

Mapping phonemic processing zones along human perisylvian cortex: an electro-corticographic investigation

Sophie Molholm · Manuel R. Mercier ·
Einat Liebenthal · Theodore H. Schwartz ·
Walter Ritter · John J. Foxe · Pierfilippo De Sanctis

Received: 11 December 2012 / Accepted: 3 May 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract The auditory system is organized such that progressively more complex features are represented across successive cortical hierarchical stages. Just when and where the processing of phonemes, fundamental elements of the speech signal, is achieved in this hierarchy remains a matter of vigorous debate. Non-invasive measures of phonemic representation have been somewhat equivocal. While some studies point to a primary role for middle/anterior regions of the superior temporal gyrus (STG), others implicate the posterior STG. Differences in stimulation, task and inter-

individual anatomical/functional variability may account for these discrepant findings. Here, we sought to clarify this issue by mapping phonemic representation across left perisylvian cortex, taking advantage of the excellent sampling density afforded by intracranial recordings in humans. We asked whether one or both major divisions of the STG were sensitive to phonemic transitions. The high signal-to-noise characteristics of direct intracranial recordings allowed for analysis at the individual participant level, circumventing issues of inter-individual anatomic and functional variability that may have obscured previous findings at the group level of analysis. The mismatch negativity (MMN), an electrophysiological response elicited by changes in repetitive streams of stimulation, served as our primary dependent measure. Oddball configurations of pairs of phonemes, spectro-temporally matched non-phonemes, and simple tones were presented. The loci of the MMN clearly differed as a function of stimulus type. Phoneme representation was most robust over middle/anterior STG/STS, but was also observed over posterior STG/SMG. These data point to multiple phonemic processing zones along perisylvian cortex, both anterior and posterior to primary auditory cortex. This finding is considered within the context of a dual stream model of auditory processing in which functionally distinct ventral and dorsal auditory processing pathways may be engaged by speech stimuli.

Electronic supplementary material The online version of this article (doi:10.1007/s00429-013-0574-y) contains supplementary material, which is available to authorized users.

S. Molholm (✉) · M. R. Mercier · W. Ritter ·
J. J. Foxe · P. De Sanctis (✉)
The Sheryl and Daniel R. Tishman Cognitive Neurophysiology
Laboratory, Departments of Pediatrics and Neuroscience,
Children's Evaluation and Rehabilitation Center (CERC), Albert
Einstein College of Medicine, Van Etten Building–Wing 1C,
1225 Morris Park Avenue, Bronx, NY 10461, USA
e-mail: sophie.molholm@einstein.yu.edu

P. De Sanctis
e-mail: pierfilippo.sanctis@einstein.yu.edu

S. Molholm · J. J. Foxe
The Cognitive Neurophysiology Laboratory, The Nathan S.
Kline Institute for Psychiatric Research, 140 Old Orangeburg
Road, Orangeburg, NY 10962, USA

E. Liebenthal
Department of Neurology, Medical College of Wisconsin,
8701 W Watertown Plank Rd, Milwaukee, WI 53226, USA

T. H. Schwartz
Department of Neurological Surgery, Weill Cornell Medical
College, New York Presbyterian Hospital, New York,
NY 10021, USA

Keywords ECoG · Phoneme · Superior temporal gyrus ·
Auditory · Speech processing · Intracranial

Introduction

Spoken language is integral to human communication, allowing us to exchange both simple and complex ideas.

Though initially sampled through the auditory sensory modality, its representation in the brain is rapidly transformed from simple acoustic features such as frequency, intensity, and duration, to more abstracted categorical units of speech (i.e., phonemes). The processing of phonemes (as compared to acoustically matched non-phonemes) engages an extended cortical network that reaches from secondary auditory cortex to frontal and parietal regions (Scott and Johnsrude 2003; Liebenthal et al. 2005; see Turkeltaub and Coslett 2010 for a meta-analysis of the related imaging data). Consistent with animal models of hierarchical organization of auditory cortex (Rauschecker 1998; Hackett 2011), patterns across imaging studies reliably indicate that the anatomy of neural representation within auditory cortex varies as a function of stimulus complexity (Peelle et al. 2010). Representations of more abstracted units such as phonemes extend outward from primary auditory cortex (i.e., Heschl's gyrus), where a veridical representation of the frequency structure of the signal is maintained. Perplexingly, while there is a clear segregation from veridical to more abstract, the specific region of auditory cortex associated with the representation of basic speech sounds varies considerably from study to study in terms of where it falls along the anterior to posterior axis of the superior temporal sulcus (STS) (Obleser and Eisner 2009). This contrasts with processing in the visual domain where there is a highly systematic and circumscribed temporal/anatomical hierarchy that corresponds with the extraction of increasing levels of image complexity and abstraction (Felleman and Van Essen 1991). This includes well-defined anatomical regions that respond best to faces (the fusiform face area: FFA; e.g. Kanwisher et al. 1997), places (the parahippocampal place area: PPA; e.g. Epstein and Kanwisher 1998), orthographic words (visual word form area; Allison et al. 1999), and body parts (occipito-temporal cortex-body part area: OTC-BPA; Orlov et al. 2010), although there is certainly some controversy regarding the organizing principles of these areas (e.g. Hasson et al. 2002; McKeef et al. 2010).

There is compelling evidence from animal physiology as well as human neuroimaging for functionally distinct ventral and dorsal auditory processing pathways. With respect to speech perception, these appear to be biased toward, respectively, the identification of speech stimuli, and sensori-motor aspects of speech processing (see e.g. Rauschecker 1998; Hickok and Poeppel 2007; Peelle et al. 2010; Chevillet et al. 2011). It is reasonable to suppose that differences in stimulation parameters, experimental design, and task demands, would engage to a greater or lesser extent these so-called ventral and dorsal streams of auditory processing, and thus account in part for the observed inconsistencies in the neural representation of prelexical speech sounds observed across studies.

In addition to the role of task and stimulus parameters, however, it is also possible that these inconsistencies partially reflect naturally occurring inter-individual variation in the anatomy of the sulci and gyri of the brain (Roland and Zilles 1998; Rademacher et al. 2002), and that specific localization of function within a cortical region may vary across individuals (Fedorenko and Kanwisher 2011). Indeed, it is well documented that there is great anatomical variation in auditory cortical structures such as Heschl's gyrus and planum temporale across individuals (see, e.g. Ide et al. 1996; Abdul-Kareem and Sluming 2008). Functional magnetic resonance imaging (fMRI) has served as the primary probe of the anatomical representation of speech because of its superior anatomical localization compared to other non-invasive approaches to mapping brain activity. However, functional neuroimaging data are almost always considered at the group level, and group-level statistical analyses may be blind to speech-related activity when the inter-subject anatomical-functional variability exceeds the unit of analysis (e.g., usually between a 3 mm and a 12 mm cubed region of cortex; see Kang et al. 2004, for a considered introduction to the issues around imaging of functional activity in auditory cortex). For example, in a previous study, phoneme-related activations in pSTS were found to vary across individuals within a 10 mm radius (Desai et al. 2005). Thus, it is possible that within a given individual there is representation of speech sounds in STS both anterior and posterior to primary auditory cortex, due to feedforward activation of both ventral and dorsal auditory processing streams. Statistically significant effects might be achieved in some studies having adequate inter-participant overlap in more anterior locations along the superior temporal gyrus (STG), and others conversely having adequate inter-participant overlap in more posterior locations along the STG.

On rare occasions, when subdural electrode implantation is clinically required, one has the opportunity to record electrical activity directly from the cortical surface of the human brain and thus acquires both high temporal and high anatomical resolution recordings of brain activity. The strength of the intracranially recorded signal allows statistical analysis of individual participant data, and thus inter-individual anatomical variability need not obscure results. Moreover, the millisecond-scale temporal resolution of the EEG signal allows consideration of the temporal dynamics of neural activity in response to speech, and thus one may assay whether speech selective signals are best attributed to sensory-perceptual processing, or if their timing is more commensurate with later post-perceptual processes (Lucan et al. 2010).

Here we took advantage of access to individuals undergoing surgical evaluation for the localization of epileptic seizures, to map the neural representation of phonemes. Neural activity in

the form of local field potentials was recorded over the surface of the left temporal, parietal, and frontal lobes where dense array electrode grids had been placed subdurally for clinical purposes. We examined the loci of the cortical representation of the speech syllables /ba/ and /da/ as compared to spectrotemporally matched non-speech stimuli and simple tones. Stimuli were presented in an oddball configuration and the so-called mismatch negativity (MMN; see e.g. Ritter et al. 1992; Picton et al. 2000; De Sanctis et al. 2009; Butler et al. 2011), which occurs in response to infrequent deviations from regularities in acoustic stimulation, was recorded. The MMN is elicited regardless of whether the participant attends the inputs (e.g., Näätänen 1992; Näätänen et al. 2007), but only in response to stimulus changes that can be perceived when the stimuli are actively attended (Amenedo and Escera 2000), and thus is tightly linked to perceptual experience. The location of the neural generators of the MMN varies with the deviating feature, and this location is considered to reflect where the deviating feature is cortically represented (Molholm et al. 2005). Thus, it was hypothesized that for a given individual, dense array electrode coverage would reveal that MMNs were generated in different cortical regions as a function of stimulus class. Further, it was expected that these data would reveal that within an individual, basic acoustic features would be represented in the regions (i.e., recording sites) sensitive to activity from Heschl's gyrus (reflected in contacts around the post-central sulcus, on either side of the sylvian fissure, and the medial STG), and complex stimuli would be preferentially processed at both anterior and posterior regions to this. The primary question of this study was where the preferential representation of speech stimuli would be observed: anterior/medial STG, posterior STG, or, consistent with a dual stream model of speech processing, in both regions.

Methods

Participants

Patients were implanted with subdural electrodes for evaluation of the foci of pharmacologically intractable epilepsy. Placement was determined on the basis of clinical need. In all three cases, a grid of electrodes was placed over the left hemisphere, centered on the sylvian fissure and reaching both posterior and anterior portions therein. For each of the participants, the placement of the borders of the grids is illustrated by a rectangular box in Figs. 2, 3, 4 (and see Supplementary Fig. 1). The participants provided written informed consent, and the procedures were approved by the Institutional Review Boards of the Nathan Kline Institute, the City College of New York, and Weill Cornell Presbyterian hospital. The conduct of this study strictly adhered to the principles outlined in the Declaration of Helsinki.

Table 1 Participant demographic and clinical characteristics

	Participants		
	AA	TK	CL
Initials	AA	TK	CL
Gender	F	M	M
Handedness	R	R	R
Age	24	23	16
Implanted hemisphere	Left	Left	Left
No. of implanted electrodes	100	100	64
Language lateralization (WADA)	Left	Left	Left
Neuropsychological exam	Average range	Below-average range	Average range

Table 1 presents demographic and clinical characteristics of the patients. Neuropsychological test ranged from mild impairment to normal. Language was left lateralized in all participants as determined by Wada testing (Wada and Rasmussen 1960).

Stimuli and design

Phoneme stimuli were created using a cascade/parallel formant synthesizer (SenSyn Laboratory Speech Synthesizer, Sensimetrics Corp., Cambridge, MA). Pitch, intensity, formant bandwidth, and formant center frequency parameters for synthesis were derived from natural utterances of the syllables /ba/ and /da/ produced by a male speaker and sampled at 44.1 kHz. By spectrally inverting the first formant (F1) and increasing the slopes of the third formant (F3) of the syllables /ba/ and /da/, the non-phoneme analogs were created (Liebenthal et al. 2005). Phoneme and non-phoneme control stimuli were 150 ms in duration with a 5 ms rise–decay envelope and presented at ~ 75 dB. Pure tones had a frequency of 720 and 800 Hz, a duration of 100 ms, a rise and fall time of 10 ms, and an intensity of 75 dB. Auditory stimuli were presented binaurally via headphones (Sennheiser HD-600). We used an oddball paradigm presenting standard and deviant auditory stimuli with a likelihood of 0.8–0.2, with the rule that at least three standards preceded a deviant. Participants ignored the sounds and watched a silent movie. For the three conditions, phoneme-MMN, matched non-phoneme-MMN, and pure tone-MMN, standard and deviant stimuli were counterbalanced across blocks within participant so that stimulus evoked activity was matched across the standard-deviant comparisons.

ECoG recording and analysis

Continuous EEG from 64 to 100 electrodes was recorded using BrainAmp amplifiers (Brain Products, Munich, Germany). The electrodes (stainless steel electrodes from AD-Tech Medical Instrument Corporation, Racine, WI)

were 2.5 mm in diameter with an interelectrode-spacing of 1 cm (note that the non-rigidity of the grid and that it is placed on a surface with changes in elevation can lead to slightly smaller interelectrode spacing). Subdural electrodes are highly sensitive to local field potentials generated within a 4.0 mm² area and much less sensitive to more distant activity (Allison et al. 1999; Sehatpour et al. 2008), which allows for considerably improved localization of the underlying current sources relative to scalp-recorded EEG. The data were band-pass filtered on-line from 0.1 to 250 Hz, and digitized at 1,000 Hz. A frontally placed intracranial electrode served as the online reference. Data were re-referenced off-line to a silent channel; that is, a channel in which no significant stimulus-related electrical brain activity was observed. Off-line, an automated artifact detection algorithm was used to reject trials with noise transients. This method is based on the standard score of signal amplitude, with *z*-values computed at every data point across time and independently for each electrode. Criterion for automatic rejection was set at $z = 4$. Artifact-free epochs, from -100 ms to $+500$ ms, were then linearly detrended, demeaned (baseline corrected over the entire epoch), and filtered. For filtering, first, a discrete Fourier transform was used to remove line noise (60 Hz and harmonics 120 and 180), and then the signal was low pass filtered at 30 Hz (two pass Butterworth filter). The first four trials of each run were excluded from the averages to avoid “novelty” effects. Single trials were averaged to compute ERPs for each of the main conditions (Phoneme, non-phoneme, and pitch), for the deviant and standard stimuli. Across all three participants, the average number of accepted trials for the standard stimuli was 797 with a range of 657–902 for phoneme trials, 847 with a range of 694–993 for non-phoneme trials, and 796 with a range of 643–894 for pitch trials. For deviant stimuli, the average number of accepted trials was 163 with a range of 144–185 for phoneme trials, 157 with a range of 114–193 for non-phoneme trials, and 161 with a range of 123–186 for pitch trials.

Derivation of the MMN response

Individual participant data were analysed for the presence of the MMN for each of the phoneme, non-phoneme, and tone conditions separately by comparing the respective standard and deviant responses. The MMN was expected to onset between 100 and 300 ms post-deviance onset, and to be seen over the temporal lobe, largely along the sylvian fissure, perisylvian regions, and the STS, as well as possibly over the parietal and frontal lobes. To determine whether amplitude differences between the standard and deviant responses were statistically significant, an unpaired randomization-permutation procedure was performed. For

this procedure, base-line corrected single trials from the two stimulus conditions (standard and deviant) were mixed and then randomly assigned to one of two separate pools of trials, with the only constraint being that the two randomized pools maintained their original size. The average of the first pool was then subtracted from the average of the second pool. By repeating this procedure 1,000 times, a distribution was built from which significance thresholds could be determined. The *p* value was computed by comparing the observed difference between the standard and deviant ERP to the surrogate distribution.

To assess differences in the MMNs between conditions (e.g., whether the MMN to the phoneme condition was larger/smaller, earlier/later than the MMN to the non-phoneme control condition), we conducted an additional unpaired randomization-permutation procedure. First, base-lined trials from the two compared conditions were shuffled (e.g., phoneme and non-phoneme control), with randomized pools keeping their original size. This randomization-permutation step was done separately for standard and deviant responses (i.e., stimulus type were not mixed, only conditions). Then the difference between deviant and standard was computed for each randomized pool condition and the resulting MMNs were used to build the surrogate distribution.

For all tests, the *p*-value threshold was set at 0.05 (two tailed). To control for multiple comparisons, only *p*-values that were significant for at least 20 ms between 100 and 300 ms post-stimulus (20 consecutive time points at 1,000 Hz)¹ were considered (Guthrie and Buchwald 1991; Pourtois et al. 2007, 2010). All data analyses were performed in Matlab (the MathWorks, Natick, MA, USA) using custom written scripts, and the Fieldtrip Toolbox (Oostenveld et al. 2011).

Localization of electrodes

For AA and TK, high-resolution presurgical MRIs, post-surgery MRIs, and CT scans were realigned and co-registered with the Montreal Neurological Institute template brain (MNI-colin 27 brain, hereafter MNI template brain). The preoperative MRI was used for its anatomical information, the post-operative CT scan provided undistorted placement of the electrode contacts, and the post-operative MRI allowed assessment of the quality of the entire co-registration process since it includes both electrodes and anatomical information. Co-registration was performed using Statistical Parametric Mapping (SPM8) developed by

¹ We reanalyzed our data without the low-pass filter and found no differences between filtered (30 Hz) and unfiltered data with regard to the statistical results for the MMNs. Data from the original analyses are presented.

The Wellcome Department of Imaging Neuroscience. The precise electrode location based on CT information was mapped onto pre/post-MRI using bioimage suite (<http://www.bioimagesuite.org>; X. Papademetris, M. Jackowski, N. Rajeevan, H. Okuda, R.T. Constable, and L.H Staib. BioImage Suite: An integrated medical image analysis suite, Section of Bioimaging Sciences, Dept. of Diagnostic Radiology, Yale School of Medicine, <http://www.bioimagesuite.org>). These placements were then projected on the MNI template brain based on these original localizations on the individual's sulci and gyri. For CL, no presurgical MRI was available and so the post-operative MRI was directly co-registered to the MNI template. For all registration processes, two of the authors (MM and PDS) independently performed all registration processes and validated the results based on inter-rater agreement and the individual anatomies. Electrode placement on the individual brains, for AA and TK, is provided as supplementary material (supplementary Fig. 1).

Mapping the MMNs by stimulus class

The following classification scheme, based on the statistical presence and absence of MMNs for the three stimulus types at a given recording site and statistical differences in the amplitude and latency of those MMNs, was used in determining specificity of the response at a given recording site:

- Greatest sensitivity to speech: MMN to the phoneme condition that is statistically greater in amplitude and/or of longer duration as compared to the other conditions (electrode placements depicted in red).
- Greatest sensitivity to complex spectral information: MMN to non-phoneme and phoneme conditions (depicted in blue), or just the non-phoneme condition (also blue).
- Greatest sensitivity to simple acoustic information: MMN to tones only, or to all types of acoustic inputs, but strongest to tones (depicted in green).
- Combinations thereof are represented by multiple colors (e.g., a half green and half blue circle).

Results

The goal of this study was to map auditory cortical areas responsive to phonemes, and to place these localizations with respect to areas most sensitive to simple tones and complex non-speech stimuli. In separate blocks, an oddball paradigm was employed using simple tones (of 720 and 800 Hz), complex non-phonemic stimuli matched spectrally and temporally to the /ba/ and /da/ phoneme tokens, and minimally contrasting phonemes /ba/ and /da/, while

subdural recordings of brain electrical activity were made over extensive portions of the temporal lobe and perisylvian regions. The resulting data provide compelling evidence for the representation of speech sounds that are most reliable in the middle/anterior regions of the STG/STS, but that are also observed over posterior regions (here pSTG and supramarginal gyrus (SMG)). First, we present a description of the composite results, followed by presentation of the individual participant data.

Composite results

Findings from the three participants, described in detail below, are projected onto the left hemisphere of the MNI template brain (Fig. 1). The clearest pattern to emerge for speech-specific regions was a cluster at anterior/medial portions of the sylvian fissure (the MTG/STG at about the central sulcus). Tone-specific regions were focused around, though not restricted to, Heschl's gyrus (and the landmark the post-central gyrus). Regions classified as most responsive to complex stimulation tended to be interspersed with tone and speech selective areas.

Individual participant data

Data are presented to highlight the anterior to posterior representation of the different classes of auditory stimulation, using the MMN as our dependent measure. The MNI y coordinate of a given channel is provided, for placement along the anterior–posterior axis. Channels are referred to by their original numbering in the grid, and these numbers

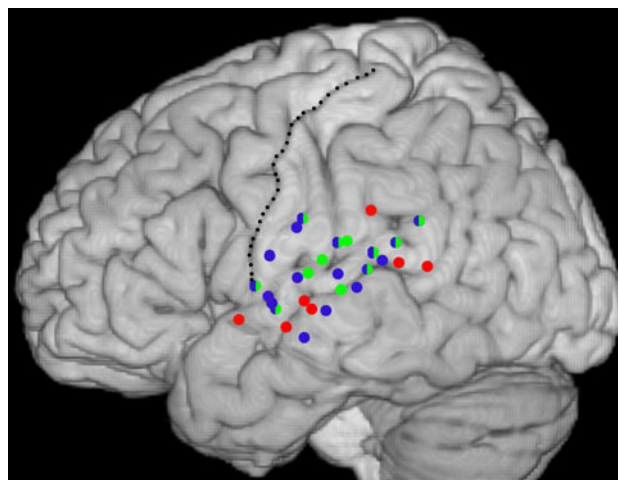
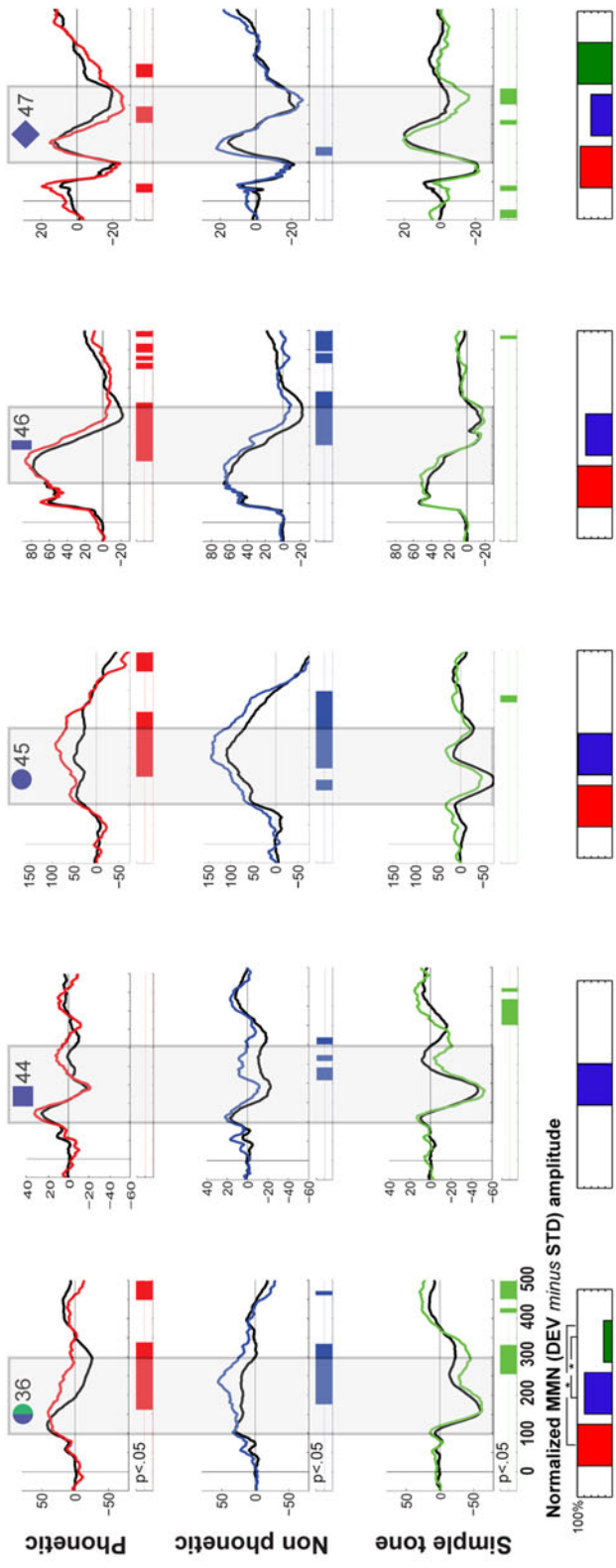


Fig. 1 Composite data from the three participants. Electrode locations that showed significant MMN responses are projected onto the MNI template brain. *Red circles* indicate electrodes classified as most responsive to speech, *blue circles* electrodes classified as responsive to complex stimulation, and *green circles* electrodes classified as most responsive to tones. Double classifications are indicated by *bi-colored circles*. The central sulcus is delineated with a *stippled line*



Participant: AA

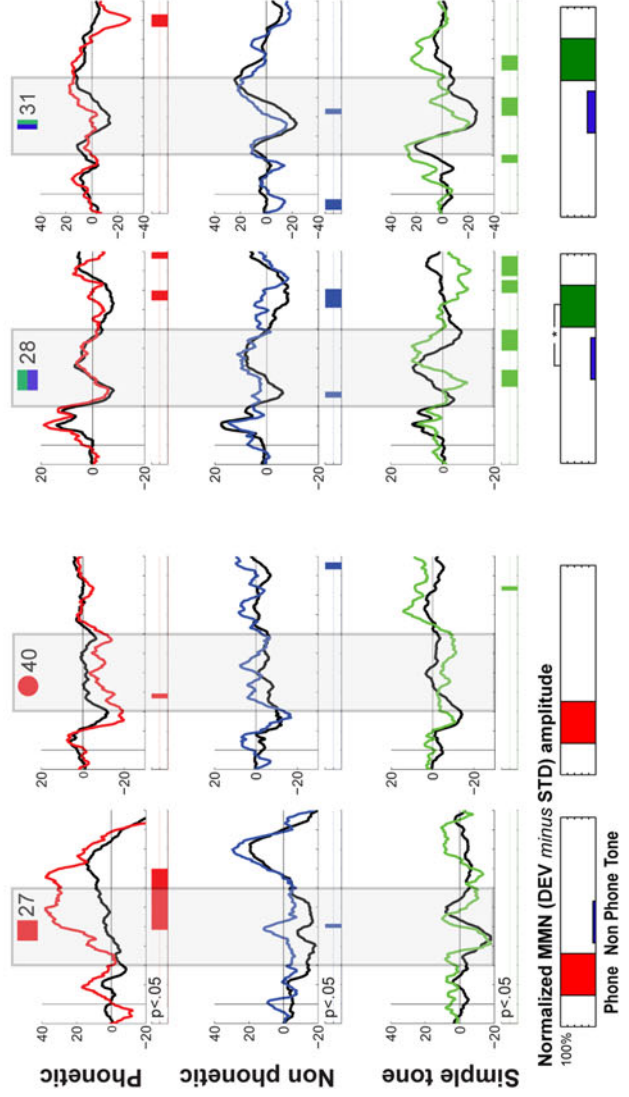
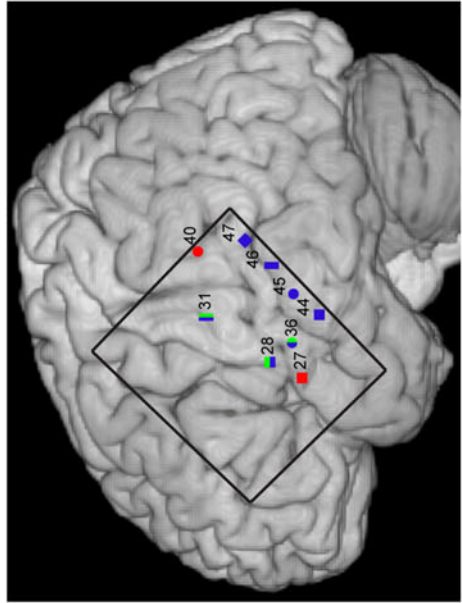


Fig. 2 Individual data for participant AA. Evoked responses to the standard (*black trace*) and deviant (*colored trace*) stimuli are shown for electrodes that exhibited significant MMN responses. Periods of significant difference are indicated by the *color strip* below the waveforms. The period of 100–300 ms post-stimulus onset is highlighted by a *gray box*. In the lower right corner, electrode positions are illustrated on the MNI template brain. Electrodes are shape and number coded to link the specific waveform data with the location. *Color coding* indicates the designation of the channel: *red shapes* indicate electrodes where the MMN was greatest in response to speech stimulation; *blue shapes* electrodes most responsive to complex stimulation; and *green shapes* electrodes most responsive to simple tones. Double classifications are indicated by *bi-colored* shapes. The *black square* on the brain shows the borders of the electrode grid. Below each column of waveform responses, a *bar graph* quantifies stimulus-specific MMN responses for each electrode site, by showing the integrated amplitude of the statistically significant effect (to therefore represent both its amplitude and duration), within the 100–300 ms post-stimulus onset window that was considered in our analyses. Values are scaled with respect to the largest at a given contact (which is set at 100 %). Statistical differences between MMNs to the different stimulus types are indicated by *asterisks*

should be considered arbitrary. The standard and deviant responses and points of significant divergence (the MMN) are depicted in Figs. 2, 3, 4 (for participants AA, TK, and CL, respectively), with contacts localized on the MNI template brain. MNI coordinates, MMN onset latency, approximate Brodmann area, and gyrus/sulcal location are presented in Table 2. The intention is to provide a description that indicates the relative mapping of MMNs by stimulus class.

Participant AA (see Fig. 2 and Table 2)

Sylvian fissure The most anterior response along the sylvian fissure was speech-specific (MMN to the speech condition only), positioned on the anterior/middle portion of the STG (ch27, $y = 9$). Moving posterior, channels on either side of the sylvian fissure and aligned with the central sulcus were responsive to both tones and complex stimulation (ch28, $y = -3$, with MMNs to both tone and non-phoneme control conditions, and the tone MMN significantly greater than the non-phoneme control MMN; ch36, $y = -5$, receiving two classifications because the different morphology of the base responses for the simple tone versus the complex stimulation conditions suggests different underlying neural generators). Channel 36 had MMNs to all conditions, which did not differ significantly between the phoneme and non-phoneme complex conditions.

Middle temporal gyrus An extended area along the middle temporal gyrus just below the STS was sensitive to complex auditory stimulation (ch44–47). At the most anterior of the four channels there was an MMN in response to the complex non-speech condition (ch44, $y = -12$). Moving posterior, channels 45 and 46 had MMNs in response to both the

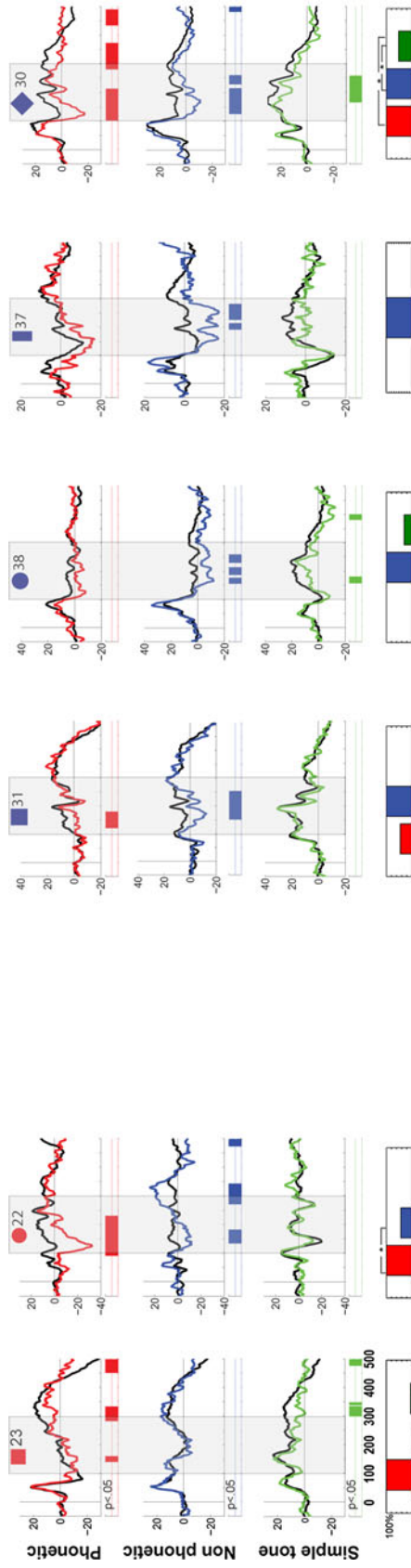
complex non-speech and phoneme stimuli, with MMNs that did not differ significantly between the two conditions ($y = -18$ and -27 , respectively). Finally, at the posterior portion of the STS, ch47 showed significant MMN responses to all conditions. The similar morphology of the responses suggested similar neuronal generators that were sensitive to complex as well as simple auditory stimulation ($y = -34$); none of the pair-wise comparisons revealed significant differences between the MMNs.

Parietal lobe A portion of the inferior parietal lobe by the post-central sulcus was sensitive to the non-phoneme condition and simple tones (ch31, $y = -14$), with MMNs to the pitch condition and non-phoneme condition that did not differ significantly, and no MMN to the speech condition. Just posterior to the post-central sulcus, on the SMG, was selective to the speech condition (ch 40, $y = -32$), with only an MMN to the phoneme condition.

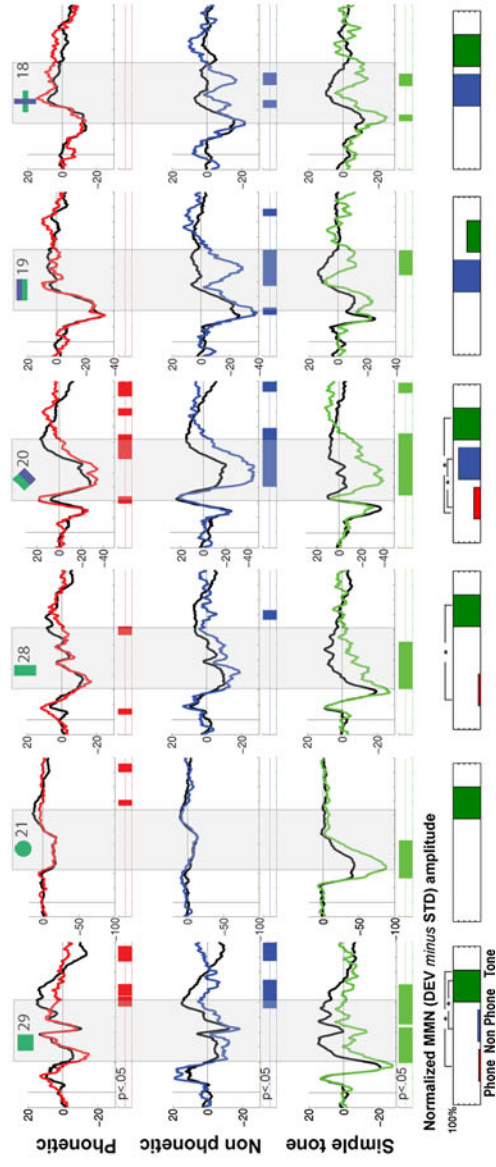
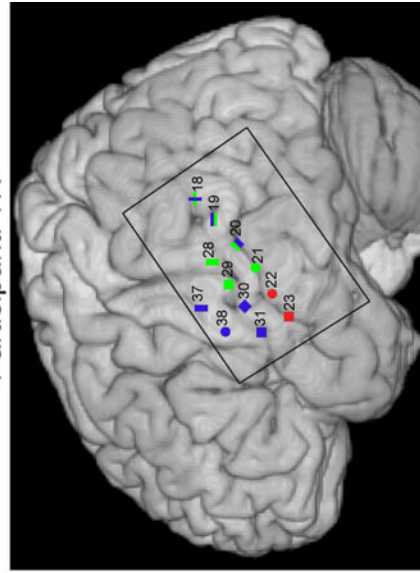
Participant TK (see Fig. 3 and Table 2)

Sylvian fissure The most anterior response along the sylvian fissure aligned with the central sulcus and was sensitive to complex acoustic stimulation (ch31, $y = -3$). MMNs were present in response to both the phoneme and non-phoneme control conditions, which did not differ significantly. An affinity for complex auditory stimulation was observed just posterior (ch30, $y = -11$), with MMNs to all three conditions, and similar MMNs to the phoneme and non-phoneme control conditions, which both onset significantly earlier than the tone-MMN. Proximal to the region of Heschl's gyrus was a region that responded best to tones: Channel 29 had a significantly greater MMN to the tone condition, as compared to the later onsetting and shorter duration MMNs to the other conditions (ch29, $y = -19$), and channel 28 showed an MMN to the tone condition (and a significantly smaller and much later MMN to the speech condition; ch28, $y = -26$). At the posterior portion of the sylvian fissure, on the pSTG, were two channels that responded similarly to the tone and non-phoneme control conditions (ch19 and ch18, $y = -40$ and -48 , respectively).

Superior temporal sulcus Speech dominant responses were located on the anterior/middle STS (channels 23 and 22, $y = -8$ and -14). For the more anterior of these, ch23, there was only a speech MMN in the time-window of consideration. Channel 22 exhibited MMNs to both the speech and non-phoneme control conditions, with the latter of shorter duration and onsetting later than the phoneme-MMN, differences that were confirmed statistically. Posterior to this, channel 21 ($y = -24$) only responded to the



Participant: TK



◀ **Fig. 3** Individual data for participant TK. Evoked responses to the standard (*black trace*) and deviant (*colored trace*) stimuli are shown for electrodes that exhibited significant MMN responses. Periods of significant difference are indicated by the *color strip* below the waveforms. The period of 100–300 ms post-stimulus onset is highlighted by a *gray box*. In the lower right corner, electrode positions are illustrated on the MNI template brain. Electrodes are shape and number coded to link the specific waveform data with the location. *Color coding* indicates the designation of the channel: *red shapes* indicate electrodes where the MMN was greatest in response to speech stimulation; *blue shapes* electrodes most responsive to complex stimulation; and *green shapes* electrodes most responsive to simple tones. Double classifications are indicated by *bi-colored shapes*. The *black square* on the brain shows the borders of the electrode grid. Below each column of waveform responses, a *bar graph* quantifies stimulus-specific MMN responses for each electrode site, by showing the integrated amplitude of the statistically significant effect (to therefore represent both its amplitude and duration), within the 100–300 ms post-stimulus onset window that was considered in our analyses. Values are scaled with respect to the largest at a given contact (which is set at 100 %). Statistical differences between MMNs to the different stimulus types are indicated by *asterisks*

simple tone condition, with negligible sensory responses to the other two conditions. The greater negativity to the deviant stimulus was restricted to the timeframe of the N1 response, and may represent a simple release from refractory effect in response to the deviant stimulus rather than the MMN (Naatanen and Picton 1987). Further posterior on the STS, sensitivity to complex acoustic stimulation and tones was observed, with MMN responses to all stimulation types. There was no significant difference between the simple tone and non-phoneme control conditions (ch20, $y = -30$), whereas the speech MMN was significantly smaller from 150 to 250 ms.

Parietal lobe The inferior parietal lobe by the central sulcus was sensitive to complex auditory stimulation. MMNs were present for the non-phoneme and tone conditions (ch38, $y = -4$), which did not significantly differ from each other, and, moving posterior, for only the non-phoneme control condition (ch37, $y = -11$).

Participant CL (see Fig. 4 and Table 2)

Sylvian fissure/superior temporal gyrus The most anterior response along the sylvian fissure, positioned just below and posterior to the central sulcus, was sensitive to complex auditory stimulation. MMNs were elicited by both the phoneme and non-phoneme control conditions (ch48, $y = -4$) that did not differ significantly. There was also a significant amplitude difference between the standard and deviant response for the tone condition, but the deviant response was uniformly of smaller amplitude than the standard, and therefore this was not considered an MMN

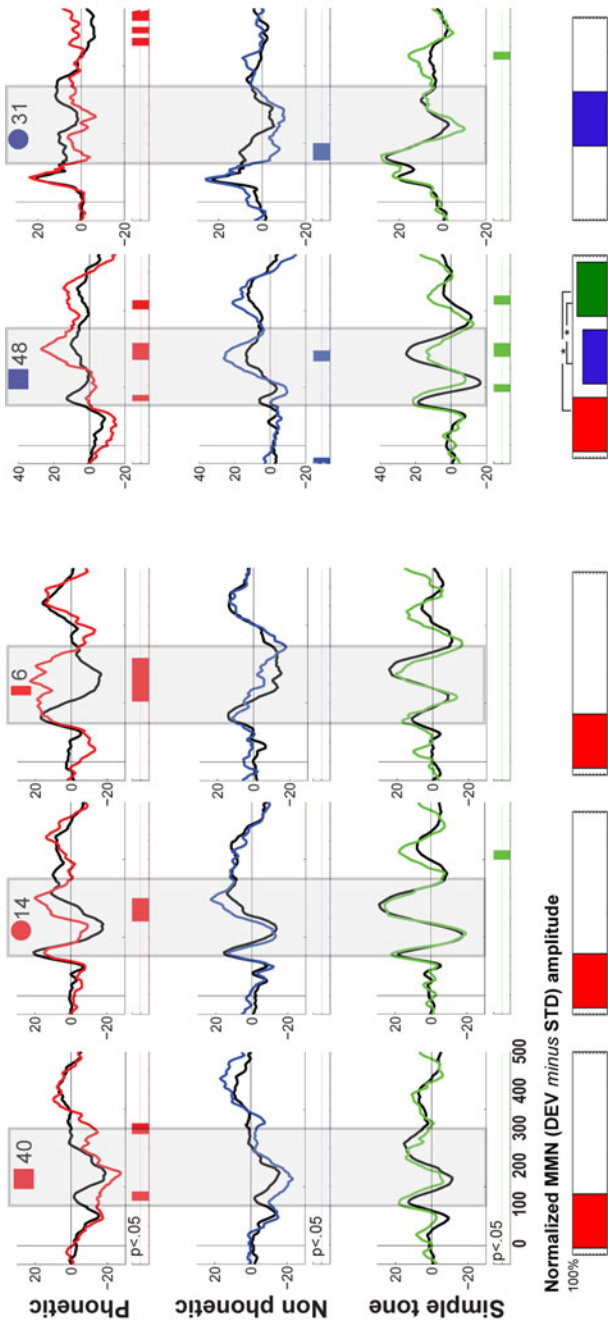
(we note nevertheless that this electrode did register differential activity in response to tone deviancy, or the frequency of the deviant tone, and thus this region was clearly sensitive to tonal information). Moving posterior, in the region of Heschl's gyrus, channel 39 was most responsive to simple auditory stimulation, with only a tone-MMN observed ($y = -14$). Further posterior, an MMN was elicited in response to non-phoneme complex acoustic stimulation but not simple tones or speech (ch31, $y = -24$). Further posterior still on the STG was responsive to both the tone and non-phoneme control condition (ch30, $y = -32$), with MMNs that did not differ significantly between the two stimulus types. At the most posterior point of the STG a speech selective region was identified, with only speech MMNs at channels 14 and 6 ($y = -41$ and -53 , respectively).

Superior temporal sulcus The anterior/medial portion of the STS showed speech specificity, with an MMN to only the phoneme condition (ch40, $y = -13$).

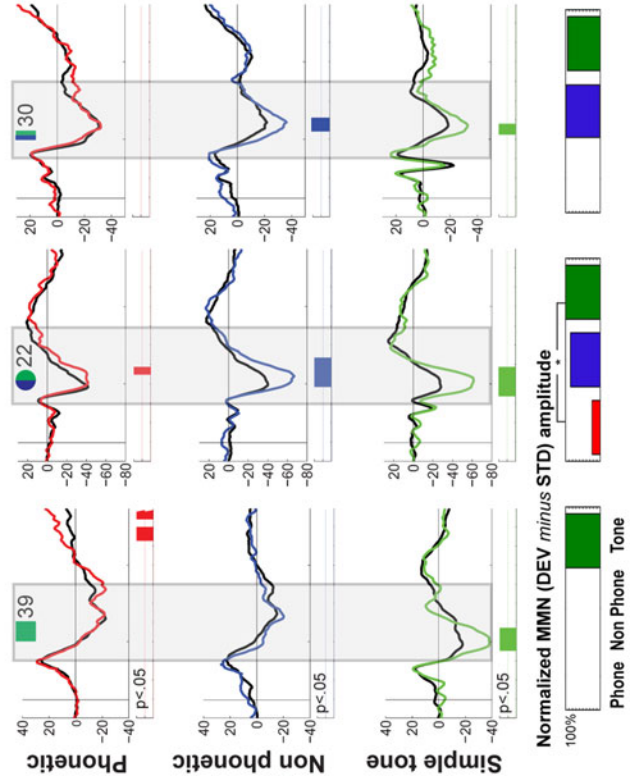
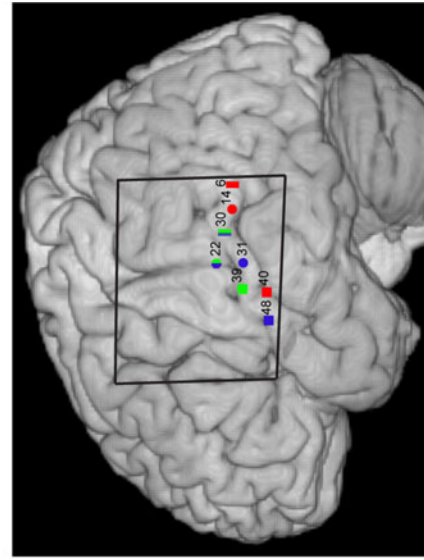
Parietal lobe On the post-central sulcus just above the sylvian fissure, all three conditions elicited MMNs. These onset earlier and were of longer duration for the tone and non-phoneme control condition, with the speech condition differing significantly from the tone condition (channel 22, $y = -24$).

Discussion

Speech is arguably the most complex acoustic signal that the auditory system decodes. Models of auditory processing, which for the most part assume hierarchical organization, would thus place the extraction of speech from the acoustic signal at the latest temporal stages of auditory information processing (for this temporal unit of the speech signal) and at auditory cortical regions peripheral to primary auditory cortex. The present data are consistent with this view, with speech information selectively processed at zones of the STG more distal from Heschl's gyrus along the sylvian fissure. Further, these data lend support to a dual stream model of decoding and representation of the speech signal. Specifically, the aSTG and pSTG/SMG speech sensitive regions correspond well with what have been referred to as the ventral and dorsal streams of auditory processing (see e.g. Rauschecker 1998, 2012; Hickok and Poeppel 2007; Peelle et al. 2010; Chevillet et al. 2011; Leavitt et al. 2011). According to this division of auditory processing, the ventral stream, moving lateral and anterior from primary auditory cortex, is biased toward sound identification, and the dorsal stream, moving first posterior from primary auditory cortex



Participant: CL



◀ **Fig. 4** Individual data for participant CL. Evoked responses to the standard (*black trace*) and deviant (*colored trace*) stimuli are shown for electrodes that exhibited significant MMN responses. Periods of significant difference are indicated by the *color strip* below the waveforms. The period of 100–300 ms post-stimulus onset is highlighted by a *gray box*. In the lower right corner, electrode positions are illustrated on the MNI template brain. Electrodes are shape and number coded to link the specific waveform data with the location. *Color coding* indicates the designation of the channel: *red shapes* indicate electrodes where the MMN was greatest in response to speech stimulation; *blue shapes* electrodes most responsive to complex stimulation; and *green shapes* electrodes most responsive to simple tones. Double classifications are indicated by *bi-colored shapes*. The *black square* on the brain shows the borders of the electrode grid. Below each column of waveform responses, a *bar graph* quantifies stimulus-specific MMN responses for each electrode site, by showing the integrated amplitude of the statistically significant effect (to therefore represent both its amplitude and duration), within the 100–300 ms post-stimulus onset window that was considered in our analyses. Values are scaled with respect to the largest at a given contact (which is set at 100 %). Statistical differences between MMNs to the different stimulus types are indicated by *asterisks*

and subsequently involving parietal and frontal regions as well, is biased toward sound localization, and the translation of action-sounds to their motor representation counterparts. In the case of speech, this involves the link between speech-sound and speech-articulation representations.

Anatomical localization of speech-specific responses

How do the current data bear on contrasting views of the anatomical representation of prelexical speech sounds, which have been largely informed by results from hemodynamic imaging of brain activity (fMRI and PET)? A number of studies indicate a primary role for middle to anterior portions of the STG in the representation of prelexical speech, a notion that the current data are clearly consistent with. In Obleser et al. (2010), BOLD responses to different consonant–vowel speech syllables obtained during a passive listening task were subjected to a multi-voxel pattern classification analysis using a support vector machine approach (Obleser et al. 2010). Simple contrast analysis resulted in speech-specific responses localized to lateral, middle, and anterior regions of the STG. Providing greater specificity, the best classification was found for left anterior STG, with the highest levels of discriminant classification among consonant–vowel speech syllables. Liebenthal and colleagues have similarly found support for anterior and middle portions of the STS/STG being favorably responsive to phonemes. In one study they presented syllable pairs and spectro-temporally matched control pairs to participants while they performed a two-alternative forced choice task (Liebenthal et al. 2005). Contrasting the hemodynamic response between the

phonetic and spectro-temporally matched non-phonetic condition resulted in a greater response to phonemes in the anterior and middle regions of the STG (see also Obleser and Eisner 2009; Liebenthal et al. 2010). Both groups argue that it is the middle to anterior portions of the STG that is primary in representing prelexical phoneme information (Liebenthal et al. 2005; Obleser and Eisner 2009; Obleser et al. 2010). A meta-analysis of 115 imaging studies performed to map the centroid of activation in response to phonemes, spoken-words, and phrase length utterances (DeWitt and Rauschecker 2012), lends further support to this view. This analysis revealed a concentration of preferential responses to speech sounds that localized to the left mid to anterior region of the STG. Additional evidence comes from a smaller meta-analysis of data in which mSTG was found to most reliably respond to speech versus non-speech control stimulation (Turkeltaub and Coslett 2010). Thus the present data, in which three of three cases had contacts in the middle/anterior portion of STG that showed speech-specific MMNs, accord with ample evidence in the literature that the middle to anterior portions of the STG, part of the ventral stream “what” pathway, play a key role in the decoding of the speech signal (e.g. Chevillet et al. 2011).

The present data also support a role, however, for the auditory dorsal stream in the processing of speech sounds, with speech-MMNs observed in this region (pSTG and SMG) in two of the three participants. The role of the auditory dorsal stream in the representation of speech has been the focus of previous intracranial EEG investigations (Chang et al. 2010; Flinker et al. 2011), which, with the use of microgrids of contacts placed over pSTS, have provided support for its key role by showing differential or discriminant responding as a function of phoneme category versus the acoustic content of the stimuli. It is important to note that in Chang et al. (2010) and Flinker et al. (2011), electrodes were not placed over anterior STG and therefore the results cannot speak to the relative optimization of anterior and posterior STG with respect to phoneme processing. In an fMRI study, Okada et al. (2010) presented participants with two types of intelligible speech stimuli (clear and noise-coded sentences), and spectro-temporally matched non-speech control stimuli, and participants indicated if the stimuli were intelligible speech. Similar to the findings of the present study, the data showed both anterior and posterior regions of the STG to be sensitive to speech compared to control stimuli. Using multivariate pattern analyses, the study further revealed that posterior regions showed greater insensitivity to acoustic differences between the two forms of intelligible speech as compared to other regions of interest (e.g., aSTS, mSTS, and core auditory regions). From this they concluded that it is pSTS that supports abstracted phonological representations.

Table 2 List of relevant electrode sites (ROI), MNI coordinates, approximate Brodmann area, and onset latency of MMN and AEP responses

Participant	Channel	MNI coordinates	Brodmann area and closest gyrus	MMN onset latency (ms)			Standard stimulus AEP onset latency (ms)		
				Phone	Non-ph	Tone	Phone	Non-ph	Tone
AA	36	-63/-5/-3	22, superior temporal gyrus	165	175	250	75	80	70
	44	-67/-12/-11	21, middle temporal gyrus	-	210	-	80	80	80
	45	-69/-19/-3	21, middle temporal gyrus	175	135	-	100	100	125
	46	-71/-27/2	21, middle temporal gyrus	160	200	-	40	35	35
	47	-65/-34/11	22, superior temporal gyrus	210	125	200	65	55	60
	27	-58/9/-4	22, superior temporal gyrus	190	200	-	280	200	150
	40	-66/-32/29	40, supramarginal gyrus	135	-	-	60	50	55
	28	-60/3/3	44, inferior frontal gyrus	-	125	150	30	30	30
	31	-64/-14/24	1, post-central gyrus	-	210	205	110	105	90
TK	23	-63/-8/-9	21, middle temporal gyrus	140	-	-	45	45	50
	22	-63/-14/-3	21, middle temporal gyrus	90	135	-	140	150	80
	31	-61/-3/0	22, superior temporal gyrus	125	150	-	120	75	70
	38	-65/-4/14	4, precentral gyrus	-	155	155	40	35	40
	37	-64/-11/22	1, post-central gyrus	-	200	-	40	40	60
	30	-64/-11/6	41, superior temporal gyrus	100	125	165	40	40	40
	29	-64/-19/13	1, post-central gyrus	-	280	100	40	35	35
	21	-67/-24/3	22, superior temporal gyrus	-	-	80	105	110	60
	28	-67/-26/19	40, supramarginal gyrus	280	-	100	50	50	40
	20	-67/-30/9	22, superior temporal gyrus	95	150	125	60	60	50
CL	19	-62/-40/17	22, superior temporal gyrus	-	90	220	50	50	55
	18	-64/-48/23	39, angular gyrus	-	155	110	50	55	55
	40	-67/-13/-1	21, middle temporal gyrus	120	-	-	55	55	50
	14	-67/-41/11	22, superior temporal gyrus	200	-	-	65	65	65
	6	-64/-53/10	39, angular gyrus	155	-	-	80	100	90
	48	-62/-4/-1	22, superior temporal gyrus	115	225	150	110	110	90
	31	-66/-24/8	41, superior temporal gyrus	-	110	-	60	60	60
	39	-63/-14/8	41, superior temporal gyrus	-	-	130	55	50	50
	22	-67/-24/19	40, supramarginal gyrus	175	145	125	75	85	75
	30	-67/-32/15	22, superior temporal gyrus	-	-	165	50	50	45

Careful consideration of their results (Fig. 5) suggests, however, that classification was highly similar for both anterior and posterior regions (see DeWitt & Rauschecker, supplementary material, for a similar conclusion).

A study by Liebenthal et al. (2005) also identified a role for pSTS, but their data argue for a more general role in which this region is sensitive to spectro-temporally complex stimuli, but does not necessarily show a differential response to long-term speech-sound representations (see also Obleser et al. 2007 for a similar conclusion). In their 2005 study, they found that posterior regions of the STG/STS responded similarly to phoneme and spectro-temporally matched non-phoneme stimuli, whereas in a later study (Liebenthal et al. 2010), when participants were trained to treat spectro-temporally matched non-speech stimuli categorically, their corresponding representations increased in

pSTS. This led the authors to suggest that pSTS is involved in the representation of newly acquired sound categories that are not yet relegated to long-term representations.

The work of Liebenthal and colleagues, and the studies reviewed above that favor speech-specific processing of pre-lexical speech along the a/mSTG, raise questions regarding the engagement of the dorsal pathway of auditory processing in decoding the speech signal. Yet in the present data we find evidence that in two of three participants the dorsal stream (pSTG, SMG) was preferentially engaged by pre-lexical speech sounds over well-matched non-speech sounds, even when no speech-related task was involved. Indeed, in a scalp electrophysiological study, visual speech inputs were found to influence auditory speech representations (Saint-Amour et al. 2007). Since the pSTG, and thus the dorsal pathway, is strongly implicated in the multisensory representation of

audio-visual inputs (Calvert and Thesen 2004; Beauchamp 2005; Saint-Amour et al. 2007), this suggests the engagement of the dorsal pathway in the decoding of the speech signal. The present data, taken in the context of the larger literature on the representation of phonetic information in auditory cortex, suggest that, while automatic engagement of the ventral auditory pathway for the selective processing of phonemes appears to be the more robust process, speech stimuli can automatically engage speech-specific processing in the dorsal auditory pathway as well. Ventral and dorsal auditory regions may represent different stages of complexity/abstraction in an information processing hierarchy, and/or streams of auditory processing that are optimized for different functions. For example, Liebenthal et al. (2010) found pSTS to be involved in the representation of newly learned categorical sounds, and more anterior regions in the representation of well-learned categorical sounds. Further, there is indication that the dorsal stream is more engaged during speech processing when the stimuli are difficult to hear, when the signal must be decoded in a noisy environment (Wise et al. 2001; Hickok and Poeppel 2007; Osnes et al. 2011). Notably, the present study was conducted in a hospital room, an often noisy auditory environment. However, it is arguably no more noisy than many of the listening conditions that one typically listens to speech. The differences across studies in activation of the auditory dorsal stream by speech might well also reflect that processing regions therein are particularly sensitive to factors that can vary between individuals (familiarity/expertise with American English, SNR of the syllables). For example, activity in left pSTG has been shown to vary with individual categorization performance (Leech et al. 2009; Desai et al. 2008). And the same has been found for the left SMG, further downstream (Raizada 2007; Desai et al. 2008). This, however, is an idea that requires testing in a dedicated study with a larger sample.

The timing of bottom-up activation of speech-sensitive regions of auditory cortex

Most studies on the anatomy of the neural representation of speech have relied on brain hemodynamics, and thus have lacked the temporal resolution to distinguish whether the timing of activations is more consonant with perceptual or post-perceptual processes. Here we found speech-specific MMNs to onset between 90 and 200 ms over both anterior/medial and posterior regions, providing a conservative ceiling on when speech-specific information was preferentially processed in the temporal information processing hierarchy. This is consistent with electrophysiological studies that have found neural responses to differentiate speech sounds on the place-of-articulation dimension at the peak of the sensory response, at about 110 ms (Chang et al. 2010), or slightly later, at about 150 ms for voicing differences (Fishman and Steinschneider 2010).

A surprising observation is that there was no clear evidence in the onset of the MMN response for a systematic temporal hierarchy in processing of increasingly complex auditory information. While mean onset latencies across contacts and individuals (using only channels with a single designation) conformed to a possible pattern in which tones are processed first, speech next, and similarly complex but less familiar stimuli later still (133 ms for the tones, 147 ms for the phoneme condition, and 163 ms for the spectro-temporally matched non-phoneme condition), at the individual participant level this pattern was only observed in two of the participants (CL and TK). In contrast, as can be assessed from Table 2, the timing of the onset of the AEPs to the tone stimuli as a function of location along the sylvian fissure (at least considering the ventral stream, along the STG from post-central sulcus to regions more anterior) was generally consistent with the anatomical hierarchy of auditory processing that has been reported in the literature. For TK, moving anterior to posterior, channels 31, 30, and 29 had onset latencies of 70, 40, and 35 ms, respectively. For CL, channels 48, 39, and 31 had latencies of 90, 50, and 60 ms, respectively. While the latter latency breaks the predicted pattern, the difference of 10 ms may be in the noise. For AA, channels 27 and 36 had onset latencies of 150 and 70 ms.

Caveats

There are limitations that must be taken into account in considering the present data: (1) these data do not speak to the lateralization of language, as only left hemisphere patients are reported. Two right hemisphere patients had sparse coverage and too few contacts with MMNs to provide a meaningful picture of the representation of speech. (2) These data come from compromised participants with epilepsy and may not be representative of the population as a whole. However, the data do fit well with the total of the neuroimaging data and thus make a compelling case for generalization to the larger healthy population. (3) Even though localization is far superior to scalp recorded data, it is delimited by the spatial sampling of 1 cm, and by clinically determined placement of the electrodes. Thus, these data can only provide a partial picture of the cortical regions that are involved in the processing and representation of prelexical speech. (4) The data are from a sample of three.

Acknowledgments The authors wish to express their sincere appreciation to the participants who kindly donated their time to this project, to Mr. Jason Adler for his contribution to the analysis of data, and to Ms. Edel Flynn for assistance with manuscript formatting. SM and JJF received support from NIMH (MH 85322) and a grant from the Nathan Ganchar Foundation, and MM from the Swiss National Science Foundation (PBELP3-123067).

References

- Abdul-Kareem IA, Sluming V (2008) Heschl gyrus and its included primary auditory cortex: structural MRI studies in healthy and diseased subjects. *J Magn Reson Imaging* 28(2):287–299
- Allison T, Puce A et al (1999) Electrophysiological studies of human face perception. I: potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cereb Cortex* 9(5):415–430
- Amenedo E, Escera C (2000) The accuracy of sound duration representation in the human brain determines the accuracy of behavioral perception. *Eur J Neurosci* 12(7):2570–2574
- Beauchamp MS (2005) See me, hear me, touch me: multisensory integration in lateral occipital-temporal cortex. *Curr Opin Neurobiol* 15(2):145–153
- Butler JS, Molholm S et al (2011) Common or redundant neural circuits for duration processing across audition and touch. *J Neurosci* 31(9):3400–3406
- Calvert GA, Thesen T (2004) Multisensory integration: methodological approaches and emerging principles in the human brain. *J Physiol Paris* 98(1–3):191–205
- Chang EF, Rieger JW, Johnson K, Berger MS, Barbaro NM, Knight RT (2010) Categorical speech representation in human superior temporal gyrus. *Nat Neurosci* 13(11):1428–1432
- Chevillet M, Riesenhuber M et al (2011) Functional correlates of the anterolateral processing hierarchy in human auditory cortex. *J Neurosci* 31(25):9345–9352
- De Sanctis P, Molholm S et al (2009) Right hemispheric contributions to fine auditory temporal discriminations: high-density electrical mapping of the duration mismatch negativity (MMN). *Front Integr Neurosci* 3:5
- Desai R, Liebenthal E et al (2005) Volumetric vs. surface-based alignment for localization of auditory cortex activation. *Neuroimage* 26(4):1019–1029
- Desai R, Liebenthal E et al (2008) Left posterior temporal regions are sensitive to auditory categorization. *J Cogn Neurosci* 20(7):1174–1188
- DeWitt I, Rauschecker JP (2012) Phoneme and word recognition in the auditory ventral stream. *Proc Natl Acad Sci USA* 109(8):E505–E514
- Epstein R, Kanwisher N (1998) A cortical representation of the local visual environment. *Nature* 392(6676):598–601
- Fedorenko E, Kanwisher N (2011) Some regions within Broca's area do respond more strongly to sentences than to linguistically degraded stimuli: a comment on Rogalsky and Hickok (2011). *J Cogn Neurosci* 23(10):2632–2635
- Felleman DJ, Van Essen DC (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex* 1(1):1–47
- Fishman YI, Steinschneider M (2010) Neural correlates of auditory scene analysis based on inharmonicity in monkey primary auditory cortex. *J Neurosci* 30(37):12480–12494
- Flinker A, Chang EF et al (2011) Sub-centimeter language organization in the human temporal lobe. *Brain Lang* 117(3):103–109
- Guthrie D, Buchwald JS (1991) Significance testing of difference potentials. *Psychophysiology* 28(2):240–244
- Hackett TA (2011) Information flow in the auditory cortical network. *Hear Res* 271(1–2):133–146
- Hasson U, Levy I et al (2002) Eccentricity bias as an organizing principle for human high-order object areas. *Neuron* 34(3):479–490
- Hickok G, Poeppel D (2007) The cortical organization of speech processing. *Nat Rev Neurosci* 8(5):393–402
- Ide A, Rodriguez E et al (1996) Bifurcation patterns in the human sylvian fissure: hemispheric and sex differences. *Cereb Cortex* 6(5):717–725
- Kang X, Bertrand O et al (2004) Local landmark-based mapping of human auditory cortex. *Neuroimage* 22(4):1657–1670
- Kanwisher N, McDermott J et al (1997) The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci* 17(11):4302–4311
- Leavitt VM, Molholm S et al (2011) “What” and “where” in auditory sensory processing: a high-density electrical mapping study of distinct neural processes underlying sound object recognition and sound localization. *Front Integr Neurosci* 5:23
- Leech R, Holt LL et al (2009) Expertise with artificial nonspeech sounds recruits speech-sensitive cortical regions. *J Neurosci* 29(16):5234–5239
- Liebenthal E, Binder JR et al (2005) Neural substrates of phonemic perception. *Cereb Cortex* 15(10):1621–1631
- Liebenthal E, Desai R et al (2010) Specialization along the left superior temporal sulcus for auditory categorization. *Cereb Cortex* 20(12):2958–2970
- Lucan JN, Foxe JJ et al (2010) Tactile shape discrimination recruits human lateral occipital complex during early perceptual processing. *Hum Brain Mapp* 31(11):1813–1821
- McKeeff TJ, McGugin RW et al (2010) Expertise increases the functional overlap between face and object perception. *Cognition* 117(3):355–360
- Molholm S, Martinez A et al (2005) The neural circuitry of pre-attentive auditory change-detection: an fMRI study of pitch and duration mismatch negativity generators. *Cereb Cortex* 15(5):545–551
- Näätänen R (1992) Attention and brain function. Lawrence Erlbaum Associates, NJ, Hillsdale
- Naatanen R, Picton T (1987) The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology* 24(4):375–425
- Näätänen R, Paavilainen P et al (2007) The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin Neurophysiol* 118(12):2544–2590
- Obleser J, Eisner F (2009) Pre-lexical abstraction of speech in the auditory cortex. *Trends Cogn Sci* 13(1):14–19
- Obleser J, Zimmerman J et al (2007) Multiple stages of auditory speech perception reflected in event-related fMRI. *Cereb Cortex* 17(10):2251–2257
- Obleser J, Leaver AM et al (2010) Segregation of vowels and consonants in human auditory cortex: evidence for distributed hierarchical organization. *Front Psychol* 1:232
- Okada K, Rong F et al (2010) Hierarchical organization of human auditory cortex: evidence from acoustic invariance in the response to intelligible speech. *Cereb Cortex* 20(10):2486–2495
- Oostenveld R, Fries P et al (2011) FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci* 2011:156869
- Orlov T, Makin TR et al (2010) Topographic representation of the human body in the occipitotemporal cortex. *Neuron* 68(3):586–600
- Osnes B, Hugdahl K et al (2011) Increased activation in superior temporal gyri as a function of increment in phonetic features. *Brain Lang* 116(2):97–101
- Peelle JE, Johnsrude IS et al (2010) Hierarchical processing for speech in human auditory cortex and beyond. *Front Hum Neurosci* 4:51
- Picton TW, Alain C et al (2000) Mismatch negativity: different water in the same river. *Audiol Neurootol* 5(3–4):111–139
- Pourtois G, Peelen MV et al (2007) Direct intracranial recording of body-selective responses in human extrastriate visual cortex. *Neuropsychologia* 45(11):2621–2625
- Pourtois G, Spinelli L et al (2010) Modulation of face processing by emotional expression and gaze direction during intracranial recordings in right fusiform cortex. *J Cogn Neurosci* 22(9):2086–2107

- Rademacher J, Burgel U et al (2002) Stereotaxic localization, intersubject variability, and interhemispheric differences of the human auditory thalamocortical system. *Neuroimage* 17(1):142–160
- Raizada RD, Poldrack RA (2007) Selective amplification of stimulus differences during categorical processing of speech. *Neuron* 56(4):726–740
- Rauschecker JP (1998) Cortical processing of complex sounds. *Curr Opin Neurobiol* 8(4):516–521
- Rauschecker JP (2012) Ventral and dorsal streams in the evolution of speech and language. *Front Evolut Neurosci* 4:7
- Ritter W, Paavilainen P et al (1992) Event-related potentials to repetition and change of auditory stimuli. *Electroencephalogr Clin Neurophysiol* 83(5):306–321
- Roland PE, Zilles K (1998) Structural divisions and functional fields in the human cerebral cortex. *Brain Res Brain Res Rev* 26(2–3):87–105
- Saint-Amour D, De Sanctis P et al (2007) Seeing voices: high-density electrical mapping and source-analysis of the multisensory mismatch negativity evoked during the McGurk illusion. *Neuropsychologia* 45(3):587–597
- Scott SK, Johnsrude IS (2003) The neuroanatomical and functional organization of speech perception. *Trends Neurosci* 26(2):100–107
- Sehatpour P, Molholm S et al (2008) A human intracranial study of long-range oscillatory coherence across a frontal-occipital-hippocampal brain network during visual object processing. *Proc Natl Acad Sci USA* 105(11):4399–4404
- Steinschneider M, Nourski KV et al (2011) Intracranial study of speech-elicited activity on the human posterolateral superior temporal gyrus. *Cereb Cortex* 21(10):2332–2347
- Turkeltaub PE, Coslett HB (2010) Localization of sublexical speech perception components. *Brain Lang* 114(1):1–15
- Wada J, Rasmussen T (1960) Intracarotid injection of sodium amytal for the lateralization of cerebral speech dominance. *J Neurosurg* 106(6):1117–1133
- Wise RJ, Scott SK et al (2001) Separate neural subsystems within ‘Wernicke’s area’. *Brain* 124(Pt 1):83–95