

Cerebral Speech Reorganization Effects on Cortical Processing

Ross K. Maddox^{1,2*}, Cyrus P. Billimoria^{1,2}, Ben P. Perrone^{1,2}, Barbara G. Shinn-Cunningham^{1,3}, Kamal Sen^{1,2}

¹ Hearing Research Center, Department of Biomedical Engineering, Boston University, Boston, Massachusetts, United States of America, ² Center for Biodynamics, Boston University, Boston, Massachusetts, United States of America, ³ Center for Computational Neuroscience and Neural Technology, Boston University, Boston, Massachusetts, United States of America

Abstract

What is spatial tuning in auditory cortex weak, even though location is important to object recognition in natural settings?

Received September 20, 2011; Accepted March 20, 2012; Published May 1, 2012

Copyright: © 2012 Maddox et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by a National Institute of Deafness and Communication Disorders grant (R01 DC007610), and by CELEST, a National Science Foundation Science of Learning Center (NSF SMA-0835976). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abbreviations: SNR, signal-to-noise ratio

* E-mail: rkmaddo@bu.edu

Introduction

Past studies of spatial effects in auditory cortex have focused on how spatial location is encoded. These studies typically find that single-unit spatial tuning in cortex is weak [1–4], not topographically organized [5,6], and not encoded independently of other perceptually important features [7]. There is good evidence for a specialized

coding scheme may provide a substrate for spatial auditory attention, as top-down modulatory control signals could selectively suppress responses of neurons favoring a masker in order to reduce competition and allow more precise analysis of a target from a desired location.

Results

Neural Responses Are Sensitive to the Locations of Competing Sources

We recorded neural responses from male zebra finches in the auditory forebrain (field L, based on stereotactic coordinates [17–19]) to stimuli from four azimuthal locations in the frontal hemifield. Target stimuli were two conspecific songs, presented either in quiet (“clean”; Figure 1A) or in the presence of a spectrally similar noise masker coming from the same or a different location as the target song (Figure 1B). We assessed neural performance using a single-trial spike-distance-based [20] nearest-neighbor classification scheme [21], calculating a percent correct

recording electrode was in the left hemisphere, loudspeaker locations were on the left side (-90° , ipsilateral to the electrode), in front (0°), halfway between front and right ($+45^\circ$), and on the right ($+90^\circ$, contralateral). These locations were flipped about the midline when recording in the right hemisphere. Henceforth, coordinates are referenced to the recording electrode, so that ipsilateral azimuths have negative signs and contralateral azimuths have positive signs. Discrimination performance was calculated for all 16 configurations and three signal-to-noise ratios (SNR; -6 dB, 0 dB, $+6$ dB). To assess the extent to which the head created an acoustical obstruction (“head shadow”) to the ear opposite the sound source, we measured sound level at both ears from all four locations using a masker token as the probe stimulus. The differences between left and right ears were 1.5 , 0.1 , -0.8 , and -1.3 dB for -90 , 0 , $+45$, and $+90^\circ$, respectively.

For the example site in Figure 1A and B, clean performance was near ceiling at all tested locations. Masked performance was much lower and varied substantially as the target was moved from the ipsilateral side (-90°) to the contralateral side ($+90^\circ$), holding the masker at -90° . Across recording sites, the masked performance varied much more than clean performance did as a function of location. To quantify this, we computed the spatial sensitivity (defined as the difference between the best and worst performance for a given experimental condition; see Materials and Methods) for each site for both clean and masked targets. Spatial sensitivity was 3-fold higher with a masker present than without ($p < .001$; Figure 1D). The driven spike rate in response to clean songs did not vary significantly with location ($r = .16$,

during syllables; see Figure 4A–C). Both types of interference have been studied before [18]; here we extended that analysis. We modeled spike trains that had only subtractions or only additions (Figure 4D; see Materials and Methods), and then calculated performance for these modeled spike trains just as we did for the measured ones.

We first validated our modeling approach by comparing predictions for modeled spike trains containing both additions and subtractions (i.e., the full effect of the masker) to measured data (see Figure 4, “modeled” rasters and performances). The example model rasters look similar to the measured masked spike trains, and target song identification performance closely matched performance using the masked spike trains. These results validate our methods for modeling additions and subtractions.

Following validation, we modeled spike trains that included only spike additions or only spike subtractions to separate their relative effects on performance. When modeling spike additions only (i.e., when no subtractions were modeled), target identification was better than for the measured response. On the other hand, performance for subtractions-only spike trains was only slightly better than the measured responses for two of the three configurations. For the target-contralateral, masker-ipsilateral configuration (right column of Figure 4), performance was essentially equal for the subtractions-only and masked spike trains. These results suggest that additions did not impair discrimination performance when the target was contralateral to the recorded site. However, including additions had some impact on the other two configurations. Overall, this analysis shows that the masker degraded performance more by preventing spikes that a clean target would have elicited than by causing additional spikes.

The times at which spikes are likely to be added by the masker tend to occur when the clean response rates are low. This can be quantified by correlating the clean stimulus response rate (Figure 4A) with the rate of subtractions (blue depths in Figure 4C) as a function of time. This correlation is significant and negative, confirming that subtractions reduce spikes the most when the likelihood of a spike in response to the clean stimulus is great ($r = -.75$, $p < .001$). In contrast, the correlation between the time-dependent spike additions (red peaks in Figure 4C) and the clean rate is weak ($r = .08$, $p < .001$). Taken together, these results suggest that the effect of removing spikes from the peaks interferes with target identification more than adding spikes. This holds true even in spatial configurations where the number of spikes added is greater than the number of spikes removed.

Discussion

Specific Experimental Paradigms Unveil Inherent Spatial Sensitivity

Here we show that, in quiet, sound source location has only a modest impact on coding of song identity in field L, an analog of auditory cortex [16]. In general, spatial tuning in brainstem is sharper than in cortex, demonstrating that cortical auditory neurons do not directly inherit the already encoded spatial information present in lower centers of the auditory processing stream [1,28]. However, our results show that the spatial configuration of competing sources strongly affects the coding of those sources' content.

Spatial effects in cortical neurons are far greater when there are competing sounds than when there is only a single source. This observation suggests that spatial information acts to modulate competition between sources, even in an anesthetized preparation. The fact that these effects arise under anesthesia is important because it shows that they are preattentive. Competition between

into orthogonal categories of spike additions, where the presence of the masker causes extra spikes (usually in the gaps between syllables), and spike subtractions, where spikes that are elicited by the target alone are reduced by the presence of the masker (usually

Ti



Spike E traction and Sorting

Extraction of action potentials (spikes) was performed off-line. First, neural traces were thresholded. The recording to 1 ms on either side of each local maximum was windowed out and considered a potential spike. These waveforms were sorted into user-defined template spike waveforms using a correlation-like coefficient:

$$r = \Sigma x_S x_T / \max(\Sigma x_S^2, \Sigma x_T^2),$$

where x_S is a spike waveform and x_T a template waveform, and the sums are taken over time. Spikes were sorted into classes based on

11. Nelken I, Fishbach A, Las L, Ulanovsky N, Farkas D (2003) Primary auditory cortex of cats: feature detection or something else? *Biol Cybern* 89: 397–406.
12. Nelken I (2004) Processing of complex stimuli and natural scenes in the auditory cortex. *Curr Opin Neurobiol* 14: 474–480.
13. Rakerd B, Hartmann WM (1985) Localization of sound in rooms, II: The effects of a single reflecting surface. *J Acoust Soc Am* 78: 524–533.
14. Ruggles D, Shinn-Cunningham B (2010) Spatial selective auditory attention in the presence of reverberant energy: individual differences in normal-hearing listeners. *J Assoc Res Otolaryngol*.
15. Kidd G, Mason CR, Brughera A, Hartmann WM (2005) The role of reverberation in release from masking due to spatial separation of sources for speech identification. *Acta Acustica United with Acustica* 91: 526–536.