Modelling of speech for oral and maxillofacial surgery (COMSPEECH)

Docent Jarmo Malinen, DSc (tech.)

Aalto University,
Dept. Mathematics and Systems Analysis

May 20th, 2016

Goals and current stage

The goals of the project are:

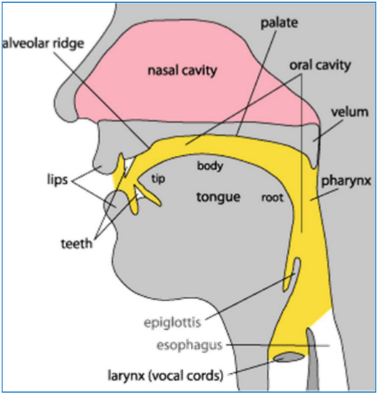
- To develope a method that can be utilized in predicting the changes in the speech production caused by the treatments affecting the anatomy of the vocal tract:
  - surgery of oral cancer,
  - reconstructive surgery,
  - orthognathic surgery,
  - prosthodontic rehabilitation of jaw defects, etc.
- To clarify anatomical and functional causes of speech disorders.
- To contribute to the speech therapy.
- To advance the basic knowledge of phonetic phenomena and speech production; in particular, the production of vowels.

The current stage of the project is in improvement and validation of the accuracy of our mathematical model "DICO" for speech production in patients undergoing orthognathic surgery.

Production of speech

Original speech

"Speech" without vocal tract



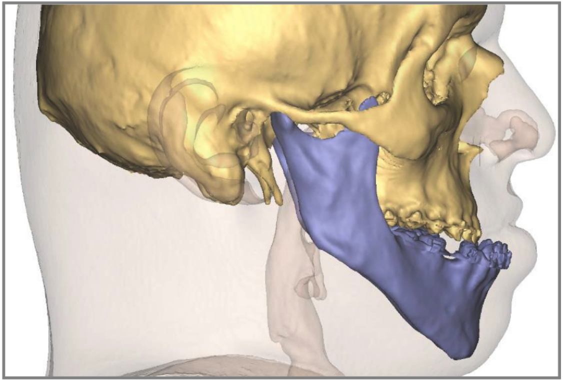
Picture: http://www.indiana.edu/~hlw/PhonUnits/vowels.htm

Vocal folds imaged at 2 kHz

Effects of maxillofacial surgery on speech

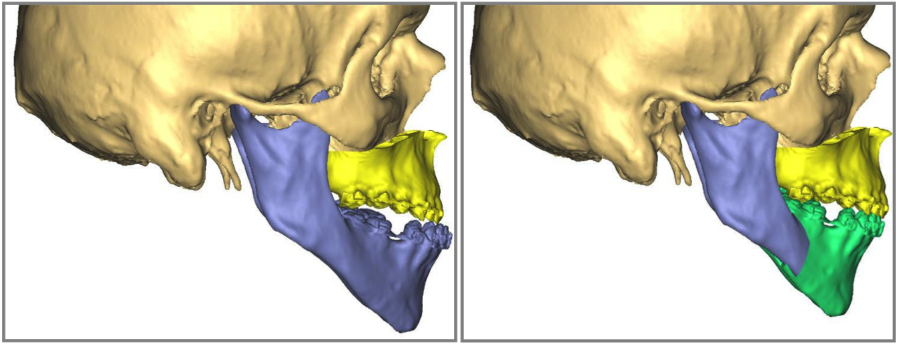
Correction of malocclusions (1)

Severe malocclusions can be treated by a combination of orthodontic and surgical therapy, i.e., by orthognathic surgery.



Correction of malocclusions (2)

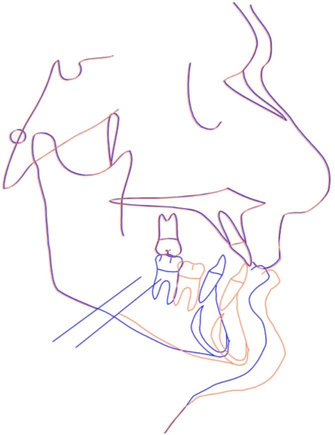
The position of the maxilla and the mandible is changed surgically in relation to the skull base.

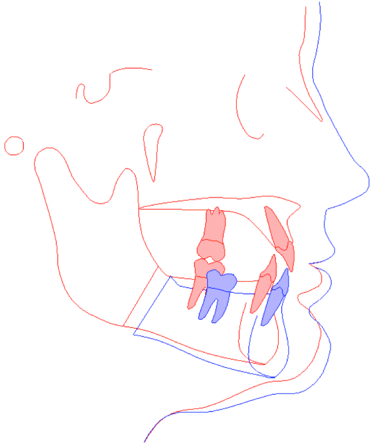


Orthodontic therapy is always an essential part of the treatment of these patients.

Correction of malocclusions (3)

Patients requiring only mandibular surgery are recruited in our validation study.





Cephalometric analyses for the advancement of the mandible.

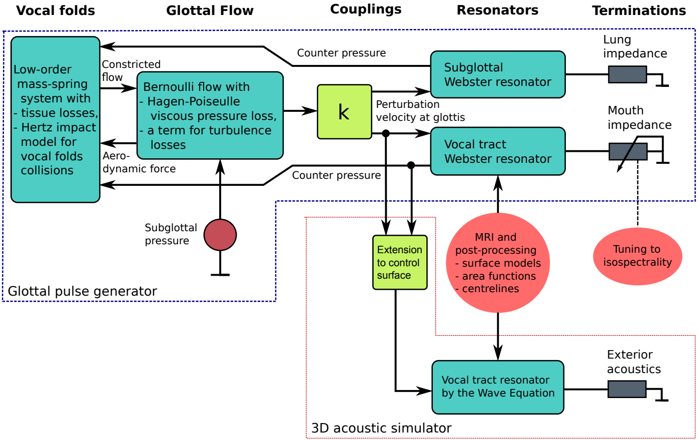
Mathematical modelling of speech acoustics

DICO -vowel production model in nutshell:

- The mathematical modelling of speech based on vocal tract 3D anatomy, obtained by Magnetic Resonance Imaging.
- Numerical simulation of speech sounds based on MRI.
- Fine-tuning and validation is based on comparison between simulated speech and speech recorded during the MRI examination.
- Model experiments using 3D prints of vocal tracts.
- The total error in modelling and data acquisition must be significantly smaller than the effect of surgery on speech.

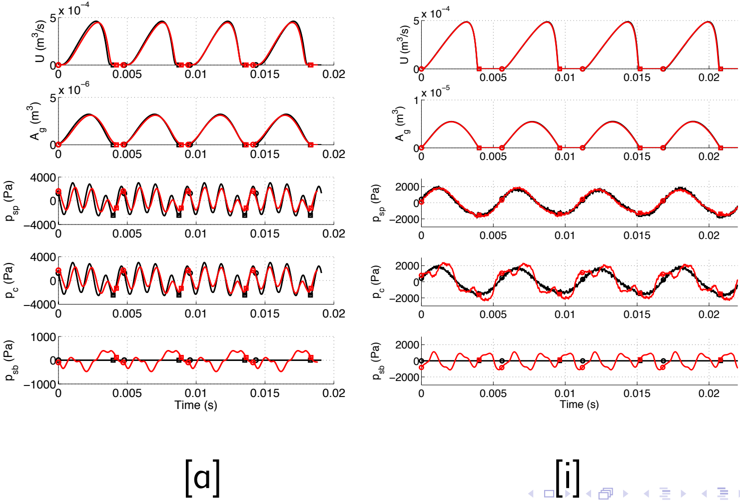
DICO (2)

In addition to the vocal tract, we need to model vocal cords, lower airways, and even the exterior space so as to simulate vowels.



DICO is an interconnected network of mechanical, flow-mechanical and acoustic partial models.

Simulated glottal opening areas and sound pressures at larynx during production of the Finnish vowels [ɑ] and [i].



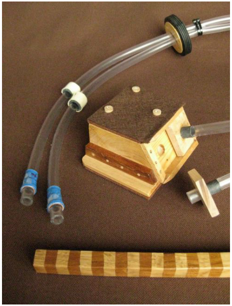
Acquisition of patient data

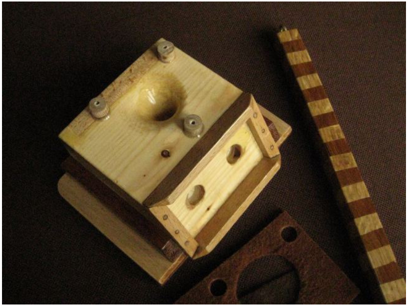
Recording speech and MRI (1)

- Modelling speech and validation of the model requires **simultaneous** recording of speech and 3D MRI from test subjects.
- Metal or electronics cannot be taken inside the MRI scanner.

Some kind of stealth technology is needed!

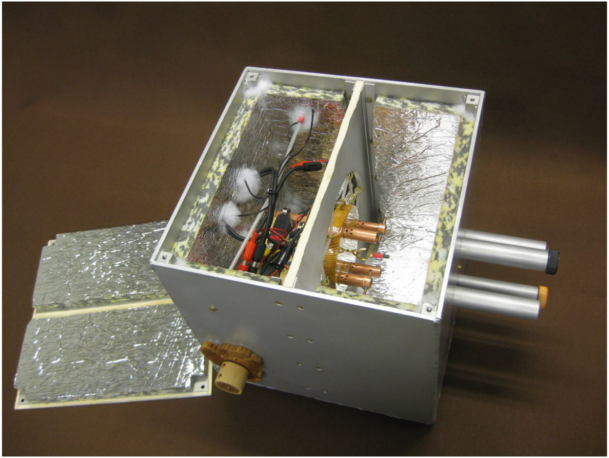
Recording speech and MRI (2)





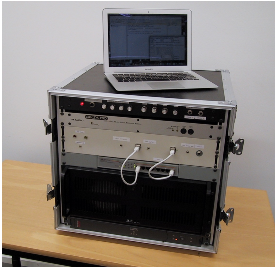
- Speech and the noise sample from MRI scanner are transmitted in "garden hose" the first 3 meters.
- Two-channel sound collector and the waveguides.
- The sound collector fits on the head and neck coils of Siemens Avanto 1.5T MRI scanner. No moving parts, immune to vibrations.

Recording speech and MRI (3)



The waveguides lead to a microphone assembly inside a sound-proof Faraday cage, beside the MRI scanner.

Recording speech and MRI (4)



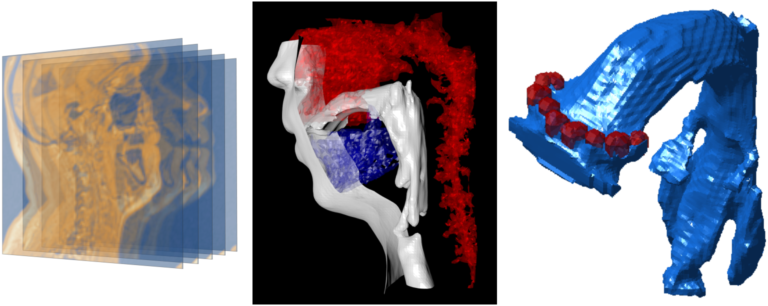
- The signals are transmitted using shielded cables from microphones to a custom-made pre-amplifier and analogue sound processing unit.
- Signals are digitised by M-Audio Delta AD-converter.
- All electronics and the computers are in a movable rack so as to speed up its installation in MRI laboratory.
- The whole experimental arrangement is (almost fully) automatised for improved efficiency.

Post-processing of MRI and sound signals

From pixels to surface models...

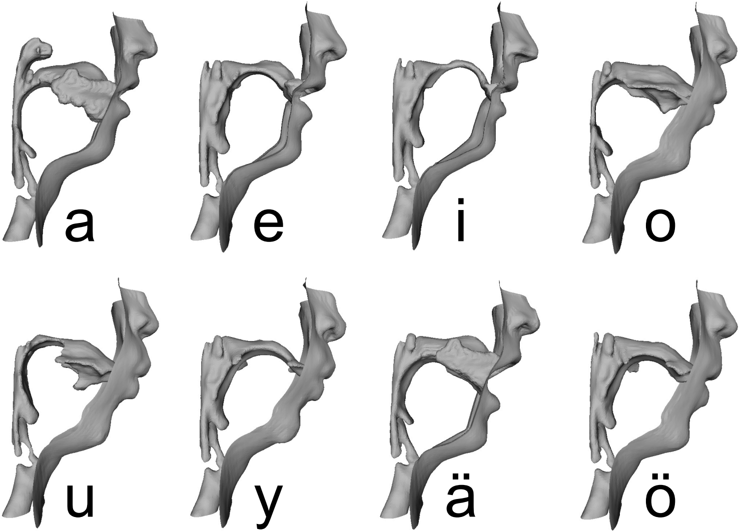
Custom 3D image processing software produces (almost) automatically surface models from the (somewhat) blurry pixel data of the MRI scanner.

The air-tissue interface is separated from other anatomic structures in surface models.



To solve acoustic equations, the surface models need be further processed to Finite Element Meshes.

... the atlas of Finnish vowels...

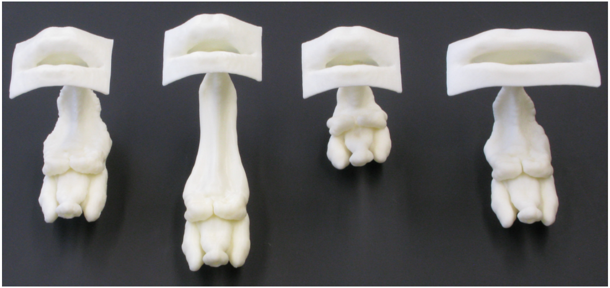


...and even to animations.

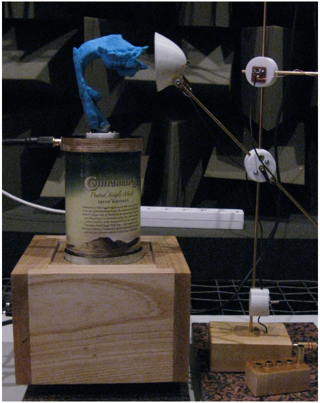
[Let us show them on full screen]

Acoustic measurements from vocal tract models produced by fast prototyping

Physical models by 3D printing (1)



Physical models by 3D printing (2)

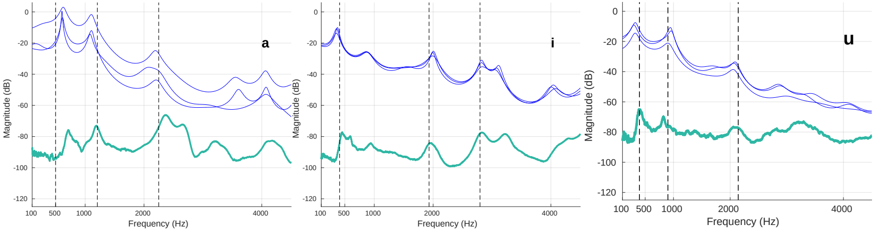


By carrying out laboratory measurements using 3D prints, it is possible to obtain independent comparison data for, e.g., model validation.

Accuracy of modelling and measurements of speech production at current stage

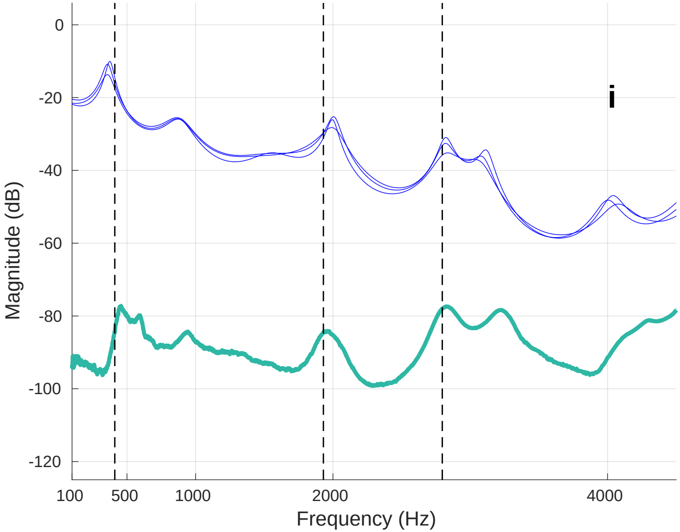
Accuracy (1)

Vowel spectrograms ([ɑ], [i], and [u]) of one test subject measured from actual speech during MRI and from 3D prints by "sweeping".



Vertical dashed lines are resonances computed from MRI numerically.

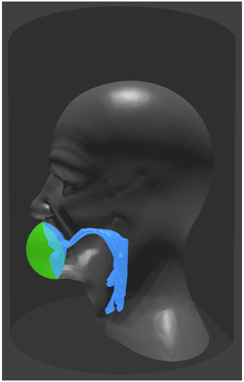
This is so nice that we better magnify it!

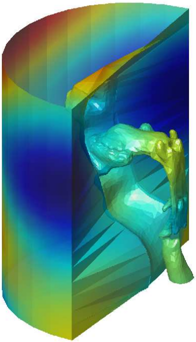


Environment acoustics
(an epilogue)

The effect of the surrounding space (1)

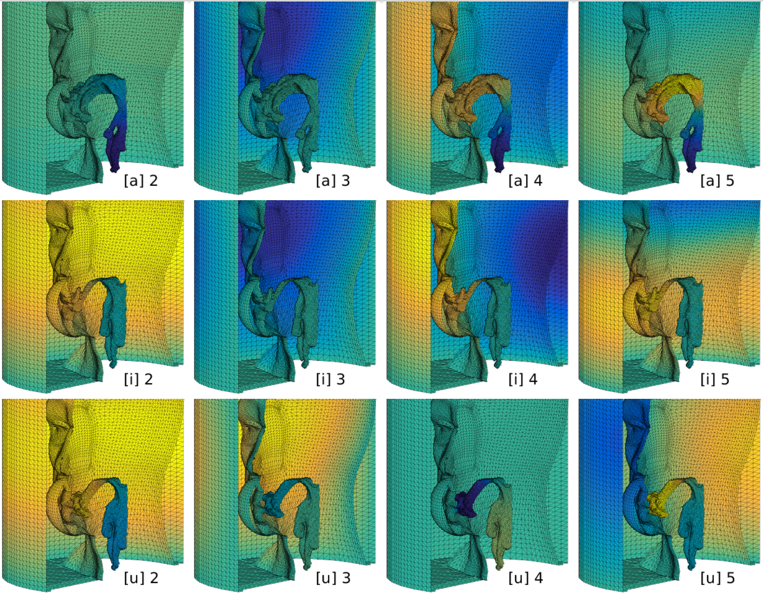
The single most significant remaining acoustic source of error is the effect of the surrounding space. This is work in progress.





The environment may be, e.g., open space or the MRI scanner coil.

The effect of the surrounding space (2)



Resonant sound pressure fields of [a], [i], [u] in a constrained environment.

Members of the COMSPEECH Research Consortium

Institute of Dentistry, University of Turku and Turku University Hospital

Institute of Behavioural Sciences, University of Helsinki

University of Alberta; Edmonton

Department of Mathematics and Systems Analysis, Aalto University, Helsinki