On the Art of Speech (and) Modelling

Antti Ojalammi

Aalto University School of Science
Department of Mathematics and Systems Analysis

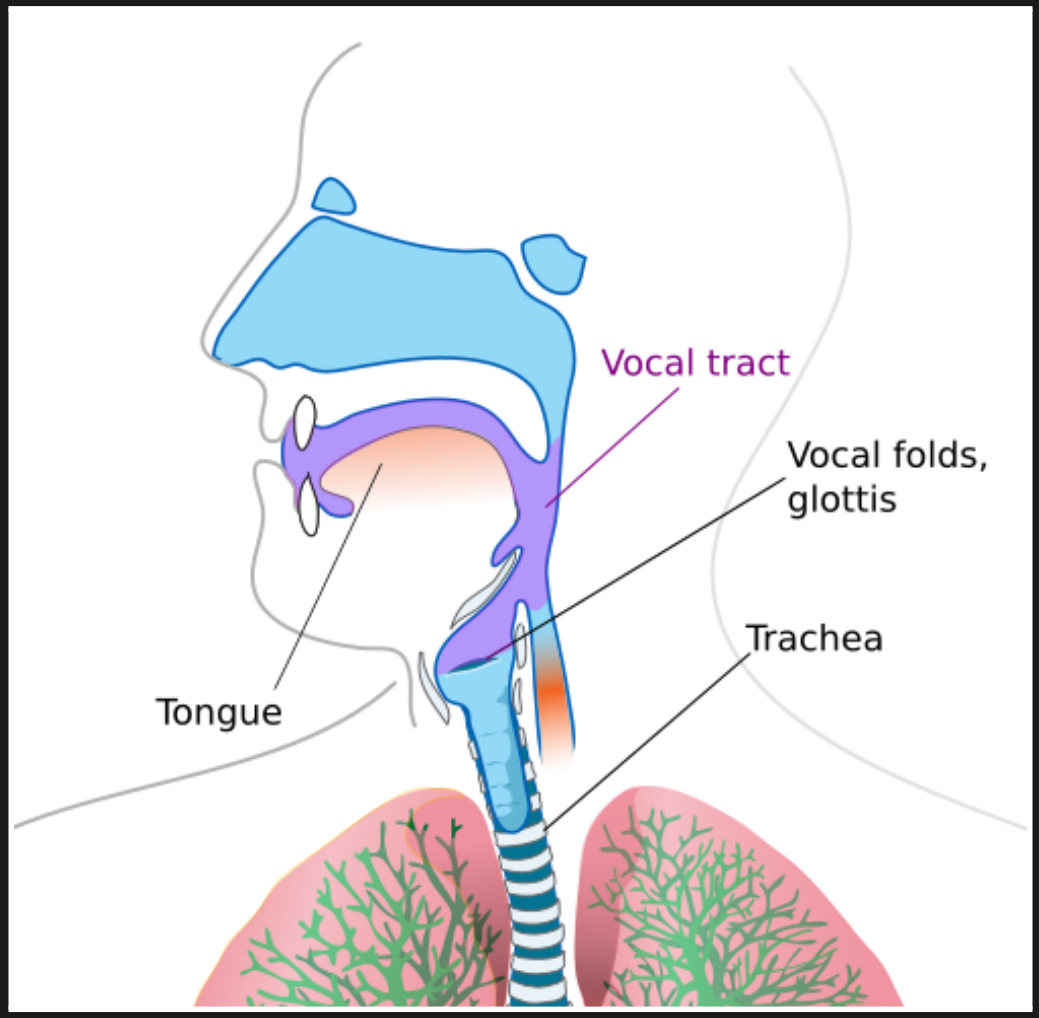
http://speech.math.aalto.fi

Mathematical Café, 26.11.2015

Aalto University
School of Science

Vocal Tract

Anatomy



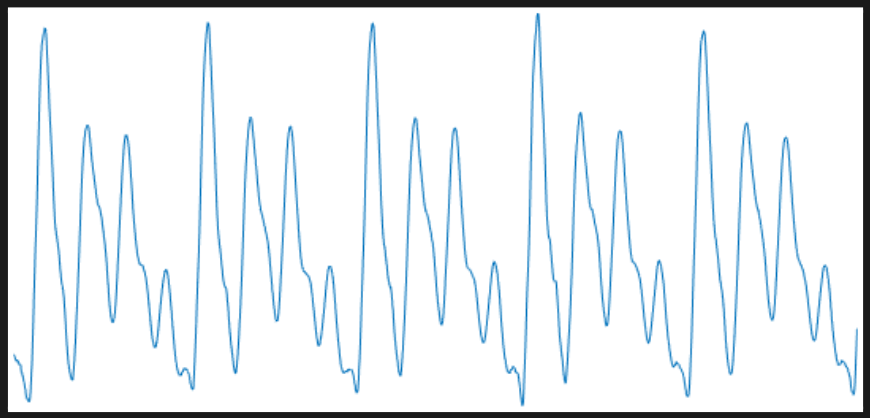
Vocal Tract

As a Filter

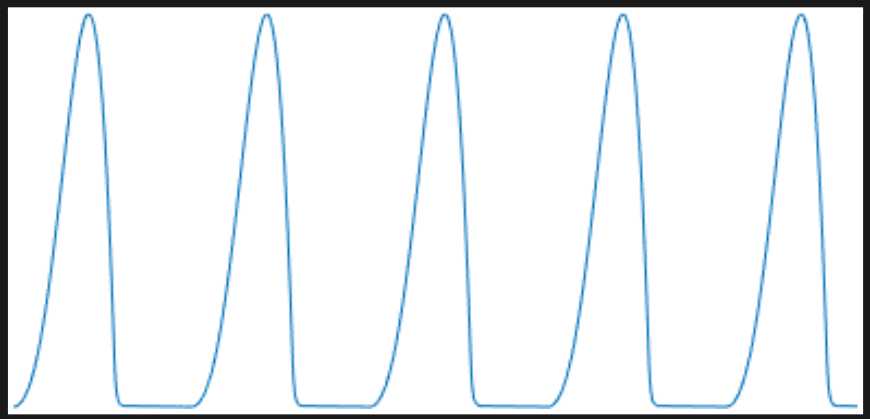
Vocal tract, [u]

Vocal Tract

As a Filter



Speech recording, vowel [u]

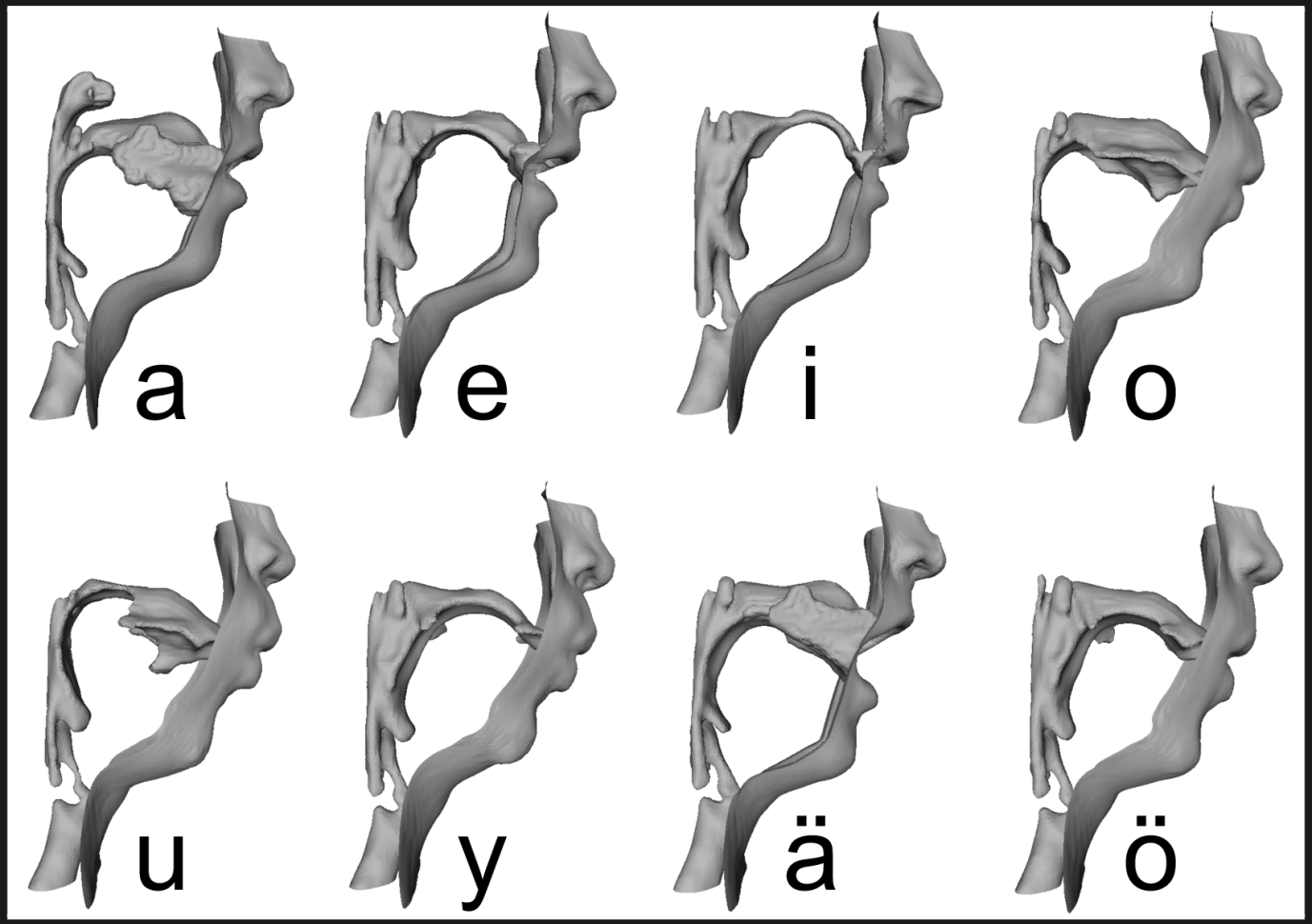


Glottal flow, simulated

Vocal tract, [u]

Introduction

Vowels



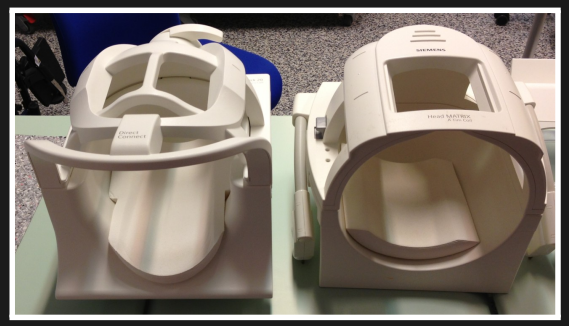
Finnish vowels

Magnetic Resonance Imaging

MRI machine

- Non-intrusive, safe 3D imaging.
- VT geometry automatically extracted from the sequence.

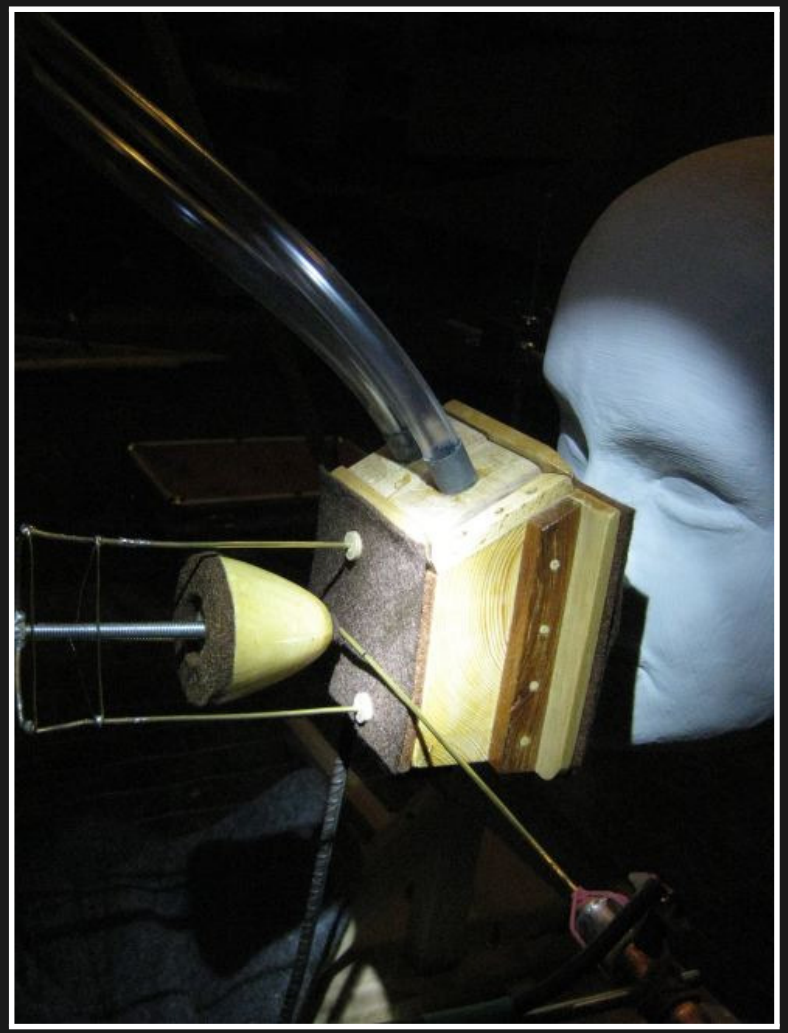
Data Acquisition



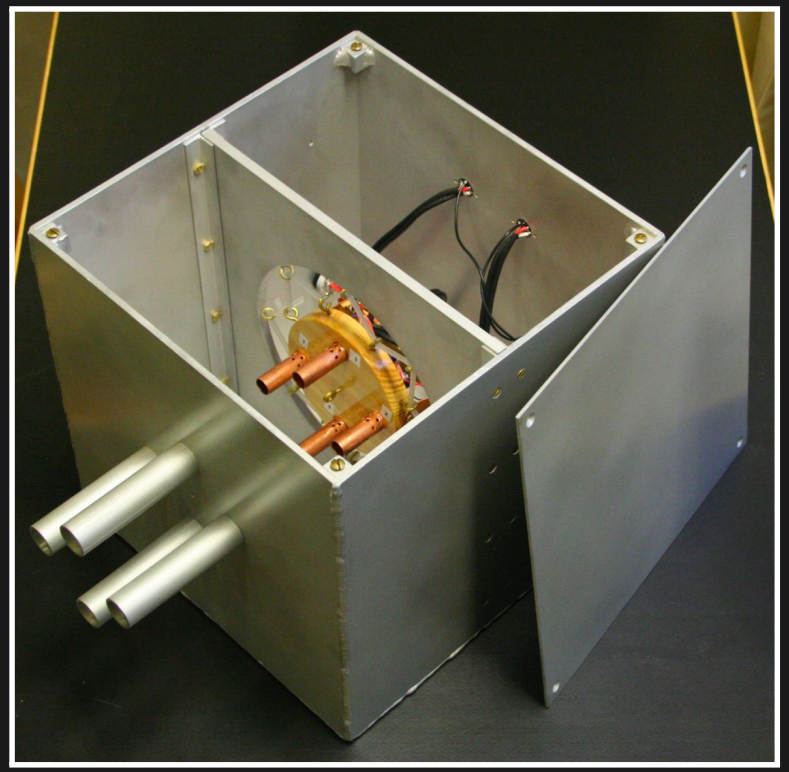
Head coil

Sagittal plane

Sound in MRI



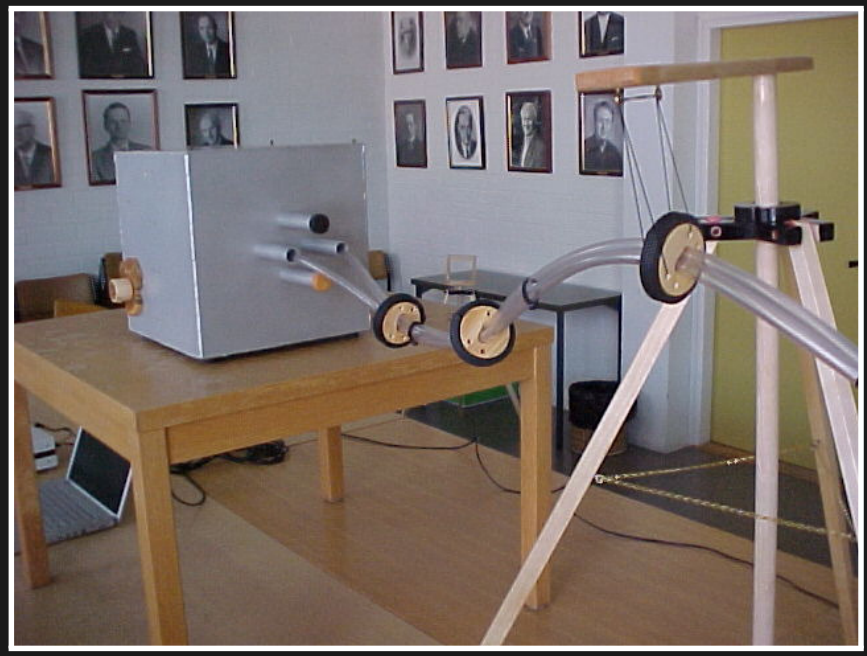
Collecting speech and noise



Faraday cage

Sound in MRI(2)

Setup demonstration

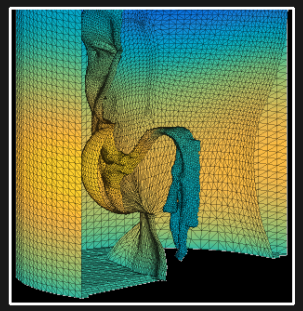


Waveguides, speech and noise channels

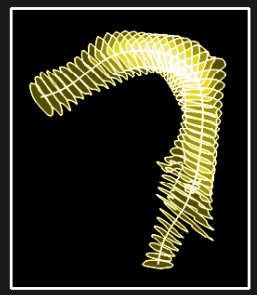
Pipeline

MRI sequence

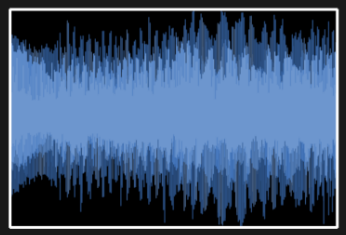
Vocal tract mesh



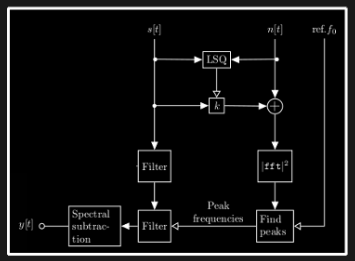
3D resonance computations



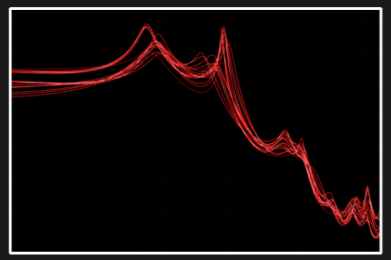
Fast 1D simulations



MRI speech

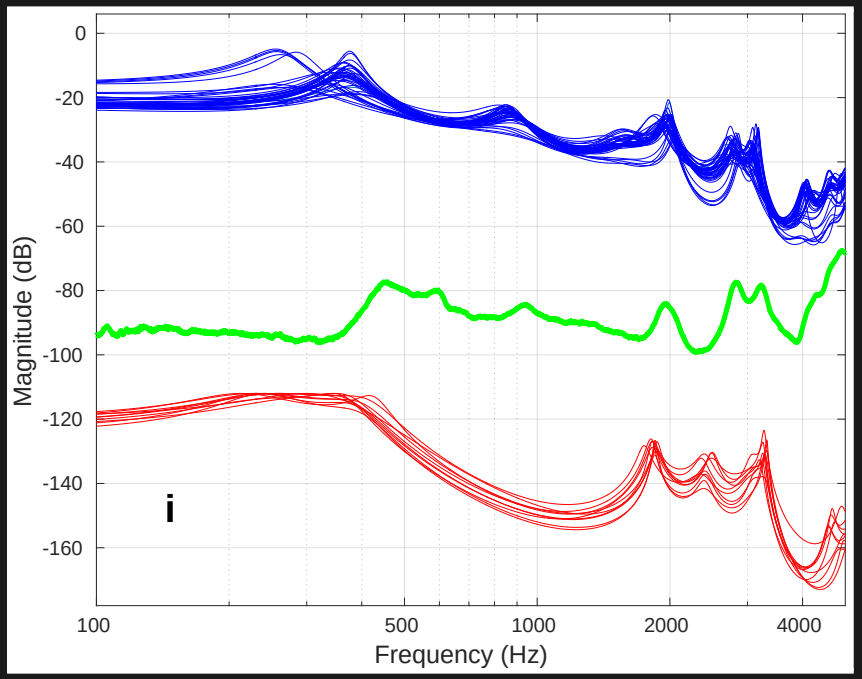


Post-processing



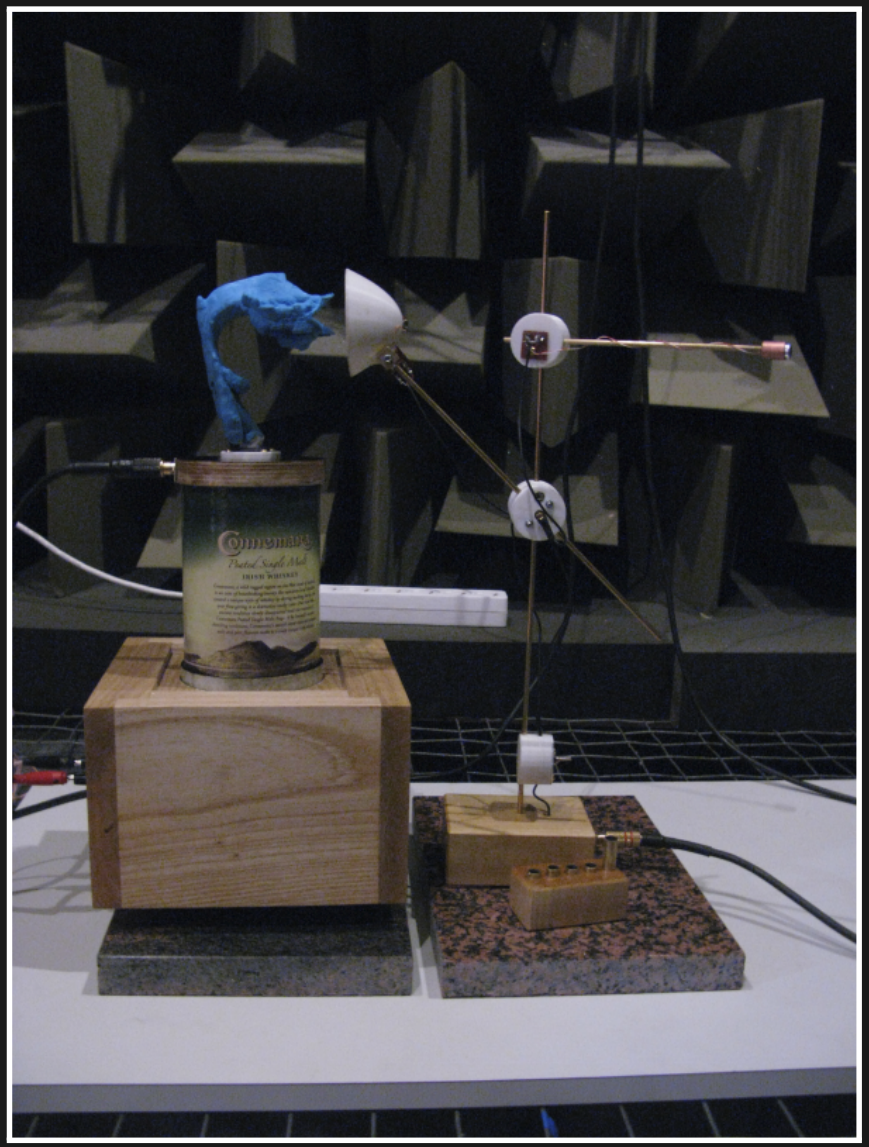
Spectral density estimation

Validation



Blue: MRI recordings, green: frequency sweep, red: anechoic recordings

- Peaks correspond to formants/resonances.
- Discrepancy between anechoic and MRI measurements.



Sweeping the frequency range

Webster's Equation

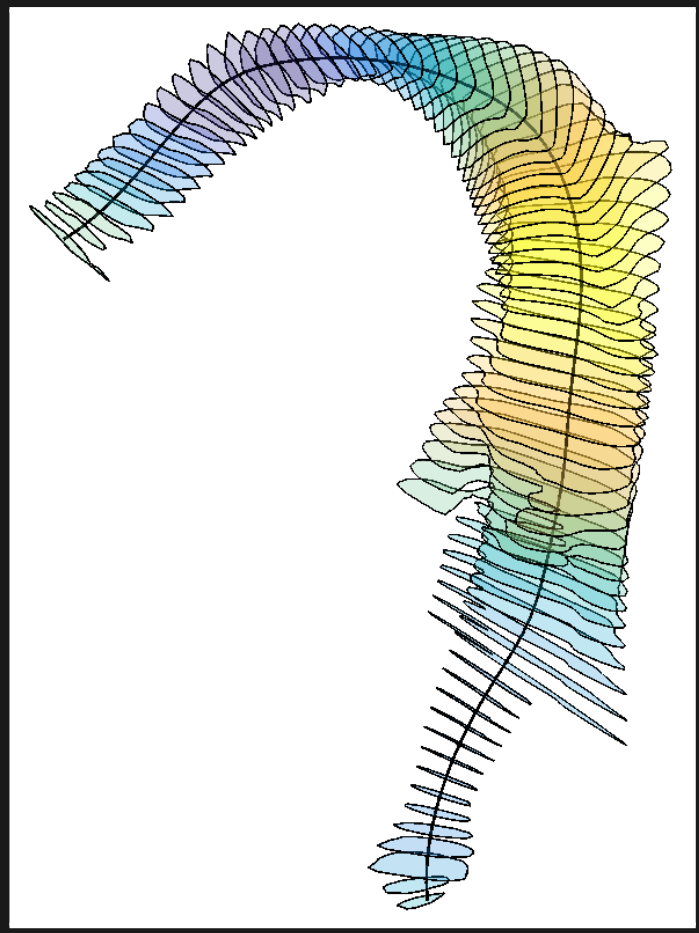
- Used for speech synthesis research at the acoustics lab.
- Parametrise the centreline by $s \in [0, 1]$

$$\frac{1}{c^2\Sigma(s)^2}\frac{\partial^2\phi}{\partial t^2} - \frac{1}{A(s)}\frac{\partial}{\partial s}\left(A(s)\frac{\partial\phi}{\partial s}\right) = 0,$$

$$\phi = \text{Velocity potential},$$

$$A(s) = \text{Area of slice at } s,$$

$$\text{Rest} = \text{Don't worry about it.}$$

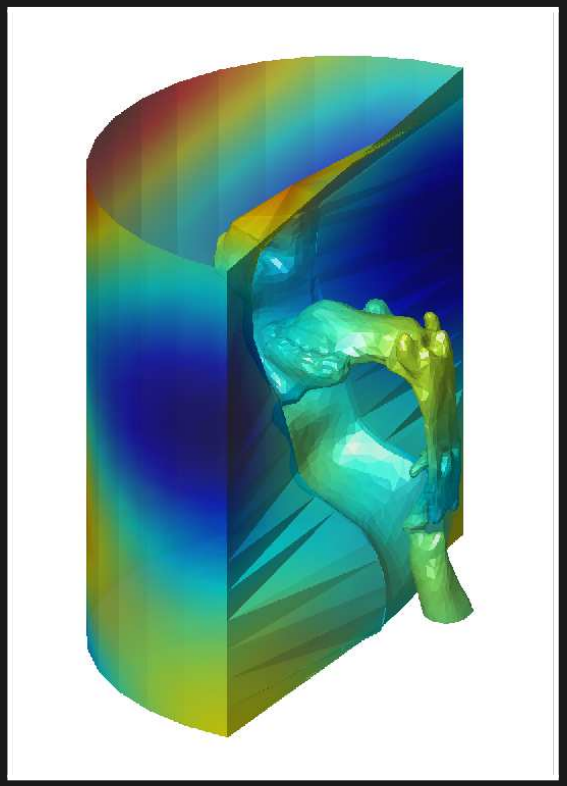


Resonances

- The resonant frequencies are related to the eigenvalue problem: Find $(\lambda, u) \in \mathbb{C} \times V$ such that

$$c^2 \Delta u = \lambda^2 u,$$

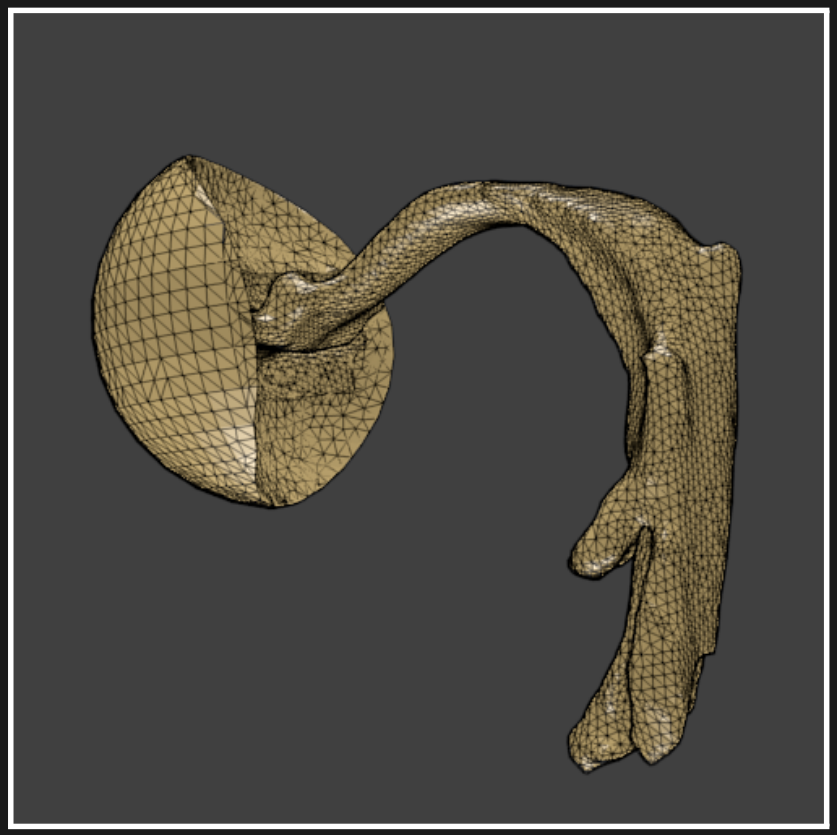
  where V is the solution space (depends on the b.c's).
- Model the head coil to account for mixed modes.



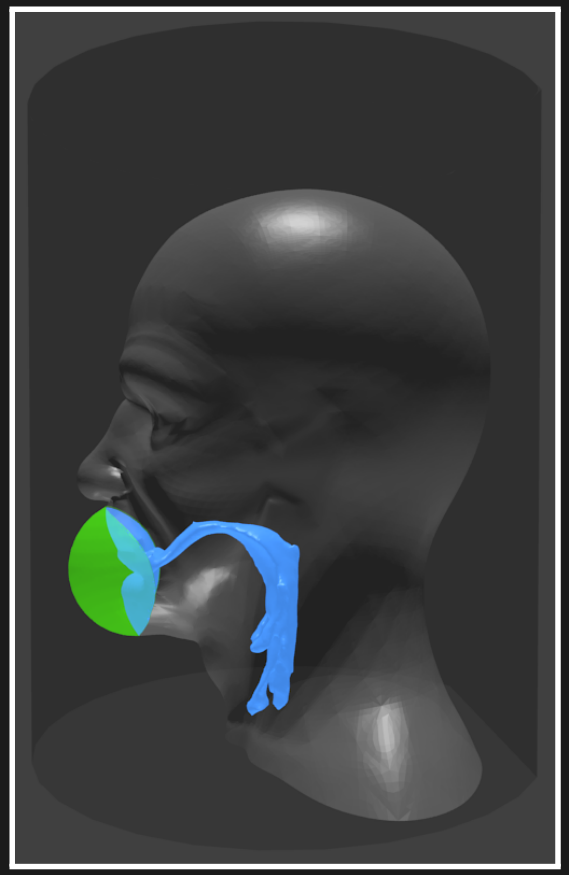
Pressure distribution for the vowel [ae]. Mixed resonance structure.

Geometries

- VT geometry and exterior acoustic space connected via a fixed interface (non-matching grids possible).
- Effect of exterior space can be pre-computed to some extent.

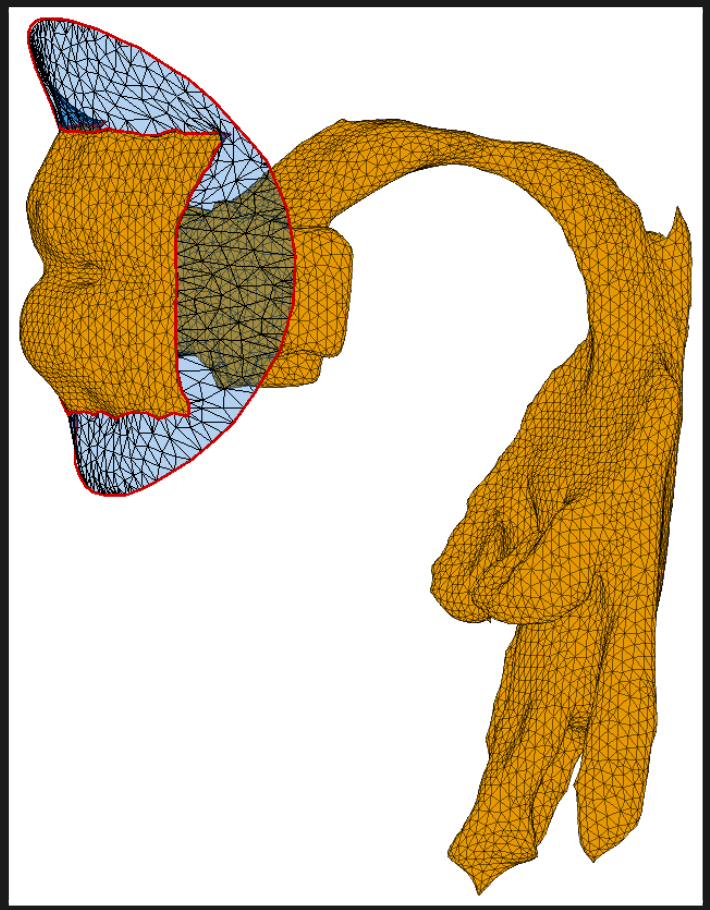


VT & interface.



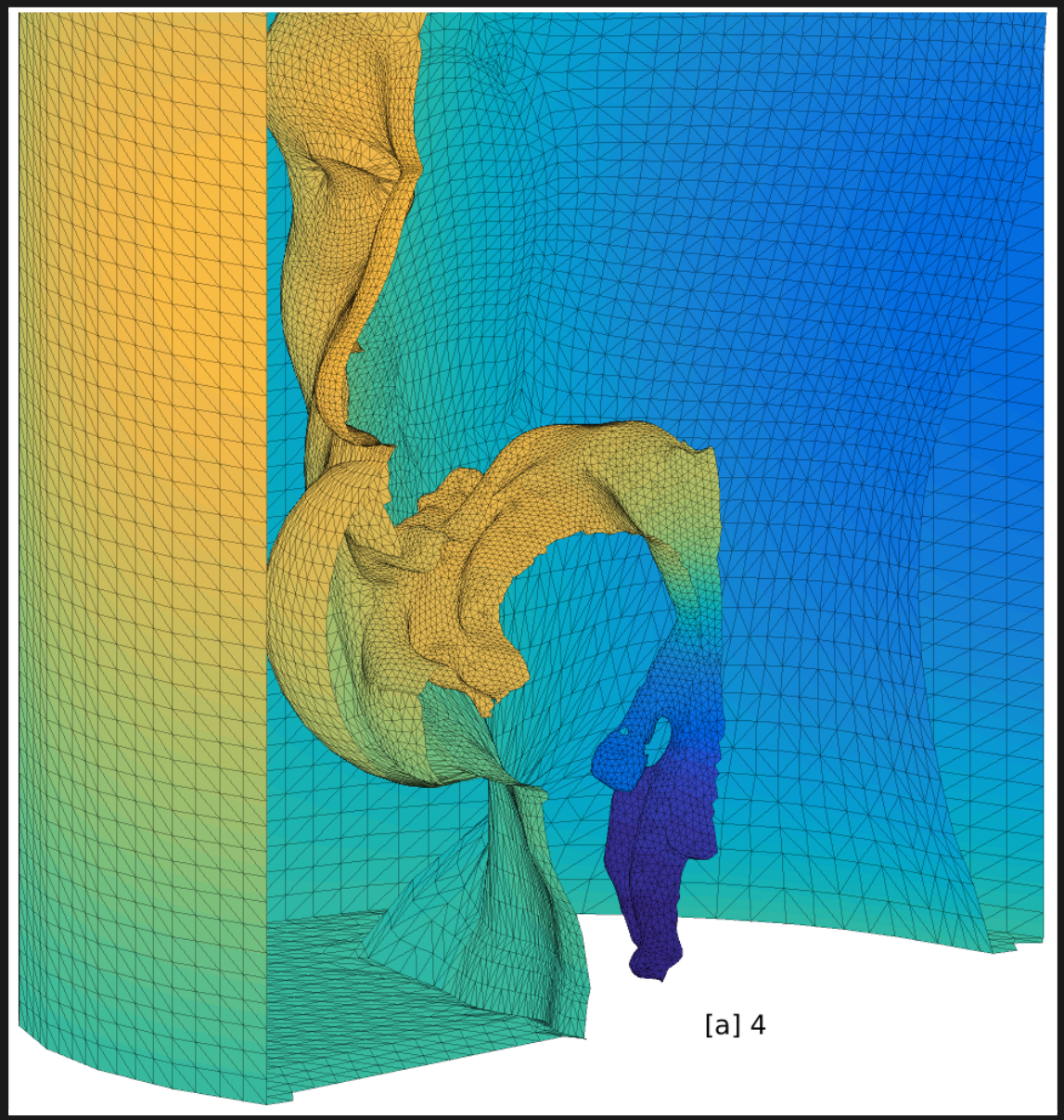
Interface in green.

Interface

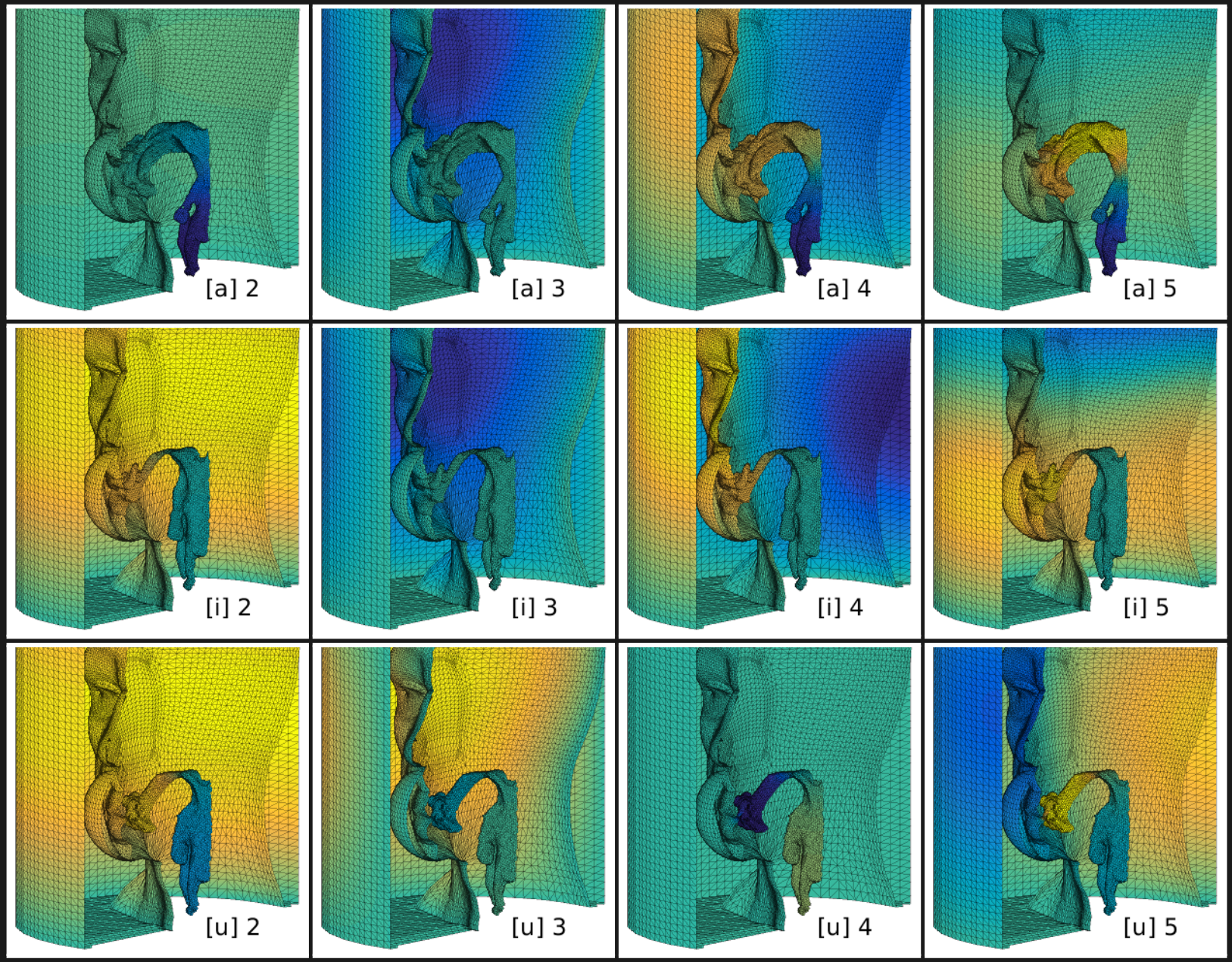


- The interface is automatically stitched to the VT geometry.
- Project the edge polygons (red) into two dimensions and triangulate.
- Solve a 2D Poisson's equation to obtain smooth depth interpolation.
- Use Nitsche's method to connect the exterior acoustic space.

Testing



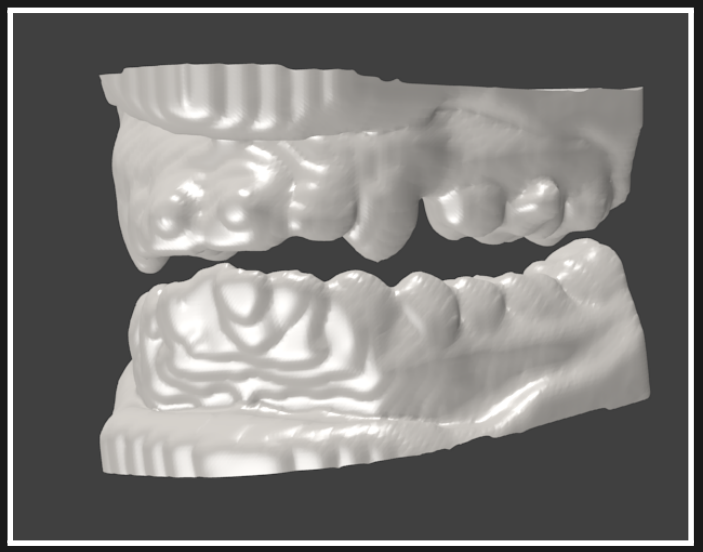
4*th* mode for [a].



Some modes for [a], [i], [u]

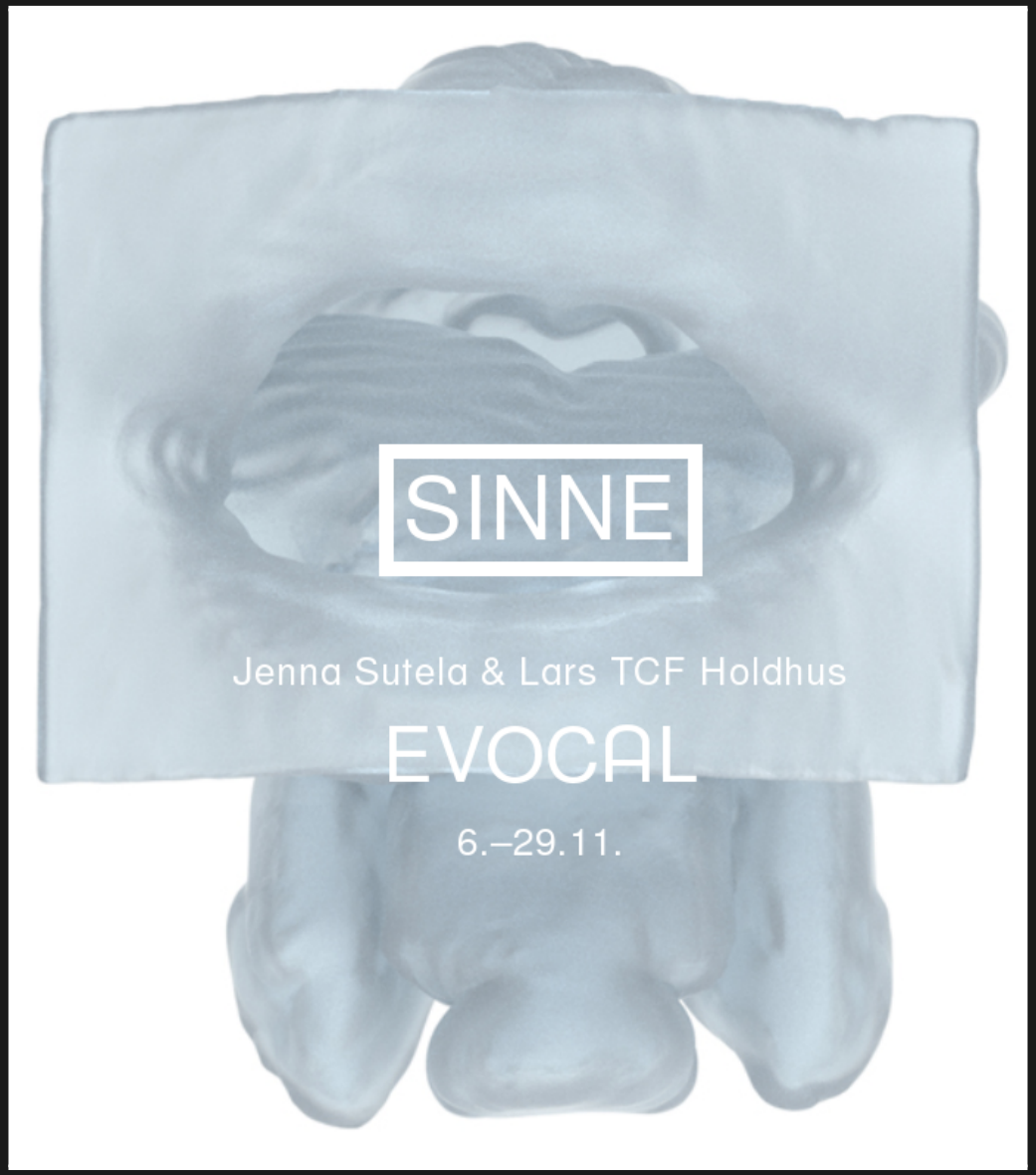
Teeth Alignment

Markers visible in MRI data

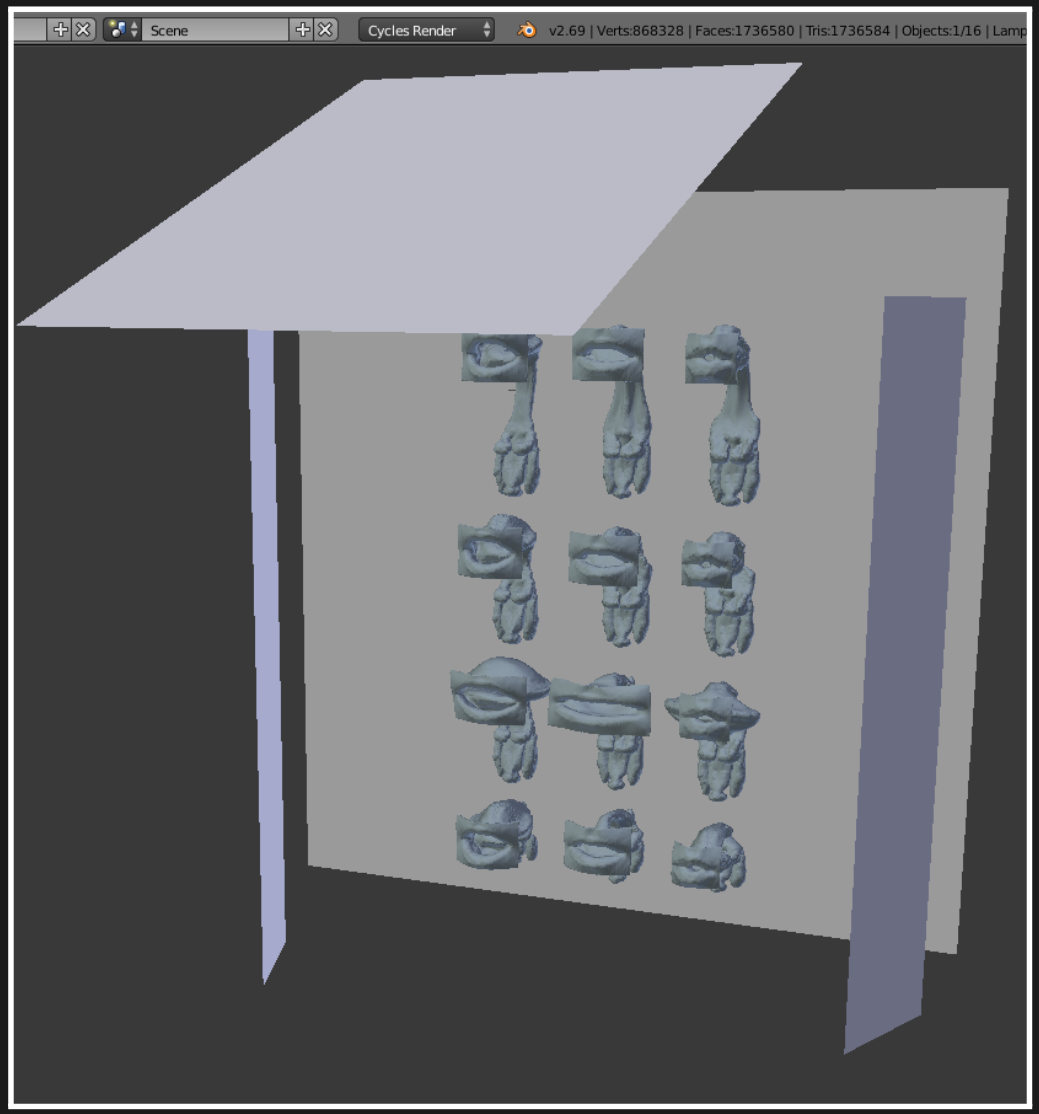


Dental mould with markers, CT scanned

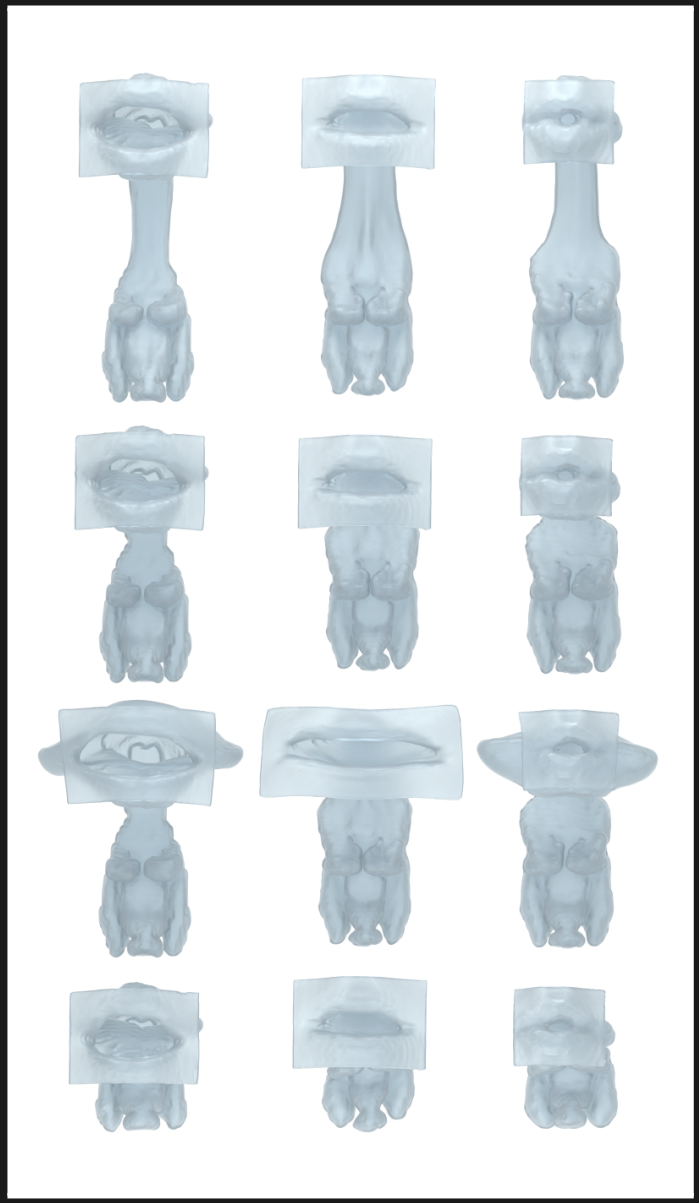
Art



Art (2)



Sculptris + Blender

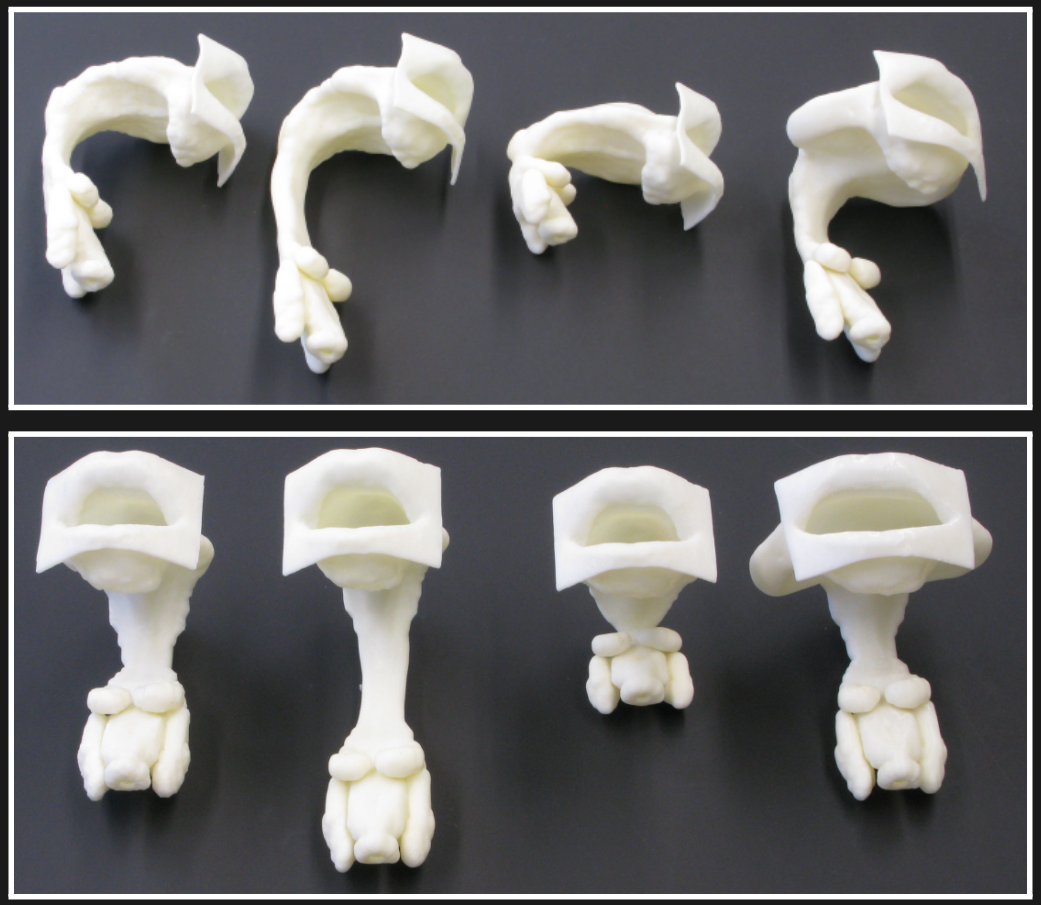


Exhibition

12 3D-printed models, different modifications

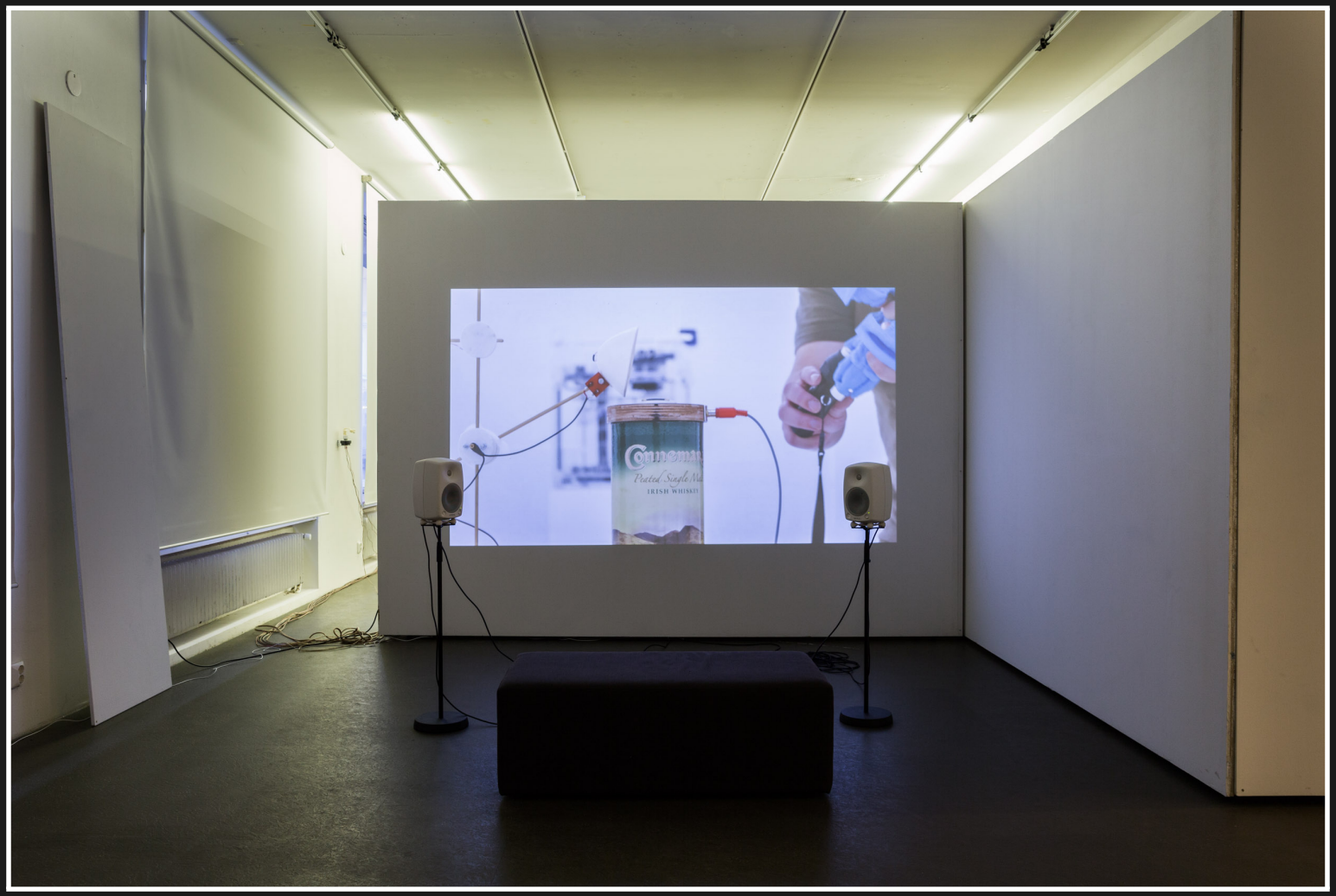
Exhibition (2)

Exhibition (3)



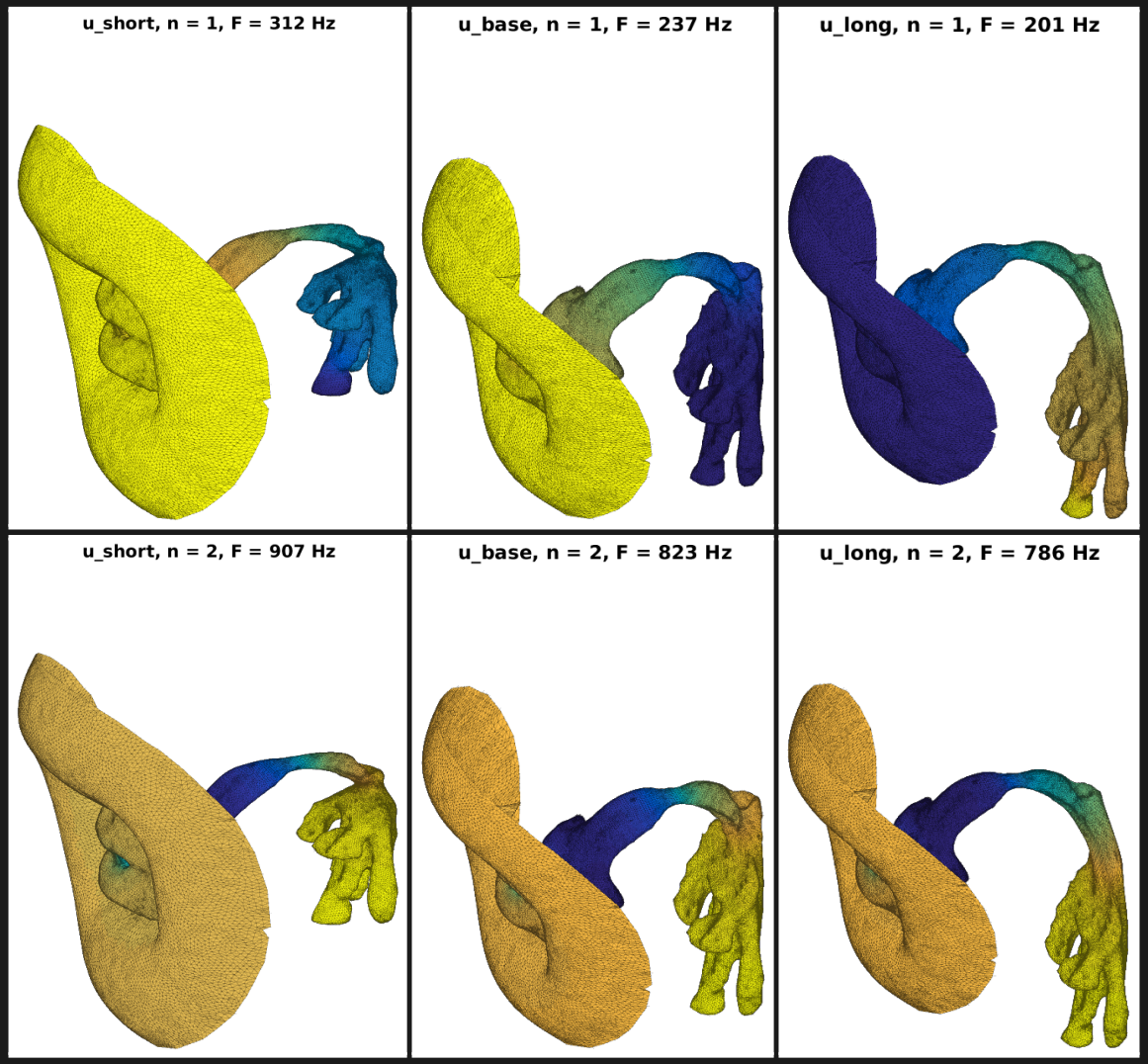
Modifications: normal, long, short, wide

Exhibition (4)



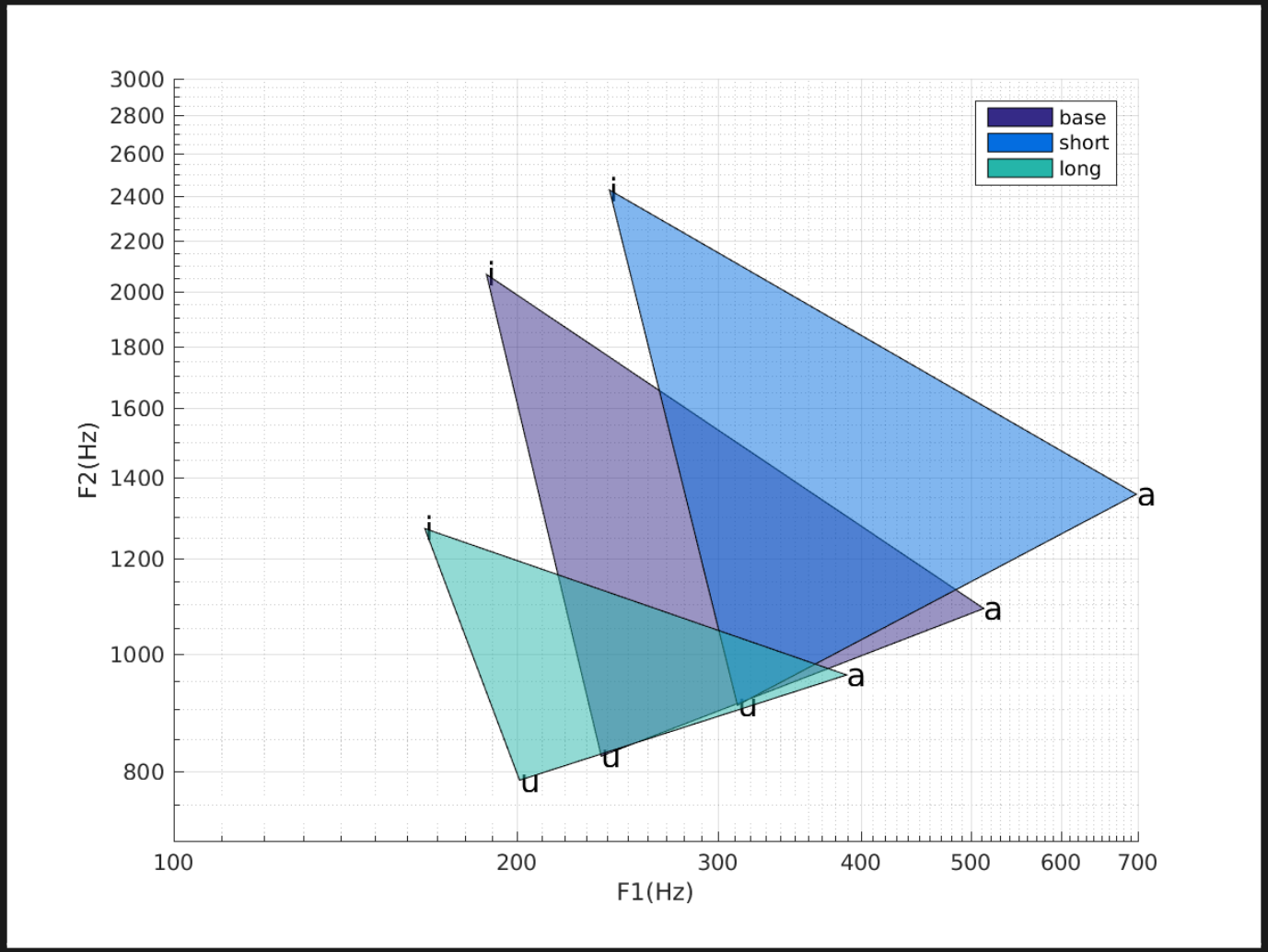
Video installation

More Resonances



Testing the effects of VT length

Vowel Space



Through the eyes of phonetics

The Future

...and Beyond

- More exhibitions, math and acoustics,
- Bigger papers and results,
- Better endings for presentations.

Thank you

http://speech.math.aalto.fi

Collaborators:

Department of Mathematics and Systems Analysis, Aalto University School of Science,
Department of Signal Processing and Acoustics, Aalto University,
Institute of Behavioural Sciences, University of Helsinki,
Department of Oral and Maxillofacial Surgery, University of Turku,
Department of Oral and Maxillofacial Diseases, Turku University Hospital, and
Medical Imaging Centre of Southwest Finland at Turku University Hospital.