Real-time speech synthesis for neural prosthesis
Robot Speech Session, ASA 2009, Portland

1,2,3 Frank Guenther and 1,4 Jonathan Brumberg

1 Cognitive and Neural Systems, Boston University, Boston, MA
2 Division of Health Sciences and Technology, Harvard University / M.I.T.
3 Research Laboratory of Electronics, Massachusetts Institute of Technology
4 Neural Signals, Inc., Duluth, GA

brumberg@cns.bu.edu

May 18, 2009
Artificial speech synthesis

- Von Kempelen: tried to mimic and replicate speech sounds with a wind “instrument”
- Fant: electrical circuits for synthesis of speech sounds

Typically used for: speech imitation by able bodied persons, theoretical investigations of speech or commercial applications

**Our Research:** To convert brain activity into a reference frame suitable for control of an artificial speech synthesis device
Study History

**Goal:** restoration of speech production for profoundly paralyzed individuals through microelectrode brain-computer interface

Subject:
- 25 year old male
- Brain stem stroke suffered during automobile accident at age 16
- Suffers from Locked-in Syndrome

Locked-in Syndrome:
- Total paralysis
- Cognition and sensation spared
- Retain weak eye or facial movements
- Causes: brain trauma, stroke, disease (i.e. ALS)
Goal: restoration of speech production for profoundly paralyzed individuals through microelectrode brain-computer interface

Subject:
- 25 year old male
- Brain stem stroke suffered during automobile accident at age 16
- Suffers from Locked-in Syndrome

Locked-in Syndrome:
- Total paralysis
- Cognition and sensation spared
- Retain weak eye or facial movements
- Causes: brain trauma, stroke, disease (i.e. ALS)
Monkeys first:
- Schwartz group: Computer cursor control and control of multidimensional robotic arm
- Nicolelis group: Control of robotic arm and legs

Humans
- Kennedy group: First human implant; computer cursor control with word and letter selection
- Donoghue group: Computer cursor control with word and letter selection

Noninvasive techniques exist primarily for word spelling: (EEG) Wolpaw, Sellers; (EMG) Impulse (Ablenet, Inc; Kennedy group)
Implant Details

**Electrode**: Neurotrophic Electrode (Kennedy et al., 1987-present)
- Extracellular microwire design with 3 wires encased in glass cone; filled with neurotrophic growth factor
- Chronic implant (4.5 years to date)

**Implanted electronics**: acquire, amplify and wirelessly transmit brain activity

![Neurotrophic Electrode](image)
Potential sites were identified using functional imaging of the brain (fMRI)

- Peak activity in the ventral motor cortex
  - Left: orange blob between red and yellow lines
  - Right: electrode wire comes out of the brain between red and yellow lines
Speech production is motor control

- We can utilize current accepted methods for motor control BCI (PVA, Linear filtering, Kalman filtering, etc)
- Implant captures motor cortical output from brain to speech articulators

Goal: decode speech motor output into reference frame suitable for artificial speech synthesis
Speech production is motor control

- We can utilize current accepted methods for motor control BCI (PVA, Linear filtering, Kalman filtering, etc)
- Implant captures motor cortical output from brain to speech articulators

Goal: decode speech motor output into reference frame suitable for artificial speech synthesis

**Problem 1**: what is the best representation of speech motor cortical activity?

**Problem 2**: animal models for speech motor control do not exist.
Control System

Speech production is motor control

- We can utilize current accepted methods for motor control BCI (PVA, Linear filtering, Kalman filtering, etc)
- Implant captures motor cortical output from brain to speech articulators

Goal: decode speech motor output into reference frame suitable for artificial speech synthesis

Problem 1: what is the best representation of speech motor cortical activity?
Problem 2: animal models for speech motor control do not exist.

Solution: DIVA model predicts acoustic information encoded by the brain region around the implant, specifically formant frequency.
Formant Representation

Formant frequencies are the resonant frequencies of the vocal tract

- Appear as peaks in the acoustic spectrum
- Only first two are needed for vowel synthesis (e.g. Fant, Klatt)
- Relatively easy for brain control signal (e.g. arm/hand kinematic BCI)
- 2D decoding is computationally more efficient
- Supported by computational model of speech production (DIVA)
Initial Progress: Verify Formants

Open-loop task for vowel sequence production in a “say-along” paradigm

Vowels IY, A, OO and UH used in a center-out and back paradigm; 8 repetitions each vowel, 64s total production

Offline ridge regression analysis found significant correlations between firing rates and formant frequency

- F1 (r=0.49, p<0.001); F2 (r=0.57, p<0.001)
- Black line: ensemble tuning curve
- Verifies DIVA prediction of formant tuning in implanted region
**Implant:** Face/vocal tract portion of motor cortex  
**Acquire:** Record signals wirelessly  
**Decoder:** Kalman filter  
**Output:** Computer synthesized speech  

Artificial speech is produced in real-time, < 50ms, and played over computer speakers
Real-time decoder performance
Quantitative Results

Within-session improvement (avg. of \(\sim 5\) trials per block):

Across-session improvement (25 sessions analyzed)

Real-time audio feedback allowed participant to improve productions with practice, both within a session and across sessions.
Work accomplished:

- First real-time BCI for continuous speech production
- First real-time control with a wireless telemetry
- First real-time control with implanted microelectrode for over 4.5 years

Immediate future step

- Increasing speech production capabilities: consonants, voicing control
- Fast, efficient articulatory vocal-tract synthesis
- Low dimensional control (2-5D)
Supported by the NIH and NSF

Collaborators

**Speechlab**
- Dr. Alfonso Nieto-Castanon
- Dr. Jason Tourville
- Misha Panko
- Robert Law

**Neural Signals, Inc.**
- Dr. Philip Kennedy
- Dinal Andreasen
- Dr. Princewill Ehrim
- Dr. Hui Mao
- E. Joe Wright
- Jess Bartels
- Steven Seibert