Data Acquisition for Sociophonetics

Pertti Palo

10 Nov 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 for ultrasound analysis and audio recording during MRI.
- I've used a bunch of different speech measurement methods and crunched data from even more.

Outline for today

- First half: Readily (locally) available methods
  - Introduction
  - Choosing a method
  - Audio
  - Video
  - Ultrasound
- Second half: More methods
  - Electromagnetic Articulography (EMA)
  - Fast Magnetic Resonance Imaging (fast MRI or real-time MRI)
  - Electropalatography (EPG)
  - Electroglottography (EGG)
  - Breathing: Air flow and other methods
  - Other methods: Motion capture, Eye tracking, Electromyography (EMG), Computed tomography (CT), cineradiography, analysis-by-synthesis, and others

First half:

Readily (locally) available methods

Introduction

- We'll concentrate on the measurement of vocal speech and its production.
- We should note that speech, which relies on sounds produced by the vocal tract, is not the only mode of human spoken language.
- If somebody is interested in sociophonetics of sign language, do tell.
- Most of the methods we will talk about are just run-of-the-mill phonetics measurement methods.
- The lists are necessarily incomplete, so you may have to (should) do some thinking and searching before selecting your method.
- Some of these methods are discussed in my Master's thesis.

Palo, P. (2006). A Review of Articulatory Speech Synthesis. Master's thesis, TKK, Helsinki.

Selecting a method

- What's available and do you have control over this?
- What's your research question and which method would suite it best?
- Does the method you need exist?
- What do colleagues and bosses say?

Audio I

- Audio is almost always a possibility nowadays. The main exception would be noisy conditions and conditions where the environment prevents use of the available hardware.
- You should practically always record audio.
- An anechoic chamber provides the best quality, a sound proofed studio is not bad, a quite room is not a catastrophe, but you may have to do the recordings in less controlled conditions and/or with less ideal equipment.

Audio II: Unidealities

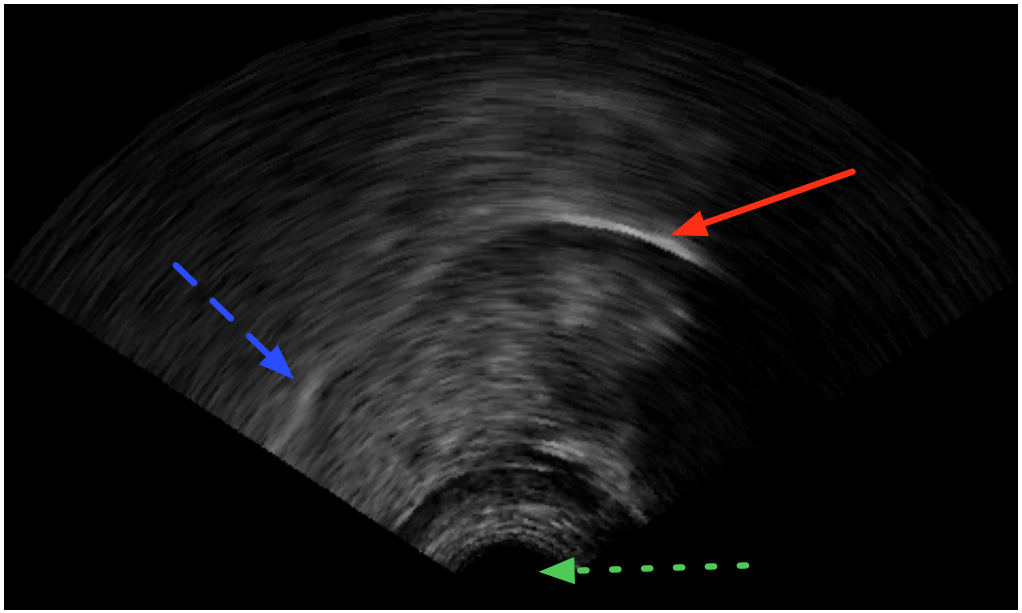
- Participants using their own hardware for recording – possibly over the interwebs or on their phones.
- Recordings are archival material and obviously can not be amended.
- The environment is part of the study and it is loud: MRI, ambulance, construction site.
- And in MRI, it is not a good idea to use ferro-magnetic parts, so a wooden sound collector might be the way to go.

Video

- Video is almost always a possibility nowadays, but is not necessarily of great quality.
- All cameras are not equal, so if possible get a good one. There is a wealth of detail in the choice.
- Thanks to webcams and phone cameras, we can also record participants without having to meet them in meat space.
- But video only gets us so far. It's good for recording facial movements including lips, but that is pretty much it.
- If we want to get data on the tongue or the larynx we need something else.

Ultrasound

- Ultrasound has been used for speech research for nearly 50 years.
- Not all ultrasound devices are suitable for speech research.
- Can be used for imaging tongue and larynx movements.
- We'll talk about ultrasound in much more detail on Thursday.



Second half:

More methods

Electromagnetic Articulography (EMA)

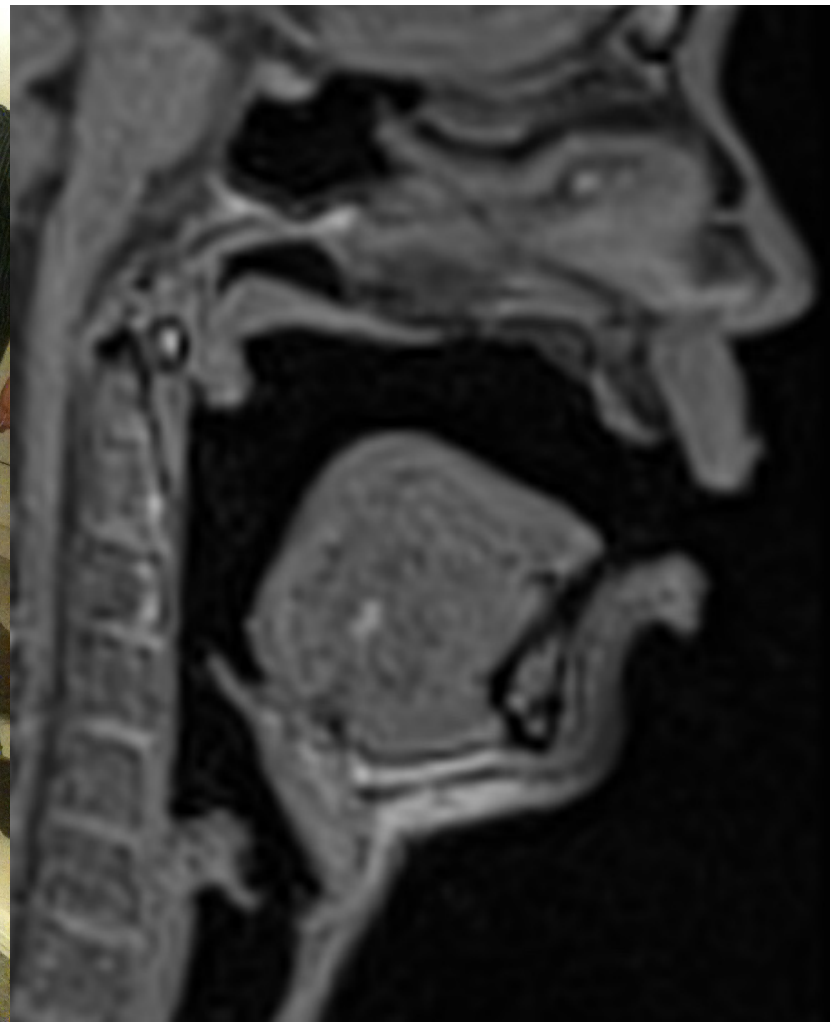
- This is a point tracking method.
- The points are small(ish) coils with wires coming out of them.
- Can be a bit strenuous for the participants, but does have the potential for long recording sessions.
- In old devices are limited to tracking the points in 2D, but modern devices track them in 5D: three spatial dimensions and two rotations.
- EMA has potentially very high temporal resolution (up to 1250 Hz) and good spatial resolution and is silent.

Fast MRI or real-time MRI

- Again the equipment sets the limitations.
- Provides potentially good detail for soft tissues, but does not show bone – like teeth – at all.
- Spatial resolution is usually good, but temporal resolution usually less so.
- Usually quite expensive to use and the recording conditions are far from ideal for studying speech.

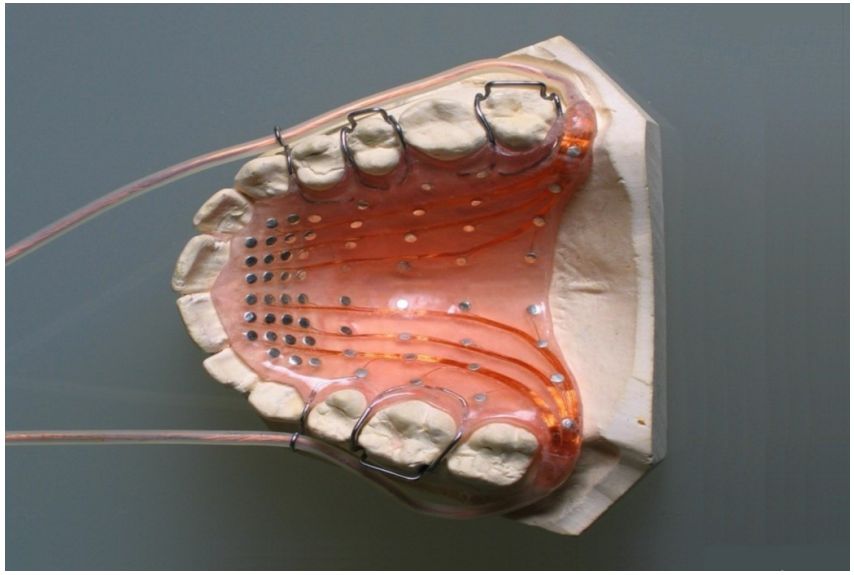
Fast MRI II

Excellent method, if we want to relate speech production to anatomical features.



Electropalatography (EPG)

- EPG measures (tongue-)palate contact at very high frequencies.
- Needs a custom fitted individual artificial palate.
- Placement of sensors on the palate provides (a kind of) normalisation across participants.
- Optical Palatography (OPG) has been under development for a long time with some recent activity giving hope that it might become available. It would provide contact measurements and distance measurements.



Electroglottography (EGG)

- EGG is a form of electrical impedance tomography.
- It provides a way to study the opening and closing of the glottis, but it should be noted that the EGG signal is *not* a direct measurement of the opening and closing of the glottis.
- Besides the more usual 1-channel EGG there are also multi-channel devices which – with proper interpretation – provide a means to evaluate when a change in the EGG signal(s) is due to glottal aperture changes and when it is, for example, due to up-down movement of the larynx.
- EGG is especially useful when we have questions about the larynx in general, changes in voice quality, and phonation.

Breathing: Air flow and other methods

- The picture shows a simultaneous air flow and ultrasound measurement.
- Air flow in speech and breathing can be recorded directly from oral flow and nasal flow.
- With a suitable device, breathing activity can also be inferred from the expanding and contracting of the chest.

Other methods I

- 3D ultrasound. Ultrasound but not just one 2D plane of imaging.
- Motion capture and other (visual) point tracking systems – not that great for vocal speech, but potentially very useful for studying sign language.
- Eye tracking can provide attention context when used with other methods.
- Electromyography (EMG) measures electric potential changes caused by nerve activation in muscles. It is not that widely available. Surface electrodes usually provide less than ideal signal and needle electrodes are just scary.

Other methods I

- Computed Tomography is silent and provides excellent imaging of both soft and hard tissues. The downside is that X-rays slightly kill you with cancer.
- Cineradiography is even worse: X-ray videos have for a long time been considered a bad idea.
- Analysis-by-synthesis – an example of doing Something Else. Record data in one modality (with one method) and extrapolate into non-imaged structures by matching an articulatory synthesiser to the data. Very difficult and very rarely done.
- There are others.

That's it!

Thank you!

And thanks to Steve Cowen for the silly and serious pictures of me, Jarmo Malinen for the one in the anechoic chamber, and to Alan Wrench for the one of himself with a modern ultrasound helmet as well as the EPG palate.