Orchestra of Speech

Daniel Formo
Musician/composer/research fellow
Norwegian University of Science and Technology

IDEAS

Ideas

Close relationship between *Speech* and *Music*

Prosody: *"the music of everyday speech"*

Improvisation as dialogue

Conversation = improvisation

Bakhtin: Speech genres correspond to *social situations*

Expression gives utterances *context* and thereby *meaning/sense:*

- Rubato : personal, intimate
- Monotonous: formal, representative
- High register, dynamic: to children (or patronizing)

Improvisation = social situation

Source material: *conversations*

Speech genres / social settings

smalltalk; telephone call; interrogation; quarrel; confession; public debate; baby talk; pillow talk; gossiping; etc.

Sources

public recordings, own recordings, reality-TV, linguistic corpora, surveillance recordings, telephone recordings, documentaries.

Context

Speech and music – long tradition

Antique rhetoric – Baroque klangrede – Electroacoustic music

Current field

Cathy Lane: lists 19 different ways speech has been used in music

Peter Ablinger; Paul Lansky; Robert Ashley; Laurie Anderson; Jaap Blonk; Alessandro Bosetti; Trevor Wishart; Steve Reich; Jason Moran; Jonathan Harvey; Glenn Gould, Aperghis, Berio, Ligeti, Schwitters, Eimert, Stockhausen; a.o.

The "Orchestra"

System for real time *"orchestration"*

Idea-driven system

System for analysing and deconstructing musical structures from speech, abstracted, transformed and orchestrated as music in real time.

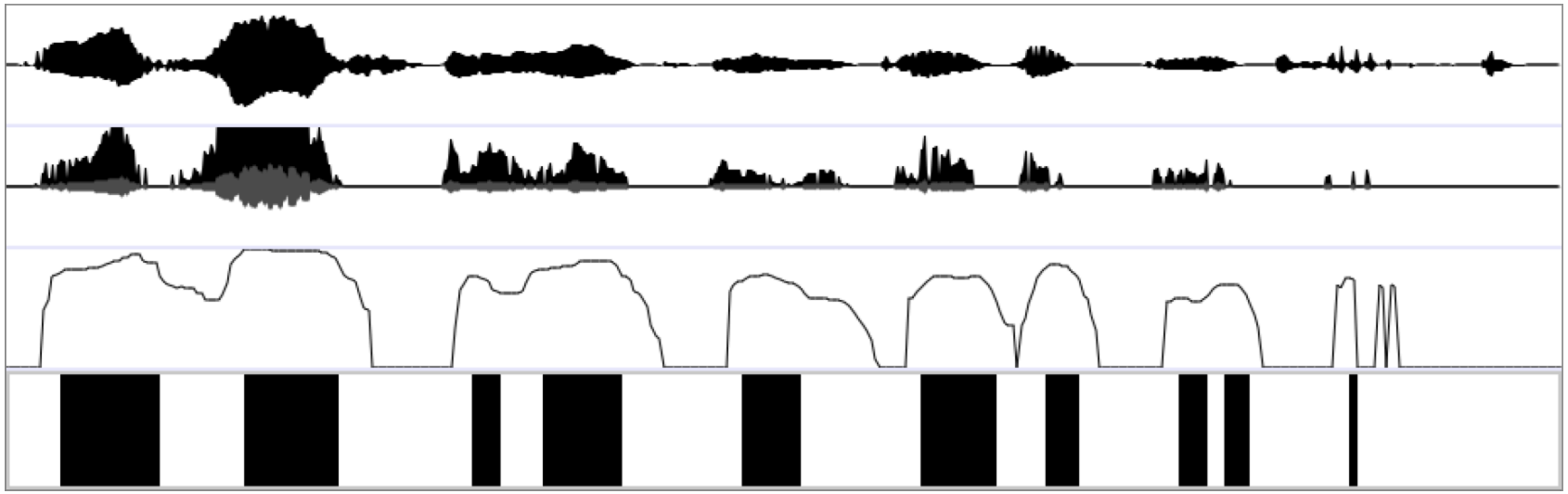
Instrument

Instrument-like system (musical and scientific)

Software environment

Developed in Max/MSP using **FTM&Co** and **Mubu** libraries from **IRCAM**

Analysis



Epoch-based syllable segmentation based on zero frequency resonator filtering
(B. Yegnanarayana and S. V Gangashetty, "Epoch-based analysis of speech signals")

Analysis

Segmentation: syllables (vowels); stress; accents; phrases; pauses; sibilants

Descriptors: Pitch (f0); centroid; spectrum (fft/lpc);
partials (harmonics, formants), voice quality (harmonic slope)

Transformation:
Cepstral smoothing, frequency/amplitude smoothing of spectrum/partials
Transposition, scaling, arpeggiating of partials
Tempo detection, tempo smoothing, quantization of segments
Secondary voices: diminution(subdivision)/augmentation etc

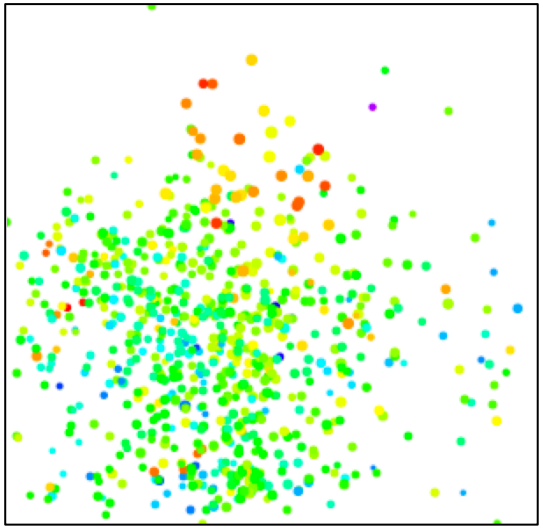
Corpus organization

Speech: organized *thought*

Music: organized *sound*

Alternative organization

- Association / Similarity
- Repetition / variation
- Fragmentation



Syllables in a space of pitch (vertical),
speech rate (horizontal), and vocal effort (colour)

Likelihood and machine learning

Machine learning in improvisation

Huge field, many approaches

Markov chains

Probable variations of sequences

Generalized phrases, no identifiable source

Probaility and improvisation

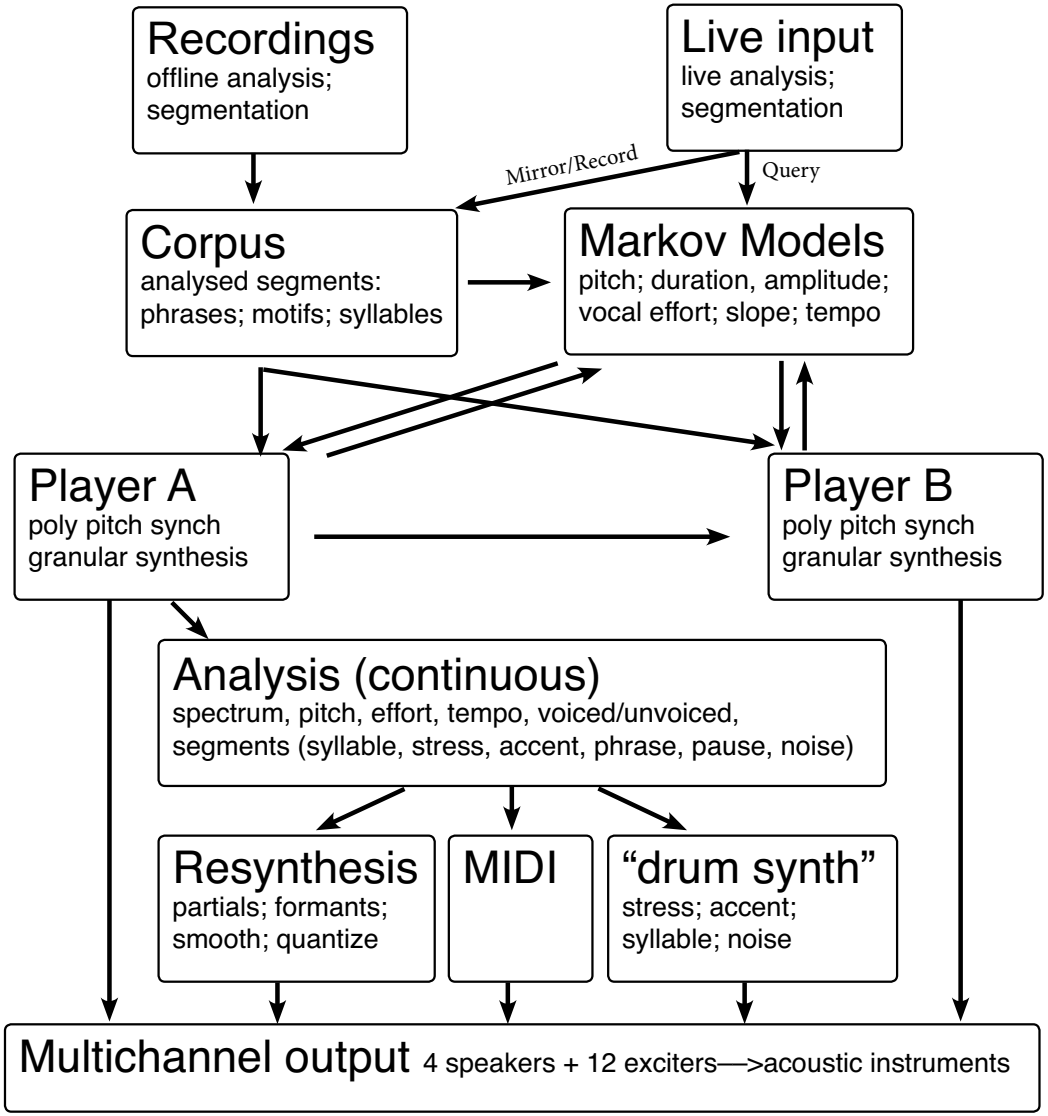
Dadaist speech-recognition system

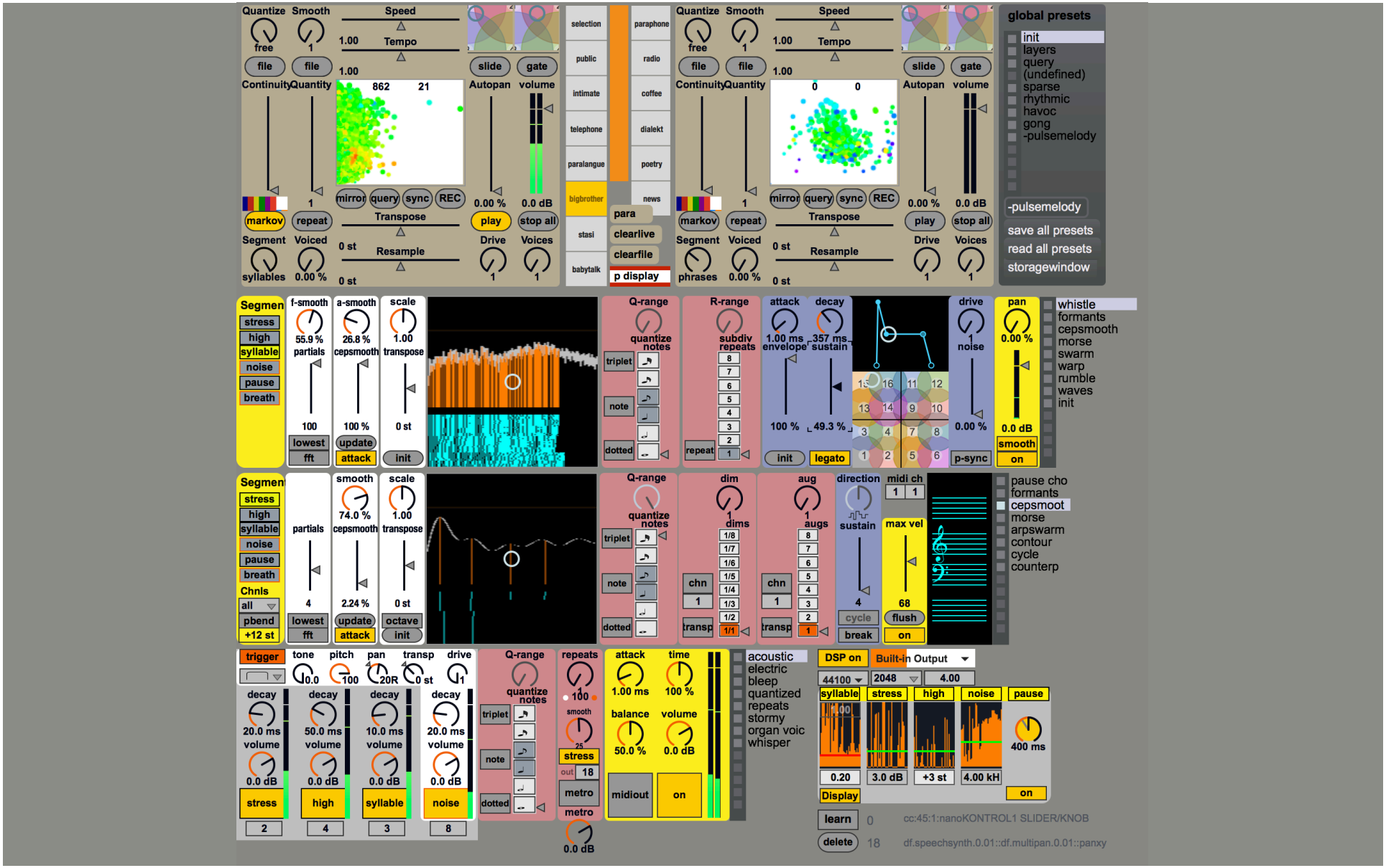
- Provides probable prosodic/musical responses/continuations
- Does not care about semantic content

Interaction/query = unknown response

- Facilitates improvisation
- Meta-dialogue between speech and music

HIWATT


Recordings
offline analysis;
segmentation
Live input
live analysis;
segmentation
Mirror/Record
Query
Corpus
analysed segments:
phrases; motifs; syllables
Markov Models
pitch; duration, amplitude;
vocal effort; slope; tempo
Player A
poly pitch synch
granular synthesis
Player B
poly pitch synch
granular synthesis
Analysis (continuous)
spectrum, pitch, effort, tempo, voiced/unvoiced,
segments (syllable, stress, accent, phrase, pause, noise)
Resynthesis
partials; formants;
smooth; quantize
MIDI
“drum synth”
stress; accent;
syllable; noise
Multichannel output 4 speakers + 12 exciters—>acoustic instruments


Quantize
Smooth
Speed
Tempo
file
Continuity
Quantity
slide
gate
Autopan
volume
mirror
query
sync
REC
markov
repeat
Transpose
play
stop all
Segment
Voiced
Resample
Drive
Voices
syllables
phrases
selection
public
intimate
telephone
paralangue
bigbrother
stasi
babytalk
paraphone
radio
coffee
dialekt
poetry
news
para
clearlive
clearfile
p display
global presets
init
layers
query
(undefined)
sparse
rhythmic
havoc
gong
-pulsemelody
save all presets
read all presets
storagewindow
Segmen
stress
high
syllable
noise
pause
breath
f-smooth
a-smooth
scale
partials
cepsmooth
transpose
lowest
update
attack
init
fft
Q-range
quantize notes
triplet
note
dotted
R-range
subdiv repeats
repeat
attack
decay
envelope
sustain
legato
drive
noise
p-sync
pan
smooth
on
whistle
formants
cepsmooth
morse
swarm
warp
rumble
waves
init
Chnls
all
pbend
+12 st
octave
dim
dims
aug
augs
chn
transp
direction
sustain
cycle
break
midi ch
max vel
flush
pause cho
arpswarm
contour
cycle
counterp
trigger
tone
pitch
transp
repeats
balance
time
midiout
acoustic
electric
bleep
quantized
repeats
stormy
organ voic
whisper
DSP on
Built-in Output
44100
2048
4.00
Display
learn
delete
cc:45:1:nanoKONTROL1 SLIDER/KNOB
df.speechsynth.0.01::df.multipan.0.01::panxy

SOUND

Sound realms

Sound = meaning

- Speaker mediated sound (virtual space) vs acoustic sound (real space)
- Voice (particular person) vs instrument (generic object)

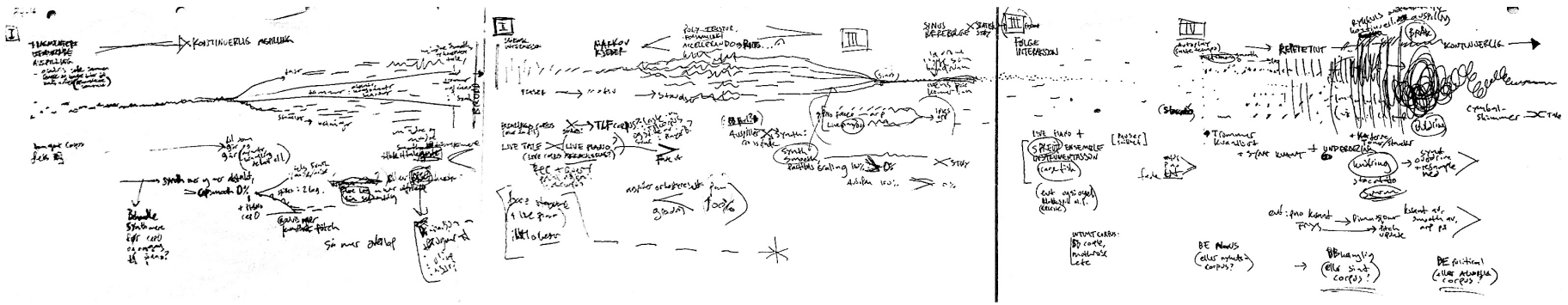
Acoustic/electroacoustic/digital hybrid

- Acoustic instrument as loudspeakers – musical framing of sound
- Embodiment of disembodied voices

Orchestra

- Both metaphorical (*"orchestrated"*) and actual orchestra

MUSIC



Musical ideas

Prosodic phenomena, musical character, instrumentation, musical ideas, technical possibilities, sound combinations, etc

Continuum between continuously fluctuating character of speech and the discrete and stylized sounds associated with formalized music (e.g. piano)

Perception

Focus of perception

- Semantic vs musical content (different brain halves?)

Time perception

- Narrative forward motion (story)
- Standstill (space rather than time)
- Attentive (*"accentuated present"*, alert interaction)
- Repetition (cyclical time)

Sound / Source

- Gesture (intent), utterance
- Speaker/voice/instrument/acoustic/synthetic/space
- Formalized/stylized music vs flowing continuous stream of speech

orchestraofspeech.com

daniel.formo@ntnu.no