Published in final edited form as:

Prog Brain Res. 2011 ; 194: 1–25. doi:10.1016/B978-0-444-53815-4.00020-0.

Making the lifetime connection between brain and machine for restoring and enhancing function

Philip Kennedy^{†,*}, Dinal Andreassen^{†,‡}, Jess Bartels[†], Princewill Ehirim[§], Hui Mao[¶], Meel Velliste^{†,||}, Thomas Wichmann[#], and Joe Wright[†]

[†]Neural Signals Inc., Duluth, GA, USA

[‡]Georgia Institute of Technology, Atlanta, GA, USA

[§]Department of Neurosurgery, Gwinnett Medical Center, Lawrenceville, GA, USA

[¶]Department of Radiology, Emory University, Atlanta, GA, USA

^{||}Department of Neurobiology, University of Pittsburgh, Pittsburgh, PA, USA

[#]Department of Neurology, Emory University, Atlanta, GA, USA

Abstract

A reliable neural interface that lasts a lifetime will lead to the development of neural prosthetic devices as well as the possibility that brain function can be enhanced. Our data demonstrate that a reliable neural interface is best achieved when the surrounding neuropil grows into the electrode tip where it is held securely, allowing myelinated axons to be recorded using implanted amplifiers. Stable single and multiunits were recorded from three implanted subjects and classified according to amplitudes and firing rates. In one paralyzed and mute subject implanted for over 5 years with a double electrode in the speech motor cortex, the single units allowed recognition of over half the 39 English language phonemes detected using a variety of decoding methods. These single units were used by the subject in a speech task where vowel phonemes were recognized and fed back to the subject using audio output. Weeks of training resulted in an 80% success rate in producing four vowels in an adaptation of the classic center-out task used in motor control studies. The importance of using single units was shown in a different task using pure tones that the same subject heard and then sung or hummed in his head. Feedback was associated with smoothly coordinated unit firings. The plasticity of the unit firings was demonstrated over several sessions first without, and then with, feedback. These data suggest that units can be reliably recorded over years, that there is an inverse relationship between single unit firing rate and amplitude, that pattern recognition decoding paradigms can allow phoneme recognition, that single units appear more important than multiunits when precision is important, and that units are plastic in their functional relationships. These characteristics of a reliable neural interface are essential for the development of neural prostheses and also for the future enhancement of human brain function.

Keywords

brain computer interfacing; brain machine interfacing; neurotrophic electrode; long-term human recording; speech prosthesis; single unit recording; multi-unit recording; local field potentials

Introduction

For obvious ethical reasons, few studies have been carried out aiming to understand the electro-physiological properties of neurons in the human cortex. Thus far, such information has come from acute, intraoperative microelectrode recordings (Breshears et al., 2009), or by using electrodes that were placed on the cortex for a few weeks prior to surgical resection of epileptic foci (Berger, 1996). An additional source of such information, which has appeared in the past decade, is data from two research groups who have reported long-term recordings using electrodes that were implanted for brain–computer interfacing (Hochberg et al., 2006; Kennedy and Bakay, 1998; Paninski et al., 2004; Serruya et al., 2002), or for speech prosthetics (Guenther et al., 2009). The available studies have largely focused on using recorded signals in neural prosthetic applications. In fact, none of the groups involved with long-term implants in humans has reported on the fundamental properties of human cortical neurons over the long-term and the functional properties of these units. This chapter focuses on the neurophysiologic properties of units recorded from the human motor cortex using data extracted from recordings in three subjects who were implanted with the neurotrophic electrode (NE) for up to 5 years, and with a cumulative experience of 13 years (Bartels et al., 2008). NEs consist of ultrathin Teflon-coated gold wires whose tips are ensheathed in a small cone-shaped glass enclosure which contains neurotrophic factors (Kennedy, 1989). As demonstrated by histological studies, neurons grow neurites into the glass compartment where they undergo myelination (Kennedy et al., 1992). Recordings can then be made from the myelinated neurites “trapped” within the electrode tip. The neuropil within the tip stabilizes after 3–4 months, allowing long-term recordings for many years (still recording over 5 years), and perhaps indefinitely. Our goal is a lifetime of recording from each electrode.

A major focus of this report presented at the brain–machine interface conference in Ystad, Sweden on August 26, 2010 is the basic aspects of the neurophysiology of the human cortex recorded over many years. Data are available from three subjects in whom the basic firing characteristics are reported. The data demonstrate that single units have a very wide range of firing rates, a wide range of amplitudes, and an inverse relationship between amplitude and firing rate. These results are similar across all subjects. In two subjects, we report on unit firings over long time periods (4 and 5 years), and in one subject, we report on data decoding techniques that used perievent time histograms (PETHs), linear discriminant analysis (LDA), flexible discriminant analysis (FDA), and support vector machine (SVM) analyses to identify over half the 38 tested English phonemes using a data set derived from several recording sessions over a few weeks from our locked-in subject implanted 2 years prior (Brumberg et al., 2011). This same subject at year 4 was able to produce vowel phonemes in a speech paradigm that involved first listening to the vowel pair and then speaking it in his head. The audio output produced the vowel sounds, and a display showed

the subject the vowel trajectories in a 2D formant frequency plane (Guenther et al., 2009). With training over many sessions, the subject could produce the vowel pair correctly 80% of the time within a session consisting of 10 trials. We also report here the ability of the subject to listen to a sound and then hum or sing it without and then with feedback of one unit. Feedback optimally consisted of a direct correlation between audio volume and unit firing rate. With this paradigm, the patterns of firings of many single units, but not multiunits, became exquisitely coordinated.

The ongoing recording from human cortical neurons for development of prostheses should produce an unprecedented wealth of basic data relevant to understanding *human* cortical neurophysiology. Obviously, a better understanding of neurophysiologic properties of the human cortex is important in efforts to build neural prosthetic devices that need control signals based on these recordings. In addition, enhancement of human brain function is predicated on understanding the basic physiology of the human brain and how it can be manipulated to optimize a connection directly to the Internet, for example. Human brain enhancement takes on some urgency when considering Ray Kurzweil's prediction that by 2045, intelligent machines will (a) surpass human intelligence, (b) be considered equal to humans, and (c) accorded the same rights as humans (Kurzweil, 2005). One way to delay this moment of singularity is to enhance the capability of the human brain by providing instantaneous access to information, increase external or internal memory storage, and increase mathematical capabilities. However, the limiting factors will likely not be technological such as a reliable neural interface, but rather our limited understanding of how the brain could assimilate and handle extraordinary amounts of information and how it can process such information into an intelligible form.

Methods

Subjects

The experiments described here were carried out in five paralyzed and communication-impaired subjects, as part of a program to develop the use of the NE to control a communication system or, more recently, a speech synthesizer. Subject ER is 26 years old and suffered a brainstem stroke at age 16. His recordings are ongoing 5 years after implantation with a view to develop a speech prosthesis (Brumberg et al., 2009; Guenther et al., 2009). Subject JR was 52 years old and also suffered a brainstem stroke. He has provided much data on brain to computer communication using single units (Kennedy et al., 2000) or local field potentials (Kennedy et al., 2004a,b). Subject DJ, 46 years old, has amyotrophic lateral sclerosis and provided data during an implantation shortened by nonclosure of the incision. Data from another subject (MH), 53 years old, provided data for a few weeks similar to that reported here (Kennedy and Bakay, 1998). The fifth subject (TT), 42 years old, had mitochondrial myopathy that affected his brain soon after implantation so little useful data were obtained despite a 4.5-year survival. Subjects were implanted with electrodes and data recording hardware, and recordings were carried out, starting several months later, and continued for over 5 years in one subject (ER) (Table 1).

Neurotrophic electrode

The “NE” is a surgically implanted electrode that enables recordings of neuronal activity from single neurons in the brain for many years. The electrode was developed for use in paralyzed human patients to provide lifetime control signals for brain–machine interfaces. A full description of the electrode and associated electronic components has recently been published along with assembly and implantation instructions (Bartels et al., 2008). Two key features will be elucidated briefly. The first is the electrode tip, shown in Fig. 1a. The electrode consists of a small glass cone whose inner surface is coated with proprietary growth factors. The tip of the electrode is 50 μm in diameter, while the upper end (where the wires enter) is 300–400 μm in diameter. Three wires, insulated up to their tips (arrows in diagram) which are spaced by 500 μm , are shown entering the glass conical tip and held in place with methacrylate glue. The electrode can contain two wires (subject JR), three wires (subject ER), or four wires (Subject DJ who also had a second 2-wire electrode). Over the course of several weeks postimplantation of the electrode, neurons from different areas of cortex send processes into the electrode which become myelinated (Kennedy et al., 1992). Electrical potentials can be recorded, from the wires, starting 2–3 weeks after electrode implantation and stabilized at 3–4 months. Since we use bipolar recording amplifiers, the location of the recorded neurite influences the shape of the recorded action potential, specifically its amplitude. The closer a neurite is to a wire, the larger its action potential. Proximity to a given wire can be discerned from the initial deflection of the action potential. Neurites close to one wire will depolarize in the direction opposite to the depolarizing direction of neurites close to the other wire. Neurites located precisely midway between the wires will not be recorded at all, as the direction of the initial depolarization is in “both” directions and hence produces a net zero change in charge. The initial depolarization direction provides some initial spatial separation and eases the task of isolating single units. The second key feature of the NE is shown at the upper end of the cone in Fig. 1a where the first coils of one wire are illustrated. All wires are coiled to allow electrode movements in all three directions, a feature that is essential for minimizing stress on the electrode, which greatly improves the stability and longevity of the recorded signals.

In vivo impedance measurements are typically in the range of 50–750 k Ω at 1 kHz and are rarely available after implantation. However, in subject ER, we had to surgically exchange his implanted electronics 2 years after implantation which required exposure of the electrode connector. Impedance measurements at that time showed an impedance of approximately 70 k Ω (at 1000 Hz) across the wires.

Implant targeting, surgical implantation, and implanted electronic package

As described in detail elsewhere (Bartels et al., 2008), the implantation target was chosen based on brain activation maps obtained from functional MRI scans during imagined active speech articulatory movements (ER) or imagined hand movements (in JR and DJ) while using rest or no movement as the control condition. The surgical implantation sites were targeted using a 3D stereotaxic method. The electronics package, containing a power induction receiving coil, regulators, one or two amplifiers, calibration circuits, and FM transmitters, was implanted subcutaneously on the skull. A coil induction system was used to power the electronic components during recording sessions, by placing a powered coil

over the subcutaneous resonant coil. The FM transmitters were used to send the brain signals to external receivers via other coils that were secured to the shaved scalp as shown in Fig. 1b.

Recording methodology, artifact rejection, and spike sorting

As described in our previous publications (Brumberg et al., 2009; Kennedy et al., 2000), the implanted custom-made amplifiers have gains of about 100×, bandpassed with filter settings of 4–4000 Hz. The recorded signals were used to FM-modulate a 30–50 MHz signal that was then transmitted through the skin, and received via Win Radios (WinRadio Corp., 15 Stamford Road, Oakleigh 3166, Australia). The modulated signals were further amplified (100×; BMA 200, CWE Inc., Ardmore, PA, USA). The signals were archived on a digital tape recorder (CDAT16, SCSI 16 channel; Cygnus, Delaware Water Gap, PA, USA) at a sampling rate of 20 kHz. The continuous data streams were archived, and a second channel digitally bandpassed at 300–6000 Hz to provide a steady baseline signal for single unit separation. Artifacts induced by discontinuities or sharp transients in the continuous data stream were detected using user-determined signal slope thresholds. In case such artifacts occur, the recording was interrupted (inputs grounded) for a user-determined duration that can last between milliseconds and minutes.

The neuronal signals were sorted into individual spike clusters using commercial software (Cheetah, Neuralynx Inc., Boseman, MO, USA). To distinguish between single and multiunits, interspike interval histograms (ISIHS) were constructed (Fig. 4). Parameters describing cluster features were saved in a file that is used at each recording session. After an initial stabilization phase, it was possible to use the same cluster parameters throughout the recording period. Clusters were nonoverlapping. In the functional data reported here for subject ER, the cluster parameter file was set up in November 2008 and the data were collected during winter and spring 2009. The cluster parameter file used for data collected in 2007 from subject ER was similar but had to be refined due to reimplantation of the implanted electronics (but not the electrode).

Data collection sessions

The data reported here for subject JR were collected during sessions when he was performing a spelling task as reported previously (Kennedy et al., 2000). The data for subject DJ were collected during a task in which he was making slight pressure movements with his right hand digits while recording from motor cortical area 4. For subject ER, the task needs to be described in full detail as follows.

Subject ER was a 26-year-old male patient, locked-in due to a brainstem stroke in 1999 (at the age of 16 years). He communicated by moving his eyes up for “yes” and down for “no.” The patient was implanted with the NE into his left motor or premotor cortex on December 22, 2004. Data collection recording sessions began on a 3 day per week schedule 3 months after implantation and continued until the fall of 2009. The subject took modafinil (Provigil, Cephalon Inc.) to enhance wakefulness 30 min before any data were collected. The session was terminated if the subject became fatigued as determined by questioning and by observing eye closure. He indicated that he enjoyed the sessions, though they were strenuous

for him, requiring considerable mental effort. To minimize fatigue, he listened to music for approximately 2–3 min after each 2–3 min of data collection. The music tended to aid his concentration. During most sessions, his father remained with him to encourage him, stretch his limbs (for comfort and to minimize spasms), and suction his throat as necessary.

In 2007, we tested his hearing using pure tones and found that he could hear up to 13 kHz. In the sessions reported here, background sounds were minimized except for the unavoidable sounds of computer fans and other electronics. During suctioning of his pharynx using a loud suction machine, recordings were discontinued.

For data collection during tone presentation, blocks of 10 trials, each lasting 31.5 s were employed. Each trial consisted of either a silent control trial or a pure tone followed by either a silent control or a “sing/hum” period where he “sang” (internally) as shown in Fig. 1c. A 750-ms computer instruction to listen was followed by a 200-ms sound to focus his attention, and then a tone was played for 20 s. After this, a “sing” instruction was given, followed by another sound, followed by a silent period during which he sang internally. Figure 1d shows the actual synchronized codes over time, with the “binary 0s and 1s” coding for the specific tone, the “Start recording” indicating recording onset for the block, and “Listen!” and “Sing!” marking their respective onsets with millisecond timing. During the internal singing period (but not the listening period), he received (1) no feedback, (2) feedback of a tone each time a recorded unit fired, or (3) feedback in the form of a tone whose volume varied directly with unit firing. He was questioned after each trial as to whether or not he sang internally, or stayed silent as requested for the control period. He indicated “yes” with upward eye movements, or moved his eyes down for “no.” These responses and other parameters were logged. Ten trials were attempted, but he sometimes showed signs of fatigue before completion of the full trial series. All transitions and codes for the presentations were sent from the instructional computer to the Neuralynx computer and recorded along with the neural data. For data collection during phoneme identification, the basic paradigm is similar to the tone data collection as depicted in Fig. 1c. The difference is that a phoneme, vowel or consonant, is substituted for the tone. In other words, the subject listened to a phoneme and then spoke the phoneme in his head. The timings differ with the “listen” period being 5 s and the “speak” period 10 s. Using a center-out task similar to that used in motor control studies, it was possible to detect four vowel phonemes. The paradigm involved the computer first sounding out two vowels such as “uh”–“ah,” or “uh”–“iy,” or “uh”–“oo.” The subject then said these in his head, and the pattern of firing was analyzed using the LDA technique. The resultant output was heard by the subject and also translated into movement of a cursor on a screen corresponding to the position of the vowels in 2D formant frequency space (that was essentially the 2D movement space, commonly used in motor control studies). The subject’s task was to move from the center “uh” to one of the corner vowels. Details are in Guenther et al. (2009).

Data analysis

The data were digitized by the Neuralynx software during the recording sessions or afterwards by playing it back from the tape. The data included the synchronized codes along with single units and continuous data streams. The cluster parameter file that sorted the

spikes online was not changed throughout the data collection period. The data were first examined using NeuroExplorer software (NEX, from Plexon Technologies, Dallas, TX). ISIHs were generated to distinguish single from multiunit recordings (Fig. 4). Signals were classified as “single units” if ISIs were at least 0.5 ms in duration. Spikes were narrow, presumably due to the fact that they originate from axons, and so can fire close together in time producing the narrow ISIH.

For the tonal and phoneme data analysis, perievent time histograms were used to average the firing rates during 10 s immediately before (listen phase) and after the ding for the “sing” phase for tone periods as well as for the silent control periods. Normalized data were averaged in each of the 10-s data segments consisting of listen, control, or speak or sing. Two tailed “*t*”-tests (at $p < 0.05$) were used to compare values within and between sessions as described in the section “Results.” To detect differences between the trends in firing rate values over a session, the firing rates were subtracted from each other across the session. These rates were converted to absolute values if the differences were negative. The values were summed, and the average of these differences and their standard deviations were found to be distributed in a Gaussian fashion and then subjected to two-tailed *t*-tests. Populations of values were compared between sessions. The null hypothesis stated that there should be no differences between the means and one standard deviation from the mean. *p* values less than 0.05 were accepted as rejecting the null hypothesis. Data reported here in JR and DJ were analyzed only for their spike rates and shapes, and not therefore subject to statistical analysis as were the functional data in ER.

Results

Unit amplitudes and firing rates

Multiunit potentials recorded with the NE range in amplitude from 15 to 50 μ Vs (Fig. 2a). Possible single units based on similar amplitudes are shown as letters in the data for subject ER (Fig. 2a, bottom right). Data from another subject (MH) had five similar small amplitude units as already reported (Kennedy and Bakay, 1998). Subject JR had 19 large single units and subject ER had 20. Subject DJ did not remain implanted long enough to give a final reliable count.

In addition to the low-amplitude signals illustrated in Fig. 2a, higher-amplitude signals were also found (Fig. 2b) with amplitudes ranging over 100 μ Vs. These were sufficiently consistent in JR who used them to control a computer cursor (Kennedy et al., 2000). As seen in the lower panel of Fig. 2b for subject ER who was implanted with the three-wire electrode (the center wire acted as reference), large units found on the tracing from one wire (channel 2) were largely not seen on the adjacent wire (channel 1). The units, though large, were seen on only one wire and hence had to be close to that wire because of the configuration of the wires shown in Fig. 1a. The largest units, however, could be detected by both wires as shown by the only unit on channel 1 for subject ER in Fig. 2b. In other segments of data shown in Fig. 2c, however, there were large units recorded by both channels. These units had to be close to the center reference wire in order to be recorded by both channels. The initial fast phases of depolarization point inward toward each other. This can only happen when the unit is common to both wires by being close to the center reference wire because

the amplifiers have fixed but opposite polarities. In addition, the amplitudes are not equal, strongly suggesting that the units were off-center to the reference wire, that is, closer to one wire or the other of the active (nonreference) wires.

Although these large units were useful in subject JR, they were not consistent enough in the other subjects to be useful prosthetic controllers. Hence the remainder of the reported data deals with the low-amplitude multiunits and their extracted single units in subject ER.

Large units often fired very slowly, while smaller units tended to show faster firing rates. Specific firing rates are illustrated for subject JR in Fig. 3a who was implanted in the hand area of motor cortex. Examples of the units are shown in a column on the right side of the figure, and the firing rates are shown over a 15-s period. Bin size is 100 ms with frequency shown on the abscissa. In contrast, the smaller units for subject ER (illustrated in Fig. 2a above) were faster firing in general and again can be divided into four groups of units, with firing rates inversely related to unit amplitudes as shown in Fig. 3b. The groups are itemized as slow firing, medium firing, fast firing, and fastest firing.

These illustrated units were single units as determined by ISIHS shown in Fig. 4. The interspike intervals of these units dropped to zero at time zero or close to zero, implying that there was little or no firing within a millisecond, implying they were single units. This only applies to those with sufficiently high firing rates. Because of the density of the multiunit data, many units could not be separated into singletons. Thus about half the units remained as multiunits.

Phoneme identification and speech production

As described in the conference talk, we identified over half the English phonemes from this one electrode using what were single units as identified by ISIH criteria. These analyses were performed offline using PETHs, SVMs, LDA, and FDA. The results were similar for all methods used as described in detail by Brumberg et al. (2011). One of these analytic techniques (LDA) was used online to detect vowel phonemes in real time. The method is described above. In moving from vowel to vowel, he achieved 80% success with practice over several months. Errors decreased, move times shortened, and the end point error decreased over the sessions. Improvements were seen within and between sessions. Detailed results have been published in Guenther et al. (2009).

An issue of great interest in the field is the functional role of single units and whether or not these are more or less important than multiunits. In subject ER, we had a unique opportunity to correlate firing rates of single and multiunits with “silent” vocalization to determine their relative contributions to tuning patterns as described in Introduction. These studies were performed over 10 days from day 1546 to 1556. The ISIH technique was applied to the data from those days (1556 and 1546), and the same single and multiunits were found from session to session as illustrated in Fig. 4. As described in the section “Methods,” the cluster definition file was not changed over this period.

Single units more important than multiunits for precision tuning

When interrogated, ER indicated that he “sang in his head” most of the time. During these “singing” studies, he was most enthusiastic when provided with loud tonal feedback of his (assumed) output. This was achieved by directly relating loudspeaker volume to unit firing rate. The issue here was not whether or not he “sang in his head,” but how the firing rates of single and multiunits compare during these “singing” periods. We chose a tone at 523 Hz (C5: one octave above middle C) because this tone appeared to produce an increase in firing rate of unit ch2-09 when tested in 2007. Additionally, we also tested tone 262 Hz (C4: middle “C”). Nine tones were initially tested without feedback to determine if there was any modulation in firing rates of any units. These tones were 110, 131, 220, 247, 262, 440, 494, 523, 880, and 988 Hz that spanned four diatonic scales as notes A–C.

On day 1556, he was provided with feedback of the tone through the speaker each time unit ch2-09 fired. The volume of the tone feedback was directly related to the firing rate. To the subject, it seemed like he was actually singing or humming (vocalizing) the tone as loudly as he could. The results for the identified single units are shown in Fig. 5a for the four classes of units. There were nine trials in each session. In each trial, as described in the section “Methods,” the firing rates were averaged over a listen period, a control period, and a “sing” period. Thus the data appear as “V” shapes in the figures. For many units, there was a progressive smoothing of this pattern as the session progressed. There was a remarkable consistency in most units, even those that were slow firing with average rates of a few Hz, as shown in the figure. The data for tone C4 (middle C) demonstrated a similar smoothing of firing rates (Fig. 5b). Many units responded differently to the different tones (e.g., ch1-01, ch1-03, ch2-13, ch2-17, ch2-19). Interestingly, unit ch2-09, which was the unit fed back as the tone, had averaged firing rates during the control (silent) periods that were higher than those during the listen or sing periods. However, toward the end of the session, the firing rates during the “sing” period increased to levels higher than the controls. In addition, the medium firing rate units (second column) appeared to be the least well tuned. Finally, for the fastest units, ch1-09 and ch1-19 have rates that drop off gradually as the session progressed.

In contrast to the single unit data, the multiunit data produce little if any smooth tuning. These dramatic findings are illustrated in Fig. 5c and d for tones 523 and 262 Hz, respectively. These data were collected at the same session as the single units. These units could not be separated into single units and are multiunits as determined by the ISIH plot (not illustrated).

Plasticity of single unit firings

These data also demonstrate the gradual smoothing of the firing rates over several sessions. This intersession improvement in tuning performance was very obvious in most units. For tone 523 Hz, three sessions preceded the most tuned session (day 1556). The first two sessions on days 1546 and 1549 were performed with no feedback, and the nontuned firing rates are illustrated in Fig. 6a. Feedback of ch2-09 unit firing was provided as a 523-Hz tone on day 1553 (third column in the figure). A pattern of tuning did begin to appear. However, the smooth tuning did not fully appear until volume-related feedback was presented (with the volume increasing as the firing rate increased). These presentations were successfully

presented later as intrasession randomized feedback versus no feedback with similar, though not so strong results. The examples represent a sample from each of the four different types of units. For tone 262 Hz, two preceding sessions provided data on intersession tuning changes. The results were similar as shown.

Statistical analysis

To statistically analyze the variability in these data, the averaged values and one SD of three units (ch2-09, ch2-13, ch2-17) were obtained as described in the section “Methods.” These values were then compared for each condition (listen, control, ing) across the sessions (days 1549, 1553, 1556) for tone 523 Hz, using a two-tailed *t*-test, and plotted in Fig. 6c. The results demonstrate statistically significant differences for all conditions and all units tested between day 1549 (no feedback) and 1556 (volume-related feedback). Statistically, significant differences were seen for most conditions when day 1553 was included.

Discussion

Cellular origins of single units

The different unit amplitudes and firing frequencies illustrated in Figs. 2 and 3 for the four different units need to be interpreted in light of the design features of the NE shown in Fig. 1. The relevant feature of the electrode is that growth into the tip occurs from upper cortical layers as well as from lower layers, as demonstrated in histological analyses of rat and monkey implants where tissue was consistently found to form a bridge throughout the glass tip and connect to the surrounding neuropil above and below the openings (Kennedy et al., 1992). Thus we should expect connections with neurons from different cortical layers. The electrode records axonal (not somatic) potentials. Because animal experiments have shown that 3 weeks or longer after implantation, all intercone fibers become myelinated (Kennedy et al., 1992), it is reasonable to assume that the signals recorded in this study were generated by myelinated axons.

Although the signal amplitude is in part dependent on the spatial separation between the axon and the electrodes, it appears likely that large amplitude units arise from large-diameter axons of corticospinal tract neurons, while small amplitude units may derive from axons of small inter-neurons. This is further supported by the observation that large-amplitude potentials fired slowly, while small-amplitude potentials fired at higher frequency: primate corticospinal tract neurons are known to fire more slowly than inter-neurons (Chang and Luebke, 2007; Gonzalez-Burgos et al., 2005; Zaitsev et al., 2009).

The four types of units illustrated in Figs. 2 and 3 are taken at rest, but the firing rates shown in Figs. 4–6 are analyzed during the vocalization (singing) task and are thus a better reflection of the active firing rates of the different neurons. We are not claiming they are maximal firing rates because we cannot be sure these are optimally task related. The results of studies of cortical neurons in primate (Chang and Luebke, 2007), cat (Chen et al., 1996), and rat (Schwindt et al., 1997) are not easily transferable to the present results in humans, as all the former studies were performed in slice preparations (including the monkey studies) using intracellular recording and stimulation. Nevertheless, they confirm the interpretation

of the rate/amplitude inverse correlation for neurons in different cortical layers, namely that fast firing units are interneurons and slow firing units are larger neurons that, presumably, give rise to the corticospinal tract.

The separation of multiunits has been hampered by the limited number of recording surfaces inside the cone tip. This disadvantage has long been recognized, and a new design under development involves microfabrication of a flexible polyimide micromachined version with 4, 8, 12, or 16 recording surfaces within the tip. This is expected to provide fewer units per channel easing the task of separation of units and, with a larger number of channels, increasing the yield per electrode. For example, with only 20 single units in ER's two-channel electrode, 160 units would be available if 10 units were available per channel in a 16-channel electrode. With fewer units per channel, separation into single units ought to be simplified.

Large-amplitude units

Units with amplitudes above 50 μV were found mainly in subject JR and used by him to drive the cursor (Kennedy et al., 2000). They were also seen in subjects DJ and ER as shown in Fig. 2b though they fired so infrequently as to be functionally useless in these subjects. They did, however, provide information during the listen/sing task discussed above. These large units were recorded at the same time as the smaller-amplitude units mentioned above in all subjects. Although the origin of these large slow-firing units is far from clear, they were stable across recording sessions and are therefore not likely to be artifacts. According to the rationale discussed above, these giant units probably originate from axons of corticospinal tract neurons. Precedence is found for these large units in monkey recordings using the NE. For example, Fig. 5 of Kennedy and Bakay (1997) illustrates that a large unit appeared only when the monkey was subject to both cortical stimulation close to the NE and simultaneously given caffeine. It is possible to speculate that similar circumstances also resulted in the appearance of the giant potentials in this study: subject JR required multiple medications which may have contributed to the consistent appearance of the large units described in the section "Results."

Stability

We distinguished single units from multiunits using ISIHs as shown in Fig. 4. The single units recorded near the beginning and at the end of the testing period for the presented data (days 1546–1556) are similar in firing rate characteristics shown in Figs. 2 and 3. The cluster parameter file was not changed during this sampling period, providing evidence for the stability of the recording. The duration of this stability is demonstrated by the ISIHs of single and multiunits acquired on day 1609, several months later (not illustrated). Data have been collected over many months during the first 5 years of implantation. It is highly unlikely that repeated detection of phonemes (Brumberg et al., 2011) and the gradual learning of vowel production over many months (Guenther et al., 2009) could have been done with unstable units, whether single or multiunits.

In addition, ISIH-identified single units had to remain stable to allow the gradual appearance of tuning over several sessions and within sessions as described in Fig. 6. It would not be

possible for such fine tuning of single units to occur if the overall system was unstable: if the system was unstable as *might* be suggested by the data from days 1546 and 1549 (Fig. 6), and the system then underwent stabilization *by chance* on days 1553 and 1556 to produce tuning of single units, then the multiunits ought to have stabilized and tuned, too. In this sense, the consistent differences between the multiunits and single unit recordings are an internal control for the stability of the recordings.

Some units within each of the four types of units preferentially demonstrated tuning as illustrated in Figs. 5a and b and described in the section “Results.” Of particular interest is that within each group, some units did not tune. One interpretation of this is that the unit tuning might be specific to the tone presented to the subject. The finding of differential responses to different tones producing tuning in the same unit tends to support this claim of specificity. An example of this effect is found by comparing unit ch2-19 in Fig. 5a responding to the 523-Hz tone with the same unit in Fig. 5b responding to the 262-Hz tone. Another interesting example is ch2-17: with tone 523 Hz, the firing rates for “listen” and sing are oppositely related to the rates produced during tone 262 Hz. It is also interesting that the rates tended to increase as the sessions progressed. This effect may have reflected adaptive changes related to recurring presentation of the stimulus, or the interest and enthusiasm of the subject perhaps. Of interest, also is the gradual reduction of firing rate in the two fastest units (ch1-09 and ch1-19) toward the end of the session when other units continued to fire vigorously and some to increase their firing rates especially during the “sing” period (examples include ch1-13, ch2-17, and ch2-09 [the unit fed back]). Of importance too is that this is a consistent effect seen for both tone presentations. One could speculate that these fast units are interneurons that are active during the tuning process, and once that process is complete, they are no longer required for effective tuning functionality so their firing rates decrease.

Feedback

Feedback is well documented in monkey studies as an important factor in learning to control devices such as virtual targets (Taylor et al., 2002) or robotic arms (Velliste et al., 2008). In the present studies, there were substantial changes in the firing rates over the four recording sessions. This plasticity appeared to be feedback dependent as shown in Fig. 6a where two sessions were devoted to no feedback and no plastic changes occurred. Presumably, the auditory feedback was directed through the external auditory nerve to the brain stem, lateral geniculate nucleus to the auditory cortex, and hence through Wernicke’s cortex to the expressive speech motor cortex, where we are presumably recording from neuritic processes of the neurons related to articulation. These are the same units that were active during phoneme detection (Brumberg et al., 2009) and vowel production (Guenther et al., 2009). The common variable between these three functions is articulation, so it seems reasonable to assume that in all studies, recordings were derived from units related to articulatory movements.

The role of unit ch2-09 in this feedback loop is of interest. The initial feedback session on day 1553 consisted of the production of the tone each time unit ch2-09 fired. The final session on day 1556 consisted of an increase in volume as unit ch2-09 fired. Even though

the activity of unit ch2-09 was used to generate the feedback, it was not the most vigorous firing unit as shown in the figure. Its rate did increase for the final trials during the 523-Hz tone presentation but not during the 262-Hz tone presentation. In addition, its firing rate during the control period was higher than during the listen or sing periods for both tone presentations. With *a priori* knowledge of the results, tone 523 Hz would have been better matched to unit ch2-17 and tone 262 Hz better matched to unit ch2-15, possibly producing even stronger modulations. One implication of these data is that the units *not* being used for feedback became more strongly modulated during the feedback sessions, suggesting they are specifically related to the articulatory movements required to produce that tone.

Auditory neurons

During the “listen” period, the subject was asked to passively listen to the tone. To remember it for the subsequent “sing” period, other processes would have to occur, including transfer of the auditory information into working memory. The data in Figs. 5 and 6 demonstrate differences in average firing rates between the listen and “sing” periods for both tones. The fact that some listen rates were also tuned may suggest that the tuning process was more than passive. This would appear to imply that “listen”-related neurons (auditory neurons) are present in cortical motor area in humans. Auditory neurons are not traditionally expected to be present in the motor area, but recent monkey data suggest that the neural dynamics of ventral premotor cortex are involved in the processing steps that link sensation and decision making during auditory discrimination (Lemus et al., 2009). In addition, multisensory units (vision and proprioception) have recently been reported in monkey area 4 (Suminski et al., 2009).

Mirror neurons

The possibility to an alternative interpretation to auditory-related neurons in motor cortex is that these units may arise from mirror neurons, and not from auditory units. The present study cannot definitively address this topic, but this possibility is raised because (a) the units may have fired in anticipation of “singing” (mirroring singing) and (b) the modulations during the “silent” control periods of the task were NOT similar, or only rarely similar, to that during active listening: if the listen period had been completely passive, listening and control firing rates should have been similar. This implies that the listening period modulations could have been something other than listening, and mirror neuronal activity is a possible explanation. There is some evidence in human imaging studies that mirror neurons exist in the human brain (Skipper et al., 2007; Turella et al., 2009). Attention and emotional factors may have played a role here as well (Kennedy, 2010).

Functionality of single versus multiunits

One enlightening basic observation was unexpectedly provided by examining a major issue in the field of neural prosthetics. This issue concerns the utility of multiunits as compared to single units. There is evidence that recorded single units are equivalent to multiunits when determining preferred directions of firing in a center-out task performed by monkeys (Frasier et al., 2009). These data were obtained from Rhesus Macaque primary motor cortex using a chronic recording system, the Utah Array (Rousche and Normann, 1998), which

recorded one or several units at the end of each of 96 tines. The task required the monkeys to move a cursor within a two-dimensional plane in eight directions under visual guidance. The recorded units were either sorted into single units or allowed to remain as multiunits. The results indicated no difference in performance between the use of single and multiunits, suggesting equivalence in the utility of such units. These results contrast with data presented here from human speech motor or pre-motor cortex recordings during the performance of listen/"sing" tasks in which we found that the firing rates of single units, but not multiunits, developed smoothly tuned patterns of firing related to auditory inputs, thus indicating non-equivalence of single and multiunits.

Neural prosthetic implications

Even though these data deal with basic electro-physiological characteristics of cortical neural activity, the results have important implications for the development of neural prosthetic devices. For prosthetic applications, the cortical control signals need to endure the lifetime of the subject, must be functionally stable, and must be trainable, that is, show plasticity. Ideally, such plasticity or adaption would occur with minimal training of the subject.

None of the existing chronic tine-type electrode designs have yet been shown capable of surviving many years of use with stable units. The 5-year life span to date of the NE is clearly a step in the right direction, however, and is substantially longer than that of other types of electrodes in humans (Hochberg et al., 2006) or monkeys (Carmena et al., 2003, Nicolelis et al., 2003; reviewed by Kennedy, 2006). Functional stability over time has been achieved with this and other electrodes, as highlighted in this chapter. It is also clear that "stability" in this context refers not simply to the ability to record units but to the ability to discriminate and hold single units to drive a function. The evidence for functional stability comes from the present results demonstrating that a much higher degree of specificity in auditory/vocalization tuning can be extracted from single neuron information than from information provided by multiple unit recordings. Thus, for functional stability in a neural prosthetic that requires *precise* control such as that of individual digits movements or speech, single units may be essential. However, for functions that do not require precise control, signals derived from multiunits, local field potentials, or EEG signals may be adequate.

Decoding algorithms work best with information from multiple simultaneously recorded signal sources. In this regard, the NE performs quite well, with 20 single units per electrode. However, design modification of the electrode presently underway will make it possible to implant many more electrodes of smaller size, record many more channels per electrode and even with fewer units per channel, the yield of single units will increase into the hundreds.

It is desirable that the recorded neurons be highly trainable, with minimal effort. Despite the long time spent training the subject in the use of the speech prosthesis (Brumberg et al., 2011; Guenther et al., 2009), the data presented here suggest that tuning can occur within a single session. This would encourage the view that training will be much briefer in future.

Thus, the NE is beginning to fulfill all the requirements for a neural prosthetic recording system. The present data provide evidence that single units continue to be recorded over years, that these units can be used to recognize phonemes, that these phonemes can be used to drive a speech synthesizer in real time, that feedback is important to produce coordinated firings of single units in a “sing” task, and that plasticity is possible. The data support the notion that long-term recording systems are more reliable when the neuropil grows into and becomes incorporated into the electrode than when an attempt is made to incorporate the electrode into the brain. Such developments in human prosthetics are expected to reveal a wealth of important basic electrophysiological cortical recording data which will drive developments of prostheses. In addition, it raises the probability of using reliable neural interfaces for neural enhancement. Thus in the future, with more electrodes and decoding paradigms, it may be possible to provide neural enhancement to both patients and nonpatients.

Acknowledgments

Support: This work was supported by a Grant from the NIDCD, NIH, R44 DC007050.

The authors would like to publicly acknowledge the continued strong support of this research by the subject and his family. We would like to particularly acknowledge the support of the subject's father who was present at all sessions reported here and who supplied encouragement to the subject and to us. The authors also acknowledge Prof. Thomas Wichmann for his critique of the chapter. We also acknowledge the role of Dr. Roy Bakay in the implantation of subject JR.

References

- Bartels JL, Andreasen D, Ehirim P, Mao H, Seibert S, Wright EJ, et al. Neurotrophic electrode: Method of assembly and implantation into human motor speech cortex. *Journal of Neuroscience Methods*. 2008; 174(2):168–176. [PubMed: 18672003]
- Berger MS. Minimalism through intraoperative functional mapping. *Clinical Neurosurgery*. 1996; 43:324–337. [PubMed: 9247814]
- Breshears J, Sharma M, Anderson NR, Rashid S, Leuthardt EC. Electrocorticographic frequency alteration mapping of speech cortex during an awake craniotomy: Case report. *Stereotactic and Functional Neurosurgery*. 2009; 88(1):11–15. [PubMed: 19940544]
- Brumberg JS, Andreasen D, Bartels JL, Ehirim P, Guenther FH, Mao H, et al. Classification of intended phoneme production from chronic intracortical microelectrode recordings in speech motor cortex. *Journal of Neural Engineering, Frontiers in Neuroscience*. 2011; 5(65):1–14.
- Brumberg, JS.; Kennedy, PR.; Guenther, FH. Proceedings of the 10th annual conference of the international speech communication association. presented at the interspeech 2009. Brighton, UK: International Speech Communication Association; 2009. Artificial speech synthesizer control by brain-computer interface.
- Carmena JM, Lebedev MA, Crist RE, O'Doherty JE, Santucci DM, Dimitrov DF, et al. Learning to control a brain-machine interface for reaching and grasping by primates. *PloS Biology*. 2003; 1(2): 193–208.
- Chang YM, Luebke JI. Electrophysiological diversity of layer 5 pyramidal cells in the prefrontal cortex of the rhesus monkey: in vitro slice studies. *Journal of Neurophysiology*. 2007; 98(5):2622–2632. [PubMed: 17804579]
- Chen W, Zhang JJ, Hu GY, Wu CP. Electro-physiological and morphological properties of pyramidal and non-pyramidal neurons in the cat motor cortex in vitro. *Neuroscience*. 1996; 73(1):39–55. [PubMed: 8783228]
- Frasier GW, Chase SM, Whitford A, Schwartz AB. Control of a brain-computer interface without spike sorting. *Journal of Neural Engineering*. 2009; 6(5):055004. [PubMed: 19721186]

- Gonzalez-Burgos G, Krimer LS, Povysheva NV, Barrioneuvo G, Lewis DA. Functional properties of fast spiking interneurons and their synaptic connections with pyramidal cells in primate dorsolateral frontal cortex. *Journal of Neurophysiology*. 2005; 93(2):942–953. [PubMed: 15385591]
- Guenther FH, Brumberg JS, Wright EJ, Nieto-Castanon A, Tourville JA, Panko M, et al. A wireless brain-machine interface for real-time speech synthesis. *PLoS One*. 2009; 4(12):e8218. [PubMed: 20011034]
- Hochberg LR, Serruya MD, Friehs GM, Mukand JA, Saleh M, Caplan AH, et al. Neuronal ensemble control of prosthetic devices by a human with tetraplegia. *Nature*. 2006; 442(7099):164–171. [PubMed: 16838014]
- 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. 3. 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?. The electrical engineering handbook series
- Kennedy PR. Changes in emotional state modulate the firing rates of human neural signals during long-term recording with the neurotrophic electrode. *Neurocase*. 2010
- Kennedy PR, Andreasen DS, Ehirim P, King B, Kirby T, Mao H, et al. Using human extra-cortical local field potentials to control a switch. *Journal of Neural Engineering*. 2004; 1:72–77. [PubMed: 15876625]
- Kennedy PR, Bakay RAE. Activity of single action potentials in monkey motor cortex during long-term task learning. *Brain Research*. 1997; 760:251–254. [PubMed: 9237542]
- 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, Moore MM, Adams K, Goldwithe J. Direct control of a computer from the human central nervous system. *IEEE Transactions on Rehabilitation Engineering*. 2000; 8(2):198–202. [PubMed: 10896186]
- Kennedy PR, Kirby T, Moore MM, King B, Mallory A. Computer control using human intra-cortical local field potentials. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. 2004; 12(3):339–344. [PubMed: 15473196]
- Kennedy PR, Mirra S, Bakay RAE. The cone electrode: Ultrastructural studies following long-term recording. *Neuroscience Letters*. 1992; 142:89–94. [PubMed: 1407726]
- Kurzweil, R. *The Singularity is Near*. Penguin Group; 2005.
- Lemus L, Hernández A, Romo R. Neural encoding of auditory discrimination in ventral premotor cortex. *Proceedings of the National Academy of Sciences of the United States of America*. 2009; 106(34):14640–14645. [PubMed: 19667191]
- Nicolelis MAL, Dimitrov D, Carmena JM, Crist R, Lehew G, Kralik JD, et al. 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]
- Paninski L, Fellows MR, Hatsopoulos NG, Donoghue JP. Spatiotemporal of motor cortical neurons for hand position and velocity. *Journal of Neurophysiology*. 2004; 91:515–532. [PubMed: 13679402]
- 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]
- Schwindt P, O'Brien JA, Crill W. Quantitative analysis of firing properties of pyramidal neurons from layer 5 of rat sensorimotor cortex. *Journal of Neurophysiology*. 1997; 77(5):2484–2498. [PubMed: 9163371]
- Serruya MD, Hatsopoulos NG, Paninski L, Fellows MR, Donoghue JP. Instant neural control of a movement signal. *Nature*. 2002; 416(6877):141–142. [PubMed: 11894084]
- Skipper JJ, Goldin-Meadow S, Nusbaum HC, Small SL. Speech-associated gestures, Broca's area, and the human mirror system. *Brain and Language*. 2007; 101(3):260–277. [PubMed: 17533001]
- Suminski AJ, Tkach DC, Hatsopoulos NG. Exploiting multiple sensory modalities in brain-machine interfaces. *The Journal of Neuroscience*. 2009; 30(50):16777–16787. [PubMed: 21159949]
- Taylor DM, Tillery SI, Schwartz AB. Direct cortical control of 3D neuroprosthetic devices. *Science*. 2002; 296(5574):1829–1832. [PubMed: 12052948]

- Turella L, Pierno AC, Tubaldi F, Castiello U. Mirror neurons in humans: Consisting or confounding evidence?. 2009; 108(1):10–21. [PubMed: 18082250]
- 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]
- Zaitsev AV, Povysheva NV, Gonzalez-Burgos G, Rotaru D, Fish KN, Krimer LS, et al. Interneuron diversity in layers 2–3 of monkey prefrontal cortex. *Cerebral Cortex*. 2009; 19(7):1597–1615. [PubMed: 19015370]

Fig. 1.

(a) Overview of the neurotrophic electrode tip. All three wires are coiled for strain relief. Neurons send neurites into and through the cone. Arrows indicate the positions of the wire tips, 500 μm apart. (b) Two FM receiving coils are shown on the left side of subject ER's head. The power induction coil is on the right (not shown). All coils are held in position with EC2 white electrode paste (Grass Electronics Inc.). (c, d) The paradigm used in these studies is described in detail in the text.

Fig. 2.

(a) Left panel: 42 ms of continuous multichannel data from subject DJ from the three channels recorded 75 days after implantation on October 25, 2002. Right panel top: 42 ms of continuous multichannel data from subject JR 122 days after implantation. Right panel bottom: examples of 42 ms of continuous multiunit data from each channel in subject ER. Possible single units are labeled as shown. The letter labels do not correspond to units in later figures and are meant only to highlight wave shapes similar in amplitude recorded on day 1556 after implantation. (b) Top panel: units from subject DJ recorded on day 75. Middle panel: units recorded from JR on day 266. Note the various amplitudes of larger units that fired

infrequently. Bottom panel: units from subject ER that fired infrequently on only one channel (blue). There is physiological crosstalk between the channels, but not with the third channel (not shown). See next figure and text for explanation. (c) Physiological crosstalk between channels 1 and 2 inside the electrode. Note the depolarization spikes face each other. For full explanation see text.

Fig. 3.

(a) Firing rate histograms of units with different amplitudes are illustrated for the large units illustrated in Fig. 2a for subject JR. Note the scale differences between these data and those in (b). (b) Firing rate histograms of units illustrated in Fig. 2a for subject ER. Data recorded at rest. Note the scale difference between unit sizes in (a) and (b).

Fig. 4. Interspike interval histograms (ISIH) of 20 units whose rates fall to zero between zero time and 1 ms, implying that each one is a single unit. Data are from day 1602 (May 22, 2009), subject ER.

(b)

Fig. 5.

(a) Averaged firing rates of the four classes of units during the listen, control, and sing periods appeared as mainly “V”-shaped triplets for tone 523 Hz. (b) Similar data presentation for firing rates for tone 262 Hz. (c) Similar data presentation for multiunit data for one 523 Hz. (d) Similar data presentation for multiunit data for 262 Hz. All data are from day 1556 in subject ER.

Fig. 6.

(a) Data from units followed over four sessions demonstrating the tuning over that period with no feedback (days 1546 and 1549), unit feedback (day 1553), and feedback of volume directly related to unit firing frequency (day 1556). The best tuned units in each of the four classes are illustrated here. Data for tone 523 Hz. (b) Similar data form days 1449 (no data for 1446), 1553, and 1556 for tone 262 Hz. (c) Plots of mean and 1 SD of the variance of firing frequencies during the listen control and sing data over the 3 days, 1549, 1553, and 1556, for tone 262 Hz. Details in text: p values are within each plot. Note the statistically significant decrease in variance between data on day 1549 compared to data on day 1556.

Table 1

List of implanted subjects

Subjects	Age	Disease	Residual capabilities	Arm hand	Speech	Date of implant	Duration	Outcome
ER	26	BSS	Eyes up slowly for "yes," down for "no"		***	Dec 22, 2004	6+ years	Ongoing speech decoding
DJ	42	ALS	Slight movement in right hand, face, eyes	***		Nov 25, 2002	3 months	Single units/EMG
TT	39	MM	Eye movements	***		Jul 24, 1999	4.5 years died	Too diseased
JR	53	BSS	Slight facial and eye movements	***		Dec 24, 1998	4 years died	First Cyborg
MH	54	ALS	Slight eye movements	***		Dec 4, 1996	76 days died	Controlled unit firings

area of motor cortex implanted.