Coding of repetitive transients by auditory cortex on posterolateral superior temporal gyrus in humans: An intracranial electrophysiology study

Kirill V. Nourski¹, John F. Brugge¹,², Richard A. Reale¹,², Christopher K. Kovach¹, Hiroyuki Oya¹, Hiroto Kawasaki¹, Rick L. Jenison², Matthew A. Howard III¹

¹Department of Neurosurgery, The University of Iowa, Iowa City, IA 52242, USA
²Department of Psychology, University of Wisconsin – Madison, Madison, WI 53706, USA

Running head: Human auditory cortex

Contact:
Kirill V. Nourski
Department of Neurosurgery
The University of Iowa
200 Hawkins Dr. 1815 JCP
Iowa City, IA 52242
Phone: (319) 335-7049
E-Mail: kirill-nourski@uiowa.edu

Key words: averaged evoked potential, click train, electrocorticography, high gamma, phase locking
ABSTRACT

Evidence regarding the functional subdivisions of human auditory cortex has been slow to converge on a definite model. In part this reflects inadequacies of current understanding of how cortex represents temporal information in acoustic signals. To address this, we investigated spatiotemporal properties of auditory responses in human posterolateral superior temporal gyrus (PLST) to acoustic click train stimuli using intracranial recordings from neurosurgical patients. Subjects were patients undergoing chronic invasive monitoring for refractory epilepsy. The subjects listened passively to acoustic click train stimuli of varying durations (160 or 1000 ms) and rates (4-200 Hz), delivered diotically via insert earphones. Multicontact subdural grids placed over the perisylvian cortex recorded intracranial electrocorticographic responses from PLST and surrounding areas. Analyses focused on averaged evoked potentials (AEPs) and high gamma (70-150 Hz) event-related band power (ERBP).

Responses to click trains featured prominent AEP waveforms and increases in ERBP. The magnitude of AEPs and ERBP typically increased with click rate. Superimposed on the AEPs were frequency-following responses (FFRs), most prominent at 50 Hz click rates, but still detectable at stimulus rates of up to 200 Hz. Loci with the largest high gamma responses on PLST were often different from those sites that exhibited strongest FFRs. The data indicate that responses of non-core auditory cortex of PLST represent temporal stimulus features in multiple ways. These include an isomorphic representation of periodicity (as measured by the FFR), a representation based on increases in non-phase-locked activity (as measured by high gamma ERBP), and spatially distributed patterns of activity.
Acoustic transients are richly represented in our environment, and the auditory system has evolved mechanisms for detecting these signals and encoding the information contained in them. These transients are brief signals, often referred to as ‘clicks’, and when they are repeated at equal intervals, the resulting percept is related to the rate at which they are presented. Clicks in a train presented at relatively low rates (<8-10 Hz) are perceived as individual events, whereas at progressively higher rates the percept evolves into one of acoustic ‘flutter’ and, eventually (>30 Hz), one of pitch. In monkeys (Lu et al., 2001; Liang et al., 2002) and other mammals (e.g. cats; Imaizumi et al., 2010) neurons in AI (primary auditory cortex) and immediately adjacent primary-like fields (e.g., fields R and RT in monkeys) employ dual mechanisms to encode such acoustic transients: an isomorphic representation by precise timing of discharges at low repetition rates and a non-isomorphic (abstracted) representation by changes in overall firing rate at faster repetition rates (Wang et al., 2003; Wang, 2007). Evidence from direct auditory cortical recording in humans suggests that similar mechanisms may be operating there as well (Brugge et al., 2009; Nourski & Brugge, 2011).

In the classical sense, auditory cortex in humans, as in other primates studied so far, comprises multiple areas located on the superior temporal gyrus (STG; Hackett, 2011). The primary and primary-like cortex of the auditory core occupies the posteromedial aspect of Heschl’s gyrus (HG). This cortical complex is surrounded by multiple anatomically distinct auditory subdivisions on anterolateral HG, planum polare, planum temporale and lateral surface of the STG (Hutsler & Gazzaniga, 1996; Rivier & Clarke, 1997; Nakahara et al., 2000; Wallace et al., 2002; Chiry et al., 2003). Based largely on evidence derived from anatomical and
physiological studies in monkey, a functional model of auditory cortex posits multiple fields interconnected in such a way as to provide for hierarchical processing of acoustic information (Kaas & Hackett, 2005; Hackett, 2011). This hierarchical model also provides a useful framework for investigating the functional organization of human auditory cortex (Rauschecker & Scott, 2009). However, with the exception of core cortex, the cross-species homologies of auditory cortical fields are still unclear (Hackett et al. 2001; Hackett, 2003, 2007, 2008, 2011; Sweet et al. 2005; Fullerton & Pandya, 2007; but see Chevillet et al., 2011).

Research performed in our laboratory employs direct intracranial recording in neurosurgical patients to study the functional organization of human auditory cortex. This method provides a unique opportunity to study multiple auditory cortical fields simultaneously and with high spatial and temporal resolution (Howard et al., 2012; Mukamel & Fried, 2012). We have previously shown that local field potentials recorded from the posteromedial aspect of HG - the putative human homolog of the core field - exhibit robust temporal locking to click-train stimuli (Brugge et al., 2008, 2009). The response is characterized by large-amplitude, short-latency averaged evoked potentials (AEPs), increases in event-related band power (ERBP) and a frequency-following response (FFR). The phase-locked FFR is most prominent when elicited by click rates of about 50-100 Hz and below but may be reliably detected at rates as high as 200 Hz (Brugge et al., 2009). The morphology of the evoked waveforms and the range of phase locking are very similar to those response properties exhibited by the auditory core of macaque monkeys studied under similar stimulus conditions (Steinschneider et al., 1998).

In humans, processing of temporal information appears to engage non-core auditory fields as well as the auditory core (Kuwada et al., 1986; Rees et al., 1986; Giraud et al., 2000; Harms & Melcher, 2002, Ross et al., 2000). An area adjacent to the putative core, on the anterolateral
aspect of HG, has been shown to exhibit low-amplitude, long-latency AEPs and ERBP responses with the FFR weak or absent (Liégeois-Chauvel et al., 1991; Brugge et al. 2008, 2009). It was suggested that these recordings were obtained from one of the numerous putative auditory cortical belt areas in humans that have been identified anatomically. Previous work in our laboratory has also identified an acoustically responsive area on the posterior portion of the lateral STG, functionally distinct from auditory areas on HG. We refer to this area as the posterolateral superior temporal auditory area (PLST; Howard et al., 2000). This area, which may comprise more than one functional field, was found to respond robustly to a wide range of stimuli, from tone and noise bursts to speech utterances (Howard et al., 2000; Brugge et al., 2008; Steinschneider et al., 2011), the latter being influenced by the visual presentation of congruent face and lip movements (Reale et al., 2007). Electrical stimulation tract tracing studies further revealed a short latency functional connection between PLST and the auditory core on HG (Brugge et al., 2003, 2005).

We also noted that certain sites in PLST exhibited robust phase locking to clicks, and have presented some of these results in abstract form (Nourski et al., 2011). Accordingly, we hypothesized that non-core auditory fields on the lateral STG would, like the auditory core, also exhibit isomorphic and non-isomorphic representations of temporal acoustic events, albeit on different time scales. Here we present details of these studies, comparing response dynamics observed within PLST to those exhibited by the auditory core and characterizing their spatial distribution. In doing so, we provide a framework for future studies aimed at establishing the role of explicit cortical representation of temporal information in the higher functions of speech perception and comprehension.
METHODS

Subjects
Experimental subjects were 16 neurosurgical patients (10 male, 6 female, age 20-56 years, median age 34.5 years). The subjects had been diagnosed with medically refractory epilepsy and were undergoing chronic invasive electrocorticogram (ECoG) monitoring to identify seizure foci prior to resection surgery. Written informed consent was obtained from each subject. Research protocols were approved by The University of Iowa Institutional Review Board.

Thirteen subjects reported themselves to be right-handed, two (L147 and R180) to be left-handed, and one (L151) to be ambidextrous. All subjects but one had left hemisphere language dominance, as determined by intracarotid amytal (Wada) test results; subject L162 had bilateral language representation. The placement of the electrode arrays was based on clinical considerations, and for that reason we were unable to record from both the left and right hemisphere in the same subject. In 7 subjects, the electrodes were implanted on the left side, while in 9 others the recordings were obtained from the right hemisphere. The side of implantation is indicated by the letter prefix of the subject code (L for left, R for right). All subjects underwent audiometric and neuropsychological evaluation before the study, and none were found to have hearing or cognitive deficits that might impact the findings presented in this study. All subjects were native English language speakers. Clinical analysis of intracranial recordings indicated that the auditory cortical areas on the STG were not involved in generation of epileptic activity in any of the subjects included in this study.

Each subject underwent whole-brain high-resolution T1-weighted structural MRI (resolution 0.78×0.78 mm, slice thickness 1.0 mm, average of 2) scans before and after electrode
implantation in order to determine recording contact locations relative to the pre-operative brain images. Pre-implantation MRIs and post-implantation thin-sliced volumetric CT scans (in-plane resolution 0.51×0.51 mm, slice thickness 1.0 mm) were co-registered using a 3D linear registration algorithm (FMRIB Linear Image Registration Tool; Jenkinson et al., 2002). Coordinates for each electrode contact obtained from post-implantation CT volumes were transferred to pre-implantation MRI volumes. Results were compared to intraoperative photographs to ensure reconstruction accuracy.

Experiments were performed in a dedicated electrically shielded suite located within the Clinical Research Unit of the University of Iowa Institute for Clinical and Translational Science. The subjects were reclining in a hospital bed or an armchair and were awake, but not actively attending to the stimuli.

**Stimuli**

Experimental stimuli were trains of acoustic clicks, digitally generated as equally-spaced rectangular pulses (0.2 ms duration) and delivered to both ears via insert earphones (ER4B, Etymotic Research, Elk Grove Village, IL) that were integrated into custom-fit earmolds. The stimuli were presented at a comfortable level, typically around 50 dB above hearing threshold.

Two stimulus configurations were used: short-duration (160 ms) click trains of 25, 50, 100, 125, 150 and 200 Hz and long-duration (1 s) click trains of 4, 8, 16, 32, 64, and 128 Hz. In 7 early experiments (R127, R129, L130, L138, R139, L140, R142), each stimulus of a particular click rate was presented 100 times after which the next click rate was chosen, and the procedure repeated until stimulus sets representing the six different click rates had been delivered. In later experiments (L147, L151, R152, R153, R154, L162, L178, R180, R186), click train stimuli of
the six rates were each presented 50 times in random order. In all experiments, the inter-train
interval was chosen randomly within a Gaussian distribution (mean interval 2 s; SD = 10 ms) to
reduce stimulus predictability and to allow more efficient AEP estimation. Stimulus delivery and
data acquisition were controlled by a TDT RP2.1 and RX5 or RZ2 real-time processor (Tucker-
Davis Technologies, Alachua, FL).

Recording

Details of electrode implantation and data collection have been described previously (Howard et
al., 2000; Brugge et al., 2008, Howard et al., 2012). In brief, filtered (1.6–1000 Hz bandpass, 12
dB/octave rolloff) and amplified (20×) ECoG data were digitally recorded (sampling rate 2034.5
Hz) from multicontact subdural grid electrodes (AdTech, Racine, WI) placed over perisylvian
cortex including the STG. The recording arrays consisted of 96 platinum-iridium disc electrodes
(2.3 mm exposed diameter, 5 mm inter-electrode distance) arranged in an 8×12 grid and
embedded in a silicon membrane. A subgaleal contact was used as a reference. Simultaneous
recordings were obtained from hybrid depth electrodes stereotactically implanted into the HG
roughly parallel to its long axis (Howard et al., 1996; Reddy et al., 2010). Data obtained from
those recordings have been reported in detail previously (Brugge et al., 2009). Recording
electrodes remained in place up to two weeks under the direction of clinical epileptologists.

Analysis

ECoG data obtained from each recording site was analyzed as the AEP and, in the time-
frequency plane, as ERBP. Data analysis was performed using custom software written in
MATLAB Version 7.13.0 programming environment (MathWorks, Natick, MA, USA). Pre-
processing of ECoG data included downsampling to 1 kHz, followed by removal of power line
noise by an adaptive notch filtering procedure. To identify peaks in the power spectrum
associated with line noise, the log Fourier power spectrum was normalized by a smoothed
baseline value. The baseline was obtained from the interpolation of an eighth-order polynomial
fitted to the log-spectrum by least-squares. Peaks in the baseline-normalized signal were
extracted using a threshold criterion. For each fundamental frequency among extracted peaks, the
entire data sequence was divided into overlapping windows of approximately 4 s length adjusted
to match as nearly as possible an integer multiple of the fundamental period. The purpose of the
latter step is to minimize spectral leakage from peaks at the targeted line noise frequency. Fourier
components at the fundamental and harmonic frequencies were then removed from the data
within the overlapping windows. Denoised data were reconstructed by applying a weighted
average, $w$, between overlapping windows, where the $k^{th}$ point in the window was weighted by
$$w[k] = \frac{1 - \cos\left(\frac{2\pi k}{N}\right)}{2},$$
where $N$ is the window length.

Additionally, single-trial (peri-stimulus) ECoG waveforms with voltage peaks or troughs
greater than 2.5 standard deviations from the mean were eliminated from the data set prior to
further analyses. This would include sporadic activity generated by electrical interference,
epileptiform spikes, high-amplitude slow-wave activity, or movement artifacts.

Single-trial waveforms obtained on the 96 cortical sites were transformed at every time
point with a spatial filter using the surface Laplacian operation (Nunez, 1981; Nunez & Pilgreen,
1991; Reale et al., 2007). The surface Laplacian is independent of the reference electrode and
reduces the effects of spatial smearing of ECoG voltage due to volume conduction in the tissue
and fluid of the brain. The surface Laplacian required an accurate representation of the spatial
distribution of potential, which was estimated using a two-dimensional thin-plate spline interpolant with clamped edges (Perrin et al., 1987; Law et al., 1993). We have shown previously that this procedure yields a noticeable increase in the spatial specificity of the recordings from PLST as compared to unprocessed data (Reale et al., 2007).

The magnitude of the AEP was characterized by the root-mean-square (RMS) amplitudes within the time interval of 300 ms after stimulus onset. Phase-locked responses to the periodicity of the click trains were visualized as the FFR by high-pass filtering AEP waveforms with a cutoff of one octave below the driving frequency using a 4th order Butterworth filter (24 dB/octave slope with zero-phase shift). FFRs were quantified using the phase-locking value (PLV) metric (Jervis et al., 1983), calculated from single-trial ECoG signals within a window of 50-250 ms after stimulus onset, tapered with a Tukey function. The PLV is an amplitude-independent metric of response synchrony and thus has the potential for accurate separation of phase-locked response components from non-phase-locked activity (Chavez et al., 2006). PLVs are often computed in the time domain using bandpass-filtered and Hilbert-transformed analytic signals (Lachaux et al., 1999). We computed PLVs in the frequency domain using the FFT of the 200 ms window. Significance of PLVs was evaluated by applying the likelihood ratio test (Pawitan, 2001) to numerically evaluated maximum likelihood fits of a von Mises distribution, wherein the null hypothesis was uniform distribution of phase angles over trials. This procedure is equivalent to the Rayleigh test (Mardia & Jupp, 2000). Correction for multiple comparisons was done by controlling false discovery rate (Benjamini & Hochberg, 1995; Benjamini et al., 2001) at $q=0.01$. Phase locking to 25 Hz often failed to reach statistical significance in our data sets. This negative finding can be attributed to the number of clicks (5) in the train in this stimulus condition, which may have been too small to allow a measurable FFR to develop. In
addition, the FFR at 25 Hz was sometimes difficult to separate from the onset AEP waveform, as their time course and dominant spectral features overlapped. As a result, significant phase-locking values at frequencies close to 25 Hz were sometimes observed in the absence of a 25 Hz driving stimulus.

Time-frequency analysis of the ECoG was performed using wavelet transforms based on complex Morlet wavelets following the approach of Oya et al. (2002). Center frequencies ranged from 20 to 200 Hz in 5 Hz increments. ERBP was calculated for each center frequency on a trial-by-trial basis and normalized to median baseline power, measured for the same center frequency within a 100 to 200 ms window prior to stimulus onset. ERBP values were then log-transformed and averaged across trials. For quantitative analysis of ERBP, we focused on the high gamma ECoG frequency band (Crone et al., 2001; Brugge et al., 2009; Edwards et al., 2009), which was defined in the present study within a range of center frequencies between 70 and 150 Hz. The wavelet constant ratio used for time-frequency analysis was defined as \( f_0/\sigma_f = 9 \), where \( f_0 \) is the center frequency of the wavelet and \( \sigma_f \) is its standard deviation in the frequency domain. At this wavelet constant value, contribution of energy from post-stimulus onset interval to the estimate of baseline power is negligible for the range of center frequencies that correspond to the high gamma frequency band (Steinschneider et al., 2011). The magnitude of high gamma ERBP was taken as the average power within the 50-250 ms time window after stimulus onset. Maps of AEP and ERBP cortical activation were smoothed using triangle-based cubic interpolation with an up-sampling factor of 16 for display purposes.

For statistical analyses, we employed the analysis of variance (ANOVA) model that is commonly used to test hypotheses concerning the effects of experimental factors on a dependent univariate response measurement (e.g. either ERBP or PLV). One set of observations consisted
of ERBP or PLV measurements obtained at unique recording sites located in either left or right hemisphere locations for each tested click-rate. Thus, for this response metric, we had one repeated-measures, within-subjects factor (CLICK RATE with 6 levels for ERBP: 25, 50, 100, 125, 150, 200 or 5 levels for PLV: 50, 100, 125, 150, 200) and one between-subjects factor (LATERALITY with 2 levels: LEFT, RIGHT). We emphasize that members of our sample are unique “recording site locations”, although referred to here in the standard statistical parlance as subjects. A second sample of observations consisted of ERBP or PLV measurements for each of 5 click-rates obtained at a unique recording site located on PLST and at a site in medial HG located in the same hemisphere. Thus, for this design, we had one repeated-measures, within-subjects factor (CLICK RATE with 5 levels: 50, 100, 125, 150, 200) and one between-subjects factor (LOCATION with 2 levels: PLST, medial HG).

We used a generalized linear model (GLM) approach (SAS 9.3 Procedure GLM) to these repeated-measures ANOVA, because it did not require the assumption of sphericity and could accommodate unequal sample sizes. The SAS GLM procedure for repeated measures for a one dependent-variable response, produced two different sets of the within-subjects hypothesis tests: one using the multivariate approach, the other using the univariate approach. Generally, both sets of tests yielded similar results. The multivariate measures provided for testing effects, as well as for any specified contrast (e.g. testing for a linear trend), included Wilks’ lambda, Pillai’s trace, Hotelling-Lawley trace, and Roy’s greatest root. We report results using only Pillai’s trace because is considered more robust to violations of the multivariate’s assumption of homogeneity of variance/covariance matrices. Nevertheless, almost identical conclusions were reached using any of the above measures.
Spatial relationship between high gamma activity and FFRs was measured in each subject using Pearson’s correlation between averaged high gamma ERBP and PLVs measured in individual recording channels, followed by a repeated-measures ANOVA on Pearson’s $R$ values.

Modulation of high gamma ERBP by the stimulus temporal envelope was evaluated for long (1 s) click train stimuli presented at relatively low rates (4, 8 and 16 Hz). By analogy with FFR measurements, PLVs were calculated from single-trial high gamma ERBP signals within a window of 250-1000 ms after stimulus onset. Significance of PLVs representing modulation of high gamma ERBP by the driving frequency of the click train stimuli was assessed with correction for multiple comparisons (false discovery rate $q=0.01$).
RESULTS

Series 1: Short click trains

Figures 1 and 2 present the cardinal features of responses (AEP, FFR and ERBP) recorded from sites on STG of the left hemisphere of one subject (Fig. 1) and the right hemisphere of another (Fig. 2). In each case, click train stimulation elicited robust responses characterized by AEP waveforms and their spatial distributions consistent with those described for PLST in our earlier reports (Howard et al., 2000; Reale et al., 2007). An “X” on each AEP map shown in Figs. 1B and 2B marks the recording site at which the AEP of greatest amplitude was recorded; a “Y” marks a second site within PLST that was some distance away and where the AEP was of lower amplitude. The AEP, FFR and ERBP derived from responses at these two sites to click trains of increasing rates are shown overlapped in Figs. 1C and 2C. The AEP and ERBP exhibited an increase and the FFR a decrease in magnitude with increase in click rate. The increase in magnitude of the AEP and ERBP was accompanied by a decrease in response latency. The growth in magnitude of the ERBP occurred across a wide range of ECoG frequencies, from about 30 Hz to at least 200 Hz. As a rule, the largest ERBP responses were observed between about 100 and 300 ms after stimulus onset. Hence, at any click rate ERBP temporally overlapped the AEP and FFR but exhibited a longer latency.

Statistical analysis of the ERBP sample population indicated that the assumption of sphericity was violated (Mauchly's Test of Sphericity), $\chi^2(14) = 668.839, p < 0.0001$. Consequently, degrees of freedom were adjusted using Greenhouse-Geisser correction. Subsequent analyses revealed that click rate had a significant main effect $[F(2.96, 1412.1) = 238.02, p < 0.0001]$, and that ERBP values tended to increase linearly (employing Pillai’s trace
statistic) with click-rate $[F(1, 477) = 494.10, p < 0.0001]$. Mean ERBP values for left hemisphere cases were observed to be lower than for right hemisphere cases, as supported by a significant main effect $[F(1, 477) = 9.52, p < 0.005]$ for laterality. Subsequent post-hoc t-tests (Tukey-Kramer adjustment for multiple comparisons: $p < 0.05$) indicated lower left-hemisphere ERPB values at all 6 click-rates. There was no significant interaction between click rate and location factors $[F(2.96, 1412) = 2.47, p=0.0785]$. FFR strength was quantified in terms of phase-locking values (PLVs). Figure 3 shows PLVs as functions of click rates for all recording sites in all subjects that exhibited a significant ($q<0.01$) PLV to at least one of the five stimuli (50, 100, 125, 150, and 200 Hz). The FFR was typically most prominent at 50 Hz with its amplitude dropping off above that rate. The total number of recording sites across the fourteen subjects that exhibited significant PLVs to 50, 100, 125, 150, and 200 Hz click trains was 153, 16, 7, 9 and 3, respectively. Phase locking to 25 Hz often failed to reach significance for reasons explained in Methods, and hence the 25 Hz stimulus condition was excluded from quantitative analyses of responses to the short click train series.

In order to better fit the assumptions of normality and homogeneity of variances, PLV data were transformed (arcsin of square root) prior to parametric statistical procedures. Transformed variables did not violate the assumption of sphericity (Mauchly's Test of Sphericity), $\chi^2 (9) = 15.37, p = 0.08$. The resulting analysis revealed that click rate had a significant main effect $[F(4, 488) = 124.70, p < 0.0001]$, while the influence of laterality was non-significant $[F(1, 122) = 1.74, p= 0.1895]$. Thus, there was no significant difference between PLVs in left and right hemisphere sites. Both multivariate $[F(4, 119) = 2.98, p=0.0218]$ and univariate $[F(4, 448) = 2.73, p = 0.0287]$ tests indicated a significant interaction effect of click rate by laterality on PLVs. However, the interaction between these two variables may be too small to have any
practical significance since the main effect for click rate was much stronger than that for this interaction.

333  **Functional maps**

Spatial maps of AEP, FFR and ERBP were derived from responses to click trains of different rates from all 16 subjects in the study. Although the maps derived from the three measures overlapped, they were not coextensive. The AEP, FFR and ERBP were not always maximal at the same locations and at any given cortical site they were often present in varying degrees. The spatial distribution of cortical responses to click trains presented at different stimulus rates is shown for two representative left-hemisphere (Fig. 4A and 4B) and two right-hemisphere cases (Fig. 5A and 5B). The map shown in Fig. 4A is derived from the data shown in Fig. 1A, and the one shown in Fig. 5A is derived from data shown in Fig. 2A. Data are plotted as interpolated cortical activation maps showing both the distribution of normalized RMS magnitude of the AEP, and of high gamma ERBP. The contour patterns associated with AEPs and ERBP were rather complex and often characterized by presence of multiple active foci that changed shape and grew in size with changes in click rate. There was no consistent evidence of an orderly representation of cortical sites preferentially responding to specific driving frequencies. While there was a certain consistency in the spatial distribution of AEPs with patterns of high gamma activity in individual subjects, the two response features were not entirely congruent.

The extent of spatial overlap between high gamma ERBP and phase-locked responses (PLVs) was quantified for each subject using Pearson’s correlation. Repeated-measures ANOVA on Pearson’s $R$ values revealed a significant effect of frequency [$F(4,56)=12.8, p<0.0005$]. A significant linear trend [$F(1,14)=47.9, p<0.0005$] indicated that the two features of the cortical
response – high gamma ERBP and the FFR – were more spatially congruent at lower click rates (Fig. 6). Furthermore, repeated-measures ANOVA revealed a significant between-subjects effect of laterality \[F(1,14)=8.96, p<0.02\] with left-hemisphere data characterized by lower Pearson’s \(R\) values. There was no significant laterality-frequency interaction \[F(4,56)=1.23, p=0.361\]. Thus, there was a greater spatial disconnect between ERBP responses and PLV in the left hemisphere cases compared to the right hemisphere subjects. Specifically, cortical sites that exhibited significant FFRs to the driving frequency (marked by circles in Figs 4 and 5) were often located just outside regions where AEPs and ERBP were of greatest magnitude.

**Relationship to responses obtained simultaneously on HG**

The majority of the subjects in this study were also implanted with multicontact depth electrodes in the superior temporal plane that targeted auditory cortex on HG. Thus, it was possible to record simultaneously click-elicited activity across a range of click rates from the lateral surface of the temporal lobe, including PLST, as well as presumed auditory core cortex and adjacent non-core fields in HG. An example of such a comparison is shown in Figure 7. Responses to click trains presented at rates between 25 and 200 Hz are shown for three cortical sites (Fig. 7A): a representative PLST site, a presumed core cortex site in medial HG and a site in a non-primary field on lateral HG, marked by X, Y and Z, respectively. Consistently with our previous studies (Howard et al, 2000; Brugge et al., 2008, 2009), AEP waveforms recorded in response to click trains had a different morphology between the three different locations (Fig. 7B). AEPs recorded from PLST often (but not always) had smaller amplitudes than those recorded from medial HG, yet were typically larger and had shorter latencies than AEPs recorded from lateral HG. Comparison of FFRs revealed across-field differences in capacity for synchronization to the
stimulus periodicity. PLST reliably exhibited FFRs to click trains presented at 50-100 Hz, and rarely to higher rates (see Fig. 3), while medial HG could typically phase-lock to 200 Hz click trains (see Brugge et al., 2009).

PLVs that characterized synchronization to the stimulus periodicity were consistently higher for medial HG sites compared to PLST (Fig. 8). For statistical analysis of this observation, the sample of PLV observations was constructed by choosing those medial HG sites for which there were also simultaneously recorded PLST sites in the same hemisphere. Only those PLST electrode sites that evidenced the largest PLV values in response to 50 Hz click trains in each tested subject were included. The assumption of sphericity could not be rejected in this sample (Mauchly's Test of Sphericity: \( \chi^2(9) = 12.294, p = 0.1972 \)), and inferences from multivariate and univariate tests agreed. The resulting analysis revealed that the click rate had a significant main effect \( F(4, 64) = 34.69, p < 0.0001 \) on transformed PLVs. This influence was also dependent on location (medial HG vs. PLST) \( F(1, 19) = 19.05, p < 0.0005 \). Subsequent post-hoc \( t \)-tests (Tukey-Kramer adjustment for multiple comparisons; \( p < 0.02 \)) indicated that PLVs in PLST were consistently lower than medial HG values for all five studied click rates. The interaction between click rate and recording location was not significant \( F(4, 64) = 1.54, p=0.2020 \).

Lateral HG only occasionally exhibited an FFR, which was considerably weaker than that in PLST and, if at all present, was limited to relatively low stimulus rates (25-50 Hz). Finally, ERBP analysis revealed differences in magnitude and latency of ECoG power changes between the three sampled cortical areas (see Fig. 7B). Putative core auditory cortex of medial HG had consistently shorter latency of ERBP changes compared to PLST and lateral HG. At low rates (e.g., 25 Hz), activity recorded from medial HG featured bursts of high gamma ERBP that were
temporally related to the FFR and followed each successive click in the train. Such bursts or
modulation of high gamma ERBP at low stimulus rates were not present in recordings made
from either PLST or lateral HG.

Series 2: Long (1 s) click trains

Results presented in Figures 1-7 were obtained using relatively short-duration (160 ms) click-
train stimuli, delivered at rates between 25 and 200 Hz. In a subset of subjects (n=6), we used a
complementary set of longer-duration (1 s) click trains, which were presented at a different range
of repetition rates (4-128 Hz). Although these stimuli were not studied as extensively as the short
duration click trains, Figure 9 presents examples of responses to these stimuli recorded from two
PLST sites in two different subjects (Fig. 9A, 9C and 9B, 9D, respectively). At a relatively low
rate of 4 Hz, each click in the train evoked a distinct AEP complex (top plots in Fig. 9A and 9B).
Here, the inter-click interval of 250 ms was long enough for each AEP to develop to each
individual click in the train. This is further illustrated in Figure 9C and 9D, which depict portions
of the AEP waveforms that correspond to responses elicited by individual clicks. Early
components of the AEP waveform that were associated with each click in the train resembled the
morphology of the onset AEP complex in response to a higher-rate (128 Hz) stimulus. As the
rate increased to 8-16 Hz, responses to successive clicks overlapped, and as the click rate
increased beyond 16 Hz, distinct on- and off-response AEP complexes emerged. Activity within
PLST, elicited by click trains presented at rates up to at least 64 Hz contained an FFR
component, which was particularly noticeable in the AEP waveform at 16-32 Hz stimulus rates.

Time-frequency analysis of ECoG revealed patterns of high gamma ERBP responses to
click trains within PLST. At low repetition rates (4-16 Hz) click trains elicited very modest
changes in high gamma power. High gamma ERBP typically was not modulated by the stimulus envelope of the click trains presented at 4-16 Hz. PLV analysis failed to reveal significant (at \( q=0.01 \)) modulation of high gamma ERBP by the driving frequency of click trains presented at these repetition rates in any PLST recording site in 5 out of 6 subjects tested using this paradigm. In the remaining subject (R152), significant PLVs characterized responses obtained from three recording sites located on the STG.

As stimulus rate increased, stronger ERBP responses emerged, peaking at about 100-150 ms after stimulus onset and then gradually decreasing over the duration of the 1 s-long train. Overall, results obtained from experiments with 1 s click trains paralleled our findings with 160 ms click train series. In both sets of experiments, we observed a dual representation of repetitive stimuli by synchronized and non-phase-locked high gamma cortical activity at low and high repetition rates, respectively.
DISCUSSION

There were three major findings related to the complex temporal and spatial patterns elicited in PLST by click trains of varying rates. First, the AEP was synchronized to individual clicks in the train for click rates below about 150-200 Hz. Phase locking to clicks tended to be maximal around 50 Hz and decline systematically with increasing click rate. Decline in phase locking was associated with an increase in the amplitude of the AEP. At any given click rate, between 50 and 200 Hz, phase locking tended to be weaker in PLST than in the putative auditory core cortex. The AEP is interpreted to represent the incoming volley of afferent activity and subsequent synaptic currents (Vaughan, 1969; Vaughan & Arezzo, 1988; Steinschneider et al. 1992, 1994).

Second, the distributions within PLST of phase locked AEP and of ERBP were overlapping, but not coextensive. Third, cortical activity in PLST in the high gamma frequencies, unlike that of the auditory core, was not temporally modulated by individual clicks, when the click trains were presented at relatively low rates (<30 Hz, i.e. below the lower limit of pitch; Krubmholz et al., 2000). Like the AEP, however, the magnitude of the ERBP generally increased with click rate. ECoG frequencies within the high gamma band have been shown to be associated with both spiking activity and hemodynamic changes (Nir et al., 2007; Whittingstall & Logothetis, 2009), which allows direct comparison of data obtained under different experimental conditions.

Phase-locked responses to click trains

Temporal representation of stimulus periodicity is a basic property of the auditory system found on all levels from the auditory nerve to the auditory cortex (Langner, 1992; Joris et al., 2004). At the cortical level, direct electrophysiological study of temporal processing has been largely
confined to the auditory core of awake monkeys listening to sinusoidally amplitude-modulated (SAM) sounds (reviewed by Wang et al., 2003; Wang, 2007). In humans, Liegeois-Chauvel et al. (2004) measured phase-locking to SAM noise bursts by calculating the power of the AEP signal at the driving modulation frequency. For core cortex, the strongest phase locking was most often found at 8 Hz, but could be seen as high as 32 Hz, whereas for lateral STG the best modulation frequency did not exceed 16 Hz. Direct comparison of findings reported by Liegeois-Chauvel et al. (2004) using AEP power with the data presented in our study is not appropriate, because we employed a power-independent metric (PLV) to characterize cortical phase-locked activity.

Ongoing cortical activity is known to exhibit the inverse power law behavior, with log power decreasing nearly linearly with increasing log frequency (Buszaki, 2006). Therefore, power in the spectra of the AEP can be expected to decrease with frequency regardless of whether an FFR is actually present or not. In a study that investigated auditory cortical responses to bursts of sinusoidal amplitude-modulated noise, Gourévitch et al. (2008) corrected for the $1/f$ ECoG characteristic and estimated the strength of the FFRs as a signal-to-noise ratio. In their study, power of the AEP signal at the driving frequency was normalized to an estimate of background activity power at the same frequency. Responses recorded from sites on the lateral STG were characterized by lower signal-to-noise ratios compared to those recorded from HG across a range of modulation frequencies between 4 and 128 Hz. Nevertheless, several sites on the lateral STG were shown to exhibit signal-to-noise ratios between 10-15 dB at 64 Hz stimulation, and between 5-10 dB in response to 128 Hz amplitude-modulated noise stimulus (Gourévitch et al., 2008). These observations are consistent with our present findings.

As shown in the present study (see Fig. 1C, 2C, 9A, 9B) and elsewhere (Brugge et al., 2009; Nourski & Brugge, 2011), cortical responses to repetitive stimuli presented at rates of
above approximately 30 Hz are characterized by increases in cortical activity in gamma and high
gamma ECoG frequency bands. This constitutes a potential confounding factor for estimation of
the FFR from the AEP power spectra to repetition rates that correspond to the frequency range of
the non-phase-locked ERBP response. AEPs are typically averaged over a relatively small
number of trials (50-100 in this study). Time-domain averaging of peristimulus epochs
minimizes the contribution of non-phase-locked components to the AEP, but does not eliminate
it completely. Therefore, the power spectrum of the AEP may feature a power increase resulting
from residual (uncancelled by time-domain averaging) non-phase-locked activity in gamma and
high gamma bands, overlapping with the driving frequency. Such circumstances may lead to a
false-positive detection of the FFR if it’s estimated from this AEP power spectrum.

This reasoning motivated us to take advantage of the phase-locking value approach in our
characterization of cortical FFRs. PLV is an amplitude-independent metric of response
synchrony and thus has the advantage of more accurate separation of phase-locked response
components from non-phase-locked activity. Also, PLVs are dimensionless units, this facilitates
comparison across subjects. A potential weakness of this approach also stems from the fact that it
is amplitude-independent. As such, it may be too sensitive in picking up far-field activity,
producing phase-locking values of spurious significance and thus yielding false-positive results.

We attempted to minimize contribution of possible common sources by applying a spline
Laplacian transform to the data. This procedure yields a noticeable increase in the spatial
specificity of the recordings compared to unprocessed data (see Reale et al., 2007, for examples).
Sharp gradients in response magnitude between adjacent (5 mm apart) recording sites can be
observed in Figures 1 and 2. Finally, relatively fast ECoG components, including those in the
high gamma frequency band and FFRs to comparable stimulus driving frequencies, are recorded
in close spatial proximity to their cortical sources and thus are highly localized (Frien et al., 2000; Liu & Newsome, 2006).

Phase-locking analysis implemented in this study provides evidence for a high capacity of human non-primary auditory cortex of PLST for explicit temporal representation of periodicity. This raises a concern about possible far-field, volume-conducted sources of phase-locked activity recorded with subdural electrodes, originating, for example, from primary auditory cortex or auditory brainstem nuclei. A clear difference in response morphology and upper limit of phase locking between recordings from PLST and medial HG makes significant contribution of activity from core auditory cortex to PLST recordings unlikely (see Fig. 7). Furthermore, if FFRs recorded with electrodes placed over the lateral STG were volume-conducted from core auditory cortex, one might expect to obtain significant PLVs mainly from sites located immediately adjacent to the Sylvian fissure. As can be observed from Figures 3 and 4, this was not the case in our study.

Upper limits of phase locking we obtained can be attributed to the abrupt onset of the clicks we employed, and hence stronger temporal locking, as compared to a more gradual ramping up of each cycle of the SAM signal (Heil, 2003). Core neurons are as remarkably precise as auditory nerve fibers in their firing to an acoustic transient (Heil & Irvine, 1997), which is a prerequisite in supporting a temporal representation of phonetically important speech components (Phillips & Farmer, 1990; Phillips & Hall, 1990; Heil & Irvine, 1997; Phillips, 1998; Heil, 2003). Underlying this precision is a delicate balance of excitation and inhibition at the thalamocortical synapse (Wehr & Zador, 2003; Tan et al., 2004).

The FFR reflects cortical activity that is phase-locked to stimulus periodicity. It may be considered the simplest, isomorphic, form of temporal sound feature representation on the
cortical level. The presence of FFRs to relatively fast (>30 Hz, i.e. above the lower limit of pitch) periodic stimuli raises questions concerning the functional utility of such isomorphic representations within non-core auditory cortex. Several lines of evidence suggest that integrity of primary auditory cortex is required for processing such segmental components of speech, which exhibit temporal variations in the millisecond to tens-of-milliseconds range (Phillips & Farmer, 1990), and the same may be said for area PLST. Phase-locked activity up to several tens of Hz may facilitate interneuronal communication, engaging large-scale networks involved in perception, and may play a critical role in multisensory interactions by, for example, synchronizing auditory and visual inputs (Senkowski et al., 2005; Doesburg et al., 2008; Hipp et al., 2011). Furthermore, animal studies demonstrate a relationship between behavioral training and phase-locked responses in the primary auditory cortex (Bao et al., 2004; Schnupp et al., 2006), suggesting that preservation of the stimulus timing at the cortical level may also be beneficial for perception of behaviorally relevant sounds. These hypotheses can be tested in future studies that employ active listening tasks to investigate the relationship between the subject’s perception of temporally dynamic stimuli and strength of phase-locked cortical responses.

**High gamma response**

Although phase locking to click trains generally decreased in strength with click rate beyond 50 Hz, auditory cortex within PLST maintained a representation of the stimuli by increases in high gamma power at higher stimulus rates. ERBP typically increased with stimulus repetition rate, at least up to about 150-200 Hz. This may be interpreted as a neural population-level correlate of a rate code for sound periodicity (e.g., Wang et al., 2008). Higher rates, above those for which
phase locking is reliably exhibited, may be encoded as discharge rate as suggested by Wang (2007). We observed also that the spatial distribution of non-phase locked gamma activity within PLST changed with click rate. This finding suggests that for stimulus rates close to or beyond the limits of cortical phase locking, information about stimulus rate may be represented in a more complex form by the spatial distribution of cortical activation.

Finally, although maps of the FFR and high gamma ERBP overlapped, their respective areas of maximal activity were not typically coextensive, particularly at higher click rates (see Fig. 6). This would suggest that within this overlapping auditory domain on posterolateral STG, the neural circuits underlying these two stimulus representations are, at least to some degree, segregated. While this spatial disconnect was seen to be more pronounced in the left hemisphere as compared to the right hemisphere, conclusions regarding laterality of the disconnect are tempered by the fact that it was not possible to record simultaneously from the two hemispheres in the same subject.

**Comparison with core auditory cortex**

The power of our experimental approach is that cortical activity may be recorded simultaneously from multiple auditory fields. We were, therefore, able to directly compare for each subject activity evoked by click trains in PLST with that evoked in presumed core and adjacent non-core areas of HG. Here we compare specifically the capacity of these areas to phase-lock to the click-train stimulus. These and other details of responses of HG core and belt cortex to click-train stimulation have been presented previously (Brugge et al. 2008, 2009). Simultaneous recordings in these areas were distinguishable by their AEP waveform complexes, ERBP and FFR. Robust phase locking to clicks was confined to HG core and PLST areas. This is despite the fact that the
auditory HG core receives its major afferent input from the ventral medial geniculate body (MGB) whereas PLST, if similar to connections in non-human primates, most likely receives a convergent input from the core and from extralemniscal ascending auditory pathways including medial geniculate complex and pulvinar (de la Mothe et al., 2012).

Responses to click trains recorded from PLST exhibited a certain degree of similarity with activity measured in core auditory cortex within posteromedial HG. In both regions, three basic response features - AEP waveform complexes, FFR and ERBP - could be identified. Qualitative changes in response properties that parallel perceptual classes of repetitive stimuli were found both in the auditory core and PLST (cf. Brugge et al., 2009; Nourski & Brugge, 2011). Specific features include changes in the shape of AEP waveforms elicited by successive clicks at and above 8 Hz, corresponding to perceptual blending of clicks into a unified flutter-like percept, and the emergence of distinct AEP on- and off- responses and high gamma ERBP at stimulus rates above around 30 Hz, thus corresponding to the lower limit of pitch in human listeners (Krumbholz et al., 2000).

Differences between core auditory cortex and PLST include overall morphology of the AEP complexes (see also Brugge et al., 2008). The strength of phase locking within PLST for a non-primary cortical region is consistently lower than that in the core (see also Brugge et al., 2009; Nourski & Brugge, 2011). In addition, while the timing of each click was typically represented by a burst in ERBP for relatively low-rate stimuli (<32 Hz) within the core auditory cortex, no bursts or modulation of high gamma ERBP at low stimulus rates was observed in recordings made from PLST. Thus, the mechanisms that sustain phase locking of the AEP in both core and PLST are not retained in the high gamma range of activity in PLST. Overall, the presence of robust responses to click trains in PLST and its high capacity for explicit temporal
representation of periodicity suggest that this cortex, while functionally distinct from the
auditory core, is relatively close to it within the hierarchy of auditory cortical fields. Nevertheless,
as we and others have suggested previously, there is still no consensus on the number of areas
present or their arrangement in human auditory cortex including this area of the posterior
superior temporal gyrus (Brugge et al., 2008; Hackett, 2011). Further studies using more
complex stimuli will further investigate the functional identity of the auditory cortex of PLST.
ACKNOWLEDGEMENTS

This work was supported by National Institute on Deafness and Other Communication Disorders at the National Institutes of Health (grant number R01-DC04290); National Center for Research Resources and the National Center for Advancing Translational Sciences at the National Institutes of Health (grant number UL1RR024979); Hearing Health Foundation (Collette Ramsey Baker Award), and the Hoover Fund. We thank Haiming Chen and Rachel Gold for help with data collection and analysis.
FIGURE LEGENDS

**FIG. 1.** Auditory cortical responses to click train stimuli recorded from the left hemisphere in a representative subject. A: Location of the 96-contact subdural grid. B: All-pass (1.6-500 Hz passband) AEP waveforms recorded from the subdural grid in response to 160 ms 100 Hz click trains. Negative voltage is plotted upwards. Sulcal patterns are indicated by grey outlines. MTG, middle temporal gyrus; sf, sylvian fissure; STG, superior temporal gyrus; sts, superior temporal sulcus. C: All-pass (1.6-500 Hz) and high-pass (cut-off 1 octave below driving frequency) AEP waveforms (blue and red traces, respectively) and time-frequency analysis of ECoG (color plots) recorded in response to click trains presented at rates between 25 and 200 Hz (top-to-bottom). Data from two recording sites, indicated by X and Y in panels A and B, are shown.

**FIG. 2.** Auditory cortical responses to click train stimuli recorded from the right hemisphere in a representative subject. See legend of Figure 1 for details.

**FIG. 3.** Phase-locking to click trains recorded from fourteen subjects. PLVs are plotted as functions of click rate. In each subject, data from sites that exhibited significant phase locking \((q<0.01)\) to at least one stimulus rate out of five are shown.

**FIG. 4.** Spatial distribution of cortical responses to click trains recorded from the left hemisphere in two representative subjects (shown in panels A and B). In each panel, location of the recording grid is shown on the top, followed by cortical activation maps across click rates (top-to-bottom rows), as measured by AEP RMS amplitude and high gamma ERBP (left and right columns,
respectively). AEP RMS amplitudes were calculated within 0-300 ms after stimulus onset; high gamma ERBP was averaged within 50-250 ms after stimulus onset for each recording site, normalized relative to the maximum values across stimuli in each subject and smoothed using cubic interpolation. Sulcal patterns are shown by grey lines. Circles indicate sites that exhibited significant PLVs at each stimulus rate.

**FIG. 5.** Spatial distribution of cortical responses to click trains recorded from the right hemisphere in two representative subjects (shown in panels A and B). See legend of Figure 4 for details.

**FIG. 6.** Correlation between high gamma ERRP and PLV at different click rates. Data from 96-contact subdural grids implanted in the left hemisphere in 7 subjects (blue circles) and in the right hemisphere in 9 other subjects (red circles). Blue and red lines represent mean Pearson’s $R$ values for left- and right-hemisphere cases, respectively. A small amount of variability was added to the abscissa values to make individual data points more visible.

**FIG. 7.** Comparison of responses to click trains between PLST and two recording locations within HG. A: Location of three exemplary recording sites (marked by X, Y, Z) in a representative subject. B: AEPs, FFRs and ERBP obtained from PLST, core auditory cortex in medial HG and a non-core field in lateral HG (top to bottom). Stimuli: 160 ms click trains, presented at rates between 25 and 200 Hz (left to right). Stimulus schematics are shown in top rows.
FIG. 8. Phase-locking in medial HG and PLST. Summary of data from 9 subjects (L145, L162, L178, R149, R152, R153, R154, R180, R186). PLVs are plotted as functions of click rate for sites within medial HG and PLST (teal and purple circles, respectively). Data from one medial HG and one PLST site that exhibited the highest PLVs at 50 Hz in each subject are shown. Teal and purple lines represent across-subject mean PLVs for medial HG and PLST sites, respectively. A small amount of variability was added to the abscissa values to make individual data points more visible.

FIG. 9. PLST responses to 1 s click trains. A, B: All-pass (1.6-500 Hz) and high-pass (cut-off 1 octave below driving frequency) AEP waveforms (blue and red traces, respectively) and ERBP (color plots), elicited by click trains of rates between 4 and 128 Hz (top to bottom). Exemplary data from two representative right-hemisphere sites in two different subjects (R152 and R180; panels A and B, respectively). C, D: AEP waveforms recorded in response to 4 and 8 Hz click trains are replotted by fragmenting and superimposing the portions of the AEP waveform that were recorded between consecutive clicks in the click train. The first 500 ms of the AEP waveforms elicited by 128 Hz click trains are plotted below.
REFERENCES


Figure 1
Figure 2
Figure 3
Figure 4
Pearson's R

Left hemisphere (n = 7)
Right hemisphere (n = 9)

Figure 6
A Lateral view of the right hemisphere

Right superior temporal plane

B Click rate (Hz):

- 100 μV
- 10 μV

Figure 7
Figure 8
Figure 9