Collaborative Vocal Puppetry

Multi-User Performative Voice Synthesis on Distributed Platforms @ eNTERFACE 2011

Nicolas d'Alessandro[1], Thierry Dutoit[2], Maria Astrinaki[2],
Johnty Wang[1], Onur Babacan[2], Àngel Calzada Defez[3]

[1] MAGIC - University of British Columbia (Vancouver)
[2] NUMEDIART Institute - University of Mons (Belgium)
[3] La Salle - Universitat Ramón Llull (Spain)

esf european social fund in the czech republic
EUROPEAN UNION
MINISTRY OF EDUCATION, YOUTH AND SPORTS
OP Education for Competitiveness
UNIVERSITY OF WEST BOHEMIA
INVESTMENTS IN EDUCATION DEVELOPMENT

Performative Speech Synthesis

providing artificial speech technologies vs. providing artificial speaking systems

- text-to-speech (TTS) based on non-uniform unit selection (NUU) is the ultimate technology for providing high-quality speech waveforms - period, thank you :)
- it leads to a well-established but rather limited set of apps where the reading of some available text is the goal;
- but it does not make the conception of an *artificial speaker* much easier than back in the days of DECTalk, well actually it makes it harder (black-box design, time scales issues);
- the *artificial speaker* as a system: realtime, reactive, interactive, listener-specific, context-aware, ubiquitous,... many aspects which require a different perspective.

esf european social fund in the czech republic
EUROPEAN UNION
MINISTRY OF EDUCATION, YOUTH AND SPORTS
OP Education for Competitiveness
UNIVERSITY OF WEST BOHEMIA
INVESTMENTS IN EDUCATION DEVELOPMENT

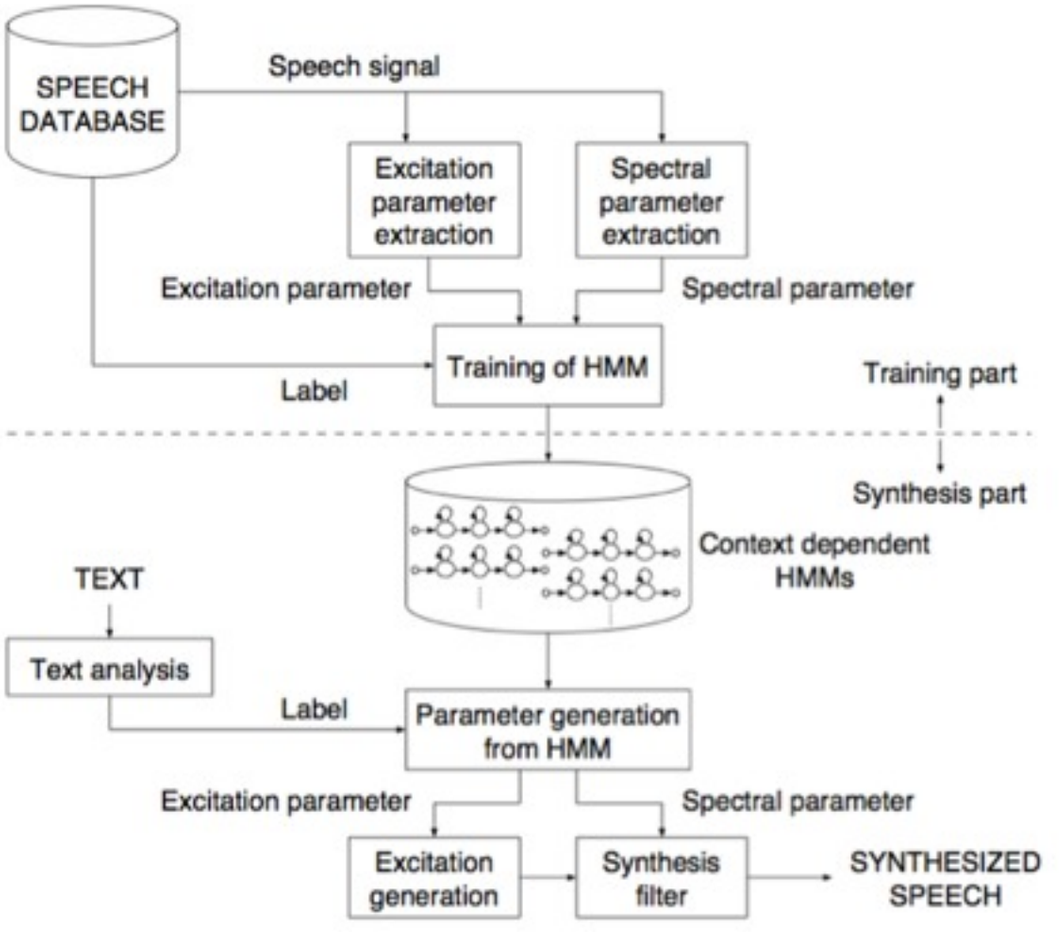
exploring... based on the two most promising approaches we have so far

1) bio-mechanical modelling
2) statistical parametric modelling

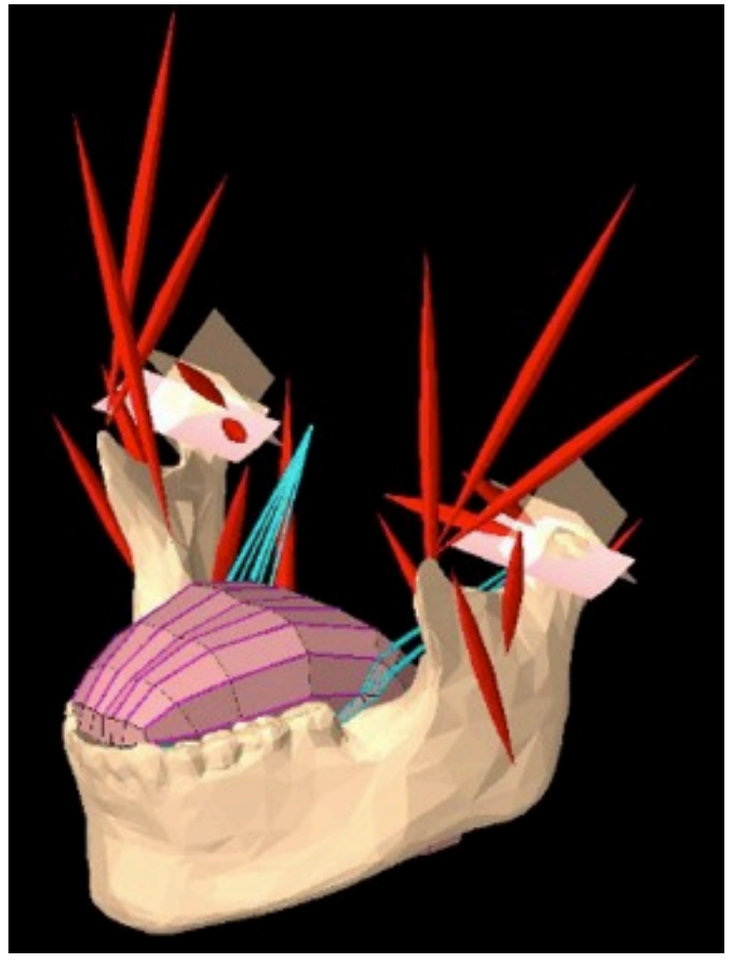
INVESTMENTS IN EDUCATION DEVELOPMENT

Starting Points

two promising but non-adapted, totally non-reactive systems: HTS and Artisynth



HTS



Artisynth

esf european social fund in the czech republic
EUROPEAN UNION
MINISTRY OF EDUCATION, YOUTH AND SPORTS
OP Education for Competitiveness
UNIVERSITY OF WEST BOHEMIA
INVESTMENTS IN EDUCATION DEVELOPMENT

I think we went over the hill...

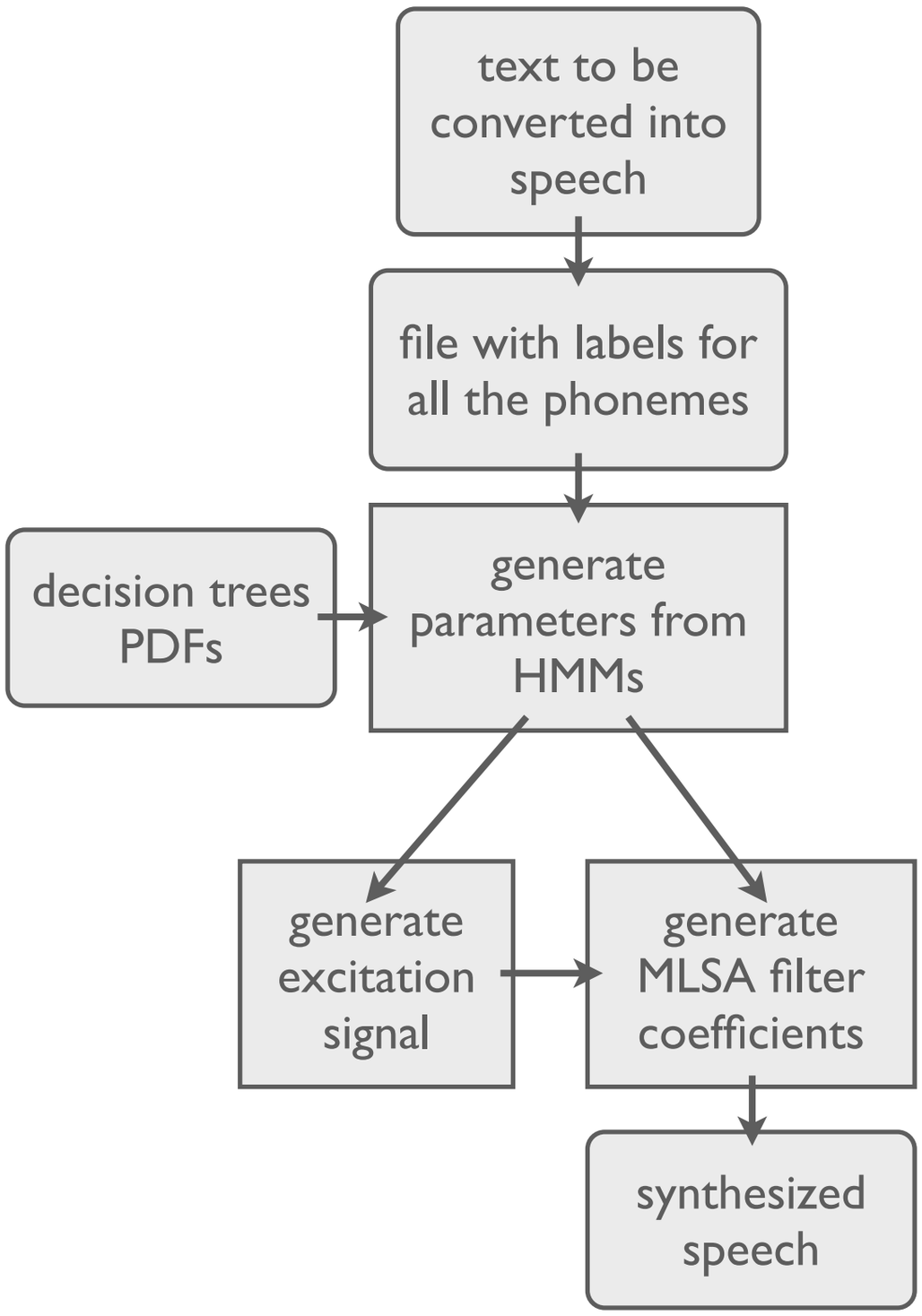
we reached convincing reactivity for both

esf european social fund in the czech republic
EUROPEAN UNION
MINISTRY OF EDUCATION, YOUTH AND SPORTS
OP Education for Competitiveness
UNIVERSITY OF WEST BOHEMIA
INVESTMENTS IN EDUCATION DEVELOPMENT

MAGE

esf european social fund in the czech republic
EUROPEAN UNION
MINISTRY OF EDUCATION, YOUTH AND SPORTS
OP Education for Competitiveness
UNIVERSITY OF WEST BOHEMIA
INVESTMENTS IN EDUCATION DEVELOPMENT

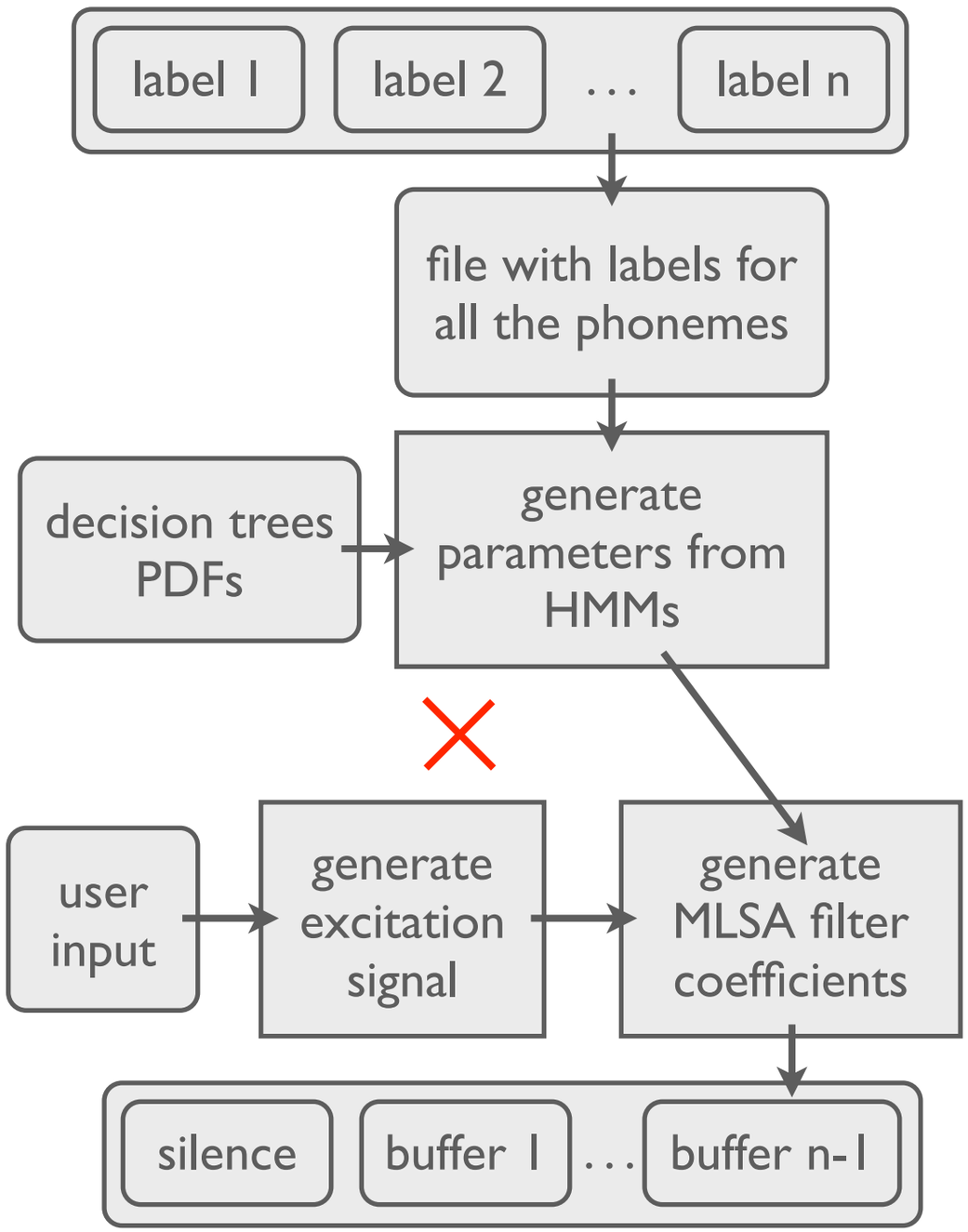
HTS: HMM-Based Speech Synthesis



- input = full text to be converted
- parsing text into labels containing full contextual linguistic information
- select created decision trees and PDFs from the database
- generate spectrum, pitch and duration trajectories, optimized for the whole label target
- generate speech in a file

esf european social fund in the czech republic
EUROPEAN UNION
MINISTRY OF EDUCATION, YOUTH AND SPORTS
OP Education for Competitiveness
UNIVERSITY OF WEST BOHEMIA
INVESTMENTS IN EDUCATION DEVELOPMENT

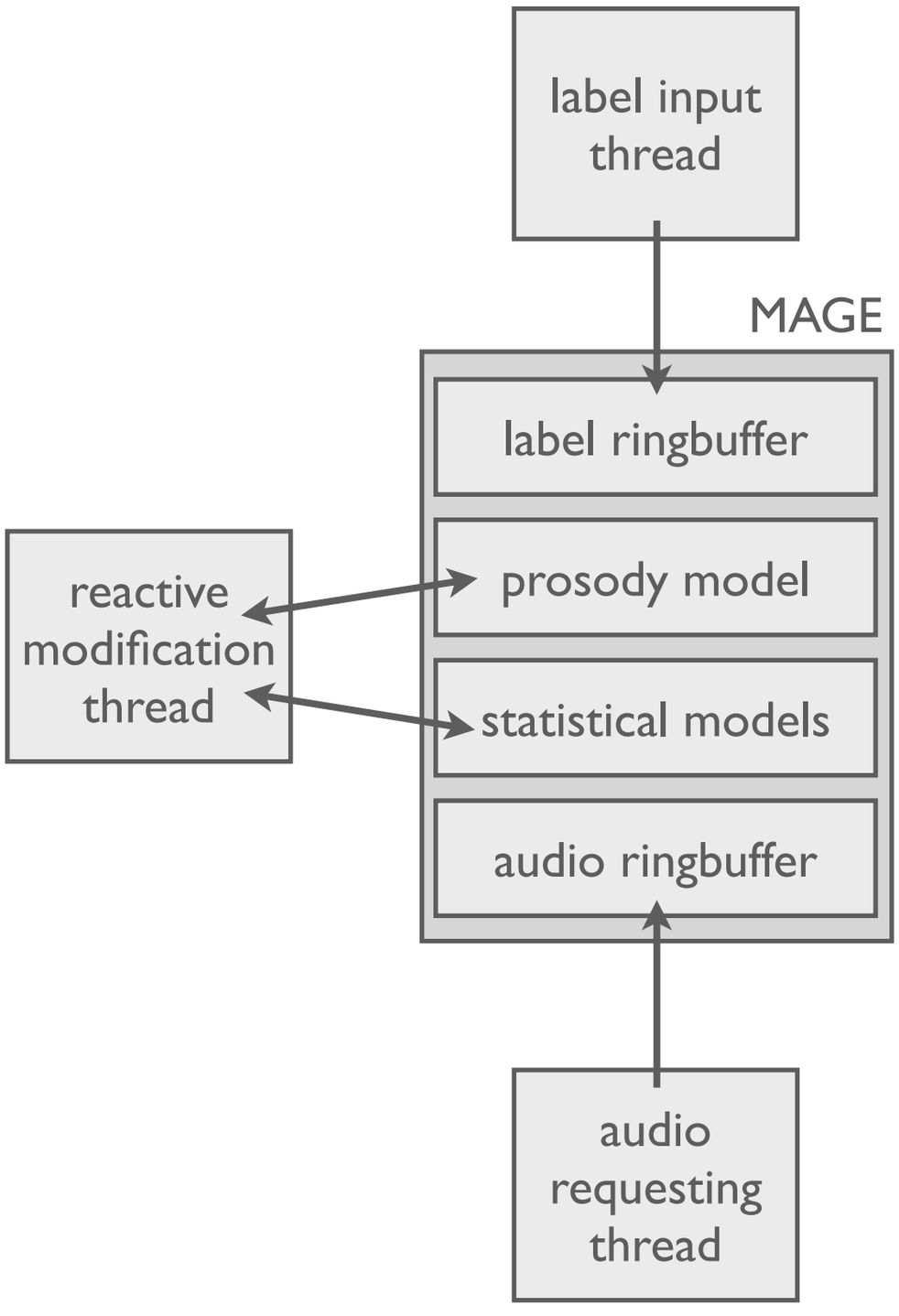
pHTS: Breaking the loop



- sequential input of labels containing minimal contextual linguistic information
- delay of only one label to preserve coarticulation
- select created decision trees and PDFs from the database
- generate spectrum, pitch and duration trajectories for one label
- generate one buffer of audio

esf european social fund in the czech republic
EUROPEAN UNION
MINISTRY OF EDUCATION, YOUTH AND SPORTS
OP Education for Competitiveness
UNIVERSITY OF WEST BOHEMIA
INVESTMENTS IN EDUCATION DEVELOPMENT

MAGE: Ready for reactive software



- thread-safe interface between the synthesizer and I/Os: labels and audio samples (lock-free)
- realtime access to the internal state of the engine: each step of the statistical modelling + prosody
- engine-independent

esf european social fund in the czech republic
EUROPEAN UNION
MINISTRY OF EDUCATION, YOUTH AND SPORTS
OP Education for Competitiveness
UNIVERSITY OF WEST BOHEMIA
INVESTMENTS IN EDUCATION DEVELOPMENT

Max: HandSketch

wacom driver

polar mapping

OSC sender

oF: chunk composer

phonetic keyboard

chunk builder

OSC sender

/chunk ...

oF: chunk turntable

OSC listener

chunk repository

chunk sequencer

OSC sender

/hs/pitch ...

/chunk ...

OSC parsing

OSC listener

prosody mapping

chunk parsing

label ringbuffer

prosody model

audio ringbuffer

audio

ofxFablio

audio callback

MAGE synthesizer

parse label

update PDFs

update MLSAs

update audio

MAGE (pHTS engine)

Reactive Natural Language Processing

HTS training

using full contextual linguistic information (from phonemes to complete utterance)

pHTS/MAGE training

using minimum contextual linguistic information reduced only to past and present phonemes and syllables

esf european social fund in the czech republic
EUROPEAN UNION
MINISTRY OF EDUCATION, YOUTH AND SPORTS
OP Education for Competitiveness
UNIVERSITY OF WEST BOHEMIA
INVESTMENTS IN EDUCATION DEVELOPMENT

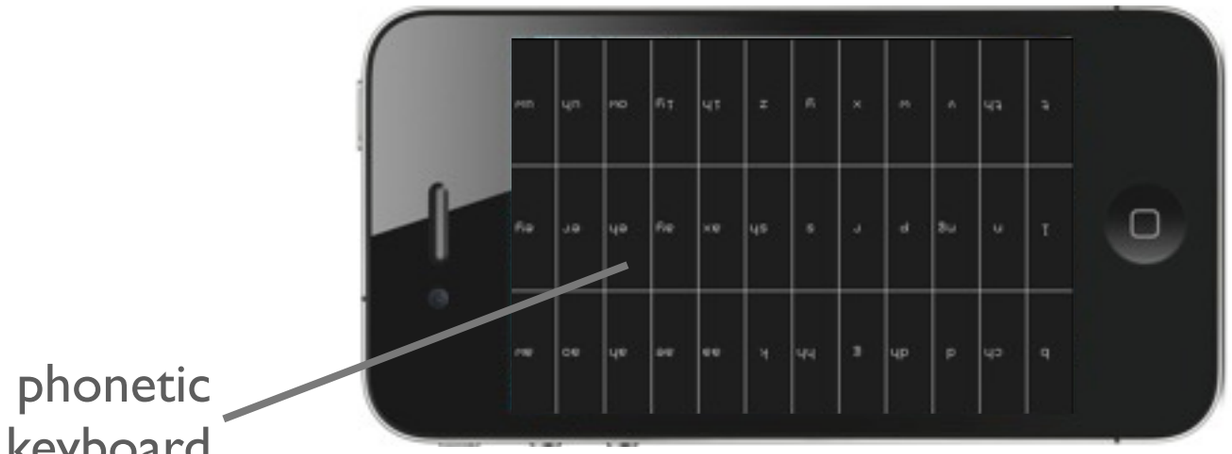
Reactive Natural Language Processing

- phoneme:
  - {before previous, previous, current, next, next next}
  - position of current phoneme in current syllable
- syllable:
  - number of phonemes at {before previous, previous, current, next, next next}
  - accent of {before previous, previous, current, next, next next}
  - stress of {before previous, previous, current, next, next next}
  - position of current syllable in current word
  - number of {preceding, succeeding} stressed syllables in current phrase
  - number of {preceding, succeeding} accented syllables in current phrase
  - number of syllables {from previous, to next} stressed syllable
  -number of syllables {from previous, to next} accented syllable
  - vowel within current syllable
- word:
  - guess at part of speech of {preceding, current, succeeding} word
  - number of syllables in {preceding, current, succeeding} word
  - position of current word in current phrase
  - number of {preceding,succeeding}content words in current phrase
  - number of words {from previous, to next} content word
- phrase:
  - number of syllables in {preceding, current, succeeding} phrase
  - position in major phrase
  - ToBI7 endtone of current phase
- utterance:
  - number of syllables in current utterance

- phoneme:
  - {before previous, previous, current, next, ~~next next~~}
  - position of current phoneme in current syllable
- syllable:
  - number of phonemes at {before previous, previous, current, next, ~~next next~~}
  - accent of {before previous, previous, current, next, ~~next next~~}
  - number of syllables {from previous, to current} accented syllable
  - number of syllables {to next} accented syllable : <u>random</u>
  - vowel within current syllable
- word:
  - No information
- phrase:
  - No information
- utterance:
  - No information

esf european social fund in the czech republic
EUROPEAN UNION
MINISTRY OF EDUCATION, YOUTH AND SPORTS
OP Education for Competitiveness
UNIVERSITY OF WEST BOHEMIA
INVESTMENTS IN EDUCATION DEVELOPMENT

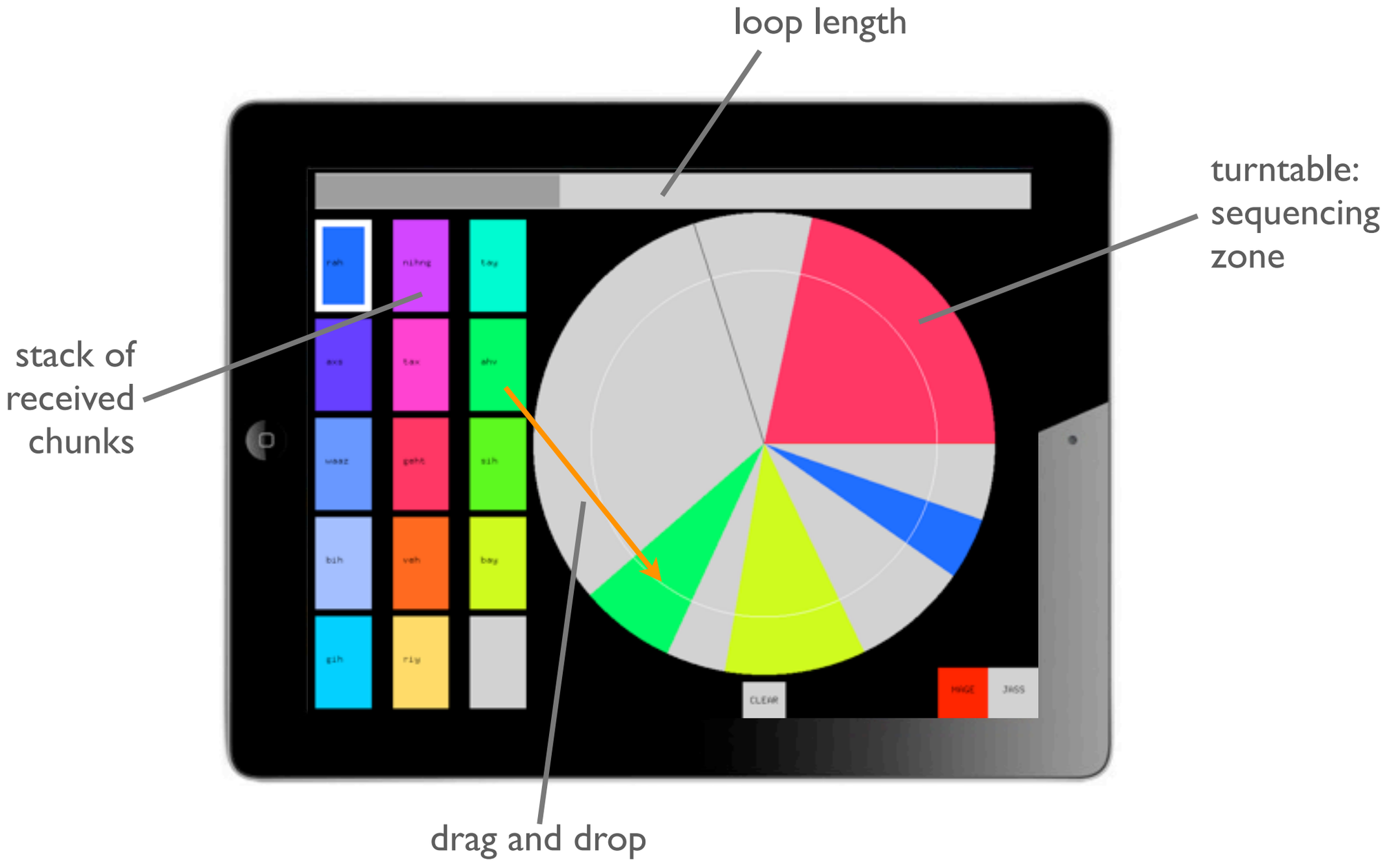
Chunk Composer



double-tap sends the chunk

esf european social fund in the czech republic
EUROPEAN UNION
MINISTRY OF EDUCATION, YOUTH AND SPORTS
OP Education for Competitiveness
UNIVERSITY OF WEST BOHEMIA
INVESTMENTS IN EDUCATION DEVELOPMENT

Chunk Turntable



double-tap stresses the vowel chunk is sent when head hits

esf european social fund in the czech republic
EUROPEAN UNION
MINISTRY OF EDUCATION, YOUTH AND SPORTS
OP Education for Competitiveness
UNIVERSITY OF WEST BOHEMIA
INVESTMENTS IN EDUCATION DEVELOPMENT

HandSketch

fan diagram: control
through angle and radius

angle ~ pitch variation
radius ~ speed variation

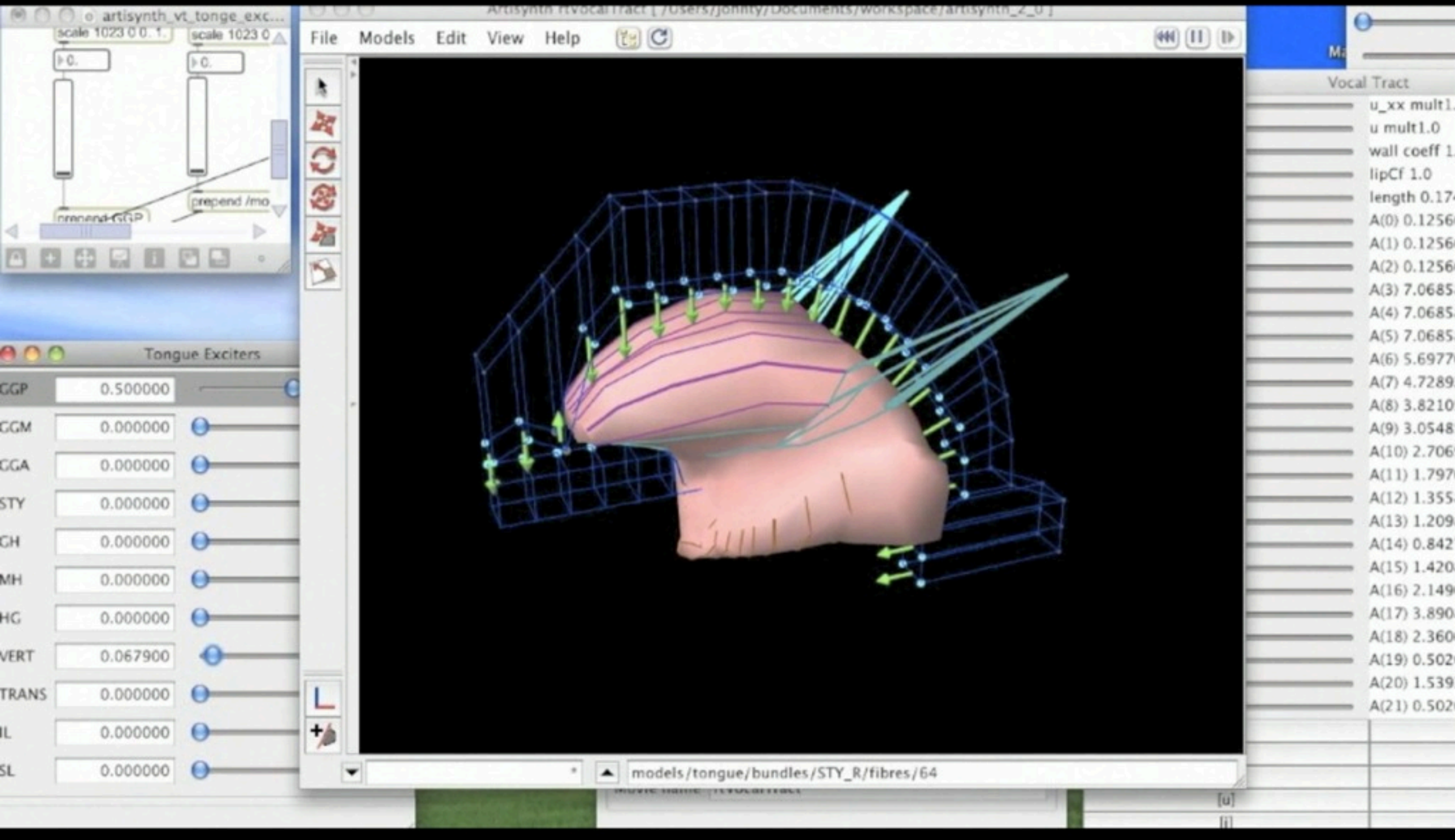
pen: pressure
and tilt + buttons

pressure ~ volume
tilt ~ vocal tract size
button #1 ~ abs/rel pitch
button #2 ~ gate tilt

esf european social fund in the czech republic
EUROPEAN UNION
MINISTRY OF EDUCATION, YOUTH AND SPORTS
OP Education for Competitiveness
UNIVERSITY OF WEST BOHEMIA
INVESTMENTS IN EDUCATION DEVELOPMENT

FEM Tongue

esf european social fund in the czech republic
EUROPEAN UNION
MINISTRY OF EDUCATION, YOUTH AND SPORTS
OP Education for Competitiveness
UNIVERSITY OF WEST BOHEMIA
INVESTMENTS IN EDUCATION DEVELOPMENT



esf european social fund in the czech republic
EUROPEAN UNION
MINISTRY OF EDUCATION, YOUTH AND SPORTS
OP Education for Competitiveness
UNIVERSITY OF WEST BOHEMIA
INVESTMENTS IN EDUCATION DEVELOPMENT

Artisynth simulation

OSC listener

Reduced FEM Tongue + Solid Bodies for Jaw, Palate and Pharynx

OSC sender

audio

audio callback

JASS simulation

OSC listener

airflow simulation

JAVA audio

oF: tract turntable

OSC listener

chunk repository

chunk sequencer

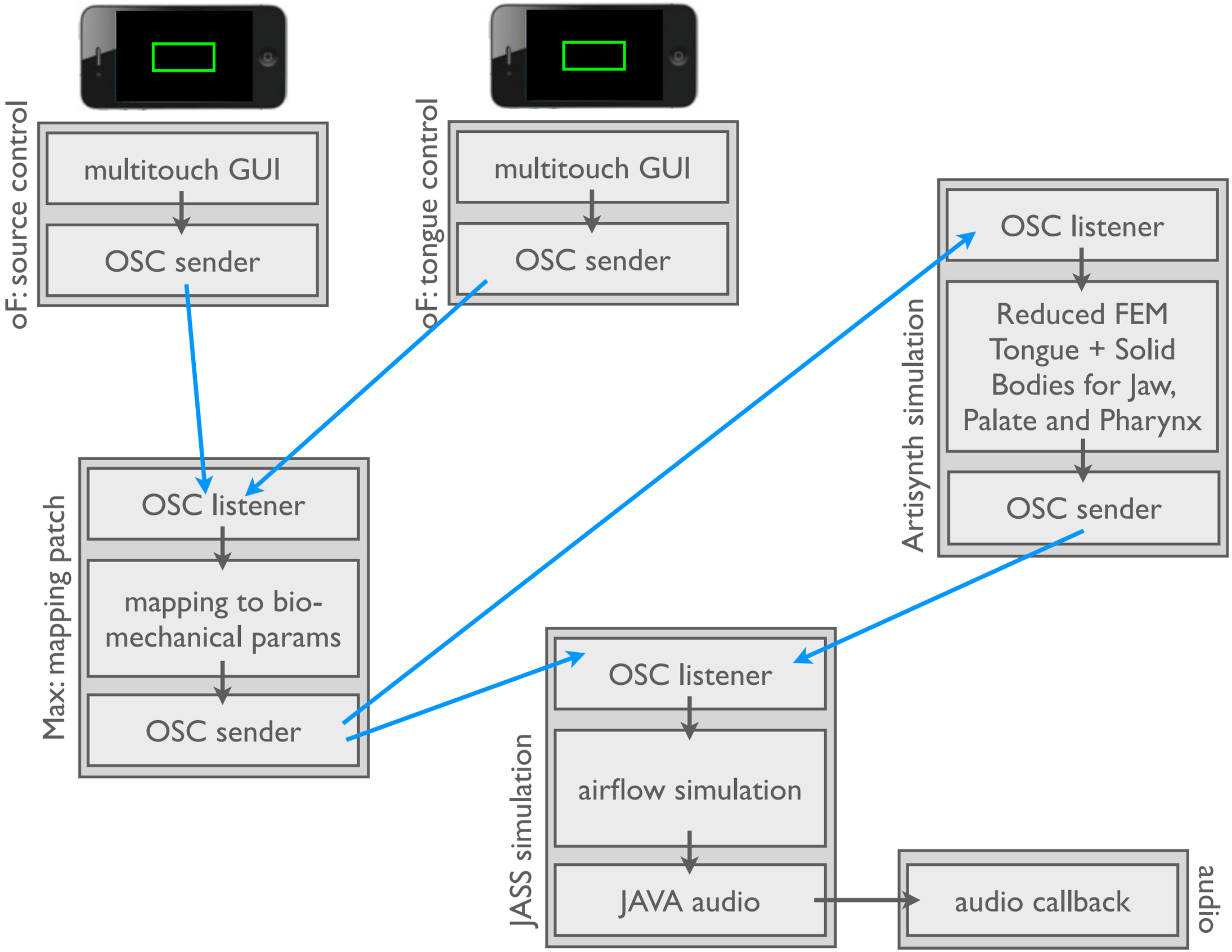
OSC sender

Max: mapping patch

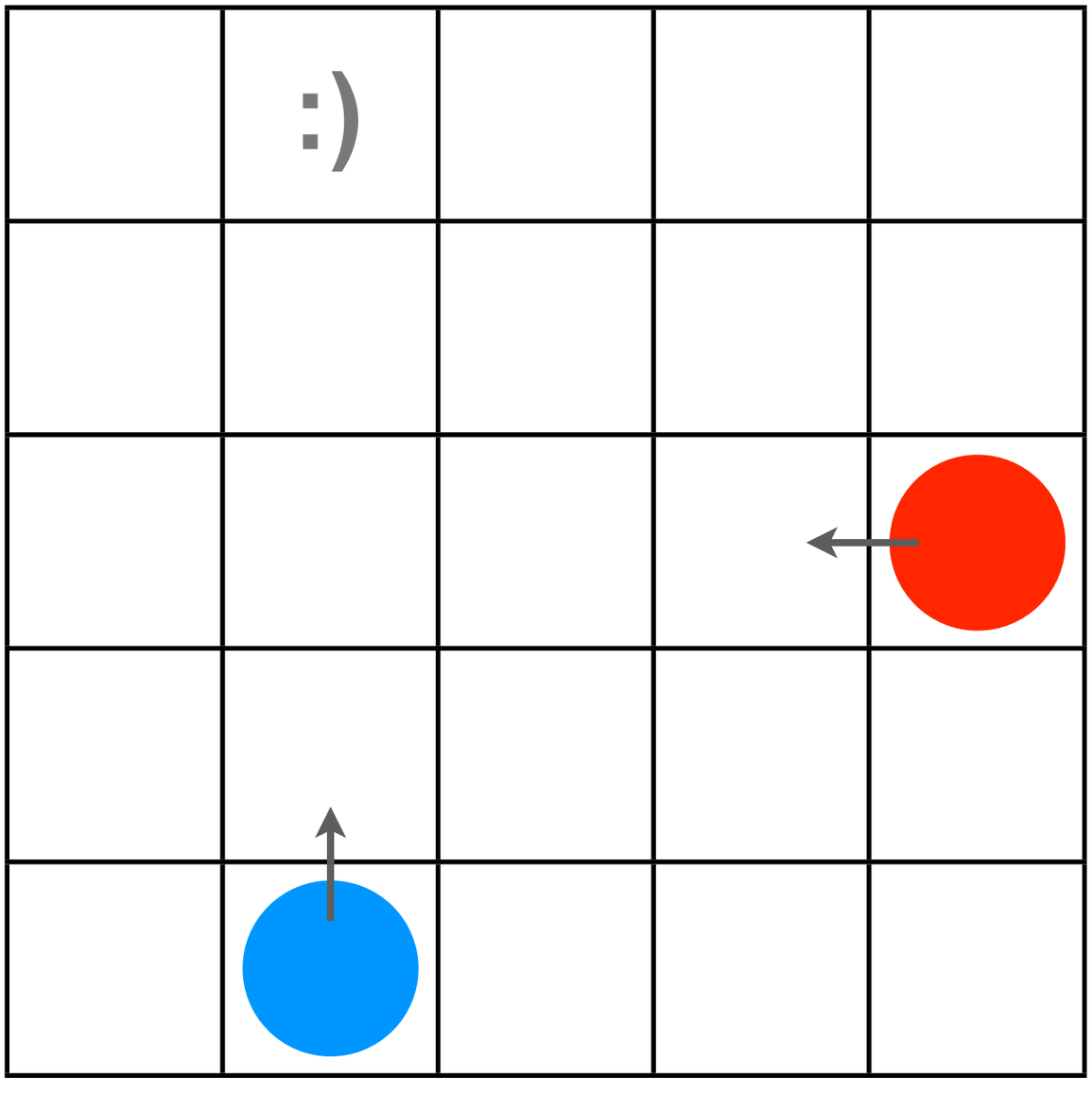
OSC listener

mapping to bio-mechanical params

OSC sender


oF: source control
multitouch GUI
OSC sender
oF: tongue control
multitouch GUI
OSC sender
Artisynth simulation
OSC listener
Reduced FEM Tongue + Solid Bodies for Jaw, Palate and Pharynx
OSC sender
Max: mapping patch
OSC listener
mapping to bio-mechanical params
OSC sender
JASS simulation
OSC listener
airflow simulation
JAVA audio
audio
audio callback

Game



- simple language to tell directions, colours and validate action or not
- one user is instructed to move coins by other who is talking through the device
- we record all trajectories achieved by the "instructor" and time to achieve the task

esf european social fund in the czech republic
EUROPEAN UNION
MINISTRY OF EDUCATION, YOUTH AND SPORTS
OP Education for Competitiveness
UNIVERSITY OF WEST BOHEMIA
INVESTMENTS IN EDUCATION DEVELOPMENT

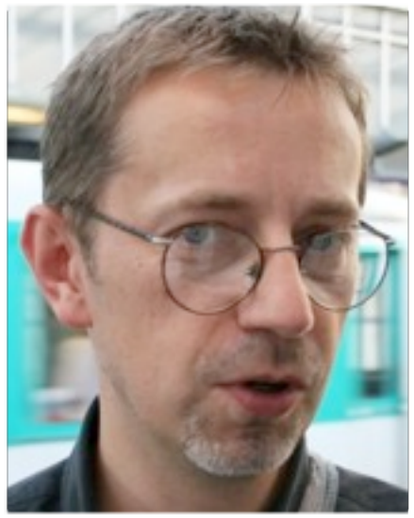
Game

WORD	MEANING	GESTURE
ale	red	back open/middle closed/front open
ela	blue	front open/middle closed/back open
lau	up	back closed/back open/back closed
lua	down	back open/back closed/back open
lea	left	front closed/front open/back open
lae	right	back closed/back open/front open
ea	yes	front closed/back open
ua	no	back closed/back open

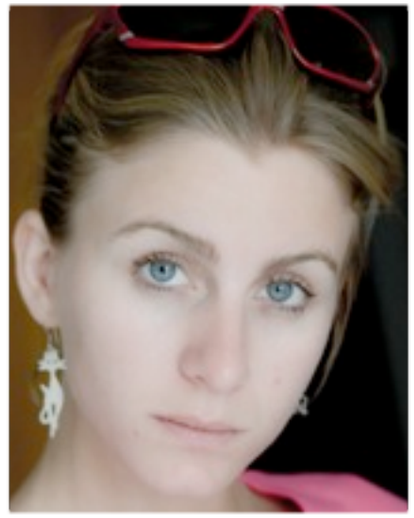
esf european social fund in the czech republic
EUROPEAN UNION
MINISTRY OF EDUCATION, YOUTH AND SPORTS
OP Education for Competitiveness
UNIVERSITY OF WEST BOHEMIA
INVESTMENTS IN EDUCATION DEVELOPMENT

Team

Nicolas
Alessandro



Thierry
Dutoit



Maria
Astrinaki

thank you guys !!!

Johnty
Wang

Àngel
Calzada

Onur
Babacan

thanks for your attention

come and see/try our demos

esf european social fund in the czech republic
EUROPEAN UNION
MINISTRY OF EDUCATION, YOUTH AND SPORTS
OP Education for Competitiveness
UNIVERSITY OF WEST BOHEMIA
INVESTMENTS IN EDUCATION DEVELOPMENT