

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/41761650>

# Brain-Computer Interfaces for Speech Communication

Article in *Speech Communication* · April 2010

DOI: 10.1016/j.specom.2010.01.001 · Source: PubMed

---

CITATIONS

249

---

READS

922

4 authors, including:



**Alfonso Nieto-Castanon**

Computational Neuroscience Research Lab

98 PUBLICATIONS 10,736 CITATIONS

SEE PROFILE



**Frank Guenther**

Boston University

227 PUBLICATIONS 12,829 CITATIONS

SEE PROFILE

Published in final edited form as:

*Speech Commun.* 2010 April 1; 52(4): 367–379. doi:10.1016/j.specom.2010.01.001.

## Brain-Computer Interfaces for Speech Communication

Jonathan S. Brumberg<sup>a,d</sup>, Alfonso Nieto-Castanon<sup>e</sup>, Philip R. Kennedy<sup>d</sup>, and Frank H. Guenther<sup>a,b,c</sup>

Jonathan S. Brumberg: brumberg@cns.bu.edu; Alfonso Nieto-Castanon: ; Philip R. Kennedy: ; Frank H. Guenther:

<sup>a</sup>Department of Cognitive and Neural Systems, Boston University, 677 Beacon Street, Boston, MA, 02215, Telephone: (617) 353- 9481, Fax Number: (617) 353-7755

<sup>b</sup>Division of Health Sciences and Technology, Harvard University - Massachusetts Institute of Technology, Cambridge, MA 02139, USA

<sup>c</sup>Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA 02129, USA

<sup>d</sup>Neural Signals, Inc., Duluth, GA 30096, USA

<sup>e</sup>StatsANC LLC, Buenos Aires, Argentina

### Abstract

This paper briefly reviews current silent speech methodologies for normal and disabled individuals. Current techniques utilizing electromyographic (EMG) recordings of vocal tract movements are useful for physically healthy individuals but fail for tetraplegic individuals who do not have accurate voluntary control over the speech articulators. Alternative methods utilizing EMG from other body parts (e.g., hand, arm, or facial muscles) or electroencephalography (EEG) can provide capable silent communication to severely paralyzed users, though current interfaces are extremely slow relative to normal conversation rates and require constant attention to a computer screen that provides visual feedback and/or cueing. We present a novel approach to the problem of silent speech via an intracortical microelectrode brain computer interface (BCI) to predict intended speech information directly from the activity of neurons involved in speech production. The predicted speech is synthesized and acoustically fed back to the user with a delay under 50 ms. We demonstrate that the Neurotrophic Electrode used in the BCI is capable of providing useful neural recordings for over 4 years, a necessary property for BCIs that need to remain viable over the lifespan of the user. Other design considerations include neural decoding techniques based on previous research involving BCIs for computer cursor or robotic arm control via prediction of intended movement kinematics from motor cortical signals in monkeys and humans. Initial results from a study of continuous speech production with instantaneous acoustic feedback show the BCI user was able to improve his control over an artificial speech synthesizer both within and across recording sessions. The success of this initial trial validates the potential of the intracortical microelectrode-based approach for providing a speech prosthesis that can allow much more rapid communication rates.

---

© 2010 Elsevier B.V. All rights reserved.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## SECTION 1: INTRODUCTION

Recent advances in electrophysiological recording technology are for the first time making viable, silent speech communication devices a reality. However, not all technologies are appropriate in every situation. When considering a method that provides artificial communication to a speech-deprived individual, one must identify the most efficient means given the nature of the individual's impairment. Some methods rely on non-vocalized articulator movements or other sub-vocalizations (Betts and Jorgensen, 2006; Fagan et al., 2008; Jorgensen et al., 2003; Jou et al., 2006; Jou and Schultz, 2009; Maier-Hein et al., 2005; Mendes et al., 2008; Walliczek et al., 2006; Wand and Schultz, 2009) which can be helpful for speech deprived individuals (e.g. laryngectomy patients). Unfortunately, these options are not feasible for profoundly paralyzed individuals such as those suffering from *locked-in syndrome*, which is characterized by near-complete paralysis. Locked-in patients often retain slow eye movement or eye blink control which can be used to answer simple yes/no questions, but they are completely deprived of any other voluntary motor behavior, including speech and movements of the speech articulators. Consequently, few speech restoration options are available for these individuals. Of these sparse options, intracortical microelectrode-based brain computer interfaces (BCIs) utilizing the electrical signals of individual neurons offer great promise for conversational speech restoration.

Current methods for silent speech available to neurologically normal individuals or those with damaged vocal tracts typically involve the placement of surface electromyographic (EMG) electrodes on the orofacial and laryngeal speech articulators (Betts and Jorgensen, 2006; Fagan et al., 2008; Jorgensen et al., 2003; Jou et al., 2006; Jou and Schultz, 2009; Maier-Hein et al., 2005; Mendes et al., 2008; Wand and Schultz, 2009). Electrical recordings are transmitted from the electrodes to an analysis computer which has been trained to recognize a small vocabulary of words based upon the speaker's EMG pattern. Under this scenario, the silent speech system effectively performs automatic speech recognition utilizing articulatory information rather than acoustic information. However, this approach fails in situations where the user has no voluntary control over the speech articulators. An alternative method was developed for paralyzed patients with minor voluntary muscle control for use in an augmentative and alternative communication (AAC) device (Wright et al., 2008). The method uses residual EMG activity from non-speaking paralyzed patients to drive a binary click operation for use in AAC devices, with expert users achieving spelling rates of over 20 characters per minute.

Unfortunately, many locked-in patients do not have reliable control of the musculature needed to successfully operate even minimal EMG communication devices. Such individuals can only be helped by true *brain computer interfaces* (BCIs); those capable of direct control of communication devices solely through observed changes in neural activity. There are many types of BCI devices for communication though they can be roughly classified into two groups: noninvasive (e.g. Birbaumer et al., 1999; Cheng et al., 2002; Donchin et al., 2000; Suppes et al., 1997; Wolpaw et al., 2000) and invasive (e.g. Hochberg et al., 2006; Kennedy et al., 2004; Leuthardt et al., 2004). Non-invasive techniques typically involve measurements of brain activity through the electroencephalogram (EEG) or magnetoencephalogram (MEG), while invasive techniques utilize the electrocorticogram (ECoG), local field potentials (LFPs) and single unit activity (SUA). Both the EEG and MEG signals are captured by a passive recording system with underlying signals based on the summed synchronous electrical activity of tens of thousands to millions of neurons in the local region near the recording sensor. Recording the ECoG, LFP and SUA require invasive surgical procedures involving a craniotomy and placement of the recording electrode on the surface of the cerebral cortex (ECoG) or insertion of the extracellular microelectrode into the cerebral cortex (LFP and SUA). ECoG signals represent the cumulative activity of hundreds to thousands neurons underneath the electrode

recording surface, the LFP signal represents the total activity of many tens of neurons in the immediate vicinity of the electrode, and the SUA represents individual neural units.

Brain computer interfaces have been developed to aid communication for impaired populations utilizing each of the neurobiological signals described above. Each signal type, and associated recording methodology, has advantages and disadvantages for successful use as a BCI for augmentative communication. In particular, the longevity of the sensing device is a crucial factor when considering a *chronic* BCI, which is required for long-term use by the target population. The remainder of this paper will discuss the state of current electrode technologies for chronic implantation and associated methods for silent communication utilizing EEG, MEG, ECoG, LFP and SUA. Particular attention will be paid to human intracortically implanted electrodes measuring the SUA, specifically the Utah microelectrode array (Maynard et al., 1997) and the Neurotrophic Electrode (Bartels et al., 2008; Kennedy, 1989). In addition, different approaches to the problem of silent speech communication will be described, including options for discrete, choice-based protocols and continuously varying artificial speech synthesis. Finally, areas of BCI improvement are discussed with particular emphasis on two main objectives: to improve signal quality and predicted speech and to convert immobile, laboratory-based equipment into robust and portable alternatives.

## SECTION 2: CURRENT ELECTRODE TECHNOLOGY

Viable chronic brain computer interfaces must utilize a recording apparatus able to record neural signals continuously over long durations (e.g. many years) and be relatively convenient or portable. Each neurobiological signal described in Section 1 is capable of being used for BCI applications but only a subset meet longevity and portability requirements. A summary of electrode technology in terms of characteristics of a successful chronic BCI is listed in Table 1, and each methodology is discussed in detail below. The study type, in Table 1, indicates whether or not the recording methodology has been used in human applications. The duration of each recording technology is taken from the literature of existing studies (Neurotrophic Electrode; Bartels et al., 2008; Utah array; Hochberg et al., 2008; Michigan array; Kipke et al., 2003; microwire arrays; Nicolelis et al., 2003), though improvements have likely been made in unpublished results. The final three categories (cross-scalp signal transmission, mechanical stability and fabrication) describe characteristics that are pertinent to the long-term use of each technology in chronic BCI applications. The mechanical stability of each technology is classified as fixed, floating, embedded or relatively movable. Fixed and floating techniques are relatively straightforward but the embedded and relatively movable classifications require clarification.

EEG and ECoG signals are obtained through similar recording techniques. Both utilize an electrode placed on the recording surface, and the summed electrical activity directly below the electrode is captured. The two technologies differ with respect to recording surface type (EEG records scalp potentials while ECoG records cortical surface potentials) though both primarily reflect the electrical activity of synchronously firing pyramidal cells. The EEG signal contains information in a relatively narrow frequency band (0–50 Hz) and is non-invasively obtained by affixing Ag/AgCl electrodes to the scalp with a conductive gel or paste. The ECoG contains information with larger bandwidth (0–500 Hz) and greater amplitudes from platinum electrodes placed directly on the surface of the cortex. Necessarily, the brain must be exposed via craniotomy. Both technologies have the potential for long-term use in a chronic BCI. For example, EEG electrodes can be placed indefinitely on the surface of the scalp (needing to be adjusted only to maintain appropriate electrode-skin conductance and impedances) and have been used for long-term continuous use in previous BCI applications (e.g. Birbaumer et al., 1999). Additionally, many research groups are investigating ECoG for BCI applications and

have claimed that the ECoG electrode grid possesses characteristics suitable for stable long-term recording (Leuthardt et al., 2004; Schalk et al., 2007, 2008).

MEG (Hamalainen et al., 1993) is related to EEG and is derived from a similar neural source, though the main component of the MEG signal is likely due to synchronous intracellular currents flowing tangentially to the cortical surface (i.e. within sulci). The MEG signal is obtained through measurement of small magnetic fields induced by intracellular ionic currents. These tiny magnetic fields are non-invasively measured using a superconducting quantum interference device (SQUID) which is placed within a liquid helium cooled helmet. In addition, the magnetic fields are so small that their detection is impossible outside of an electromagnetically sealed environment. The MEG apparatus can record continuously without adjustment, but requires that the patient or subject remain completely still within the sensor helmet. Due to its cost and inconvenience (e.g. electromagnetically sealed environment, subject movement constraints, etc.), MEG is a sub-optimal solution for chronic BCI applications.

The LFP and SUA, for BCI applications, are both measured utilizing intracortical extracellular microelectrodes, which are inserted into the cerebral cortex after exposing the brain through a craniotomy. Microelectrode technology has a long history in neuroscience applications but is only recently becoming specialized for simultaneous recording of multiple neural units over long periods of time. These electrodes represent a promising area of current and future chronic BCI research.

Many electrode technologies exist for recording extracellular potentials from groups of neurons in the cerebral cortex. However, most of these electrodes are simply not suitable for chronic human implantation or generally for use in clinical applications. Extracellular electrode design formerly was aimed at precise identification and isolation of single neurons; viable long-term recording of electrical potentials for scientific study was not considered. These single-contact electrodes are limited in the number of simultaneously recorded single units to one or two, though the isolation of the single units is often quite precise and unambiguous. Utilizing these electrodes, it is possible to identify many neurons by recording from the same animal subject over the course of several months, though recordings from any one neuron last only on the order of minutes to hours. Common recording paradigms involve repeated insertions into the cortical surface of interest of the subject animal. The electrode is then driven slowly into the cortex in very small increments (5–15 nm per step) using a microdrive, leading to the discovery of new neurons at different cortical depths (Mountcastle et al., 1975). Unfortunately, the limitation on the number of simultaneously recorded neurons and the relative size of the entire electrode drive renders movable single electrodes insufficient for BCI designs.

Alternatively, extracellular electrodes permit the simultaneous recording of multiple single units, or multi-units, and have long been utilized in chronic recording paradigms (Schmidt et al., 1976). Multi-unit electrodes are typically capable of simultaneous recordings of multiple single units. These simultaneously recorded potentials represent additive combinations of neuroelectrical activity observed at the electrode tip. Acquisition of these signals is analogous to the effect observed in a cocktail-party problem; in the case of multi-units the individual neural sources are the many speakers in the cocktail-party room and the electrode tip is the microphone. Unfortunately, multi-unit recordings require intensive processing to correctly identify distinct single units (i.e. SUA), which is not always possible. Despite this issue, extracellular multi-unit electrodes are currently the only type of electrode capable of simultaneous long-term recordings of many (e.g. tens to hundreds) neural units. In addition, extracellular multi-unit electrodes are capable of recording the LFP signal, which represent the summed activity of a local population of neural units around the electrode tip (Mitzdorf, 1985). As such, they are now being extensively implanted in animal subjects (Carmena et al., 2003; Kennedy et al., 1992a; Nicolelis et al., 2003; Taylor et al., 2002; Velliste et al., 2008;

Wessberg et al., 2000) and in some human volunteers (Bartels et al., 2008; Guenther et al., in press; Hochberg et al., 2006; Kennedy et al., 2004; Kim et al., 2007) for experimental BCI investigations to restore artificial movement and communication.

Recent advances in multi-unit electrode technology have improved upon two major design components for use in human BCI devices: *stability* and *longevity*. Both are critical for sustained functionality of a chronic neural implant for speech restoration. In this context, stability refers to the consistent isolation and identification of recorded single units, while longevity refers to recording viability over long durations (many years). Chronic implantation BCI investigations in monkeys and humans currently comprise four major classes of multi-unit electrodes: microwire arrays (Nicoletis et al., 2003; Taylor et al., 2002; Williams et al., 1999), silicon planar arrays (Hoogerwerf and Wise, 1994; Wise et al., 1970), silicon etched arrays (Jones et al., 1992; Maynard et al., 1997; Rousche and Normann, 1998) and neurotrophic microwires (Kennedy, 1989; Kennedy et al., 1992b). These four electrode types are shown in Figure 1 (images reprinted with permission). However, only two chronic electrodes have reported implantation durations in human subjects of two or more years, the silicon Utah microelectrode array for 2+ years (Hochberg et al., 2008) and the Neurotrophic Electrode for 4+ years (Bartels et al., 2008; Guenther et al., in press; Kennedy, 2006).

Of the multi-unit electrodes described above, the Neurotrophic Electrode device is the only chronic microelectrode platform for human subjects that makes use of implantable hardware for signal pre-amplification utilizing wireless power induction and transcutaneous wireless signal transmission over standard frequency modulated (FM) radio. Implanted electronics are a necessity for practical, portable and permanent usage of multi-unit electrodes for silent communication BCIs. Wireless signal transmission via FM radio eliminates the need for percutaneous wired connectors, reducing the risk of infection by allowing the scalp incision over the implanted area to completely heal. In addition, a wireless architecture is more aesthetically pleasing; there are no wires protruding from a plug in the BCI recipient's head. Though the benefits of silent communication devices likely outweigh the aesthetic costs, improvements in outward appearance of the device are also worthy of consideration.

A summary of all invasive and non-invasive recording systems is given in Table 1. Of the invasive options, the properties of the Neurotrophic Electrode implant make it uniquely suited for application to a brain computer interface for speech production. It has been used for chronic recording studies for over four years in human (Bartels et al., 2008; Kennedy, 2006) until patient death from underlying ailments unrelated to the implantation procedure. The Neurotrophic Electrode is currently being used in an intracortical neural prosthesis for speech restoration involving a 26 year old locked-in subject (Brumberg et al., 2009; Guenther et al., in press). At the time of this writing, the original implanted electrode remains viable after over 4 years of continuous use.

### SECTION 3: BRAIN-COMPUTER INTERFACES FOR COMMUNICATION

As mentioned in previous sections, silent speech communication for profoundly paralyzed individuals can be achieved utilizing scalp surface-based electrodes (e.g. EEG), cortical surface electrodes (e.g. ECoG) or intracortical microelectrodes (e.g. Utah array or Neurotrophic Electrode). Additionally, the methods for communication can differ between discrete (e.g. typewriter) or continuous control paradigms. In the discrete paradigm, the user can select between two or more discrete "choices", for example choosing a particular key on a "virtual keyboard" on a computer screen. In a continuous control paradigm, a small number of continuous kinematic variables (for example, the x and y coordinates of the cursor position on the screen, or the values of the first two formant frequencies for a speech prosthesis) are controlled by the user.



Much prior work has been accomplished for BCI augmented communication applications using EEG: slow cortical potentials (SCP; Birbaumer et al., 1999, 2000, 2003; Hinterberger et al., 2003; Kübler et al., 1999), P300 signal (Donchin et al., 2000; Krusienski et al., 2006, 2008; Sellers et al., 2006; Vaughan et al., 2006), sensorimotor rhythms (SMR; Vaughan et al., 2006; Wolpaw et al., 2000; Wolpaw and McFarland, 2004) and steady state visual evoke potentials (SSVEP; Allison et al., 2008; Cheng et al., 2002; Trejo et al., 2006). Though these techniques provide silent communication to paralyzed users, they are not currently capable of operating at rates fast enough for conversational or near-conversational speech, with production rates on the order of just one word or fewer per minute. In addition, each of the EEG interfaces described above require accurate visual perceptual abilities for the user to choose letters on a graphical display through a finite choice visual feedback system. The displays typically consist of keyboard letters presented on a screen while the user changes their electrophysiological activity to indicate a “yes” choice. Such a visually dependent system is entirely impractical for users with such severe paralysis that visual perception is unreliable.

Each of the EEG signals used for the interfaces discussed also requires “mapping” or translation into the speech domain, though others have conducted preliminary investigations of direct prediction using speech related potentials with varying degrees of success (DaSalla et al., 2009; Porbadnigk et al., 2009; Suppes et al., 1997). For instance, users of the SCP-based devices must be trained to modulate their SCP for use in the binary selection task. The SCP themselves do not necessarily represent the value of the intended selection; rather it is simply a neurophysiological signal which has been shown useful to perform binary selection tasks. Similarly, the P300 and SSVEP responses do not encode the meaning of the items selected by the user. Instead, both indicate the presence of a visual stimulus to which the user is attending. The EEG SMR, though, is directly related to the activity of the motor cortex. These rhythms, also obtained through ECoG, have been shown to represent the movements of peripheral and orofacial end-effectors (e.g. arm, hand, tongue, jaw, etc.) and have been successful for BCIs providing one and two dimensional cursor control (Leuthardt et al., 2004; Schalk et al., 2007, 2008). Such an interface can only be used in an AAC device by selection of items on a computer screen (Vaughan et al., 2006). Discrete “typewriter” approaches through mouse cursor control have also been implemented using intracortical electrodes, as discussed below.

Kennedy and colleagues were the first to implant a human subject with a chronic microelectrode for the sole purpose of restoring communication to a paralyzed volunteer by intracortical BCI (Kennedy and Bakay, 1998). Their first subject implanted with the Neurotrophic Electrode, MH, was able to make binary choice selections until her death 76 days after implantation, though no attempt at “typewriter” control was made. In 1998 another locked-in subject, JR, was implanted in the hand area of his primary motor cortex (Kennedy et al., 2004). JR learned to control the 2D position of a mouse cursor over a virtual keyboard, and to select desired characters via neural activity related to imagined movements of his hand. Though this technique for typewriter control is quite intuitive, JR was only able to produce 3 characters per minute (Kennedy et al., 2004).

Donoghue and colleagues have also chronically implanted human volunteers with their Utah microelectrode array-based system (Hochberg et al., 2006; Kim et al., 2007) in a continuous control scheme. At least two subjects learned to control the position of a mouse cursor on a computer screen. Cursor position was determined by mapping the neural firing rates of single and multi-unit neural sources into movement kinematics, specifically the position and velocity of the mouse cursor. The system is designed to allow learned artificial mouse control as an interface for virtual devices such as keyboards (similar to Kennedy et al., 2004) and other AAC applications.

Neural decoding of arm kinematics has been studied for nearly 30 years beginning with Georgopoulos et al. (1982) and recently by Velliste et al. (2008) in monkeys. Kinematic decoding is well-suited to artificial mouse control as well as control of arm prosthetic devices. Unfortunately, cursor control is a clumsy and slow interface for augmented communication when restricted to prediction of hand kinematics. We sought to address this constraint while exploiting the decades of research of kinematic prediction from motor cortical activity. Our key insight here lies in understanding that speech production can be characterized as a continuous motor control task, albeit a complex one, and that speech motor cortical signals can be used to control “movements” of an artificial speech synthesizer. Our recent research utilizing unit activity obtained from intracortical microelectrode recordings aims to overcome the hurdles faced by traditional BCIs for augmented speech and provide a means for silent communication interfaces capable of producing computer synthesized conversational speech.

### Decoding speech with the Neurotrophic Electrode implant

In December 2004, a locked-in brain stem stroke volunteer, ER, was implanted with the Neurotrophic Electrode in speech motor cortex (Bartels et al., 2008; Kennedy, 2006) with the primary goal of decoding the neural activity related to speech production and providing an alternative means for communication (Bartels et al., 2008; Brumberg et al., 2009; Guenther et al., in press). The implantation procedure was approved by the Food and Drug Administration (IDE G960032), Neural Signals, Inc. Institutional Review Board, and Gwinnett Medical Center Institutional Review Board. Informed consent was obtained from the participant and his legal guardian prior to implantation.

The specific implanted region of the motor cortex was very important for the success of the speech prosthesis implant. We sought to implant an area of motor cortex which was related to the movements of the speech articulators. In this way, we hypothesized that control of the continuous speech production BCI could be driven by speech articulation-related activity in the motor cortex rather than non-speech related neural activity (cf. EEG P300, SSVEP, SCP, etc.). Localization of the Neurotrophic Electrode implant involved first conducting a pre-operative fMRI study in which ER participated in imagined picture naming and word repetition tasks in an fMRI protocol. The task revealed increased BOLD response in much of the normal speech production network, and the implantation site was chosen as the area of peak activity on the ventral precentral gyrus (location of the speech motor cortex). Details of the implantation procedure can be found elsewhere (Bartels et al., 2008, Guenther et al., in press).

For the current subject, ER, single and multi-units were identified from the multi-unit extracellular potential recorded from the Neurotrophic Electrode. Briefly, the extracellular potentials were first bandpass filtered (300–6000 Hz), then a voltage threshold was applied ( $\pm 10\mu\text{V}$ ) using the Cheetah data acquisition system (Neuralynx, Inc., Bozeman, MT). Threshold crossings were taken as putative action potentials, and an approximately 1 ms (or 32-point; hardware-dependent) data segment sampled at 30 kHz around each crossing was saved for classification analysis. All spike waveforms were classified on-line using a convex-hull technique (SpikeSort3D, Neuralynx, Inc.) according to manually defined regions obtained from previous offline analysis. These cluster regions, once stabilized, were reused for each recording session. In the current study, approximately 56 units were identified across two recording channels ( $N_1=29$ ,  $N_2=27$ ), although this is likely an overestimate as some clusters may represent the same parent neural source.

We then developed a real-time neural prosthesis for control of a formant frequency-based speech synthesizer (Figure 2) using unit firing rates via a continuous filter approach, which is detailed elsewhere (Brumberg et al., 2009, Guenther et al., in press) but briefly summarized below. The neural decoder obtained firing rate activity from many neural sources and performed a mapping into formant frequency space utilizing a Kalman filter (Kalman, 1960)



based decoding algorithm. Specifically, the position (i.e. the first two formant frequencies, or simply *formants*) and velocities (i.e. 1<sup>st</sup> derivative of formants) were decoded from normalized unit firing rates. A similar continuous filter decoder was developed by Kim and colleagues (Kim et al., 2007) to decode hand movement kinematics from human subjects. The formant frequencies were then used to drive an artificial speech synthesizer which played the synthesized vowel waveform from the computer speakers with a total system delay of less than 50 ms from neural firing to sound output. Low system delays are necessary for fluent speech, as auditory feedback delays more than 200 ms are known to disrupt normal speech production (MacKay, 1968).

Control of the continuously varying speech BCI is analogous to two-dimensional cursor movement control previously accomplished with SUA in monkeys and humans. The main difference lies in the nature of both the underlying neural signal and the control modality, both of which are represented in the auditory domain rather than the visuo-spatial domain. Speech production and perception are naturally acoustic tasks; as such auditory feedback is much more informative regarding our ongoing speech movements than visual feedback (which is completely lacking during self-generated speech). Interestingly, visual feedback can be important for speech perception, but typically this refers to perception of external vocal articulators (e.g. jaw, lips and tongue tip) and is of most use when the acoustic speech signal is degraded (e.g. conversing in a noisy environment). This type of visual feedback may be useful in the speech prosthesis to help users visualize neural decoding algorithm predictions but will likely be secondary to fast auditory feedback.

Formant frequencies are a natural choice for speech representations in the speech motor cortex for two reasons. First, the Neurotrophic Electrode was implanted on the border of the left primary and premotor cortex, in a location that may be involved in planning upcoming utterances, as proposed by a neurocomputational model of speech production, the Directions into Velocities of Articulators (DIVA) model (Guenther, 1994, 1995; Guenther et al., 2006). Specifically, the DIVA model hypothesizes speech motor trajectories are planned in the premotor cortex utilizing acoustic targets. Second, the formant frequencies of speech are highly correlated with movements of the vocal tract (e.g., changes in first formant frequency are related to forward and backward movements of the tongue). Therefore, we used a formant frequency approach as they are a low-dimensional representation of speech sounds and are related to the gross movements of the jaw and tongue. We believe that a formant frequency approach should work similarly to an articulatory approach due to this relationship, and the low-dimensionality needed for BCI control of formants should aid ER in initial production attempts.

In the speech production study (Brumberg et al., 2009; Guenther et al., in press), ER was instructed to perform two paradigms. The first paradigm required the subject to listen to artificially synthesized vowel sequences, played over computer speakers, consisting of repetitions of three different vowels (/AA/ [hot], /IY/ [heat] and /UW/ [hoot]) interleaved with a neutral vowel sound (/AH/ [hut]). The vowels and vowel-transitions were synthesized using a formant synthesizer according to predetermined formant trajectories. The subject was asked to covertly speak along with the vowel sequence stimulus while he listened. The data obtained in this paradigm was used for offline calibration of the real-time Kalman filter neural decoder. Parameters for the Kalman filter decoder were estimated by performing a least squares regression of unit firing rates and the vowel sequence formant trajectories. An offline analysis of the training data showed statistically significant correlations between the ensemble unit firing rates and formant frequencies (Guenther et al., in press). Individual unit tuning preferences were computed according to the method in Guenther et al. (in press) and a few examples are shown in Figure 3. The tuning curves summarize individual units' preferred formant frequencies defined as the correlation between unit firing rates and training sample

formant frequency trajectories. The tuning strength is defined as the magnitude of the correlation coefficient. The mean tuning strength (Figure 3, black curve; 95% confidence intervals in gray) represents the units' preference for formant frequencies in the shown direction relative to the center vowel /AH/. Figures 3(a) and (b) show the tuning curves for units with mostly F2 preference while Figures 3(c) and (d) show tuning curves for units with mixed F1 and F2 preference.

The second paradigm again required ER to listen to artificially synthesized vowel sequences. However, in this paradigm he was instructed to listen only during stimulus presentation. In addition, these stimuli were limited to two vowels (/V1 V2/) where V1 was always /AH/ and V2 was randomly selected between the three vowels /AA/, /IY/ and /UW/. A production period followed the listen period in which the subject was instructed to attempt to produce the vowel sequence. During the production period, the real-time neural decoder was activated and new formant frequencies were predicted from brain activity related to the production attempt every 15 milliseconds.

To date, ER has been able to perform the vowel production task using the BCI with reasonably high accuracy, attaining 70% correct production on average after approximately 15–20 practice attempts per session over 25 sessions (Brumberg et al., 2009; Guenther et al., in press). Table 2 illustrates the within-session learning effect. Production trials were grouped into blocks (roughly four blocks of six trials per session) and analyzed for endpoint production accuracy and error. Early trials (Block 1) show relatively poor performance which statistically significantly increases by Block ( $p < 0.05$ ; t-test of zero slope as a function of block). Vowel sequence endpoint error, defined as the Euclidean distance from the endpoint formant pair to the target vowel, significantly decreased from the session start to termination ( $p < 0.05$ ; t-test of zero slope as a function of block). A summary of trial production attempts is shown in Figure 4. In both figures, a two-dimensional histogram was computed from predicted formant trajectories and smoothed for ease of visualization. The red trajectories correspond to /AH AA/ trials, green for /AH UW/ trials and blue for /AH IY/ trials. Target vowel formant regions are indicated by the gray regions for each of the three endpoint vowels. The target region acted both to determine the categorical accuracy of current productions as well as provide an attractor, based on distance from target center, to help ER maintain steady vowel productions. Figure 4 (a) shows the predicted formant trajectories for all trials while Figure 4(b) is restricted to only correct productions. A detailed description of the methods and results of the closed-loop BCI speech production study can be found elsewhere (Guenther et al., in press). Though preliminary, these results show it is possible for a human subject to use the Neurotrophic Electrode for real-time continuous speech production through formant frequency speech synthesis.

The formant frequency speech BCI suffers one major drawback, namely a high degree of difficulty in production of consonants compared to vowels. Therefore, we are developing methods for continuous articulatory decoding and synthesis as well as discrete phoneme prediction. The articulation-based system permits a greater range of productions including both vowels and consonants, in contrast to the formant based system which is suitable primarily for vowels.

In addition to these continuous control systems, we are also implementing a discrete control system that is intended to determine specific phonemes, syllables, and/or words from neural firing rates. Note that the discrete system is proposed to be different than previous virtual keyboard attempts involving the selection of characters by visual feedback. Rather, this approach will attempt to decode phoneme information much like the automatic speech recognition techniques using non-audible signals described in Section 1, but directly from neural activity associated with their attempted production.

## SECTION 4: FUTURE CHALLENGES

This section addresses important areas for future development of chronically implanted intracortical electrode speech BCIs in order to achieve near-conversational speech rates. Silent communication by intracortical electrode brain computer interface does not yet permit synthesizer “movements” that are fast and accurate enough to achieve normal speaking rates, though this rate limitation can be improved upon by increasing the quantity of microelectrodes and improving decoding algorithms for estimating the speech signal from neural activity. Furthermore, the neural prosthesis system must be portable and robust for use by the recipient in real-world environments, rather than limited to the laboratory.

### Increased number of electrodes

Both the Utah array and Neurotrophic Electrode are capable of recording from dozens of neurons, but this number of neurons is minute compared to the number of cells involved in speech production. Previous electrophysiological investigations of arm and hand kinematics have shown that relatively few ( $N < 50$ ) single units are needed to predict the movements of arm and hand with reasonable accuracy, allowing for control of robotic devices via BCIs (Taylor et al., 2002; Velliste et al., 2008). However, the motor control required for speech production is less like gross arm movements and more like fine hand movements used in handwriting. It is likely that the fine, precise movements of handwriting and speech may require hundreds of single units rather than dozens. It has been shown that prediction accuracy is positively correlated with the number of recorded units (Carmena et al., 2003; Wessberg et al., 2000). Accordingly, future versions of the Neurotrophic Electrode will take advantage of microfabrication techniques to increase the number of contact points per electrode wire (Siebert et al., 2008), increasing the number of possible recorded neural signals while improving single unit discrimination. Furthermore, several Neurotrophic Electrodes will be implanted in future volunteers (with more wires per electrode), rather than the single three-wire electrode utilized in the current subject, again greatly increasing the number of potentially recorded neural sources.

### Electrodes in multiple brain regions

Placement of additional electrodes requires careful study to make the most out of the increased number of electrical recordings. Speech production in neurologically normal subjects spans many brain regions, such as ventral primary and premotor cortex, supplementary motor area, Broca’s area and higher order auditory and somatosensory regions (Guenther et al., 2006). Though it is currently only possible to implant four or five Neurotrophic Electrodes due to implanted hardware size constraints, they can be intelligently placed, for example, in the different orofacial regions of primary motor cortex to ensure recording of neural activity related to movements of different speech articulators, including the lips, jaw, tongue and larynx. Speech prosthesis control would likely improve significantly with additional specific information about the major speech articulators. Additionally, electrode recordings of Broca’s area (the left posterior inferior frontal gyrus) and/or the supplementary motor area may increase prediction accuracy as these areas have long been understood to be crucial to normal speech production (Penfield and Roberts, 1959).

### Improved algorithms for decoding neural signals

The algorithms currently being used to decode motor cortical activity into speech representations were originally designed for other, non-neurophysiological tasks. Though they perform quite well in limited applications so far, further refinement should be undertaken to account for the specific characteristics of both speech and neurophysiological signals. For instance, both Wiener (Gelb, 1974) and Kalman (Kalman, 1960) filtering are popular techniques for prediction of an unknown system state (e.g. speech output) given known discrete

measurements (e.g. neural activity) and have been used to decode arm movement kinematics (Hochberg et al., 2006; Kim et al., 2007) as well as formant frequency trajectories (Brumberg et al., 2009; Guenther et al., in press). However, neither method was designed with electrophysiology in mind. Furthermore, neural decoding algorithms, in general, have been developed under the assumption that the known measurement states are neural firing rates (spikes/sec) of recorded neurons. Rate estimation, as such, is simply the result of kernel-based smoothing of the neural spike train. Unfortunately, the smoothing operation eliminates useful information contained within the individual spike arrival times. Specifically, the mean is retained but spiking variance is ignored. Therefore, algorithms dedicated to the analysis of point processes (i.e. neural spike trains) with specific modifications for speech representations (i.e. formants or articulator trajectories) should lead to improved decoding (Brown et al., 2004; Truccolo et al., 2005, 2008).

Combined continuous and discrete control systems will likely improve prosthesis performance as well (Kim et al., 2007). Our group has already demonstrated the feasibility of discrete decoding of speech information utilizing nonlinear and neural network techniques (offline, Miller et al., 2007; online, Wright et al., 2007). An immediate discrete classification goal is to predict the onset of vocalization during speech production attempts (Matthews et al., 2008) and to integrate such a “voicing detector” into the current continuous speech decoding protocol.

### Development of portable systems

For the most part, silent communication BCIs are currently restricted to usage within the laboratory. Though some EEG methods are available for home-bound, real-world environment use, these methods do not accomplish conversation-rate speech. In contrast, intracortical BCIs have a greater intrinsic potential given the richness of the information source (specifically, the ability to separate the contributions of individual neurons from motor cortex) and have enjoyed success for motor-cortical BCIs for mouse and robotic arm control. Therefore, considerable effort must be made to translate laboratory-grade intracortical BCI hardware and software systems to robust portable systems for real-world use. One important translational milestone has already been met by Kennedy and colleagues, who have developed a wireless telemetry system for transmission of recorded extracellular potentials rather than wired percutaneous connections. Wireless transmission is a key requirement to the development of portable systems as it permits a relatively easy approach to connecting and disconnecting the user from the prosthetic device. In addition, wireless transmission allows for complete healing of the scalp wound, which is preferable for long-term use in uncontrolled environments as it avoids infection. Further development of the BCI systems discussed in this manuscript can be guided in part by lessons learned from other commercial BCIs, such as cochlear implants, where the task of converting laboratory equipment for portable use has already been successfully accomplished.

A key factor in the development of portable systems from laboratory-grade research systems for locked-in individuals is to accommodate bulky hardware devices commonly used in BCI applications. For instance, typical hardware systems for intracortical microelectrode data acquisition are rack-mounted, making their transportation difficult. The hardware and software systems required to acquire neural signals and decode predicted behavior states can be integrated into the wheelchair power supply and chassis. The intracortical continuous and discrete conversational speech production BCIs described in Section 3 have further advantages for portable design manufacturing; they require no additional hardware for acoustic output. Artificial computer speech synthesis algorithms can be implemented on off-the-shelf computers without additional effort. These computers simply require rugged encasings and integration with the user's wheelchair.

## SECTION 5: CONCLUDING REMARKS

This article has described the current state of silent communication BCI devices based on neural signals, with a focus on intracortical electrode designs. Though EEG brain computer interfaces have been successfully used by locked-in patients in discrete finite letter choice spelling paradigms, typing rates are quite slow, with full paragraphs taking hours to produce. Furthermore, we claim auditory feedback is the most natural feedback domain for speech production and a more effective feedback modality than visual feedback (as in typing paradigms) for locked-in users. An alternative to EEG and ECoG BCI methodologies involves intracortical microelectrode implantation of the speech motor cortex with the goal of continuous control of a speech synthesizer which acts as a “prosthetic vocal tract,” treating speech as a motor control problem rather than using a keyboard-based solution.

Four major extracellular electrodes were discussed, outlining their usefulness in long-term brain computer interfaces, with particular emphasis on the Utah array and Neurotrophic Electrode as they are already being used in human BCI applications. Both electrode designs have remained viable in human cortex for durations on the order of years, demonstrating their feasibility for long-term prostheses. Recent work by our lab has extended previous non-speech movement kinematics electrophysiological research to the speech domain, indicating how speech production can be represented effectively in low-dimensional acoustic or articulatory domains, making speech restoration via an invasive brain-computer interface a tractable problem.

Finally, future directions of neural interfaces for speech restoration were discussed. Four major design considerations were outlined, including increasing the number of electrodes, utilizing smarter and more accurate electrode localization strategies, improving decoding algorithms with specific emphasis on neurophysiology and speech science, and the translation of bulky laboratory-based systems to portable designs suitable for use outside the laboratory.

## Acknowledgments

This research was supported by the National Institute on Deafness and other Communication Disorders (R01 DC007683; R01 DC002852; R44 DC007050-02) and by CELEST, an NSF Science of Learning Center (NSF SBE-0354378). The authors thank the participant and his family for their dedication to this research project, Rob Law and Misha Panko for their assistance with the preparation of this manuscript and Tanja Schultz for her helpful comments.

## REFERENCES

- Allison BZ, McFarland DJ, Schalk G, Zheng SD, Jackson MM, Wolpaw JR. Towards an Independent Brain - Computer Interface Using Steady State Visual Evoked Potentials. *Clinical neurophysiology* : official journal of the International Federation of Clinical Neurophysiology 2008;119(2):399–408. [PubMed: 18077208]
- Bartels JL, Andreasen D, Ehirim P, Mao H, Seibert S, Wright EJ, Kennedy PR. Neurotrophic electrode: Method of assembly and implantation into human motor speech cortex. *Journal of Neuroscience Methods* 2008;174(2):168–176. [PubMed: 18672003]
- Betts BJ, Jorgensen C. Small Vocabulary Recognition Using Surface Electromyography in an Acoustically Harsh Environment. *Interacting with Computers* 2006;18(6):1242–1259.
- Birbaumer N, Ghanayim N, Hinterberger T, Iversen I, Kotchoubey B, Kubler A, Perelmouter J, Taub E, Flor H. A spelling device for the paralysed. *Nature* 1999;398(6725):297–298. [PubMed: 10192330]
- Birbaumer N, Kubler A, Ghanayim N, Hinterberger T, Perelmouter J, Kaiser J, Iversen I, Kotchoubey B, Neumann N, Flor H. The thought translation device (TTD) for completely paralyzed patients. *IEEE Transactions on Rehabilitation Engineering* 2000;8(2):190–193. [PubMed: 10896183]



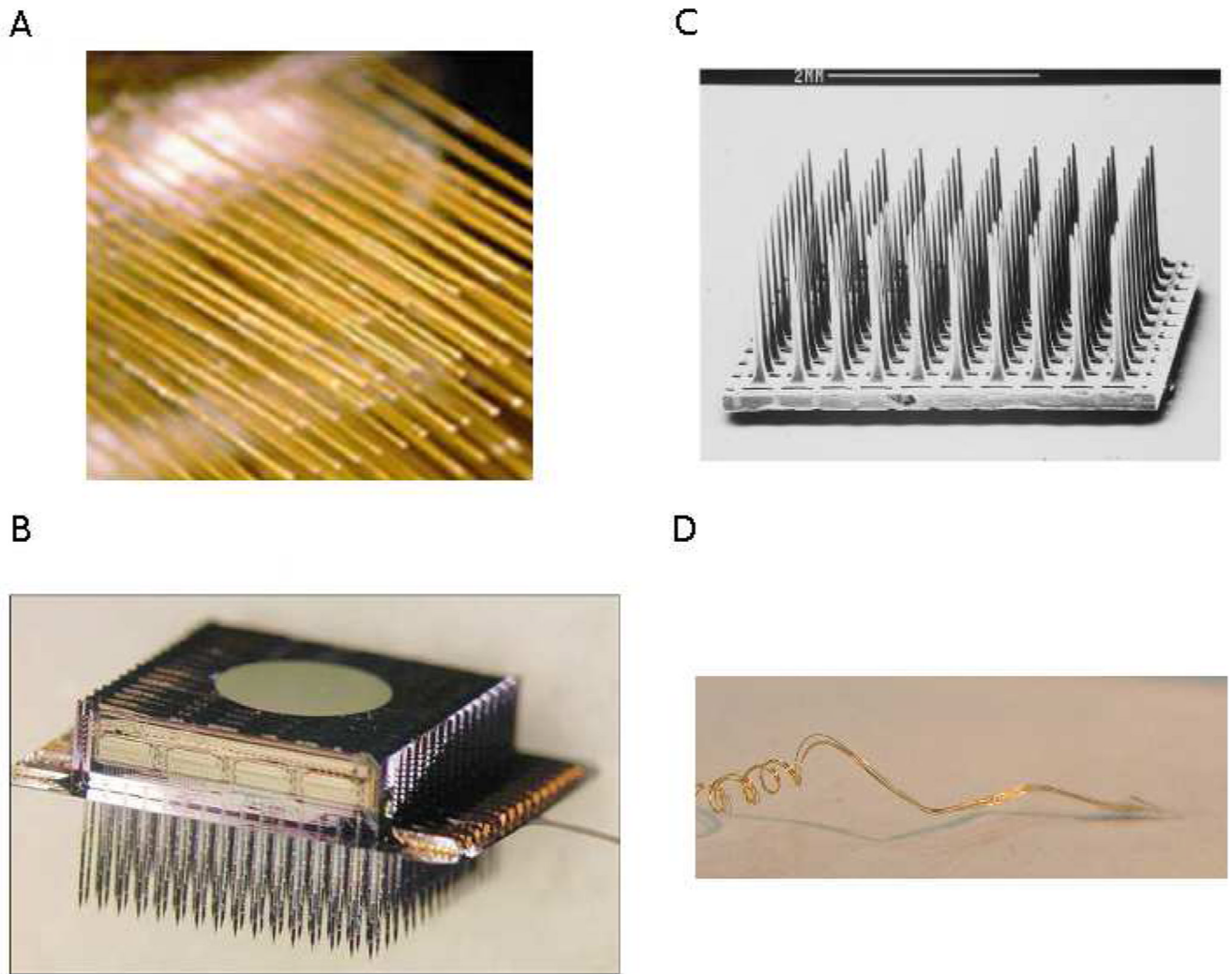
- Birbaumer N, Hinterberger T, Kübler A, Neumann N. The thought-translation device (TTD): neurobehavioral mechanisms and clinical outcome. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 2003;11(2):120–123. [PubMed: 12899251]
- Brown, EN.; Barbieri, R.; Eden, UT.; Frank, LM. Likelihood methods for neural spike train data analysis. In: Feng, J., editor. *Computational Neuroscience: A Comprehensive Approach*. Vol. Vol. 7. CRC Press; 2004. p. 253-286.
- Brumberg, JS.; Kennedy, PR.; Guenther, FH. Artificial Speech Synthesizer Control by Brain-Computer Interface; Proceedings of the 10th Annual Conference of the International Speech Communication Association; Brighton, U.K: International Speech Communication Association; 2009.
- Carmena JM, Lebedev MA, Crist RE, O'Doherty JE, Santucci DM, Dimitrov DF, Patil PG, Henriquez CS, Nicolelis MAL. Learning to control a brain-machine interface for reaching and grasping by primates. *PloS Biology* 2003;1(2):193–208.
- Cheng M, Gao X, Gao S, Xu D. Design and implementation of a brain-computer interface with high transfer rates. *IEEE Transactions on Biomedical Engineering* 2002;49(10):1181–1186. [PubMed: 12374343]
- DaSalla CS, Kambara H, Sato M, Koike Y. Single-trial classification of vowel speech imagery using common spatial patterns. *Neural Networks* 2009;22(9):1334–1339. [PubMed: 19497710]
- Donchin E, Spencer K, Wijesinghe R. The mental prosthesis: assessing the speed of a P300-based brain-computer interface. *IEEE Transactions on Rehabilitation Engineering* 2000;8(2):174–179. [PubMed: 10896179]
- Donoghue JP, Nurmikko A, Black M, Hochberg LR. Assistive technology and robotic control using motor cortex ensemble-based neural interface systems in humans with tetraplegia. *Journal of Physiology* 2007;579(3):603–611. [PubMed: 17272345]
- Fagan M, Ell S, Gilbert J, Sarrazin E, Chapman P. Development of a (silent) speech recognition system for patients following laryngectomy. *Medical Engineering & Physics* 2008;30(4):419–425. [PubMed: 17600751]
- Gelb, A. *Applied Optimal Estimation*. The MIT Press; 1974.
- Georgopoulos AP, Kalaska JF, Caminiti R, Massey JT. On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. *Journal of Neuroscience* 1982;2(11):1527–1537. [PubMed: 7143039]
- Guenther FH. A neural network model of speech acquisition and motor equivalent speech production. *Biological Cybernetics* 1994;72(1):43–53. [PubMed: 7880914]
- Guenther FH. Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review* 1995;102(3):594–621. [PubMed: 7624456]
- Guenther FH, Brumberg JS, Wright EJ, Nieto-Castanon A, Tourville JA, Panko M, Law R, Siebert SA, Bartels JL, Andreasen DS, Ehirim P, Mao H, Kennedy PR. A wireless brain-machine interface for real-time speech synthesis. *PLoS ONE*. (in press).
- Guenther FH, Ghosh SS, Tourville JA. Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language* 2006;96(3):280–301. [PubMed: 16040108]
- Hamalainen M, Hari R, Ilmoniemi RJ, Knuutila J, Lounasmaa OV. Magnetoencephalography•theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics* 1993;65(2):413.
- Hinterberger T, Kübler A, Kaiser J, Neumann N, Birbaumer N. A brain-computer interface (BCI) for the locked-in: comparison of different EEG classifications for the thought translation device. *Clinical Neurophysiology* 2003;114(3):416–425. [PubMed: 12705422]
- Hochberg LR, Serruya MD, Fiehs GM, Mukand JA, Saleh M, Caplan AH, Branner A, Chen D, Penn RD, Donoghue JP. Neuronal ensemble control of prosthetic devices by a human with tetraplegia. *Nature* 2006;442(7099):164–171. [PubMed: 16838014]
- Hochberg, LR.; Simeral, JD.; Kim, S.; Stein, J.; Fiehs, GM.; Black, MJ.; Donoghue, JP. More than two years of intracortically-based cursor control via a neural interface system. *Neuroscience Meeting Planner* 2008, Program No. 673.15; Washington, DC. 2008.
- Hoogerwerf A, Wise K. A three-dimensional microelectrode array for chronic neural recording. *IEEE Transactions on Biomedical Engineering* 1994;41(12):1136–1146. [PubMed: 7851915]



- Jones K, Campbell P, Normann R. A glass/silicon composite intracortical electrode array. *Annals of Biomedical Engineering* 1992;20(4):423–437. [PubMed: 1510294]
- Jorgensen, C.; Lee, D.; Agabont, S. Sub auditory speech recognition based on EMG signals; Proceedings of the International Joint Conference on Neural Networks; 2003. p. 3128–3133.
- Jou, SC.; Schultz, T.; Walliczek, M.; Kraft, F.; Waibel, A. Towards Continuous Speech Recognition Using Surface Electromyography; INTERSPEECH-2006; 2006. (pp. 1592–Mon3WeS.3)
- Jou, SS.; Schultz, T. Biomedical Engineering Systems and Technologies. 2009. Automatic Speech Recognition Based on Electromyographic Biosignals; p. 305–320.
- Kalman RE. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering* 1960;82(1):35–45.
- Kennedy PR, Kirby T, Moore MM, King B, Mallory A. Computer control using human intracortical local field potentials. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 2004;12(3):339–344. [PubMed: 15473196]
- Kennedy PR. The cone electrode: a long-term electrode that records from neurites grown onto its recording surface. *Journal of Neuroscience Methods* 1989;29:181–193. [PubMed: 2796391]
- Kennedy, PR. The Biomedical Engineering Handbook, The electrical engineering handbook series. 3rd ed. Vol. 1. Boca Raton: CRS/Taylor and Francis; 2006. Comparing electrodes for use as cortical control signals: Tiny tines, tiny wires or tiny cones on wires: Which is best?.
- Kennedy PR, Bakay RAE. Restoration of neural output from a paralyzed patient by direct brain connection. *NeuroReport* 1998;9:1707–1711. [PubMed: 9665587]
- Kennedy PR, Bakay RAE, Sharpe SM. Behavioral correlates of action potentials recorded chronically inside the Cone Electrode. *NeuroReport* 1992a;3:605–608. [PubMed: 1421115]
- Kennedy PR, Mirra SS, Bakay RAE. The cone electrode: ultrastructural studies following long-term recording in rat and monkey cortex. *Neuroscience Letters* 1992b;142:89–94. [PubMed: 1407726]
- Kim, S.; Simeral, JD.; Hochberg, LR.; Donoghue, JP.; Friebs, GM.; Black, MJ. Multi-state decoding of point-and-click control signals from motor cortical activity in a human with tetraplegia; *Neural Engineering*, 2007. CNE'07. 3rd International IEEE/EMBS Conference; 2007. p. 486–489.
- Kipke D, Vetter R, Williams J, Hetke J. Silicon-substrate intracortical microelectrode arrays for long-term recording of neuronal spike activity in cerebral cortex. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 2003;11(2):151–155. [PubMed: 12899260]
- Krusienski DJ, Sellers EW, Cabestaing F, Bayouth S, McFarland DJ, Vaughan TM, Wolpaw JR. A comparison of classification techniques for the P300 Speller. *Journal of Neural Engineering* 2006;3(4):299–305. [PubMed: 17124334]
- Krusienski DJ, Sellers EW, McFarland DJ, Vaughan TM, Wolpaw JR. Toward enhanced P300 speller performance. *Journal of Neuroscience Methods* 2008;167(1):15–21. [PubMed: 17822777]
- Kübler A, Kotchoubey B, Hinterberger T, Ghanayim N, Perelmouter J, Schauer M, Fritsch C, Taub E, Birbaumer N. The thought translation device: a neurophysiological approach to communication in total motor paralysis. *Experimental Brain Research* 1999;124(2):223–232.
- Leuthardt EC, Schalk G, Wolpaw JR, Ojemann JG, Moran DW. A brain-computer interface using electrocorticographic signals in humans. *Journal of Neural Engineering* 2004;1(2):63–71. [PubMed: 15876624]
- MacKay DG. Metamorphosis of a Critical Interval: Age-Linked Changes in the Delay in Auditory Feedback that Produces Maximal Disruption of Speech. *The Journal of the Acoustical Society of America* 1968;43(4):811–821. [PubMed: 5645830]
- Maier-Hein, L.; Metze, F.; Schultz, T.; Waibel, A. Session independent non-audible speech recognition using surface electromyography; Automatic Speech Recognition and Understanding, 2005 IEEE Workshop on; 2005. p. 331–336.
- Matthews, BA.; Clements, MA.; Kennedy, PR.; Andreasen, DS.; Bartels, JL.; Wright, EJ.; Siebert, SA. Automatic detection of speech activity from neural signals in Broca's Area. *Neuroscience Meeting Planner* 2008, Program No. 862.6; Washington, DC. 2008.
- Maynard EM, Nordhausen CT, Normann RA. The Utah Intracortical Electrode Array: A recording structure for potential brain-computer interfaces. *Electroencephalography and Clinical Neurophysiology* 1997;102(3):228–239. [PubMed: 9129578]

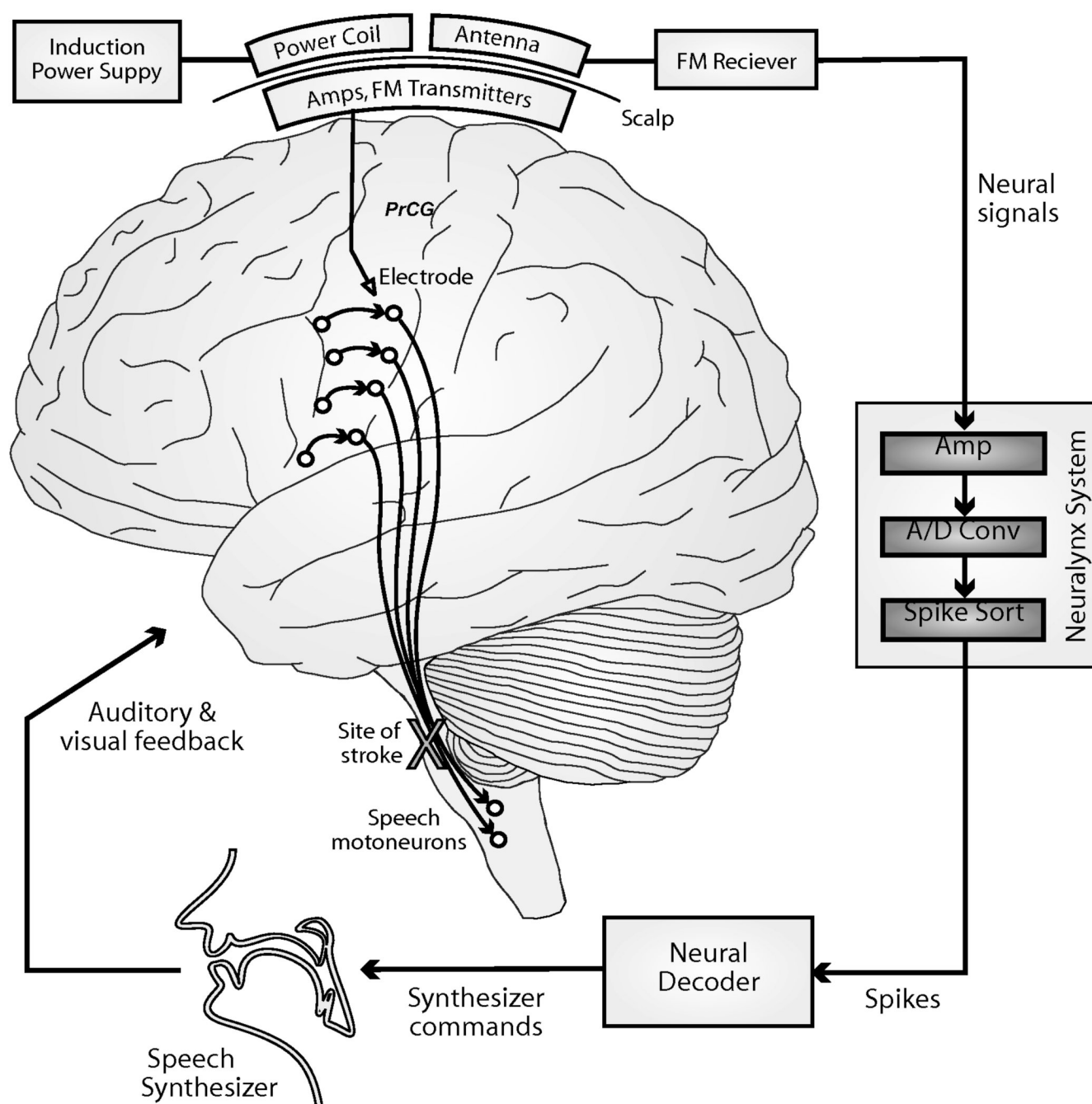
- Mendes, J.; Robson, R.; Labidi, S.; Barros, A. Subvocal Speech Recognition Based on EMG Signal Using Independent Component Analysis and Neural Network MLP; Image and Signal Processing, 2008, CISP '08. Congress on; 2008. p. 221-224.
- Miller, LE.; Andreasen, DS.; Bartels, JL.; Kennedy, PR.; Robesco, J.; Siebert, SA.; Wright, EJ. Human speech cortex long-term recordings [4]: bayesian analyses. Neuroscience Meeting Planner 2007, Program No. 517.20; San Diego, CA. 2007.
- Mitzdorf U. Current source-density method and application in cat cerebral cortex: investigation of evoked potentials and EEG phenomena. *Physiological Reviews* 1985;65(1):37-100. [PubMed: 3880898]
- Mountcastle VB, Lynch JC, Georgopoulos A, Sakata H, Acuna C. Posterior parietal association cortex of the monkey: command functions for operations within extrapersonal space. *Journal of Neurophysiology* 1975;38(4):871-908. [PubMed: 808592]
- Nicolelis MAL, Dimitrov D, Carmena JM, Crist R, Lehew G, Kralik JD, Wise SP. Chronic, multisite, multielectrode recordings in macaque monkeys. *Proceedings of the National Academy of Sciences of the United States of America* 2003;100(19):11041-11046. [PubMed: 12960378]
- Penfield, W.; Roberts, L. Speech and brain-mechanisms. Princeton, NJ: Princeton University Press; 1959.
- Porbadnigk, A.; Wester, M.; Calliess, JP.; Schultz, T. Biosignals. Porto, Portugal: 2009. EEG-based speech recognition – impact of temporal effects; p. 376-381.
- Rousche PJ, Normann RA. Chronic recording capability of the Utah Intracortical Electrode Array in cat sensory cortex. *Journal of Neuroscience Methods* 1998;82(1):1-15. [PubMed: 10223510]
- Schalk G, Kubánek J, Miller KJ, Anderson NR, Leuthardt EC, Ojemann JG, Limbrick D, Moran D, Gerhardt LA, Wolpaw JR. Decoding two-dimensional movement trajectories using electrocorticographic signals in humans. *Journal of Neural Engineering* 2007;4(3):264-275. [PubMed: 17873429]
- Schalk G, Miller KJ, Anderson NR, Wilson JA, Smyth MD, Ojemann JG, Moran DW, Wolpaw JR, Leuthardt EC. Two-dimensional movement control using electrocorticographic signals in humans. *Journal of Neural Engineering* 2008;5(1):75-84. [PubMed: 18310813]
- Schmidt E, Bak M, McIntosh J. Long-term chronic recording from cortical neurons. *Experimental Neurology* 1976;52(3):496-506. [PubMed: 821770]
- Sellers EW, Krusienski DJ, McFarland DJ, Vaughan TM, Wolpaw JR. A P300 event-related potential brain-computer interface (BCI): The effects of matrix size and inter stimulus interval on performance. *Biological Psychology* 2006;73(3):242-252. [PubMed: 16860920]
- Siebert, SA.; Bartels, JL.; Shire, D.; Kennedy, PR.; Andreasen, D. Advances in the development of the Neurotrophic Electrode. Neuroscience Meeting Planner 2008, Program No. 779.9; Washington, DC. 2008.
- Suner S, Fellows MR, Irwin VC, Nakata GK, Donoghue JP. Reliability of signals from a chronically implanted, silicon-based electrode array in non-human primate primary motor cortex. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 2005;13(4):524-541. [PubMed: 16425835]
- Suppes P, Lu Z, Han B. Brain wave recognition of words. *Proceedings of the National Academy of Sciences of the United States of America* 1997;94(26):14965-14969. [PubMed: 9405722]
- Taylor DM, Tillery SI, Schwartz AB. Direct cortical control of 3D neuroprosthetic devices. *Science* 2002;296(5574):1829-1832. [PubMed: 12052948]
- Trejo LJ, Rosipal R, Matthews B. Brain-computer interfaces for 1-D and 2-D cursor control: designs using volitional control of the EEG spectrum or steady-state visual evoked potentials. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 2006;14(2):225-229. [PubMed: 16792300]
- Truccolo W, Eden UT, Fellows MR, Donoghue JP, Brown EN. A point process framework for relating neural spiking activity to spiking history, neural ensemble and extrinsic covariate effects. *Journal of Neurophysiology* 2005;93:1074-1089. [PubMed: 15356183]
- Truccolo W, Friehs GM, Donoghue JP, Hochberg LR. Primary Motor Cortex Tuning to Intended Movement Kinematics in Humans with Tetraplegia. *Journal of Neuroscience* 2008;28(5):1163-1178. [PubMed: 18234894]

- Vaughan T, McFarland D, Schalk G, Sarnacki W, Krusienski D, Sellers E, Wolpaw J. The wadsworth BCI research and development program: at home with BCI. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 2006;14(2):229–233. [PubMed: 16792301]
- Velliste M, Perel S, Spalding MC, Whitford AS, Schwartz AB. Cortical control of a prosthetic arm for self-feeding. *Nature* 2008;453(7198):1098–1101. [PubMed: 18509337]
- Walliczek, M.; Kraft, F.; Jou, SC.; Schultz, T.; Waibel, A. Sub-Word Unit Based Non-Audible Speech Recognition Using Surface Electromyography. *INTERSPEECH-2006*; Pittsburgh, PA. 2006. (pp. 1596-Wed1A20.3)
- Wand, M.; Schultz, T. Towards Speaker-Adaptive Speech Recognition based on Surface Electromyography. *International Conference on Bio-inspired Systems and Signal Processing*; Porto, Portugal. 2009.
- Wessberg J, Stambaugh CR, Kralik JD, Beck PD, Laubach M, Chapin JK, Kim J, Biggs SJ, Srinivasan MA, Nicolelis MAL. Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. *Nature* 2000;408(6810):361–365. [PubMed: 11099043]
- Williams JC, Rennaker RL, Kipke DR. Long-term neural recording characteristics of wire microelectrode arrays implanted in cerebral cortex. *Brain Research Protocols* 1999;4(3):303–313. [PubMed: 10592339]
- Wise K, Anderson D, Hetke J, Kipke D, Najafi K. Wireless implantable microsystems: high-density electronic interfaces to the nervous system. *Proceedings of the IEEE* 2004;92(1):76–97.
- Wise KD, Angell JB, Starr A. An Integrated-Circuit Approach to Extracellular Microelectrodes. *IEEE Transactions on Biomedical Engineering* 1970;17(3):238–247. [PubMed: 5431636]
- Wolpaw J, McFarland D, Vaughan T. Brain-computer interface research at the Wadsworth Center. *IEEE Transactions on Rehabilitation Engineering* 2000;8(2):222–226. [PubMed: 10896194]
- Wolpaw JR, McFarland DJ. Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans. *Proceedings of the National Academy of Sciences of the United States of America* 2004;101(51):17849–17854. [PubMed: 15585584]
- Wright, EJ.; Andreasen, DS.; Bartels, JL.; Brumberg, JS.; Guenther, FH.; Kennedy, PR.; Miller, LE.; Robesco, J.; Schwartz, AB.; Siebert, SA.; Velliste, M. Human speech cortex long-term recordings [3]: neural net analyses. *Neuroscience Meeting Planner* 2007, Program No. 517.18; San Diego, CA. 2007.
- Wright, EJ.; Siebert, SA.; Kennedy, PR.; Bartels, JL. Novel method for obtaining electric bio-signals for computer interfacing. *Neuroscience Meeting Planner* 2008, Program No. 862.5; Wash. 2008.



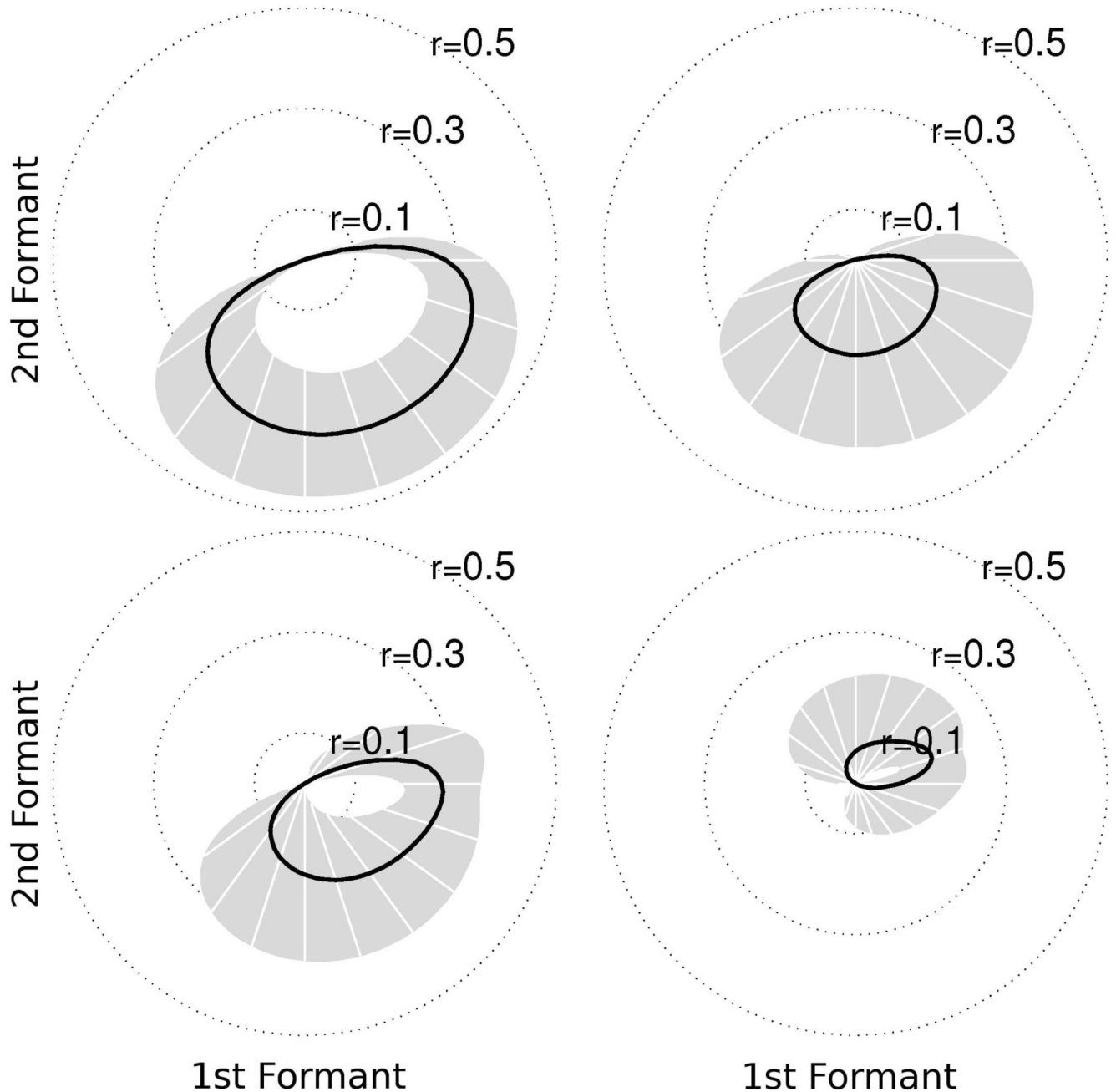
**Figure 1.**

Examples of the four intracortical electrodes discussed in this paper for use in chronic BCI applications. (A) Microwire array; Nicolelis et al. (2003). Chronic, multisite, multielectrode recordings in macaque monkeys. *Proceedings of the National Academy of Sciences of the United States of America*, 100(19), 11041–11046. Copyright (2003) National Academy of Sciences, U.S.A., (B) Michigan microelectrode array (Wise et al., 2004), reproduced with permission (© 2004 IEEE), (C) Utah microelectrode array; Reprinted from *Journal of Neuroscience Methods*, Vol. 82 (1), Rousche and Normann, Chronic recording capability of the Utah Intracortical Electrode Array in Cat Sensory Cortex, 1–15, Copyright (1998) with permission from Elsevier, (D) Neurotrophic Electrode; Reprinted from *Journal of Neuroscience Methods*, Vol. 174 (2), Bartels et al., Neurotrophic Electrode: Method of assembly and implantation into human motor speech cortex, 168–176, Copyright (2008) with permission from Elsevier.



**Figure 2.** Schematic of the continuous neural decoder for speech synthesis. Neural signals are obtained via the Neurotrophic Electrode and acquired utilizing the Cheetah acquisition system (Neuralynx, Inc., Bozeman, MT). Formant frequencies or articulatory trajectories are decoded from neural firing rates.

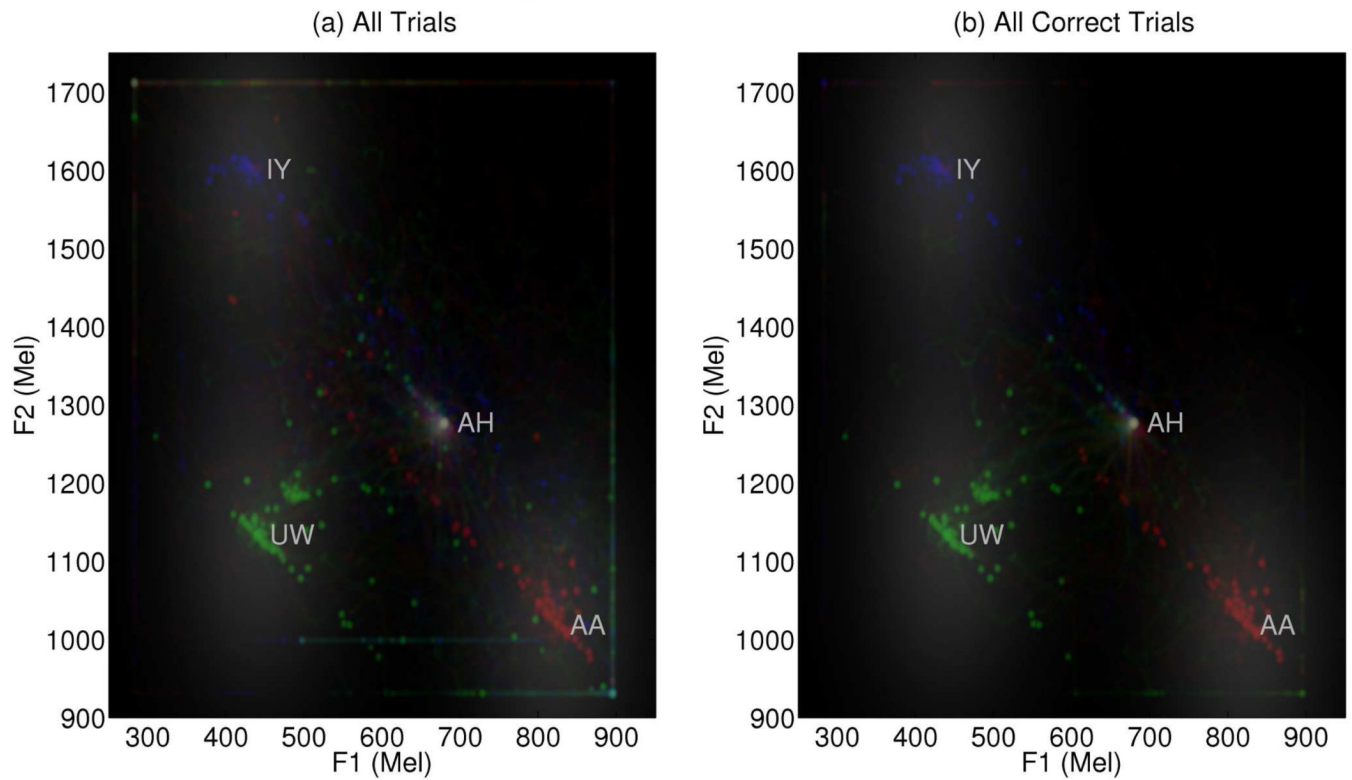




**Figure 3.**

Two dimensional (F1/F2) tuning preferences for sample units determined from offline formant analysis in Guenther et al. (in press). Tuning preference determined by correlation between unit firing rates and formant frequency trajectories. Mean tuning preference indicated by the black curve with 95% confidence intervals in gray. (a,b) tuning curves for units with primarily F2 preference. (c,d) tuning curves for units with mixed F1 – F2 preference. The direction and magnitude of each tuning curve indicates the preference for formant frequencies relative to the center vowel /AH/.





**Figure 4.**

The two dimensional histogram of predicted formant trajectories found during the closed-loop BCI study. Trajectories for /AH AA/ trials are shown in red, /AH UW/ in green and /AH IY/ in blue. Formant regions used for vowel classification and attractors for steady vowel production shown in gray. (a) All trials from 25 closed-loop recording sessions. (b) Correct trials only from all recording sessions.

Table 1

Summary of neurobiological recording technologies in the context of chronic brain-computer interface (BCI) applications. Each of the microelectrode types listed (microwire array, Michigan array, Utah array and Neurotrophic Electrode) are capable of recording both local field potentials (LFPs) and single-unit activity (SUA).

	Study Type	Duration	Implant Characteristics		Fabrication
				Cross-scalp Signal Transmission	
Electrode Sensor Type	Microwire Array	Non-human	18 months	Wired; Percutaneous	Hand-made
	Michigan Array	Non-human	1+ years	Wired; Percutaneous (recent efforts for transcutaneous wireless)	Automated
	Utah Array	Human	2+ years	Wired; Percutaneous	Automated
	Neurotrophic Electrode	Human	4+ years	Wireless; Transcutaneous	Hand-made
	EEG	Human	indefinite	N/A	Automated
	MEG	Human	indefinite	N/A	Automated
	ECoG	Human	Days to weeks (longevity currently unknown)	Wired	Automated

**Table 2**

A summary of within-session vowel production accuracy in terms categorical and continuous error measures for subject ER using the closed-loop speech production BCI. Sessions (N=25) were broken into roughly four blocks per session (six trials per block). Endpoint vowel success and distance to target center were computed for each trial and averaged within blocks across all sessions. Both measures improve: hit rate increases and formant error decreases with block number ( $p < 0.05$ ; t-test of zero slope). Mean values shown with 95% confidence intervals.

	<b>Block 1</b>	<b>Block 2</b>	<b>Block 3</b>	<b>Block 4</b>
Endpoint Hit Rate	0.45 (0.09)	0.58 (0.09)	0.53 (0.13)	0.70 (0.19)
Endpoint Error (Hz)	436 (72)	358 (78)	301 (87)	236 (128)