Published in final edited form as:

Neuroscience. 2012 July 12; 214: 28–35. doi:10.1016/j.neuroscience.2012.04.029.

Dynamics of Phase-Independent Spectro-Temporal Tuning in Primary Auditory Cortex of the Awake Ferret

D.A. Depireux¹, H.D. Dobbins², P. Marvit³, and B. Shechter¹

¹Institute for Systems Research, University of Maryland, College Park

²Physics Education Research Group, University of Maryland, College Park

³Department of Biology, University of Maryland College Park

Abstract

Tuning of cortical neurons is often measured as a static property, or during a steady-state regime, despite a number of studies suggesting that tuning depends on when it is measured during a neuron's response (e.g., onset vs. sustained vs. offset). We have previously shown that phase-locked tuning to feature transients evolves as a dynamic quantity from the onset of the sound. In this follow-up study, we examined the phase-independent tuning during feature transients. Based on previous results, we hypothesized phase-independent tuning should evolve on the same timescale as phase-locked tuning. We used stimuli of constant level, but alternating between flat spectro-temporal envelope and a modulated envelope with well defined spectral density and temporal periodicity. This allowed the measure of changes in tuning to novel spectro-temporal content, as happens during running speech and other sounds with rapid transitions without a confounding change in sound level. For 95% of neurons, tuning changed significantly from the onset, over the course of the response. For a majority of these cells, the change occurred within the first 40 ms following a feature onset, often even around 10–20ms. This solidifies the idea that tuning can change rapidly from onset tuning to the sustained, steady-state tuning.

Keywords

Tuning dynamics; Auditory Cortex; Auditory grating; Transient Auditory Grating; Spectro-Temporal Receptive Field; Adaptive Filtering

Introduction

Neurons in auditory cortex are often characterized in a steady-state. This involves the implicit assumption that their response properties are largely static. A number of studies show these response characteristics change on various timescales: minutes (Fritz et al., 2005, Fritz et al., 2007, Gourevitch and Eggermont, 2008), hours (Ohl and Scheich, 1996, Weinberger, 1998, Weinberger and Bakin, 1998), and days to months (Weinberger et al., 1993, Mrsic-Flogel et al., 2003, Bergan et al., 2005, Chang et al., 2005, Moucha et al., 2005,

© 2012 IBRO. Published by Elsevier Ltd. All rights reserved.

Contact Information. Didier A. Depireux, Institute for Systems Research, University of Maryland, College Park, College Park, MD 20742, depireux@umd.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Turner et al., 2005, Polley et al., 2006). In many of these studies, the changes result from learning or some modification of the auditory pathway. Scarcer are studies examining how tuning arises in the first place and its stability (Shechter and Depireux, 2007, Shechter et al., 2009). While these studies describe the neural response to spectral content, they differ in analysis timescale. Some are concerned with onset responses (when a sound is turned on), while others consider the steady-state response to on-going sounds.

In a recent study (Shechter et al., 2009), we disambiguated two aspects of sound: namely, sound level and spectro-temporal content. These are normally represented concurrently in cortical responses. We defined the terms *level transient* and *feature transient* to describe changes that occur on a short timescale with respect to sound level and spectro-temporal content. The difficulty in characterizing the encoding of level transients versus feature transients lies in disambiguating the two components of the response. For instance, the introduction of a broadband sound from silence, even spectrally flat noise, induces a response because of the level transient. The poor predictability in neural responses to the onset of a complex sound is due at least in part to an entanglement of the *level transient* response with the *feature transient* response (Kowalski et al., 1996b).

In Shechter et al. (2009), we showed that *phase-locked* tuning to feature transients is not instantaneous; it develops as a function of time after the onset of a spectro-temporal feature. We found that the majority of cells in primary auditory cortex (AI) (~70%) shifted their tuning to higher spectral and temporal modulation frequencies. Similar tuning dynamics have also been show in the visual system: Bredfeldt and Ringach (2002) measured the dynamics of spatial frequency tuning in visual cortex by reverse correlation with rapidly changing spatial luminance gratings. They found that tuning in most cells becomes more selective over the course of the response, and tuning shifts from lower to higher preferred spatial frequencies.

In this follow-up study, we tested the hypothesis that the rate tuning to feature onsets in broadband stimuli is dynamic. As before, we measured tuning in auditory cortex from the *onset* of spectro-temporal features and examined whether it changed significantly during the time-course of the response. However, instead of measuring the phase-locked component to the features of the transient, we measured the change in phase-independent tuning (i.e., overall spike rate) during the feature.

Experimental Procedures

All procedures used in these experiments have been approved by and were performed in accordance with the University of Maryland Institutional Animal Care and Use Committee and are in compliance with National Institutes of Health guidelines on the care and use of laboratory animals.

An in-depth description of the surgery and data acquisition are provided in Dobbins et al. (2007) and will only be summarized here. Data were collected from 3 domestic ferrets (*Mustela putorius furo*) implanted for several months. For surgery, induction of a deep plane of anesthesia was achieved with 3.5% halothane. The level was adjusted to keep heart rate, end-tidal CO₂ and SpO₂ within limits (typically 1.75 to 2.25% maintenance). A custom-made multi-electrode array device (MEAD) was surgically implanted within a craniotomy over visualized AI. One MEAD contained 8 to 12 individually moveable 2–6 M Ω tungsten electrodes (Micro Probe, Inc, Gaithersburg, MD). The MEAD was anchored to the skull using stainless steel screws and dental cement.

After a 2–4 hour period for recovery from surgery, the electrodes were lowered through the dura for anchoring, as measured by a sudden drop in the impedance of each electrode.

Neural recordings began three days later. For each daily recording session, the animals were placed into a custom made holder inside a sound proof booth (IAC). The electrodes were lowered individually until cells were found. Then, stimuli (described below) were presented free field from an overhead speaker (Manger Transducer, Manger, Germany) located at zenith relative to the animal's head. Sounds were delivered using a TDT sound delivery system while neural activity was recorded with a Neuralynx recording system (Neuralynx, Tucson AZ).

Neural activity was recorded and assigned to single neurons in two steps. During the recording, the electrode signal was band-pass filtered with low and high cutoff frequencies of 300Hz and 3kHz, respectively. Events were captured when the amplitude exceeded a threshold derived from the average power of the recorded signal; this threshold was set low enough to capture all spikes, but it also captured large excursions of the evoked potential. Event times were assigned by position of the peak. After a recording session, spikes were sorted into single-unit classes and a “miscellaneous” class using a modification of the MClust package with the automated cutter KlustaKwik (Harris and Redish, 2004). KlustaKwik uses a CEM algorithm (*Conditional Expectation Maximization*) for which we used as inputs the Fourier transform, first and second principal components and energy of each event to classify spikes. Our low threshold and conservative sorting typically yielded a large number of rejected events, which were placed in the “miscellaneous” class and not considered neural spikes (and therefore not analyzed here).

Auditory Stimuli

The transient auditory grating stimuli used in this study are a modification of the auditory grating we have previously used for estimating the steady-state spectro-temporal receptive field (Depireux et al., 2001). Briefly, auditory gratings are periodic and broadband stimuli with a spectro-temporal profile modulated sinusoidally in log-spectrum and in time. The modulation of a grating is characterized in spectrum by its spectral density (Ω , cycles/octave), in time by its temporal periodicity (w , Hz), and in amplitude by its excursions away from the mean level of the stimulus (modulation depth ΔA). In a grating stimulus, the amplitude $S(x, t)$ of each tone component is given by

$$S(x, t) = L[1 + \Delta A \cdot \cos(2\pi(\Omega \cdot x + w \cdot t) - \phi)], \quad (1)$$

for linear modulation. The frequency of each tone that makes up the grating is given by x , where $x = \log_2(f/f_0)$; f_0 is the lower edge of the spectrum. L is related to the intensity of the stimulus and ϕ is the starting phase of the grating. When both Ω and w are positive, the envelope drifts towards the low frequencies. The tones f that make up the grating are logarithmically spaced, so that the pitch of the overall stimulus is indeterminate.

Transient Auditory Gratings—We used transient grating stimuli to probe the dynamics of tuning to spectro-temporal content on a millisecond timescale following feature onsets (Shechter et al., 2009). These stimuli alternate between robust spectro-temporal features (auditory gratings) and flat spectrum broadband noise, but have an almost constant sound level throughout. With the 100% modulation depth used in this study, the level changed by 3dB. This enabled us to dissociate the responses to feature transients as opposed to level transients (which often occur concurrently), as there were only weak level transients during the sound. As mentioned earlier, a typical cortical cell will respond to changes in level, regardless of its preferred frequency content (i.e. how it responds to auditory features).

A typical stimulus spectro-temporal envelope is flat except for eight 50 ms intervals (transients) of modulation randomly distributed throughout the stimulus duration. Each transient consisted of 50 ms of an auditory grating with specific spectral density, temporal

periodicity and starting phase. In a given waveform, the 8 transients had the same density and temporal periodicity, but starting phases were chosen from a random permutation of $\{2\pi \cdot x/8, x = 0, 1, \dots, 7\}$. Taken together, these 8 transients consist in at least 2 full cycles of gradient modulation (Shechter et al., 2009). Random inter-transient interval (ITI) durations were chosen from a normal distribution with a mean of 150 ms and a standard deviation of 50 ms, hard-limited between 75 ms and 225 ms. The first transient began 50 ms after the stimulus onset, and the ITIs were used to determine when subsequent transients began. A 3 ms ramp was applied to the onset and offset of the spectro-temporal transient envelope, to avoid the perception of a click at the beginning and the end of the transients. In other words, the modulation depth changed between 0 and its maximal value over a period of 3 ms. An example spectrogram of a transient auditory grating envelope is illustrated in Fig. 1.

These spectro-temporal envelopes were used to determine the amplitude of 100 tones per octave over 5 octaves as a function of the time. These carrier tones were in random temporal phase. The tones were added together to form a sound of almost constant power with no level onsets, but with a series of well-defined spectro-temporal feature transients. There were 63 transient grating stimuli: Spectral density of the transients ranged from -2 cyc/oct to 2 cyc/oct in steps of 0.5 cyc/oct and temporal periodicity from 0 Hz to 30 Hz in steps of 5 Hz; all transients were modulated with 100% modulation depth. All 63 sounds were played ten times for each recording position, with a variable interstimulus interval from 300ms to 900ms, at a mean sound level typically between 60 and 75 dB SPL.

Measuring Tuning Dynamics

Spike rates were measured in eight consecutive, non-overlapping 10ms windows, beginning at the onset of each grating feature ($\tau_1 = 0-10$ ms, $\tau_2 = 10-20$ ms, ...). Recall that each waveform had grating features of a given w - Ω combination, and that the eight grating segments per waveform correspond to eight different onset phases. The spike counts for each lag τ_i were summed across phases. Thus, for every w - Ω combination, we had 8 measurements of the spike rate at 8 different lags in which to examine the preference for different stimulus parameters after feature onset (illustrated in Fig. 1). This resulted in a spike rate for each combination of w , Ω , and for each lag —i.e., the response was a function $r(w, \Omega, \tau)$.

Combining positive and negative Ω responses

Visual inspection indicated that for the majority of cells, the response (i.e. total number of spikes summed across all phases), was symmetric about the $\Omega = 0$ cyc/oct axis. In other words, the spike rate was the same for upward and downward moving spectral profiles, which correspond to negative and positive Ω s respectively. We therefore quantified the symmetry about the $\Omega = 0$ cyc/oct axis. For the 93% of cells where the response was symmetric, we combined the spike rates for $+\Omega$ and $-\Omega$. Those remaining 7% of cells which were not symmetric were not included in this study.

To assess symmetry, we computed at each lag the cross-correlation between the responses to up-moving gratings and the responses to down-moving gratings, normalized by their auto correlations. In order to determine significance of the correlation, we computed a surrogate distribution by shuffling the responses to down-moving gratings 50 times. 50 surrogate correlations were then computed between the shuffled down-going responses and the original up-going responses, normalized by their auto-correlations. If the true correlation was greater than one standard deviation of its surrogate mean, it was deemed significant and we thus considered the response symmetric.

Determining tuning significance

We next determined which responses were tuned to the transient gratings. For a given lag τ , the tuning was significant if there was any transient grating which elicited a spike rate at lag τ that was significantly above the baseline rate. Since we were not interested in the response to level or intensity transients, the baseline rate was computed during the flat-spectrum noise portion of the stimulus. We computed the mean rate in the 10ms preceding each of the 8 feature onsets. Since we presented 63 stimuli at 10 sweeps each, we obtained 630 measurements of the background spike rate. We fit this distribution to a Gaussian curve, which we call the background Gaussian.

We next determined a response Gaussian: From the 10 sweeps we presented per stimulus, we computed 500 bootstrap estimates of the response, which were then fit to a Gaussian curve. This curve represents the response distribution at lag τ following the onset of a spectro-temporal feature of spectral density Ω and temporal periodicity w .

We measured the area over which the two Gaussians overlapped (OL), such that the overlap values had the range $0 < OL \leq 1$. When the background and response Gaussians completely overlapped ($OL = 1$), there was no appreciable response that could be distinguished from background. Smaller OL values meant a greater the separation between the response and background; any response for which $OL \leq 0.05$ was called significant.

Determining tuning dynamics

To determine how the tuning to feature transients, measured as a spike rate, evolved as a function of lag, we tracked those spectral densities and temporal periodicities eliciting the maximal response at each lag τ . Since the response is subject to trial by trial variability, we needed to determine how stable peak response was (in cortex, and particularly in the awake animal, repeated presentations of the same stimulus can elicit a variable number of spikes with varying timing of those spikes).

To assess stability of the peak response, we computed 10 jittered responses. We first fit the 10 spike count measurements (from 10 sweeps) using a Poisson distribution with $\lambda = \text{mean spike count}$. This Poisson distribution was then used to compute 10 additional “jittered” responses, such that the mean rate of the process was consistent with the actual data. These jittered responses enabled us to determine whether a change in the w - Ω location of the peak response may have occurred due solely to the variance in the Poisson process. We considered spectral density tuning and temporal periodicity separately, and computed the mean and standard deviation of the peak position in w and in Ω based on the jittered responses.

For a change in peak position to be significant, it had to be at least 2 standard deviations greater than the change in jittered peak response position. In other words, if the mean change in peak w response between the jittered sets was 5 Hz with a standard deviation of 2.5 Hz, then the actual peak had to change its w coordinate more than 10 Hz to be considered significant. Based on this measure of significance, we examined two values: 1) The amount of time for a significant change in tuning after the first significant response, and 2) The amount of change in peak position between the two delays with the maximal responses (measured as the total response power).

Results

We present data from 103 cells for which we found significant, symmetric tuning to the transient gratings. Fig. 1 shows a raster plot of the neural response to one of the transient auditory gratings.

Tuning Onsets

We measured tuning onset as the first lag τ for which there was a significant response above background (i.e., $OL = 0.05$). The majority of cells (59%, 63 cells) had an onset lag between 20 – 40ms. We show the distribution of onset lag in Fig. 2. We then categorized the responses based on whether the peak location changed with respect to lag following the onset of the feature. For the vast majority of cells (76%, $n=78$), the peak response changed in both w and Ω . For another 14% of cells ($n=15$), the peak changed only in w , whereas for 5% of cells ($n=5$), the peak changed only in Ω . Only 5% of cells ($n=5$) did not show tuning dynamics during the response to the transient gratings. We show the lag-dependent responses to the transient gratings for a number of cells demonstrating these tuning dynamics (Fig. 3).

We next examined the time interval between the first significant response and the first significant change in tuning with respect to peak location (Fig. 4). Most cells' tunings changed within 40 ms or less. In fact, of the cells for which w -tuning changed (irrespective of whether there was a change in peak- Ω), this change occurred in under 40ms for 95% of cells; similarly for cells with change in Ω -tuning, this occurred in under 40ms for 82%. The magnitude of these changes is depicted in Fig. 4B.

We next measured the change in peak tuning at the two lags with maximal responses. Maximal response was defined in terms of the overall power of the response. Fig. 5 shows the lag at which maximum response occurred (Fig. 5A) and the time interval to the penultimate maximum (Fig. 5B). Not surprisingly, the two maximal responses often occur at consecutive lags (88%, $n=93$). We found that for 30% of cells ($n=32$), the tuning changed both in w and Ω ; 37% of cells ($n=39$) only changed tuning in w ; and 3% of cells ($n=3$) only changed tuning in Ω . Tuning did not significantly change between maximum and penultimate maximum responses for the remaining 30% of cells ($n=32$).

Discussion

In this study, we tested the hypothesis that in auditory cortex, phase-independent tuning to fixed spectro-temporal content, measured by spiking rate, dynamically evolves as a function of time elapsed, or lag, from the introduction of novel spectro-temporal content (independent of any confounding changes in sound level). We have previously shown that *phase-locked* tuning is dynamic following such a change in spectro-temporal content (Shechter et al., 2009). This temporal dynamics of tuning is rapid, and can be measured over timescales of tens of milliseconds. By extension, the cortical response to any complex sound of varying spectro-temporal content reflects a dynamic process. Specifically, we found that, for the majority of cells, tuning to the spectro-temporal content of a sound changes significantly within 40 ms (and often less) of the introduction of new spectro-temporal content. The two most important observations are that: 1) there is a significant tuning within 20–40 ms of the feature onset, and 2) the tuning subsequently changes significantly for 95% of the cells. Over 80% of the cells that changed their tuning did so within 40 ms of the onset of new spectro-temporal content. Interestingly, we did not find any significant differences in the evolution of tuning as a function of recording depth. These results further suggest that tuning of cortical cells can be selective for different auditory features in the onset and steady-state components of an on-going auditory stimulus.

Previous studies characterizing the response of auditory cortex neurons (e.g., Kowalski et al. (1996b, 1996a), Klein et al. (2000)) specifically discarded the first 100 ms or so of the "onset" response, because of the confounding change in level. (Indeed, these studies were unable to use their characterization to successfully predict the response at stimulus onset). In this respect, they characterized only the steady-state response to an ongoing sound of

unchanging spectro-temporal content, lasting several seconds. The onset response thus masks any response to feature transients. There is also the possibility of forward suppression: however, we found the magnitude of the onset response (initial 50ms of noise) to be significantly greater than the magnitude of the response to the transients. This is unlikely due to forward masking, as the response magnitudes to the transients in aggregate were statistically indistinguishable from each other as a function of their ordering. The response during the intervening zero-modulation segments (i.e., just the noise carrier, or zero-modulation part of the stimulus) was smaller in magnitude compared to that of the transients. Forward masking would have manifested itself as a lower magnitude response to the first transient than to later transients.

At some basic level, the modulations present in the features transients we used can be viewed as level transients on a very local scale. This view is important because AI neurons integrate acoustic features over a comparatively narrow range of frequency. However, the transition from complete silence to stimulus is much more drastic than the transition from flat to modulated noise—even locally. In addition, level onsets—even at a local level—are often fast transitions from silence to stimulus level, and therefore possess energy at all frequencies (the so-called spectral splatter). Such a broad frequency range would evoke network wide activity, thereby masking any response to local features. The feature transitions we study are continuous, and therefore, well constrained in frequency (even if they very locally change the level).

Other studies (deCharms et al., 1998, Valentine and Eggermont, 2004) derived receptive fields from sounds whose spectro-temporal content was continuously and rapidly changing, making it hard to dissociate dynamics in tuning from changes in the stimulus (since the stimulus did not contain any periods of fixed spectro-temporal content). Our results imply that the receptive fields estimated using continuously changing content ought to be different from receptive fields obtained using the auditory grating or ripple method. Experiments using both methods on the same neurons should further elucidate the possible multiplicity of neural tunings to novel versus slowly- or non-changing spectro-temporal content of broadband sounds, and the transitions between the different regimes that occur over short timescales.

It has been shown that tuning can change on many different time scales. In its most basic, our measurement of dynamics must relate to tuning onset latency. However, most measurement of latency are with respect to sound presentation from silence and not specifically to the feature transients reported on here (Qin et al., 2003, Sakai et al., 2009). While previous studies have differentiated regimes of tonal response based on latency (i.e., onset vs sustained vs offset), they do not look directly at dynamics of tuning within these regimes, nor at the encoding of spectro-temporal features (Bizley 2005, Wang 2005). As a reflection of cortical plasticity, there are many permanent tuning changes that take place over the course of days to months (Weinberger et al., 1993, Mrsic-Flogel et al., 2003, Bergan et al., 2005, Chang et al., 2005, Moucha et al., 2005, Turner et al., 2005, Polley et al., 2006, Atiani et al., 2009). On shorter time scales, Ohl and Scheich (1996) showed that conditional training can change the mean value of a cortical evoked response over a few hours, and that the response changes are stable long-term. Fritz et al. (2005, 2007, Atiani et al., 2009) demonstrated that the steady-state receptive field can change over the course of minutes due to behavioral and attentional effects. We have previously shown that even without external influences, the phase-locked tuning changes from its onset (Shechter et al., 2009) and—to a lesser degree—even during its steady-state (Shechter and Depireux, 2007). Gourevitch and Eggermont (2008) also show such adaptive changes to occur over the course of minutes. In this study, we establish that the measured rate response, phase-independent tuning changes over the course of the response within milliseconds.

We previously showed that a class of cells in the central nucleus of the inferior colliculus responds with a lagged temporal response profile, which might occur as a result of cortical feedback (Lim and Anderson, 2006, Simon et al., 2007, Shechter et al., 2010). This dynamic loop of processing between subcortical and cortical stations could give rise to the dynamics of the tuning that follows the first cortical response as we describe here. Onset tuning in AI would be a result of feed forward propagation of processing; as cortical feedback influences collicular responses, these collicular responses would in turn modify tuning in AI. In this respect, dynamics of tuning could arise as a direct result of cortical feedback on lower auditory stations (Polley et al., 2006, Bajo et al., 2007). This in turn implies that the tuning of collicular neurons might also evolve in a similar fashion to what we measure in AI. The collicular source of dynamics could be investigated under reversible inactivation of the cortico-collicular feedback pathways.

In this study, we found most cells respond symmetrically in their phase-independent response to feature transients. The DC spike rate can be spectrally symmetric, while the modulation following response is not. In Shechter et al (2009), we showed the phase-dependent response is spectrally symmetric at the onset of a feature transient. Thus from Shechter et al, we conclude that the spectral asymmetry as observed in the steady state (Depireux et al., 2001), develops after a certain lag after any feature onset. This might suggest that direction of envelope is coded by spike timing solely (Shechter et al., 2009), whereas spectro-temporal tuning is encoded by both rate and timing.

The summed evidence indicates that we must think of cortical tuning as a complex and time-evolving process, rather than a static and time-independent property of the cell. Previous studies in the awake auditory cortex in response to pure tone showed evidence of a sustained cortical response separate from the onset response with unique properties (Bendor and Wang, 2005), and this study demonstrates the non-static nature of cortical tuning. In the visual cortex, Bredfeldt and Ringach (2002) similarly demonstrated the existence of dynamics in the tuning to spatial frequency during peri-stimulus time: Tuning becomes more selective over the course of the response, and the preferred spatial frequency shifts from low to higher frequencies. This reflects the design of an ideal filter, which first detects a change in energy and subsequently identifies that change. These changes were shown to happen over tens of milliseconds. Mazer et al. (2002) showed in addition that orientation tuning changes on the same timescales. The evoked firing of a neuron depends not only on the spectral content of the stimulus presented, but very much also its contextual location (i.e., temporal position) within the stimulus.

The primary auditory cortex plays a key role in the representation of auditory objects (Nelken et al., 2003) such as speech. Some studies have found cortical correlates of features typical of speech such as voice onset time (Steinschneider et al., 1994, Eggermont, 1995), and it has been shown that cortex plays an important role in the categorization of phonemes (Guenther et al., 2004, Poeppel et al., 2004, Luo et al., 2005). In addition, very short duration changes at the onset of a sound can greatly alter psychophysical perception (Denham and Winkler, 2006).

In humans, there is increased sensitivity to masking following auditory cortex strokes. Tramo et al. (2002) describe “heightened sensitivity to masking” in bilateral AI lesions. Their descriptions strongly suggest that patients with damage to AI have difficulty extracting features typically found in casual speech, implying that AI therefore allows for a unique, noise-robust time-frequency representation of complex sounds. It naturally follows that AI has complex tuning dynamics with characteristics of adaptive filtering. Models of auditory processing, especially at the level of auditory cortex, (including our current and earlier results (Shechter et al., 2009)) should thus be interpreted in such a context.

Highlights

Neural tuning in auditory cortex has dynamics over short time scales
 We show tuning is different between onset and sustained part of a sound
 Tuning has dynamics over tens of ms, when presented new spectro-temporal content
 There is an interesting region of *dynamical tuning* post spectro-temporal transients
 We need to revisit our concept of tuning in continuous cases such as running speech

Acknowledgments

We thank Yadong “KK” Ji for extensive help in animal care and data acquisition. This research was funded by NIH/NIDCD 1 RO1 DC005937 awarded to DAD. PM also received support from training grant NIH/NINDS 2T32NS007375-11

Abbreviations

AI	Primary auditory cortex
ITI	inter-transient interval
MEAD	multi-electrode array device
OL	overlap value

References

- Atiani S, Elhilali M, David SV, Fritz JB, Shamma SA. Task difficulty and performance induce diverse adaptive patterns in gain and shape of primary auditory cortical receptive fields. *Neuron*. 2009; 61:467–480. [PubMed: 19217382]
- Bajo VM, Nodal FR, Bizley JK, Moore DR, King AJ. The Ferret Auditory Cortex: Descending Projections to the Inferior Colliculus. *Cereb Cortex*. 2007
- Bendor D, Wang X. The neuronal representation of pitch in primate auditory cortex. *Nature*. 2005; 436:1161–1165. [PubMed: 16121182]
- Bergan JF, Ro P, Ro D, Knudsen EI. Hunting increases adaptive auditory map plasticity in adult barn owls. *J Neurosci*. 2005; 25:9816–9820. [PubMed: 16237185]
- Bredfeldt CE, Ringach DL. Dynamics of spatial frequency tuning in macaque V1. *J Neurosci*. 2002; 22:1976–1984. [PubMed: 11880528]
- Chang EF, Bao S, Imaizumi K, Schreiner CE, Merzenich MM. Development of spectral and temporal response selectivity in the auditory cortex. *Proc Natl Acad Sci U S A*. 2005; 102:16460–16465. [PubMed: 16263924]
- deCharms RC, Blake DT, Merzenich MM. Optimizing sound features for cortical neurons. *Science*. 1998; 280:1439–1443. [PubMed: 9603734]
- Denham SL, Winkler I. The role of predictive models in the formation of auditory streams. *J Physiol Paris*. 2006; 100:154–170. [PubMed: 17084600]
- Depireux DA, Simon JZ, Klein DJ, Shamma SA. Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex. *J Neurophys*. 2001; 85:1220–1234.
- Dobbins HD, Marvit P, Ji Y, Depireux DA. Chronically recording with a multi-electrode array device in the auditory cortex of an awake ferret. *J Neurosci Methods*. 2007; 161:101–111. [PubMed: 17134761]
- Eggermont JJ. Representation of a voice onset time continuum in primary auditory cortex of the cat. *J Acoust Soc Am*. 1995; 98:911–920. [PubMed: 7642830]

- Fritz JB, Elhilali M, David SV, Shamma SA. Does attention play a role in dynamic receptive field adaptation to changing acoustic salience in A1? *Hear Res.* 2007; 229:186–203. [PubMed: 17329048]
- Fritz JB, Elhilali M, Shamma SA. Differential dynamic plasticity of A1 receptive fields during multiple spectral tasks. *J Neurosci.* 2005; 25:7623–7635. [PubMed: 16107649]
- Gourevitch B, Eggermont JJ. Spectro-temporal sound density-dependent long-term adaptation in cat primary auditory cortex. *The European journal of neuroscience.* 2008; 27:3310–3321. [PubMed: 18598269]
- Guenther FH, Nieto-Castanon A, Ghosh SS, Tourville JA. Representation of sound categories in auditory cortical maps. *J Speech Lang Hear Res.* 2004; 47:46–57. [PubMed: 15072527]
- Harris K, Redish AD. MClust: a spike-sorting toolbox, freely available software. 2004
- Klein DJ, Depireux DA, Simon JZ, Shamma SA. Robust spectrotemporal reverse correlation for the auditory system: optimizing stimulus design. *Journal of Computational Neuroscience.* 2000; 9:85–111. [PubMed: 10946994]
- Kowalski N, Depireux DA, Shamma SA. Analysis of dynamic spectra in ferret primary auditory cortex. I. Characteristics of single-unit responses to moving ripple spectra. *J Neurophysiol.* 1996a; 76:3503–3523. [PubMed: 8930289]
- Kowalski N, Depireux DA, Shamma SA. Analysis of dynamic spectra in ferret primary auditory cortex. II. Prediction of unit responses to arbitrary dynamic spectra. *J Neurophysiol.* 1996b; 76:3524–3534. [PubMed: 8930290]
- Lim HH, Anderson DJ. Antidromic activation reveals tonotopically organized projections from primary auditory cortex to the central nucleus of the inferior colliculus in guinea pig. *J Neurophysiol.* 2006
- Luo H, Husain FT, Horwitz B, Poeppel D. Discrimination and categorization of speech and non-speech sounds in an MEG delayed-match-to-sample study. *Neuroimage.* 2005; 28:59–71. [PubMed: 16023868]
- Mazer JA, Vinje WE, McDermott J, Schiller PH, Gallant JL. Spatial frequency and orientation tuning dynamics in area V1. *Proc Natl Acad Sci U S A.* 2002; 99:1645–1650. [PubMed: 11818532]
- Moucha R, Pandya PK, Engineer ND, Rathbun DL, Kilgard MP. Background sounds contribute to spectrotemporal plasticity in primary auditory cortex. *Exp Brain Res.* 2005; 162:417–427. [PubMed: 15616812]
- Mrsic-Flogel TD, Schnupp JW, King AJ. Acoustic factors govern developmental sharpening of spatial tuning in the auditory cortex. *Nat Neurosci.* 2003; 6:981–988. [PubMed: 12910241]
- Nelken I, Fishbach A, Las L, Ulanovsky N, Farkas D. Primary auditory cortex of cats: feature detection or something else? *Biol Cybern.* 2003; 89:397–406. [PubMed: 14669020]
- Ohl FW, Scheich H. Differential frequency conditioning enhances spectral contrast sensitivity of units in auditory cortex (field A1) of the alert Mongolian gerbil. *Eur J Neurosci.* 1996; 8:1001–1017. [PubMed: 8743748]
- Poeppel D, Guillemin A, Thompson J, Fritz J, Bavelier D, Braun AR. Auditory lexical decision, categorical perception, and FM direction discrimination differentially engage left and right auditory cortex. *Neuropsychologia.* 2004; 42:183–200. [PubMed: 14644105]
- Polley DB, Steinberg EE, Merzenich MM. Perceptual learning directs auditory cortical map reorganization through top-down influences. *J Neurosci.* 2006; 26:4970–4982. [PubMed: 16672673]
- Qin L, Kitama T, Chimoto S, Sakayori S, Sato Y. Time course of tonal frequency-response-area of primary auditory cortex neurons in alert cats. *Neuroscience research.* 2003; 46:145–152. [PubMed: 12767477]
- Sakai M, Chimoto S, Qin L, Sato Y. Differential representation of spectral and temporal information by primary auditory cortex neurons in awake cats: relevance to auditory scene analysis. *Brain research.* 2009; 1265:80–92. [PubMed: 19368805]
- Shechter B, Depireux DA. Stability of spectro-temporal tuning over several seconds in primary auditory cortex of the awake ferret. *Neuroscience.* 2007; 148:806–814. [PubMed: 17693032]
- Shechter B, Dobbins HD, Marvit P, Depireux DA. Dynamics of spectro-temporal tuning in primary auditory cortex of the awake ferret. *Hear Res.* 2009; 256:118–130. [PubMed: 19619629]

- Shechter B, Marvit P, Depireux DA. Lagged cells in the inferior colliculus of the awake ferret. *Eur J Neurosci.* 2010; 31:42–48. [PubMed: 20092554]
- Simon JZ, Depireux DA, Klein DJ, Fritz JB, Shamma SA. Temporal symmetry in primary auditory cortex: implications for cortical connectivity. *Neural Comput.* 2007; 19:583–638. [PubMed: 17298227]
- Steinschneider M, Schroeder CE, Arezzo JC, Vaughan HG Jr. Speech-evoked activity in primary auditory cortex: effects of voice onset time. *Electroencephalogr Clin Neurophysiol.* 1994; 92:30–43. [PubMed: 7508851]
- Tramo MJ, Shah GD, Braidia LD. Functional role of auditory cortex in frequency processing and pitch perception. *J Neurophysiol.* 2002; 87:122–139. [PubMed: 11784735]
- Turner JG, Hughes LF, Caspary DM. Affects of aging on receptive fields in rat primary auditory cortex layer V neurons. *J Neurophysiol.* 2005; 94:2738–2747. [PubMed: 16000522]
- Valentine PA, Eggermont JJ. Stimulus dependence of spectro-temporal receptive fields in cat primary auditory cortex. *Hear Res.* 2004; 196:119–133. [PubMed: 15464309]
- Weinberger NM. Physiological memory in primary auditory cortex: characteristics and mechanisms. *Neurobiol Learn Mem.* 1998; 70:226–251. [PubMed: 9753599]
- Weinberger NM, Bakin JS. Learning-induced physiological memory in adult primary auditory cortex: receptive fields plasticity, model, and mechanisms. *Audiol Neurotol.* 1998; 3:145–167. [PubMed: 9575382]
- Weinberger NM, Javid R, Lapan B. Long-term retention of learning-induced receptive-field plasticity in the auditory cortex. *Proc Natl Acad Sci U S A.* 1993; 90:2394–2398. [PubMed: 8460150]

Figure 1. Measuring the response to transient auditory gratings

(Top) Spectro-temporal envelope of a transient grating stimulus. The stimulus has a flat envelope with eight 50 ms transient segments of grating modulation interspersed. Each transient has the same density (Ω) and temporal periodicity (w), but starts at a different phase (ϕ). (Middle) The response raster of a neuron to the stimulus shown at top, aligned to the stimulus onset time: each line corresponds to one sweep and each dot corresponds to a spike event. We overlay the peri-stimulus time histogram (gray line). (Bottom) We count the number of spikes in 8 consecutive, non-overlapping 10ms windows beginning at the onset of each grating modulation ($\tau_1, \tau_2, \tau_3 \dots$). The spike counts are summed for corresponding windows across all 8 phases of the grating modulation. The resulting spike counts are shown in the plots at right, where each plot corresponds to a different lag τ after the feature onset.

Figure 2. Distribution of Tuning Onset Lag

Tuning onset is defined as the first lag for which at least one stimulus elicited a response that was significantly above background ($OL = 0.05$). The majority of cells (59%, 63 cells) had an onset lag between 20 – 40ms.

Change in tuning to spectral density (Ω) and temporal periodicity (w) Change in tuning only to temporal periodicity (w) Change in tuning only to spectral density (Ω) **Figure 3. Examples of Tuning Dynamics**

Phase-independent response for 5 example cells as a function of lag after feature onset. Each row corresponds to a single cell, where the first three cells showed dynamics in both preferred spectral density and preferred temporal periodicity. The next two cells showed dynamics only in preferred temporal periodicity or preferred spectral density, respectively. Each column corresponds to a different lag τ (determined by the window in which spikes were counted after the feature onset). The significant maxima, after collapsing down- and up-moving envelopes (positive and negative are shown with a white circle). For clarity, the axes are labeled for a single plot, but are the same across plots.

Figure 4. Timescale of Tuning Dynamics

We measured the time interval between the tuning onset and the first significant change in tuning with respect to preferred temporal periodicity (**A**) and preferred spectral density (**B**). Strikingly, 70% of cells showed dynamics in their response to temporal periodicity on the order of 10–20 ms. While the timescale of tuning dynamics with respect to spectral density were broader, changes in spectral tuning still occurred mostly within 40 ms or less. The magnitudes of these changes in preferred temporal periodicity and spectral density are shown in **C** and **D**, respectively.

Figure 5. Lag to Maximum Response

We measured the change in tuning at the two lags with maximal responses. Maximal response was defined in terms of the overall power of the response. **A.** Distribution of lags τ at which the maximal response was obtained. **B.** Distribution of time interval between maximal and penultimate maximum responses. The two maximal responses often occur at consecutive lags (88%, $n=93$). We compared the tuning at the two lags with maximal responses and found that for 30% of cells ($n=32$), the tuning changed both in w and Ω ; 37% of cells ($n=39$) only changed tuning in w ; and 3% of cells ($n=3$) only changed tuning in Ω . Tuning did not significantly change between maximum and penultimate maximum responses for the remaining 30% of cells ($n=32$).