



Differential activation of human core, non-core and auditory-related cortex during speech categorization tasks as revealed by intracranial recordings

Mitchell Steinschneider^{1*}†, Kirill V. Nourski^{2†}, Ariane E. Rhone², Hiroto Kawasaki², Hiroyuki Oya² and Matthew A. Howard III²

¹ Departments of Neurology and Neuroscience, Albert Einstein College of Medicine, Bronx, NY, USA

² Human Brain Research Laboratory, Department of Neurosurgery, The University of Iowa, Iowa City, IA, USA

Edited by:

Einat Liebenthal, Medical College of Wisconsin, USA

Reviewed by:

Christopher I. Petkov, Newcastle University, UK

Dana Boatman-Reich, Johns Hopkins School of Medicine, USA

*Correspondence:

Mitchell Steinschneider, Department of Neurology, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY, 10461, USA
e-mail: mitchell.steinschneider@einstein.yu.edu

† These authors have contributed equally to this work

Speech perception requires that sounds be transformed into speech-related objects with lexical and semantic meaning. It is unclear at what level in the auditory pathways this transformation emerges. Primary auditory cortex has been implicated in both representation of acoustic sound attributes and sound objects. While non-primary auditory cortex located on the posterolateral superior temporal gyrus (PLST) is clearly involved in acoustic-to-phonetic pre-lexical representations, it is unclear what role this region plays in auditory object formation. Additional data support the importance of prefrontal cortex in the formation of auditory objects, while other data would implicate this region in auditory object selection. To help clarify the respective roles of auditory and auditory-related cortex in the formation and selection of auditory objects, we examined high gamma activity simultaneously recorded directly from Heschl's gyrus (HG), PLST and prefrontal cortex, while subjects performed auditory semantic detection tasks. Subjects were patients undergoing evaluation for treatment of medically intractable epilepsy. We found that activity in posteromedial HG and early activity on PLST was robust to sound stimuli regardless of their context, and minimally modulated by tasks. Later activity on PLST could be strongly modulated by semantic context, but not by behavioral performance. Activity within prefrontal cortex also was related to semantic context, and did co-vary with behavior. We propose that activity in posteromedial HG and early activity on PLST primarily reflect the representation of spectrotemporal sound attributes. Later activity on PLST represents a pre-lexical processing stage and is an intermediate step in the formation of word objects. Activity in prefrontal cortex appears directly involved in word object selection. The roles of other auditory and auditory-related cortical areas in the formation of word objects remain to be explored.

Keywords: electrocorticography, Heschl's gyrus, high gamma, prefrontal cortex, semantics, speech, superior temporal gyrus

INTRODUCTION

Speech perception requires that incoming sounds be transformed into word objects. It is unclear at what level in the auditory pathways this transformation occurs. Some data suggest that primary auditory cortex principally represents acoustic sound attributes (Mesgarani et al., 2008; Poeppel et al., 2008; Steinschneider et al., 2013). Other data suggest that primary auditory cortex is more directly involved in sound object representation (Nelken, 2008; Nelken and Bar-Yosef, 2008). It is also unclear what role non-primary auditory cortex, located on the posterolateral superior

temporal gyrus (PLST), plays in object formation. PLST is critical for acoustic-to-phonetic transformations (Boatman, 2004; Poeppel et al., 2008; Chang et al., 2010; Steinschneider et al., 2011; Mesgarani et al., 2014). This process could be interpreted as a remapping of the speech signal from one encoding acoustic attributes to one representing its phonemic components. By extension, it could be argued that this process remains a precursor to the formation of word objects. In this scheme, word object formation would be expected to take place at higher levels in auditory and auditory-related cortex (Griffiths and Warren, 2004; Griffiths et al., 2012).

Multiple studies have examined the transformation of neural activity associated with the representation of sound attributes to a representation of sound objects (Griffiths and Warren, 2004; Winkler et al., 2006; Shinn-Cunningham, 2008; Alain and Winkler, 2012; Griffiths et al., 2012; Simon, 2014). At the object

Abbreviations: ECoG, electrocorticography; ERBP, event-related band power; ERP, event-related potential; fMRI, functional magnetic resonance imaging; HG, Heschl's gyrus; IFG, inferior frontal gyrus; MFG, middle frontal gyrus; MRI, magnetic resonance imaging; MTG, middle frontal gyrus; PLST, posterolateral superior temporal gyrus; POA, place of articulation; STG, superior temporal gyrus; VOT, voice onset time.

formation processing stage, neural activity associated with a specific object must be distinct from that associated with other sound objects. Further, the neural representation of an object must be relatively invariant to variations in the detailed acoustics of the sounds. For instance, the representation of a specific word and its meaning must remain stable despite variations in acoustic characteristics that occur when a given word is spoken by different talkers. Given these requirements, object formation can be evaluated by utilizing tasks that require classifying words into semantic categories (Shahin et al., 2006; Hon et al., 2009).

Intracranial electrophysiological recordings in humans offer a unique opportunity for studying task-related activity in auditory cortex that accompanies semantic processing of speech. The technique combines exquisite spatial and temporal resolution beyond that offered by non-invasive methods such as neuro-magnetic responses and functional magnetic resonance imaging (MRI) (e.g., Lachaux et al., 2012). An excellent example of the sensitivity and specificity provided by intracranial recordings in humans is the study demonstrating that competing speech signals can be segregated according to speaker through analysis of cortical activity recorded from PLST during selective attention tasks (Mesgarani and Chang, 2012). The neural activity associated with the attended stream was enhanced, while activity associated with the unattended stream was suppressed. In a related study, target detection tasks led to enhanced neural activity to target tone stimuli on PLST when compared to responses obtained during passive listening and responses to non-target tone stimuli (Nourski et al., 2014a). These effects occurred during later portions of the neural responses. Early activity was minimally affected by the task requirement and appeared to represent the acoustic attributes of the tones. Similarly, minimal effects were noted in activity simultaneously recorded from posteromedial Heschl's gyrus (HG), the putative location of core auditory cortex. These findings suggest that activity generated within posteromedial HG and early activity from PLST reflect acoustic encoding rather than the representation of non-speech and speech-related objects at the phonemic level. It remains unclear from these studies, however, if this region of auditory cortex will also be involved in the formation of speech-related objects at the level of words and their semantic meaning.

The current study focused on high gamma responses (70–150 Hz) generated during target detection tasks using both speech and non-speech stimuli. High gamma activity has been shown to be a sensitive and specific indicator of auditory cortical activation and has been successfully used to define organizational features of human auditory cortex (e.g., Crone et al., 2001; Steinschneider et al., 2008, 2011; Flinker et al., 2010; Mesgarani and Chang, 2012; Mesgarani et al., 2014; Nourski et al., 2014b). Tasks of the current study included detecting words belonging to specific semantic categories or talker gender, as well as the detection of tones intermixed with the word sequences. Words were consonant-vowel-consonant exemplars from the semantic categories of animals, numbers and colors, as well as non-sense syllables, each spoken by different male and female talkers. Therefore, neural activity associated with target detection should not be based solely on acoustic attributes and instead should be related to semantic categorization and, consequently, word

object formation. We predicted that the tone detection task would not engage speech-related object formation, as this task only required differentiating the sound objects based on their acoustic attributes. In contrast, tasks that required the subject to detect words from a specific target category necessitated that words be decoded and categorized as sound objects belonging to specific semantic categories. Detection of talker gender provided an intermediate control condition. If the successful completion of the task was solely dependent upon decoding the fundamental frequencies typically encountered across gender (e.g., Hillenbrand et al., 1995), then, we hypothesized, sound object formation would not engage word-specific processing. If, however, formation of word objects incorporated representation of gender, then response profiles should be similar to that observed when words were categorized along semantic categories.

We also examined neural activity within auditory-related cortical areas that have been shown to be critical components of the neural network subserving speech perception (e.g., Rauschecker and Scott, 2009). Neural activity from inferior frontal gyrus (IFG) in the language-dominant hemisphere measured with intracranial recordings has been shown to represent lexical, grammatical and phonological aspects of speech (e.g., Sahin et al., 2009). In the present study, responses from the portion of IFG that overlaps with classically defined Broca's area were compared with activity recorded from HG and PLST. Additionally, contributions from middle temporal gyrus (MTG) and middle frontal gyrus (MFG) were examined, as these higher-order cortical regions may also be involved in word object formation (Griffiths and Warren, 2004; Poeppel et al., 2008). Simultaneous recordings from multiple regions including core, non-core and auditory-related cortex provided a unique opportunity to examine the role of each of these areas in word object formation during target detection tasks with high temporal and spatial detail.

METHODS

SUBJECTS

Experimental subjects were three neurosurgical patients diagnosed with medically refractory epilepsy and undergoing chronic invasive electrocorticographic (ECoG) monitoring to identify potentially resectable seizure foci. The subjects were 38 (L258), 30 (L275), and 40 (L282) years old. All subjects were male, right-handed and left hemisphere language-dominant, as determined by intracarotid amyltal (Wada) test results. Recordings were obtained from the left hemisphere in all three subjects. Research protocols were approved by the University of Iowa Institutional Review Board and the National Institutes of Health. Written informed consent was obtained from all subjects. Research participation did not interfere with acquisition of clinically required data, and subjects could rescind consent at any time without interrupting their clinical evaluation.

All subjects underwent audiometric evaluation before the study, and none was found to have hearing deficits that should impact the findings presented in this study. Subjects L258 and L282 were native English speakers, and subject L275 was a native Bosnian speaker who learned German at the age of 10 and English at the age of 17. Neuropsychological testing of L258 was normal except for mild deficiencies in verbal working memory. Subject

L275 had grossly intact conversational language comprehension, though formal neuropsychological testing showed non-localizing cognitive function deficits. Subject L282 had 13 years earlier undergone anterior temporal lobectomy that spared auditory cortex on the superior temporal gyrus. This subject was found to have mild deficits in verbal memory, fluency and naming. However, all three subjects had comparable performance in all experimental tasks both in terms of target detection accuracy and reaction times. This indicates that their performance of the tasks was not limited by any cognitive deficits identified during formal neuropsychological testing. Intracranial recordings revealed that auditory cortical areas were not epileptic foci in any subject.

Experiments were carried out in a dedicated electrically-shielded suite in The University of Iowa General Clinical Research Center. The room was quiet, with lights dimmed. Subjects were awake and reclining in an armchair.

STIMULI

Experimental stimuli were consonant-vowel-consonant syllables [cat], [dog], [five], [ten], [red], [white], [res], and [tem] from TIMIT (Garofolo et al., 1993) and LibriVox (<http://librivox.org/>) databases. Non-word syllables were excised from words using SoundForge 4.5 (Sonic Foundry Inc., Madison, WI). A total of 20 unique exemplars of each syllable were used in each experiment: 14 spoken by different male and 6 by different female speakers. Additionally, the stimulus set included complex tones with fundamental frequencies of 125 (28 trials) and 250 Hz (12 trials), approximating the average voice fundamental frequencies of male and female speakers, respectively. All stimuli were normalized to the same root-mean-square amplitude and edited to be 300 ms in duration using SoundForge with 5 ms rise-fall times. They were presented with an inter-stimulus interval chosen randomly within a Gaussian distribution (mean interval 2 s; $SD = 10$ ms) to reduce heterodyning in the recordings secondary to power line noise. Stimuli were delivered via insert earphones (ER4B, Etymotic Research, Elk Grove Village, IL) that were integrated into custom-fit earmolds. Stimulus delivery was controlled using Presentation software (Version 16.5 Neurobehavioral Systems, <http://www.neurobs.com/>).

The same stimuli were presented in random order in multiple target detection tasks. The target stimuli were either complex tones (presented as first block in each subject), speech stimuli spoken by female talkers, or words belonging to specific semantic categories such as animals or numbers. The subjects were instructed to use the index finger of their left hand (ipsilateral to the recording hemisphere) to push the response button whenever they heard a target sound. Prior to data collection, the subjects were presented with a random-sequence preview of stimuli to ensure that the sounds were presented at a comfortable level and that they understood the task requirements.

RECORDINGS

ECoG recordings were simultaneously made from HG and lateral cortical surface using multicontact depth and subdural grid electrodes, respectively. Details of electrode implantation have been described previously, and more comprehensive details regarding recording, extraction and analysis of high gamma cortical activity

are available for the interested reader (Howard et al., 1996, 2000; Reddy et al., 2010; Nourski et al., 2013; Nourski and Howard, 2014). In brief, hybrid depth electrode arrays were implanted stereotactically into HG, along its anterolateral to posteromedial axis. In subject L258, a hybrid depth electrode was used, which contained 4 cylindrical platinum macro-contacts, spaced 10 mm apart, and 14 platinum micro-contacts, distributed at 2–4 mm intervals between the macro contacts. In subjects L275 and L282, a depth electrode with 8 macro-contacts, spaced 5 mm apart, was used. Subdural grid arrays were implanted over the lateral surface of temporal and frontal lobes in subjects L258 and L275. The grid arrays consisted of platinum-iridium disc electrodes (2.3 mm exposed diameter, 5 mm center-to-center inter-electrode distance) embedded in a silicon membrane. The electrodes were arranged in an 8×12 grid, yielding a 3.5×5.5 cm array of 96 contacts. A subgaleal contact was used as a reference. Electrode arrays were placed solely on the basis of clinical requirements, and were part of a more extensive set of recording arrays meant to identify seizure foci. Electrodes remained in place under the direction of the patients' treating neurologists.

Subjects underwent whole-brain high-resolution T1-weighted structural MRI scans (resolution 0.78×0.78 mm, slice thickness 1.0 mm) before electrode implantation to locate recording contacts. Two volumes were averaged to improve the signal-to-noise ratio of the MRI data sets and minimize the effects of movement artifact on image quality. Pre-implantation MRIs and post-implantation thin-sliced volumetric computerized tomography scans (resolution 0.51×0.51 mm, slice thickness 1.0 mm) were co-registered using a linear co-registration algorithm with six degrees of freedom (Jenkinson et al., 2002). Locations of recording sites were confirmed by co-registration of pre- and post-implantation structural imaging and aided by intraoperative photographs.

Data acquisition was controlled by a TDT RZ2 real-time processor (Tucker-Davis Technologies, Alachua, FL). Collected ECoG data were amplified, filtered (0.7–800 Hz bandpass, 12 dB/octave rolloff), digitized at a sampling rate of 2034.5 Hz, and stored for subsequent offline analysis. Behavioral responses to the target stimuli were recorded using a Microsoft SideWinder game controller. The timing of the button-press events was recorded and stored for analysis along with ECoG data.

DATA ANALYSIS

ECoG data obtained from each recording site were downsampled to a rate of 1000 Hz. To minimize contamination with power line noise, ECoG waveforms were de-noised using an adaptive notch filtering procedure (Nourski et al., 2013). Prior to calculation of high gamma event-related band power (ERBP), individual trials were screened for possible contamination from electrical interference, epileptiform spikes, high amplitude slow wave activity, or movement artifacts. To that end, individual trial waveforms with voltage exceeding 2.5 standard deviations from the mean were rejected from further analysis. Data analysis was performed using custom software written in MATLAB Version 7.14 programming environment (MathWorks, Natick, MA, USA).

Quantitative analysis of the ERBP focused on the high gamma ECoG frequency band. High gamma ERBP was calculated for

each recording site. Single-trial ECoG waveforms were bandpass filtered between 70 and 150 Hz (100th order finite impulse response filter) and squared. The resultant high gamma power waveforms were smoothed using a moving average filter with a span of 25 ms, log-transformed, normalized to power in a pre-stimulus reference (250–50 ms prior to stimulus onset), and averaged across trials. To assess the presence and timing of task-related modulation of high gamma activity on representative cortical sites, single-trial high gamma ERBP was first averaged in 50 ms-wide consecutive windows to decrease the number of multiple comparisons. Next, for each window from 0–50 to 950–1000 ms, a two-sample one-tailed *t*-test was performed on single-trial windowed ERBP values to compare responses to stimuli presented in the non-target (tones task) and target condition. Finally, *p*-values were corrected for multiple comparisons (i.e., recording sites and time windows) using false discovery rate by controlling the false discovery rate following the method of Benjamini and Hochberg (1995) and Benjamini et al. (2001) with a threshold of $q = 0.01$.

RESULTS

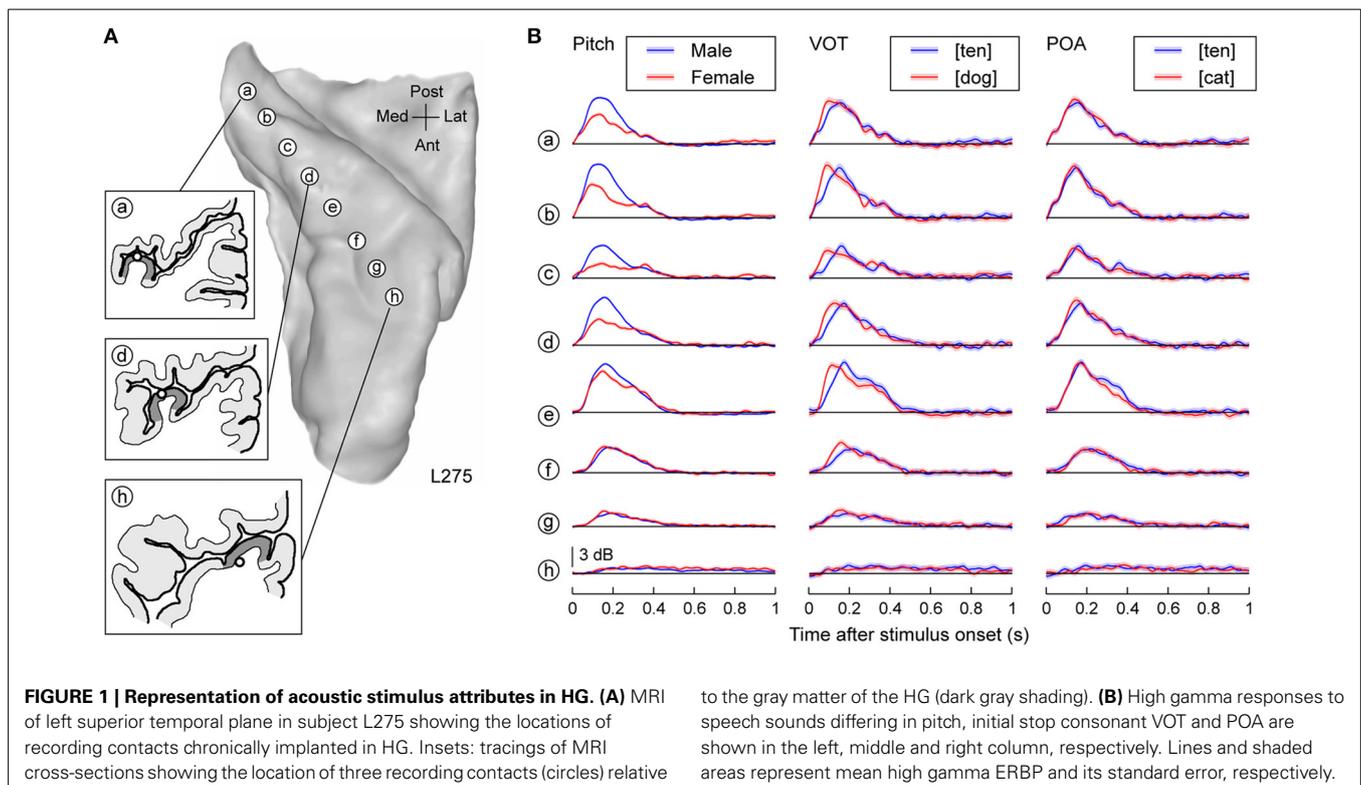
HG

Neural activity on HG primarily represented acoustic attributes of the speech stimuli (Figure 1). Figure 1A illustrates the location of the eight recording contacts that targeted HG along its long axis in subject L275. Mean high gamma power elicited by three acoustic attributes of speech is shown for each recording site (Figure 1B). Responses to the speech stimuli spoken by male talkers were consistently larger compared to those elicited

by female talkers (Figure 1B, left column), reflecting differences in their fundamental frequency (male talkers: mean 125 Hz, *SD* 25 Hz; female talkers: mean 202 Hz, *SD* 36 Hz). These differences represent a contribution in the high gamma responses of phase locking to the lower fundamental frequency of the male talkers within posteromedial HG [sites (a) through (d)] (cf. Nourski and Brugge, 2011; Steinschneider et al., 2013).

Voice onset time of the initial stop consonants was also differentially represented in the high gamma activity. In general, high gamma activity peaked earlier for initial consonants with short voice onset times (VOTs) (i.e., [dog]) relative to those with more prolonged VOTs (i.e., [ten]). This effect was maximal in more central portions of HG compared to the observed effect of pitch on neural activity [sites (e), (f); Figure 1B, middle column]. Differences based upon initial consonant place of articulation (POA) were more subtle, likely due to the overlap in spectral content across the stimulus exemplars (e.g., site (d); Figure 1B, right column). These patterns of activity within HG were also observed in the other two subjects (Supplementary Figures 1, 2).

Whereas activity along most of HG was strongly modulated by the acoustic attributes of the sounds, responses in the high gamma range were only weakly affected by the target detection tasks (Figure 2). The left column in Figure 2 compares neural activity to the same set of stimuli (female voices) in three blocks: when they were targets, when they were non-targets in the tone detection block, or when they were non-targets in a semantic task (numbers). A low-amplitude increase in high gamma was seen beginning within 600–650 ms after stimulus onset when female voices were the targets [site (a)], overlapping in time



with the subject's behavioral response. A similar effect was seen for responses to the animals and numbers when they were the targets. However, the onset of the task-related high gamma modulation in these semantic categorization conditions was even slower than that occurring during voice identification task ($q < 0.01$ at 750–800 ms after stimulus onset; middle and right columns of **Figure 2**).

A different pattern was observed within the most anterolateral portion of HG outside of presumed core cortex [site (h) in **Figure 2**]. Here the response was delayed relative to the activity on posteromedial HG and was specifically associated with target stimuli. Importantly, this task-related activity preceded task-related changes that were observed on posteromedial HG. These task-related increases, however, were variable across subjects. In the other two subjects, no significant task-related effects were observed at the level of either posteromedial or anterolateral HG (Supplementary Figures 3, 4). Thus, in total, task-related changes in HG were, as we will show, modest, when compared to those changes observed on PLST and in auditory-related cortex.

PLST

More complex response profiles were observed on PLST (**Figures 3, 4**) when compared with profiles simultaneously recorded from HG (see **Figure 2**). There was a rapid and large increase in high gamma ERBP occurring within 200 ms after stimulus onset. This early activity was variably affected by the task [e.g. sites (a), (b), and (c) in **Figures 3, 4**]. When female voices were targets, a modest but significant increase in high gamma power was observed as early as 50–100 ms after stimulus onset. Peak activity at 150–200 ms was only marginally affected by the task. Later activity was more variable across recording sites. Both enhancement of high gamma activity to the target syllables beginning prior to their offsets [e.g. sites (a), (b), and (c) in **Figure 3**] and minimal modulation of later activity related to the task (see **Figure 3B**) were observed in this region. Task-related high gamma activity was earlier than that occurring in HG (cf. **Figure 2**) and preceded the subject's behavioral response.

Responses to non-target words were also modulated by the specific task requirements. For instance, late high gamma activity to non-target words spoken by females was enhanced when the target detection tasks required words to be categorized relative to the task where complex tones were the targets (see **Figure 3**, green and blue plots, respectively). Responses to female voices when they were target stimuli were consistently larger than when they were non-targets, even though the subject was engaged in cognitively more demanding tasks (detecting numbers or animals) (see **Figure 3**, red and green plots, respectively). The difference in task difficulty can be inferred from behavioral response times, which were significantly shorter when the target was female voices (median response time 672 ms) relative to either task requiring semantic classification (animals: median response time 866 ms; numbers: median response time 815 ms) ($p < 0.001$, Mann-Whitney rank-sum tests).

Enhancement of high gamma power was also observed when the targets were animals (**Figure 4**). Once again, targets elicited the largest responses when compared to when they were non-targets presented in a tone detection task [see **Figure 4**, sites

(a) and (c)]. While variable across sites [cf. site (b) in **Figure 4**], enhanced activity could occur early and remain elevated even during the time period of the behavioral response. Responses to non-target animal words presented in a different semantic categorization task (detecting numbers) were intermediate in magnitude. The behavioral reaction times were comparable in the animals and numbers detection tasks ($p = 0.71$, Mann-Whitney rank-sum test). Therefore, it is reasonable to conclude that these differences between target and relevant non-target were not based solely on task difficulty. Importantly, increases in high gamma activity observed during either semantic categorization task began prior to the offset of the syllables, suggesting that these increases were not directly related to word classification, and likely were reflecting lower-level phonological processing, a prerequisite for semantic classification (cf. Boatman, 2004).

In subject L258, task-related enhancement was not observed from sites located on PLST [Supplementary Figure 5; sites (a), (b), and (c)]. This negative finding may reflect in part differences in placement of the electrode grids, where the anterior limit of the temporal recording grid was anatomically more posterior than that in subject L275. Additionally, responses from L275 were averaged over a larger number of trials, improving signal-to-noise ratio, and subject L275 was generally more enthusiastic about performing the behavioral tasks compared to L258. However, responses were modulated by the task on sites overlying the MTG [e.g. sites (d) and (e); Supplementary Figure 5], similar to that seen on PLST in subject L275. Specifically, late responses to target stimuli were larger than responses in the tone detection task, reaching significance on site (e) in the gender identification task ($q < 0.01$), and were marginally significant ($q < 0.05$) in the semantic categorization tasks on sites (d) and (e) (significance bars are not shown). Additionally, there was a trend for non-target words to elicit larger late responses during semantic categorization tasks compared to tone detection (green and blue plots, respectively, in Supplementary Figure 5).

AUDITORY-RELATED CORTEX: IFG AND MFG

Task-related changes in high gamma activity were not restricted to the temporal lobe and were observed in IFG and MFG in both subjects with frontal lobe electrode coverage (**Figure 5**, Supplementary Figure 5). Targets elicited larger responses compared to when the same words were presented in a tone detection task in both IFG and MFG (purple bars in **Figure 5**). Minimal activity in both regions was observed in response to non-target speech stimuli when tones were targets, and phonemic and semantic processing were not necessary for task performance. In contrast, both targets and non-targets relevant to the task elicited responses in IFG in both subjects and MFG in subject L258 (red and green plots). Responses within MFG in subject L275 were restricted to target stimuli and had onset latencies longer than those observed at sites overlying either the superior temporal gyrus (STG), MTG or IFG, but were comparable to the timing of the late high gamma increases seen on posteromedial HG. These late increases in high gamma activity always preceded the subjects' behavioral responses (horizontal box plots in **Figure 5** and Supplementary Figure 5), which elicited high gamma activity within both pre- and post-central gyrus (see **Figures 3, 4**).

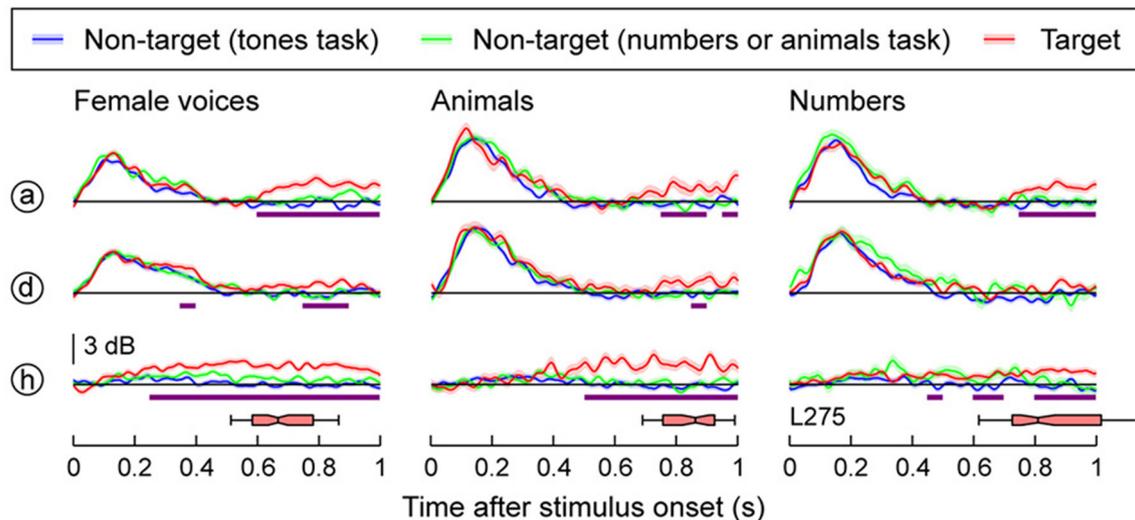


FIGURE 2 | Task effects on responses to speech stimuli in HG. Responses to three types of stimuli (female voices, animals, numbers; left, middle and right column, respectively) are shown for three representative recording sites in HG (rows). See **Figure 1A** for location of the recording sites. Colors (blue, green, and red) represent different task conditions. Lines and shaded areas

represent mean high gamma ERBP and its standard error, respectively. Purple bars denote time windows where responses to the target stimuli were significantly larger than those to the same stimuli in the tones task ($q < 0.01$). Horizontal box plots denote the timing of behavioral responses to the target stimuli (medians, 10th, 25th, 75th, and 90th percentiles).

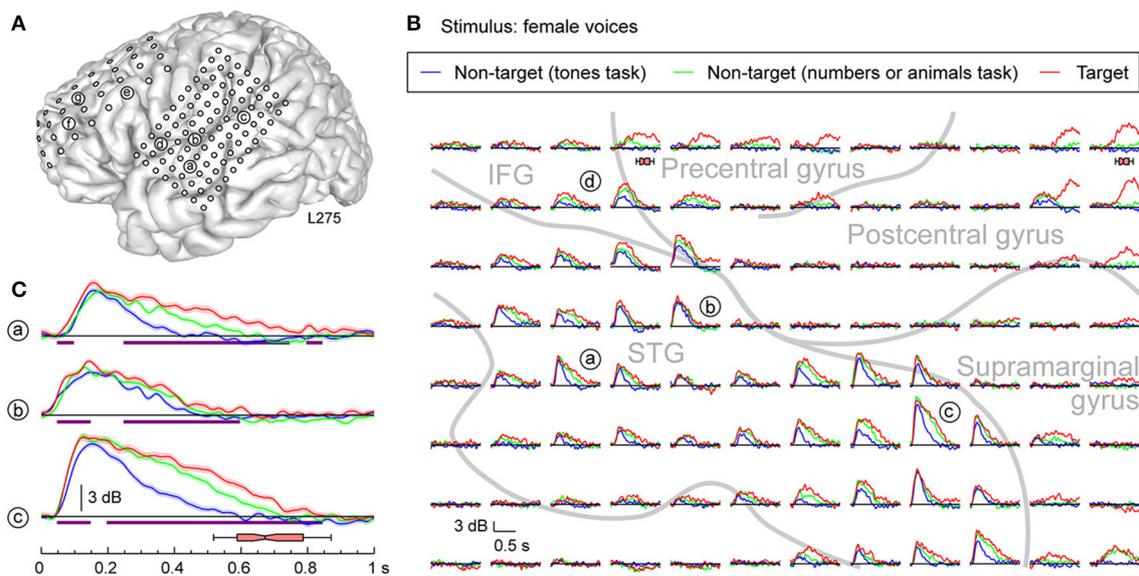


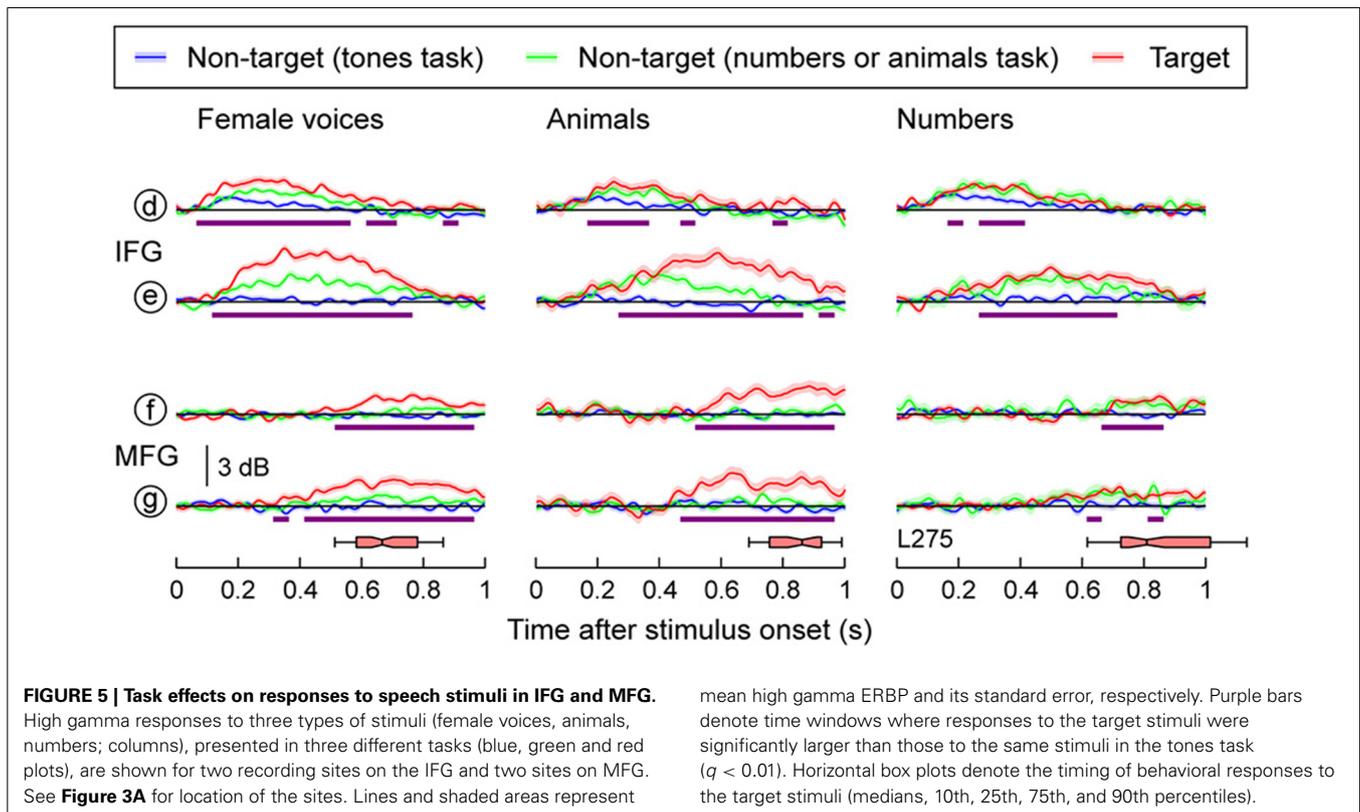
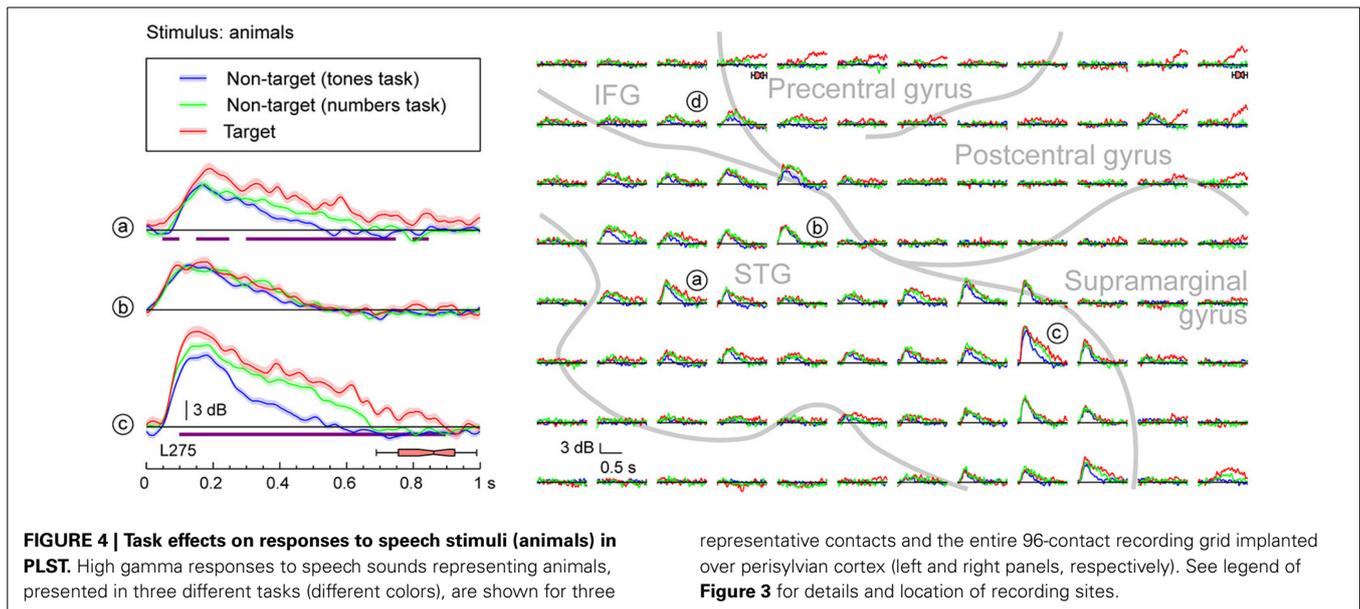
FIGURE 3 | Task effects on responses to speech stimuli (female voices) in PLST. (A) MRI of the left hemisphere in subject L275 showing the locations of chronically implanted subdural grid contacts. **(B)** High gamma responses to syllables spoken by females, presented in three different tasks (different colors), are shown for the 96-contact recording grid implanted over perisylvian cortex. Gray lines represent approximate boundaries of STG, IFG, pre- and post-central gyri covered by the recording grid. **(C)** High gamma

ERBP time course replotted for three recording sites on PLST. Lines and shaded areas represent mean high gamma ERBP and its standard error, respectively. Purple bars denote time windows where responses to the target stimuli were significantly larger than those to the same stimuli in the tones task ($q < 0.01$). Horizontal box plot denotes the timing of behavioral responses to the target stimulus (median, 10th, 25th, 75th, and 90th percentiles).

DIFFERENTIAL RESPONSE PATTERNS TO TARGET STIMULI: PLST vs. IFG

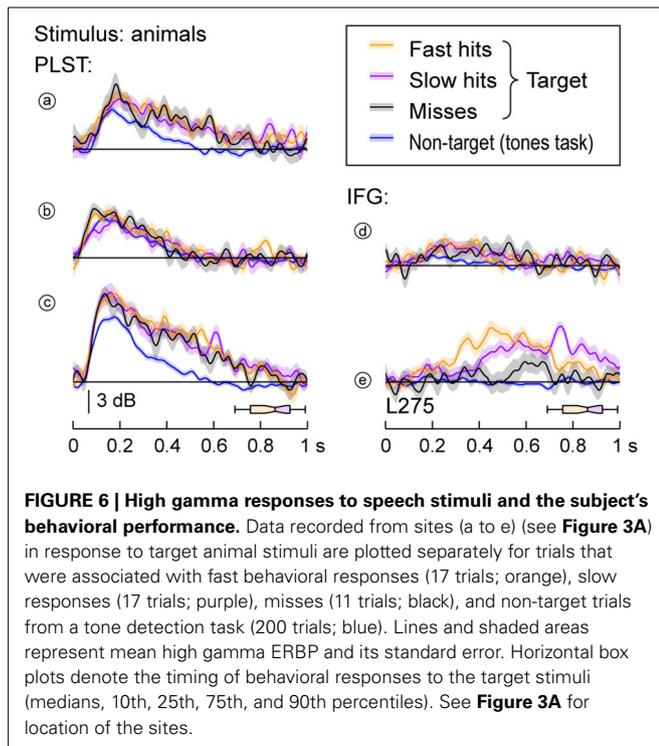
Different response patterns elicited by target stimuli were noted between activity simultaneously recorded from PLST and IFG in subject L275. High gamma activity on PLST elicited by

target stimuli (animals) did not significantly vary as a function of whether the subject responded rapidly or slowly or when the target was missed altogether (**Figure 6**, left column). In comparison, the same words when they were not relevant



non-targets (tone detection task) elicited comparable early activity, but markedly diminished responses later in time [sites (a) and (c) in **Figure 6**]. In contrast to activity on PLST, activity within pars opercularis of IFG could be significantly modulated by the presence and timing of the behavioral response. This finding is exemplified at site (e) located on the dorsal portion of the pars opercularis (see **Figure 6**), where faster response

times were associated with earlier peaks of activity when contrasted with slower behavioral responses. Additionally, misses were associated with markedly decreased responses compared to hits, and there was no response when the same stimulus was presented as a non-relevant, non-target during a tone detection task. For subject L258, parcelation of single-trial high gamma activity based on behavioral performance did not reveal consistent



differences between PLST and IFG. This was due to highly variable responses and low response magnitudes, particularly in IFG.

DISCUSSION

POSTEROMEDIAL HG

As expected from previous studies, activity within posteromedial HG was highly sensitive to the acoustic characteristics of speech (e.g., Nourski et al., 2009; Steinschneider et al., 2013). In general, high gamma activity was greater for male talkers than female talkers. This finding reflects contribution from phase-locked responses to the lower fundamental frequency of male talkers relative to female talkers and was most prominently observed in the most posteromedial aspect of HG. This property is not unique to speech, as this region exhibits reliable phase-locked responses elicited by click trains at repetition rates of up to 200 Hz (Brugge et al., 2008, 2009; Nourski and Brugge, 2011). VOT was reflected in the timing of high gamma activity as a delay in the peak of high gamma response. This effect was most prominent in more central areas of HG, contrasting with the temporal representation of the voice fundamental. This apparent spatial differentiation may be a consequence of the tonotopic organization, wherein phase locking would most likely occur in high best frequency areas of the HG, whereas VOT would be represented in low frequency regions, due to the later onset of low frequency energy associated with voicing onset (Steinschneider et al., 1994). The absence of single and double-on responses previously reported (e.g., Steinschneider et al., 2013) can be attributed to the temporal smearing inherent to averaging of responses to unique and naturally-elicited speech exemplars characterized by different VOTs. Finally, responses reflecting differences in stop

consonant POA were more subtle, and were likely a result of spectral smearing due to averaging of responses to 20 different exemplars of [cat] and [ten] across multiple talkers and the location of the recording sites with reference to the tonotopic organization of HG.

Activity within posteromedial and central HG was not strongly modulated by task requirements in all three subjects, and if it occurred (e.g., L275), it was later than task-related modulations in all other regions studied. Thus, current findings do not support the premise that human primary auditory cortex is the location where auditory object formation occurs. In contrast, studies in primary auditory cortex of experimental animals have shown robust responses reflecting auditory object formation, task-related activity, and reward expectancy (e.g., Fritz et al., 2003; Nelken and Bar-Yosef, 2008; Brosch et al., 2011; Niwa et al., 2012). The difference between the current observations and those in animals may reflect species differences and the relative complexity of auditory cortical organization in humans (Hackett, 2007). This complexity would be paralleled by greater functional specialization for primary and non-primary areas as the demands for vocal learning and auditory sequence learning become progressively more complex (Petkov and Jarvis, 2012).

Our findings in HG are consistent with several magnetoencephalography and event-related potential (ERP) studies (Shahin et al., 2006; Gutschalk et al., 2008; Sabri et al., 2013; Simon, 2014; but see Bidet-Caulet et al., 2007). One study observed that during selective attention to one speech stream over another, the M100, but not M50 component of the neuromagnetic response, was modulated by the attended stream (Simon, 2014). This finding is consistent with our negative results, as the M50 component is dominated by generators in or near primary auditory cortex, while the M100 component reflects generators from multiple non-primary areas, particularly those in planum temporale (Liégeois-Chauvel et al., 1994). Another study sorted magnetoencephalography data according to whether or not target tones were detected in a multi-tone cloud background capable of producing informational masking of the targets (Gutschalk et al., 2008). Detected targets elicited an M100-like component that was not present when the target sounds were not detected. In contrast, both detected and undetected tones evoked auditory middle-latency and steady-state responses whose generators likely include prominent contributions from the primary auditory cortex on HG. It should be noted, however, that other studies utilizing auditory detection paradigms failed to find modulation of the N100 component (Shahin et al., 2006; Sabri et al., 2013). This negative result is not restricted to the auditory modality and has been observed in early cortical activity during visual target detection tasks (Bansal et al., 2014).

The minimal modulation of early high gamma activity that we observed replicates the findings in a previous intracranial study, where no effect was observed in the magnitude or timing of high gamma activity within posteromedial HG during a tone detection task relative to passive listening (Nourski et al., 2014a). Finally, functional neuroimaging studies have not shown consistent task-related changes in HG (Pugh et al., 1996; Leicht et al., 2010). When present, attention-related modulations occurred mainly in non-primary auditory cortex lateral to core areas (Petkov et al.,

2004). This latter finding is consistent with task-related modulations currently seen in the most anterolateral portion of HG in one subject (see **Figure 2**). It must be acknowledged, however, that limited sampling inherent to human HG recordings may be responsible for the lack of consistent task-related effects seen in the three subjects studied here.

PLST

Early activity on PLST, occurring within 200 ms after stimulus onset, was not strongly modulated by task requirements, mirroring a result seen in different subjects performing a tone detection task (Nourski et al., 2014a). Studies have demonstrated that early high gamma activity reflects more automatic processing that helps represent specific spectral characteristics of tone stimuli (Nourski et al., 2014a,b), as well as the remapping of acoustic speech characteristics to those representing phonetic categories (Chang et al., 2010; Travis et al., 2013; Mesgarani et al., 2014). In contrast, later high gamma activity on PLST could be strongly modulated by task requirements. Findings such as these are neither unique to humans nor restricted to the auditory system. For instance, during visual object detection tasks, single unit activity from neurons within areas V4 and IT of the monkey showed limited modulation as a function of the target stimulus in the initial response component, yet were strongly dependent on the specific target in later response segments (Chelazzi et al., 1998, 2001). The authors suggested that these later effects were based on feedback from higher visual centers involved in working memory, and reflected response bias toward the behaviorally relevant objects. A similar “top-down” mechanism that biases responses toward task-relevant stimuli may also be responsible for the currently observed effects in PLST.

Several studies have shown that neural patterns of activity in auditory cortex independently encode speaker identity and phonemic content of verbal speech (“Who” is saying “what”; e.g., Formisano et al., 2008; Mesgarani and Chang, 2012). We examined whether similar patterns independently encoding voice vs. speech content would emerge during the performance of the current target detection tasks, but found no clear differences. It should be noted, however, that in the study of Formisano et al. (2008), subjects passively listened to only three vowels spoken by three talkers. Here, subjects actively listened to 180 unique word exemplars spoken by an almost equal number of different talkers presented during semantic classification tasks and control conditions that included gender identification. Furthermore, the brain regions associated with gender identification were primarily located over the non-dominant right hemisphere and distributed on the lateral portion of HG and Heschl’s sulcus, as well as portions of the superior temporal sulcus (Formisano et al., 2008). The current study examined the dominant left hemisphere with limited sampling of HG, and did not sample neural activity in Heschl’s sulcus or the superior temporal sulcus. In the study by Mesgarani and Chang (2012), the subjects were performing a different behavioral task (selective attention), and the neural activity only had to be capable of discriminating sentences spoken by two talkers (one male). It thus remains to be determined whether high gamma power, at least within PLST, is capable of independently determining multiple speaker identities

(or gender) and phonemic content (e.g., Obleser and Eisner, 2009).

Response enhancement on PLST began prior to word offset during the semantic classification tasks (see **Figure 6**). The timing of response enhancement indicates that the effect was not driven by processes directly reflecting semantic classification, but instead represented the phonemic processing that must by necessity occur earlier in order to accurately decode the words. Further, the target words elicited a larger response than non-target words. As pointed out by Hon et al. (2009), any target enhancement that occurs within early sensory regions when a semantic target is detected must originate from higher-level brain areas providing relevant feedback to the lower areas. In the present study, subjects had been primed to know that the same two exemplar words for each semantic category would be presented in each successive recording block. This priming would allow subjects to know that, for instance, in the animals task, /d/ and /k/ would be the first phonemes in the target words ([dog] and [cat]) and thus provide additional information useful for the completion of the semantic task.

Response enhancement on PLST was also independent of task performance accuracy and reaction time. The same effect has been observed on PLST in a different subject performing a tone detection task, thus replicating current findings (Nourski et al., 2014a). Object-based detection tasks require two sequential processes, object formation followed by object selection (Shinn-Cunningham, 2008). The independence of the neural responses from behavioral measures are consistent with PLST being involved in the process of semantic object formation, yet not directly tied to the process of object selection. Similar observations have been made in the lateral belt field AL in macaque auditory cortex when performing a discrimination task using consonant-vowel syllables (Tsunada et al., 2011). In that study, single-cell responses reflected the categorization of the syllable (i.e., object formation), but did not vary as a function of the animal’s behavioral performance (i.e., object selection). Activity that does not vary with behavioral performance likely reflects processes that precede sound object formation.

Even in the subject where later activity was strongly modulated (L275), effects were not uniform and showed site-by-site variability. This variability may partly explain why task-related modulation on PLST was not seen in subject L258 (see Supplementary Figure 5). Additionally, electrode array placement was more posterior along the STG in subject L258 when compared to the placement in L275. Electrical stimulation in subjects with epilepsy while they participated in various auditory and speech-related tasks has demonstrated the functional heterogeneity of the STG (Boatman, 2004), indicating that differences in electrode placement can be a major source of inter-subject variability. Finally, language processing skills of the subjects and effort necessary for successful performance of the task, may have also been a significant factor contributing to the inter-subject variability observed in this study.

AUDITORY-RELATED CORTEX

Multiple brain regions outside of the classically defined auditory cortex were differentially activated during the target detection

tasks. For instance, task-related activity was shown within MTG, and enhancement of later activity was observed in responses to targets and non-targets in the semantic categorization tasks. Similar activation of MTG immediately adjacent to the superior temporal sulcus in response to speech has been reported (**Figure 3** in Flinker et al., 2010). This region has been shown to be important in lexical processing, and is activated even during passive presentation of words (Dronkers et al., 2004; Indefrey and Cutler, 2004; Hagoort, 2005; Hickok, 2009). Unfortunately, sampling was too limited to better describe these modulations outside of observing that they had latencies comparable with those seen on PLST and frontal regions.

The IFG of the dominant left hemisphere was also activated during target detection tasks. High gamma activity was observed when stimuli were targets, and, to a lesser degree, non-targets. Findings are in keeping with other auditory target detection studies. Bilateral activation of the IFG occurred during an auditory detection task using positron emission tomography when targets were words, consonant-vowel syllables, or tone triplets (Fiez et al., 1995). Activation of the left IFG was observed in a study by Shahin et al. (2006) that combined functional MRI (fMRI) and ERP and used two target detection paradigms similar to that used in the current study: (1) a semantic task of detecting infrequent word targets denoting animals in a stream of words denoting non-animate objects; (2) a voice gender task detecting infrequent tokens spoken by males in a stream of words spoken by females. Results from fMRI were used to constrain possible anatomical source generators of the ERP. Activation of the IFG in the left hemisphere was seen in the semantic task performed with fMRI, and was associated with negative ERP components to both target and non-target words. Further, responses to targets were larger than responses to non-targets. Peak latencies of these negative ERP components were 450 and 600 ms, respectively, and overlap in time with the high gamma activity observed in the IFG in the present study. These results obtained from neurologically-normal subjects are all concordant with current results, despite the fact that all of our subjects were epileptic patients, and one subject (L275) was trilingual and had non-localizing cognitive deficiencies.

An important distinction between the responses located on PLST and IFG is that activity within pars opercularis of the IFG could vary as a function of behavioral performance (see **Figure 6**). Activity recorded during correctly identified targets was larger than when the target was missed. Further, activity during trials with shorter reaction times peaked earlier than activity during trials when reaction times were longer. This relationship with behavioral performance mirrors that seen in ventrolateral prefrontal cortex of macaques performing a phonemic discrimination task (Russ et al., 2008), and, as discussed above, contrasts with neural activity observed in field AL (Tsunada et al., 2011). The transformation in response characteristics from temporal to frontal lobe is parsimonious with the view that PLST is involved in the process of word object formation, while IFG is involved in the process of word object selection (Shinn-Cunningham, 2008).

MFG appears to also be involved in object selection, as it too responded only to targets (see **Figure 5**) and relevant non-targets

during semantic categorization tasks (see Supplementary Figure 5). This activity began later than that in STG and IFG, yet preceded behavioral responses. Activation of the left MFG during a semantic target detection task has been reported using fMRI (Shahin et al., 2006). Variability in responses to targets and relevant non-targets has also been shown in detection tasks using visual stimuli (Kirino et al., 2000; Kiehl et al., 2001; Bledowski et al., 2004; Hampshire et al., 2007; Hon et al., 2012). To varying degrees, MFG as well as IFG were shown to respond either selectively to visual targets or to both targets and relevant non-targets. Additional work will be required to determine the sources of variability that characterized responses during the semantic classification tasks in IFG and MFG.

Strong task-related modulation of high gamma power outside classically defined auditory cortex is consistent with that seen in both the auditory and visual modalities in human ERP and fMRI studies (Sabri et al., 2013; Bansal et al., 2014). In the one study that compared responses to detected vs. undetected sound targets (Sabri et al., 2013), greater activation (as revealed by fMRI) was noted in the parietal lobe, thalamus and basal ganglia. While these regions were not examined in the current study, present results indicate that activity within IFG and MFG (as revealed by high gamma ERBP) is also related to the behavioral outcomes of the task, including the presence of the behavioral response and its timing.

CONCLUDING REMARKS

The response patterns described here reflect multiple processing stages of word object formation that constitute lexical encoding. At a neuroanatomical level, it does not appear that object formation occurs in posteromedial HG. Responses within this region are dominated by representation of the acoustic attributes of speech, and are therefore prelexical. Activity on PLST is also prelexical, but, in contrast to posteromedial HG, can also be strongly modulated by higher-order areas subserving lexical and semantic processing. The modulation on PLST during semantic classification tasks indicates that this region represents an early stage in word object formation.

It should be acknowledged that the subjects that participated in this study are patients who have neurologic deficits, including those in the language domain, and who have been treated with multiple anticonvulsant drugs over long periods of time. This calls into question as to whether findings in this population can be generalized to subjects without neurologic deficits. Despite this limitation, intracranial investigations of neurosurgical patients have been highly fruitful in defining organizational features of auditory and auditory-related cortex (e.g., Crone et al., 2001; Sahin et al., 2009; Chang et al., 2010; Mesgarani and Chang, 2012; Mesgarani et al., 2014). Findings described in the present report confirm and extend our own previous intracranial results demonstrating that PLST exhibits task-related modulation of high gamma activity regardless of behavioral outcome (Nourski et al., 2014a). Finally, results are congruent with non-invasive human studies (e.g., Pugh et al., 1996; Shahin et al., 2006; Gutschalk et al., 2008; Obleser and Eisner, 2009; Leaver and Rauschecker, 2010; Leicht et al., 2010; Simon, 2014) and relevant investigations using experimental animals (e.g., Russ et al.,

2008; Brosch et al., 2011; Tsunada et al., 2011; David et al., 2012; Steinschneider et al., 2013; Sutter et al., 2013).

Future intracranial studies must corroborate current observations and extend them by examining task-related activity in other brain regions known to be important for sound processing. Specifically, investigation of response profiles in anterolateral HG, planum temporale, anterior STG, superior temporal sulcus and MTG will help identify additional stages of word object formation. Similarly, additional work will be needed to further characterize the roles of IFG and MFG in both dominant and non-dominant hemispheres in word object selection. Finally, future studies should include investigation of dynamic interactions between cortical regions, including feedback from higher-order cortices onto sensory areas. This will likely require examination of long-range phase coherence at multiple frequency bands (e.g., theta-gamma) that are likely important in long-range interactions between spatially disparate regions. As we continue investigation of these circuits, our conclusions will undoubtedly be refined and, hopefully, translationally relevant for the understanding of normal speech processing and its dysfunction occurring in developmental language disorders, and acquired disorders such as stroke and normal aging.

ACKNOWLEDGMENTS

We thank Haiming Chen, Rachel Gold and Christopher Kovach for help with data collection and analysis. This study was supported by NIH R01-DC04290, UL1RR024979, Hearing Health Foundation and the Hoover Fund.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fnins.2014.00240/abstract>

REFERENCES

- Alain, C., and Winkler, I. (2012). "Recording event-related brain potentials: application to study auditory perception," in *The Human Auditory Cortex*, eds D. Poeppel, T. Overath, A. N. Popper, and R. R. Fay (New York, NY: Springer Science+Business Media, LLC), 199–224.
- Bansal, A. K., Madhavan, R., Agam, Y., Golby, A., Madsen, J. R., and Kreiman, G. (2014). Neural dynamics underlying target detection in the human brain. *J. Neurosci.* 34, 3042–3055. doi: 10.1523/JNEUROSCI.3781-13.2014
- Benjamini, Y., Drai, D., Elmer, G., Kafkari, N., and Golani, I. (2001). Controlling the false discovery rate in behavior genetics research. *Behav. Brain Res.* 125, 279–284. doi: 10.1016/S0166-4328(01)00297-2
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* 57, 289–300.
- Bidet-Caulet, A., Fischer, C., Besle, J., Aguera, P. E., Giard, M. H., and Bertrand, O. (2007). Effects of selective attention on the electrophysiological representation of concurrent sounds in the human auditory cortex. *J. Neurosci.* 27, 9252–9261. doi: 10.1523/JNEUROSCI.1402-07.2007
- Bledowski, C., Prvulovic, D., Hoehstetter, K., Scherg, M., Wibral, M., Goebel, R., et al. (2004). Localizing P300 generators in visual target and distractor processing: a combined event-related potential and functional magnetic resonance imaging study. *J. Neurosci.* 24, 9353–9360. doi: 10.1523/JNEUROSCI.1897-04.2004
- Boatman, D. (2004). Cortical bases of speech perception: evidence from functional lesion studies. *Cognition* 92, 47–65. doi: 10.1016/j.cognition.2003.09.010
- Brosch, M., Selezneva, E., and Scheich, H. (2011). Representation of reward feedback in primate auditory cortex. *Front. Syst. Neurosci.* 5:5. doi: 10.3389/fnsys.2011.00005
- Brugge, J. F., Nourski, K. V., Oya, H., Reale, R. A., Kawasaki, H., Steinschneider, M., et al. (2009). Coding of repetitive transients by auditory cortex on Heschl's gyrus. *J. Neurophysiol.* 102, 2358–2374. doi: 10.1152/jn.91346.2008
- Brugge, J. F., Volkov, I. O., Oya, H., Kawasaki, H., Reale, R. A., Fenoy, A., et al. (2008). Functional localization of auditory cortical fields of human: click-train stimulation. *Hear. Res.* 238, 12–24. doi: 10.1016/j.heares.2007.11.012
- Chang, E. F., Rieger, J. W., Johnson, K., Berger, M. S., Barbaro, N. M., and Knight, R. T. (2010). Categorical speech representation in human superior temporal gyrus. *Nat. Neurosci.* 13, 1428–1432. doi: 10.1038/nn.2641
- Chelazzi, L., Duncan, J., Miller, E. K., and Desimone, R. (1998). Responses of neurons in inferior temporal cortex during memory-guided visual search. *J. Neurophysiol.* 80, 2918–2940.
- Chelazzi, L., Miller, E. K., Duncan, J., and Desimone, R. (2001). Responses of neurons in macaque area V4 during memory-guided visual search. *Cereb. Cortex* 11, 761–772. doi: 10.1093/cercor/11.8.761
- Crone, N. E., Boatman, D., Gordon, B., and Hao, L. (2001). Induced electrocorticographic gamma activity during auditory perception. *Clin. Neurophysiol.* 112, 565–582. doi: 10.1016/S1388-2457(00)00545-9
- David, S. V., Fritz, J. B., and Shamma, S. A. (2012). Task reward structure shapes rapid receptive field plasticity in auditory cortex. *Proc. Natl. Acad. Sci. U.S.A.* 109, 2144–2149. doi: 10.1073/pnas.1117717109
- Dronkers, N. F., Wilkins, D. P., Van Valin, R. D. Jr., Redfern, B. B., and Jaeger, J. J. (2004). Lesion analysis of the brain areas involved in language comprehension. *Cognition* 92, 145–177. doi: 10.1016/j.cognition.2003.11.002
- Fiez, J. A., Raichle, M. E., Miezin, F. M., Petersen, S. E., Tallal, P., and Katz, W. F. (1995). PET studies of auditory and phonological processing: effects of stimulus characteristics and task demands. *J. Cogn. Neurosci.* 7, 357–375. doi: 10.1162/jocn.1995.7.3.357
- Flinker, A., Chang, E. F., Barbaro, N. M., Berger, M. S., and Knight, R. T. (2010). Sub-centimeter language organization in the human temporal lobe. *Brain Lang.* 117, 103–109. doi: 10.1016/j.bandl.2010.09.009
- Formisano, E., De Martino, F., Bonte, M., and Goebel, R. (2008). "Who" is saying "what?" Brain-based decoding of human voice and speech. *Science* 322, 970–973. doi: 10.1126/science.1164318
- Fritz, J., Shamma, S., Elhilali, M., and Klein, D. (2003). Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nat. Neurosci.* 6, 1216–1223. doi: 10.1038/nn1141
- Garofolo, J. S., Lamel, L. F., Fisher, W. M., Fiscus, J. G., Pallett, D. S., Dahlgren, N. L., et al. (1993). *TIMIT Acoustic-Phonetic Continuous Speech Corpus*. Philadelphia, PA: Linguistic Data Consortium.
- Griffiths, T. D., Micheyl, C., and Overath, T. (2012). "Auditory object analysis," in *The Human Auditory Cortex*, eds D. Poeppel, T. Overath, A.N. Popper, and R.R. Fay (New York, NY: Springer Science+Business Media, LLC), 199–224.
- Griffiths, T. D., and Warren, J. D. (2004). What is an auditory object? *Nat. Rev. Neurosci.* 5, 887–892. doi: 10.1038/nrn1538
- Gutschalk, A., Micheyl, C., and Oxenham, A. J. (2008). Neural correlates of auditory perceptual awareness under informational masking. *PLoS Biol.* 6:e138. doi: 10.1371/journal.pbio.0060138
- Hackett, T. A. (2007). "Organization and correspondence of the auditory cortex of humans and nonhuman primates," in *Evolution of Nervous Systems Vol. 4: Primates*, ed J. H. Kaas (New York, NY: Academic Press), 109–119.
- Hagoort, P. (2005). On Broca, brain, and binding: a new framework. *Trends Cogn. Sci.* 9, 416–423. doi: 10.1016/j.tics.2005.07.004
- Hampshire, A., Duncan, J., and Owen, A. M. (2007). Selective tuning of the blood oxygenation level-dependent response during simple target detection dissociates human frontoparietal subregions. *J. Neurosci.* 27, 6219–6223. doi: 10.1523/JNEUROSCI.0851-07.2007
- Hickok, G. (2009). The functional neuroanatomy of language. *Phys. Life Rev.* 6, 121–143. doi: 10.1016/j.plrev.2009.06.001
- Hillenbrand, J., Getty, L. A., Clark, M. J., and Wheeler, K. (1995). Acoustic characteristics of American English vowels. *J. Acoust. Soc. Am.* 97, 3099–3111. doi: 10.1121/1.411872
- Hon, N., Ong, J., Tan, R., and Yang, T. H. (2012). Different types of target probability have different prefrontal consequences. *Neuroimage* 59, 655–662. doi: 10.1016/j.neuroimage.2011.06.093

- Hon, N., Thompson, R., Sigala, N., and Duncan, J. (2009). Evidence for long-range feedback in target detection: detection of semantic targets modulates activity in early visual areas. *Neuropsychologia* 47, 1721–1727. doi: 10.1016/j.neuropsychologia.2009.02.011
- Howard, M. A. 3rd., Volkov, I. O., Granner, M. A., Damasio, H. M., Ollendieck, M. C., and Bakken, H. E. (1996). A hybrid clinical-research depth electrode for acute and chronic *in vivo* microelectrode recording of human brain neurons. Technical note. *J. Neurosurg.* 84, 129–132. doi: 10.3171/jns.1996.84.1.0129
- Howard, M. A., Volkov, I. O., Mirsky, R., Garell, P. C., Noh, M. D., Granner, M., et al. (2000). Auditory cortex on the human posterior superior temporal gyrus. *J. Comp. Neurol.* 416, 79–92. doi: 10.1002/(SICI)1096-9861(2000103)416:1<79::AID-CNE6>3.0.CO;2-2
- Indefrey, P., and Cutler, A. (2004). “Prelexical and lexical processing in listening,” in *The Cognitive Neurosciences III*, ed M. Gazzaniga (Cambridge, MA: MIT Press), 759–774.
- Jenkinson, M., Bannister, P., Brady, M., and Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17, 825–841. doi: 10.1006/nimg.2002.1132
- Kiehl, K. A., Laurens, K. R., Duty, T. L., Forster, B. B., and Liddle, P. F. (2001). Neural sources involved in auditory target detection and novelty processing: an event-related fMRI study. *Psychophysiology* 38, 133–142. doi: 10.1111/1469-8986.3810133
- Kirino, E., Belger, A., Goldman-Rakic, P., and McCarthy, G. (2000). Prefrontal activation evoked by infrequent target and novel stimuli in a visual target detection task: an event-related functional magnetic resonance imaging study. *J. Neurosci.* 20, 6612–6618.
- Lachaux, J. P., Axmacher, N., Mormann, F., Halgren, E., and Crone, N. E. (2012). High-frequency neural activity and human cognition: past, present and possible future of intracranial EEG research. *Prog. Neurobiol.* 98, 279–301. doi: 10.1016/j.pneurobio.2012.06.008
- Leaver, A. M., and Rauschecker, J. P. (2010). Cortical representation of natural complex sounds: effects of acoustic features and auditory object category. *J. Neurosci.* 30, 7604–7612. doi: 10.1523/JNEUROSCI.0296-10.2010
- Leicht, G., Kirsch, V., Giegling, I., Karch, S., Hantschk, I., Möller, H. J., et al. (2010). Reduced early auditory evoked gamma-band response in patients with schizophrenia. *Biol. Psychiatry* 67, 224–231. doi: 10.1016/j.biopsych.2009.07.033
- Liégeois-Chauvel, C., Musolino, A., Badier, J. M., Marquis, P., and Chauvel, P. (1994). Evoked potentials recorded from the auditory cortex in man: evaluation and topography of the middle latency components. *Electroencephalogr. Clin. Neurophysiol.* 92, 204–214. doi: 10.1016/0168-5597(94)90064-7
- Mesgarani, N., and Chang, E. F. (2012). Selective cortical representation of attended speaker in multi-talker speech perception. *Nature* 485, 233–236. doi: 10.1038/nature11020
- Mesgarani, N., Cheung, C., Johnson, K., and Chang, E. F. (2014). Phonetic feature encoding in human superior temporal gyrus. *Science* 343, 1006–1010. doi: 10.1126/science.1245994
- Mesgarani, N., David, S. V., Fritz, J. B., and Shamma, S. A. (2008). Phoneme representation and classification in primary auditory cortex. *J. Acoust. Soc. Am.* 123, 899–909. doi: 10.1121/1.2816572
- Nelken, I. (2008). Processing of complex sounds in the auditory system. *Curr. Opin. Neurobiol.* 18, 413–417. doi: 10.1016/j.conb.2008.08.014
- Nelken, I., and Bar-Yosef, O. (2008). Neurons and objects: the case of auditory cortex. *Front. Neurosci.* 2, 107–113. doi: 10.3389/neuro.01.009.2008
- Niwa, M., Johnson, J. S., O’Connor, K. N., and Sutter, M. L. (2012). Activity related to perceptual judgment and action in primary auditory cortex. *J. Neurosci.* 32, 3193–3210. doi: 10.1523/JNEUROSCI.0767-11.2012
- Nourski, K. V., and Brugge, J. F. (2011). Representation of temporal sound features in the human auditory cortex. *Rev. Neurosci.* 22, 187–203. doi: 10.1515/rns.2011.016
- Nourski, K. V., Brugge, J. F., Reale, R. A., Kovach, C. K., Oya, H., Kawasaki, H., et al. (2013). Coding of repetitive transients by auditory cortex on posterolateral superior temporal gyrus in humans: an intracranial electrophysiology study. *J. Neurophysiol.* 109, 1283–1295. doi: 10.1152/jn.00718.2012
- Nourski, K. V., and Howard, M. A. 3rd. (2014). Invasive recordings in human auditory cortex. *Hand. Clin. Neurol.* (in press).
- Nourski, K. V., Reale, R. A., Oya, H., Kawasaki, H., Kovach, C. K., Chen, H., et al. (2009). Temporal envelope of time-compressed speech represented in the human auditory cortex. *J. Neurosci.* 29, 15564–15574. doi: 10.1523/JNEUROSCI.3065-09.2009
- Nourski, K. V., Steinschneider, M., Oya, H., Kawasaki, H., and Howard, M. A. 3rd. (2014a). Modulation of response patterns in human auditory cortex during a target detection task: an intracranial electrophysiology study. *Int. J. Psychophysiol.* doi: 10.1016/j.ijpsycho.2014.03.006. [Epub ahead of print].
- Nourski, K. V., Steinschneider, M., Oya, H., Kawasaki, H., Jones, R. D., and Howard, M. A. 3rd. (2014b). Spectral organization of the human lateral superior temporal gyrus revealed by intracranial recordings. *Cereb. Cortex* 24, 340–352. doi: 10.1093/cercor/bhs314
- Obleser, J., and Eisner, F. (2009). Pre-lexical abstraction of speech in the auditory cortex. *Trends Cogn. Sci.* 13, 14–19. doi: 10.1016/j.tics.2008.09.005
- Petkov, C. I., and Jarvis, E. D. (2012). Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates. *Front. Evol. Neurosci.* 4:12. doi: 10.3389/fnevo.2012.00012
- Petkov, C. I., Kang, X., Alho, K., Bertrand, O., Yund, E. W., and Woods, D. L. (2004). Attentional modulation of human auditory cortex. *Nat. Neurosci.* 7, 658–663. doi: 10.1038/nn1256
- Poeppl, D., Idsardi, W. J., and van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 1071–1086. doi: 10.1098/rstb.2007.2160
- Pugh, K. R., Shaywitz, B. A., Shaywitz, S. E., Fulbright, R. K., Byrd, D., Skudlarski, P., et al. (1996). Auditory selective attention: an fMRI investigation. *Neuroimage* 4, 159–173. doi: 10.1006/nimg.1996.0067
- Rauschecker, J. P., and Scott, S. K. (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat. Neurosci.* 12, 718–724. doi: 10.1038/nn.2331
- Reddy, C. G., Dahdaleh, N. S., Albert, G., Chen, F., Hansen, D., Nourski, K., et al. (2010). A method for placing Heschl gyrus depth electrodes. *J. Neurosurg.* 112, 1301–1307. doi: 10.3171/2009.7.JNS09404
- Russ, B. E., Orr, L. E., and Cohen, Y. E. (2008). Prefrontal neurons predict choices during an auditory same-different task. *Curr. Biol.* 18, 1483–1488. doi: 10.1016/j.cub.2008.08.054
- Sabri, M., Humphries, C., Verber, M., Mangalathu, J., Desai, A., Binder, J. R., et al. (2013). Perceptual demand modulates activation of human auditory cortex in response to task-irrelevant sounds. *J. Cogn. Neurosci.* 25, 1553–1562. doi: 10.1162/jocn_a_00416
- Sahin, N. T., Pinker, S., Cash, S. S., Schomer, D., and Halgren, E. (2009). Sequential processing of lexical, grammatical, and phonological information within Broca’s area. *Science* 326, 445–449. doi: 10.1126/science.1174481
- Shahin, A. J., Alain, C., and Picton, T. W. (2006). Scalp topography and intracerebral sources for ERPs recorded during auditory target detection. *Brain Topogr.* 19, 89–105. doi: 10.1007/s10548-006-0015-9
- Shinn-Cunningham, B. G. (2008). Object-based auditory and visual attention. *Trends Cogn. Sci.* 12, 182–186. doi: 10.1016/j.tics.2008.02.003
- Simon, J. Z. (2014). The encoding of auditory objects in auditory cortex: insights from magnetoencephalography. *Int. J. Psychophysiol.* doi: 10.1016/j.ijpsycho.2014.05.005. [Epub ahead of print].
- Steinschneider, M., Fishman, Y. I., and Arezzo, J. C. (2008). Spectrotemporal analysis of evoked and induced electroencephalographic responses in primary auditory cortex (A1) of the awake monkey. *Cereb. Cortex* 18, 610–625. doi: 10.1093/cercor/bhm094
- Steinschneider, M., Nourski, K. V., and Fishman, Y. I. (2013). Representation of speech in human auditory cortex: is it special? *Hear. Res.* 305, 57–73. doi: 10.1016/j.heares.2013.05.013
- Steinschneider, M., Nourski, K. V., Kawasaki, H., Oya, H., Brugge, J. F., and Howard, M. A. 3rd. (2011). Intracranial study of speech-elicited activity on the human posterolateral superior temporal gyrus. *Cereb. Cortex* 21, 2332–2347. doi: 10.1093/cercor/bhr014
- Steinschneider, M., Schroeder, C. E., Arezzo, J. C., and Vaughan, H. G. Jr. (1994). Speech-evoked activity in primary auditory cortex: effects of voice onset time. *Electroencephalogr. Clin. Neurophysiol.* 92, 30–43. doi: 10.1016/0168-5597(94)90005-1
- Sutter, M., O’Connor, K. N., Downer, J., Johnson, J., and Niwa, M. (2013). Hierarchical effects of attention on amplitude modulation encoding in auditory cortex. *J. Acoust. Soc. Am.* 134, 4085. doi: 10.1121/1.4830920

- Travis, K. E., Leonard, M. K., Chan, A. M., Torres, C., Sizemore, M. L., Qu, Z., et al. (2013). Independence of early speech processing from word meaning. *Cereb. Cortex* 23, 2370–2379. doi: 10.1093/cercor/bhs228
- Tsunada, J., Lee, J. H., and Cohen, Y. E. (2011). Representation of speech categories in the primate auditory cortex. *J. Neurophysiol.* 105, 2634–2646. doi: 10.1152/jn.00037.2011
- Winkler, I., van Zuijlen, T. L., Sussman, E., Horváth, J., and Näätänen, R. (2006). Object representation in the human auditory system. *Eur. J. Neurosci.* 24, 625–634. doi: 10.1111/j.1460-9568.2006.04925.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 29 May 2014; accepted: 22 July 2014; published online: 11 August 2014.
Citation: Steinschneider M, Nourski KV, Rhone AE, Kawasaki H, Oya H and Howard MA III (2014) Differential activation of human core, non-core and auditory-related cortex during speech categorization tasks as revealed by intracranial recordings. *Front. Neurosci.* 8:240. doi: 10.3389/fnins.2014.00240

This article was submitted to Auditory Cognitive Neuroscience, a section of the journal *Frontiers in Neuroscience*.

Copyright © 2014 Steinschneider, Nourski, Rhone, Kawasaki, Oya and Howard. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.