Spectral Resolution of Monkey Primary Auditory Cortex (A1) Revealed With Two-Noise Masking

Yonatan I. Fishman1 and Mitchell Steinschneider1,2
1Department of Neurology and 2Department of Neuroscience, Albert Einstein College of Medicine, Rose F. Kennedy Center, Bronx, New York

Submitted 2 March 2006; accepted in final form 30 May 2006

Fishman, Yonatan I. and Mitchell Steinschneider. Spectral resolution of monkey primary auditory cortex (A1) revealed with two-noise masking. J Neurophysiol 96: 1105–1115, 2006. First published May 31, 2006; doi:10.1152/jn.00124.2006. An important function of the auditory nervous system is to analyze the frequency content of environmental sounds. The neural structures involved in determining psychophysical frequency resolution remain unclear. Using a two-noise masking paradigm, the present study investigates the spectral resolution of neural populations in primary auditory cortex (A1) of awake macaques and the degree to which it matches psychophysical frequency resolution. Neural ensemble responses (auditory evoked potentials, multiunit activity, and current source density) evoked by a pulsed 60-dB SPL pure-tone signal fixed at the best frequency (BF) of the recorded neural populations were examined as a function of the frequency separation (ΔF) between the tone and two symmetrically flanking continuous 80-dB SPL, 50-Hz-wide bands of noise. ΔFs ranged from 0 to 50% of the BF, encompassing the range typically examined in psychoacoustic experiments. Responses to the signal were minimal for ΔF = 0% and progressively increased with ΔF, reaching a maximum at ΔF = 50%. Rounded exponential functions, used to model auditory filter shapes in psychoacoustic studies of frequency resolution, provided excellent fits to neural masking functions. Goodness-of-fit was greatest for response components in lamina 4 and lower lamina 3 and least for components recorded in more superficial cortical laminae. Physiological equivalent rectangular bandwidths (ERBs) increased with BF, measuring nearly 15% of the BF. These findings parallel results of psychoacoustic studies in both monkeys and humans, and thus indicate that a representation of perceptual frequency resolution is available at the level of A1.

INTRODUCTION

An important function of the auditory nervous system is to analyze the spectral content of complex sounds. The importance of this function is exemplified by the maintenance of a tonotopic representation of stimulus frequency throughout the mammalian auditory pathway (e.g., Popper and Fay 1992). The spectral resolution of the auditory system has been examined using a variety of psychoacoustic methods in both humans and animals. Traditionally, auditory frequency resolution has been quantified by determining the frequency separation between sound components beyond which an abrupt perceptual change occurs, such as in loudness, roughness, or detection or discrimination thresholds (for reviews, see Greenwood 1991; Moore 1993; Scharf 1970). More contemporary studies have examined frequency selectivity using methods that include masking of a tone signal centered in a spectral notch in a wide-band noise or between two flanking tones or narrow bands of noise as a function of the frequency separation between the maskers and the tone signal (Glasberg and Moore 1990; Glasberg et al. 1984; Greenwood 1961; Ishigami et al. 1995; Moore 1993; Patterson et al. 1982; Rabinowitz et al. 1980). Results of these studies indicate that the auditory system consists of an array of overlapping band-pass filters, with only those components of a masker that pass through a filter being effective in masking a signal presented at the center frequency of the filter (Fletcher 1940; Greenwood 1961; Patterson 1974). The resolution bandwidths of these auditory filters have been quantified in terms of “critical bandwidths” or “equivalent rectangular bandwidths” (ERBs), which typically range from 10 to 20% of the center frequency in humans (Glasberg and Moore 1990; Greenwood 1991; Moore 1993). Thus the absolute bandwidth of frequency resolution increases with center frequency.

Humans display an auditory perceptual frequency resolution for simultaneously presented stimulus components similar to that of other mammals, including monkeys (Gourevitch 1970; Serafin et al. 1982), chinchillas (McGee et al. 1976), and cats (Pickles 1975, 1979), thus suggesting that some of the neural substrates underlying auditory spectral resolution in humans may be investigated in experimental animal models. The physiological determinants of auditory perceptual frequency resolution are unclear. A potential cochlear origin is suggested by the demonstration of two-tone inhibition (suppression of a response to a tone by the presence of another simultaneously presented tone of slightly different frequency) in auditory nerve fibers (Sachs and Kiang 1968) and by the good correspondence found between behavioral critical bandwidths and the effective bandwidth of frequency-threshold tuning curves of auditory nerve fibers in guinea pigs (Evans et al. 1989). In contrast, other investigations in the cat indicate that critical bandwidths of auditory nerve fibers underestimate psychophysical critical bandwidths by as much as threefold (Pickles 1975, 1979; Pickles and Comis 1976). These studies thus suggest a role for central auditory structures in determining the frequency selectivity found in behavioral experiments (Pickles 1975, 1979; Pickles and Comis 1976). Within central auditory structures, correlates of psychophysical critical bandwidths have been reported for single neurons in the inferior colliculus of anesthetized cats (Ehret and Merzenich 1985, 1988). Similar to the behavioral studies, the physiological studies used a narrow band of noise of variable bandwidth to just mask the response to a tone at the charac-
characteristic frequency of the recorded neurons. Keeping the noise
centered at the characteristic frequency and the spectrum level
constant, the bandwidth of the noise was progressively reduced
until a neural response to the tone reappeared. The bandwidth
of the noise masker at which the response to the tone was 20%
higher than the baseline firing rate operationally defined the
critical bandwidth (Ehret and Merzenich 1985, 1988). Physio-
logical critical bandwidths increased with center frequency in
parallel with behavioral critical bandwidths (Nienhuys and
Clark 1979; Pickles 1975, 1979). Although physiological crit-
ical bandwidths recorded in the inferior colliculus are closer to
behavioral values than at the level of the auditory periphery,
they are still about twice the value of behavioral critical bandwidths measured in the same species. Similar noise band-
narrowing techniques were used to measure neural critical bandwidths in primary auditory cortex (A1) of anesthetized
cats (Ehret and Schreiner 1997). As in the inferior colliculus,
critical bandwidths of auditory cortex neurons were about
twice the values obtained in behavioral studies.

In humans, noninvasive magnetoencephalography (MEG)
recordings of the N100m component of responses to tones
masked by flanking bands of noise (“notched noise”) demon-
strate masking bandwidths that increase with test tone fre-
cquency, in accordance with psychoacoustic findings (Sams and
Salmelin 1994). However, only two frequencies were tested in
these studies (1 and 2 kHz) and tuning bandwidths were
considerably larger than values obtained psychophysically in
humans for these frequencies (psychoacoustic tuning band-
widths at 1 and 2 kHz are 133 and 241 Hz, respectively
(Glasberg and Moore 1990), whereas N100m tuning band-
widths at 1 and 2 kHz are 247 and 602 Hz, respectively].
Therefore to date, a close correspondence between behavioral
and physiological frequency resolution has not been estab-
lished.

A recent study in awake monkeys demonstrated critical
bandlike patterns of neuronal ensemble responses to complex
tones in A1 (Fishman et al. 2000). Specifically, responses to
three-component harmonic complexes showed abrupt in-
creases in amplitude when the spacing between the compo-
nents exceeded 10 to 20% of the center frequency, in agree-
ment with behavioral values of critical bandwidth in monkeys
and humans (Gourevitch 1970; Moore 1993). Although this
suggests that critical band-related information is available at
the cortical level, this study did not examine whether physio-
logical masking functions are consistent with the “rounded
exponential” shapes of auditory filters derived in psychophys-
ical masking studies (e.g., Glasberg and Moore 1990; Glasberg
et al. 1984; Patterson 1976). Thus it remains unclear whether
an accurate representation of frequency resolution, as mea-
sured psychophysically, is present at the level of auditory
cortex.

To address this issue, the present study uses a two-noise
masking paradigm to examine the frequency resolution of
neural populations in A1 of awake monkeys and the degree to
which it corresponds to auditory frequency resolution mea-
sured psychophysically. Specifically, the response to a pulsed
tone at the best frequency (BF; the pure-tone frequency elic-
ting the largest neural response) of the recorded neural popu-
lation is examined as a function of the frequency separation
(ΔF) between the tone and two symmetrically flanking contin-
uous bands of noise. Similar stimuli have been used in psy-
choacoustic experiments to measure auditory frequency selec-
tivity in humans and monkeys (Glasberg et al. 1984; Goure-
vitch 1970; Greenwood 1961; Ishigami et al. 1995; Rabinowitz
et al. 1980), thereby allowing comparisons to be made between
physiological and psychophysical data. Moreover, the present
study simultaneously examines multunit activity (MUA), au-
ditory evoked potential (AEP), and current source density
(CSD) measures, which may facilitate the interpretation of
noninvasive MEG recordings measuring frequency selectivity
in humans.

METHODS

Two adult male macaque monkeys (Macaca fascicularis) were
studied using previously described methods (Fishman et al. 2000;
Steinschneider et al. 2003). Animals were housed in our AAALAC-
accredited Animal Institute under daily supervision of laboratory and
veterinary staff. All experimental procedures were reviewed and
approved by the AAALAC-accredited Animal Institute of Albert
Einstein College of Medicine and were conducted in accordance with
institutional and federal guidelines governing the experimental use of
primates. To minimize the number of monkeys used, other auditory
experiments were conducted in the same animals during each record-
ing session.

Surgical procedure

Animals were acclimated to the recording environment while sit-
ting in custom-fitted primate chairs before surgery. Under pentobar-
bital anesthesia and using aseptic techniques, holes were drilled
bilaterally into the dorsal skull to accommodate matrices composed of
18-gauge stainless steel tubes glued together in parallel. Tubes served
to guide electrodes toward A1 for repeated intracortical recordings.
Matrices were stereotaxically positioned to target A1. They were
oriented at a 30° anterior–posterior angle and with a slight medial–
lateral tilt to direct electrode penetrations perpendicular to the superior
surface of the superior temporal gyrus, thereby satisfying one of the
major technical requirements of one-dimensional CSD analysis
(Vaughan and Arezzo 1988). Matrices and Plexiglas bars, used for
painless head fixation during the recordings, were embedded in a
pedestal of dental acrylic secured to the skull with inverted bone
screws. Peri- and postoperative antibiotic and antiinflammatory med-
ications were routinely administered.

Recordings began after a 2-wk postoperative recovery period.
Recordings were conducted in an electrically shielded, sound-attenu-
ated chamber, with the animals awake and with arms comfortably
restrained. Frequent visits by experimenters into the recording cham-
er between stimulus blocks and delivery of fruit juice rewards
helped to maintain the animals in an alert state throughout the
recordings. However, the possibility of intermittent drowsiness during
recordings cannot be excluded.

Neurophysiological recordings

Intracortical recordings were performed using linear-array multi-
contact electrodes containing 14 recording contacts, evenly spaced at
150-μm (±10%) intervals (Barna et al. 1981). Individual contacts
were constructed from 25-μm-diameter stainless steel wires, each
with an impedance of about 200 kΩ. An epidural stainless steel guide
tube positioned over the occipital cortex served as a reference elec-
trode. Field potentials were recorded using unity-gain headstage
preamplifiers, and subsequently amplified 5,000 times by differential
amplifiers (Grass) with a frequency response down 6 dB at 3 Hz and
at 3 kHz. Signals were digitized on-line at 3.4 kHz and averaged by
computer (Neuroscan software and hardware; Neurosoft) to yield
auditory evoked potentials. To derive multunit activity (MUA),

J Neurophysiol • VOL 96 • SEPTEMBER 2006 • www.jn.org
signals were simultaneously high-pass filtered at 500 Hz, amplified an additional eight times, full-wave rectified, and then low-pass filtered at 600 Hz before digitization and averaging to prevent signal aliasing (see Super and Roelfsema 2005 for a methodological review). MUA is a measure of the envelope of summed action potential activity of neuronal aggregates within a sphere of about 100 μm in diameter surrounding each recording contact (Brosch et al. 1997; Legatt et al. 1980; Super and Roelfsema 2005; Vaughan and Arezzo 1988). MUA and single-unit techniques have been shown to yield similar results (Super and Roelfsema 2005) and MUA displays greater response stability than single-unit activity (Nelken et al. 1994).

One-dimensional CSD analyses characterized the laminar pattern of net current sources and sinks within A1 generating the AEPs. CSD was calculated using a three-point algorithm that approximates the second spatial derivative of voltage recorded at each recording contact (see Freeman and Nicholson 1975; Nicholson and Freeman 1975). Current sinks represent net inward transmembrane current flow associated with local depolarizing excitatory postsynaptic potentials or passive, circuit-completing current flow associated with hyperpolarizing potentials at adjacent sites. Current sources represent net outward transmembrane currents associated with active hyperpolarization or passive current return associated with adjacent depolarizing potentials. The corresponding MUA profile helps to distinguish these possibilities: current sinks coincident with increases in MUA reflect net synaptic depolarization, whereas current sources coincident with reductions in MUA from baseline levels likely reflect neuronal hyperpolarization.

Electrodes were moved with a microdrive and guided by on-line examination of click-evoked potentials. Test stimuli were delivered when the electrode channels bracketed the inversion of early AEP components and the largest MUA, typically occurring during the first 50 ms, was situated in the middle channels. For each stimulus condition, evoked responses to 50 presentations of the stimuli were averaged with an analysis window of 300 ms (including a 25-ms prestimulus baseline interval).

At the end of the recording period, monkeys were deeply anesthetized with sodium pentobarbital and transcardially perfused with 10% buffered formalin. Tissue was sectioned in the coronal plane (80 μm thickness) and stained for acetylcholinesterase (Bakst and Amaral 1984; Nissl substance to reconstruct the electrode tracks and to identify A1 according to previously published criteria (Hackett et al. 1998; Merzenich and Brugge 1973; Morel et al. 1993; Wallace et al. 1991). The earliest sink/source configuration was used to locate lamina 4 (Steinschneider et al. 1992). Other laminar locations were determined by their relationship to lamina 4 and the measured widths of laminae within A1 for each electrode penetration.

Stimuli

All stimuli were generated and delivered at a sample rate of 100 kHz by a PC-based system using RP2 modules (Tucker Davis Technologies). Frequency response functions (FRFs), based on pure-tone responses, were used to characterize the spectral tuning of the cortical sites. Pure tones used to generate the FRFs ranged from 0.2 to 170 kHz, were 175 ms in duration (including 10-ms linear rise/fall ramps), and were presented with a stimulus onset-to-onset time of 658 ms. Pure tones were monaurally delivered at 60 dB SPL by a dynamic headphone (MDR-7502, Sony) to the ear contralateral to the recorded hemisphere. Sounds were introduced to the ear through a 3-in-long, 60-m plastic tube attached to the headphone. Sound intensity was measured with a Bruel & Kjær sound level meter (type 2236) positioned at the opening of the plastic tube. The frequency response of the headphone was flattened (±3 dB) from 0.2 to 170 kHz by a graphic equalizer (GE-60, Rane).

Two-noise–masking stimuli consisted of a pulsed 60-dB SPL pure-tone signal, which was fixed at the BF of the recorded neural population, symmetrically flanked by two continuous 50-Hz-wide bands of noise presented at a combined intensity of 80 dB SPL (Fig. 1). The pure tone was 50 ms in duration, including 5-ms linear rise/fall ramps. The independent variable was the relative ΔF between the near spectral edges of the noise bands and the frequency of the pure tone. ΔF ranged from 0 to 50% of the BF in 10% increments. Spectral edges of the noise bands had slopes of 96 dB/octave. Stimulus intensities and ΔFs were within the range of those examined in psychoacoustic experiments using notched noise, two-noise masking, and two-tone masking to assess frequency selectivity (Glasberg and Moore 1990; Glasberg et al. 1984; Ishigami et al. 1995; Oxenham and Shera 2003; Patterson et al. 1982). Although formal psychoacoustic tests were not conducted in the present study, based on informal listening by experimenters, the pure-tone signal was undetectable when ΔF was 0% (no spectral gaps between the noise band edges and the pure-tone signal), but readily heard when ΔF was 50%. Although neural response thresholds were not measured, responses to the signal tone were either minimal or absent for all sites when there was no spectral gap between the noise bands (i.e., when ΔF = 0%).

Data analysis

The BF of each cortical site was defined as the pure-tone frequency eliciting the largest peak MUA in lamina 4 and lower lamina 3 within the first 50-ms poststimulus onset. The expected anterior–lateral to posterior–medial topographic gradients of low to high BF representation were found in both animals (Merzenich and Brugge 1973; Morel et al. 1993; Recanzone et al. 2000). MUA results are based on the average of MUA recorded in two adjacent electrode channels centered within lower lamina 3 in the thalamorecipient zone, as indicated by the presence of prominent
MUA and large-amplitude initial current sinks (see Fishman et al. 2000; Steinschneider et al. 1992, 1994). Previous studies localized the initial sinks to thalamorecipient zone layers of A1 (Metherate and Cruikshank 1999; Müller-Preuss and Mitzdorf 1984; Steinschneider et al. 1992; Sukov and Barth 1998). This MUA is likely dominated by action potentials of larger pyramidal cells, but potentially includes contributions from nonpyramidal cells and thalamocortical afferents (Steinschneider et al. 1992). For all sites examined, MUA recorded in the two thalamorecipient zone channels had the same BF.

Six stereotypical A1 response components, also identified in previous work (e.g., Fishman et al. 2000; Steinschneider et al. 1992, 1994), were examined in the study to allow translational comparison with auditory evoked responses recorded noninvasively in humans: 1) MUA centered in lower lamina 3 (both peak amplitude and area, i.e., summed activity within 0- to 50-ms posttone onset), 2) P28 component of the superficial AEP, 3) N60 component of the superficial AEP, 4) Lower lamina 3 CSD Sink, 5) Upper lamina 3 CSD Sink, and 6) P28 CSD Source, occurring within superficial laminae.

RESULTS

Results are based on 22 sites in A1 of 2 monkeys. BFs were evenly distributed and ranged from 0.6 to 13.0 kHz (Fig. 2). All sites displayed FRFs with a single BF peak. The mean FRF, expressed as a function of the percentage deviation from the BF, is shown in Fig. 3 (bin width = 10%). This illustrates the average selectivity of pure-tone frequency tuning of MUA recorded at the 22 sites. The 50% down points of the mean FRF are between 20 and 25% of the BF. Average spectral tuning bandwidth of MUA is comparable to that obtained in a previous study using similar methods (Steinschneider et al. 2005) and to frequency tuning bandwidths of single neurons in A1 of squirrel monkeys (Cheung et al. 2001).

At all sites examined, amplitude of MUA evoked by the pulsed tone at the BF increased with increasing ΔF between the edges of the flanking noise bands and the frequency of the pulsed tone. This finding is illustrated in Fig. 4, which shows MUA from six sites (three from each of the two animals) with different BFs as a function of ΔF. FRFs of the sites based on the area of MUA within 0- to 50-ms poststimulus onset are plotted in the left-hand column. The solid vertical line superimposed on the FRFs indicates the BF, whereas dotted vertical lines superimposed on the FRFs indicate frequencies corresponding to the inner edges of the two noise maskers when ΔF was ±0.5 times the BF. Normalized amplitudes of MUA evoked by the BF tone are plotted as a function of ΔF in the right-hand column (filled symbols: peak MUA; open symbols: MUA area). These curves, designated as “masking response functions,” approximate sigmoid shapes. Comparison between FRFs and masking response functions reveals opposite trends: whereas FRFs decrease with increasing frequency distance from the BF, masking response functions increase with increasing ΔF.

AEP and CSD response components display increases in amplitude with increasing ΔF that parallel MUA. Figure 5 depicts average MUA, superficial AEP, and CSD waveforms recorded at laminar depths at which stereotypical A1 response components were identified. The solid lines represent response waveforms averaged across the 22 sites; the extent of shading above and below the lines represents ± SE. The onset and peak of lower lamina 3 MUA occur about 2 ms after the onset and peak of the early CSD sink in lower lamina 3 (“Lower Lamina 3 Sink”), which indexes initial synaptic depolarizations within A1. A slightly later current sink in upper lamina 3 (“Upper Lamina 3 Sink”) is balanced by a coincident superficial source in laminae 1 and 2, peaking at about 28 ms poststimulus onset (“P28 Source”). This superficial current source/sink pair represents the main generator of the superficially recorded AEP P28 component, indicated in the top row of AEP waveforms (Arezzo et al. 1986; Steinschneider et al. 1992, 1994). The P28 and subsequent N60 component of the AEP are of interest because they may be homologous to components of the human surface-recorded AEP P50 and N100, respectively (Eggermont 2001; Steinschneider et al. 1994). The persistence of response components after averaging across sites indicates that these components are highly consistent and synchronized across the A1 sites examined.

Average masking response functions for all MUA, AEP, and CSD response components are shown in Fig. 6. Response amplitudes were normalized to the maximum response at each site before averaging across sites. All response components display statistically significant increases in amplitude with ΔF (ANOVA, P < 0.00001). With the exception of the N60
component of the AEP, on average, response amplitudes are near zero at \( \Delta F = 0\% \) and near 100\% at \( \Delta F = 50\% \).

To compare the frequency resolution of these response components with psychophysical measures in monkeys and humans, amplitudes of normalized response components were converted to logarithmic units (dB) and fitted with the following "rounded exponential" auditory filter equation typically used to characterize frequency selectivity in human psychoacoustic experiments (Glasberg and Moore 1990; Glasberg et al. 1984; Moore 1993; Patterson et al. 1982):

\[
\text{Response amplitude (in dB)} = K - 10 \log_{10} \left[ (1 - r)(1 + pg) \exp^{-r g} + r \right]
\]

where \( K, p, \) and \( r \) are fitting parameters, and \( g = \Delta F \) (expressed as a proportion of the pure-tone "signal" frequency, which was set equal to the BF). The rationale for this conversion is based on the fact that psychophysical thresholds in studies examining auditory filter selectivity are expressed in decibels before fitting with the rounded exponential equation used to define auditory filter shapes (Moore 1993). Because the maskers in the present study were two narrow bands of noise, to a reasonable approximation, the auditory filter equation used in two-tone masking studies can be fitted directly to the present data (Glasberg et al. 1984).

To illustrate this fitting procedure and the derivation of corresponding auditory filter shapes, Fig. 7 shows normalized MUA data from the same sites as those depicted in Fig. 4, converted to logarithmic units and fitted by the auditory filter equation (top row of graphs). Parameter values were determined by a standard least-squares minimization procedure (Glasberg and Moore 1990; Moore 1993). Parameter values are indicated within the plots. Goodness-of-fit \( (R^2) \) values exceeded 0.90 for all of the sites shown. The corresponding auditory filter shapes based on the best-fit values of the auditory filter parameters are shown in the bottom row of graphs.

FIG. 4. Lower lamina 3 MUA recorded from 6 sites (3 from each animal) evoked by BF tones as a function of relative frequency separation (\( \Delta F/BF \)). Range of BFs of the sites shown is representative of the range examined in the study. FRFs of the neural populations based on MUA area from 0- to 50-ms poststimulus onset are shown in the left-hand column. Note that the range of the abscissa has been adjusted for each plot according to BF to enhance the display resolution of the FRFs. Solid vertical lines indicate the BF, the value of which is included in the plots; dotted vertical lines indicate frequencies of the inner edges of the noise bands when the \( \Delta F \) of the masker was \( \pm 0.5BF \). Normalized masking response functions, shown in the right-hand column, represent MUA as a function of relative frequency separation. Filled symbols represent peak MUA; open symbols represent MUA area. For all sites, MUA amplitude increases with increasing \( \Delta F \).
The ERB of an auditory filter is defined as $4f/p$, where $p$ is a fitting parameter of the auditory filter equation that determines both the bandwidth and the slope of the auditory filter and $f$ is the signal frequency in Hertz (Glasberg and Moore 1990; Moore 1993). ERBs (with $f/BF$) derived from the fitting parameter $p$ are indicated within the plots. Physiological ERBs increase with increasing BF.

The distribution of goodness-of-fit ($R^2$) values corresponding to the fit between physiological data at each recording site and the auditory filter equation for each of the response components examined is shown in Fig. 8. With the exception of one site for the N60 component of the AEP (represented by the filled symbol), all sites and response components displaying $R^2$ values $<0.75$ had 95% confidence intervals for the $p$ parameter of the auditory filter equation that included a value of zero. In contrast, all sites and response components displaying $R^2$ values $>0.75$ had 95% confidence intervals for the $p$ parameter of the auditory filter equation that were above zero. The boundary of $R^2$ values at 0.75 thus sharply differentiated sites at which responses were systematically modulated by $\Delta F$ (where values of $p$ were statistically different from zero, which would yield auditory filters with infinite bandwidths) from those at which responses were not systematically modulated by $\Delta F$ (where the value of $p$ was not statistically different from zero). Therefore an $R^2$ value of 0.75 was used as a criterion by which to demarcate sites showing good fits between the audi-
The rounded exponential auditory filter functions provided excellent fits to the MUA and lower lamina 3 masking response functions, with $R^2$ values exceeding 0.90 for all but one recording site (which still displayed an $R^2$ value >0.75). Although most sites displayed $R^2$ values >0.75 for the response components examined, the spread of $R^2$ values was greater for components recorded superficial to thalamorecipient zone laminae. Out of the 22 sites, 20 displayed good fits between physiological data and auditory filter functions for the upper lamina 3 sink and P28 source components of the CSD. $R^2$ values for P28 and N60 components of the superficial AEP were <0.75 for five and nine of the 22 recording sites, respectively.

Derived ERBs for all response components are plotted as a function of BF in Fig. 9. Only those sites displaying $R^2$ values >0.75 corresponding to the fit between physiological data and auditory filter functions are included in the graphs (N refers to the number of sites satisfying this criterion). Power regression lines fitted to the ERBs derived for each response component are shown superimposed on the graphs (dotted lines). The equation for the power regression line, which appears as a straight line on these log–log axes, is mathematically equivalent to a linear regression equation of the form, $\log \text{ERB} = a \log \text{BF} + b$. To examine whether the spectral resolution of A1 responses parallels psychoacoustic data, behavioral critical bandwidths in macaque monkeys (Macaca nemestrina), obtained using two-tone and narrow-band noise maskers, are plotted as a function of center frequency for comparison (square symbols; Gourevitch 1970). Monkey behavioral data are fitted by a power regression line (top solid line). The bottom solid curve represents a regression fitted to ERB values derived in psychophysical experiments using notched-noise masking in human subjects (Glazberg and Moore 1990; regression equation: $\text{ERB} = 24.7(4.37F + 1)$, where F is frequency in kHz). Physiological ERBs are significantly correlated with BF for all response components examined, in accordance with psychoacoustic findings (Pearson correlation coefficients $r$ and associated probability values $P$ [null hypothesis: $r = 0$], are included in plots of Fig. 9). Physiological ERBs display a close match to macaque psychoacoustic critical bandwidths, which are generally larger than corresponding ERBs derived in human psychoacoustic experiments.

**DISCUSSION**

The frequency resolution of neural populations in A1 derived using a two-noise–masking paradigm parallels that observed in psychophysical experiments with human and macaque subjects. Physiological ERBs increase with BF, similar to the increase in psychoacoustic critical bandwidths and ERBs.
with center frequency in humans and macaques. Physiological ERBs closely match corresponding behavioral critical bandwidths in macaques (Gourevitch 1970). Excellent fits ($R^2 > 0.90$) were obtained for 21 of the 22 sites between masking response functions for lower lamina 3 MUA and CSD and auditory filter functions used to characterize the frequency selectivity of the auditory system in human psychophysical studies.

The present results confirm parallels reported between physiological critical bandwidths obtained for single-unit responses in A1 of anesthetized cats and behavioral measures of frequency selectivity (Ehret and Schreiner 1997). The closer match observed between physiological and behavioral measures in the awake state in the present study suggests the possibility that barbiturate anesthesia or measures of single neurons contributed to the elevated physiological critical bandwidths in the studies of Ehret and Schreiner (1997). The present findings are also consistent with MEG results in humans demonstrating an increase in the amplitude of the surface-recorded N100m component with increasing ΔF. The N100 component in humans is thought to be partly generated by activity of neural populations in Heschl’s gyrus, portions of which correspond to A1 in humans (Godey et al. 2001; Howard et al. 2000; Liegeois-Chauvel et al. 1991, 1994; Lütkenhöner and Steinstrater 1998; Scherg and Von Cramon 1986; Yvert et al. 2005). These MEG findings thus suggest that responses in human A1 may reflect the frequency resolution of the auditory system. On the other hand, ERBs derived from MEG data were considerably larger than those reported in psychophysical experiments in humans for the signal frequencies tested. This suggests that the N100m component is not a sensitive index of psychoacoustic auditory frequency selectivity. Consistent with this possibility, only 13 of the 22 sites examined in the present study showed good fits between the auditory filter function and the amplitude of the N60 component of the intracortical AEP in the monkey, which is thought to be homologous to the N100 recorded in humans (Arezzo et al. 1986; Steinschneider et al. 1994; Vaughan and Arezzo 1988). The reduced fit and greater variance of derived ERB values may reflect the lower sensitivity and specificity of superficially recorded intracortical AEP components, which likely include volume conducted activity from adjacent sites. Similarly, because surface-recorded EEG and MEG components reflect the summed activation of multiple neural generators within the superior temporal gyrus (Lütkenhöner 2003; Lütkenhöner and Steinstrater 1998; Yvert et al. 2005), the frequency selectivity of the N100m component may include contributions from both primary and nonprimary auditory cortical fields, which might exhibit greater excitatory tuning bandwidths and correspondingly reduced frequency resolution (Rauschecker and Tian 2004).

Masking response functions were inversely related to the pure-tone FRFs. Specifically, whereas response amplitudes increased with increasing ΔF between the noise masker edges and the BF, pure-tone response amplitudes decreased with increasing frequency distance away from the BF. This inverse relationship indicates that the noise maskers are most effective in suppressing responses to the pure-tone signal when their spectra overlap the excitatory response area of the recorded neural population. The mechanisms underlying this response suppression cannot be determined with the present data. However, several studies suggest a possible contribution of GABAergic inhibition or a reduction in neuronal membrane input impedance to suppression of neuronal responses to sequentially and simultaneously presented sounds, both at cortical and subcortical levels (Backoff et al. 1997; Cox et al. 1992; de Ribaupierre et al. 1972; Frisina 2001; Lu and Jen 2001; Metherate 1998; Metherate and Ashe 1994, 1995a,b; Spirou et al. 1999; Tan et al. 2004; Vater et al. 1992; Volkov and Galazyuk 1992; however, see Wehr and Zador 2005). Because the noise maskers used in the present study were presented continuously while the pure-tone signal was periodically gated, whatever suppressive mechanisms are involved must extend beyond the onset response and persist throughout the duration of the noise stimulation.

Several potential limitations of the present study require consideration. The relation between perceptual and physiological measures of frequency selectivity is tempered by the fact that behavioral experiments were not conducted in parallel...
with neurophysiological recordings in the same animals. Moreover, because neural response thresholds were not determined, and stimulus intensities were kept fixed across conditions and electrode penetrations, it is possible that responses evoked by the BF signal were either “overmasked” or “undermasked” by the noise maskers at some sites. For instance, the presence of above-baseline MUA at some sites under the $\Delta F = 0\%$ condition indicates that the masker intensity used was insufficient to fully suppress the responses. Undermasking (or overmasking) may influence not only the dynamic range of the masking functions, but also the slope of the filter functions and corresponding ERB values. However, the effects of undermasking on MUA results are likely to be rather limited because tone-evoked MUA responses were generally small or negligible under the $\Delta F = 0\%$ condition.

Auditory filter bandwidths measured in psychoacoustic studies have been shown to increase markedly with increasing masker spectrum level (Moore 1993; Moore and Glasberg 1987). Thus using a different masker level adjusted to the neural response threshold at each electrode penetration introduces an additional source of variability into the data that renders comparison with psychoacoustic studies using a fixed masker level potentially problematic. Maintaining signal and masker intensities fixed across electrode penetrations and at a level corresponding to the middle of the dynamic range of masked threshold functions obtained in psychoacoustic studies thus represents a compromise between the aforementioned considerations. Finally, because the present findings are based on neuronal population measures, it is unclear to what extent they are representative of filter bandwidths of individual A1 neurons. Therefore the present results are perhaps best viewed as reflecting an integrated bandwidth of spectral resolution in A1.

It is unclear at what level of the ascending auditory pathway physiological frequency resolution begins to consistently match that measured in psychoacoustic experiments. Several studies indicate that the inferior colliculus plays a role in shaping the spectral selectivity of the auditory system to conform to behavioral measures (Ehret and Merzenich 1985, 1988). The fact that physiological ERBs for early MUA and CSD response components in the thalamorecipient zone displayed better matches to behavioral critical bandwidths than later, more superficial response components is consistent with this possibility. Thus although A1 may contribute to shaping the spectral resolution of the auditory system, the most conservative interpretation of the present findings is that the spectral resolution of A1 responses is a passive reflection of the frequency selectivity of brain stem structures and thalamic inputs to the cortex.

Moreover, it is still an open question whether the spectral resolution of A1 responses described here is used by macaques in other stimulus contexts. For example, O’Connor et al. (2000) found that rhesus macaques display a comparatively poor ability to discriminate spectral envelopes of sine-profile–modulated stimuli relative to the spectral resolution found in the present study and in behavioral experiments using two-tone masking stimuli (Gourevitch 1970; Serafin et al. 1982). This suggests that the relatively good spectral resolution demonstrated in the present study is not necessarily used by the animals to discriminate spectral profiles of other stimuli. One possible explanation for this discrepancy is that the present stimuli were characterized by sharply delineated spectra (e.g., the noise masker slopes were 96 dB/octave), whereas the sine-profile–modulated stimuli had more gradual spectral transitions. The comparatively good performance found for humans in the study of O’Connor et al. (2000) suggests that there are genuine species differences between monkeys and humans under conditions where spectral envelopes are not sharply defined. These might include differences in spectral integration bandwidths of neurons in nonprimary auditory cortical areas, which display considerably broader tuning than A1 neurons (O’Connor et al. 2000; Rauschecker and Tian 2004; Rauschecker et al. 1995). What is clear, however, is that information related to auditory spectral resolution measured in perceptual studies using stimuli with sharply delineated spectral components is available at the level of A1.

ACKNOWLEDGMENTS

We are grateful to Dr. Steven Walkley, M. Huang, L. O’Donnell, and S. Seto for providing technical, secretarial, and histological assistance. Three anonymous reviewers provided constructive comments on a previous version of the manuscript. We also thank Drs. C. J. Moore and Andrew J. Oxenham for helpful correspondence.

GRANTS

This research was supported by National Institute of Deafness and Other Communications Disorders Grant DC-00657.

REFERENCES


