

# UC Berkeley

## UC Berkeley Electronic Theses and Dissertations

### Title

Development and Plasticity in the Primary Auditory Cortex

### Permalink

<https://escholarship.org/uc/item/59r4d83x>

### Author

Kim, Heesoo

### Publication Date

2011

Peer reviewed|Thesis/dissertation

Development and Plasticity in the Primary Auditory Cortex

By

Heesoo Kim

A dissertation submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy

in

Neuroscience

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Shaowen Bao, Chair

Professor Yang Dan

Professor Daniel E Feldman

Professor Fei Xu

Fall 2011

Copyright  
Heesoo Kim, 2011  
All rights reserved.

## **ABSTRACT**

Development and Plasticity in the Primary Auditory Cortex

by

Heesoo Kim

Doctor of Philosophy in Neuroscience

University of California, Berkeley

Professor Shaowen Bao, Chair

The early acoustic environment plays a crucial role in how the brain represents sounds and how language phonemes are perceived. Human infants are born with the capacity to distinguish phonemes from virtually all languages, but very quickly change their perceptual ability to match that of their primary language. This has been described as the Perceptual Magnet Effect in humans, where phoneme tokens are perceived to be more similar than they physically are, leading to decreased discrimination ability.

Early development is marked by distinct critical periods, when cortical regions are highly plastic and particularly sensitive to sensory input. These lasting alterations in cortical sensory representation may directly impact the perception of the external world. My thesis is comprised of three different studies, all of which investigate the role of the developmental acoustic environment on cortical representation and the behavioral consequence of altered cortical representation.

Passive exposure to pure-tone pips during the auditory critical period can lead to over-representation of the exposure tone frequency in the primary auditory cortex (A1) of rats. This over-representation is associated with decreased discrimination ability of that frequency, similar to the Perceptual Magnet Effect in humans. Another hallmark of human language is categorical perception. Using a computational model of A1, I show that certain representation patterns (which may be achieved with passive exposure to two distinct pure-tone pips) in A1 can lead to categorical perception in rats. This suggests that cortical representation may be a mechanism that drives categorical perception.

Rodents are socially vocal animals whose con-specific calls are often presented in bouts in the ultrasonic frequency range. These calls are vocalized at ethologically relevant repetition rates. I show that pure-tone pips that are presented at the ethological repetition rate (but not slower or faster rates) during the auditory critical period lead to over-representation of the pure-tone frequency. A certain subclass of ultrasonic vocalizations, the pup isolation calls, occurs during the auditory critical period. I show that there is over-representation of ultrasonic vocalization frequencies in the rat A1. This preferential representation is experience-dependent and is associated with higher discrimination ability.

## TABLE OF CONTENTS

Table of Contents.....	i
List of Figures & Tables.....	ii
Acknowledgements.....	iv
Vita.....	v
Chapter 1: Introduction.....	1
Chapter 2: Distributed representation of perceptual categories in the auditory cortex.....	8
Abstract.....	8
Introduction.....	9
Methods.....	11
Results.....	15
Discussion.....	19
Figures.....	22
References.....	32
Chapter 3: Selective increase in representation of sounds repeated at an ethological rate.....	36
Abstract.....	36
Introduction.....	37
Methods.....	38
Results.....	41
Discussion.....	44
Figures.....	47
References.....	51
Chapter 4: Over-representation of ultrasonic vocalization frequencies in the rat primary auditory cortex.....	54
Abstract.....	54
Introduction.....	55
Methods.....	57
Results.....	60
Discussion.....	65
Figures.....	69
References.....	84

## LIST OF FIGURES & TABLES

Chapter 2:	
Figure 2.1: Tuning curves of the model neurons in the model naïve primary auditory cortex.....	22
Figure 2.2: Example of distributions (probability density functions) of the differences between the decoded frequencies.....	23
Figure 2.3: Discrimination performances of the model naïve AI and animals as a function of differences between the input frequencies.....	24
Figure 2.4: Influences of neuronal response properties on the discrimination threshold of model naïve AI.....	25
Figure 2.5: Impaired discrimination of the over-represented frequencies.....	26
Figure 2.6: Comparison of naïve animal performances with model naïve AI performances in a frequency identification task.....	27
Figure 2.7: Frequency identification by the model naïve AI and model 7.1-kHz-exposed AI.....	28
Figure 2.8: Representations of two perceptual categories.....	29
Figure 2.9: Categorical population responses.....	30
Figure 2.10: Influences of neuronal response properties on categorical representation of frequencies.....	31
Chapter 3:	
Table 3.1: Response properties of A1 neurons.....	46
Figure 3.1: Characterization of rat calls.....	47
Figure 3.2: Influences of stimulus presentation rate on spectral plasticity.....	48
Figure 3.3: Over-representation of sounds repeated at the ethological rates.....	49
Figure 3.4: Effects of tonal exposure on cortical temporal response properties.....	50

Chapter 4

Table 4.1: Ultrasonic vocalization in the rat.....	67
Table 4.2 Sound stimulus does not capture complete tuning curve for most low frequency neurons.....	68
Figure 4.1: Ultrasonic vocalization in the rat.....	69
Figure 4.2: Vocalization analysis methods.....	70
Figure 4.3: Representation of ultrasonic frequencies.....	71
Figure 4.4: Isolation of response properties.....	72
Figure 4.5: Representation of ultrasonic frequencies and response properties.....	73
Figure 4.6: Population Fisher Information.....	74
Figure 4.7: Discrimination of ultrasonic frequencies.....	75
Figure 4.8: Development of ultrasonic frequencies.....	76
Figure 4.9: Response properties in developing animals.....	77
Figure 4.10: Ultrasonic representation requires early experience.....	78
Figure 4.11: Response properties of deafened animals.....	79
Figure 4.12: Increased representation of ultrasonic frequencies in inferior colliculus.....	80
Figure 4.13: Response properties in inferior colliculus.....	81
Figure 4.14: Gender.....	82
Figure 4.15: Gender Response Properties.....	83

## **ACKNOWLEDGEMENTS**

Words cannot describe the gratitude I feel toward Shaowen. His ability to see the bigger picture is something I still admire and aspire to achieve. I thank Shaowen for his dedication to all of us graduate students; whether it was help with mapping at odd hours, setting up rearing boxes at the last minute, or venturing in analysis, we could always be certain Shaowen would be there for us. I feel so fortunate to have spent my graduate career as Shaowen's student.

The members of my qualifying exam committee and thesis committee played such an important role in my development as a scientist. I am so thankful for their continual support and all the insight that they provided through the years. In particular, a big thank you to Dan Feldman for countless letters of recommendation, many requested at the last minute.

I thank the Helen Wills Neuroscience Institute for giving me this opportunity to pursue neuroscience research. My experience in the program would not have been the same without John Ngai and Kati Markowitz's advocacy and support of us graduate students. We are all indebted to you both.

It has been such a rewarding experience to be a part of the Bao lab. Everyone's enthusiasm for science was truly a blessing to be around. In particular I thank Hania and Liberty – I'm so glad to have found such great friends.

My parents always believed in me and sacrificed so much to give me every opportunity. I still can't imagine how they managed to get their respective Ph.D.s with two kids running around the house. Thank you both for everything.

And finally, to Kevin: thank you for all the big moments and little details that we share... And yes, my dear, we do talk about science far more than we should.

Chapter 2, in full, is a reprint of the material as it appears in:  
Heesoo Kim & Shaowen Bao. Distributed representation of perceptual categories in the auditory cortex. *Journal of Computational Neuroscience*. (2008) 24(3): 277-290.

Chapter 3, in full, is a reprint of the material as it appears in:  
Heesoo Kim & Shaowen Bao. Selective Increase In Representations Of Sounds Repeated At An Ethological Rate. *The Journal of Neuroscience*. (2009) 29(16): 5163-5169.

Chapter 4 is original work in preparation.

## VITA

- 2000-2005 B.S., Symbolic Systems Program, Human Computer Interaction  
M.S., Symbolic Systems Program, Neuroscience  
Leland Stanford Jr. University, Stanford, California
- 2006-2011 Ph.D., Neuroscience  
University of California, Berkeley

## PUBLICATIONS

- Nadine Gaab, Paula Tallal, Heesoo Kim, Jermaine J. Archie, Kala Lakshminarayanan, Gary, H. Glover and John D.E. Gabrieli. (2005). Neural Correlates of Rapid Temporospectral Processing in Musicians and Nonmusicians. *Annals of the New York Academy of Sciences*, 1060: 82-88.
- Noa Ofen, Yun-Ching Kao, Peter Sokol-Hessner, Heesoo Kim, Susan Whitefield-Gabrieli, and John D.E. Gabrieli. (2007). Development of the declarative memory system in the human brain. *Nature Neuroscience*, 10: 1198-1205.
- Heesoo Kim and Shaowen Bao. (2008). Distributed Representation of Perceptual Categories in the Auditory Cortex. *Journal of Computational Neuroscience*, 24: 277-290.
- Dafna Bergerbest, John D.E. Gabrieli, Susan Whitfield-Gabrieli, Heesoo Kim, Glenn T. Stebbins, David A. Bennet, and Debra A. Fleischman. (2009). Age-Associated Reduction of Asymmetry in Prefrontal Function and Preservation of Conceptual Repetition Priming. *Neuroimage*, 45:237-246.
- Heesoo Kim and Shaowen Bao. (2009). Selective Increase in Representation of Sound Repeated at an Ethological Rate. *Journal of Neuroscience*, 29(16): 5163-5169.
- Michele Insanally, Hania Koeber, Heesoo Kim, and Shaowen Bao. (2009). Feature-Dependent Sensitive Periods in the Development of Complex Sound Representation. *Journal of Neuroscience*, 29 (17): 5456-5462.

## AWARDS

- 2007-2009 National Science Foundation Graduate Research Service Award
- 2009-2011 Ruth L. Kirschstein National Research Service Award
- 2010 Elizabeth Roboz Einstein Fellowship for Studies in Neurosciences and Human Development
- 2010 SfN Travel Award for the Federation of Neuroscience Societies Forum

## **CHAPTER 1: Introduction**

The early acoustic environment plays a crucial role in shaping how sounds are perceived in adulthood. For example, individuals who grow in an environment where they exclusively hear Japanese often have difficulty distinguishing certain English phonemes, such as /ra/ and /la/ (Miyawaki et al., 1975). This reduction of perceptual sensitivity near extensively experienced sounds, often called the perceptual magnet effect, has been thoroughly studied with comparisons between several languages (Iverson and Kuhl, 1995; Kuhl, 2000). This effect is not genetic, but is highly dependent on the early developmental acoustic and language environment of the individual.

Early development is often associated with high levels brain plasticity. During these critical or sensitive periods, various cortical areas are particularly sensitive to their respective inputs. Artificially altered sensory inputs during a critical period can often lead to lasting changes in cortical representation. However, the same sensory inputs presented in a later time in life often have little effect. This was first demonstrated in the cat and monkey visual cortex by the seminal work of David Hubel and Torsten Wiesel (Wiesel and Hubel, 1965; Hubel and Wiesel, 1970; Hubel et al., 1977) and later demonstrated in the rodent model (Fagiolini et al., 1994). Critical period plasticity has also been described in the rodent barrel cortex (Fox, 1992) and the rodent primary auditory cortex (Zhang et al., 2001). These changes in cortical representation may directly impact behavioral perception (Han et al., 2007; Kim and Bao, 2008).

This thesis explores the relationship between developmental plasticity and perception. This introduction will give a brief overview of rodent primary auditory cortex (A1) critical period plasticity; the relationship between cortical representation and perception; and rodent vocalizations. The original work for this dissertation is presented in three chapters, two of which have been previously published. Chapter 2 is a computational modeling study investigating the role of cortical representation on categorical perception (Kim and Bao, 2008). Specific patterns of cortical representation are proposed to lead to categorical perception. In chapter 3, the role of temporal rate in spectral plasticity is explored (Kim and Bao, 2009). Specifically, trains of tone pips that are presented at ethologically relevant repetition rates (but not slower or faster rates) lead to expansion of representation of the carrier frequency in A1. Finally, chapter 4 will explore the cortical representation of ethologically relevant vocalization frequencies. Preferential representation of ultrasonic frequencies in A1 is found to be experience-dependent and is accompanied by increased discrimination ability.

### *Rodent A1 critical period plasticity*

The rat auditory cortex has an orderly tonotopic gradient with low to high frequencies represented smoothly along the caudal-rostral axis (Sally and Kelly, 1988; Rutkowski, 2003). This orderly tonotopy can be altered in adults through intense training (Polley et al., 2006) or with co-presentation of neural modulators (Kilgard and Merzenich, 1998; Bao et al., 2001). However, during the developmental critical period, mere passive exposure to a single frequency tone can lead to substantial over-representation of that particular frequency (Zhang et al., 2001; Han et al., 2007; Kim and Bao, 2009). Although the critical period window for frequency plasticity in rat A1 has

been defined to be between post-natal day 11 (P11) and P13 (de Villers-Sidani et al., 2007), typically longer windows are utilized in sound exposure experiments.

Pure tone pips are not the only stimulus that can elicit changes in tonotopic representation. Broad-band noise bursts have been shown to disrupt proper development of frequency representation (Zhang et al., 2002; Insanally et al., 2009). Exposure of frequency modulated (FM) sweeps in a period of time that includes the tonal critical period also can result in changes in tone representation. When exposed to downward FM sweeps early in development (P8-P15), rat A1 appears to lose representation of low frequencies (Insanally et al., 2009). Interestingly, exposure to the same stimulus, but at different points in development can lead to differential plasticity effect. For example, exposure to downward FM sweep or broadband noise between P16 and P23 results to broader tuning bandwidth, but normal tonotopic representation (Insanally et al., 2009, 2010). This suggests that critical period of different response properties exist at different times in development.

Efforts to see plasticity effects with multiple tones have lead to results that are more difficult to interpret. Passive exposure to multi-tone sequences disrupts tonotopic representation, but the changes in representation do not appear to be predictable by the stimulus (Nakahara et al., 2004). Similarly in adult animals, while nucleus basalis (NB) stimulation paired with single frequency tone pips lead to large increase in representation of that frequency (Kilgard and Merzenich, 1998), NB stimulation paired with multiple tone presentations result in normal tonotopic maps (Kilgard et al., 2001). Further studies are required to understand the role of complex multi-tone environments on cortical plasticity.

The mechanisms underlying A1 critical period plasticity are largely unknown, although many mechanisms have been proposed to drive plasticity in cortex (Feldman, 2009). In primary visual cortex (V1) the inhibitory circuit has been proposed to play an important role in controlling critical period plasticity (Hensch, 2005). Recently, two studies have found the excitatory-inhibitory balance in rat A1 neurons changes during early development, suggesting that similar critical period plasticity mechanisms may exist between A1 and V1 (Dorrn et al., 2010; Sun et al., 2010). Utilizing pharmacology or genetic model organisms such as mice (Barkat et al., 2011) will help us further understand the mechanisms underlying A1 critical period plasticity.

### *Cortical representation and perception*

Intense behavioral training has been shown to alter cortical representation of sounds (Recanzone et al., 1993; Bao et al., 2004; Polley et al., 2006). Generally speaking, these studies suggest increased discrimination ability is associated with increased cortical representation. However, it has also been shown that increased cortical representation via developmental exposure to 7 kHz is associated with decreased discrimination ability around 7 kHz (Han et al., 2007). Although seemingly contradictory to previous results, closer examination of the tuning curve distribution reveals an interesting explanation: many of the tuning curves have a peak around 7 kHz with a conspicuous lack of tuning curves with peaks in neighboring frequencies (Han et al., 2007). It has previously been suggested that discrimination is easiest along the slope of tuning curve (Butts and Goldman, 2006), much in line with the behavioral results of this study. Furthermore, this representation pattern and behavioral results are

quite similar to the perceptual magnet effect described in human speech perception (Iverson and Kuhl, 1995). Chapter 2 explores the proposal that specific distribution of tuning curves in A1 can give rise to categorical perception, one of the hallmarks of speech perception (Kim and Bao, 2008).

Biases in perception are not unique to human language perception. For example, human display increased discrimination of cardinal orientations (Girshick et al., 2011). Efficient representation of the visual world can explain human's increased discrimination ability for cardinal orientations (Ganguli and Simoncelli, 2011). Similar mechanisms have been proposed to explain biases in owl sound localization (Fischer and Peña, 2011). In addition, developmentally driven changes in rat A1 has been proposed to alter perceptual ability (Kim and Bao, 2008; Köver and Bao, 2010). This is explored further in Chapter 4, by investigating the representation of ethologically relevant con-specific vocalization frequencies and it's relationship to behavioral perceptual ability.

### *Rodent vocalizations*

Rodents are highly vocal animals that have con-specific communication calls that are largely in the ultrasonic frequency range. Rat vocalizations have been broadly grouped by frequency into three categories: (1) 22 kHz alarm calls, (2) 40 kHz pup isolation calls and (3) 50 kHz adult calls that are often associated with positive affective states (Brudzynski et al., 1993, 1999; Knutson et al., 2002; Brudzynski, 2005; Portfors, 2007). In adult rats, calls of different frequencies are associated with specific behaviors, such as fighting, feeding and running, further separating the 50 kHz adult calls into separate behaviorally relevant groups (Takahashi et al., 2010). These ultrasonic vocalizations are capable of eliciting specific behaviors. Mothers will actively search for pups when presented with 40 kHz pup isolation calls (Ehret and Haack, 1981; Ehret et al., 1987; Hahn and Lavooy, 2005) and female rodents are more receptive to sexual encounters after exposure to male encounter calls (McIntosh and Barfield, 1978). These calls play an extraordinarily important role in the social behaviors and survival of rodents.

Investigating how con-specific vocalizations are represented in cortex is important in helping us understand how these communication calls can elicit specific behaviors (Wang, 2000). Representation of vocalizations can change: maternal experience has been shown to alter cortical responses to vocalizations and other synthetic sounds (Liu et al., 2006; Cohen et al., 2011). One interesting characteristic of animal communication calls is that they are often presented in bouts rather than in isolation (Liu et al., 2003; Schnupp et al., 2006; Kim and Bao, 2009). Previous reports have shown that auditory cortex neurons are sensitive to the temporal structure of con-specific calls (Wang and Kadia, 2001). In order to efficiently represent ethologically relevant spectral frequencies, A1 may utilize the temporal information of external auditory stimuli to help filter enhanced representation of specific frequencies (Kim and Bao, 2009). This hypothesis is explored in Chapter 3.

Rat communication calls primarily occur in the ultrasonic frequency range. However, the vast majority of rat A1 studies have only investigated responses to frequencies up to 32 kHz (Zhang et al., 2001; Han et al., 2007; Polley et al., 2007; de Villers-Sidani et al., 2007; Insanally et al., 2009). The representation of ultrasonic frequencies in A1 extends to 64 kHz (Rutkowski, 2003; Kim and Bao, 2009). In Chapter

4, the representation of ultrasonic vocalization frequencies in rat A1 is investigated. Preferential representation of ultrasonic frequencies is found to be experience-dependent, and this over-representation is associated with enhanced discrimination ability.

## **References**

- Bao S, Chan VT, Merzenich MM (2001) Cortical remodelling induced by activity of ventral tegmental dopamine neurons. *Nature* 412:79–83.
- Bao S, Chang EF, Woods J, Merzenich MM (2004) Temporal plasticity in the primary auditory cortex induced by operant perceptual learning. *Nature neuroscience* 7:974–981.
- Barkat TR, Polley DB, Hensch TK (2011) A critical period for auditory thalamocortical connectivity. *Nature Neuroscience* 14:1189–1194.
- Brudzynski SM (2005) Principles of rat communication: quantitative parameters of ultrasonic calls in rats. *Behavior Genetics* 35:85–92.
- Brudzynski SM, Bihari F, Ociepa D, Fu XW (1993) Analysis of 22 kHz ultrasonic vocalization in laboratory rats: long and short calls. *Physiology & Behavior* 54:215–221.
- Brudzynski SM, Kehoe P, Callahan M (1999) Sonographic structure of isolation-induced ultrasonic calls of rat pups. *Developmental Psychobiology* 34:195–204.
- Butts D a, Goldman MS (2006) Tuning curves, neuronal variability, and sensory coding. *PLoS Biology* 4:e92.
- Cohen L, Rothschild G, Mizrahi A (2011) Multisensory Integration of Natural Odors and Sounds in the Auditory Cortex. *Neuron* 72:357–369.
- Dorn AL, Yuan K, Barker AJ, Schreiner CE, Froemke RC (2010) Developmental sensory experience balances cortical excitation and inhibition. *Nature* 465:932–936.
- Ehret G, Haack B (1981) Categorical perception of mouse pup ultrasound by lactating females. *Die Naturwissenschaften* 68:208–209.
- Ehret G, Koch M, Haack B, Markl H (1987) Sex and parental experience determine the onset of an instinctive behavior in mice. *Die Naturwissenschaften* 74:47.
- Fagiolini M, Pizzorusso T, Berardi N, Domenici L, Maffei L (1994) Functional postnatal development of the rat primary visual cortex and the role of visual experience: dark rearing and monocular deprivation. *Vision Research* 34:709–720.
- Feldman DE (2009) Synaptic mechanisms for plasticity in neocortex. *Annual Review of Neuroscience* 32:33–55.
- Fischer BJ, Peña JL (2011) Owl's behavior and neural representation predicted by Bayesian inference. *Nature Neuroscience* 14:1061–1066.
- Fox K (1992) A critical period for experience-dependent synaptic plasticity in rat barrel cortex. *The Journal of Neuroscience* 12:1826–1838.
- Ganguli D, Simoncelli EP (2011) Implicit encoding of prior probabilities in optimal neural populations. *Processing*:6–9.
- Girshick AR, Landy MS, Simoncelli EP (2011) Cardinal rules: visual orientation perception reflects knowledge of environmental statistics. *Nature Neuroscience* 14:926–932.
- Hahn ME, Lavooy MJ (2005) A review of the methods of studies on infant ultrasound production and maternal retrieval in small rodents. *Behavior Genetics* 35:31–52.
- Han YK, Köver H, Insanally MN, Semerdjian JH, Bao S (2007) Early experience impairs perceptual discrimination. *Nature Neuroscience* 10:1191–1197.
- Hensch TK (2005) Critical period plasticity in local cortical circuits. *Nature Reviews Neuroscience* 6:877–888.

- Hubel D, Wiesel TN (1970) The period of susceptibility to the physiological effects of unilateral eye closure in kittens. *The Journal of physiology* 206:419.
- Hubel DH, Wiesel TN, LeVay S (1977) Plasticity of ocular dominance columns in monkey striate cortex. *Philosophical transactions of the Royal Society of London* 278:377–409.
- Insanally MN, Albanna BF, Bao S (2010) Pulsed noise experience disrupts complex sound representations. *Journal of Neurophysiology* 103:2611–2617.
- Insanally MN, Köver H, Kim H, Bao S (2009) Feature-dependent sensitive periods in the development of complex sound representation. *The Journal of Neuroscience* 29:5456–5462.
- Iverson P, Kuhl PK (1995) Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. *Journal of the Acoustical Society of America* 97:553–562.
- Kilgard MP, Merzenich MM (1998) Cortical Map Reorganization Enabled by Nucleus Basalis Activity. *Science* 279:1714–1718.
- Kilgard MP, Pandya PK, Vazquez J, Gehi A, Schreiner CE, Merzenich MM (2001) Sensory input directs spatial and temporal plasticity in primary auditory cortex. *Journal of Neurophysiology* 86:326–338.
- Kim H, Bao S (2008) Distributed representation of perceptual categories in the auditory cortex. *Journal of Computational Neuroscience* 24:277–290.
- Kim H, Bao S (2009) Selective increase in representations of sounds repeated at an ethological rate. *The Journal of Neuroscience* 29:5163–5169.
- Knutson B, Burgdorf J, Panksepp J (2002) Ultrasonic vocalizations as indices of affective states in rats. *Psychological Bulletin* 128:961–977.
- Kuhl PK (2000) A new view of language acquisition. *Proceedings of the National Academy of Sciences of the United States of America* 97:11850–11857.
- Köver H, Bao S (2010) Cortical plasticity as a mechanism for storing bayesian priors in sensory perception. *PloS One* 5:e10497.
- Liu RC, Linden JF, Schreiner CE (2006) Improved cortical entrainment to infant communication calls in mothers compared with virgin mice. *The European Journal of Neuroscience* 23:3087–3097.
- Liu RC, Miller KD, Merzenich MM, Schreiner CE (2003) Acoustic variability and distinguishability among mouse ultrasound vocalizations. *The Journal of the Acoustical Society of America* 114:3412.
- McIntosh TK, Barfield RJ (1978) Ultrasonic vocalisations facilitate sexual behaviour of female rats. *Nature* 272:163–164.
- Miyawaki K, Strange W, Verbrugge R (1975) An effect of linguistic experience: The discrimination of [ r ] and [ ɹ ] by native speakers of Japanese and English. *Perception & Psychophysics* 18:331–340.
- Nakahara H, Zhang LI, Merzenich MM (2004) Specialization of primary auditory cortex processing by sound exposure in the “critical period”. *Proceedings of the National Academy of Sciences of the United States of America* 101:7170–7174.
- Polley DB, Read HL, Storace D a, Merzenich MM (2007) Multiparametric auditory receptive field organization across five cortical fields in the albino rat. *Journal of Neurophysiology* 97:3621–3638.

- Polley DB, Steinberg EE, Merzenich MM (2006) Perceptual learning directs auditory cortical map reorganization through top-down influences. *The Journal of Neuroscience* 26:4970–4982.
- Portfors CV (2007) Types and functions of ultrasonic vocalizations in laboratory rats and mice. *Journal of the American Association for Laboratory Animal Science* 46:28–34.
- Recanzone GH, Schreiner CE, Merzenich MM (1993) Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *The Journal of Neuroscience* 13:87–103.
- Rutkowski R (2003) Characterisation of multiple physiological fields within the anatomical core of rat auditory cortex. *Hearing Research* 181:116–130.
- Sally SL, Kelly JB (1988) Organization of auditory cortex in the albino rat: sound frequency. *Journal of Neurophysiology* 59:1627–1638.
- Schnupp JWH, Hall TM, Kokelaar RF, Ahmed B (2006) Plasticity of temporal pattern codes for vocalization stimuli in primary auditory cortex. *The Journal of Neuroscience* 26:4785–4795.
- Sun YJ, Wu GK, Liu B-H, Li P, Zhou M, Xiao Z, Tao HW, Zhang LI (2010) Fine-tuning of pre-balanced excitation and inhibition during auditory cortical development. *Nature* 465:927–931.
- Takahashi N, Kashino M, Hironaka N (2010) Structure of rat ultrasonic vocalizations and its relevance to behavior. *PLoS One* 5:e14115.
- de Villers-Sidani E, Chang EF, Bao S, Merzenich MM (2007) Critical period window for spectral tuning defined in the primary auditory cortex (A1) in the rat. *The Journal of Neuroscience* 27:180–189.
- Wang X (2000) On cortical coding of vocal communication sounds in primates. *Proceedings of the National Academy of Sciences* 97:11843–11849.
- Wang X, Kadia SC (2001) Differential representation of species-specific primate vocalizations in the auditory cortices of marmoset and cat. *Journal of Neurophysiology* 86:2616.
- Wiesel TN, Hubel DH (1965) Comparison of the effects of unilateral and bilateral eye closure on cortical unit responses in kittens. *Journal of neurophysiology* 28:1029–1040.
- Zhang LI, Bao S, Merzenich MM (2001) Persistent and specific influences of early acoustic environments on primary auditory cortex. *Nature Neuroscience* 4:1123–1130.
- Zhang LI, Bao S, Merzenich MM (2002) Disruption of primary auditory cortex by synchronous auditory inputs during a critical period. *Proceedings of the National Academy of Sciences of the United States of America* 99:2309–2314.

## **CHAPTER 2: Distributed representation of perceptual categories in the auditory cortex**

Heesoo Kim & Shaowen Bao

Originally published in:

Journal of Computational Neuroscience, 24(3): 277-290

June 2008

### **Abstract**

Categorical perception is a process by which a continuous stimulus space is partitioned to represent discrete sensory events. Early experience has been shown to shape categorical perception and enlarge cortical representations of experienced stimuli in the sensory cortex. The present study examines the hypothesis that enlargement in cortical stimulus representations is a mechanism of categorical perception. Perceptual discrimination and identification behaviors were analyzed in model auditory cortices that incorporated sound exposure-induced plasticity effects. The model auditory cortex with over-representations of specific stimuli exhibited categorical perception behaviors for those specific stimuli. These results indicate that enlarged stimulus representations in the sensory cortex may be a mechanism for categorical perceptual learning.

## **Introduction**

While sensory stimuli may vary continuously along their physical dimensions, the behaviorally significant events that they represent are often discrete. Through a process called categorical perception, the sensory system maps continuous stimulus spaces to discrete perceptual spaces (Harnad, 2003). For instance, lights of gradually changing wavelength are perceived as having discrete hues (Bornstein et al., 1976). Gradual shift in sound frequencies may lead to categorical changes of the perceived musical intervals (Burns and Ward, 1978). Categorically perceived stimuli may be recognized more quickly in the presence of distortions and contextual interferences. This efficient sensory processing provides the bases for higher-level cognitive functions such as verbal communication and music appreciation (Harnad, 1987).

Categorical perception was first discovered in speech research and was thought to involve language-specific, higher-level brain mechanisms, but not the basic sensory processing mechanisms of auditory system (Liberman et al., 1967; Liberman et al., 1957). Later research indicated that categorical perception occurs in a variety of non-speech sounds (Ehret, 1992; Ehret and Haack, 1981; Nelson and Marler, 1989; Wyttenbach et al., 1996). In addition, speech sounds are also categorically perceived by animals of many species (Kluender et al., 1987; Kuhl and Miller, 1975; Kuhl and Padden, 1982; Kuhl and Padden, 1983). These findings suggest that categorical perception may be an auditory, rather than a purely phonetic, process and may be mediated by the auditory sensory system.

Neural mechanisms underlying categorical perception are not well understood. Investigations of such mechanisms often involve searching for categorical neurons — those that respond preferentially to all stimuli in one category, but not to any of the other categories, showing sigmoidal stimulus selectivity. These categorical neurons have been found in the frontal cortex (Freedman et al., 2001; Romo et al., 1997). Although behavioral and psychophysical evidence suggest that sensory systems may mediate categorical perception, the neurons in the sensory cortex, which typically respond to a broad range of stimuli and exhibit bell-shaped tuning curves, are not considered categorical.

Categorical perception may arise both through innate mechanisms and as a result of sensory experiences and learning (Livingston et al., 1998). Some human speech sounds, for instance, are categorically perceived in newborn human infants (Eimas, 1974) and in some model animals that have never been exposed to the speech sounds (Ehret and Haack, 1981; Kluender et al., 1987; Kuhl and Miller, 1975; Nelson and Marler, 1989; Wyttenbach et al., 1996). It has been suggested that the auditory systems of both humans and the model animals are innately sensitive to the acoustic distinctions of those speech sounds, and our vocal communication system simply exploits this sensitivity (Holt et al., 2004; Steinschneider et al., 2003). On the other hand, language experience can also alter the perceptual sensitivity of the auditory system to speech sounds and change their categorical boundaries (Lasky et al., 1975; MacKay et al., 2001; Williams, 1977). This language-specific reshaping of the phonetic perceptual categories occurs in the first year of life (Kuhl et al., 1992), presumably as a result of acoustic exposure to the speech sound environment. Categorical perception of pitch is also shaped by musical experiences (Burns and Ward, 1978).

Sensory experience in a limited window of early life has a profound influence on the development of the cortical sensory representations (Wiesel, 1982). Recent studies indicate that repeated exposure to a stimulus results in enlarged cortical representations of the experienced stimulus—i.e., more neurons becoming selectively responsive to the stimulus (Chang and Merzenich, 2003; Erickson et al., 2000; Sengpiel et al., 1999; Zhang et al., 2001). Similar preferential representations of experienced speech sounds and musical notes have also been shown in humans (Naatanen et al., 1997; Pantev et al., 1998). Given the profound impact of early experience both on categorical perception of speech sounds and musical notes, and on cortical sound representations, it is possible that experience-driven reorganization of the auditory cortex plays a role in forming perceptual categories (Crozier, 1997; Lasky et al., 1975; MacKay et al., 2001; Takeuchi and Hulse, 1993; Williams, 1977). In this study, we construct models of acoustic representations of the primary auditory cortex, and examine the effects of experience-induced reorganization of acoustic representation on perceptual discrimination and identification performances of the model primary auditory cortex. We show that categorical perception may arise as a result of enlarged cortical representation induced, for instance, by early experience.

## Methods

### *Modeling the frequency representations in the primary auditory cortex.*

The parameters of the model were chosen based on properties of the primary auditory cortical neurons documented in the literature and our unpublished observations. The firing rates of the neurons in the auditory cortex exhibit significant variability. We have recorded response magnitude to tone pips from 121 AI neurons and obtained a mean Fano factor value of  $0.98 \pm 0.21$  (SD), suggesting that the neuronal firing may be modeled with a Poisson process. An earlier study showed more reliable responses of auditory cortical neurons (DeWeese et al., 2003). The difference in observations may be due to the use of different stimuli, experimental conditions, and recording techniques. We modeled the frequency representations in the primary auditory cortex with a population of Poisson-firing neurons. Each neuron had a Gaussian-shaped, response-frequency tuning curve as

$$R = \alpha e^{\frac{-(\text{frequency} - \mu)^2}{2\sigma^2}} + \delta \quad (1)$$

The maximum response magnitude,  $\alpha$ , was set to 1 spike/tone recorded in a 50-ms response window for all neurons. Spontaneous firing rate of the model neurons was set at 1 spike/second, corresponding to  $\delta = 0.05$  spike in the 50-ms response window. The tuning bandwidth of the model neurons, defined here as two standard deviations of the Gaussian-shaped tuning curve, was set at 1 octave (i.e.,  $\sigma = 0.5$  octave). These response parameters were chosen for simplicity, and are consistent with properties of recorded neurons (Bao et al., 2001; Bao et al., 2004). We also examined in later sections how changes in these parameters impacted perception behaviors. For the model naïve AI, the best frequencies (BFs), corresponding to  $\mu$  in equation 1, were equally spaced along the logarithmic frequency scale from 1 to 50 kHz (Figure 2.1A). To model the AI of a 7.1-kHz-tone-exposed animal, best frequencies of model neurons in a range of  $7.1 \text{ kHz} \pm 1$  octave were shifted to have a Gaussian distribution centered at 7.1 kHz and a standard deviation of 0.1 octave (Figure 2.1B).

### *Modeling frequency discrimination.*

The response of the  $i^{\text{th}}$  neuron of the model AI to a tone of frequency  $f$  was denoted as  $R_i$ —the number of spikes in the response window. As the model neurons fire spikes in a Poisson-random fashion, the probability of the neuron responding to  $f$  with response  $R_i$  is

$$P(R_i | f) = \frac{T_i(f)^{R_i}}{R_i!} e^{-T_i(f)} \quad (2)$$

where  $T_i$  is the neuron's response-frequency tuning curve. In practice,  $R_i$  is simulated with a Poisson-random number with a mean of  $T_i(f)$ .

The responses of all the  $N$  model neurons (1, 2, ...  $N$ ) to a tone of an unknown frequency  $f$  were simulated as  $R_1, R_2, \dots, R_N$ . We obtained the maximum likelihood estimate of  $f$ , denoted as  $F$ , by maximizing the following log-likelihood ( $LL$ ) function (Jazayeri and Movshon, 2006; Seung and Sompolinsky, 1993)

$$\begin{aligned}
LL(F) &= \sum_{i=1}^N \ln P(R_i | F) = \sum_{i=1}^N \ln \left( \frac{T_i(F)^{R_i}}{R_i!} e^{-T_i(F)} \right) \\
&= \sum_{i=1}^N R_i \ln T_i(F) - \sum_{i=1}^N T_i(F) - \sum_{i=1}^N \ln(R_i!) \quad (3)
\end{aligned}$$

using a sequential quadratic programming method available in Matlab toolboxes (Powell, 1977).

The  $F$  may be regarded as the “percept” of  $f$  by the model AI. As the model neuronal responses are stochastic, the estimated frequency  $F$  may deviate from the true frequency  $f$ . Such variability of “perception” of the model AI determines its frequency decoding precision and frequency discrimination capability. To estimate this variability, we presented the model AI with pairs of tones of the same frequency (i.e.,  $f_1 = f_2$ ). We then calculated the estimates of the two frequencies,  $F_1$  and  $F_2$ , and their difference,  $\Delta F_0 = |F_1 - F_2|$ . This was done 100 times to obtain the distribution of the  $\Delta F_0$ s (Figure 2.2). The 50th percentile of all these  $\Delta F_0$ s was chosen as the threshold, so that random chance-level performance would be 50%. To determine discrimination performance for a pair of different frequencies (i.e.,  $f_1 \neq f_2$ ), the difference of the estimates ( $\Delta F$ ) was calculated 100 times to obtain the distribution of the  $\Delta F$ . A  $\Delta F$  greater than the  $\Delta F_0$  threshold indicated a pair of tones was discriminated by the model AI. The percentage of successful discriminations was used as the performance level. Presented in all graphs were means of 200 performance levels in each testing condition. The variability of the performance was measured with 95% confidence intervals, which cover the range of 2.5th and 97.5th percentile of the performance levels.

### *Modeling frequency identification.*

In a typical behavioral identification task, the subject is presented with an unknown stimulus ( $f_x$ ) and asked to make a forced choice on which of two fixed stimuli ( $f_1$  and  $f_2$ ) is more likely to be the unknown stimulus. In our simulation, the model AI was presented with an unknown frequency ( $f_x$ ). The response of model AI to  $f_x$  was denoted as  $\mathbf{R}_x$ . The task was to determine which of two known frequencies ( $f_1$  and  $f_2$ ) was more likely to be the one that activated  $\mathbf{R}_x$ . We modeled the perceptual decision process in the frequency identification task with a stochastic process and a deterministic process.

The stochastic model assumes that the decision-making is a Bernoulli-stochastic process with two alternative outcomes—i.e., the model AI chooses either the low frequency  $f_1$  or the high frequency  $f_2$ . We first calculated the log of the ratio of the likelihood that  $\mathbf{R}_x$  was activated by  $f_1$  over the likelihood that it was activated by  $f_2$  as

$$LLR(f_x, f_1, f_2) = \ln \left( \frac{P(\mathbf{R}_x | f_1)}{P(\mathbf{R}_x | f_2)} \right) = \sum_{i=1}^N \left( R_i \ln T_i(f_1) - T_i(f_1) - R_i \ln T_i(f_2) + T_i(f_2) \right) \quad (4)$$

The probability that the model AI selected  $f_1$  as supposed to  $f_2$  is determined as

$$P(\text{choosing } f_1) = \frac{LLR(f_x, f_1, f_2) - LLR(f_2, f_1, f_2)}{LLR(f_1, f_1, f_2) - LLR(f_2, f_1, f_2)} \quad (5)$$

in which  $LLR(f_1, f_1, f_2)$  is the log-likelihood ratio that a  $f_1$ -activated response is activated by  $f_1$  as suppose to  $f_2$ , and  $LLR(f_2, f_1, f_2)$  is the log-likelihood that a  $f_2$ -activated response is activated by  $f_1$  as suppose to  $f_2$ . With fixed  $f_1$  and  $f_2$ ,  $LLR(f_1, f_1, f_2)$  and  $LLR(f_2, f_1, f_2)$  are also fixed and were calculated. Thus, the probability of categorizing a frequency  $f_x$  as  $f_1$  is a scaled linear function of the  $LLR(f_x, f_1, f_2)$ , which is a function of the unknown frequency  $f_x$ . The probability P is bounded in the range of 0 to 1.

The deterministic model assumes that the decision-making in the frequency identification task is determined solely by the perceptual process. The following likelihood ratio

$$LR(f_x, f_1, f_2) = P(\mathbf{R}_x | f_1) / P(\mathbf{R}_x | f_2) = \exp\left(\sum_{i=1}^N (R_i \ln T_i(f_1) - T_i(f_1) - R_i \ln T_i(f_2) + T_i(f_2))\right) \quad (6)$$

was used to directly determine which frequency to choose in the frequency identification task— $f_1$  was chosen if the ratio was greater than one, and  $f_2$  was chosen otherwise. This method is referred to as the likelihood ratio (LR) method. In addition, we have also modeled the stimulus identification process with a maximum likelihood estimation method (Equation 2), in which  $f_x$  was decoded from  $\mathbf{R}_x$ , denoted as  $F_x$ . The frequency (either  $f_1$  or  $f_2$ ) that was closer to  $F_x$  was chosen. This method is referred to as the maximum likelihood estimation (MLE) method.

Each simulation was run 100 times, and the percentage that the model AI chose  $f_1$  was used as the identification index. Each point in all graphs was the mean of 200 individually calculated identification indices in the specific testing condition. The variability of the performance was measured with 95% confidence intervals, which cover the range of 2.5th and 97.5th percentile of the identification indices.

#### *Testing stimulus discrimination in adult rats.*

All procedures are approved by Animals Care and Use Committee of University of California, Berkeley. Five female Sprague-Dawley rats (200-300 gm, over 2 months old) were trained in a tonal frequency discrimination task. Training and testing took place in a wire cage located in an anechoic sound-attenuation chamber. Upon automatic initiation of a trial, tone pips of 100-ms duration and of a standard frequency were played five times per second through a calibrated speaker. All tones were played at a 60-dB sound pressure level. After a random duration of 5-35 seconds, tone pips of a target frequency were played in the places of every other standard tone pips. Rats were trained to detect the frequency difference and make a nose-poke in a nosing hole within 3 seconds after the first target tone, which was scored as a hit and rewarded with a food pellet. False alarm rate was determined as the percentage responses to probe trials, in which the target frequency was the same as the standard frequency. In each training day, an animal was allowed to achieve 200-300 hits. The difference between the target and the standard tone pips were varied. The animal's performance may be influenced by its motivational states and its internal response criteria. To account for these factors and to estimate the animals' *perceptual* ability, we used the discrimination index (Grier, 1971; Pollack and Norman, 1964):

$$A' = \frac{1}{2} + \frac{[(h - fa)(1 + h - fa)]}{[4h(1 - fa)]} \quad (7)$$

in which  $h$  is the hit rate and  $fa$  is the false alarm rate.  $A'$  varies from 0.5 to 1, which allows comparison with the discrimination probability of the cortical model (see below).

*Testing stimulus identification in adult rats.*

Animals were first trained to recognize two prototype tonal frequencies. In each trial, 100-ms tone pips of a prototype frequency were played at a rate of 5 pips per second and at 60 dB SPL. The animal was trained to make a nose-poke in one of two nosing holes (either on the left or on the right) depending on which one of the two prototype frequencies (6 kHz or 12 kHz) was being played—i.e. a identification task. A nose-poke in the correct hole within ten seconds from the onset of the sound was considered a “hit” and rewarded with a food pellet. A nose-poke in the wrong hole or inaction in the 10-s period was a miss and not rewarded. It takes approximately 10 days for naive animals to reach an asymptotic performance level of approximately 80% correct recognition. Then, we tested how animals perceived and categorized a series of 9 tones of intermediate frequencies. These frequencies were logarithmically equally spaced between the two prototype frequencies. The prototype sounds were tested in regular trials (80% of all trials). The intermediate sounds were tested in probe trials (20%), in which the animal did not receive a food pellet regardless of the animal’s response. We did not reinforce the animals in these trials to avoid biasing their responses, which could interfere with the perceptual tests. To keep animals motivated with food pellet reward, we included 80% regular trials in which correct responses to prototype stimuli were rewarded. The percentage of trials that animals made nose-poke in the left nosing hole (corresponding to the lower frequency) was used to construct the identification function.

## **Results**

### **Psychometric function of the model AI**

We first examined the model performances as a function of the input frequency difference and the total number of neurons in the model AI. As shown in Figure 2.3, the psychometric performance-difference function was approximately sigmoidal. Having more model neurons improved the model performance, as indicated by a leftward shift of the psychometric function. The shape of the psychometric function, however, did not change with the neuron numbers. As predicted (Seung and Sompolinsky, 1993), the discrimination threshold of the model AI, as measured with the half-height frequency difference, was inversely proportional to the square root of the number of neurons (Figure 2.4A).

We examined animal performance in a frequency discrimination task, in which discrimination of various frequency differences was tested in adult rats that have not been exposed to specific sound (hereafter referred to as naïve animals, in contrast to sound-exposed animals with altered frequency representations). The psychometric function of naïve rats was sigmoidal, similar to that of the model AI. Furthermore, the performances of the model AI with 800 neurons fitted well with the animal performances. The total number of neurons in the primary auditory cortex of the rat (1-2 mm<sup>2</sup> in size) is on the order of 100,000, including local and inhibitory neurons (Cherniak, 1990; O'Kusky and Colonnier, 1982). The relatively small number of neurons required for the model to reach the performance levels of the animals is consistent with earlier modeling results (Paradiso, 1988). All simulations presented in the subsequent sections used model AIs with 800 neurons.

The tuning bandwidth, response magnitude and spontaneous firing rate of the model neurons were also varied to examine how these properties influence perceptual discrimination behaviors of the model AI. Frequency discrimination threshold decreases with greater response magnitude, narrower tuning bandwidth and lower spontaneous firing rate (Figure 2.4, B-D). These results provide constraints for further comparison between model and animals performances.

### **Perceptual discrimination by sound-exposed model AI.**

One of the two behavioral traits of categorical perception is that the perceptual discrimination ability is worse within a category than between different categories. If a perceptual category forms around the experienced stimulus, perceptual discrimination would be relatively poor within the category. We constructed a sound-exposed model AI, incorporating sound exposure-induced plasticity effects: over-representation of the experienced frequency and under-representation of neighboring frequencies in the range of  $\pm 1$  octave (See Figure 2.1B and Chang and Merzenich, 2003). Simulation results indicate that discrimination of 0.1-octave frequency differences in the over-represented frequency range was significantly impaired. By contrast, discrimination of neighboring frequencies was improved (Figure 2.5).

These results may be understood in terms of the amount of Fisher information the model neurons provide for frequency decoding (Dayan and Abbott, 2001). Sensory neurons contribute to stimulus decoding by changing their firing rates (Bala et al., 2003; Luna et al., 2005; Paradiso, 1988). Two similar stimuli that are near the center of a

Gaussian-shaped tuning curve of a neuron will elicit similar firing rates (close to the maximum response magnitude). However, two similar stimuli that fall on the slopes of a neuron's tuning curve, where firing rate is most sensitive to stimulus differences, will elicit responses of very different firing rates. In the sound-exposed AI, a large number of neurons become tuned near the experienced frequency. These retuned neurons are less sensitive to changes in frequencies near the experienced tone, because those frequencies fall near the center of their tuning curves. Instead, these neurons become sensitive to frequency changes in the neighboring frequency bands, where the slopes of the tuning curves are located. The limit of decoding accuracy set by Fisher information measure can be attained by maximum likelihood estimation, when a large number of neurons are involved in coding (Dayan and Abbott, 2001). Thus, discrimination thresholds derived from Fisher information should be similar to those calculated with MLE.

### **Perceptual identification by sound-exposed model AI.**

The second behavioral trait of categorical perception is the sigmoidal identification function where stimuli on one side of a categorical boundary are classified as members of the same category. Behaviorally, it is often tested with an identification task, in which animals are required to classify a series of equally spaced stimuli into two categories. We performed frequency identification test in naive animals, and observed a near-linear frequency identification function (Figure 2.6). Using this result as a constraint, we explored three methods to model the stimulus identification process—a Bernoulli-stochastic process method, a likelihood-ratio threshold method and a maximum-likelihood estimation method (see the Methods section for details). Among the three methods, only the Bernoulli random process method produced near linear identification function for naïve model AI. The performances of the likelihood-ratio threshold (LR) and maximum-likelihood estimation (MLE) methods were almost identical, and were pooled together (Figure 2.6). The LR/MLE methods produced an inverted sigmoidal identification function that diverges from the corresponding animal behavior. The identification function generated with these two methods shows a complete categorical transition within a 0.2-octave frequency distance, similar to the frequency discrimination threshold shown in Figure 2.3. This is not surprising because the methods essentially perform frequency decoding, and then make perceptual decisions based on the decoded frequency. The result that the model AI performed equally well in identification and discrimination tasks when LR/MLE methods are used is inconsistent with experimental findings that animals generally perform worse in identification than in discrimination tasks (For a discussion, see Massaro, 1987), suggesting that the LR/MLE methods are inappropriate as models of the perceptual identification processes. The difference between the Bernoulli-stochastic and LR/MLE methods is likely due to their different assumptions about the decision-making process—the Bernoulli stochastic method assumes that the decision-making is stochastic, and the LR/MLE methods assume that the decision-making is deterministic (see Methods).

Comparison of likelihood measures has been proposed as a model of the perceptual decision processes (Green and Swets, 1966). In simple stimulus difference detection tasks (e.g., stimulus discrimination), subjects may compare a likelihood of

having perceived stimulus differences with a threshold value to make a perceptual decision (as in the frequency discrimination process described above). Thus the performance is limited by the frequency decoding ability. In the perceptual identification task, however, the stimulus differences are often supra-threshold—i.e.,  $f_x$  is perceived as different from both  $f_1$  and  $f_2$ . Deciding which one of  $f_1$  and  $f_2$  is closer to the unknown frequency  $f_x$  is likely a probabilistic process, not by simply comparing an index value to a fixed threshold. The notion that the discrimination and identification tasks involve different perceptual decision processes is consistent with the findings that performances are generally worse in identification than in discrimination tasks (Massaro, 1987). Figure 2.6 indicates that the performances of MLE/LR methods are as good as the performances of the model AI in a discrimination task, but deviate from the animal performance. Instead, a Bernoulli-random process with the choice probabilities described by the linearly scaled log-likelihood ratio may capture some aspects of the perceptual identification behaviors in an identification task.

We analyzed perceptual identification behaviors of the model 7.1-kHz-exposed AI using the Bernoulli-random process method. The results showed that the tone-exposed AI consistently classified frequencies near 7.1 kHz as the lower one (i.e., 5.9 kHz) of the two prototypes (Figure 2.7). This behavior, together with the reduced discrimination performance near 7.1 kHz, indicates that frequencies near 7.1 kHz were grouped into a perceptual category. It is a result of the sound exposure, because it only occurred near the exposed frequency, but not for the frequencies above 8.3 kHz.

### **Representations of two perceptual categories.**

The above results indicate that exposure to a single stimulus may shape a perceptual category near the stimulus. In psychophysical studies, categorical perception is typically defined between at least two categories by a peaked discrimination function and a sigmoidal identification function. We have also simulated cortical plasticity effects of exposure to two tones of different frequencies that were either 2 octaves or 0.5 octave apart (Figure 2.8). The characteristics of the plasticity effects were similar to those in earlier sections of the simulations—neurons that used to be tuned to within 1 octave of the exposed frequencies were retuned closer to the exposed frequencies, and the retuned best frequencies had a Gaussian distribution with a 0.1-octave standard deviation (same as those in previous sections, see first section of the Method). The neurons that used to be tuned to the frequencies in-between the two exposure frequencies were split equally between the two frequencies. Other neuronal response properties (tuning bandwidths, maximum response magnitudes and spontaneous firing rate) were unchanged. It should be noted that the specific forms of two-tone-induced cortical plasticity used in our simulation are hypothetical, simplified and extrapolated from single tone-induced effects (Chang and Merzenich, 2003; Zhang et al., 2001).

Simulation results indicate that when the two experienced frequencies were 2 octaves apart, the model two-tone-exposed AI showed categorical perceptual behaviors—a sigmoidal identification function and a peaked discrimination function. The discrimination function is similar to that of categorical discrimination of phonemes observed in animals (Kuhl and Padden, 1983). These results indicate that categorical perception may be mediated by populations of neurons with bell-shaped tuning curves. In addition, the prototypes of the categorically perceived stimuli are over-represented—

e.g., more neurons were tuned to the categorically perceived frequencies near 3.5 and 14 kHz as shown in Figure 2.8 A-B. Interestingly when the two frequencies were 0.5 octave apart, no categorical perception was observed. Categorical perception would be established in this case if the tuning bandwidths of the neurons become narrower (data not shown). These results suggest that the properties of the cortical circuits constrain the categorical learning processes. Certain stimuli may be more learnable as categorical prototypes than the others.

Figure 2.9 illustrates the population categorical responses of the model AI in comparison to non-categorical responses. The tones of the same frequencies activated overlapping and gradually shifting populations of neurons in the model naïve AI (Figure 2.9A). In the model AI that had over-representations of 3.5 and 14 kHz tones, two distinctive populations of neurons were activated by tones in the two different categories (Figure 2.9B). In the model AI with over-representations of 5.9 and 8.3 kHz (Figure 2.9C), the same population of neurons was activated by a range of frequencies near the over-represented ones. These activation patterns are consistent with the results in Figure 2.8, showing categorical perception when the over-represented frequencies are two octaves apart.

We also varied neuronal properties—i.e., tuning bandwidth, response magnitude and spontaneous firing rate—and examined how they influence categorical sound representation. The model AI used was the same as described in Figure 2.8A, having over-representations of 3.5 and 14 kHz. Altering tuning bandwidth had a profound impact on categorical representation (Figure 2.10, A and D). When the bandwidth was in the range between one and two octaves, we observed two perception categories at the two over-represented frequencies. When bandwidth was four octaves, the 0.1-octave frequency difference was perceived equally poorly across the tested frequency range, and the identification function was close to linear, suggesting that there was no perceptual categories. When bandwidth was 0.5 octave, there appeared to be three categories. Changing response magnitude altered discrimination performances, but not frequency identification performances (Figure 2.10, B and E). Altering the level of spontaneous had impact only on frequency discrimination, but not on frequency identification performances (Figure 2.10, C and F).

## **Discussion**

Categorical perception may be learned by exposure to specific stimuli during early development, or by extensive training in adulthood (Goldstone, 1994; Lasky et al., 1975; MacKay et al., 2001; Williams, 1977). After learning, the stimuli within a stimulus category are perceived as being more similar, and stimuli from different categories are perceived as being more different. These two forms of perceptual alterations are referred to as acquired perceptual equivalence and distinctiveness, respectively (Liberman et al., 1957). They are believed to underlie categorical perceptual behaviors—e.g., peaked discrimination functions and sigmoidal identification functions. Recently electrophysiological studies have revealed that sensory exposure and perceptual training often enlarge cortical representations of the relevant stimuli by retuning of neuronal selectivity to the stimuli. In the present study, we examined the possibility that enlargement in cortical representation is a cortical mechanism of categorical perception. Our computational simulation results indicate that the perceptual contrast of the over-represented stimuli may be reduced, analogous to acquired perceptual equivalence, and the perceptual contrast of the neighboring under-represented stimuli may be enhanced, resulting in acquired perceptual distinctiveness. Thus, a perceptual category may form for the over-represented stimuli. Further analysis of the model AI with two over-represented stimulus ranges revealed behaviors characteristic of categorical perception—a peaked discrimination function and a sigmoidal identification function. These results support the notion that enlargement in cortical representation mediates learned categorical perception.

Previous electrophysiological studies have investigated neural mechanism of categorical perception by identifying categorical neurons—those respond to all members of one category but not to any members of other categories. These neurons may be regarded as the category readout neurons. It is still unclear what kind of transformation of sensory information gives rise to this category-selectivity and where the transformations take place. Results of the present study suggest that experience-dependent reorganization of stimulus representations in the primary sensory cortex could provide the transformation underlying learned categorical perception. In the sensory cortex, sensory information and hence perceptual categories are represented in populations of neurons, each of which shows graded responses to a large range of stimuli. There must be readout mechanisms to transform this distributed categorical representation into categorical responses in single neurons. In the present study, we obtained categorical perceptual behaviors in the models of AI using analyses of likelihood measures. Whether and how the neural systems perform likelihood analysis is still under active investigations, and some models have been proposed (Jazayeri and Movshon, 2006; Zhang et al., 1998). These models may provide the needed readout mechanisms to transform distributed categorical representations into categorical responses in single neurons.

Several computational models of categorical learning have been investigated in earlier studies such as unsupervised, auto associative feedback networks (Anderson et al., 1977) and supervised, multi-layered networks with a hidden layer and back-propagating error signals (Harnad et al., 1991). The construction of these models was primarily based on theoretical considerations, and the biological plausibility of some of the mechanisms (e.g., the back-propagation of error signal) is unclear. In the present

study, the model auditory representations were based on findings of electrophysiological studies—e.g., more neurons become tuned to more frequently experienced frequency. We only considered the cortical decoding capacity and how it would influence animals' perceptual performances. We did not provide accounts on how the experience-altered cortical decoding capacity can be transformed into categorical neuronal response and guide perceptual behaviors (i.e., the readout problem). The shaping of categorical perception with sensory exposure described in the present study is similar to the learning of perceptual categories by the auto-associative network in that both are unsupervised learning and the learned perceptual categories are represented in distributed population responses (Anderson et al., 1977). The acoustic representations modeled in the present study may also be analogous to the hidden layers of the multilayer network models, which may be altered by experience in animals, and by learning in the multilayer network models (Harnad et al., 1991). Studies of sensory plasticity may provide insights for constructing biologically plausible models of categorical learning.

The results of this study provide some insights into cortical mechanisms of perceptual learning. Enlarged cortical representations of relevant stimuli have been observed after extensive training of adult animals to discriminate tonal frequencies (Recanzone et al., 1993), sound levels (Polley et al., 2004; Polley et al., 2006), temporal modulation rates (Bao et al., 2004), or somatosensory stimuli (Recanzone et al., 1992). Some of the studies show that representational sizes are highly correlated with tonal frequency discrimination performances after perceptual training (Recanzone et al., 1993). These results lead to the notion that greater cortical representations are the neural basis for better perceptual discrimination performance. Such a simplistic view, however, has been challenged by opposite results showing that perceptual discrimination training sometimes does not alter the cortical feature representational map (Brown et al., 2004). Furthermore, animals with cortical representations of certain tonal frequencies enlarged by intracortical electrical stimulations did not show any improvement in stimulus discrimination performances in the over-represented frequency range (Talwar and Gerstein, 2001). These results suggest that perceptual discrimination capability may be determined by many cortical neuronal properties, and not just by representational sizes. This is consistent with the simulation results of the present study, which shows that enlarged representations of a very narrow frequency range may cause impaired discrimination of the over-represented frequencies. Our modeling results also indicate that over-represented frequencies may be discriminated better if the tuning bandwidths of the neurons become narrower (Figure 2.10A), or if a large range of frequencies are over-represented (not shown). These results help to reconcile the seemingly contradicting results reviewed above.

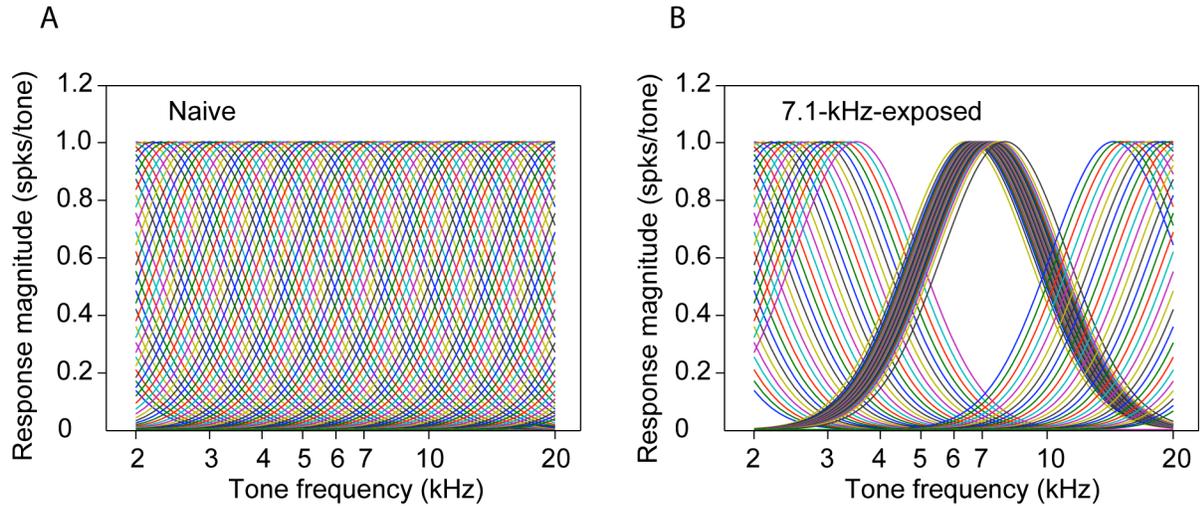
Maximum likelihood estimation is an optimal population decoding method. It is not a considered biologically realistic decoding mechanism, although certain neuronal architectures are thought to be able to perform similar computations (Jazayeri and Movshon, 2006; Zhang et al., 1998). In the limit of large numbers of encoding neurons and for Poisson firing rate distributions, its performance saturates the Cramer-Rao bound of the variance of estimate, and sets the upper limit of the performance of the biological systems (Dayan and Abbott, 2001; Seung and Sompolinsky, 1993). In essence, maximum likelihood estimation measures the maximum decoding *capacity* of

a representational system. It has been used to model visual discrimination processes (Paradiso, 1988). Although the successful applications of the method do not imply that brain decodes sensory information using a similar maximum likelihood decoding method, it does indicate that perceptual behaviors are correlated with stimulus decoding capacity of the neuronal network revealed by the method. We followed the same rationale in our analysis of the impact of cortical plasticity effects on perceptual discrimination performance.

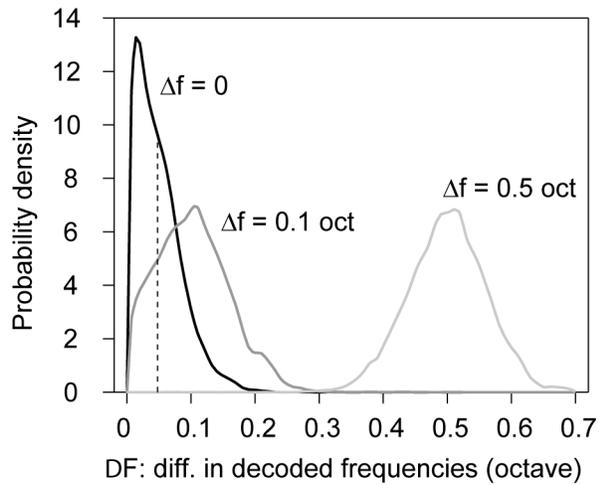
The information processing events underlying the perceptual identification behavior is unknown. The traditional view is that both discrimination and identification are mediated by the same perceptual processes so that their performances should match each other. Later experiments showed that the stimulus identification performance is generally worse than what would be predicted from discrimination functions (Massaro 1987). In the present study, animals showed a nearly linear identification function across a large frequency range. Such a linear identification function is inconsistent with a purely discrimination-based identification process, which would have yielded sigmoidal identification function like that of the MLE/LR group in Figure 2.6. We modeled identification behaviors in two steps—first, the choice probability is determined with the log-likelihood ratio, and, second, a Bernoulli random process determines the identification choices. The two steps may correspond to the two separate processes underlying identification behaviors—sensory decoding and decision-making.

In this study, we simplified neuronal tuning properties—all neurons have the same firing rate, tuning bandwidth and spontaneous firing rate. Essentially same results were obtained with model neurons whose properties have the same distributions as those of recorded neurons (data not shown). The sound exposure-induced cortical plasticity effects were also simplified in this study, and only changes in the tuning frequencies were included in the analysis. Other neuronal response properties, such as the shapes of the tuning curves, the maximum response magnitudes, spontaneous firing rates, and spike timing/correlation can also be altered either by sound exposure or by perceptual learning (Bao et al., 2001; Beitel et al., 2003; Blake et al., 2006; Brown et al., 2004; Chang and Merzenich, 2003; Chowdhury and Suga, 2000; Edeline and Weinberger, 1993; Engineer et al., 2004; Fritz et al., 2003; Kilgard and Merzenich, 1998; Kilgard et al., 2001; Ma and Suga, 2003; Ohl and Scheich, 1996; Polley et al., 2004; Recanzone et al., 1993; Schoups et al., 2001; Zhang et al., 2001). Those forms of cortical plasticity effects could also contribute to the learning of categorical perception. Nevertheless, our analysis demonstrates that the enlargement of cortical representations could be a mechanism for categorical perception. Systematic examinations of categorical perception in animals that have been exposed to controlled sensory input would provide new insights into the neural mechanisms of categorical perceptual learning.

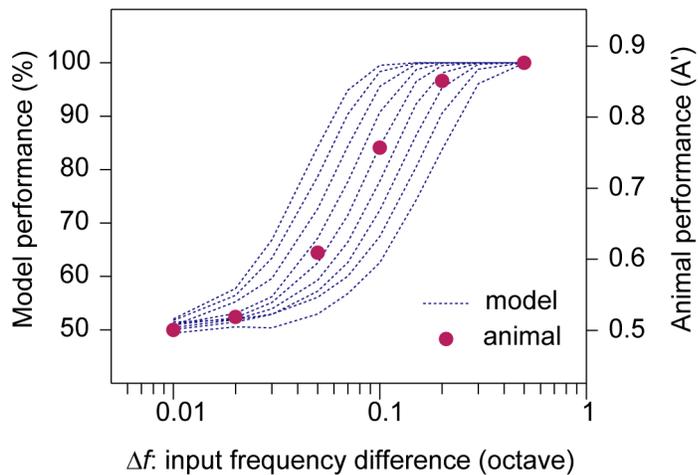
## Figures



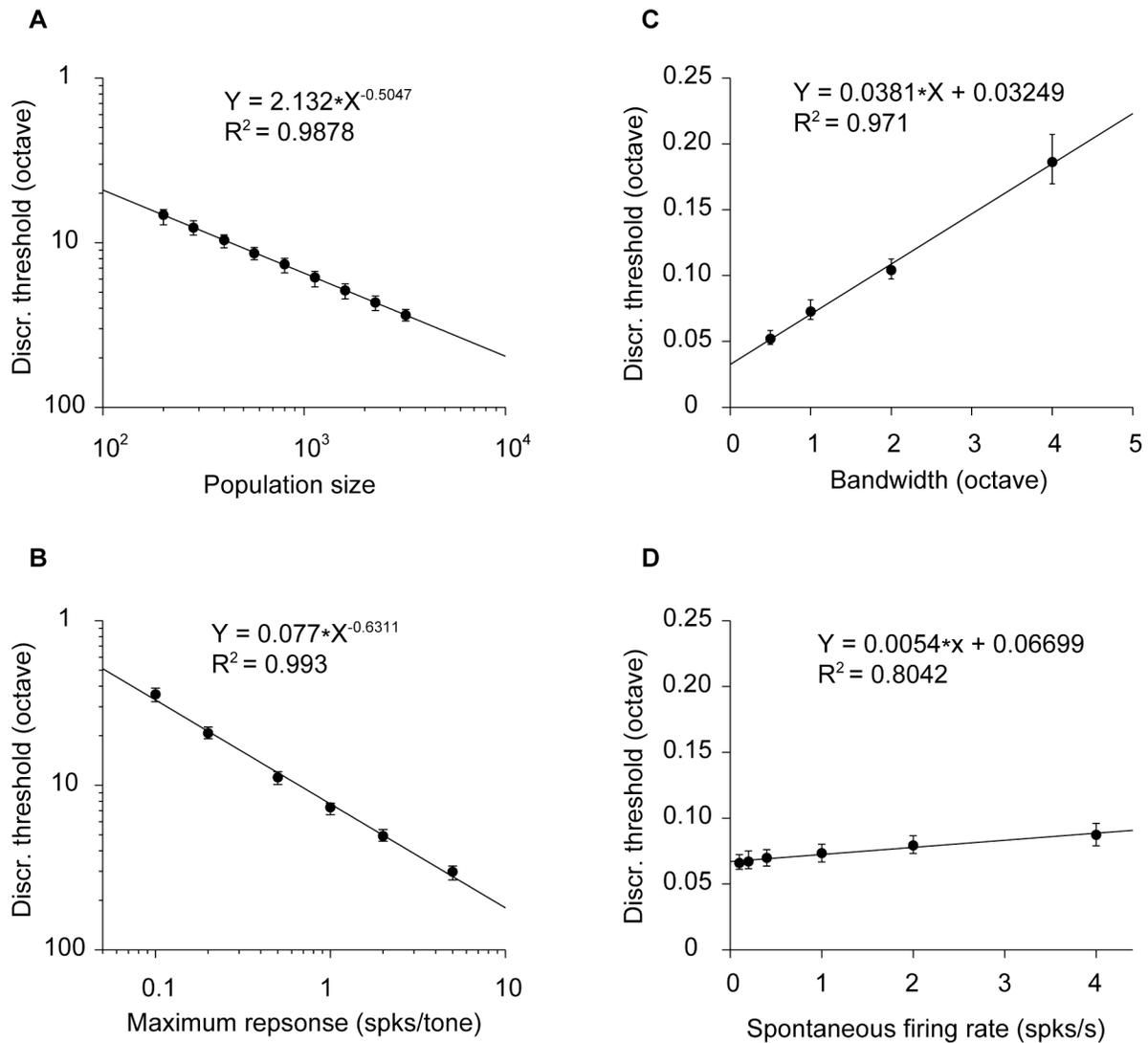
**Figure 2.1.** Tuning curves of the model neurons in the model naïve primary auditory cortex (**A**) and the model 7.1-kHz-exposed primary auditory cortex (**B**). All model neurons have a maximum response magnitude of one spike per tone and a tuning bandwidth of one octave. Spontaneous firings of the model neurons are not shown in the tuning curves.



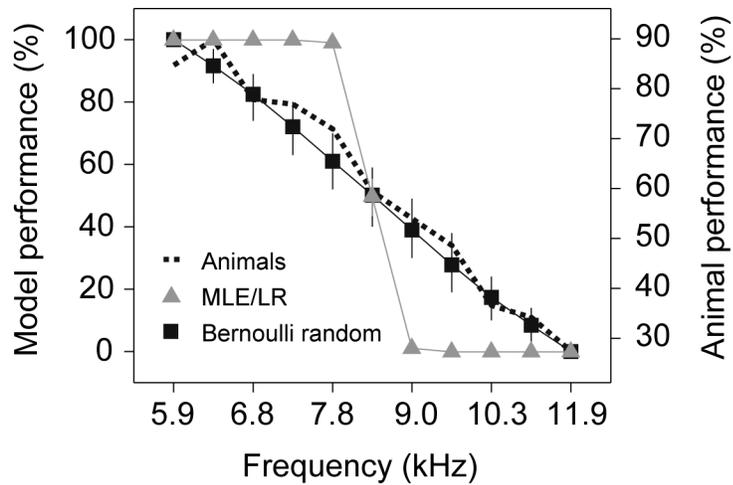
**Figure 2.2.** Example of distributions (probability density functions) of the differences between the decoded frequencies (i.e.,  $\Delta F$ , where  $F$  denotes the decoded frequency) with the input frequencies being 0, 0.1 and 0.5 octave apart in a model naïve AI.  $\Delta F$  may be considered as perceived frequency difference by the model AI. The distribution is shifted towards the right with increasingly greater differences in the input frequencies ( $\Delta f$ ). The vertical dashed line marks the 50<sup>th</sup> percentile  $\Delta F$  value of the distribution with  $\Delta f = 0$ . This value was used as the discrimination threshold. Any two tones that produced  $\Delta F$ s greater than this threshold value were considered discriminated.



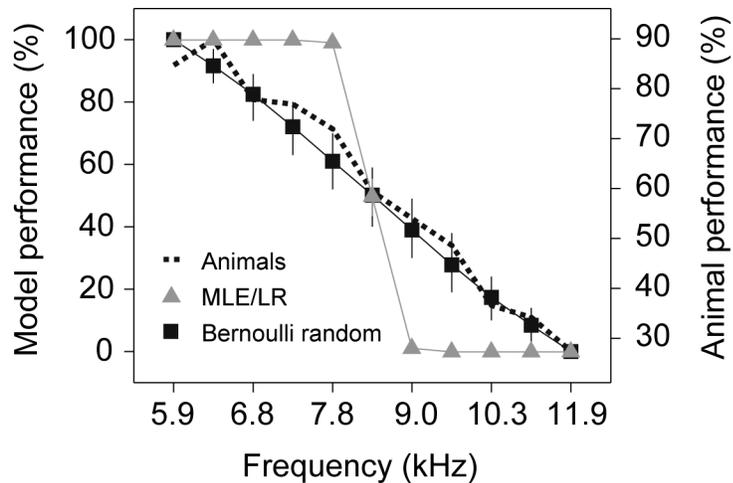
**Figure 2.3.** Discrimination performances of the model naïve AI and animals as a function of differences between the input frequencies. The model performance was quantified with the proportion of all trials in which two frequencies were discriminated. The dashed lines show the performances of nine different models each using a different number of neurons. From left to right, the number of model neurons are 3200, 2262, 1600, 1131, 800, 565, 400, 282, 200. Animal performance, shown as circles, was quantified with the discrimination index  $A'$ . The performance of the model with 800 neurons matched that of the naïve animals closely. Therefore, 800 model neurons were used in all subsequent simulations.



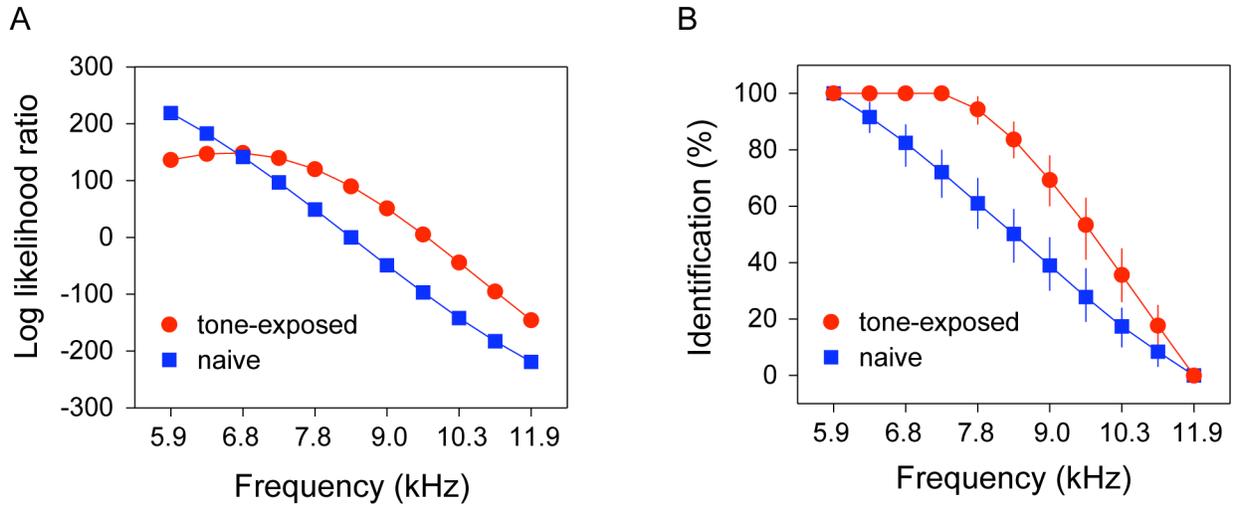
**Figure 2.4.** Influences of neuronal response properties on the discrimination threshold of model naïve AI. Discrimination threshold was determined as the frequency difference at which the model AI performance is 75% (see figure 3). Discrimination threshold decreases with the total number of neurons in the model AI (**A**) and with the maximum response magnitude—i.e., the peak height of the Gaussian tuning curve (**B**). It increases with neuronal tuning bandwidth (**C**) and with spontaneous firing rate (**D**). A linear regression indicates that the threshold is inversely proportional to the square root of the population size. The parameters that were not systematically varied were given the following values: Tuning bandwidth, 1 octave; Response magnitude, 1 spike/tonne; spontaneous activity 1 spike/second.



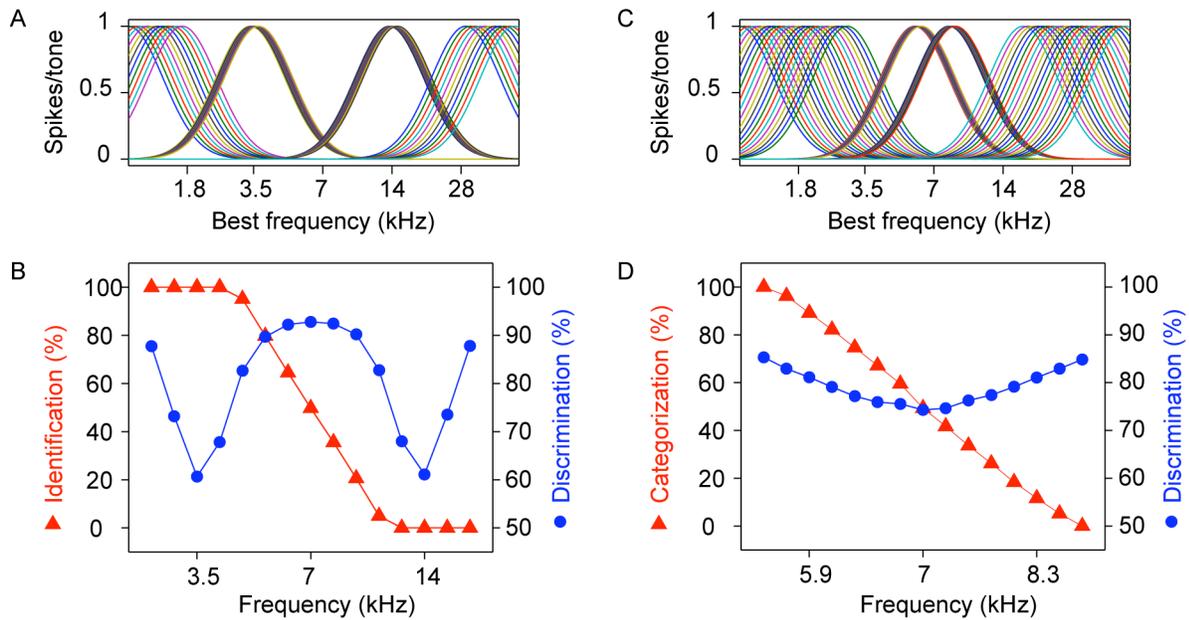
**Figure 2.5.** Impaired discrimination of the over-represented frequencies. The model naïve AI and 7.1-kHz-exposed AI were tested on frequency discrimination ability of various tone-pip pairs with  $\Delta f = 0.1$  octaves. The model naïve AI performed the frequency discrimination task with the same success rate across all frequencies (dashed line). Model 7.1-kHz-exposed AI, which had enlarged representations near 7.1 kHz, showed impaired discrimination performance near 7.1 kHz, and improved performance in the neighboring 11.9-kHz frequency range. The performance was quantified using the proportion of trials in which the two frequencies were discriminated.



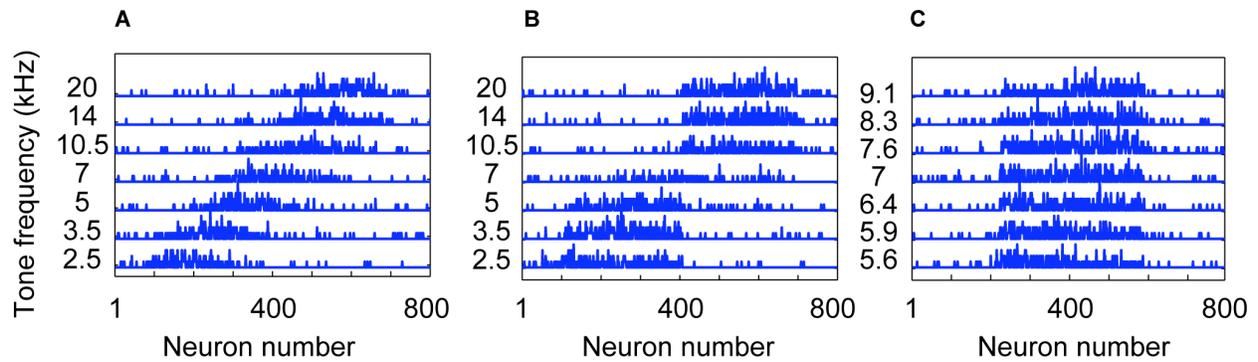
**Figure 2.6.** Comparison of naïve animal performances with model naïve AI performances in a frequency identification task. The animals showed a near linear identification function (dashed line). Naïve model AI performance in the identification task was simulated with three methods, likelihood ratio, maximum likelihood estimation and Bernoulli stochastic process (for details, see Methods). The first two methods yield identical results, which differ from the animal performance. The Bernoulli stochastic choice method produced near linear identification function, and was used for later analysis of model identification behaviors. The performance was measured with the percentage of trials in which the tonal frequency was identified as the low frequency of 5.9 kHz.



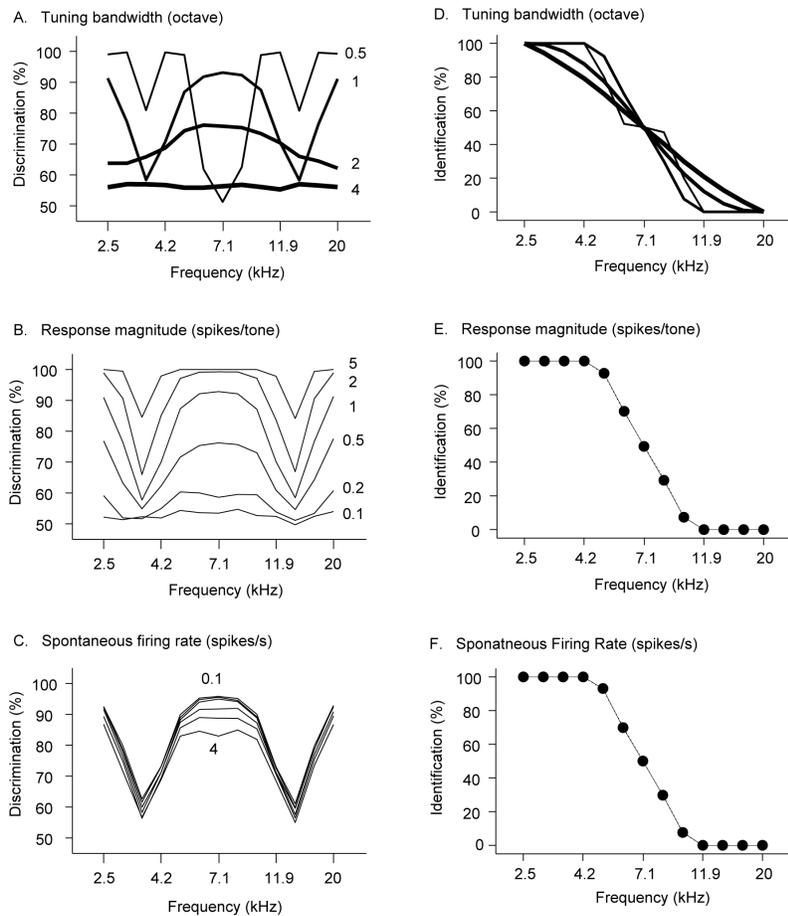
**Figure 2.7.** Frequency identification by the model naïve AI and model 7.1-kHz-exposed AI. **(A)** The ratio of the log-likelihood that an unknown frequency is 5.9 kHz vs. the log-likelihood that the frequency is 11.9 kHz. **(B)** The percentage of trials that a frequency in the range from 5.9 to 11.9 kHz is identified as 5.9 kHz by a naïve model AI and a 7.1-kHz-exposed AI is shown (see Methods for details). The 7.1-kHz-exposed AI consistently identifies several tones as 5.9 kHz, building what looks like a perceptual category for lower-frequency tones.



**Figure 2.8.** Representations of two perceptual categories. **(A)** Model AI with over-representations of 3.5-kHz and 14-kHz tones. The neighboring frequencies were under-represented. The over- and under-representations were similar to those described in Figure 1B. **(B)** The model AI described in **A** showed categorical perception of the frequencies with categorical centers at 3.5 and 14 kHz. The model AI showed a sigmoidal identification function and better discrimination of 0.1-octave frequency difference at the categorical boundary at 7 kHz. **(C)** Model AI with over-representations of 5.9- and 8.3-kHz tones. **(D)** The model AI described in **C** did not show categorical perception.



**Figure 2.9.** Categorical population responses. Activity of various model AIs, each having 800 model neurons. Each row shows the responses of all neurons to a single presentation of a tone. The frequency of the tone is indicated on the left. The bar represents the number of spikes that a neuron discharged in response to the tone. Neurons are arranged by best frequency such that the neuron with the lowest best frequency is to the left on the x-axis. **(A)** Activity of model naïve AI. Tones of different frequencies activated different but overlapping populations of neurons. **(B)** Activity of the model AI with over-representations of 3.5- and 14-kHz tones. The response patterns activated by 2.5-, 3.5- and 5-kHz tones were similar, and were much different from those activated by 10.5-, 14- and 20-kHz tones. **(C)** Activity of the model AI with over-representation of 5.9- and 8.3-kHz tones. Tones of frequencies from 5.6 to 9.1 kHz all activated the same population of neurons.



**Figure 2.10.** Influences of neuronal response properties on categorical representation of frequencies. We investigated how changes in tuning bandwidth, response magnitude and spontaneous activity influence discrimination of 0.1-octave frequency differences (A-C) and identification of frequencies (E-F) in a model with over-representations of 3.5 and 14 kHz. (A) and (D). Performances of models with four different neuronal tuning bandwidths: 0.5, 1, 2 and 4 octaves. Thicker line indicates broader tuning bandwidth. Tuning bandwidth has a strong influence on how the model discriminates and identifies frequencies. (B-C) and (E-F). Response magnitude and spontaneous firing rate influence discrimination but not identification performances. The response magnitudes are indicated in the figure in units of spikes/ton. Spontaneous firing rates were 0.1, 0.2, 0.4, 1, 2, 4 spikes/second. Frequency discrimination performance was generally better with higher response magnitude and lower spontaneous firing rates. Frequency identification performances were not changed by response magnitude or spontaneous firing rate—the performances overlapped completely. The parameters that were not systematically varied were given the following values: Tuning bandwidth, 1 octave; Response magnitude, 1 spike/ton; spontaneous activity 1 spike/second.

## References:

- Anderson, J. A., Silverstein, J. W., Ritz, S. A., and Jones, R. S. (1977). "Distinctive features, categorical perception, and probability learning: Some applications of a neural model." *Psychol Rev*, 84, 413-51.
- Bala, A. D., Spitzer, M. W., and Takahashi, T. T. (2003). "Prediction of auditory spatial acuity from neural images on the owl's auditory space map." *Nature*, 424(6950), 771-4.
- Bao, S., Chan, V. T., and Merzenich, M. M. (2001). "Cortical remodelling induced by activity of ventral tegmental dopamine neurons." *Nature*, 412(6842), 79-83.
- Bao, S., Chang, E. F., Woods, J., and Merzenich, M. M. (2004). "Temporal plasticity in the primary auditory cortex induced by operant perceptual learning." *Nat Neurosci*, 7(9), 974-81.
- Beitel, R. E., Schreiner, C. E., Cheung, S. W., Wang, X., and Merzenich, M. M. (2003). "Reward-dependent plasticity in the primary auditory cortex of adult monkeys trained to discriminate temporally modulated signals." *Proc Natl Acad Sci U S A*, 100(19), 11070-5.
- Blake, D. T., Heiser, M. A., Caywood, M., and Merzenich, M. M. (2006). "Experience-dependent adult cortical plasticity requires cognitive association between sensation and reward." *Neuron*, 52(2), 371-81.
- Bornstein, M. H., Kessen, W., and Weiskopf, S. (1976). "The categories of hue in infancy." *Science*, 191(4223), 201-2.
- Brown, M., Irvine, D. R., and Park, V. N. (2004). "Perceptual learning on an auditory frequency discrimination task by cats: association with changes in primary auditory cortex." *Cereb Cortex*, 14(9), 952-65.
- Burns, E. M., and Ward, W. D. (1978). "Categorical perception--phenomenon or epiphenomenon: Evidence from experiments in the perception of melodic musical intervals." *J Acoust Soc Am*, 63(2), 456-68.
- Chang, E. F., and Merzenich, M. M. (2003). "Environmental noise retards auditory cortical development." *Science*, 300(5618), 498-502.
- Cherniak, C. (1990). "The bounded brain: toward quantitative neuroanatomy." *J Cogn Neurosci*, 2, 58-68.
- Chowdhury, S. A., and Suga, N. (2000). "Reorganization of the frequency map of the auditory cortex evoked by cortical electrical stimulation in the big brown bat." *J Neurophysiol*, 83(4), 1856-63.
- Crozier, J. B. (1997). "Absolute pitch: practice makes perfect, the earlier the better." *Psychology of Music*, 25, 110-9.
- Dayan, P., and Abbott, L. F. (2001). *Theoretical Neuroscience*, The MIT Press, Cambridge, MA.
- DeWeese, M. R., Wehr, M., and Zador, A. M. (2003). "Binary spiking in auditory cortex." *J Neurosci*, 23(21), 7940-9.
- Edeline, J. M., and Weinberger, N. M. (1993). "Receptive field plasticity in the auditory cortex during frequency discrimination training: selective retuning independent of task difficulty." *Behav Neurosci*, 107(1), 82-103.
- Ehret, G. (1992). "Categorical perception of mouse-pup ultrasounds in the temporal domain." *Anim. Behav.*, 43(3), 409-416.

- Ehret, G., and Haack, B. (1981). "Categorical perception of mouse pup ultrasound by lactating females." *Naturwissenschaften*, 68(4), 208-9.
- Eimas, P. D. (1974). "Auditory and linguistic processing of cues for place of articulation by infants." *Percept Psychophys*, 16, 564-70.
- Engineer, N. D., Percaccio, C. R., Pandya, P. K., Moucha, R., Rathbun, D. L., and Kilgard, M. P. (2004). "Environmental enrichment improves response strength, threshold, selectivity, and latency of auditory cortex neurons." *J Neurophysiol*, 92(1), 73-82.
- Erickson, C. A., Jagadeesh, B., and Desimone, R. (2000). "Clustering of perirhinal neurons with similar properties following visual experience in adult monkeys." *Nat Neurosci*, 3(11), 1143-8.
- Freedman, D. J., Riesenhuber, M., Poggio, T., and Miller, E. K. (2001). "Categorical representation of visual stimuli in the primate prefrontal cortex." *Science*, 291(5502), 312-6.
- Fritz, J., Shamma, S., Elhilali, M., and Klein, D. (2003). "Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex." *Nat Neurosci*, 6(11), 1216-23.
- Goldstone, R. (1994). "Influences of categorization on perceptual discrimination." *J Exp Psychol Gen*, 123(2), 178-200.
- Green, D. M., and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*, John Wiley, New York.
- Grier, J. B. (1971). "Nonparametric indexes for sensitivity and bias: computing formulas." *Psychol Bull*, 75(6), 424-9.
- Harnad, S. (2003). "Categorical perception." *Encyclopedia of Cognitive Science*, L. Nadel, ed., Macmillan, London.
- Harnad, S., Hanson, S. J., and Lubin, J. (1991). "Categorical perception and the evolution of supervised learning in neural nets." *Working papers of the aaai spring symposium on machine learning of natural language and ontology*, L. Reeker, ed., Stanford, CA, 65-74.
- Harnad, S. R. (1987). *Categorical Perception: The Groundwork of Cognition*, Cambridge University Press, Cambridge.
- Holt, L. L., Lotto, A. J., and Diehl, R. L. (2004). "Auditory discontinuities interact with categorization: implications for speech perception." *J Acoust Soc Am*, 116(3), 1763-73.
- Jazayeri, M., and Movshon, J. A. (2006). "Optimal representation of sensory information by neural populations." *Nat Neurosci*, 9(5), 690-6.
- Kilgard, M. P., and Merzenich, M. M. (1998). "Cortical map reorganization enabled by nucleus basalis activity." *Science*, 279(5357), 1714-8.
- Kilgard, M. P., Pandya, P. K., Vazquez, J., Gehi, A., Schreiner, C. E., and Merzenich, M. M. (2001). "Sensory input directs spatial and temporal plasticity in primary auditory cortex." *J Neurophysiol*, 86(1), 326-38.
- Kluender, K. R., Diehl, R. L., and Killeen, P. R. (1987). "Japanese quail can learn phonetic categories." *Science*, 237(4819), 1195-7.
- Kuhl, P. K., and Miller, J. D. (1975). "Speech perception by the chinchilla: voiced-voiceless distinction in alveolar plosive consonants." *Science*, 190(4209), 69-72.

- Kuhl, P. K., and Padden, D. M. (1982). "Enhanced discriminability at the phonetic boundaries for the voicing feature in macaques." *Percept Psychophys*, 32(6), 542-50.
- Kuhl, P. K., and Padden, D. M. (1983). "Enhanced discriminability at the phonetic boundaries for the place feature in macaques." *J Acoust Soc Am*, 73(3), 1003-10.
- Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., and Lindblom, B. (1992). "Linguistic experience alters phonetic perception in infants by 6 months of age." *Science*, 255(5044), 606-8.
- Lasky, R. E., Syrdal-Lasky, A., and Klein, R. E. (1975). "VOT discrimination by four to six and a half month old infants from Spanish environments." *J Exp Child Psychol*, 20(2), 215-25.
- Lieberman, A. M., Cooper, F. S., Shankweiler, D. P., and Studdert-Kennedy, M. (1967). "Perception of the speech code." *Psychol Rev*, 74(6), 431-61.
- Lieberman, A. M., Harris, K. S., Hoffman, H. S., and Griffith, B. C. (1957). "The discrimination of speech sounds within and across phoneme boundaries." *J Exp Psychol*, 54(5), 358-68.
- Livingston, K. R., Andrews, J. K., and Harnad, S. (1998). "Categorical perception effects induced by category learning." *J Exp Psychol Learn Mem Cogn*, 24(3), 732-53.
- Luna, R., Hernandez, A., Brody, C. D., and Romo, R. (2005). "Neural codes for perceptual discrimination in primary somatosensory cortex." *Nat Neurosci*, 8(9), 1210-9.
- Ma, X., and Suga, N. (2003). "Augmentation of Plasticity of the Central Auditory System by the Basal Forebrain and/or Somatosensory Cortex." *J Neurophysiol*, 89(1), 90-103.
- MacKay, I. R., Flege, J. E., Piske, T., and Schirru, C. (2001). "Category restructuring during second-language speech acquisition." *J Acoust Soc Am*, 110(1), 516-28.
- Massaro, D. W. (1987). "Categorical partition: A fussy logical model of categorical behavior." *Categorical perception: the groundwork of cognition*, S. Harnad, ed., Cambridge University Press, Cambridge, UK, 254-83.
- Naatanen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R. J., Luuk, A., Allik, J., Sinkkonen, J., and Alho, K. (1997). "Language-specific phoneme representations revealed by electric and magnetic brain responses." *Nature*, 385(6615), 432-4.
- Nelson, D. A., and Marler, P. (1989). "Categorical perception of a natural stimulus continuum: birdsong." *Science*, 244(4907), 976-8.
- Ohl, F. W., and Scheich, H. (1996). "Differential frequency conditioning enhances spectral contrast sensitivity of units in auditory cortex (field AI) of the alert Mongolian gerbil." *Eur J Neurosci*, 8(5), 1001-17.
- O'Kusky, J., and Colonnier, M. (1982). "A laminar analysis of the number of neurons, glia, and synapses in the adult cortex (area 17) of adult macaque monkeys." *J Comp Neurol*, 210(3), 278-90.
- Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., and Hoke, M. (1998). "Increased auditory cortical representation in musicians." *Nature*, 392(6678), 811-4.
- Paradiso, M. A. (1988). "A theory for the use of visual orientation information which exploits the columnar structure of striate cortex." *Biol Cybern*, 58(1), 35-49.

- Pollack, L., and Norman, D. A. (1964). "A non-parametric analysis of recognition experiments." *Psychonomet Sci*, 1, 125–6.
- Polley, D. B., Heiser, M. A., Blake, D. T., Schreiner, C. E., and Merzenich, M. M. (2004). "Associative learning shapes the neural code for stimulus magnitude in primary auditory cortex." *Proc Natl Acad Sci U S A*, 101(46), 16351-6.
- Polley, D. B., Steinberg, E. E., and Merzenich, M. M. (2006). "Perceptual learning directs auditory cortical map reorganization through top-down influences." *J Neurosci*, 26(18), 4970-82.
- Powell, M. J. D. (1977). "A Fast Algorithm for Nonlinearly Constrained Optimization Calculations." Numerical Analysis, G. A. Watson, ed., Springer Verlag, New York.
- Recanzone, G. H., Merzenich, M. M., Jenkins, W. M., Grajski, K. A., and Dinse, H. R. (1992). "Topographic reorganization of the hand representation in cortical area 3b owl monkeys trained in a frequency-discrimination task." *J Neurophysiol*, 67(5), 1031-56.
- Recanzone, G. H., Schreiner, C. E., and Merzenich, M. M. (1993). "Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys." *J Neurosci*, 13(1), 87-103.
- Romo, R., Merchant, H., Zainos, A., and Hernandez, A. (1997). "Categorical perception of somesthetic stimuli: psychophysical measurements correlated with neuronal events in primate medial premotor cortex." *Cereb Cortex*, 7(4), 317-26.
- Schoups, A., Vogels, R., Qian, N., and Orban, G. (2001). "Practising orientation identification improves orientation coding in V1 neurons." *Nature*, 412(6846), 549-53.
- Sengpiel, F., Stawinski, P., and Bonhoeffer, T. (1999). "Influence of experience on orientation maps in cat visual cortex." *Nat Neurosci*, 2(8), 727-32.
- Seung, H. S., and Sompolinsky, H. (1993). "Simple models for reading neuronal population codes." *Proc Natl Acad Sci U S A*, 90(22), 10749-53.
- Steinschneider, M., Fishman, Y. I., and Arezzo, J. C. (2003). "Representation of the voice onset time (VOT) speech parameter in population responses within primary auditory cortex of the awake monkey." *J Acoust Soc Am*, 114(1), 307-21.
- Takeuchi, A. H., and Hulse, S. H. (1993). "Absolute pitch." *Psychol Bull*, 113(2), 345-61.
- Talwar, S. K., and Gerstein, G. L. (2001). "Reorganization in awake rat auditory cortex by local microstimulation and its effect on frequency-discrimination behavior." *J Neurophysiol*, 86(4), 1555-72.
- Wiesel, T. N. (1982). "Postnatal development of the visual cortex and the influence of environment." *Nature*, 299(5884), 583-91.
- Williams, L. (1977). "The perception of stop consonant voicing by Spanish-English bilinguals." *Percept Psychophys*, 21, 289-97.
- Wytenbach, R. A., May, M. L., and Hoy, R. R. (1996). "Categorical perception of sound frequency by crickets." *Science*, 273(5281), 1542-4.
- Zhang, K., Ginzburg, I., McNaughton, B. L., and Sejnowski, T. J. (1998). "Interpreting neuronal population activity by reconstruction: unified framework with application to hippocampal place cells." *J Neurophysiol*, 79(2), 1017-44.
- Zhang, L. I., Bao, S., and Merzenich, M. M. (2001). "Persistent and specific influences of early acoustic environments on primary auditory cortex." *Nat Neurosci*, 4(11), 1123-30.

## **CHAPTER 3: Selective Increase In Representations Of Sounds Repeated At An Ethological Rate**

Heesoo Kim & Shaowen Bao

Originally published in:

The Journal of Neuroscience, 29(16):5163-5169

April 22, 2009

### **Abstract**

Exposure to sounds during early development causes enlarged cortical representations of those sounds, leading to the commonly held view that the size of stimulus representations increases with stimulus exposure. However, representing stimuli based solely on their prevalence may be inefficient, because many frequent environmental sounds are behaviorally irrelevant. Here we show that cortical plasticity depends not only on exposure time, but also on the temporal modulation rate of the stimulus set. We examined cortical plasticity induced by early exposure to 7-kHz tone pips repeated at a slow (2 Hz), fast (15 Hz) or ethological (6 Hz) rate. Certain rat calls are modulated near 6 Hz. We found that spectral representation of 7-kHz increased only in the ethological-rate-reared animals, whereas improved entrainment of cortical neurons was seen in animals reared in the slow- and fast-rate condition. This temporal rate dependence of spectral plasticity may serve as a filtering mechanism to selectively enlarge representations of species-specific vocalizations. Further, our results indicate that spectral and temporal plasticity can be separately engaged depending on the statistical properties of the input stimuli.

## **Introduction**

Cortical sensory representations can be reorganized during early development and in adulthood (Diamond and Weinberger, 1986; Edeline et al., 1993; Recanzone et al., 1993; Bakin et al., 1996; Irvine and Rajan, 1996; Kilgard and Merzenich, 1998; Bao et al., 2001; Zhang et al., 2001; Beitel et al., 2003; Polley et al., 2004; Blake et al., 2006; Norena et al., 2006; Polley et al., 2006; de Villers-Sidani et al., 2008; Zhou et al., 2008). Such cortical plasticity processes are believed to enlarge representations of behaviorally important stimuli, thereby optimizing the processing capacity for these stimuli. Consistent with such a view, behaviorally important sounds, such as species-specific vocalizations, are preferentially represented in the auditory cortex of many species (Rauschecker et al., 1995; Wang et al., 1995; Ohlemiller et al., 1996; Tian et al., 2001; Wang and Kadia, 2001). While cortical plasticity in adult animals is induced by behaviorally important sensory stimuli associated with activity in the neuromodulatory systems, plasticity in developing animals can be induced by passive sensory exposure (Zhang et al., 2001; Chang and Merzenich, 2003; de Villers-Sidani et al., 2007; de Villers-Sidani et al., 2008; Zhou and Merzenich, 2008). A potential problem with exposure-induced plasticity is that both behaviorally important and irrelevant stimuli are present in the sensory environment. In some environments, behaviorally irrelevant stimuli may even dominate the sensory input. Thus, representation of stimuli based solely on frequency of occurrence or acoustic power could be highly inefficient.

Natural animal vocalizations are often repeated in bouts (Liu et al., 2003; Schnupp et al., 2006). The temporal repetition rate of vocalizations within these bouts is an important feature that may distinguish animal vocalizations from other environmental sounds. For instance, mouse vocalization calls are typically produced 5-10 times per second, whereas insect chirps may be repeated at much higher rates (Liu et al., 2003; Schnupp et al., 2006). A plausible mechanism that could allow for selective representations of species-specific calls is temporal filtering of the sensory input so that only sounds modulated near an ethologically relevant modulation rates induce experience-dependent plasticity.

In the present study, we investigated how cortical plasticity depends on temporal repetition rate. We characterized rat calls, and showed that they are typically repeated at 3-10 Hz. We then exposed rat pups to brief tones repeated at 2, 6 or 15 pips per second, and subsequently examined spectral frequency and temporal rate representations. Our results indicate that exposure to tones repeated at an ethologically relevant rate, but not a slower or faster rate, enlarged cortical representations of the exposure frequency. Although spectral representation was not changed for animals reared in the faster or slower rate, temporal rate representation was improved.

## **Methods**

### *Recording and analysis of animal vocalization.*

All procedures used in this study were approved by the University of California Berkeley Animal Care and Use Committee. To record rat pup isolation calls, individual rat pups were placed on a platform located in an anechoic chamber where the ambient temperature is maintained at 21.5 °C. A ¼-inch Bruel and Kjaer (B&K) model 4135 microphone was connected to a B&K 2669 preamplifier and B&K 2690 conditioning amplifier, and the output signal was digitized with a 16-bit A/D converter (National Instrument) at 200 kHz. Adult encounter calls were recorded after an adult female rat was introduced to the home cage of a single adult male. Five post-natal 11 (P11), six P15 rat pups and two adult pairs were used.

Visual examination of all recorded rat calls indicated that all pup calls were in the frequency range from 25 to 50 kHz and all adult encounter calls were in the frequency range from 25 to 70 kHz (Liu et al., 2003; Brudzynski et al., 1999; Brudzynski and Pniak, 2002). Thus, we band-pass-filtered all calls to obtain signals in the ranges of 25-50 kHz for pup calls or 25-75 kHz for adult calls, for further automatic identification of the calls based on their amplitude envelopes. The start of a call was defined as an upward crossing of a threshold of six standard deviations above the mean amplitude of a no-call period, and the end occurred when the amplitude envelope was below the threshold for at least 40 ms. Calls less than 5ms (or 10ms) long were automatically excluded for the pups (or adults). Call-onset asynchrony (COA) was defined as the time between the start of two consecutive calls.

### *Acoustic rearing of young rat pups.*

Four groups (Ethological, Slow, Fast and Mixed) of Sprague Dawley rat pups were placed with their mothers in anechoic sound-attenuation chambers from P8 to P30. This time range covers the critical period for spectral plasticity in AI and has been used previously (Zhang et al., 2001; de Villers-Sidani et al., 2007 Han et al., 2007). The Ethological, Fast and Slow rat pups experienced tone pips (7.071 kHz, 60 dB SPL, 25 ms) presented at one of three repetition rates 24 hours a day. The Ethological and Fast groups heard trains of six tone pips presented at the rate of 6 Hz and 15 Hz, respectively, with one train every 1.5 seconds (Fig. 3.1D). A pair of tone pips were played to the Slow group every 1.5 seconds, with 0.5 seconds onset asynchrony between the tone pips. To ensure that the Slow group receives the same amount of acoustic energy as the other groups, the duration of the tone pips was set at 75 ms (Fig. 3.1D). The Mix litter was exposed to trains of 15-kHz tone pips (60dB SPL, 25 ms) presented at the Ethological rate (6 Hz) and trains of 5-kHz tone pips (60dB SPL, 25ms) presented at the Fast rate (15 Hz). The respective trains were presented once every 3 seconds and interleaved so they never overlapped (Fig. 3.3A). After sound exposure, rats were moved to a regular animal room environment until they were mapped (typically 4-20 days after the end of rearing). A Control litter was reared in a regular animal room environment.

### *Electrophysiological recording procedure.*

The primary auditory cortex (AI) of sound-reared and control rats were mapped at comparable ages from P34 to P52. Care was taken to ensure that animals in

different groups were recorded at comparable ages (Control:  $47.9 \pm 18.4$  days; Ethological:  $39.8 \pm 6.3$ ; Fast:  $39.2 \pm 6$ ; Slow:  $37 \pm 2.2$ ). Rats were pre-anesthetized with buprenorphine (0.05 mg/kg, subcutaneous) a half hour before they were anesthetized with sodium pentobarbital (50 mg/kg, followed by 10-20 mg/kg supplements as needed). Atropine sulfate (0.1 mg/kg) and dexamethasone (1 mg/kg) were administered once every 6 hours. The head was secured in a custom head-holder that left the ears unobstructed, and the cisterna magna was drained of cerebrospinal fluid. The right auditory cortex was exposed through a craniotomy and duratomy, and was kept under a layer of silicone oil to prevent desiccation. Sound stimuli were delivered to the left ear through a custom-made speaker that had been calibrated to have less than 3% harmonic distortion and flat output in the entire frequency range.

Cortical responses were recorded with tungsten microelectrodes (FHC Inc.). Recording sites were chosen to evenly and densely sample the primary auditory cortex while avoiding surface blood vessels, and were marked on an amplified digital image of the cortex. Microelectrodes were lowered orthogonally into the cortex to a depth of 450-600 microns where responses to noise bursts could be found. Multiunit responses to 25 ms tone pips of 51 frequencies (1-32 kHz, 0.1-octave spacing, 5-ms cosine-squared ramps) and eight sound pressure levels (0-70 dB SPL, 10-dB steps) were recorded to reconstruct the frequency-intensity receptive field. In two control animals, 12 additional frequencies (32-74 kHz, 0.1-octave spacing) were included to quantify representations of high ultrasonic frequencies up to 74 kHz.

Responses to trains of tone pips and noise bursts were recorded in two additional rats per group using 4x4 silicon polytrodes, with approximately 1-m $\Omega$  impedance (NeuroNexus Technologies, N2T). After finding AI by coarse mapping with tungsten microelectrodes, a polytrode was lowered into cortex. Six noise bursts or pure tone pips were presented in trains at six different presentation rates (3, 6, 9, 12, 15, 18 Hz). The noise bursts and tone pips were 25-ms long (with 5-ms cosine-squared ramps) and presented at a sound pressure level of 50 dB. Each carrier-rate combination was repeated 10 times and presented in a pseudo-randomized order. One train was presented once every 3 seconds.

### *Data Analysis.*

The characteristic frequency (CF) was defined as the frequency at which responses are evoked at threshold—the lowest sound pressure level that activate the neuron. The bandwidth at 30 dB above threshold (BW30) measures the width of the receptive field (in octaves). The CF, threshold, and BW30 for each penetration site were determined visually. AI was functionally defined by well-tuned neurons and fine tonotopic gradient with increasing CFs going rostrorodorsally. Penetrations that were not in AI were removed leaving a total of 1590 AI recording sites (389 from Control, 368 from Ethological, 343 from Fast, 157 from Slow, 194 from Mix, and 139 from ultrasonic recordings). Cortical area representing a specific frequency was measured using voronoi tessellation (Matlab, Mathworks).

Repetition rate transfer functions (RRTFs), normalized responses as functions of presentation rates, were calculated as follows. First, only trials in which the response to the first noise burst (or pure tone) was greater than two standard deviations above mean spontaneous spike rate were included. The normalized response was calculated

by taking the average response of the last 5 sound presentations (response being the number of spikes triggered 7 to 40 ms after onset of the noise/tone) and dividing it by the response to the first sound. A normalized response greater than one indicates that the unit responded better to subsequent sounds than to single noise pulse/tone pip. All reported statistics are two-tailed t-tests unless indicated otherwise.

## **Results**

### **Repetition rate of rat vocalizations**

The repetition rate of rat vocalizations were measured by recording pup isolation calls and adult encounter calls. A total of 1610, 1210 and 1063 calls were extracted for adults, P15s and P11s, respectively. A bout was defined as a series of successive calls with call onset asynchronies (COAs) less than one second. Only calls that were produced in bouts were included for further analysis, resulting in 1410 (88%), 724 (60%) and 819 (77%) calls over 295, 197 and 209 bouts. An analysis of pup isolation and adult encounter vocalizations revealed that calls within bouts occur between 3-10 Hz (Fig. 3.1). The repetition rate of adult calls is faster than that of the pup calls (Fig. 3.1A, B). The COA distribution of the P15 group showed a very clear peak (240 ms, ~4.2 Hz), as the isolation calls from P15 pups were very stereotypical and repeated regularly. The repetitions were less regular for the P11 and the adult calls, resulting in the larger spreads. The COA distribution of the P11 group has a peak at 270 ms, corresponding to ~3.7 Hz while that of adults has a peak at 160 ms, corresponding to 6.3 Hz.

### **Effect of presentation rate on spectral plasticity**

To investigate the impact of repetition rates on cortical plasticity, we exposed three groups of rat pups to trains of 7-kHz tone pips, with tone trains presented once every 1.5 seconds. Within each train, tone pips were repeated at a rate of 2 Hz for the Slow group, 6 Hz for the Ethological group and 15 Hz for the Fast group. The Slow and Fast rates are below and above the range of the ethological repetition rates of rat vocalizations (Fig. 3.1B). The duration of the tone pips was increased for the Slow group so that the total acoustic energy of tone pips experienced by the animals was the same for all three groups (Fig. 3.1C).

Previous studies have shown that exposure to a tone increases cortical representations of that tone (Zhang et al., 2001; de Villers-Sidani et al., 2007; Han et al., 2007). In this study, we mapped the auditory cortex of several animals for each group (Control:  $n = 10$ ; Ethological:  $n = 6$ ; Fast:  $n = 6$ ; Slow:  $n = 4$ ), and found enlarged representations of the exposure frequency in animals reared with the Ethological rate, but not with the Slow or Fast rate (Fig. 3.2). A 4-conditions by 9-frequencies ANOVA showed no differences across condition ( $p = 0.95$ ) and a significant interaction ( $p < 0.04$ ). One-way ANOVAs across the frequency bins showed significant differences between conditions only in the 7-kHz ( $p < 0.037$ ) and 12-kHz bins ( $p < 0.041$ ). Animals reared in the Ethological rate (6 Hz) showed a significant increase in the cortical area representing 7-kHz ( $\pm 0.2$  octaves) when compared to the naïve Control, Slow and Fast groups ( $p < 0.013$ , 0.041, 0.001, respectively, one-tailed t-test). Animals reared with the Ethological rate showed smaller representations of 12-kHz tone when compared to the Control ( $p < 0.034$ ) and Fast ( $p < 0.018$ ) groups (Fig. 3.2C).

An additional group of rat pups were exposed to two different carrier frequencies presented at two different rates—trains of 15-kHz tone pips were presented at the Ethological rate (6 Hz) and trains of 5-kHz tone pips were presented at the Fast rate (15 Hz). The respective trains were presented once every 3 seconds and were interleaved so that one train was played every 1.5 seconds (Fig. 3.3A). A comparison with the naive control animals showed an increase in representation of 15-kHz ( $p < 0.05$ ) and a

decrease in the representation of the neighboring 20-kHz ( $p < 0.001$ ). No changes were observed around the representation of 5-kHz ( $p = 0.58$ ) (Fig. 3.3B). These results confirm our finding that sounds that are repeated at an ethological rate are over-represented.

The average threshold, response latency, BW30, and recording depth are shown in Table 3.1. One-way ANOVAs comparing across the five groups showed no significant differences for threshold ( $p > 0.2$ ), latency ( $p > 0.05$ ), BW30 ( $p > 0.2$ ) or recording depth ( $p > 0.5$ ).

### **Over-representation of rat vocalization frequencies**

The above results suggest that the representation of the frequency range of ultrasonic rat vocalizations should also be enlarged, because rat calls are mostly repeated at ethological rates. The ultrasonic (up to 74 kHz) region of the AI was mapped in two naïve virgin animals. Both animals showed large representation of ultrasonic frequencies (Fig. 3.4C). To facilitate statistical analysis, penetrations were separated into five one-octave-sized frequency bins (lowest bin covering 1.56-3.13 kHz and the highest bin covering 25-50 kHz, frequencies higher than 50 kHz were not included leaving 106 penetrations). A Chi-Square test showed that the representation of frequencies from 25 to 50 kHz was significantly greater than those of the other frequency ranges ( $p < 0.001$ , see Fig. 3.4Ciii).

### **Temporal response plasticity**

Plasticity in temporal response properties was tested in cortical neurons by measuring their responses to trains of tone pips presented at various rates. Only units that responded reliably to 7-kHz pure tones (see Methods) were included for the analysis (Control:  $n = 26$ , Ethological:  $n = 28$ , Fast:  $n = 37$ , Slow:  $n = 42$ ). The 7-kHz-tone-derived repetition rate transfer function (RRTF) of the Control and Ethological groups overlapped (Fig. 3.4A). A group x rate analysis of variance (ANOVA) showed significant main effects for group ( $p < 0.001$ ) and rate ( $p < 0.0001$ ) and a significant interaction ( $p < 0.00001$ ). Post-hoc t-tests revealed that the normalized response at the 12 Hz repetition rate was significantly larger for the Fast group when compared to all the other groups ( $p < 0.05$ ). In addition, the Slow group had greater normalized responses at the repetition rate of 6 Hz when compared to all the other groups ( $p < 0.001$ ). Although neurons in the Fast group showed slightly better following responses than the Control groups at the repetition rate of 15 Hz, the rearing rate, the difference was not significant ( $p > 0.2$ ). The Slow group also showed slightly enhanced following responses at 3 Hz, but it was not different from that of the Ethological group ( $p > 0.05$ ).

Of the 133 units analyzed above, 116 were responsive to noise bursts (Control:  $n = 24$ , Slow:  $n = 25$ , Fast:  $n = 31$ , Slow:  $n = 36$ ). A one-way ANOVA revealed significant differences between groups at the repetition rates of 9 Hz ( $p = 0.0002$ ) and 18 Hz ( $p < 0.01$ , Fig. 3.4B). Post-hoc t-tests showed that the Fast group did not entrain to noise burst as well as the Control or Slow groups at the repetition rate of 9 Hz ( $p < 0.002$  for both), and that the Slow group did not entrain at 18 Hz as well as the other groups ( $p < 0.02$ , for all). However, we did not observe enhanced responses to noise bursts repeated at 12 Hz in the Fast group, or at 6 Hz in the Slow group (Fig. 3.4B).

We also examined cortical responses to repeated tones of various carrier frequencies (4.5, 5.6, 8.9, 11.2, 14.1, or 17.8 kHz) in the Control, Ethological, Fast and

Slow groups ( $n = 21, 23, 32$  and  $40$ , respectively). The carrier frequencies were chosen to reliably activate the units and only units used in the 7-kHz analysis were included. One-way ANOVAs across the different repetition rates showed significant differences only for the 15 and 18 Hz ( $p < 0.01$ , for both, Fig 4C). At the repetition rate of 15 Hz, the Fast group did not entrain to tone pips as well as the Control or Ethological group ( $p < 0.02$ , for both); while in at the repetition rate of 18 Hz, the Slow group showed lower normalized responses compared to all other groups ( $p < 0.04$ , for all). Thus the enhanced responses to sounds repeated at 12 Hz in the Fast group and enhanced response to sounds repeated at 6 Hz in the Slow group were specific to the 7-kHz carrier frequency of the exposure tone.

## **Discussion**

In the present study, we tested the hypothesis that temporal repetition rates influence how sound experiences shape cortical representations of the sound. All animals in the Fast, Slow, Ethological groups of animals were exposed to a 7-kHz tone of the same total acoustic energy, and yet the sensory experiences had completely different effects—a 40% increase in 7-kHz representations for the Ethological group, but not for the Fast or the Slow group. Furthermore, exposing developing animals to two tones (5 and 15 kHz) presented at two different rates (Fast and Ethological, respectively) lead to the overrepresentation of only the tone presented at the Ethological rate. These results indicate that temporal repetition rates of sensory stimuli have a strong impact on experience-dependent plasticity. Earlier studies of sound exposure-induced cortical plasticity mostly used repetition rates similar to our Ethological rate, and robust increases in representations of the exposed stimulus were observed (Zhang et al., 2001; Han et al., 2007). In contrast, reduced representations were observed for stimuli that are constantly present in the environment without temporal modulation (de Villers-Sidani et al., 2008; Zhou et al., 2008). Here we show that exposure to tone pips that are repeated at 2 or 15 Hz does not result in greater representations of the tone. Such a temporal filtering mechanism would enlarge representations of stimuli that are repeated/modulated near the ethological rates of species-specific vocalization, and would suppress representations of other potentially irrelevant stimuli that are modulated at other rates.

Repetition rates of the rat calls have considerable variability. However, the majority of calls were repeated at rates from 3 to 10 Hz (Fig. 3.1). Less than 1% of the calls had repetition rate higher than 15 Hz, whereas 26% had repetition rate  $\leq$  2 Hz. Thus, a large number of rat calls could pass through the presumptive temporal filter and shape cortical acoustic representations. The COAs that were  $\leq$  0.5 s ( $\leq$  2 Hz) were mostly between two calls in different bouts. The enlarged representation of the frequency range of the rat vocalizations is consistent with the notion of a temporal filter for selective representation of sounds repeated at ethological rates. However, it could also be attributable to other experience-independent mechanisms.

The neural mechanisms of the temporal filter in cortical plasticity are unknown. It is well known that cortical neurons respond differently to sound repeated at different rates. For instance, cortical neurons in anesthetized rats do not respond well to sounds repeated  $>10$  times per second (Kilgard & Merzenich, 1998b), whereas auditory thalamic neurons are capable of responding to much faster rates (Wehr & Zador, 2005). In the awake preparation, multi-unit clusters have been shown on average to synchronize to click rate of 72 Hz, but normalized responses show a decrease around 10 Hz (Anderson et al., 2006). Such temporal response properties may contribute to the lack of spectral plasticity in the Fast group. However, such cortical temporal response properties cannot account for the lack of spectral plasticity in the Slow group, because cortical neurons respond well to slow-rate sounds.

Although exposure to Fast-rate tone pips did not enlarge representations of the tone, it did improve entrainment of responses to quickly repeating tone pips in neurons of the Fast group. Similarly, exposure to Slow-rate tone pips enhanced cortical responses to slowly repeating tone pips. In previous studies, temporal plasticity was induced either with noise bursts (Kilgard et al., 2001; Chang and Merzenich, 2003; Bao

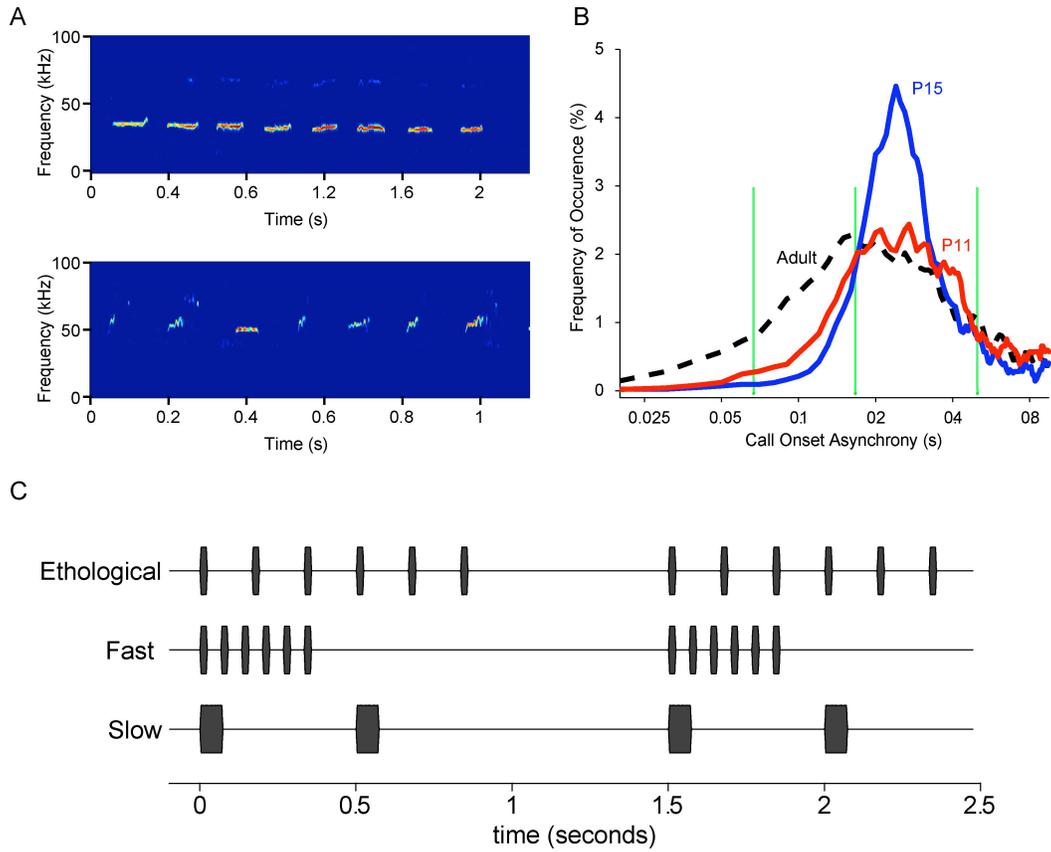
et al., 2004; de Villers-Sidani et al., 2008; Zhou and Merzenich, 2008) or tone pips of several carrier frequencies (Kilgard et al., 2001). Tone pips of a single frequency were previously found ineffective in inducing temporal plasticity (Kilgard et al., 2001). We show here that temporal plasticity can be induced with a single-frequency tone, and can be specific to the tonal stimulus. The discrepancy between the earlier and the present results may be related to differences in the experimental methods. The earlier study used stimulation of Nucleus Basalis to induce plasticity in adult animals, whereas we simply exposed young animals to the sound stimuli. Further, we restricted our analysis only to neurons that responded reliably to the exposure tone, which might be necessary to reveal frequency-specific plasticity effects. These results are consistent with earlier findings of enhanced entrainment of cortical responses in mother rats to trains that are spectrally and temporally similar to pup calls (Liu et al., 2006). Our results suggest that both spectral and temporal information of specific stimuli can be represented in the same population of neurons. They also indicate that spectral and temporal plasticity can be separately engaged depending on the characteristics of input stimuli, such as the temporal modulation rate.

It has long been hypothesized that efficient representations of sensory stimuli depends on the stimulus statistics (Barlow, 1961; Lewicki, 2002; Singh and Theunissen, 2003), and that the learning and plasticity processes that shape sensory representations must be sensitive to the statistics of the sensory input (Kilgard et al., 2001; Maye et al., 2002; Toro and Trolabon, 2005). Our results support this hypothesis by showing that cortical circuits are sensitive to temporal rates, and may use this feature to selectively represent sounds that are likely to be behaviorally relevant.

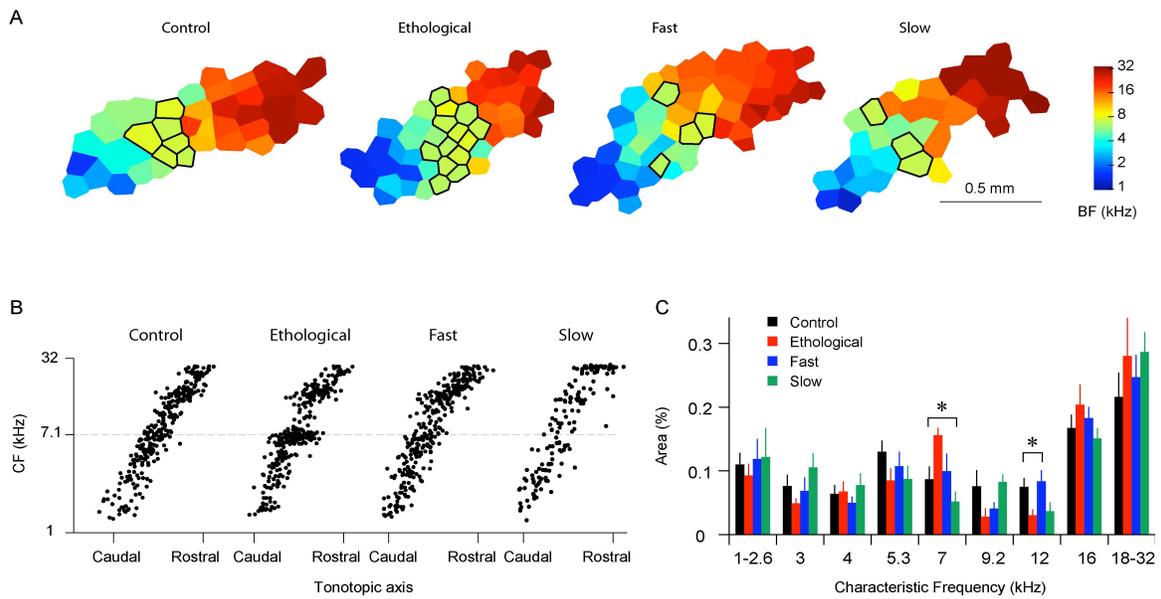
**Table 3.1.** Response properties of AI neurons with SEM in parentheses.

	Threshold (dB)	Latency (ms)	BW30 (oct)	Depth (microm)
Control	37.34 (2.63)	17.93 (0.62)	1.40 (0.073)	550.8 (10.9)
Ethological	32.42 (3.65)	17.44 (0.46)	1.31 (0.064)	563.8 (13.1)
Fast	35.29 (1.33)	17.30 (0.41)	1.33 (0.093)	555.9 (15.4)
Slow	26.2 (2.54)	20.28 (1.07)	1.13 (0.040)	540.0 (5.8)
Mix	32.88 (2.23)	18.34 (0.37)	1.22 (0.017)	536.1 (3.2)

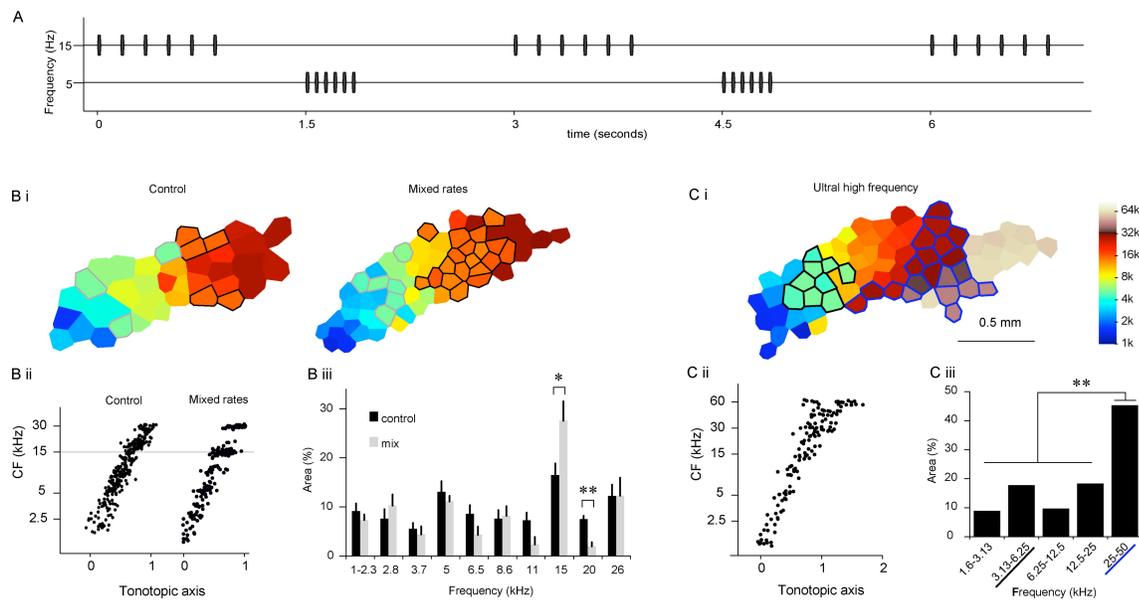
## Figures



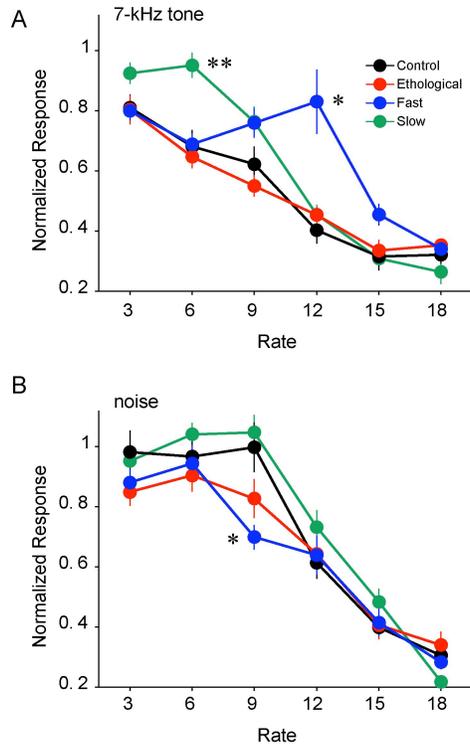
**Figure 3.1.** Characterization of rat calls. **A**, Example spectrograms of a bout of rat pup isolation calls (above) and adult encounter calls (below). **B**, Distributions of call onset asynchronies (COA) within a bout. Green vertical lines indicate the COAs of the three experiments condition: Fast (15 Hz, COA = 0.067 s, far left), Ethological (6 Hz, COA = 0.167 s, middle), and Slow (2 Hz, CAO = 0.5 s, far right). **C**, A schematic of the stimuli used in the Ethological, Fast, and Slow experimental rearing conditions.



**Figure 3.2.** Influences of stimulus presentation rate on spectral plasticity. **A**, Representative cortical CF maps of the Control, Ethological, Fast, and Slow groups. The areas representing  $7 \text{ kHz} \pm 0.2$  octaves were outlined in black. **B**, Distributions of CFs along the tonotopic axis. **C**, Sizes of cortical areas representing different frequency bands. Significant differences were seen for the 7-kHz and 12-kHz bands. Error bars indicate standard error of the mean. \* indicates  $p < 0.05$ .



**Figure 3.3.** Over-representation of sounds repeated at the ethological rates. **A**, A schematic of “mix-rate” rearing stimuli. A train of 15-kHz tone consisted of six tone pips presented at the Ethological rate (6 Hz), and a train of 5-kHz tone consisted of six tone pips presented at the Fast rate (15 Hz). Trains of the two repetition rates were interleaved such that one train was heard every 1.5 seconds. **B**, CF map reorganization resulted from the mixed-rate rearing. **Bi**, Example maps of control and mixed-rate animals. Control animal is the same as seen in Figure 2A. Area represented 5 kHz  $\pm$  0.2 octaves are outlined in gray while area representing 15 kHz  $\pm$  0.2 octaves are outlined in black. **Bii**, Distributions of CFs along the tonotopic axis. **Biii**, Sizes of cortical areas representing different frequency bands. There was a significant increase in representation at 15 kHz and a significant decrease at 20 kHz. Error bars indicate standard error of the mean. \* indicates  $p < 0.05$ , \*\* indicates  $p < 0.001$ . **C**, Cortical representation of ultrasonic frequencies. **Ci**, An example CF map from a control animal mapped up to 74 kHz. Areas representing 25-50 kHz are outlined in blue while areas representing 3.13-6.25 kHz are outlined in black. **Cii**, Distribution of CFs along the tonotopic axis. **Ciii**, Sizes of cortical areas representing one-octave frequency bands. The representation of the 25-50 kHz band was significantly larger than those of the other bands.



**Figure 3.4.** Effects of tonal exposure on cortical temporal response properties. **A-B**, Repetition rate transfer functions characterized with trains of 7-kHz tone pips (**A**) and noise bursts (**B**). Error bars indicate standard error of the mean. \* indicates  $p < 0.05$  and \*\* indicates  $p < 0.001$ , all for comparison to the Control.

## **References**

- Bakin JS, South DA, Weinberger NM (1996) Induction of receptive field plasticity in the auditory cortex of the guinea pig during instrumental avoidance conditioning. *Behav Neurosci* 110:905-913.
- Bao S, Chan VT, Merzenich MM (2001) Cortical remodelling induced by activity of ventral tegmental dopamine neurons. *Nature* 412:79-83.
- Bao S, Chang EF, Woods J, Merzenich MM (2004) Temporal plasticity in the primary auditory cortex induced by operant perceptual learning. *Nat Neurosci* 7:974-981.
- Barlow HB (1961) Possible principle underlying the transformation of sensory messages. In: *Sensory Communication* (W. R, ed), pp 217-234. Cambridge: MIT press.
- Beitel RE, Schreiner CE, Cheung SW, Wang X, Merzenich MM (2003) Reward-dependent plasticity in the primary auditory cortex of adult monkeys trained to discriminate temporally modulated signals. *Proc Natl Acad Sci U S A* 100:11070-11075.
- Blake DT, Heiser MA, Caywood M, Merzenich MM (2006) Experience-dependent adult cortical plasticity requires cognitive association between sensation and reward. *Neuron* 52:371-381.
- Brudzynski SM, Kehoe P, Callahan M (1999) Sonographic structure of isolation-induced ultrasonic calls of rat pups. *Dev Psychobiol* 34:195-204
- Brudzynski SM, Pniak A (2002) Social contacts and production of 50-kHz short ultrasonic calls in adult rats. *J Comp Psychol* 116:73-82
- Chang EF, Merzenich MM (2003) Environmental noise retards auditory cortical development. *Science* 300:498-502.
- de Villers-Sidani E, Chang EF, Bao S, Merzenich MM (2007) Critical period window for spectral tuning defined in the primary auditory cortex (A1) in the rat. *J Neurosci* 27:180-189.
- de Villers-Sidani E, Simpson KL, Lu YF, Lin RC, Merzenich MM (2008) Manipulating critical period closure across different sectors of the primary auditory cortex. *Nat Neurosci* 11:957-965.
- Diamond DM, Weinberger NM (1986) Classical conditioning rapidly induces specific changes in frequency receptive fields of single neurons in secondary and ventral ectosylvian auditory cortical fields. *Brain Res* 372:357-360.
- Edeline JM, Pham P, Weinberger NM (1993) Rapid development of learning-induced receptive field plasticity in the auditory cortex. *Behav Neurosci* 107:539-551.
- Han YK, Kover H, Insanally MN, Semerdjian JH, Bao S (2007) Early experience impairs perceptual discrimination. *Nat Neurosci* 10:1191-1197.
- Irvine DR, Rajan R (1996) Injury- and use-related plasticity in the primary sensory cortex of adult mammals: possible relationship to perceptual learning. *Clin Exp Pharmacol Physiol* 23:939-947.
- Kilgard MP, Merzenich MM (1998) Cortical map reorganization enabled by nucleus basalis activity. *Science* 279:1714-1718.
- Kilgard MP, Pandya PK, Vazquez J, Gehi A, Schreiner CE, Merzenich MM (2001) Sensory input directs spatial and temporal plasticity in primary auditory cortex. *J Neurophysiol* 86:326-338.
- Lewicki MS (2002) Efficient coding of natural sounds. *Nat Neurosci* 5:356-363.

- Liu RC, Linden JF, Schreiner CE (2006) Improved cortical entrainment to infant communication calls in mothers compared with virgin mice. *Eur J Neurosci* 23:3087-3097.
- Liu RC, Miller KD, Merzenich MM, Schreiner CE (2003) Acoustic variability and distinguishability among mouse ultrasound vocalizations. *J Acoust Soc Am* 114:3412-3422.
- Maye J, Werker JF, Gerken L (2002) Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82:B101-111.
- Norena AJ, Gourevitch B, Aizawa N, Eggermont JJ (2006) Spectrally enhanced acoustic environment disrupts frequency representation in cat auditory cortex. *Nat Neurosci* 9:932-939.
- Ohlemiller KK, Kanwal JS, Suga N (1996) Facilitative responses to species-specific calls in cortical FM-FM neurons of the mustached bat. *Neuroreport* 7:1749-1755.
- Polley DB, Steinberg EE, Merzenich MM (2006) Perceptual learning directs auditory cortical map reorganization through top-down influences. *J Neurosci* 26:4970-4982.
- Polley DB, Heiser MA, Blake DT, Schreiner CE, Merzenich MM (2004) Associative learning shapes the neural code for stimulus magnitude in primary auditory cortex. *Proc Natl Acad Sci U S A* 101:16351-16356.
- Rauschecker JP, Tian B, Hauser M (1995) Processing of complex sounds in the macaque nonprimary auditory cortex. *Science* 268:111-114.
- Recanzone GH, Schreiner CE, Merzenich MM (1993) Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *J Neurosci* 13:87-103.
- Schnupp JW, Hall TM, Kokelaar RF, Ahmed B (2006) Plasticity of temporal pattern codes for vocalization stimuli in primary auditory cortex. *J Neurosci* 26:4785-4795.
- Singh NC, Theunissen FE (2003) Modulation spectra of natural sounds and ethological theories of auditory processing. *J Acoust Soc Am* 114:3394-3411.
- Tian B, Reser D, Durham A, Kustov A, Rauschecker JP (2001) Functional specialization in rhesus monkey auditory cortex. *Science* 292:290-293.
- Toro JM, Trobalon JB (2005) Statistical computations over a speech stream in a rodent. *Percept Psychophys* 67:867-875.
- Wang X, Kadia SC (2001) Differential representation of species-specific primate vocalizations in the auditory cortices of marmoset and cat. *J Neurophysiol* 86:2616-2620.
- Wang X, Merzenich MM, Beitel R, Schreiner CE (1995) Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. *J Neurophysiol* 74:2685-2706.
- Zhang LI, Bao S, Merzenich MM (2001) Persistent and specific influences of early acoustic environments on primary auditory cortex. *Nat Neurosci* 4:1123-1130.
- Zhou X, Merzenich MM (2008) Enduring effects of early structured noise exposure on temporal modulation in the primary auditory cortex. *Proc Natl Acad Sci U S A* 105:4423-4428.

Zhou X, Nagarajan N, Mossop BJ, Merzenich MM (2008) Influences of un-modulated acoustic inputs on functional maturation and critical-period plasticity of the primary auditory cortex. *Neuroscience* 154:390-396.

## **CHAPTER 4: Over-representation of ultrasonic vocalization frequencies in the rat primary auditory cortex**

### **Abstract**

Species-specific vocalizations play a critical role in communication and guiding appropriate behaviors. It is unclear whether the representation and perception of conspecific vocalization sounds are innate or experience dependent. We find that ethologically relevant rat communication calls primarily occur in the ultrasonic frequency range (above 30kHz). Virtually all previous studies of the rat primary auditory cortex (A1) have studied the representation of frequencies strictly below 32kHz. We find nearly 40% of A1 is devoted to representing vocalization frequencies (32kHz – 64kHz), compared to less than 20% for any other octave-band below 32kHz. This increase in representation of ultrasonic frequencies is accompanied by enhanced frequency discrimination around 32kHz, suggesting this form of representation is an efficient method of representing ethologically relevant frequencies. We show that this preferential representation of ultrasonic frequencies depends on early auditory experience. Animals reversibly deafened at an early age show a distribution of frequency representation similar to immature developing animals. Representation of ultrasonic frequencies is also enhanced in the inferior colliculus (IC) of mature rats. We find that preferential and efficient representation of ethologically relevant ultrasonic frequencies is dependent on early experience with rodent vocalizations.

## **Introduction**

Species-specific vocalizations play a critical role in guiding appropriate social behaviors and promoting survival. In rodents, socially relevant ultrasonic vocalizations (USVs) exclusively occur in a frequency range well beyond what humans are capable of hearing. In rats, three distinct categories of USVs have been reported: 40 kHz pup isolation calls, 50 kHz positive affect calls and 22 kHz alarm calls (Brudzynski et al., 1993, 1999; Knutson et al., 2002; Portfors, 2007; Takahashi et al., 2010). Pup isolation calls are thought to be elicited when the body temperature of the pups drop, as would occur when a pup is way from its nest. These pup isolation calls will prompt a mother to leave her nest and retrieve the vocalizing pup (Hahn and Lavooy, 2005). On the other hand, USVs around 50 kHz are often associated with positive affective state (Knutson et al., 2002; Brudzynski, 2005). Exposure to adult encounter calls emitted by males will prime female rodents to be more receptive to sexual encounters (McIntosh and Barfield, 1978). We propose that these ethologically relevant vocalizations should be preferentially represented in the rat primary auditory cortex (A1).

The representation of the frequencies encompassing these ethologically relevant calls has not been studied. The vast majority of publications on the rat A1 have focused exclusively on pure-tone pips below 32kHz (Zhang et al., 2001; Chang and Merzenich, 2003; Chang et al., 2005; Han et al., 2007; Polley et al., 2007; Engineer et al., 2008; de Villers-Sidani et al., 2008; Zhou and Merzenich, 2008; Insanally et al., 2009). We verified that rat USVs, specifically pup isolation calls and adult encounter calls, occur in the ultrasonic range above 30kHz, and mapped the primary auditory cortex utilizing pure tone pips from 1 to 74kHz, increasing the stimulus set by 1.2 octaves, or approximately 20%. We find a clear over-representation of vocalization frequencies in the rat A1. Nearly 40% of A1 is devoted to representing vocalization frequencies (32kHz – 64kHz), compared to less than 20% for any other octave-band below 32kHz.

The distribution of representation and the shape of tuning curves have direct implications on perception (Schoups et al., 2001; Butts and Goldman, 2006; Kim and Bao, 2008; Fischer and Peña, 2011; Ganguli and Simoncelli, 2011; Girshick et al., 2011). In addition to more neurons being tuned to ultrasonic frequencies, we find a systematic decrease in tuning bandwidth with higher frequency tuning, theorized as an optimal organization for representing a sensory environment of vocalizations (Ganguli and Simoncelli, 2011). This narrowing of tuning curves would also suggest enhanced discrimination ability in the ultrasonic frequency range (Schoups et al., 2001; Girshick et al., 2011). We indeed found that rats were better able to discriminate frequencies in the ultrasonic range. These results suggest that the representation of ultrasonic frequency is efficient and optimized to represent vocalizations (Ganguli and Simoncelli, 2011).

Auditory cortex neurons develop in a progressive manner with certain response properties maturing before others. Response properties such as response latency, bandwidth, threshold, temporal rate following and FM sweep response magnitude have all been shown to mature at different points of development (Chang et al., 2005; de Villers-Sidani et al., 2007; Insanally et al., 2009). Utilizing frequencies from 1-32kHz, it has previously been reported that A1 reaches mature size and spectral representation on either P14 or P16, with representation of low and high frequencies being absent in younger animals (de Villers-Sidani et al., 2007; Insanally et al., 2009). Studies in

subcortical auditory responses have also shown a similar time-course in spectral development (Romand and Gunter, 1990). We investigated the representation of ultrasonic frequencies in developing animals and found mature, adult-like overrepresentation appears only after P21.

It is unknown whether this over-representation of USV frequencies with narrow tuning bandwidth is innate or experience dependent. Previous reports have documented the importance of experience on auditory cortex responses to vocalizations (Cheung et al., 2005; Liu et al., 2006; Cohen et al., 2011). In addition, we have demonstrated that A1 may utilize second order statistic of their acoustic environment to guide plasticity. Specifically, pure tone pips that are presented at an ethologically relevant temporal repetition rate will lead to increased spectral representation of that tone (Kim and Bao, 2009). Tones presented at faster or slower rates do not show an expansion in representation. Since vocalizations, by definition, are repeated at an ethological rate, we previously hypothesized that vocalization frequencies be over-represented in A1. However, if this over-representation is due to experience or is an innate feature of A1 is still an open question. We reversibly deafened developing animals via ear canal ligation (Popescu and Polley, 2010) at post-natal day 10 (P10), well before the auditory critical period (de Villers-Sidani et al., 2007) and the onset of hearing (Geal-Dor et al., 1993). These animals show less representation of very high ultrasonic frequencies and look very similar to young developing animals. This suggests that representation of ultrasonic frequency is experience dependent.

We find that this preferential representation of ultrasonic frequencies is not specific to cortex. The inferior colliculus also showed an over-representation of these vocalization frequencies. Further experiments are necessary to determine whether the representation in cortex is merely a feed-forward readout of subcortical auditory regions or feedback circuits from the cortex are utilized to shape the representation of the IC. Female rats are the intended recipients of the calls we studied: pup isolation calls to mothers (Ehret et al., 1987) and encounter calls to potential mates (McIntosh and Barfield, 1978). However, we find little consistent difference in representation between male and female rats, suggesting that early experience plays a role in shaping the representation of ultrasonic frequencies.

We find an experience-dependent over-representation of species-specific ultrasonic vocalization frequencies in the rat. This specialized ultrasonic region has significantly narrow bandwidths and a corresponding enhanced behavioral discrimination ability. We hypothesize that the auditory and vocalization system of the rat co-evolved to utilize ultrasonic frequencies.

## **Methods**

### *Recording and Analysis of Animal Vocalizations.*

All procedures were approved by the University of California Berkeley Animal Care and Use Committee. Animal vocalizations were recorded as previously described (Kim and Bao, 2009). Briefly, Sprague-Dawley rats were placed in an anechoic chamber with a ¼ inch Bruel and Kjaer (B&K) model 4135 microphone connected to a B&K 2669 preamplifier and B&K 2690 conditioning amplifier. The signal was digitized with a 16-bit analog-to-digital converter (National Instruments) at 200 kHz. Over 2 hours of vocalizations were recorded with five postnatal day 11 (P11), six P15 rat pups and two pairs of adults.

Examples of call analysis can be seen in Figure 4.2. The experimenter visually identified 1098, 2113 and 1798 (P11, P15 and adult, respectively) calls, indicating for each call the approximate start time and center frequency. To increase the signal-to-noise ratio, a bandpass filter ( $\pm 10$  kHz) was applied around the experimenter-defined frequency for each call. The envelope of the filtered call (Hilbert transformation) was used to identify the start and end of each call. To be defined as a call, the signal had to cross the threshold set at 6X the standard deviation of the residual noise (the signal 25-50 ms preceding the experimenter-defined start time). The end of a call was marked when the signal dropped below threshold for at least 40 ms. The frequency profile or distribution of an individual call was defined by collapsing the spectrogram of the filtered call across the duration of the call. The center frequency of the call was defined as the peak of the frequency profile of the call. The bandwidth of the call was the full width at half maximum of the frequency distribution.

### *Behavioral Testing*

Frequency discrimination ability was tested as previously described (Han et al., 2007). Adult female animals were food deprived and maintained at 90 to 95% of initial body weight. All training and testing took place in a wire cage within a sound attenuated box. Trials were automatically initiated with 100ms pure tone pips of the *standard* frequency presented at five pips per second for 5 to 35 seconds, after which the *standard* frequency alternated with a *target* frequency. Rats indicated the presence of the *target* frequency with a nose poke within 3 seconds to receive a food pellet award. Nose pokes outside this three-second window were discouraged through time-outs. Rats were trained on the task for up to four days where the difference between the *standard* and *target* frequencies ( $\Delta f$ ) was one octave and the response window was greater than 3 seconds. Testing was separated into two frequency ranges: normal (4-32 kHz, 9 bins) and ultrasonic (20-67 kHz, 9 bins). Animals were tested for up to three days both ranges, with the order counterbalanced between animals. On testing days,  $\Delta f$  was set to 30% of 0.5, 0.3 and 0.1 octaves and 10% of 0 octaves (to measure false alarms).

### *Electrophysiological Recording Procedure*

The primary auditory cortex (A1) was mapped as previously described (Kim and Bao, 2009). Inferior colliculus (IC) recordings were adapted from (Popescu and Polley, 2010). Rats were pre-anesthetized with buprenorphine (0.05 mg/kg, s.c.) one half hour

before they were anesthetized with sodium pentobarbital (50 mg/kg, i.p.). Atropine sulfate (0.1 mg/kg, s.c.) and dexamethasone (1 mg/kg, s.c.) were administered once every 6 hours and lactated Ringer's solution (0.5 – 1.0 mL, s.c.) was administered once every 4 hours. The head was secured in a custom head-holder that left the ears unobstructed, and the cisterna magna was drained of CSF. For cortical recordings, the right auditory cortex was exposed and kept under a layer of silicone oil to prevent desiccation. For IC recordings, a craniotomy was made above the right primary visual cortex at the interaural line extending at least 2mm lateral to the midline. Multiunit responses to 25 ms pure tone pips (1-74 or 2-74 kHz, 0.1 octave increments at 0-70 dB, 10 dB increments) were recorded using tungsten microelectrodes (FHC) or silicon polytrodes (NeuroNexus). References to neurons or units in this text refer to multi-unit responses recorded extracellularly. Cortical responses were recorded in the thalamorecipient layer of cortex (450-600 microns). The central nucleus of the IC was identified when electrodes/polytrodes had passed through visual cortex and reliable, tonotopically-organized responses to tones were seen (starting approximately 3mm below surface of V1). Tones were presented at three pips per second into the left ear and each frequency-dB combination was repeated 3 times. Cortical penetration locations were recorded on a high resolution image and care was taken to avoid blood vessels. IC depth was recorded off of readings from a hydraulic pump and depth along the silicon polytrode.

#### *Manipulation of Developmental Environment*

Procedure for reversible deafening by ear ligation were adapted from (Popescu and Polley, 2010). Small incisions were made below each pinna and the external meatus was ligated using polyester sutures in P10 rat pups under isoflourance anesthesia. Following recovery from surgery, rats were returned to their home cages for a minimum of two weeks. The quality of the ligation was verified around P21 with either Auditory Brainstem Responses (ABRs) and/or visual inspection to ensure adequate closure and bilateral deafening of each animal. Prior to cortical recordings from the deafened animals, the left outer ear was removed and any buildup was carefully removed until the tympanic membrane could be visualized. In most animals, ABRs were taken before and after the ear removal to confirm the recovery of hearing.

#### *Auditory Brainstem Responses (ABRs)*

ABRs were recorded to verify successful deafening and hearing recovery in a subset of animals. A 0.005" diameter, half hard, stainless steel bare wires were inserted behind the pinna of both ears and the vertex of the skull. Pure tone pips (3ms duration) were presented at 19 pips per second with an average of 500 repetitions for every attenuation-dB combination. Data acquisition and sound presentation were done using BioSigRP software on a Tucker Davis Technology Sys3 recording rig.

#### *Data Analysis*

The receptive fields and response properties were isolated utilizing custom-made programs in Matlab. For each unit, isolation of the receptive field required calculation of the response latency. First the peri-stimulus time histogram (PSTH) for all sound stimuli was convolved with a uniform 5 ms window. The peak of the PSTH within 7 and 30 ms

after onset of stimuli was the most reliable measure and was defined as the response latency as seen in later figures. The baseline firing rate was taken as the mean firing rate in the 47 ms preceding the stimuli. The start of the response window was defined as the point of time at which the PSTH exceeded the baseline firing rate at least 7ms after the onset of stimuli. The end of the response window was defined as the point of time at which the PSTH was less than the baseline firing rate at least 10 ms after the peak of the PSTH. The spikes that occurred between the start and end of the response window were counted to reconstruct an appropriate receptive field (as seen in Figure 4.4). This penetration-specific re-windowing was important since there is a frequency-dependent shift in PSTH timing (see Figure 4.5F) and also an age-dependent shift in response latency (Figure 4.9D).

Receptive fields were isolated utilizing a thresholding and filtering algorithm (Insanally et al., 2009). These isolated receptive fields (Figure 4.4, middle panels) were the basis for the extraction of the following response properties. The characteristic frequency (CF) of a neuron was defined as the center of mass of the isolated receptive field. The threshold of the neuron was the lowest dB level that elicited responses in the isolated receptive field. The maximum response magnitude was the maximal number of spikes seen for a single frequency-dB combination. Since each frequency-dB combination was repeated 3 times, the average of those 3 responses was taken. The tuning curve of each neuron was calculated by collapsing the responses to the top two dB levels of the isolated receptive field (Figure 4.4, bottom panel). To find the bandwidth (BW) of the tuning curves, first the tuning curves were smoothed with a uniform 0.4 octave window and then the full width at half maximum (FWHM) was calculated. The same tuning curves were utilized to calculate the Fisher information (FI). Initially, the tuning curves were fitted to a Gaussian distribution. To calculate the FI, the square of the first derivative of the fitted tuning curve was divided by the tuning curve (Seung and Sompolinsky, 1993).

All error bars indicate standard errors of the mean (SEM). Appropriate statistical tests were applied as necessary (e.g. t-tests, ANOVA, etc).

## **Results**

### **Rat Vocalizations are Predominately Ultrasonic.**

To understand the representation of vocalization frequencies, we initially quantified the distribution of frequencies utilized in rat species-specific communications. Postnatal day 11 (P11) and P15 pup isolation and adult encounter calls were recorded in a sound attenuated chamber (Figure 4.1A). Individual calls were identified by hand, while bandpass filters and amplitude envelopes were utilized to extract call statistics (Figure 4.2, see methods). The distribution of the calls' center frequency and frequency profile show all calls were above 20kHz, with the pup isolation calls and adult encounter calls encompassing distinct portions of the frequency space (Figure 4.1B, Table 4.1). Pup isolation call distribution show a peak between 35 and 40kHz while adult encounter calls peak around 60kHz.

The bandwidth of the P11, P15 pup isolation calls and adult encounter calls were not significantly different (Figure 4.1C, Table 4.1). The duration of individual calls were significantly different between the adult group and the two pup groups (Table 4.1). Interestingly, the duration of the individual calls were significantly inversely correlated with the center frequency of a call: calls of higher frequencies had significantly shorter duration ( $r = -0.65$ ,  $p < 0.001$  – collapsing all groups, Figure 4.1D, Table 4.1). The difference between the distribution of the center frequency and the frequency profile for the P15 calls (Figure 4.1B) can be explained by this effect: although 32% of the P15 calls had a center frequency higher than 50kHz, these higher frequency calls were substantially shorter in duration.

The vast majority studies mapping the primary auditory cortex (A1) of rats used pure tone pips below 32kHz (Zhang et al., 2001; Chang and Merzenich, 2003; Chang et al., 2005; Han et al., 2007; Polley et al., 2007; Engineer et al., 2008; de Villers-Sidani et al., 2008; Zhou and Merzenich, 2008; Insanally et al., 2009) (but see (Sally and Kelly, 1988; Rutkowski, 2003; Kim and Bao, 2009)). As previously reported, we confirm that ethologically relevant rat vocalizations primarily occur above 30kHz (Knutson et al., 2002; Hahn and Lavooy, 2005), indicating that most rat A1 studies have been investigating a frequency range that is ethologically irrelevant to rodents. To address this, we mapped six adult female rats (all older than P150) using 1 to 74kHz tones (and additional 1.2 octave compared to most studies).

### **Ultrasonic Frequencies are Preferentially Represented in A1**

Increasing the stimulus set by 20% lead to a near doubling in frequency-response areas of A1 (Figure 4.3 and 4.5B). Similar results were previous reported by this lab (Kim and Bao, 2009). The general tonotopic gradient is maintained throughout the new ultrasonic frequencies (Figures 4.3A and 4.5A). We find that neurons tuned to higher frequencies show narrower tuning bandwidths. When binning the neurons into octave-sized bins, we see a gradual linear decrease in bandwidth across all animals, with the exception of the two lowest frequency bins (Figure 4.5C). This is most likely due to the fact that these low-frequency neurons are capable of responding to frequencies lower than 1kHz, therefore our bandwidth measures are often an underestimate for the lowest octave band. We find that nearly 75% and 45% of neurons with CF lower than 4kHz have a receptive field that extends beyond 1kHz (Table 4.2).

Linear regression analysis of neurons with receptive fields completely contained within 1 and 74kHz show a highly significant negative correlation between CF and BW for all six animals (average  $r = -0.77$  with  $p < 0.0001$  for all animals).

The receptive field threshold was significantly lower for middle-frequency tuned neurons compared to low and high-tuned neurons (Figure 4.5D), as previously reported (Sally and Kelly, 1988). Maximal response magnitude varied significantly with CF (Figure 4.5E), but no post-hoc pairwise comparisons were significant. However, a significant negative correlation was found between response magnitude and CF ( $r = -0.57$ ,  $p < 0.001$ ), suggesting ultrasonic neurons will on average fire fewer action potentials in response to a tone. We find no significant difference of response latency between frequency bins (Figure 4.5F).

### **Discrimination Ability Correspond with Population Fisher Information**

It has been suggested that to efficiently represent the distribution of sensory variable, there would be an optimal allocation of neurons and spikes in a population (Ganguli and Simoncelli, 2011). For all units, we recorded characteristic V-shaped receptive fields and calculated the tuning curve by collapsing across the highest two dB levels (Figure 4.4). The narrow bandwidths of the ultrasonic neurons are readily apparent, even in an individual animal (Figure 4.6A). Given there are substantially more ultrasonic neurons (Figures 4.3B and 4.5B) and these neurons have narrower tuning bandwidth (Figure 4.5C), it is no surprise we also find that the population tuning curve (an average of all tuning curves) has uniform distributions (4.5A – turquoise line). This would suggest that any given pure-tone will elicit the same number of action potentials from A1. To automatically assess the Fisher information (FI) of these neurons, we fitted a Gaussian to each tuning curve before calculation the FI (single animal example: Figure 4.6B). The population FI (the sum of the individual FIs) is enhanced in the ultrasonic frequency range (Figure 4.6B – purple line), as expected given the narrow bandwidth of these neurons. These effects – uniform population tuning curve and higher FI in the ultrasonic region – were consistent between animals. In addition, the peak of the population FI corresponded highly with the distribution of frequencies found in the USV (Figure 4.6C).

The higher FI values in the ultrasonic region would suggest increase discrimination ability for those frequencies (Schoups et al., 2001; Han et al., 2007; Kim and Bao, 2008). We tested the discrimination ability of 20 adult female rats (half with breeding experience, half naïve virgins) utilizing a protocol previously described (Han et al., 2007). Animals were trained to detect a transition between a train of single-frequency pure-tone pips (the standard frequency) to alternating pure-tone pips (the standard and target frequencies) (Figure 4.7A). The difference between the standard and target frequencies ( $Df$ ) and the standard frequency were both varied. No difference was seen between the naïve animals and those with breeding experience, and the results were combined. We find an enhanced discrimination ability of ultrasonic frequencies with a peak around 32kHz (Figure 4.7B), which highly corresponds to the peak in population FI (Figure 4.6C). This improvement in discrimination was not due to an improved detection ability, as the false alarm rate remained stable around 15% across all frequencies tested.

### **Ultrasonic Representation Develops Relatively Late**

To investigate the development of ultrasonic representation, we mapped A1 of several animals across various developmental stages (Figure 4.8). We find that developing animals show responses to mid-level frequencies, expanding representation of the full spectral range with age, as previously reported (de Villers-Sidani et al., 2007; Insanally et al., 2009). Ultrasonic representation appears to mature at P21, with younger animals showing very little ultrasonic responses. When comparing the amount of cortical area devoted to specific frequency bands between the young (P15-20) and old (P21-26) animals, we find the older animals have significantly more area responding to frequencies above 45kHz, while the younger animals have more area responding to frequencies between 20 and 27kHz (Figure 4.8C).

Response properties were compared between the young and old animals (Figure 4.9). For threshold and response magnitude, there were no significant differences between the two age groups ( $p > 0.10$  for both). However, a highly significant age effect was seen for response latency, with younger developing animals showing longer latency values. The younger group was found to have a marginally narrower bandwidth than the older group ( $p = 0.055$ ), similar to previous reports (de Villers-Sidani et al., 2007; Insanally et al., 2009).

### **Early Experience is Critical for Ultrasonic Representation**

Rat pup isolation calls generally occur between P4 and P16 (Noirot, 1968), during the critical period for tonal representation in the rat (de Villers-Sidani et al., 2007; Insanally et al., 2009). We hypothesized that this early exposure to the 40kHz pup isolation calls could be the mechanism driving the over-representation of ultrasonic frequencies. To test this hypothesis, we reversibly deafened animals at P10 by ligating their ear canals (Popescu and Polley, 2010). Adequate deafening was confirmed by ABRs. After an average of 17.7 days after deafening, six animals (P23, P24, P28, P29, P31, P31) were put under anesthesia and the left ear was removed to recover hearing (confirmed with ABR or visualization of tympanic membrane) while the right auditory cortex was mapped. Previously, it has been reported that rearing in continuous noise (to mask out environmental sounds) will result in a delay in the development of the auditory cortex (Chang and Merzenich, 2003). We find that masking out environmental sounds via deafening leads to substantially normal tonotopy (Figure 4.10A). Similar to that of developing animals, we find that the ligated animals show less representation of frequencies above 45 kHz, but more representation of frequencies between 23 and 32kHz (Figure 4.10B).

When comparing response properties between the ligated animals and our controls, we find no significant differences in bandwidth ( $p = 0.25$ ), threshold ( $p = 0.87$ ), response magnitude ( $p = 0.97$ ), or response latency ( $p = 0.19$ ) (Figure 4.11). However, we do see a trend where ligated animals have slightly longer response latency, similar to that which was seen in the young developing animals. From these results, we presume that early experience of ultrasonic vocalizations is critical to the proper development of representation of ultrasonic frequencies.

## Subcortical Representation of Ultrasonic Frequencies

We find that preferential representation of ultrasonic frequencies is not specific to the cortex. We recorded pure-tone responses from the central nucleus of the inferior colliculus (IC) in 3 adult rats (P45, P46, P50). In one animal, tungsten electrodes were used, while the remaining two, 1x16 polytrodes were used. For one animal, two independent tracks through IC were taken, allowing for 4 sets of data. The superficial layers of IC were tuned to low-frequencies with progressively higher-frequency tuning for lower layers of IC (Figure 4.12A). Much like what we found in A1, high-frequency neurons had narrower bandwidths (Figure 4.12C). To characterize the distribution of frequencies along the depth of the IC, neurons were separated into those with CF smaller than 32kHz (“low frequency neurons”) or greater than 32kHz (“high frequency neurons”). Linear regression of the high, ultrasonic frequency neurons showed significantly shallower slopes than the slopes of the lower frequency neurons (Figure 4.12B). These results indicate that the over-representation of ultrasonic frequencies we find in A1 is not specific to cortex and can be found in subcortical auditory regions.

Responses properties were generally very similar between the IC and A1. We find non-significant differences between IC and A1 in terms of response threshold ( $p = 0.56$ ) and maximal response magnitude ( $p = 0.82$ ) (Figure 4.13). However, a significant difference for bandwidth was found ( $p = 0.003$ ), along with a significant interaction ( $p = 0.0054$ ). We find that IC neurons with CFs between 4 and 32 kHz show significantly broader tuning bandwidths compared to similar A1 neurons (Figure 4.13A). In addition, IC neurons had a marginally significantly faster latency to response compared to A1 neurons ( $p = 0.077$ ), as is expected given that the IC is one of the obligatory relay stations in the lemniscal pathway.

## Gender and Ultrasonic Representation

Since pup isolation and adult encounter USVs can elicit specific behavioral responses in female rats (McIntosh and Barfield, 1978; Ehret and Haack, 1981; Ehret et al., 1987), we were curious if this preferential representation of ultrasonic frequencies was specific to female rats. The full A1, including ultrasonic frequencies, of four juvenile male rats (P30, P32, P36, P39) and two age-matched female littermate rats (P35 and P37) were mapped. Both male and female rats show orderly tonotopy and overrepresentation of ultrasonic frequencies (Figure 4.14). When comparing the male rats to their female littermates, we find no significant effect for gender ( $p = 0.72$ ) and a marginally significant interaction ( $p = 0.060$ ). The marginal significance could be due to the small sample size of the female littermates. When comparing the four males to the six original female adult controls, we again find no significant effect for gender ( $p = 0.88$ ), but a significant interaction ( $p = 0.031$ ). This interaction is most likely due to the fact that the male rats show slightly different frequency distributions between 16 and 32 kHz compared to the female controls (Figure 4.14B).

Comparing response properties between the male animals and control females show many significant differences. The male rats' bandwidth was found to be broader than the female controls (Figure 4.15A;  $p = 0.00011$ ). However, when comparing against the littermate control females, the male rats were found to have narrower bandwidth ( $p = 0.015$ ). Thus, this difference may be due to differences between litters. In addition, we find that male rats have higher response magnitude ( $p < 0.0001$ ) and

longer response latency ( $p = 0.0029$ ) compared to the control females (Figure 4.15C,D). However, when comparing against littermate controls, we find that the males have significantly lower response magnitudes and slower response latencies (data not shown). Again, this is likely due to differences between litters rather than a consistent effect of gender. Threshold was found to not be different between the male animals and either female group (Figure 4.15B).

## **Discussion**

We find that ethologically relevant conspecific vocalization frequencies are preferentially representing the rat primary auditory cortex (Figure 4.3). These ultrasonic frequencies are represented efficiently, leading to an increased behavioral discrimination ability of higher frequencies (Figure 4.7). Interestingly, this overrepresentation of ultrasonic frequencies develops relatively late, around P21 (Figure 4.8). Also, this representation is dependent on early experience with vocalizations: reversibly deafened animals lack representation of very high ultrasonic frequencies (Figure 4.10). In addition, we find that this representation is not specific to cortex (Figure 4.12) and is not specific to gender (Figure 4.14).

### *Is the ultrasonic region a specialized region or a part of A1?*

The most thorough investigation of the organization of the five tonotopically organized auditory regions reports a conspicuous gap between the high-frequency regions of A1 and the anterior auditory field (Polley et al., 2007). However, this study only used frequencies up to 32 kHz, so likely missed was unable to find the tonal receptive field of these neurons. Previous studies that utilize higher frequencies (up to 50, 64, or 73 kHz) show that tonotopy is maintained beyond 32 kHz (Sally and Kelly, 1988; Rutkowski, 2003; Wu et al., 2006; Kim and Bao, 2009; Popescu and Polley, 2010). In one case an orderly reversal is reported between the border of A1 and AAF (Rutkowski, 2003), suggesting the ultrasonic region is definitely a part of A1. Even though the neurons in the highest frequency band have significantly lower bandwidths, higher thresholds and faster response latencies, these differences are part of a gradual continuum that starts well below the ultrasonic range (Figure 4.5). In many of our recordings, we would find that the tuning of the ultrasonic neurons would reverse, indicating a border between A1 and AAF (data not shown). These putative AAF neurons were excluded from all analyses.

### *Ultrasonic representation: comparison between species*

Similar to rats, the primary auditory cortex of cats and mice have been shown to have orderly tonotopy going from low to high frequencies on the caudal-rostral axis in A1 with a reversal of frequencies at the anterior auditory field (AAF) (Andersen et al., 1980; Reale and Imig, 1980; Stiebler et al., 1997; Carrasco and Lomber, 2009; Hackett et al., 2011). Although neurons in the cat A1 are capable of responding to very high ultrasonic frequencies (close to 64kHz), careful analysis have shown that the cat “A1 showed a uniform distribution in the gradient of characteristic frequencies” (Carrasco and Lomber, 2009). Unlike cats, mice A1 and AAF appear to be devoid of very high ultrasonic frequencies. The reversal point between A1 and AAF has been reported to be between 32 and 45 kHz (Stiebler et al., 1997; Hackett et al., 2011), substantially lower than what has been previously reported in rat and what our data would suggest. We propose that the preferential representation of ultrasonic frequencies we find the rat is specialized to this species and plays an important role in the perception of vocalizations.

Interestingly, mouse vocalizations have been reported to be quite similar to rat calls, with the exception of being substantially higher in frequency (Liu et al., 2003; Hahn and Lavooy, 2005; Portfors, 2007), therefore one would assume that

representation of these frequencies are critical for mice. Mice appear to have a separately specialized region dubbed the ultrasonic field (UF), which represents frequencies between 40 and 80 kHz (Stiebler et al., 1997). One possibility is that mouse vocalizations are processed through a specialized region, such as the UF in cortex. Alternatively, it has also been reported that mouse inferior colliculus neurons show robust responses to mouse USVs, even when the vocalization frequencies are substantially higher than the neuron's preferred frequency range (Portfors et al., 2009). Similar results have been found for mouse A1 neurons in our lab (Hamilton et al., personal communication), however the response of rat cortical neurons to vocalizations is still unknown.

### *Future Directions*

Although we find an over-representation of vocalization frequencies in the primary auditory cortex, it is unknown how these neurons will respond to ultrasonic vocalizations. An important future direction in line of research will be to record responses to vocalization stimuli in A1 neurons. It will be interesting to see if the coding strategy apparent in mice (Portfors et al., 2009) is consistent in rats.

In addition to preferential representation of ultrasonic frequencies in the auditory cortex, we find enlarged representation in the inferior colliculus as well (Figure 4.12). The over-representation of ultrasonic frequencies is not specialized or specific to cortex. However, we do not know if the cortical representation is merely a feed-forward readout of subcortical representation, or if cortex is playing an active role with feedback to shape the representation in IC (Winer et al., 2002; Winer, 2006). To further understand this relationship, it will be necessary to map out the development of frequency representation in the IC. If the IC reaches mature representation before P21, we would assume that cortical representation is a feed-forward readout of subcortical representation (cortex reaches mature representation patterns at P21 – Figure 4.8). However, if the IC reaches mature representation after P21, we would assume that cortical feedback plays a critical role in shaping the representation in IC. Further studies are planned to address this question.

Early acoustic experience plays a critical role in the development of appropriate ultrasonic representation (Figure 4.10). Artificial manipulation of the early acoustic environment has also been shown to shape cortical representation (Zhang et al., 2001; Chang and Merzenich, 2003; Han et al., 2007; de Villers-Sidani et al., 2007, 2008; Insanally et al., 2009; Kim and Bao, 2009). In addition, manipulation of vocal output has also been shown to change cortical representation (Cheung et al., 2005). We presume that early exposure to pup isolation calls directs the cortical representation of ultrasonic frequencies, since the deafening of animals leads to immature representation (Figure 4.10). However, it is possible that broadly limiting acoustic input can lead to immature representation (Chang and Merzenich, 2003). Therefore, it may be necessary to leave the animal's hearing intact while changing their vocalization patterns (Cheung et al., 2005). Such an experiment can help define whether or not early experience of ultrasonic vocalizations is necessary for over-representation of ultrasonic frequencies.

## Tables

	P11	P15	Adult
mean center frequency (STD) in kHz **	41.2 (5.16)	47.7 (12.95)	57.8 (9.56)
mean of frequency distribution in kHz	40.8	39.3	58.1
mean BW (STD) in octaves	0.107 (0.030)	0.103 (0.042)	0.103 (0.085)
mean duration (STD) in ms **	117.5 (72.5)	118.6 (112.8)	47.0 (70.4)
<i>r</i> for duration vs. frequency	-0.48 ( $p < 0.001$ )	-0.68 ( $p < 0.001$ )	-0.46 ( $p < 0.001$ )

\*\*  $p < 0.001$  for one-way ANOVA

### **Table 4.1. Ultrasonic vocalization in the rat**

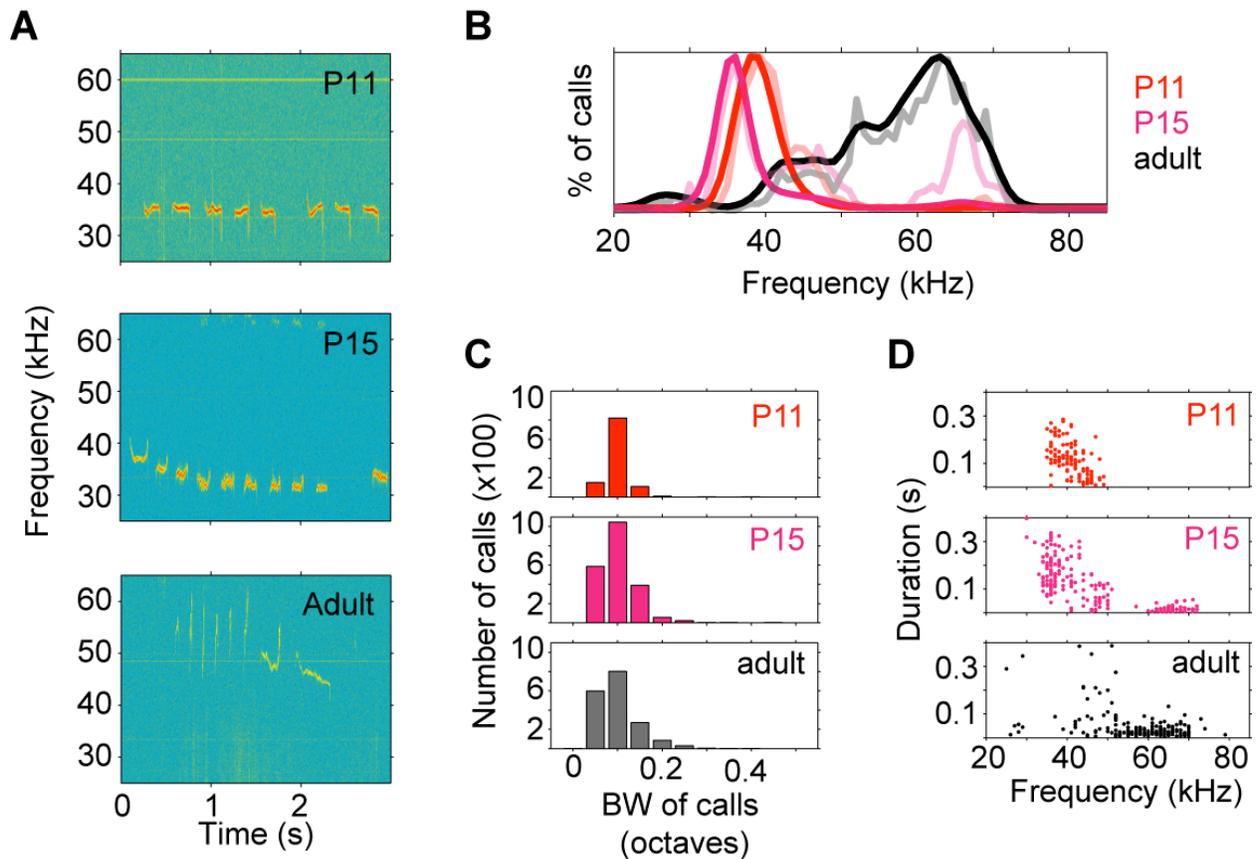
Call properties (frequency, bandwidth and duration) for three different vocalization groups. Although the mean of the center frequency (corresponding to the muted lines in Figure 4.1B) shows a very significant difference between all three group ( $F(2, 5006) = 935, p < 0.001$ ), the mean of the frequency distribution is very similar between the two pup groups (the saturated lines in Figure 4.1B). Bandwidth was not found to be different between the groups. The durations of the calls are significantly different between the pup isolation calls and the adult encounter call. (One-way ANOVA:  $F(2, 5006) = 352, p < 0.001$ ; post-hoc comparisons show no significant difference between the P11 and P15 groups ( $p > 0.5$ )). A significant negative correlation was found between the duration and center frequency of the calls for all animals.

Frequency bin (kHz)	1-2	2-4	4-8	8-16	16-32	32-64
Percentage of tuning curves extending beyond 1 or 74 kHz	73% (26%)	43% (40%)	9.3% (11%)	0% (0%)	1% (3%)	13% (10%)

**Table 4.2. Sound stimulus does not capture complete tuning curve for most low frequency neurons.**

The mean percentage (and standard deviation) of neurons whose tuning curves extended beyond 1 or 74 kHz are reported. The bandwidths of the tuning curves are likely underestimated for most low neurons.

## Figures



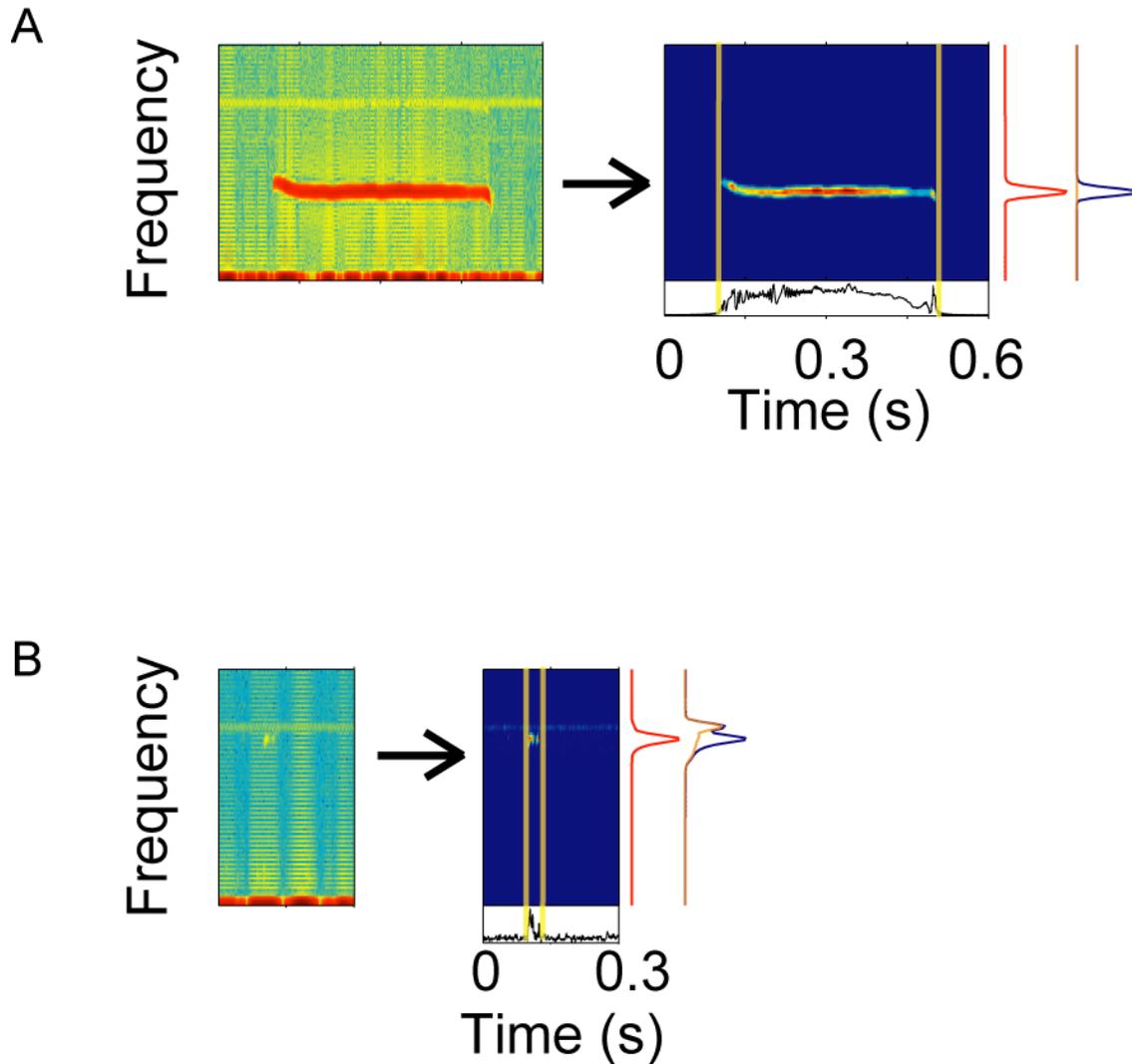
### Figure 4.1. Ultrasonic vocalization in the rat

(A) Example pup isolation (postnatal day 11 and P15) and adult encounter calls.

(B) Peak frequency distribution (muted lines) and average frequency profile (saturated lines) for calls recorded. A one-way ANOVA between P11, P15 and adult calls show a highly significant difference in call frequencies ( $F(2, 5006) = 935, p < 0.001$ ).

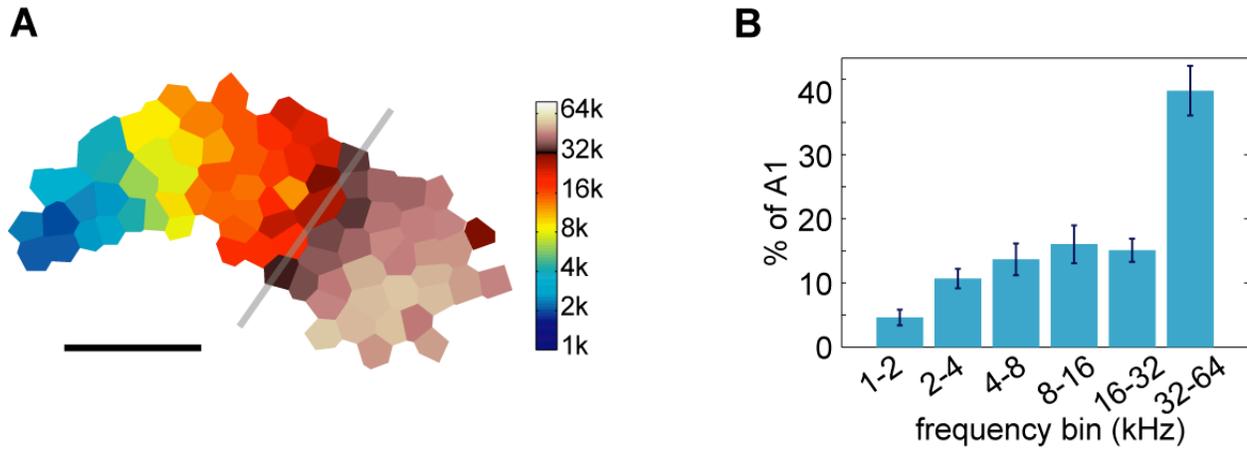
(C) Distribution of call bandwidths (BW) were not different between the three groups (one-way ANOVA:  $F(2, 5006) = 1.56, p = 0.21$ ).

(D) Relationship between center frequency and duration of calls (10% of data shown). Higher frequency calls have shorter durations. This accounts for the difference seen between the distribution of peak frequency and the average frequency profile seen in the P15 pup isolation calls (Figure 1B, pink curves).



#### Figure 4.2. Vocalization analysis methods

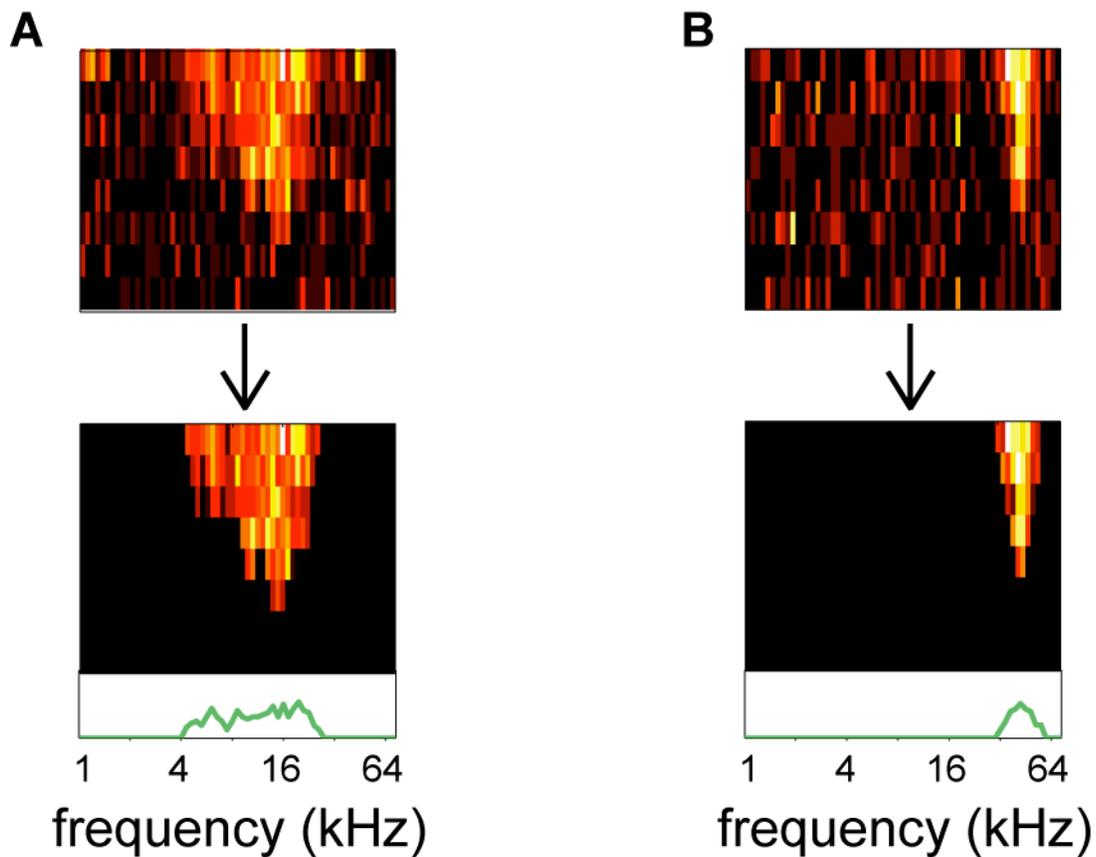
Schematic of how each individual call was analyzed for a pup isolation call (A) and adult encounter call (B). The experimenter visually identified the call's approximate frequency and start time (left). A bandpass filter was applied around the approximate frequency to exclude residual noise (right). A Hilbert transform (black curve) was applied to identify the start and end of each call (indicated by yellow lines). The frequency profile (red curve) was calculated by summing across time along the frequency axis for the duration of the call (blue curve) and subtracting profile of the residual noise (orange curve – taken from the 25-50 ms preceding the start time). Peak frequency and bandwidth were determined by the frequency profile.



**Figure 4.3. Representation of ultrasonic frequencies**

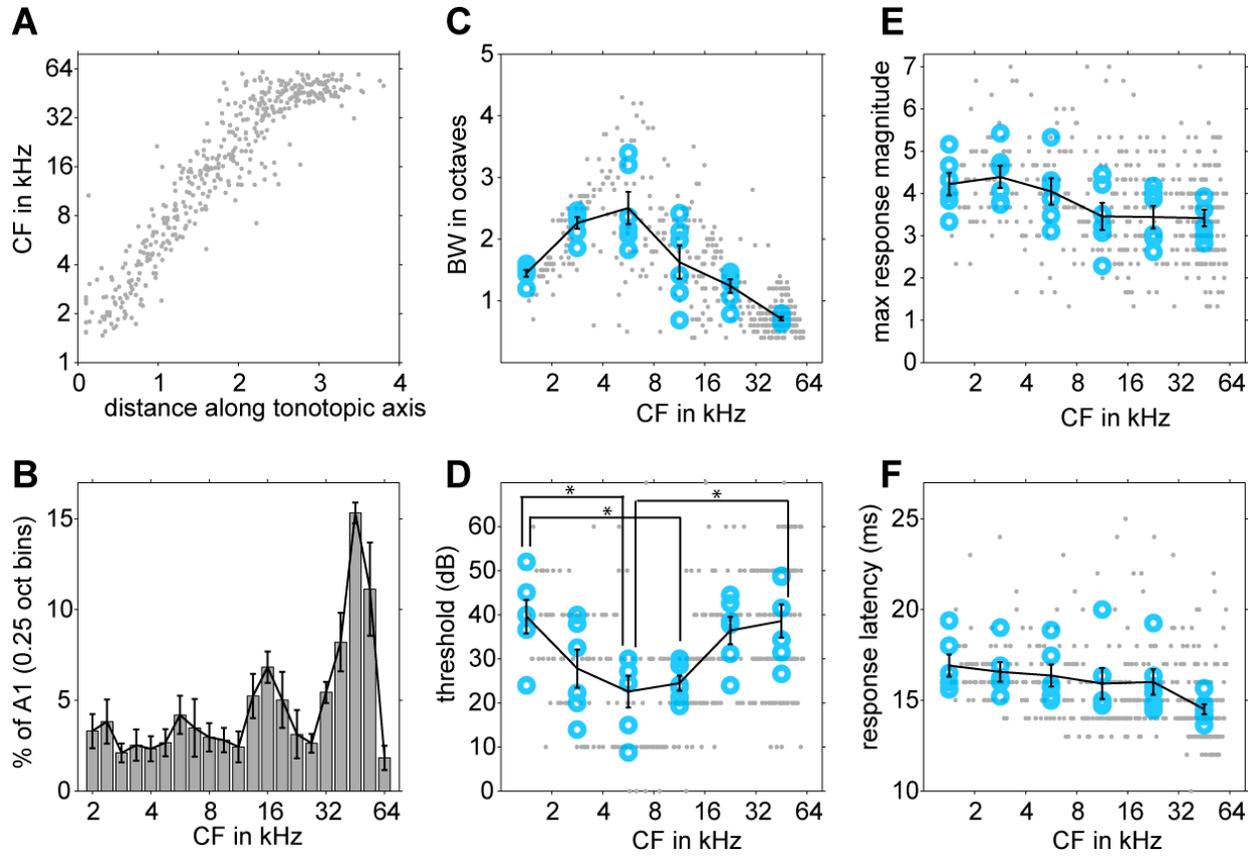
(A) Representative cortical characteristic frequency map, when recording up to 74kHz. Gray line indicates where typical maps would end when recording only to 32kHz. Scale bar: 0.5 mm.

(B) Amount of cortical area representing one-octave frequency bands. Representation of the 32-64kHz band is significantly larger than the other bands.



#### Figure 4.4. Isolation of response properties

Example receptive fields for a mid-frequency (A) and high-frequency (B) neuron. Top panels: raw receptive fields with frequency represented on the x-axis and decibel level represented on the y-axis. Note the V-shaped receptive field commonly found in A1. The maximum firing rate was taken from this receptive field. Middle panels: receptive fields were isolated utilizing a thresholding and filtering algorithm. The characteristic frequency was calculated as the center of mass of this receptive field; the threshold was calculated as the lowest dB level that was still a part of the receptive field. Bottom panels: tuning curves could be derived either by collapsing the receptive field over the top two dB levels. Bandwidth and Fisher Information were calculated using the green tuning curves.



### Figure 4.5. Representation of ultrasonic frequencies and response properties

(A) Distribution of characteristic frequency (CF) across the tonotopic axis.

(B) Percentage of cortical area representing 0.25 octave sized bins.

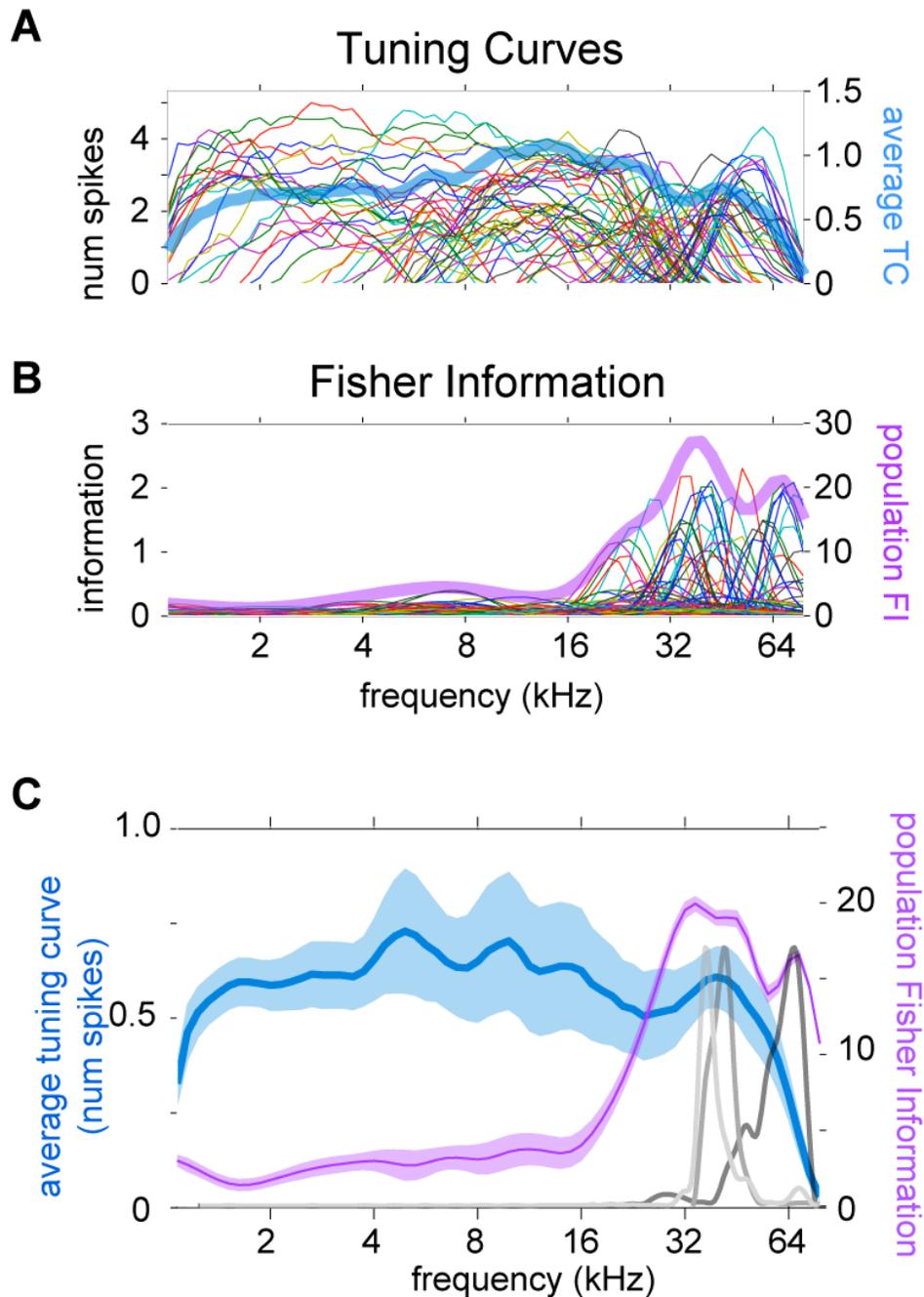
(C,D,E,F) Bandwidth, threshold, maximal response magnitude and response latency as a function of CF. Small gray points indicate individual recording sites; larger black dots represents the octave-band average for each animal; the blue line indicates the average between animals (error bar: SEM).

(C) Neurons with higher CFs have narrower bandwidths. The narrow bandwidth of the lowest two frequency bins is due an edge effect. One-way ANOVA including all data shows a significant CF-dependency on BW ( $F(5,30) = 15.82, p = 1.2 \times 10^{-7}$ ).

(D) Lower thresholds are generally seen for middle-frequency tuned neurons. One-way ANOVA across frequency bands:  $F(5,30) = 4.69, p=0.0028$ . \* indicates  $p<0.05$ , post-hoc comparisons.

(E) For each recording site, the frequency-dB combination that elicited the highest number of action potentials was used. Response magnitude indicates the average number of spikes observed across three repetitions. Although a one-way ANOVA is significant ( $F(5,30) = 2.6, p = 0.046$ ), no pairwise comparisons came out significant in post-hoc analyses. However, a significant negative correlation was found between response magnitude and CF ( $r = -0.57, p < 0.001$ ).

(F) Peak response latency did not show a CF-dependent difference (one-way ANOVA  $F(5,30) = 1.8, p = 0.14$ ).

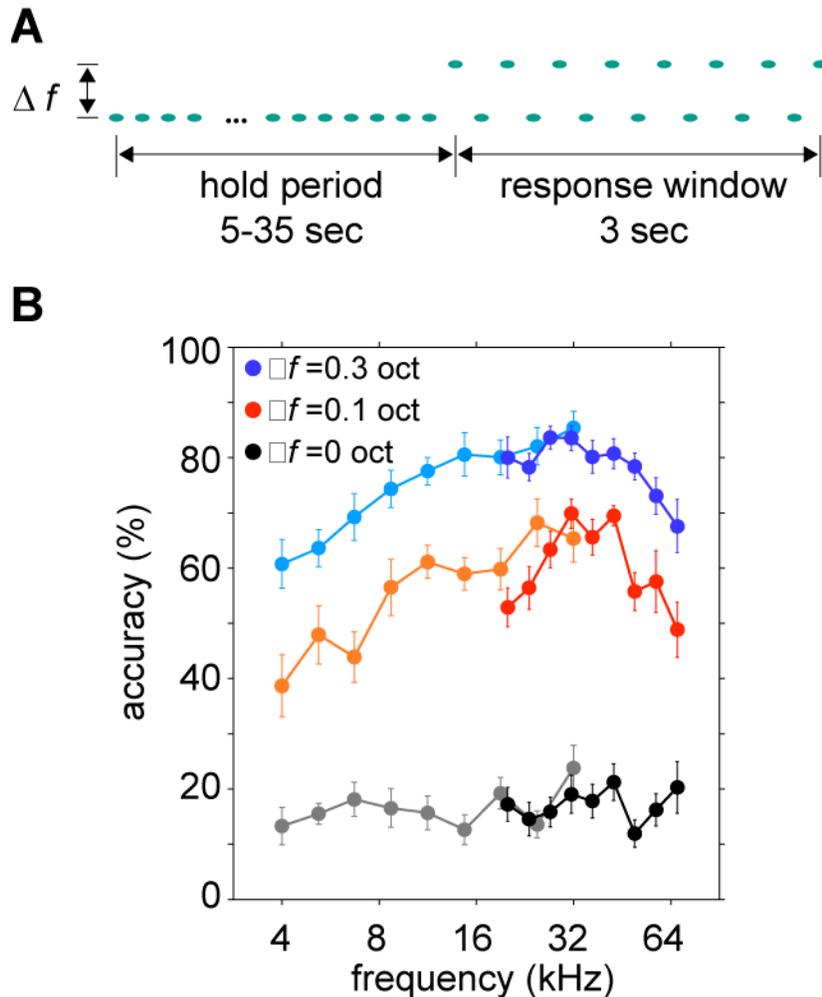


**Figure 4.6. Population Fisher Information**

(A) Tuning curves of all neurons in example map in 4.3A. Thick turquoise line indicates the average tuning curve for this map.

(B) Fisher Information (FI) of all neurons in example map in 4.3A. Thick purple line indicates the population FI this map.

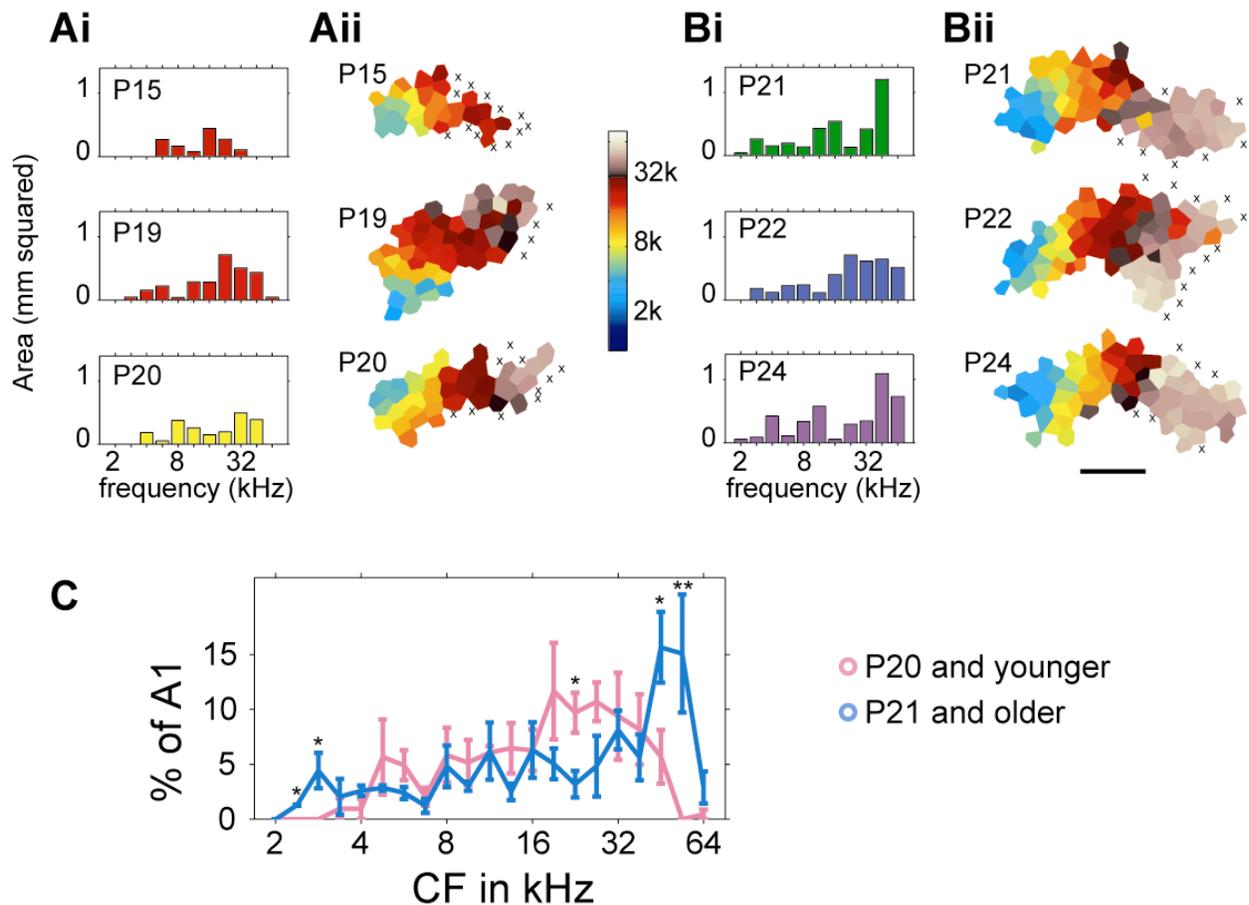
(C) Turquoise and purple curves indicate the average tuning curve and average population FI for all control animals. Error bar = SEM. Gray curves indicate the distribution of frequencies of rat ultrasonic vocalizations as seen in Figure 4.1B.



**Figure 4.7. Discrimination of ultrasonic frequencies**

(A) Schematic of behavioral test. Tone pips (100 ms) were presented at 5 pips per second. During the hold period, a single reference tone was played for 5 to 35 seconds, after which the reference tone alternates with a target tone. The difference between the reference and target tone ( $\Delta f$ ) was varied.

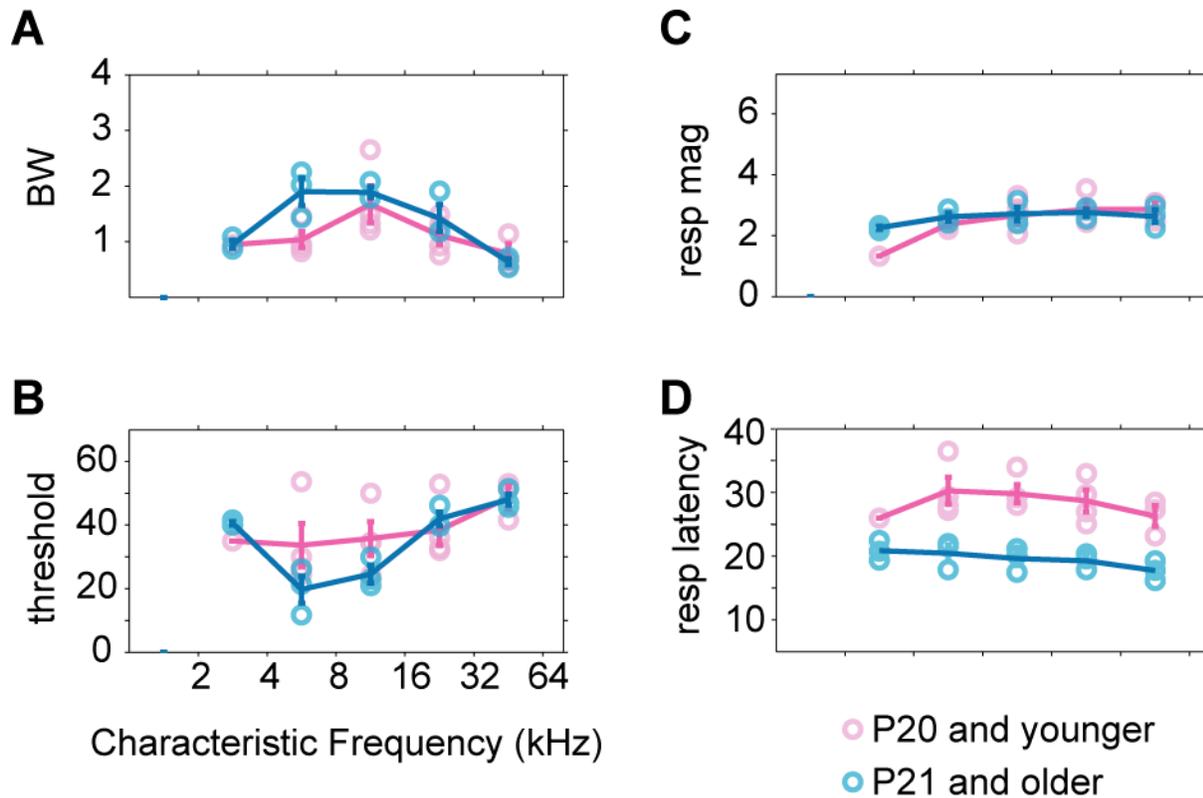
(B) Behavioral results. Animals were trained at  $\Delta f = 1.0$  octave and were tested at 0.5 (not shown), 0.3 (blue traces), 0.1 (red traces) and 0 (black traces) octaves. Animals were first tested on the two ranges (4-32 kHz and 20-64 kHz) on different days, indicated by the lighter and darker colors. Animals showed a steady false alarm rate ( $\Delta f = 0$  octaves) of 10-20% across all frequency levels. The task was consistently more difficult for smaller  $\Delta f$ . Animals also showed improved discrimination ability for frequencies near 32 kHz.



### Figure 4.8. Development of ultrasonic frequencies

(A) Animals younger than P21. (B) Animals P21 and older. (Ai) & (Bi) The distribution of cortical area representing 2-75kHz in young animals ranging from P15 to P24. (Aii) & (Bii) Corresponding maps associated with the histograms in Ai and Bi. x's indicate sites that were not sound responsive; scale bar = 0.5 mm.

(C) Average percentage of representation in 0.25 octave bins. Animals were collapsed as either young ( $n = 4$ , ages: P15, P17, P19, P20) or old ( $n = 4$ , ages: P21, P22, P24, P26). Two-way ANOVA (age x frequency bin) finds a no significant effect for age ( $F(1,126) = 0.01$ ,  $p = 0.94$ ), but highly significant effects for frequency bin ( $F(20,126) = 5.3$ ,  $p < 0.00001$ ) and a highly significant interaction ( $F(20,126) = 3.79$ ,  $p < 0.00001$ ). Interaction is likely due to a significant increase in representation of frequencies above 32kHz in the older group. In addition, the younger group shows a systematic increase in representation of frequencies between 16 and 32 kHz. (\*  $p < 0.05$ , \*\*  $p < 0.01$ , 2-sample t-tests uncorrected).



### Figure 4.9. Response properties in developing animals

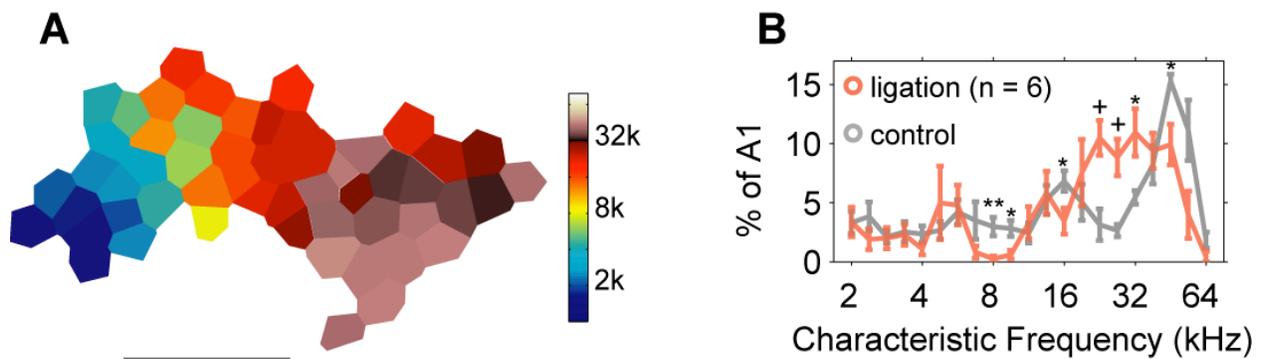
Response properties are calculated in the same way as Figure 4.5. Animals are the same as that of Figure 4.8C. It should be noted that some of the younger animals did not have frequency responses to the lowest and highest bins. These data points were removed from the ANOVAs.

(A) Bandwidth in octaves. A two-way ANOVA (age x frequency bin) reveals a significant effect for frequency bin ( $F(4,26) = 4.93$ ,  $p = 0.0043$ ), but not for age ( $F(1,26) = 4.05$ ,  $p = 0.055$ ). We do not find a significant interaction ( $F(4,26) = 1.58$ ,  $p = 0.208$ ).

(B) Threshold in dB. We a significant effect for frequency ( $F(4,26) = 5.4$ ,  $p = 0.0027$ ), but not for age ( $F(1,26) = 2.44$ ,  $p = 0.13$ ) and there was no significant interaction ( $F(4,26) = 0.92$ ,  $p = 0.47$ ).

(C) Maximal response magnitude in spikes. A two-way ANOVA reveals a significant effect for frequency bin ( $F(4,26) = 6.5$ ,  $p = 0.0009$ ), but not for age ( $F(1,26) = 1.51$ ,  $p = 0.23$ ) and no significant interaction ( $F(4,26) = 1.45$ ,  $p = 0.25$ ).

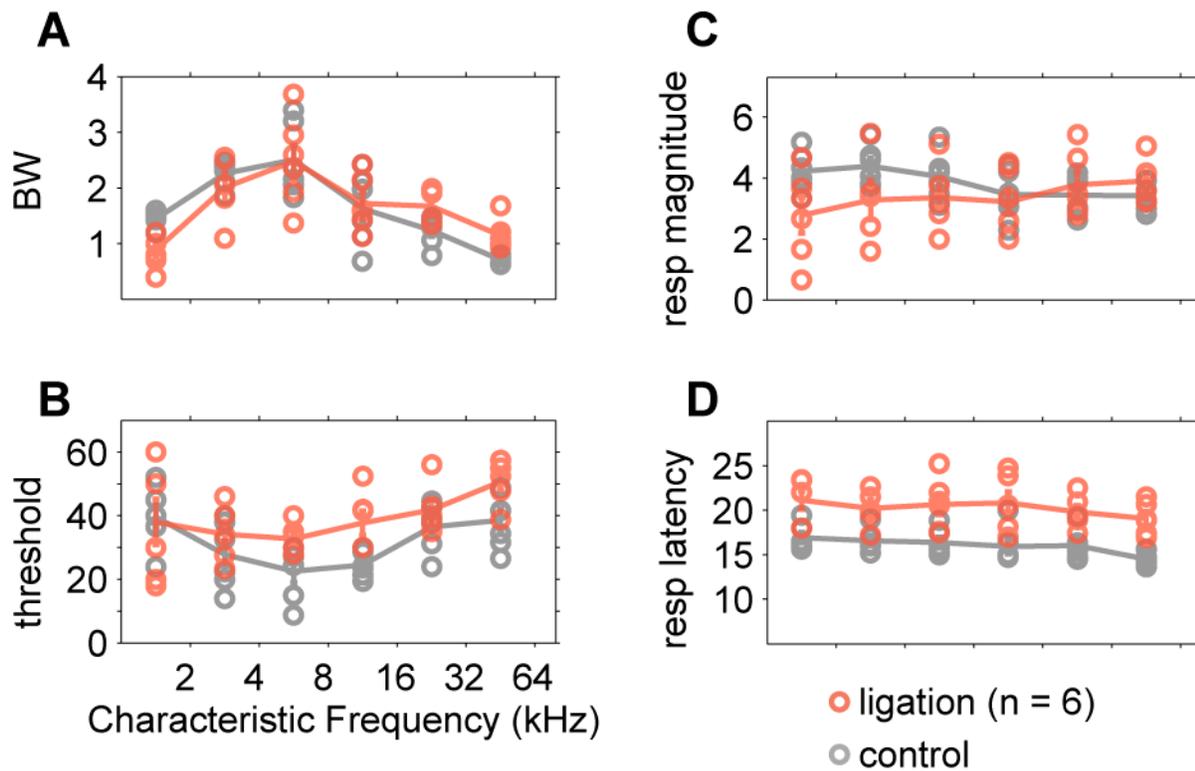
(D) Peak response latency in ms. A two-way ANOVA reveals a significant effect for age ( $F(1,26) = 85.7$ ,  $p < 0.00001$ ), but not for frequency bin ( $F(4,26) = 1.29$ ,  $p = 0.30$ ). We do not find a significant interaction ( $F(4,26) = 0.55$ ,  $p = 0.70$ ). Response latency is significantly different for all testable frequencies ( $p < 0.01$  for the four highest bins).



**Figure 4.10. Ultrasonic representation requires early experience.**

(A) Example map of animal deprived of normal hearing experience from P10 until mapping. Scale bar = 0.5 mm.

(B) Average percentage of representation in 0.25 octave bins. Gray control animals are the same as shown in Figure 4.5B. Two-way ANOVA finds a no significant effect for age ( $F(1,21) = 0.03, p = 0.87$ ), but highly significant effects for frequency bin ( $F(20,210) = 9.7, p < 0.00001$ ) and a highly significant interaction ( $F(20,210) = 3.51, p < 0.00001$ ). Interaction is likely due to a significant increase in representation of frequencies above 32kHz in the older group. In addition, the younger group shows a systematic increase in representation of frequencies between 16 and 32 kHz. (\*  $p < 0.05$ , \*\*  $p < 0.01$ , +  $p < 0.005$ , 2-sample t-tests uncorrected).



### Figure 4.11. Response properties of deafened animals.

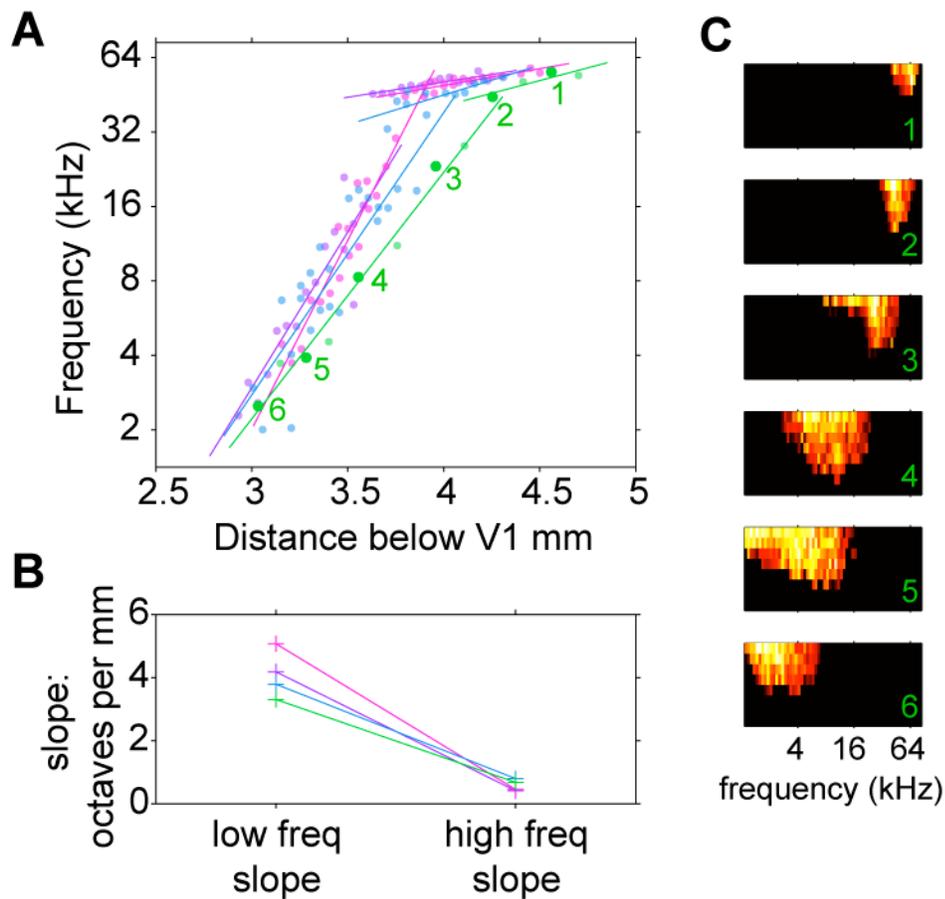
Response properties are calculated in the same way as Figure 4.5. Animals are the same as that of Figure 4.10B. 2-way ANOVAs (condition: ligation group vs control x frequency bin) were conducted as in Figure 4.5

(A) Bandwidth in octaves. We find no main effect for condition ( $F(1,59) = 1.33, p = 0.25$ ), but a significant effect for frequency bin ( $F(5,59) = 11.3, p < 0.0001$ ) and a significant interaction ( $F(5,59) = 10.5, p < 0.0001$ ).

(B) Threshold in dB. No main effect is seen for condition ( $F(1,59) = 0.03, p = 0.87$ ) and there is no interaction ( $F(5,59) = 1.83, p = 0.12$ ). There is a main effect for frequency bin ( $F(5,59) = 7.4, p < 0.0001$ ).

(C) Maximal response magnitude in spikes. Again, there is no main effect is seen for condition ( $F(1,59) = 0, p = 0.97$ ) and there is no interaction ( $F(5,59) = 0.45, p = 0.81$ ). There is a main effect for frequency bin ( $F(5,59) = 2.8, p = 0.035$ ).

(D) Peak response latency in ms. There is a trend towards the ligated group having longer latency ( $F(1,59) = 1.76, p = 0.19$ ). A significant effect is seen for frequency bin ( $F(5,59) = 16.89, p < 0.0001$ ), but no interaction ( $F(5,59) = 0.24, p = 0.94$ ).



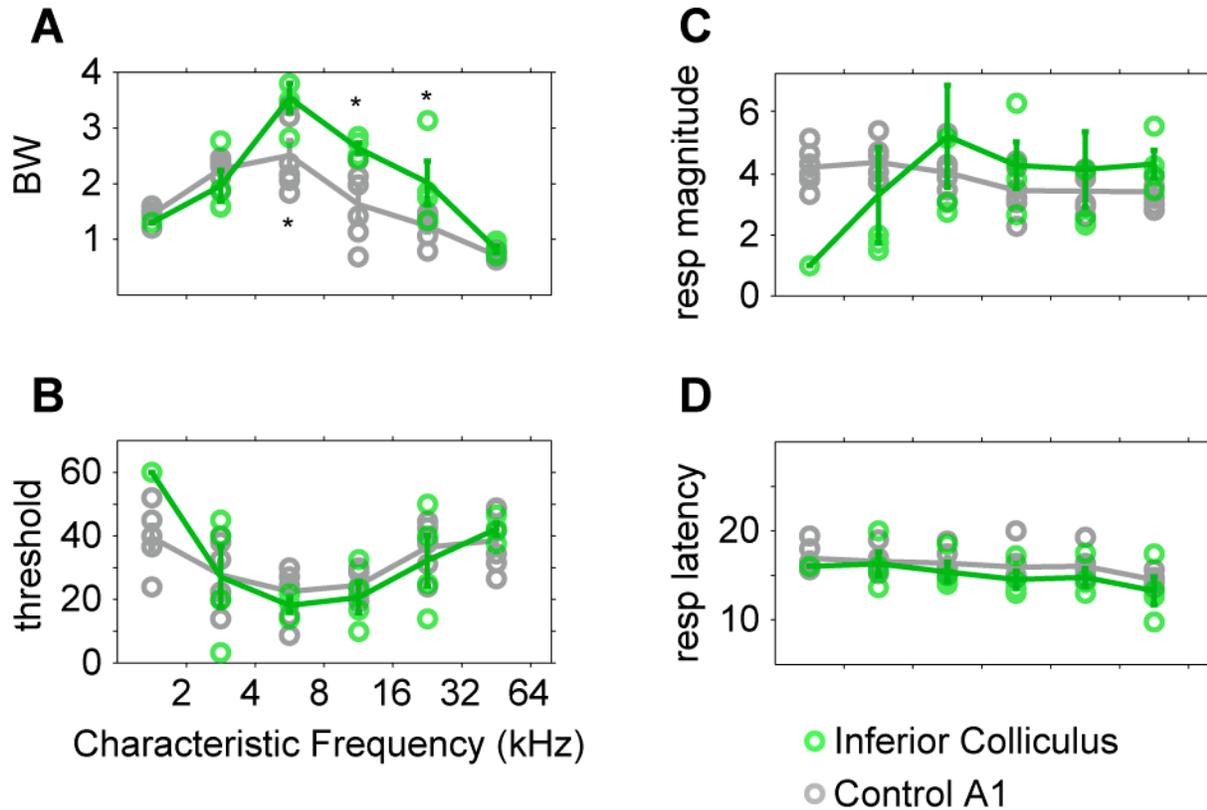
**Figure 4.12. Increased representation of ultrasonic frequencies in inferior colliculus**

(A) CF distribution as a function of distance below V1 for four independent tracts (across 3 animals). Lines of best fit are shown for sites tuned to below or above 32kHz. Numbers indicate sites with example receptive fields in C.

(B) Slope of lines of best fit for low frequency (<32kHz) and high frequency (>32kHz) neurons.

The decrease in slope indicates an over-representation of high, ultrasonic neurons. ( $p = 0.0042$ , paired t-test)

(C) Example receptive fields as indicated in A.



### Figure 4.13. Response properties in inferior colliculus

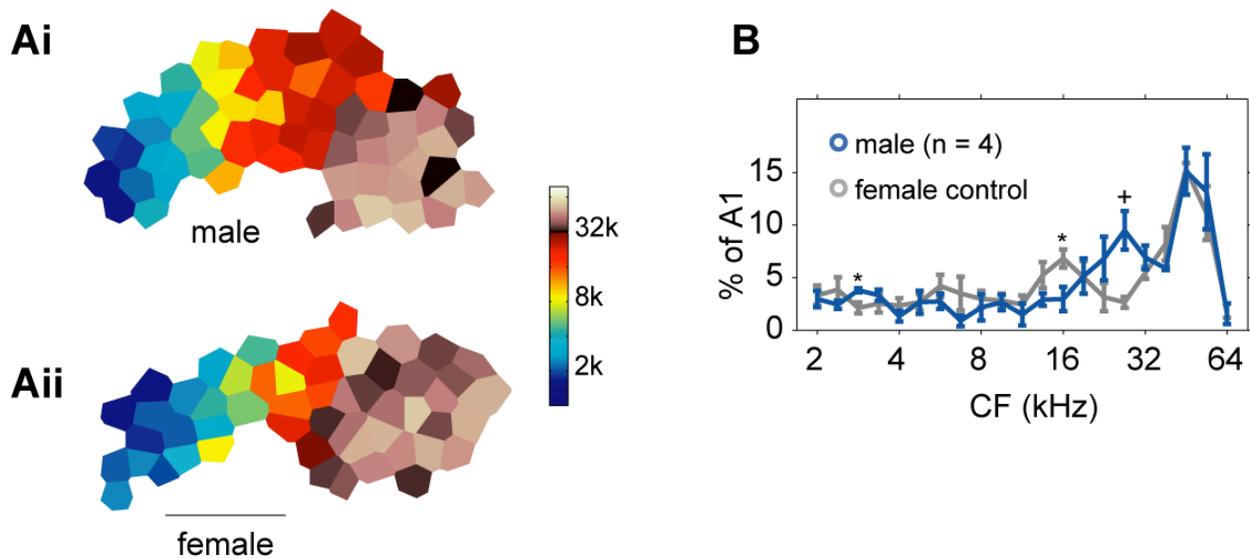
Response properties are calculated in the same way as Figure 4.5. Animals are the same as that of Figure 4.12. 2-way ANOVAs was conducted (condition: IC vs A1 x frequency bin) as in Figure 4.5.

(A) Bandwidth in octaves. We find significant main effects for condition ( $F(1,45) = 9.82$ ,  $p = 0.003$ ), frequency ( $F(5,45) = 27.85$ ,  $p < 0.0001$ ), and an interaction ( $F(5,45) = 3.86$ ,  $p = 0.0054$ ). Interaction due to the fact that neurons in IC tuned between 4kHz and 32kHz have broader bandwidths (\*  $p < 0.05$ ; uncorrected t-test).

(B) Threshold in dB. No effect was found for condition ( $F(1,45) = 0.35$ ,  $p = 0.556$ ) and no interaction ( $F(5,45) = 1.01$ ,  $p = 0.42$ ) was found. There was a significant effect for frequency ( $F(5,45) = 7.77$ ,  $p < 0.0001$ ).

(C) Maximal response magnitude in spikes. No main effects (condition:  $F(1,45) = 0.05$ ,  $p = 0.82$ ; frequency:  $F(5,45) = 0.96$ ,  $p = 0.45$ ) or interaction ( $F(5,45) = 1.65$ ,  $p = 0.17$ ) were found.

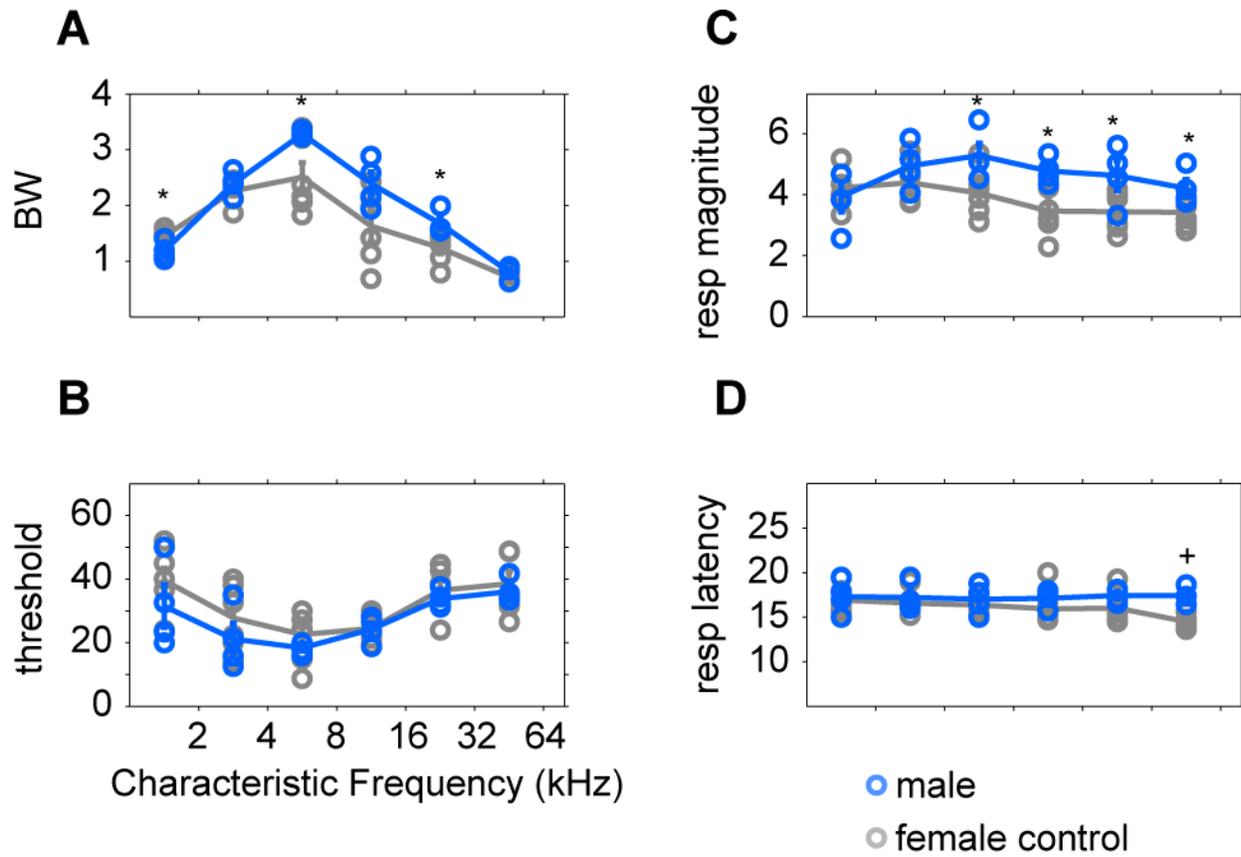
(D) Peak response latency in ms. Response latency was found to be marginally faster in the IC group ( $F(1,45) = 3.27$ ,  $p = 0.077$ ), with a marginal main effect for frequency ( $F(5,45) = 2.23$ ,  $p = 0.068$ ). No interaction was found ( $F(5,45) = 0.11$ ,  $p = 0.99$ ).



#### Figure 4.14. Gender

(A) Example maps of male (Ai) and female (Aii) animals. No obvious difference seen. Scale bar = 0.5 mm.

(B) Average percentage of representation in 0.25 octave bins. Gray control animals are the same as shown in Figure 4.5B. Two-way ANOVA finds a no significant effect for gender ( $F(1,168) = 0.02, p = 0.88$ ), but highly significant effects for frequency bin ( $F(20,168) = 14.75, p < 0.00001$ ) and a significant interaction ( $F(20,168) = 1.74, p = 0.031$ ). Interaction is likely due to small differences seen between the male and female groups between 16 and 32 kHz. (\*  $p < 0.05$ , +  $p < 0.005$ , 2-sample t-tests uncorrected).



### Figure 4.15. Gender Response Properties

Response properties are calculated in the same way as Figure 4.5. Animals are the same as that of Figure 4.14. 2-way ANOVAs was conducted (condition: male vs female x frequency bin) as in Figure 4.5.

(A) Bandwidth in octaves. Significant main effects (condition:  $F(1,48) = 12.17, p = 0.0011$ ; frequency:  $F(5,48) = 45.91, p < 0.0001$ ) and interaction ( $F(5,48) = 3.28, p = 0.0125$ ) was found. Males having broader bandwidths only in mid-frequency neurons likely drove interaction.

(B) Threshold in dB. A marginally significant effect was found for condition ( $F(1,48) = 3.63, p = 0.063$ ) and highly significant main effect for frequency ( $F(5,48) = 7.71, p < 0.0001$ ). No interaction was found ( $F(5,48) = 0.31, p = 0.90$ ).

(C) Maximal response magnitude in spikes. A highly significant effect was found for condition ( $F(1,48) = 17.84, p = 0.0001$ ) and marginally significant main effect for frequency ( $F(5,48) = 2.36, p = 0.054$ ). No interaction was found ( $F(5,48) = 1.74, p = 0.14$ ).

(D) Peak response latency in ms. A significant effect for condition ( $F(1,48) = 9.84, p = 0.0029$ ) was found with no effect for frequency ( $F(5,48) = 0.67, p = 0.65$ ) and no interaction ( $F(5,48) = 0.97, p = 0.45$ ). The main effect for latency was likely driven by the males having a longer response latency for the highest frequency bin.

## **References**

- Andersen RA, Knight PL, Merzenich MM (1980) The thalamocortical and corticothalamic connections of AI, AII, and the anterior auditory field (AFF) in the cat: Evidence of two largely segregated systems of connections. *The Journal of Comparative Neurology* 194:663–701.
- Brudzynski SM (2005) Principles of rat communication: quantitative parameters of ultrasonic calls in rats. *Behavior Genetics* 35:85–92.
- Brudzynski SM, Bihari F, Ociepa D, Fu XW (1993) Analysis of 22 kHz ultrasonic vocalization in laboratory rats: long and short calls. *Physiology & Behavior* 54:215–221.
- Brudzynski SM, Kehoe P, Callahan M (1999) Sonographic structure of isolation-induced ultrasonic calls of rat pups. *Developmental Psychobiology* 34:195–204.
- Butts D a, Goldman MS (2006) Tuning curves, neuronal variability, and sensory coding. *PLoS Biology* 4:e92.
- Carrasco A, Lomber SG (2009) Differential modulatory influences between primary auditory cortex and the anterior auditory field. *The Journal of Neuroscience* 29:8350–8362.
- Chang EF, Bao S, Imaizumi K, Schreiner CE, Merzenich MM (2005) Development of spectral and temporal response selectivity in the auditory cortex. *Proceedings of the National Academy of Sciences* 102:16460–16465.
- Chang EF, Merzenich MM (2003) Environmental noise retards auditory cortical development. *Science* 300:498–502.
- Cheung SW, Nagarajan SS, Schreiner CE, Bedenbaugh PH, Wong A (2005) Plasticity in primary auditory cortex of monkeys with altered vocal production. *The Journal of Neuroscience* 25:2490–2503.
- Cohen L, Rothschild G, Mizrahi A (2011) Multisensory Integration of Natural Odors and Sounds in the Auditory Cortex. *Neuron* 72:357–369.
- Ehret G, Haack B (1981) Categorical perception of mouse pup ultrasound by lactating females. *Die Naturwissenschaften* 68:208–209.
- Ehret G, Koch M, Haack B, Markl H (1987) Sex and parental experience determine the onset of an instinctive behavior in mice. *Die Naturwissenschaften* 74:47.
- Engineer CT, Perez C a, Chen YH, Carraway RS, Reed AC, Shetake J a, Jakkamsetti V, Chang KQ, Kilgard MP (2008) Cortical activity patterns predict speech discrimination ability. *Nature Neuroscience* 11:603–608.
- Fischer BJ, Peña JL (2011) Owl's behavior and neural representation predicted by Bayesian inference. *Nature Neuroscience* 14:1061–1066.
- Ganguli D, Simoncelli EP (2011) Implicit encoding of prior probabilities in optimal neural populations. *Processing*:6–9.
- Geal-Dor M, Freeman S, Li G, Sohmer H (1993) Development of hearing in neonatal rats: air and bone conducted ABR thresholds. *Hearing Research* 69:236–242.
- Girshick AR, Landy MS, Simoncelli EP (2011) Cardinal rules: visual orientation perception reflects knowledge of environmental statistics. *Nature Neuroscience* 14:926–932.
- Hackett T a, Barkat TR, O'Brien BMJ, Hensch TK, Polley DB (2011) Linking topography to tonotopy in the mouse auditory thalamocortical circuit. *The Journal of Neuroscience* 31:2983–2995.

- Hahn ME, Lavooy MJ (2005) A review of the methods of studies on infant ultrasound production and maternal retrieval in small rodents. *Behavior Genetics* 35:31–52.
- Han YK, Köver H, Insanally MN, Semerdjian JH, Bao S (2007) Early experience impairs perceptual discrimination. *Nature Neuroscience* 10:1191–1197.
- Insanally MN, Köver H, Kim H, Bao S (2009) Feature-dependent sensitive periods in the development of complex sound representation. *The Journal of Neuroscience* 29:5456–5462.
- Kim H, Bao S (2008) Distributed representation of perceptual categories in the auditory cortex. *Journal of Computational Neuroscience* 24:277–290.
- Kim H, Bao S (2009) Selective increase in representations of sounds repeated at an ethological rate. *The Journal of Neuroscience* 29:5163–5169.
- Knutson B, Burgdorf J, Panksepp J (2002) Ultrasonic vocalizations as indices of affective states in rats. *Psychological Bulletin* 128:961–977.
- Liu RC, Linden JF, Schreiner CE (2006) Improved cortical entrainment to infant communication calls in mothers compared with virgin mice. *The European Journal of Neuroscience* 23:3087–3097.
- Liu RC, Miller KD, Merzenich MM, Schreiner CE (2003) Acoustic variability and distinguishability among mouse ultrasound vocalizations. *The Journal of the Acoustical Society of America* 114:3412.
- McIntosh TK, Barfield RJ (1978) Ultrasonic vocalisations facilitate sexual behaviour of female rats. *Nature* 272:163–164.
- Noirot E (1968) Ultrasounds in young rodents. II. Changes with age in albino rats. *Animal Behaviour* 16:129–134.
- Polley DB, Read HL, Storace D a, Merzenich MM (2007) Multiparametric auditory receptive field organization across five cortical fields in the albino rat. *Journal of Neurophysiology* 97:3621–3638.
- Popescu MV, Polley DB (2010) Monaural deprivation disrupts development of binaural selectivity in auditory midbrain and cortex. *Neuron* 65:718–731.
- Portfors CV (2007) Types and functions of ultrasonic vocalizations in laboratory rats and mice. *Journal of the American Association for Laboratory Animal Science* 46:28–34.
- Portfors CV, Roberts PD, Jonson K (2009) Over-representation of species-specific vocalizations in the awake mouse inferior colliculus. *Neuroscience* 162:486–500.
- Reale RA, Imig TJ (1980) Tonotopic organization in auditory cortex of the cat. *The Journal of Comparative Neurology* 192:265–291.
- Romand R, Gunter E (1990) Development of tonotopy in the inferior colliculus. I. Electrophysiological mapping in house mice. *Developmental Brain Research* 54:221–234.
- Rutkowski R (2003) Characterisation of multiple physiological fields within the anatomical core of rat auditory cortex. *Hearing Research* 181:116–130.
- Sally SL, Kelly JB (1988) Organization of auditory cortex in the albino rat: sound frequency. *Journal of Neurophysiology* 59:1627–1638.
- Schoups A, Vogels R, Qian N, Orban G (2001) Practising orientation identification improves orientation coding in V1 neurons. *Nature* 412:549–553.
- Seung HS, Sompolinsky H (1993) Simple models for reading neuronal population codes. *Proceedings of the National Academy of Sciences* 90:10749–10753.

- Stiebler I, Neulist R, Fichtel I, Ehret G (1997) The auditory cortex of the house mouse: left-right differences, tonotopic organization and quantitative analysis of frequency representation. *Journal of Comparative Physiology* 181:559–571.
- Takahashi N, Kashino M, Hironaka N (2010) Structure of rat ultrasonic vocalizations and its relevance to behavior. *PLoS One* 5:e14115.
- de Villers-Sidani E, Chang EF, Bao S, Merzenich MM (2007) Critical period window for spectral tuning defined in the primary auditory cortex (A1) in the rat. *The Journal of Neuroscience* 27:180–189.
- de Villers-Sidani E, Simpson KL, Lu Y-F, Lin RCS, Merzenich MM (2008) Manipulating critical period closure across different sectors of the primary auditory cortex. *Nature Neuroscience* 11:957–965.
- Winer J (2006) Decoding the auditory corticofugal systems\*. *Hearing Research* 212:1–8.
- Winer J a, Chernock ML, Larue DT, Cheung SW (2002) Descending projections to the inferior colliculus from the posterior thalamus and the auditory cortex in rat, cat, and monkey. *Hearing Research* 168:181–195.
- Wu GK, Li P, Tao HW, Zhang LI (2006) Nonmonotonic synaptic excitation and imbalanced inhibition underlying cortical intensity tuning. *Neuron* 52:705–715.
- Zhang LI, Bao S, Merzenich MM (2001) Persistent and specific influences of early acoustic environments on primary auditory cortex. *Nature Neuroscience* 4:1123–1130.
- Zhou X, Merzenich MM (2008) Enduring effects of early structured noise exposure on temporal modulation in the primary auditory cortex. *Proceedings of the National Academy of Sciences* 105:4423.