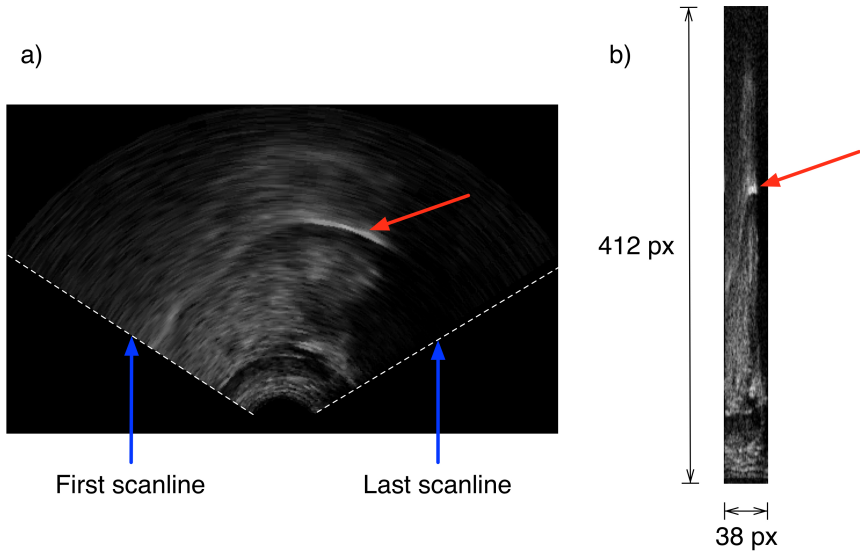
That's it!

Thank you!

And thanks to Felix Schaeffler for some of the images and text in these slides, and to Steve Cowen for the picture of me.