Published in final edited form as:

Proc Meet Acoust. ; 19(1): . doi:10.1121/1.4799206.

Auditory processing for contrast enhancement of salient communication vocalizations

Alex G. Dunlap, Frank Lin, and Robert Liu*

Abstract

In a natural acoustic environment, coherent representations of auditory objects and sources are streamed from the myriad sounds that enter our ears. Features of those sounds that are familiar and behaviorally salient to us are detected and discriminated into invariant precepts that inform us about our external world. Research into how this occurs is increasingly converging on the idea that there is a transformation from the auditory periphery wherein an initial acoustically faithful representation by neurons becomes progressively altered to enhance the population neural representation of perceptually relevant aspects of the sound. How this occurs may vary for sounds whose meanings are acquired in different ways, perhaps depending on what actions and decisions must be executed upon recognition. We have investigated this process in a natural social context in which mouse mothers “learn” about the meaning of pup ultrasound vocalizations through their maternal care. Here we discuss our recent studies in awake mice using electrophysiological, behavioral, immunohistochemical and computational methods. Our results suggest that experience with natural vocalizations may alter core auditory cortical neural responses so that the contrast in activity across the neural population enhances the detection and discrimination of salient calls.

INTRODUCTION

In a natural acoustic environment, coherent representations of auditory objects and sources are streamed from the myriad sounds that enter our ears. Features of those sounds that are familiar and behaviorally salient to us are detected and discriminated into invariant precepts that inform us about our external world. Research into how this sound recognition occurs is increasingly converging on the idea that there is a transformation from the auditory periphery wherein an initial acoustically faithful representation by neurons becomes progressively altered to enhance the population neural representation of perceptually relevant aspects of the sound. How this occurs may vary for sounds whose meanings are acquired in different ways, perhaps depending on what actions and decisions must be executed upon recognition. We have investigated this process in a natural social context in which mouse mothers “learn” about the meaning of pup ultrasonic vocalizations through their maternal care.

A mouse pup, when displaced from its nest, emits ultrasonic calls, which serve as a communication signal to its mother to search and retrieve the pup. The calls consist mostly

of single frequency whistles in a range above 40 kHz (Figure 1a). Behavioral studies have shown that even though pup retrieval employs many sensory modalities (sound, smell, and tactile input), the calls played back alone can elicit the mother's search and retrieval, provided the mother is maternally motivated to respond (Haack, Markl et al. 1983). What features of these calls enable that recognition? As can be seen from Figure 1a, individual calls are variable from one pup to another, and even between calls from the same pup. Nevertheless, their acoustic parameters are not random, and in fact naturally cluster (Figure 1b) along certain simple dimensions like call frequency and duration (Liu, Miller et al. 2003). This could assist mothers in recognizing these calls versus other sounds in their environment (like adult ultrasound vocalizations, also pictured in Figure 1b), in a manner similar to how a vowel phoneme is recognized based on stereotypic clustering in formant frequency space (Peterson and Barney 1952). Indeed, as with speech phonemes for humans, mother mice have been shown to categorically perceive ultrasounds along discriminating acoustic dimensions (ultrasound frequency, duration and frequency bandwidth) by preferentially approaching pup-like ultrasound models over lower frequency, behaviorally unmeaningful pure tones (Ehret and Haack 1982, Ehret 1992). Hence, this mouse model offers a golden opportunity to study how different acoustic parameters important for the behavioral recognition of sound categories are represented throughout the auditory system.

Importantly, the mother's preferential approach to pup-like ultrasounds is acquired, thereby making this model system attractive for revealing how the neural representation is altered as behavioral relevance is gained. Virgin mice with no exposure to pups do not preferentially approach ultrasonic pup calls, but both mothers as well as pup-experienced virgin mice do (Ehret, Koch et al. 1987). The degree to which this acquisition is facilitated by the hormonal changes associated with motherhood is being actively explored. Clearly then, this system provides a rich platform for investigating the neural processing and plasticity that occurs when the meaning of vocalizations is "learned." This animal model has many other benefits that make it ideal for this purpose. For example, the natural ultrasonic calls are much simpler to parameterize acoustically than human or non-human primate vocalizations, making these easier to work with in designing acoustic control stimuli for neurophysiological studies. Using mice also opens the door to powerful genetic techniques that allow us to better dissect underlying cellular mechanisms (Liu 2006), and potentially manipulate auditory circuitry. For these reasons the mouse ultrasonic communication model is becoming a powerful tool for researchers investigating neural plasticity and vocalization perception in the auditory system. We have been exploiting the system primarily to understand neural coding and plasticity for acoustic communication at the level of the auditory cortex. Some of our recent findings are discussed below in a larger context of how auditory cortex may participate in the transformation from acoustic input to perceptual objects.

FINDINGS

Auditory input is represented at the eighth-nerve in an acoustically faithful fashion. As an indication of this, models of neural firing based only on the characterization of how a neuron responds to acoustic parameters can be quite successful at predicting auditory peripheral responses. This is the case not only for multi-parameter biophysical models (Zhang, Heinz et al. 2001), but also for characterizations based on just receptive fields, like the Wiener kernel

(Temchin, Recio-Spinoso et al. 2005). Even simply knowing the shape of a tonal sound's amplitude envelope is sufficient to accurately predict the latency of an auditory nerve fiber's first spike (Neubauer and Heil 2008). Hence, changes in acoustic parameters like amplitude envelope, frequency trajectory, etc. will generate neural responses at the peripheral level that accurately reflect this variation. However, faithfulness purely to the acoustics would also complicate how different natural variants from a class of calls could be recognized by the auditory system as belonging to the same perceptual category. Instead, acoustic features of calls relevant for categorization should be identified and encoded at some point along the auditory pathway in a way that is more tolerant of natural variability.

We have begun to explore the nature of this hypothesized coding transformation at the level of auditory cortex. If such a transformation were underway at this stage, we would expect that acoustic parameters that are faithfully represented at the periphery might be less so at the auditory cortex, where behavioral salience might instead bias responses. With this in mind, we investigated the sensitivity of cortical single units in awake, head-restrained mice to one particular variable acoustic parameter, the amplitude envelope onset of ultrasound vocalizations. Since pup calls are fairly tonal in nature, the model of Neubauer and Heil (Neubauer and Heil 2008), which used nonlinear sound envelope integration to successfully predict auditory nerve fiber latencies, can be straightforwardly extended to test whether cortical single units show the same type of onset sensitivity. We found that most cortical units are not that well predicted by such a model (Lin and Liu 2010), suggesting that cortical responses are not just driven by the acoustics at vocalization onset. Importantly though, the model does not simply fail because cortical activity is generally unpredictable from the acoustics. Indeed, a physiologically distinct subgroup of cortical neurons was actually found to have vocalization responses that were fairly well predicted by the model. These units had shorter latencies, fired robustly and transiently to the onsets of most sounds, had higher spontaneous activity, and were, curiously, all putative inhibitory interneurons based on their action potential waveforms. In contrast, those units that were most poorly predicted by the model had much longer latencies, fired more selectively to fewer ultrasound calls, and were predominantly putative pyramidal cells, which would provide the projection of the auditory cortical neural representation to downstream processing areas. These units' longer latencies may indicate additional processing that could allow these poorly predicted neurons to discriminate higher order acoustic features beyond sound onset, such as frequency modulations that occur along a call's trajectory. This leads to a key question in our investigation of coding transformations along the auditory pathway. If such a higher-order feature were to provide a basis for discriminating a sound category that gains behavioral salience, would responses of poorly predicted units change to reflect that increase in salience, as we might expect if the neural representation were being transformed to better encode behaviorally relevant auditory objects?

An experiment to test this would be fairly straightforward in our mouse model. Such higher-order features may be essential for a mother mouse to unequivocally distinguish pup from adult calls, which have their own behavioral meaning for adult females (Shepard and Liu 2010). As evident from Figure 1b, these two categories of calls overlap in simple acoustic parameters like median frequency and duration, so some vocalizations could be perceptually confused if frequency and duration were the only acoustic features by which to categorize

ultrasounds. Fortunately though, adult calls have much more frequency modulation, and neurons sensitive to this higher order feature could be decisive for telling apart each type of call. This ability is presumably more important for a mother that has learned that pup ultrasounds are potentially indicative of a lost pup in need of retrieval. Therefore, if in neurophysiology experiments we were to present both pup and adult calls matched for their simple acoustic features but differing in their frequency trajectories, would we see encoding differences in poorly predicted neurons between mothers and virgins? If the auditory cortical representation of higher-order features is indeed modulated by behavioral salience, then we would expect to see an increase in mothers in the activity of poorly predicted, putative pyramidal neurons in response to pup calls over adult calls, irrespective of natural variability in the simple acoustic features shared by both sound categories. Whether this is indeed the case is being actively investigated.

Besides increasing the excitatory activity of projection neurons encoding behaviorally relevant features of salient vocalizations, another way the auditory cortex might contribute to transforming the neural representation natural vocalizations is to actively suppress the activity of neurons responding to acoustic features that are not behaviorally relevant. From a population encoding perspective, this would effectively enhance the neural contrast in detecting whether features of behaviorally relevant calls are present in the incoming sound stream. Our studies of single unit recordings from awake, head-restrained mother and virgin mice have found evidence in support of this idea. Specifically, we found that inhibition of neurons in auditory cortical areas with best frequency responses tuned below those frequencies that are typical of pup calls (>50 kHz) was in fact different between mothers and virgins. Mother mice had call evoked inhibition that was earlier, longer, stronger, and more stereotyped when compared to virgins (Galindo-Leon, Lin et al. 2009). Because of its rapid onset we speculate that this inhibition may reflect a bottom up mechanism for facilitating the detection and discrimination of pup calls for mothers. In principle, this enhanced, so-called lateral band call evoked inhibition alters the population level representation in such a way that downstream neurons pooling across the population of auditory cortex would be less influenced by features that do not represent pup calls. This would help the neural activity representing pup calls to stand out more clearly, particularly in the presence of background noise. The fact that this mechanism is strengthened after the salience of pup calls is gained in mothers further supports the idea that the auditory cortex is not simply responding to the acoustics of individual vocalizations, but instead biasing their neural representation to better recognize categories of behaviorally relevant sounds.

Acknowledgments

We thank Katy Shepard for assisting with preparation of ultrasound call spectrograms. We also thank NIDCD (008343) and NSF CBN (IBN-9876754) for funding.

References

- Ehret G. Categorical perception of mouse-pup ultrasounds in the temporal domain. *Anim Behav.* 1992; 43(3):409–416.
- Ehret G, Haack B. Ultrasound recognition in house mice: key-stimulus configuration and recognition mechanism. *J Comp Physiol [A].* 1982; 148(2):245–251.

- Ehret G, Koch M, Haack B, Markl H. Sex and parental experience determine the onset of an instinctive behavior in mice. *Naturwissenschaften*. 1987; 74(1):47. [PubMed: 3561521]
- Galindo-Leon EE, Lin FG, Liu RC. Inhibitory Plasticity in a Lateral Band Improves Cortical Detection of Natural Vocalizations. *Neuron*. 2009; 62(5):705–716. [PubMed: 19524529]
- Haack, B.; Markl, H.; Ehret, G. Sound communication between parents and offspring. In: Willott, JF., editor. *The auditory psychobiology of the mouse*. Springfield, IL: Charles C. Thomas; 1983. p. 57-97.
- Lin FG, Liu RC. Subset of thin spike cortical neurons preserve the peripheral encoding of stimulus onsets. *J Neurophysiol*. 2010; 104(6):3588–3599. [PubMed: 20943946]
- Liu RC. Prospective contributions of transgenic mouse models to central auditory research. *Brain Res*. 2006; 1091(1):217–223. [PubMed: 16574081]
- Liu RC, Miller KD, Merzenich MM, Schreiner CE. Acoustic variability and distinguishability among mouse ultrasound vocalizations. *J Acoust Soc Am*. 2003; 114(6):3412–3422. [PubMed: 14714820]
- Neubauer H, Heil P. A physiological model for the stimulus dependence of first-spike latency of auditory-nerve fibers. *Brain Res*. 2008; 1220:208–223. [PubMed: 17936252]
- Peterson GN, Barney HL. Control methods used in a study of the vowels. *J Acoust Soc Am*. 1952; 24:175–184.
- Shepard KN, Liu RC. Experience restores innate female preference for male ultrasonic vocalizations. *Genes Brain Behav*. 2010
- Temchin AN, Recio-Spinoso A, van Dijk P, Ruggero MA. Wiener kernels of chinchilla auditory-nerve fibers: verification using responses to tones, clicks, and noise and comparison with basilar-membrane vibrations. *J Neurophysiol*. 2005; 93(6):3635–3648. [PubMed: 15659530]
- Zhang XD, Heinz MG, Bruce IC, Carney LH. A phenomenological model for the responses of auditory-nerve fibers: I. Nonlinear tuning with compression and suppression. *Journal of the Acoustical Society of America*. 2001; 109(2):648–670. [PubMed: 11248971]

FIGURE 1.

(a) The two panels illustrate the spectrograms of pup call bouts emitted by two different isolated pups. Calls are predominantly single frequency whistles, although there are also calls where the frequency discontinuously jumps (end of first call in top spectrogram). (b) Probability distribution of pup call duration and median frequency. Pup calls form two clusters, but due to the hearing range of mice the cluster centered at approximately 65 kHz is likely to be the most relevant for mouse ultrasonic communication. Also shown is a cluster for ultrasound vocalizations emitted by adult male mice in the presence of female mice, which is complementary to the pup call clusters. Panel (b) adapted from (Liu, Miller et al. 2003).