Chapter 43. Intracranial electrophysiology in language research

Adeen Flinker, Vitória Piai, and Robert T. Knight

43.1 Introduction

Intracranial electrophysiological recording in humans has been a long standing technique in neurosurgical treatment for epilepsy and have served as an important window in to how the human brain processes language. This chapter is aimed to introduce the reader to the technique, how it historically contributed to language mapping, its advantages and disadvantages as a research tool, and analysis techniques that have provided novel findings and approaches in the area of language processing.

43.2 History of epilepsy surgery

The pioneering work of Paul Broca and Carl Wernicke localizing speech deficits to the inferior frontal gyrus (IFG) (Broca, 1861) and posterior temporal lobe (Wernicke, 1874) sparked a new era in the neurobiology of language, as well as greatly influencing subsequent approaches to neurosurgical treatment of epilepsy. Broca’s observation together with the finding that electrical stimulation of cortex can produce specific motor responses (Hitzig, 1900) paved the road to consider epilepsy a form of irritated tissue and treatable by resection of the disruptive tissue (Feindel et al., 2009). Cortical resection became a safe and effective treatment for focal epilepsy based on stimulation mapping of motor function as well as epileptic foci (Horsley, 1909; Krause & Thorek, 1912). These stimulation techniques were augmented after the discovery of Electroencephalography (EEG) by Hans Berger (Berger, 1929; 1931) and its acceptance by the international community (Adrian & Matthews, 1934; Feindel et al., 2009). EEG proved to be a powerful tool in the classification and diagnosis of epileptic activity (Gibbs et al., 1936; Jasper & Kershman, 1941) and was first recorded intraoperatively in humans under local anesthesia while subjects were awake and stimulation mapping was performed (Feindel et al., 2009; Foerster & Altenburger, 1935).

Electrocorticography (ECoG) or intracranial EEG (iEEG) soon became a routine technique in aiding neurosurgeons during cortical mapping (Jasper et al., 1951; Penfield & Erickson, 1941). During these early surgeries stimulation mapping provided great insight into the topographic organization of motor and somatosensory cortices culminating in the famous cortical homunculus diagram (Penfield & Boldrey, 1937). While mapping of motor cortex was prevalent since the turn of the twentieth century (Cushing, 1909), Penfield and colleagues in the Montreal Neurological Institute (MNI) were the first to systematically extend stimulation mapping to language and devised protocols that remain largely unchanged to this day (Penfield & Roberts, 1959). During a typical electrical stimulation mapping (ESM) procedure, the patient is awakened from anesthesia while cortex is still exposed and is asked to perform the same tasks Penfield first employed over 80 years ago: counting and picture naming. While the patient engages in counting or picture naming, different cortical sites are repeatedly stimulated. If the applied electrical current reliably disrupts the patient’s speech output, or causes naming errors with intact speech output, the cortical site is deemed to be critical for language and is spared from resection. ESM has been invaluable in sparing neo cortical sites critical for language, but these sites have been found over distributed regions of frontal, temporal, and parietal cortex and with a high degree of intersubject variability (Ojemann et al., 1989; Sanai et al., 2008). Even though counting and picture naming tasks tap into a relatively small subset of language functions, sparing of sites identified by cortical stimulation dramatically reduces postoperative deficits and the procedure remains the gold standard in the field (; Chang et al., 2014; Sanai et al., 2008).

43.3 Intracranial EEG monitoring

ESM and definition of the epileptogenic zone are mostly performed in the operating room, as pioneered by Penfield and colleagues. However, this constrains the time available for testing and monitoring the patients. In cases where the epileptic zone is unclear, patients are implanted with an array of electrodes for a period
spanning several days to weeks (Ojemann, 1987; Wieser & Elger, 2012). During this time, intracranial EEG is recorded from the electrodes to more clearly define epileptic zones as well as map function via electrical stimulation. After clinical assessment has completed, the electrodes are removed, and epileptic tissue is resected. In addition to providing a more refined mapping of cognitive function for clinical purposes, the extended time these patients spend in the monitoring unit provides ample opportunity to engage with patients in research studies (Crone, 2000). During lulls in clinical treatment, patients may consent to different cognitive and language tasks at their convenience in a bedside setting. The acquisition of electrical signals directly from cortex provides a valuable opportunity to explore language function and while electrode locations are solely guided by the clinical necessity of the patient, they often sample perisylvian regions. During the late 1990s there was a surge of research performed in the chronic bedside setting. The most common procedure in North America is the implantation of subdural grids and strips, referred to as electrocorticography (ECoG), which provide a wide spatial coverage on the surface of cortex but are typically limited to one hemisphere and the surface of cortex surrounding the area of the craniotomy (Gonzalez Martinez et al., 2013; Risinger & Gummit, 1995; Widdess Walsh et al., 2007). In cases where deeper structures, such as the hippocampus, amygdala, orbitofrontal cortex, or insula, need to be monitored for epileptic activity depth electrodes with several contacts are implanted as well (Ojemann, 1987). An alternative approach was developed in the Paris school of Talairach and colleagues whereby depth electrodes are implanted stereotactically, stereotactic EEG or sEEG, enabling the monitoring of epileptic activity across the depth of cortical structures in both hemispheres and is more widely adapted in Europe (Crandall & Babb, 1993; Lüders, 2008; Talairach et al., 1958). Both have clinical advantages and disadvantages (Wellmer et al., 2012) and can be complementary approaches depending on the hypothesized epileptogenic zone (Mullin et al., 2016).

43.4 Methodological considerations

Due to the nature of epilepsy, that often involves temporal lobe structures, as well as the clinical necessity to map critical language sites prior to resection of tissue, electrode coverage often involves perisylvian cortices. This cortical sampling as well as the combined temporal and spatial resolution provided by intracranial recordings offer a unique opportunity to address language processing. Compared with non-invasive electrophysiology (EEG, MEG), intracranial recordings do not suffer from volume conduction outside of cortex providing a much higher signal to noise ratio, a spatial resolution limited only by the electrode spacing (typically 1 cm; Lachaux et al., 2003) as well as a lower susceptibility to muscle and eye movement artifacts (Crone, 2000). Both sEEG and ECoG are largely immune to muscle artifacts from eye, jaw, lip and tongue making them ideal for speech production studies (Llorens et al., 2011) as well as research on reading. Nevertheless, contamination from eye movements has been reported especially in the vicinity of the temporal pole and orbitofrontal cortex (Ball et al., 2009; Jerbi et al., 2009; Kovach et al., 2011), requiring some care classifying electrodes that may be prone to the effect. While intracranial recordings provide an excellent combined temporal and spatial resolution which cannot be established by noninvasive electrophysiology (EEG, MEG) or neuroimaging (fMRI, PET) alone, they do hold several drawbacks. Typical ECoG coverage involves a grid of electrodes (typically 8 × 8 with 1 cm spacing) as well as depth electrodes and supplementary strips depending on the clinical necessity (Fig. 43.1, right). The insertion of these electrodes requires a craniotomy and is limited to one hemisphere (although some limited sampling of the other hemisphere may be provided by insertion of depth or strip electrode via burr holes). This limits the spatial sampling to mainly one hemisphere and constrains within subject approaches for hypotheses concerning function of both hemispheres, such as hemispheric asymmetry. Conversely, in sEEG the typical coverage involves depth electrodes bilaterally (burr holes bilaterally without a craniotomy) but the sampling is sparse and limited to the number of depth probes inserted (Fig. 43.1, left). Lastly, intracranial research is most commonly recorded from patients with epilepsy that may have had some degree of cortical reorganization during development to compensate for epileptic tissue. To address this issue, it is important
to remove all epileptic electrodes from research analysis and include multiple subjects with varying epileptic sources. Signals acquired directly from cortex in animal models allow for varying degree of invasiveness as well as insertion of electrode probes on multiple scales. In human intracranial research the most common electrodes provide a sampling of several millimeters of cortex. The typical clinical electrode is 4.0 mm in diameter and has 2.3 mm exposed to cortex with 1 cm interelectrode spacing (smaller and higher density pediatric grids are also used with 0.4–0.5 mm interelectrode spacing). These electrodes record activity from large neuronal populations beneath the area of cortex exposed to the electrode (Wieser & Elger, 2012). In some cases, small microwire electrodes (20–50 µm diameter) are used which produce a signal typically referred to as a Local Field Potential (LFP)—although the nomenclature is freely used to reflect various recording techniques from cortex (Buzsáki et al., 2012) and the extent to which LFP is spatially limited is still debated (Kajikawa & Schroeder, 2011). The LFP signal is extremely rich with both broadband information (i.e., delta, theta, alpha, beta, gamma rhythms and high gamma activity >70 Hz and up to 250 Hz) as well as activity in the kHz range which can be used to extract multiunit activity and in some instances individual neurons (single unit activity). The use of microwire recordings is rare in language research although there have been several reports from anterior temporal cortex, including neuronal responses in the superior and middle temporal gyri to specific phonetic cues in auditory sentences (Creutzfeldt et al., 1989), spatially organized tonotopic responses in Heschel’s gyrus (Howard et al., 1996a), robust tracking of compressed speech (Nourski et al., 2009), as well as visual word recognition in the ventral temporal lobe (Halgren et al., 2006; 2015). Microwire electrodes are typically part of a larger clinical depth electrode with microcontacts at the tip using a hybrid approach (Howard et al., 1996b).

Fig. 43.1 sEEG and ECoG typical coverage. Superior view of 3D reconstructions of computed tomography (CT) scans in a sEEG (left) and ECoG (right) implantation of intracranial electrodes.

43.5 Intracranial event related potentials

Early intracranial recordings and analysis were limited to evoked potentials. Fried et al. recorded intraoperative event related potentials (ERP) during presentation of pictures. They found that premotor sites (likely IFG) showed a sustained negativity that was specific to a rhyming task compared with a visual angle discrimination task (Fried et al., 1981). In contrast to ECoG surface and depth recordings, sEEG provides routine coverage of both hemispheres. Liégeois Chauvel et al. localized early auditory evoked responses to specific regions of Heschel’s gyrus (Liégeois Chauvel et al., 1991). Additionally, the authors
provided direct electrophysiological evidence for a left hemisphere specificity to voice onset time (Liégeois Chauvel et al., 1999) and for a hemispheric difference in processing temporal modulations (Liégeois Chauvel et al., 2004). Similarly, deep brain regions such as the medial temporal lobe, and the hippocampus in particular, can only be reached with depth electrodes. Previous studies have focused on different research questions regarding the involvement of medial temporal lobe regions in language. McCarthy et al. tested patients reading correct sentences and sentences ending with a semantic violation (e.g., “I ordered a ham and cheese scissors.”) (McCarthy et al., 1995). A negative deflection was observed in the ERPs for the semantic violations relative to the correct sentences, peaking around 400–500 ms post target word onset. This effect was focal, found in contacts in the anterior medial temporal lobe bilaterally. In another study (Nobre & McCarthy, 1995), participants saw word(-like) stimuli presented individually on a screen. Their task was to press a button when a word was of a particular category (e.g., a body part). In one experiment, the stimuli presented consisted of different types of words: orthographically illegal non words (e.g., gtprlm), pseudowords (e.g., glubbalt), function words (e.g., hence), and content words (e.g., truck). In another experiment, all stimuli were content words that were paired as prime target words. Primes could either be semantically related or unrelated to the target words. Content words (including the primes) elicited a negative deflection in the ERPs peaking around 400 ms post stimulus onset in the anterior medial temporal lobe. The amplitude of this negative deflection decreased for pseudowords, and further decreased for function words and non-words. An attenuation of this negative deflection was also observed for semantic ally primed target words. In both studies, the N400 like potentials measured in contacts in the anterior medial temporal lobe likely reflects activity from the anterior fusiform and parahippocampal gyri, but not from the hippocampus proper. Meyer et al. had their participants read three types of sentences: correct sentences (e.g., “The door was being closed”), sentences containing semantic violations (e.g., “The ocean was being closed”), and sentences containing syntactic violations (i.e., a phrase structure violation, e.g., “The shop was being on closed”) (Meyer et al., 2005). Recordings from the rhinal cortex revealed a negative deflection in the ERPs between 200 and 500 ms post target word onset for the correct and semantically incorrect sentences that was absent in the sentences with syntactic violations. By contrast, hippocampal recordings revealed a negative deflection between 500 and 800 ms for the syntactic violations that was absent in the other two types of sentences.

43.6 Intracranial time frequency signatures

In a series of ECoG papers Crone et al., investigated how different frequency bands can index and track cognitive function across a range of language (Crone et al., 1994; 2001a) and motor tasks (Crone et al., 1998a; 1998b). Alpha power decreases were found in three language production tasks, namely picture naming, word reading, and word repetition. These effects were found in electrodes over sensorimotor regions, superior temporal gyrus, and basal temporal occipital cortex. The early studies of Crone and colleagues focused on the alpha and beta frequency bands given their well-known relation to the sensorimotor cortex. In the memory domain, however, the theta band has been consistently linked to mnemonic processes in medial temporal lobe regions (Buzsáki & Moser, 2013; Lega et al., 2012; Rutishauser et al., 2010). Based on this evidence, Piai and colleagues focused on the theta frequency band in the medial temporal lobe during a language task. They employed a sentence completion task in which participants named pictures that completed semantically constrained (e.g., “She swept the floor with a,” [picture: BROOM]) or neutral sentences (Piai et al., 2016). Piai et al. observed increases in theta power for semantically constrained relative to neutral sentences during sentence processing, preceding picture presentation. Figure 43.2 shows this effect for 10 different patients. This effect was found not only in the hippocampus proper, but also in the parahippocampal gyrus and entorhinal cortex. These results provided evidence that medial temporal lobe structures contribute to language processing online, relating words in the sentence to stored semantic knowledge. Moreover, this process seems to be supported by the same neuronal computations performed by these structures for memory function, as reflected in theta oscillations.
43.7 Discovery of high gamma activity

One of the key findings in the studies of Crone and colleagues was the existence of high gamma activity (70–150 Hz) that tracked cortical processing and provided a more robust and localized within subject index compared with low frequency desynchronization and ERPs (Crone, 2000; Crone et al., 2001b; 2006). While the extent to which high gamma power reflects band limited oscillations or broadband activity has been debated, it has been successfully employed across cognitive domains to track cortical activity (Jacobs et al., 2010; Crone et al., 2011; Miller, 2010). High gamma activity has been shown to be correlated with the spiking rate of underlying neurons as well as coupled to the hemodynamic BOLD response in both animal (Allen et al., 2007; Belitski et al., 2008; Ray et al., 2008) and human cortex (Mukamel et al., 2005; Nir et al., 2007). These findings have been followed by a growth of intracranial research elucidating cognitive function while leveraging this high frequency band (HFB) signal which is hard to detect outside the skull. The high frequency signal (referred to as high gamma, HFB and sometimes broadband signal) has become one of the most common cortical indices used to map cognitive function as well as track the perception and production of language.
Fig. 43.2 Context effect (constrained vs. neutral) time locked to picture presentation. Significant theta effect is shown in stronger colors (multiple comparisons corrected). Trial events are shown at the bottom. The timing of each word position is indicated by the continuous lines. The left end of each line indicates the earliest possible word onset. The right end indicates the latest possible word offset (and next word onset). Median word onset (and previous word offset) is indicated by the orange vertical bars.

Adapted from Vitória Piai, Kristopher L. Anderson, Jack J. Lin, Callum Dewar, Josef Parvizi, Nina F. Dronkers, and Robert T. Knight, Direct brain recordings reveal hippocampal rhythm underpinnings of language processing, Proceedings of the National Academy of Sciences of the United States of America, 113 (40), pp. 11366–71, Figure 3, doi: 10.1073/pnas.1603312113 ©2016 Vitória Piai, Kristopher L. Anderson, Jack J. Lin, Callum Dewar, Josef Parvizi, Nina F. Dronkers, and Robert T. Knight. This work is licensed under the Creative Commons

The high gamma signal allows for a unique window into language function as it provides a signature for neural activity that is robust on the level of single trials with high temporal resolution as well as activation signatures similar to fMRI. In non-invasive electrophysiology, the high gamma signal is virtually inaccessible due to volume conductance effects and a sharp drop in power in higher frequencies when passing the skull. Furthermore, unlike neuroimaging and noninvasive electrophysiology, the position of electrodes directly on
(or in) the brain renders intracranial signals minimally contaminated by motor artifacts from speech production or movement. Using a picture naming task and sEEG recordings from the medial temporal lobe, Hamamé et al., observed increases in high gamma power between picture presentation and initial articulation in the hippocampus (Hamamé et al., 2014). The peak latency of the high gamma activity correlated with the participants’ picture naming latencies. Finally, tip of the tongue states (i.e., when a speaker recognizes an object but cannot retrieve its name) were associated with no increases in hippocampal high gamma activity relative to a pre stimulus baseline. Recordings from the surface of the temporal cortex have been mostly focused on acoustic, phonetic, and lexical levels of processing. Studies have established strong high gamma responses to auditory stimulation in the superior temporal gyrus (STG) when presented with non-speech stimuli (Crone et al., 2001a; Brugge et al., 2009; Edwards, 2005), phonemes (Chang et al., 2011; Crone et al., 2001a; Flinker et al., 2010; Fukuda et al., 2010) and words (Brown et al., 2008; Canolty et al., 2007; Crone et al., 2001b; Edwards et al., 2010; Flinker et al., 2011; Pei et al., 2011). Sites on the lateral surface of the STG respond more robustly as the hierarchy of linguistic input is increased, that is, phonemes > tones (Crone et al., 2001a), words > phonemes (Flinker et al., 2011) and words > non words (Canolty et al., 2007). One of the striking aspects of high gamma responses is the ability to track activity within a subject, both temporally and spatially on the level of single trials. Cortical activity while listening to words (contrasted with an acoustic control) has been shown to propagate across the STG to the superior temporal sulcus (STS) (Canolty et al., 2007) and when production is required follows to the IFG motor cortices (Brown et al., 2008; Edwards et al., 2010; Flinker et al., 2015; Fukuda et al., 2010; Pei et al., 2011; Towle et al., 2008). Flinker et al., investigated auditory word repetition and found that Broca’s area was active as early as 200 ms poststimulus onset and activity was absent by the time the word was articulated (Fig. 43.3). This early pattern of activity in Broca’s area, commencing prior to articulation, and fading by the time of articulation, was evident across a range of tasks including repetition of auditory mono syllabic words, repetition of auditory multisyllabic words and overt word reading (Flinker et al., 2015). Early activity in Broca’s area has also been reported during lexical and inflectional processing of words without overt articulation (Sahin et al., 2009). The monosyllabic words in employed by Flinker et al., were comprised of both real words (e.g., hope) as well as pronounceable pseudowords matched for phonotactic probabilities (e.g., yode). Articulation of a novel sequence of phonemes elicited more activity in Broca’s area compared with real words but this was not the case in motor cortex. These findings taken together with the early window of activity implicate Broca’s area in the formation of an articulatory plan rather than coordination of the articulators themselves. The authors also leveraged the high gamma signals together with Granger causal connectivity analyses to show the directional flow of cortical activity during word repetition. During perception of the auditory word there was a peak of Granger causal flow from STG to Broca’s area followed by reciprocal feedback from Broca’s area to STG. This feedback influence from Broca’s area onto STG was evident until 200 ms prior to articulation onset and in parallel there was Granger causal flow from Broca’s area to motor cortices lasting up to articulation onset as well as feedback from motor cortices onto Broca’s area. These temporal dynamics reveal a key role Broca’s area plays in manipulating and forwarding cortical representations as a heard word is processed and transformed into an articulatory code. Individual electrodes over the posterior STG (pSTG) have often shown responses to a linguistic unit regardless of input modality when multiple tasks were employed, such as word reading and auditory repetition (Crone et al., 2001b; Flinker et al., 2010) as well as
Fig. 43.3 Repetition of monosyllabic words in a representative subject. (A) Event related spectral perturbations (ERSPs), averaged across trials, and locked to the onset of auditory word stimulus. Cortical activation indexed by power increases in high frequencies is first apparent in STG during word perception, subsequently in Broca’s area, and finally extends to motor cortex during word production (vertical lines mark mean articulation onset). (B) High frequency power (γhigh, 70–150 Hz) traces, averaged across trials, and locked to word stimulus onset are shown for STG (blue), Broca (green), and motor (red) electrodes. The first electrode in every pair is marked by a black circle and corresponds to the ERSP plotted on the left. The shaded gray area marks the distribution of articulation onset for this subject (1 SD in each direction).

picture naming (Crone et al., 2001b). Similarly, this invariability has been reported across languages in multilingual patients performing a picture naming task where a majority of STG responses were common to both L1 and L2 speech (albeit some sites with L2 specificity) (Cervenka et al., 2011). An interesting finding arising from intracranial research has been the rich spatial variability of responses in STG whether exhibiting modality invariance or selectivity to linguistic hierarchy. Flinker et al. found neighboring sites 4 mm apart with distinct functional responses: one site responding to both syllables and words and its neighbor selective to words (Flinker et al., 2011). Figure 43.4 shows the distribution of responses over STG in one subject listening to words or phonemes (CV syllables). Similarly, some sites in STG showed a typical suppression during production of speech while neighboring sites showed a selectivity for speech output (Flinker et al., 2010; 2011). While STG reduced activity during speech production has been consistently reported (Crone et al., 2001b; Edwards et al., 2010; Flinker et al., 2010; Fukuda et al., 2010; Pei et al., 2011; Towle et al., 2008), the topography of suppression can vary across the STG during auditory feedback as well as pitch perturbation (Chang et al., 2013; Flinker et al., 2010; Greenlee et al., 2011). This rich topography of responses is sometimes only visible when employing higher density electrode arrays. Chang et al., showed that the spatial topography of pSTG responses encoded the phonetic category of CV syllables rather than
the acoustic continuum used to synthesize them (Chang et al., 2010). Similarly, Mesgarani et al. used high density ECoG arrays to decode responses during auditory sentences that spanned the English phonetic lexicon. They show that individual electrodes are highly selective to different phonemes and encode acoustic phonetic features. Similarly, high density recordings from Rolandic cortex (pre and postcentral gyri) have shown a somatotopical organization of speech articulators which are phonetically grouped (Bouchard et al., 2013) as well as a surprising acoustic receptive field representation within motor cortex (Cheung et al., 2016). The discovery of the high gamma signal together with leveraging computational approaches has sparked avenues of research that were previously mostly limited to animal research. For example, the spectrotemporal receptive field (STRF) has been traditionally limited to modeling the relationship between acoustic input and neuronal firing rates. Pasley et al., were the first to use high gamma activity in STRF modeling providing evidence for a robust acoustic representation in STG which consistently reconstructs neural responses based on the acoustic input (Pasley et al., 2012). Similar techniques have been used to elucidate the acoustic phonetic organization in STG (Mesgarani et al., 2014) and how auditory cortex enhances speech intelligibility by rapid tuning of the underlying receptive fields (Holdgraf et al., 2016). In addition to these encoding models (predicting neural activity based on stimulus features) there has been great interest in decoding activity from cortex (predicting a stimulus class based on the neural activity). Pasley et al. provided a reconstruction of the auditory sentences played to the patients based solely on the neural activity in STG (Pasley et al., 2012). Similar approaches are being used to decode imagined speech (Martin et al., 2016a; 2016b) and try to drive a speech prosthesis (Leuthardt et al., 2011). Decoding techniques have also been employed in decoding sensory motor transformations providing evidence that speech transformations occur bilaterally (Cogan et al., 2014).
43.8 Future directions

Intracranial electrophysiology provides an exciting opportunity to explore language processing in the human cortex with high temporal and spatial resolution. While much progress has been made, there still remain many unresolved questions and unaddressed areas in language processing. High gamma has provided an unprecedented index of cortical activity, but its relationship to low frequency and ERP findings in the non-invasive language literature remains understudied. Moreover, a great number of intracranial studies focus on sublexical and lexical processing and there is a lack of research studying higher order linguistic operations. Recent studies have leveraged the high gamma signal in advanced machine learning techniques to reconstruct signals based on sets of features. Such approaches are valuable and illuminating but very often are limited to acoustic or sublexical features. For example, acoustic STRF modeling (also known as a type of an encoding model) predicts the neural signal based on the auditory input (e.g., speech) the patient heard. When the model is robust then one can reliably reconstruct the neural signal given a new set of speech stimuli, but such models are limited to the time frequency acoustics of the signal and do not necessarily account for phonetic, sublexical, lexical, semantic, and grammatical structure. That said, such modeling techniques could easily incorporate higher order linguistic structures. Similarly, machine learning decoding techniques that try to predict a target class or variable (e.g., acoustic spectrogram, phoneme, and so on) given a set of neural signals could be augmented to classify lexical, semantic, and syntactic representations. Critically these models have to be driven by and interpreted within current (or novel) theoretical frameworks of language processing. There has been increasingly interest in the role of the medial temporal lobe in language (Covington & Duff, 2016; Duff & Brown Schmidt, 2017), partly fueled by recent findings from sEEG research (Hamamé et al., 2014; Jafarpour et al., 2017; Llorens et al., 2016; Piai et al., 2016). The combination of excellent temporal and spatial resolution of iEEG recordings is particularly important for understanding the role of these deeper regions, as they cannot be easily recorded from with scalp based techniques and fMRI approaches are limited in temporal resolution necessary to track the dynamics of speech and language. Future evidence from iEEG recordings will likely greatly expand our understanding of the relations between the language and memory systems. The finding that Broca’s area is not involved during articulation per se but rather coordinating articulatory planning is a striking example of the unique advantage intracranial studies provide. Previous non-invasive electrophysiological studies were limited by both the spatial specificity of neural generators as well as motor artifacts during speech production. In contrast, neuroimaging studies of speech production could resolve activity within Broca’s area but could not elucidate the exact timing and stage of recruitment. The combined temporal and spatial resolution provided by intracranial recordings, together with minimal contamination from motor artifacts, ideally situates it to elucidate the role of Broca’s area during speech production. Nevertheless, such an endeavor requires more than simple speech production tasks such as word repetition and overt reading. Ideally, spontaneous speech and longer, more complex utterances should be investigated. Such an endeavor is not trivial, as it requires novel speech production paradigms that prompt continuous segments of spontaneous speech that are experimentally controlled and are optimally recorded in tandem with neural recordings. One approach that is likely to benefit both speech production research as well as impact patient care is employing decoding and machine learning techniques on large corpora of speech together with the intracranial signals in Broca’s area and motor cortices. Such an approach could reveal receptive field properties in frontal cortices while providing a stepping stone toward a neural speech production thesis. While the ability to reliably reproduce speech based on intracranial signals is still far in the future, it could provide invaluable insight into the speech production network and its temporal dynamics.
References


