

**Experience Dependent Coding of Intonations by Offsets in Mouse Auditory Cortex**

A Dissertation  
Presented to  
The Academic Faculty

By

Kelly K. Chong

In Partial Fulfillment  
of the Requirements for the Degree  
Doctor of Philosophy in the  
School of Biomedical Engineering

Georgia Institute of Technology  
Emory University  
May 2019

Copyright © 2019 by Kelly K. Chong

## Experience Dependent Coding of Intonations by Offsets in Mouse Auditory Cortex

Approved by:

Dr. Robert C. Liu, Advisor  
Department of Biology  
*Emory University*

Dr. Dieter Jaeger  
Department of Biology  
*Emory University*

Dr. Lynne C. Nygaard  
Department of Psychology  
*Emory University*

Dr. Samuel J. Sober  
Department of Biology  
*Emory University*

Dr. Garrett B. Stanley  
School of Biomedical Engineering  
*Georgia Institute of Technology*

Date Approved:  
December 6, 2018

## ACKNOWLEDGEMENTS

The process of completing this work could not have been possible without the support and guidance of many people along the way. First and foremost, I would like to thank my advisor, Dr. Robert Liu, for his advice and guidance throughout my time in the lab. He would constantly push me to become something beyond simply a method-savvy lab technician and helped me grow both as a person and as a scientist. I also would like to thank Katy Shepard, for helping to shape my understanding of the scientific field of our study. I would also like to thank Tamara Ivanova and Chip Mappus for their patience and technical expertise as they trained me in the handling of live animals and surgical techniques.

I would also like to thank Alex Dunlap, for the endless intellectually stimulating discussions we have had and the various side projects and collaborations in our studies throughout the years. I also would like to thank Amielle Moreno, for her infectious enthusiasm and support in histology and molecular work. I would also like to thank Liz Anne Amadei and Jim Kwon for their support, and Lily Shen for all her help over the years. I would like to thank Gerald Wong and Dori Kascoh for their patience as my first undergraduate mentees.

I would also like to thank Dr. Dieter Jaeger for allowing me to share his lab space in the pursuit of voltage imaging, and for generously lending use of a high-speed camera and transgenic mouse strain. I would like to thank Dr. Lynne Nygaard for her guidance and advice in the drafting and approval process of human studies to extend our mouse auditory findings. I also would like to thank my other committee members, Dr. Samuel Sober and Dr. Garrett Stanley for their helpful advice, time, and consideration.

Additionally, I would like to thank my family and friends, who have helped me retain my sanity throughout the process. I would like to thank my parents, Jocelyn and Chenlook Chong, and my siblings, Kevin and Kyle Chong for all their support. I would also like to thank the Georgia Tech fighting game community for their endless enthusiasm, determination, and perseverance that inspired me to soldier onward during times of discouragement in the program.

## TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	iii
LIST OF TABLES.....	x
LIST OF FIGURES.....	xi
LIST OF SYMBOLS AND ABBREVIATIONS.....	xiii
SUMMARY.....	xvii
<b>CHAPTER 1: INTRODUCTION.....</b>	<b>1</b>
1.1 The Auditory Periphery.....	1
1.2 Classes of Auditory Stimuli.....	3
1.2.1 Pure Tones.....	3
1.2.2 Linear Frequency Modulation.....	3
1.2.3 Two-tone Stimuli.....	4
1.2.4 STRFs and Ripple Stimuli.....	5
1.2.5 Sinusoidally Frequency Modulated Tones.....	7
1.3 On and Offset Responses.....	9
1.3.1 Early Awake Anecdotal Reports of Offset Responses.....	10
1.3.2 On and Offset Response Properties.....	11
1.4 Core and Noncore Auditory Cortex.....	13
1.4.1 Noncore Auditory Cortex Physiological Properties.....	13
1.5 Natural Vocalizations and Auditory Cortical Plasticity.....	15
1.5.1 The Maternal Mouse Communication Model.....	16

1.5.2 Cortical Plasticity in the Maternal Model .....	16
1.6 Summary and Significance .....	19
<b>CHAPTER 2: Experience Dependent Coding of Intonations by Offsets in Mouse</b>	
<b>Auditory Cortex .....</b>	<b>21</b>
2.1 Background .....	21
2.2 Materials and Methods .....	24
2.2.1 Experimental Details .....	24
2.2.1.1 Experimental model - mouse .....	24
2.2.1.2 Sinusoidally Frequency Modulated (sFM) USV Modeling .....	24
2.2.1.3 Sound Stimulus Playback .....	25
2.2.1.4 Single Unit Awake Electrophysiology .....	26
2.2.2 Data Analysis .....	28
2.2.2.1 Analysis of $A_{fm}$ and $f_{fm}$ Tuning around Best Frequency .....	28
2.2.2.2 Pure Tone Integration Model for Prediction of $A_{fm}$ Stimulus Spike Rates .....	29
2.2.2.3 sFM Matched Noise and Linear FM Analysis .....	29
2.2.2.4 Spike Density Plot .....	30
2.2.2.5 On and Offset Prevalence Analysis .....	30
2.2.2.6 Generalized Linear Modeling .....	31
2.2.2.7 Logistic Regression Modeling for Classification of sFM Parameter Combinations as Pup-like vs Adult-like .....	32
2.2.2.8 $A_{fm}$ Tuning Analysis .....	32

2.3 Results .....	34
2.3.1 SUs can show tuning to $A_{fm}$ .....	34
2.3.2 SUs can show tuning to $A_{fm}$ and $f_{fm}$ specifically in their Offset response .....	36
2.3.3 SU Offset responses are more likely to prefer sFM stimuli over pure tone compared to On responses .....	38
2.3.4 SUs show Heterogeneity in timing of responses to USVs .....	42
2.3.5 Strength of On Responses to USVs decreases with experience .....	44
2.3.6 Prevalence of Offset Response to Pup USVs increases with experience .....	45
2.3.7 Maternal SUs are capable of discriminating between USV categories with matched onset properties for longer calls only .....	46
2.3.8 Acoustic Features Encoded by On and Offset Responses via Generalized Linear Modeling .....	49
2.3.9 Maternal SUs respond to the same number of pup USVs on average as nonmaternal SUs .....	51
2.3.10 Population of Maternal and Nonmaternal SUs equally likely to respond at On or Offset of all USVs .....	52
2.3.11 Sinusoidally Frequency Modulated (sFM) Tones can be used as a model of Natural Vocalization Frequency Trajectories .....	53
2.3.12 sFM Parameters can be used to classify call types using a Logistic Regression Model .....	55
2.3.13 A2 On-only and Offset-only responses are suppressed in Maternal units for Adult-like Pup USVs .....	56
2.3.14 A2 Offset responses in Maternal Units are tuned to higher $A_{fm}$ .....	57

2.3.15 Pup-like Calls have Higher $A_{fm}$ Values .....	60
2.4 Discussion .....	61
<b>CHAPTER 3: CONCLUSIONS</b> .....	<b>64</b>
3.1 Frequency Trajectory Encoding in the Auditory Cortex .....	64
3.2 The Mouse Model for Experience Dependent Plasticity .....	66
3.3 Offset Responses in the Auditory Cortex.....	66
3.4 The Role of Secondary Auditory Cortex in Frequency Trajectory Encoding.....	68
3.5 Future Directions.....	69
APPENDIX A: List of Single Units used in Experiments.....	71
A.1 List of Units in On and Offset Prevalence Analysis .....	71
A.2 List of Units in $A_{fm}$ $f_{fm}$ 8x8 Tuning around Pure Tone .....	74
A.3 List of Units in $A_{fm}$ Tuning around USV .....	76
APPENDIX B: Sound Stimulus Parameters.....	78
B.1 Pure Tone Tuning Stimulus .....	78
B.2 Natural Mouse Ultrasonic Vocalization Stimulus .....	79
B.3 $A_{fm}$ and $f_{fm}$ Tuning Stimulus (8x8).....	80
B.4 $A_{fm}$ and $f_{fm}$ Tuning Stimulus Matched Bandwidth Noise Control .....	82
B.5 $A_{fm}$ Tuning Stimulus around USV .....	83
B.6 $A_{fm}$ Matched Bandwidth Noise Stimulus around USV .....	84
APPENDIX C: Yoked Cage Design .....	86
APPENDIX D: Immunohistochemistry .....	89

D.1 Alignment of Atlas and Physiologically Defined Auditory Regions .....	89
D.2 Parvalbumin Expression in Secondary Auditory Cortex .....	92
D.3 Parvalbumin Staining Protocol.....	94
APPENDIX E: Human Psychophysics .....	95
E.1 Introduction .....	95
E.2 Experimental Design .....	95
E.3 Results .....	98
APPENDIX F: Stimulus Specific Adaptation (SSA).....	102
F.1 Verifying SSA in Awake Auditory Cortex .....	102
F.2 Investigating the Presence of Category Specific Adaptation (CSA) .....	104
APPENDIX G: Voltage Sensitive Dyes and Fluorescent Protein Imaging .....	108
G.1 Voltage Sensitive Dye Imaging .....	108
G.2 Voltage Sensitive Fluorescent Protein Imaging .....	110
APPENDIX H: Photoidentification of Neuronal Populations (PINP) and Viral Vectors..	114
H.1 PINP .....	115
H.2 PINP with Transsynaptic Tracing .....	119
REFERENCES .....	125

## LIST OF TABLES

Table 2.1: GLM Parameter Estimates when fitting sFM parameters of USVs to spike rate responses to USVs .....	50
Table A.1: List of Units in On and Offset Prevalence Analysis .....	71
Table A.2: List of Units in $A_{fm}$ $f_{fm}$ 8x8 Tuning around Pure Tone.....	75
Table A.3: List of Units in $A_{fm}$ Tuning around USV .....	76
Table B.1: Pure tone tuning stimulus parameters .....	78
Table B.2: Natural USV stimulus sFM parameters.....	79
Table B.3: sFM 8x8 $A_{fm}$ and $f_{fm}$ tuning around BF stimulus parameters .....	81
Table B.4: Matched bandwidth noise control stimulus parameters.....	82
Table B.5: $A_{fm}$ tuning around best USV stimulus parameters .....	84
Table H.1: Laser power calibration .....	116

## LIST OF FIGURES

Figure 1.1: The Central Auditory Pathway .....	2
Figure 1.2: Sample Spectrotemporal Receptive Field .....	6
Figure 1.3: Neuron On and Offset responses can be tuned to different frequencies .....	11
Figure 1.4: Mouse Auditory Cortex .....	14
Figure 2.1: sFM model and sample single unit $A_{fm}$ tuning .....	35
Figure 2.2: $A_{fm}$ and $f_{fm}$ tuning in a sample single unit .....	37
Figure 2.3: Population sFM Tuning Data .....	40
Figure 2.4: USV On and Offset Responses across units .....	43
Figure 2.5: On and Offset Responses to USVs across Region and Experience .....	45
Figure 2.6: Maternal SUs discriminate USVs with matched onset properties for long calls .....	47
Figure 2.7: sFM six parameter distribution of pup and adult USVs .....	49
Figure 2.8: sFM works as a model for USVs and can be used to discriminate call categories using a Logistic Regression model .....	54
Figure 2.9: sFM Parameter Distribution and Unit Responses to USVs classified by ROC .....	57
Figure 2.10: $A_{fm}$ tuning around USVs changes with experience .....	59
Figure B.1: Distribution of $A_{fm}$ in natural USVs .....	83
Figure C.1: Cartoon schematic of yoked cage setup .....	87
Figure C.2: Cage divider specifications, outer .....	88

Figure D.1: Craniotomy High Resolution Image with Electrode Penetrations and Histological Section .....	90
Figure D.2: Cell Counting and Quantification methods .....	91
Figure D.3: Parvalbumin expression across Physiologically Defined Auditory Cortical Regions .....	92
Figure E.1: Sample Images of Human Psychophysics Task .....	96
Figure E.2: Proportion of Vocalizations Correctly Classified as Familiar by Subjects ....	99
Figure E.3: High False Positive Reporting as Familiar by Subjects .....	99
Figure E.4: Noise Loudness Rating Reports.....	101
Figure F.1: Verifying SSA for Pure Tones in Awake Auditory Cortex .....	104
Figure F.2: Mouse USVs selected for SSA and CSA stimulus .....	105
Figure F.3: No SSA or CSA seen for Mouse USVs .....	107
Figure G.1: Auditory Cortical Voltage Sensitive RH-1691 Dye Imaging .....	109
Figure G.2: Histological images of VSFPB expression in Mouse Auditory Cortex.....	112
Figure G.3: Time course images of VSFPB in Awake Auditory Cortex .....	113
Figure H.1: Image of PINP electrode and optic fiber over Mouse Left Auditory Cortex	117
Figure H.2: Sample light-driven response.....	117
Figure H.3: Histological Section showing localization of ChR2 expression in Auditory Cortex .....	119
Figure H.4: Histological Confirmation of AAV Transsynaptic activity .....	121
Figure H.5: Sample Projection Receiving PINP-identified Single Unit .....	122

## LIST OF SYMBOLS AND ABBREVIATIONS

cm	Centimeters
dB	Decibel
dB SPL	Decibel sound pressure level
dBa	Decibels of Attenuation
h	Hour
Hz	Hertz
kg	Kilogram
kHz	Kilohertz
mg	Milligram
mm	Millimeter
ms	Millisecond
mW	Milliwatt
MΩ	Mega Ohm
nm	Nanometer
rad	Radians
s	Second
μm	Micrometers
A1	Primary Auditory Cortex
A2	Secondary Auditory Cortex
AAF	Anterior Auditory Field
AAFA1	Primary Auditory Cortex to Anterior Auditory Field Transition
AAV	Adeno Associated Virus

AC	Auditory Cortex
$A_{fm}$	Amplitude of Frequency Modulation
AM	Amplitude Modulation
anim	Animal
Au1	Auditory Cortex, Primary (Atlas)
AUC	Area Under Curve
AuD	Auditory Cortex, Dorsal (Atlas)
AuV	Auditory Cortex, Ventral (Atlas)
BF	Best Frequency
CB	Calbindin
ChR2	Channelrhodopsin 2
CI	Confidence Interval
Ctrl	Control
CSA	Category Specific Adaptation
Dev	Deviant
DP	Dorsoposterior Field
dur	Duration
EMX1	Empty Spiracles Homeobox 1 (Pyramidal Neuron Marker)
EYFP	Enhanced Yellow Fluorescent Protein
$f_0$	Onset Frequency
$f_{fm}$	Frequency of Frequency Modulation
$f_{slope}$	Slope of Linear Frequency Modulation
FDR	False Discovery Rate
FM	Frequency Modulation
Freq	Frequency

GEVI	Genetically Encoded Voltage Indicator
GFP	Green Fluorescent Protein
GLM	Generalized Linear Model
HSD	Honestly Significantly Different
IACUC	Institutional Animal Care and Use Committee
IFM	Linear Frequency Modulation
IP	Intraperitoneal
M	Maternal
MGB	Medial Geniculate Nucleus
MU	Multiunit
Nm	Nonmaternal
P	Postnatal
PINP	Photoidentification of Neuronal Populations
Pref	Preference
PSTH	PeriStimulus Time Histogram
PT	Pure Tone
PV	Parvalbumin
Q40	Tuning Bandwidth 40dB above response threshold
Resp	Response
ROC	Receiver Operating Curve
sFM	Sinusoidal Frequency Modulation
SNR	Signal to Noise Ratio
SSA	Stimulus Specific Adaptation
SST	Somatostatin
std	Standard Deviation

stim	Stimulus
STRF	Spectrotemporal Receptive Field
SU	Single Unit
UF	Ultrasound Field
USV	Ultrasonic Vocalization
V	Voltage
VIP	Vasoactive Intestinal Peptide
VSD	Voltage Sensitive Dye
VSFP	Voltage Sensitive Fluorescent Protein
VSFPB	Voltage Sensitive Fluorescent Protein Butterfly 1.2

## SUMMARY

Frequency modulations are an inherent feature of many behaviorally relevant sounds, including vocalizations and music. Changing trajectories in a sound's frequency often convey meaningful information, which can be used for differentiating sound categories, such as in the case of intonations in tonal languages. However, how frequency modulations are encoded within the auditory system is still an open question. In particular, it is not clear what features of the neural responses in what parts of the auditory pathway (e.g. primary or secondary auditory cortex) might be more important for conveying information about behaviorally relevant frequency modulations, and how these responses may change with experience.

Our work utilizes a natural paradigm in which mouse mothers learn the behavioral significance of pup ultrasonic vocalizations during maternal experience to study how mice use frequency trajectory to discriminate vocalization categories. We model the whistle-like mouse vocalizations using a parameterized sinusoidally frequency modulated (sFM) tone such that the vocalization's entire frequency trajectory is captured by a set of six parameters. We employed a combination of in-vivo head-fixed awake single unit electrophysiology and modeling of the natural mouse vocalization repertoire to explore neural sensitivity in frequency trajectory parameter space.

We obtained recordings across core (A1, AAF, UF) and secondary auditory (A2) regions, from animals with and without maternal experience. We found that neurons across the entire hearing range can show tuning to sFM parameters such as amplitude and frequency of frequency modulation. Offset responses, or spiking occurring after the end of stimulus playback, were more likely to prefer sFM stimuli over pure tones compared to On responses that occurred during stimulus playback. In addition, we found an increased prevalence of Offset responses accompanied by a decreased prevalence

and strength of On responses to natural vocalizations after maternal experience in A2, while no changes were seen in A1/AAF/UF, although changes in subsets of Core auditory neurons may be obscured at the population level. Moreover, in maternal A2 units, there was a bias in both On and Offset responses that favored vocalizations that have pup-like sFM parameters, as defined by a nominal logistic regression model. This bias can be explained by a shift in the tuning of maternal A2 neurons towards parameter values that are more characteristic of the pup vocalization category. This work furthers our understanding of how nonprimary auditory cortex attunes to features in acoustic space, demonstrating that offset tuning to frequency trajectory in secondary auditory cortex plays a role in natural sound category learning.

## Chapter 1

### Introduction

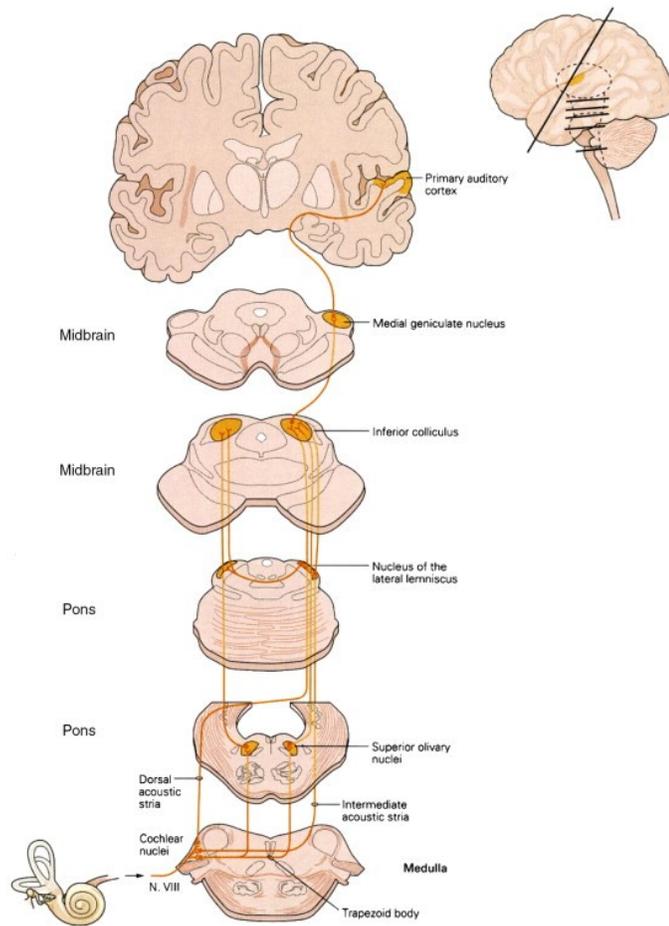
In order to navigate the environment, organisms must utilize various sensory cues given off by their surroundings. Objects in motion can distort the air, creating pressure waves in the form of sound. Over the course of evolution, organisms have adapted sensory organs capable of detecting the frequencies of vibration in the air, known as the sense of hearing. The ability to discriminate the frequency content within a sound provides useful information about the source of the sound waves. As organisms became more complex, organisms began to take advantage of the sense of hearing to produce communicative sounds that are informative and beneficial to survival. As these communication signals became more complex, organisms must have had to learn to discriminate the different types of informative sounds in the environment and respond appropriately.

#### 1.1 The Auditory Periphery

The most primitive form of terrestrial hearing was present in early amphibious creatures, which utilized a primitive oval window, or otic capsule. The otic capsule conducts pressure waves generated by objects moving under water, as seen in lungfish (Christensen-Dalsgaard et al., 2011). The hearing organ present in most terrestrial animals, which consists of an additional middle and outer ear, presumably evolved as a consequence of air being a less potent conductor of sound vibrations compared to water. The middle ear in terrestrial animals, including mammals, is capable of amplifying vibrations from the air by up to a thousand-fold, which enhances the sensitivity of the hearing organ. It consists of the outer ear, where the pinna captures incoming sound, and sound waves are transduced through the ear canal to the tympanic membrane.

Sound will cause the tympanic membrane to vibrate, and three small bones in the middle ear, the malleus, incus, and stapes, also known as the ossicles, mechanically amplify the vibrations experienced by the tympanic membrane and transduce them to the oval window of the inner ear, which is connected to the cochlea.

The cochlea is the primary hearing organ and resides in the inner ear. The sensory neurons of the cochlea are known as inner hair cells, and will vibrate depending on the frequencies present within the sound, following the principle of resonance. From the cochlea, canonically, the information is transmitted to the cochlear nucleus in the brain stem, followed by the superior olivary complex, nucleus of the lateral lemniscus, inferior colliculus, medial geniculate nucleus of the thalamus, and the auditory cortex (Fig 1.1).



**Figure 1.1** The Central Auditory Pathway. Depicts auditory processing brain nuclei starting from the cochlea ascending upwards to the auditory cortex. Upper right image depicts the whole brain and the location of individual slices that depict each auditory processing node. Reproduced from (Graven et al., 2008).

Additional non-canonical auditory pathways exist (Flores et al., 2015), although the work in this thesis will focus on the pathway leading to the auditory cortex. All areas of the auditory system are generally responsive to sound, although response properties will vary at different processing nodes along the pathway (Aitkin et al., 1971; Moller et al., 1982; Romand et al., 1990; Thornton, 1976; Whitfield, 1956). Researchers have studied the response properties of the auditory system are by using a variety of different types of auditory stimuli and measuring neural responses.

## **1.2 Classes of Auditory Stimuli**

### **1.2.1 Pure Tones**

Pure tones have been a useful tool in the study of the auditory cortex, being simple stimuli with only a single frequency that does not change across time. Many important characteristics of the auditory cortex and its neural response properties have been characterized using pure tones, including tonotopy, sound intensity sensitivity, tuning bandwidth, and the timing of response onset to tones (Heil, 1997a; Reale et al., 1980; Schreiner et al., 1992; Sutter et al., 1995). However, at the same time, pure tones are not a commonly occurring auditory stimulus in the natural environment; natural sounds are often much more complex in frequency content and envelope compared to pure tones.

### **1.2.2 Linear Frequency Modulation**

Studies also looked into linear or logarithmic frequency modulation in auditory stimuli. In studies of the more peripheral auditory system, upon presentation of

frequency modulated stimuli, up or down sweep preference was observed in the cat auditory nerve fiber (Sinex et al., 1981) and cochlear nucleus (Britt et al., 1976), as well as in the bat cochlear nucleus (Suga, 1964). Upon reaching more central areas of auditory processing, in addition to seeing sweep directional preference, sensitivity to the speed of the frequency modulation was also observed. Sweep velocity sensitivity was absent from the earlier auditory areas, suggesting that further auditory processing gives rise to this response feature. Sweep velocity sensitivity was well-documented in the primary auditory cortex across multiple species, including cat (Heil et al., 1992), bat (Shannon-Hartman et al., 1992), ferret (Nelken et al., 2000), squirrel monkey (Godey et al., 2005), owl monkey (Atencio et al., 2007), and chinchilla (Brown et al., 2009).

Others have attempted to utilize neural responses to frequency modulation to try to predict responses to natural vocalizations. Using a generalized linear nonlinear model as a framework, neural responses of rat primary auditory cortex (A1) to bouts of vocalizations could be predicted using only the frequency modulation and amplitude as input parameters (Carruthers et al., 2013). A little over half of the recorded neurons exhibited a high prediction accuracy using this model, which is impressive given the simplicity of the model. However, overall frequency modulation and single-tone amplitude do not entirely capture the acoustic variability in natural vocalizations, and responses from neurons that had poor prediction accuracy could be better captured by other parameters that may not have been included in the model.

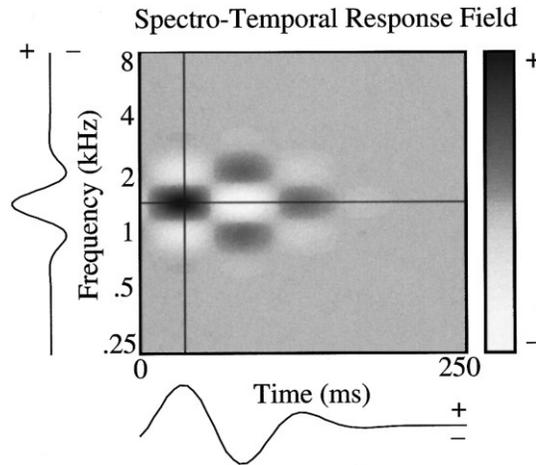
### ***1.2.3 Two-tone Stimuli***

In parallel, other methods of exploring a more complex auditory response space were investigated using a library of sounds containing more complex frequency content. Two-tone stimuli, for example, have been used to demonstrate that a primary auditory

cortical neuron's response to a single tone can be altered if a second tone is played simultaneously, beyond what might be expected from the linear sum of responses to each individual tone. The second tone does not necessarily have to be within the frequency range the neuron responds to when played alone. The presence of a second tone outside its response area can still have facilitative or suppressive effects on the neuron's responses to a tone that falls within its response area (Nelken et al., 1994; Shamma et al., 1993). This illustrates that the responses of the auditory cortex to stimuli with multiple frequencies cannot be entirely explained by the neuron's responses to each of the individual frequencies alone. The authors of the two-tone studies posited that there may be higher order interactions involved beyond just a summation of excitatory pure tone responses when multiple frequencies are present in a stimulus.

#### ***1.2.4 STRFs and Ripple Stimuli***

As another method of characterizing auditory neural responses to sounds, the idea of the spectrotemporal receptive field (STRF) was introduced (Aertsen et al., 1981). The concept of the STRF is similar to the spike-triggered average. In theory, by playing back a large library of sounds with varying spectral properties, the average spectral content of a stimulus over time that precedes each individual spike from a neuron can be used to reconstruct a neuron's STRF. The STRF is a representation of the precise spectrotemporal area that the neuron will respond to in time and frequency space (**Fig 1.2**).



**Figure 1.2** Image of an idealized STRF, with collapsed time and frequency domains depicted along their respective axes. Reproduced from (Depireux et al., 2001).

Studies have demonstrated that complex STRFs can be constructed from primary auditory cortical neurons, using both pure tone and auditory grating or ripple stimuli, which consist of multiple tones sweeping up or downwards. In primate, it was found that only 5% of primary auditory cortical neurons showed an STRF with only a single region of increased activity without any regions of inhibition; all other neurons showed more complex STRFs (deCharms et al., 1998). In this same study on primates, stimuli that theoretically should maximally excite a given neuron could be constructed from its STRF. The study found that the STRF is capable of inducing prolonged, tonic activity in neurons for about 50% of the recorded A1 neurons. Prolonged activity is typical for a sensory neuron when its best or preferred stimulus is presented, as seen in the visual cortex (Maunsell et al., 1992) as well as the auditory cortex (deCharms et al., 1996; Wang et al., 2005). This demonstrates that the STRF is indeed useful for describing the response characteristics of neurons, and could find a “preferred stimulus” for about half of the neurons in A1. However, the other half is not well-characterized by the STRF, and their responses may involve more complex stimulus features that were not captured by STRFs constructed from pure tones or tone gratings. In ferrets,

depending on whether the ripple stimulus is sweeping upwards or downwards, the constructed STRF for the same neuron can be different (Depireux et al., 2001; Kowalski et al., 1996). Of neurons in A1, most neurons have this asymmetric property depending upon its sensitivity to sweep direction, which illustrates the nonlinearity of auditory cortical responses.

Others have extended the construction of STRFs to natural stimuli libraries, such as in birdsong (Theunissen et al., 2000). In urethane-anesthetized zebra finches, neurons from Field L, the bird primary auditory cortex analog, STRFs that are derived from natural sound libraries were shown to be superior to those generated from a set of pure tones. However, even natural sound derived STRFs did not capture all the response dynamics seen in auditory cortical neurons, particularly for the neurons with nonlinear behavior. The STRF itself is limited as it assumes that there is a linear mapping between the stimulus and neural response, and cannot capture short term changes such as stimulus specific adaptation (Zhao et al., 2011). Although useful, STRFs do not perfectly describe neural response properties. Rather than creating a single, static image of what an auditory cortical neuron responds best to using an STRF, it may be useful to also investigate how neurons respond, and are perhaps tuned, to different individual acoustic features that describe natural sounds. Higher order features of the sound such as its temporal envelope or frequency modulation can be captured by parameterized modeling of these features in natural sounds.

### ***1.2.5 Sinusoidally Frequency Modulated Tones***

Frequency modulations are an important feature of natural sounds. For example, in the discrimination of different speech phonemes, frequency modulation can play a critical role (Shannon et al., 1995; Stickney et al., 2005; Zeng et al., 2005). In addition, deficits in frequency modulation processing have been associated with impairments in

language learning (Tallal et al., 1996). The way in which frequency modulations are encoded cannot be ascertained with pure tone stimuli alone, given the nonlinearities present in auditory cortical responses.

Starting from the periphery, the encoding of frequency modulation is thought to be more faithful to the acoustics of the stimulus. The auditory nerve, for example, phase locks to frequency or amplitude modulations present in a stimulus (Johnson, 1980; Joris et al., 1992; Palmer, 1982). However, as one moves up the auditory processing pathway, the modulation rates that auditory processing areas are capable of phase locking to become progressively lower. In the inferior colliculus and auditory thalamus, phase locking can be seen in the range of 100 – 1000 Hz (Creutzfeldt et al., 1980; Frisina et al., 1990; L. F. Liu et al., 2006), whereas at the auditory cortex, this drops around tens of Hz in cat (Eggermont, 1991, 1994; Schreiner et al., 1988), rat (Gaese et al., 1995), and mouse (R. C. Liu et al., 2006). Rather than using phase-locking, or a temporal code, the cortex begins to represent faster modulation rates with increased firing rates, or a rate code (Lu et al., 2001).

The auditory cortex also shows a preference for frequency modulated sounds, such as linear or sinusoidal frequency modulation, over pure tones. This preference for frequency modulation, as well as amplitude modulation, over pure tones has been observed in human fMRI imaging studies, manifesting as greater activation of auditory fields including Heschl's gyrus and the lateral supratemporal plane (Hall et al., 2002; Hart et al., 2003). Preference for temporally modulated tones over pure tones was also shown through electrophysiology in awake marmosets (Liang et al., 2002). In this study, either frequency modulation (FM) or amplitude modulation (AM) was introduced to 1-second pure tones presented to marmosets while recording in A1. The study first found the best AM or FM depth for each neuron, and then would vary the frequency of AM or FM at that fixed depth to obtain each neuron's response curves. As a result, the depth of

AM or FM modulation was not the focus of the study, although interesting results regarding modulation depth were still reported. Researchers found a nonmonotonic variation in firing rate as FM depth was varied. Units tended to respond best to a specific value of FM depth that varied depending on individual unit, ranging anywhere from 1/20 to 1/2 octave. Amplitude modulation depth, on the other hand, appeared to elicit responses in a monotonic fashion, and for the purposes the study, was always fixed at 100% amplitude modulation to elicit the best response. In this study, marmoset A1 was found to respond best to AM or FM frequencies of about 16 to 32 Hz, and responses to FM or AM stimuli were relatively similar. However, as mentioned earlier, this work did not explicitly consider tuning to FM depth, which is another important feature of natural sounds, and does not consider how experience may affect the encoding of FM.

### **1.3 On and Offset Responses**

Early electrophysiological studies in auditory cortex primarily reported Onset responses, or responses that occur immediately after stimulus playback begins. Most of these studies were conducted in the cat primary auditory cortex (Heil, 1997a, 1997b; Hind, 1953; Mendelson et al., 1997; Merzenich et al., 1975; Phillips et al., 1990; Phillips et al., 1994; Schreiner et al., 1992; Sutter et al., 1995), under barbiturate anesthesia. From these studies, they were still able to demonstrate tonotopy of the primary auditory cortex (Hind, 1953). In addition, the anesthetized work conducted important characterization of multiple auditory cortical response properties, including tuning bandwidths in the primary auditory cortex (Schreiner et al., 1992), how variation in sound intensity affects tuning (Sutter et al., 1995), first spike latency (Heil, 1997a; Mendelson et al., 1997), as well as amplitude modulation (Heil, 1997b). Of note is that in all these studies conducted under barbiturate anesthesia, primarily Onset responses were observed, whereas Offset responses were absent.

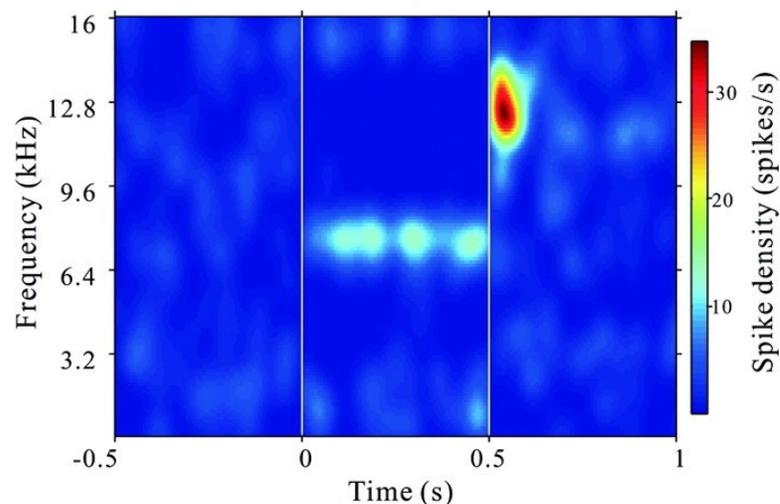
In contrast, early human studies in awake subjects that utilize techniques such as magnetoencephalography have reported significant Onset as well as Offset responses in the coils localized to the putative auditory cortex (Hari et al., 1987). Since then, other reports of the presence of auditory cortical activity locked specifically to the Offset of stimulus presentation have appeared using the auditory-evoked potential (Takahashi et al., 2004) and functional magnetic resonance imaging (Harms et al., 2005; Okada et al., 2004). The reports of Offset responses in this awake human literature in contrast to electrophysiological work done in anesthetized cat illustrates the possibility that the anesthesia preparation may be altering the responses observed. This then motivates electrophysiological studies done in an awake preparation.

### ***1.3.1 Early Awake Anecdotal Reports of Offset Responses***

Electrophysiological studies in non-barbiturate anesthetized animals, and later in awake animals show responses occurring at the Offset of stimulus playback, leading to the conclusion that barbiturate anesthesia specifically was eliminating the presence of Offset responses. Offset responses have been reported when using ketamine or halothane anesthesia in the cat auditory cortex (Moshitch et al., 2006; Volkov et al., 1991). Sustained responses to tones as well as responses at the Offset of stimulus playback were reported in awake cats (Chimoto et al., 2002; Evans et al., 1964) and monkeys (Brugge et al., 1973; Pfingst et al., 1981; Recanzone, 2000). However, in all these studies, the responses at the Offset of tone playback was simply anecdotally observed but was not thoroughly characterized or compared to Onset responses. Some early studies have concluded that one of the sources of Offset responses was simply post-inhibitory rebound following the onset excitation response (Phillips et al., 2002). For some time, Offset responses were not considered as its own neural response, and their role in encoding was set aside.

### 1.3.2 On and Offset Response Properties

It was not until more recently that auditory studies began to directly compare On and Offset responses. In the awake cat primary auditory cortex, a single neuron's On and Offset responses can be tuned to different frequencies (**Fig 1.3**) (Qin et al., 2007). Others have reported similar disparities in the response properties of an auditory cortical neuron early versus late after stimulus Onset or Offset in the macaque (Fishman et al., 2009; Tian et al., 2013).



**Figure 1.3** Sample neuron showing On and Offset responses tuned to different frequencies. Spectrotemporal spike activity for this neuron was obtained from tones played at 70 dB SPL, in awake cat primary auditory cortex (Qin et al., 2007).

With the distinct physiological response properties in Offset responses brought into focus, studies recently have begun to investigate the mechanism by which On and Offset responses arise. In ketamine-anesthetized rats, whole cell patch clamp recordings were conducted to show that distinct synaptic inputs give rise to On and Offset responses in the rat auditory cortex (Scholl et al., 2010). In their set of n=26 recorded neurons, they note that none of the Offset responses observed were due to post-

inhibitory rebound, but rather both On and Offset responses were driven by synaptic activation. In their recorded cells, the Offset response was tuned 1-2 octaves above the On response, and due to this difference in frequency tuning, the balance of synaptic excitation and inhibition comes from different sets of inputs. They showed that On responses can suppress subsequent On responses through depression of the synapses that contribute to the On response, but activation of the Offset response does not suppress On responses. This provides additional evidence that the Offset response is a distinct phenomenon and should be considered a sound-driven response that is separate from the On response.

Others have looked at how On and Offset responses change over the course of natural development. In the mouse auditory cortex, for young mice (15-23 days old), individual neurons' On and Offset receptive fields are largely overlapping, whereas adult mice (>60 days old) show divergence of the On and Offset receptive fields (Sollini et al., 2018). In addition, the arrangement of an individual neuron's On and Offset receptive fields can account for its frequency modulation directional sensitivity. These results further illustrate the importance of On and Offset responses in the auditory processing of temporal modulations. In the study of auditory cortical encoding of temporal modulation, separating the On and Offset response is important as the two responses come from distinct sources, can be tuned to different spectral regions, and thus have different response properties. Combining the two responses complicates the interpretation of results. For the work in this thesis, the On and Offset response components will be considered separately.

Most studies looking at On and Offset responses have primarily been looking within A1, which is the first node of auditory cortical processing that receives information directly from the auditory thalamus. However, the auditory cortex consists of multiple regions that play a role in auditory processing that have yet to be thoroughly explored.

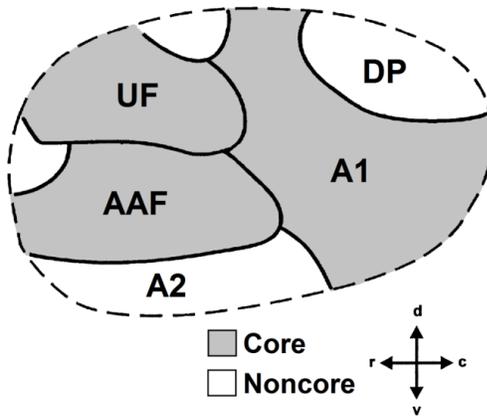
## **1.4 Core and Noncore Auditory Cortex**

Functional imaging studies in humans have supported the idea that the auditory cortex is organized into hierarchical regions that process increasingly complex aspects of sound. Generally, earlier areas in the core, such as primary auditory cortex, respond well to simple sounds such as pure tones. Regions moving outwards from the core auditory cortex, known as the belt or noncore auditory regions, are thought to process sounds with more complex spectrotemporal structure such as speech or music (Chevillet et al., 2011; Wessinger et al., 2001). In nonhuman primates, the noncore regions of the auditory cortex were found to respond better to bandpass noise or natural sounds compared to pure tones (Kaas et al., 1999; Kikuchi et al., 2010; Rauschecker et al., 1995). Additionally, differential frequency modulation preferences were found across the three macaque auditory anterolateral, mediolateral, and caudolateral belt areas (Tian et al., 2004), where anterolateral cortex responded best to slower frequency modulation rates associated with natural sounds. In cat auditory cortex, it was shown that across pure tones, noise, or natural vocalization stimuli, the latency of responses in the core are consistently shorter than in noncore areas (Carrasco et al., 2011). Compared to the core auditory cortex, far fewer studies have focused on noncore areas. Almost all literature cited in the previous sections discussing the different classes of auditory stimuli used in the study of the auditory cortex, as well as On and Offset responses looked within the primary auditory cortex. More studies in the noncore auditory areas are needed to better understand its role in auditory processing.

### ***1.4.1. Noncore Auditory Cortex Physiological Properties***

The noncore auditory regions have been reported to show distinct physiological response characteristics from the core. In the mouse, noncore auditory regions such as

secondary auditory cortex (A2) are also situated in the dorsal and ventral borders away from the center of the auditory cortex (**Fig 1.4**).



**Figure 1.4** Mouse Auditory Cortex. Core regions include primary auditory cortex (A1), anterior auditory field (AAF), and ultrasound field (UF). Noncore regions include dorsoposterior field (DP), and secondary auditory cortex (A2). Modified from (Stiebler et al., 1997). It is worth noting that the existence of UF is debated and some studies will instead refer to this region as the dorsomedial field (DM) (Tsukano et al., 2017).

Mouse A2 is typically responsive to a wide range of frequencies, and has been found not to be tonotopic using electrophysiological techniques (Guo et al., 2012; Stiebler et al., 1997). However, frequency gradients running from low to high in the dorsoventral direction in mouse A2 have also been reported using higher spatial resolution techniques such as intrinsic or calcium imaging (Issa et al., 2014; Kubota et al., 2008; Tsukano et al., 2015; Tsukano et al., 2016; Tsukano et al., 2017). This discrepancy is likely due to the lower spatial sampling density in electrophysiological recordings compared to imaging techniques.

In looking at multi-unit response properties, neurons in A2 tend to show later average response latencies, slightly higher spontaneous rates, higher best frequencies compared to A1 and anterior auditory field (AAF), and narrower Q40 frequency tuning bandwidths (Joachimsthaler et al., 2014). Additionally, fewer phasic responses are observed in A2, whereas there are very few tonic or phasic-tonic responses observed in

the core auditory regions. This has also been observed in gerbils (Schulze et al., 1997) using both tones and linear frequency modulated stimuli. A separate study using guinea pigs (Wallace et al., 2000) also reports the ventrorostral belt showing later response onset latencies and more sustained responses to pure tones. Within the ferret auditory cortex, the ventrorostral belt was observed to be nontonotopic and preferentially responded to broadband noise with increased response latencies (Bizley et al., 2005). All these studies observe primarily On-responses in the core auditory cortex, as opposed to prolonged or later response latencies in noncore auditory cortex. However, within the mouse, less work has been done on noncore representations of natural vocalizations, despite the observed responsiveness of the secondary cortex to speech or vocalization sounds in primates (Kaas et al., 1999; Kikuchi et al., 2010; Rauschecker et al., 1995).

### **1.5 Natural Vocalizations and Auditory Cortical Plasticity**

As further motivation for investigating natural vocalizations, sounds perceived in a social context, such as vocalizations, play a communicative role and often drive neural activity in ways that are distinct from non-social contexts (Bennur et al., 2013; Ehret, 2005). For example, in the mouse, natural ultrasonic vocalizations have been shown to be more intrinsically arousing than other types of non-vocal ultrasounds, thereby engaging limbic areas differentially (Geissler et al., 2002). Work in non-human primates has further shown that there is specialization in auditory processing areas for vocalizations. Within anterior regions of the temporal lobe in rhesus macaques (*Macaca mulatta*), neurons have been observed to be selectively responsive to conspecific vocalizations rather than heterospecific vocalizations or sounds that are not vocalizations (Perrodin et al., 2011). These "voice cells" were likened to "face cells" found in the fusiform gyrus, a visual processing area. These examples highlight the

importance of using vocalizations and studying social contexts in order to better understand normal auditory processing.

### **1.5.1 The Maternal Mouse Communication Model**

The mouse has become a good mammalian model to study the neurobiology of maternal sensory processing involving auditory cues that acquire behavioral relevance. When separated from the nest, mouse pups emit isolation calls, which are a class of ultrasonic vocalization at frequencies greater than ~50kHz (Liu et al., 2003). Mouse mothers naturally and reliably respond to pup isolation calls with phonotaxis, which involves approaching and retrieving the pup back to the nest. Female mice without any prior pup exposure do not conduct sound-guided retrieval (Ehret et al., 1984). Such females can learn to retrieve pups when housed as a co-carer with a mother and her pups and can begin to display phonotaxis to pup call playback within five days after pup birth (Ehret et al., 1987). Importantly, both mothers and recent co-carers will preferentially approach the sound of a pup isolation call as opposed to a behaviorally neutral sound such as a frequency matched pure tone (Ehret et al., 1987; Lin et al., 2013), demonstrating the efficacy of the vocalizations and not just other multimodal pup cues in eliciting maternal behavior.

### **1.5.2 Cortical Plasticity in the Maternal Model**

A large literature beginning with electroencephalography studies by Galambos and colleagues (Galambos, 1956) has demonstrated that neural plasticity within primary auditory cortex provides a detectable trace of a sound's learned behavioral relevance (Shepard et al., 2013; Weinberger, 2004). These studies have focused largely on making simple pure tones behaviorally relevant through laboratory conditioning and training paradigms. Generally, this has revealed a retuning of the excitatory receptive

field of core auditory cortical neurons, leading to a topographical expansion in the cortical area tuned to the newly relevant sound frequency (Recanzone et al., 1993).

Given this history of prior research, when considering the nature of maternal auditory cortical plasticity, cortical map expansion was first on our list to investigate as a trace of the acquired behavioral importance of these calls for mothers. Mouse strains typically feature an ultrasound field (UF) within the core auditory cortex that is specifically tuned to pure-tone frequencies above 50kHz (Stiebler et al., 1997). As pup USVs reside within the high frequency range, one might predict expansion of the tonotopic size of UF in maternal animals compared to non-maternal animals. However, interestingly, unlike the laboratory conditioning paradigms in which tones are paired with a reward or shock, no maternal expansion in the size of the UF itself was observed (Shepard, Lin, et al., 2015). The lack of map expansion in this natural context does not imply that such expansion cannot occur in other paradigms, such as developmental sound exposure (Han et al., 2007; Shepard, Liles, et al., 2015). However, this result does suggest that in realistic learning situations, auditory cortical map expansion per se need not be a long-term memory trace, even if it might have a function during auditory learning itself (Reed et al., 2011).

Even though map expansion is not the form that auditory cortical plasticity takes for the long-term memory of infant cries (as it has been shown to be absent on the order of one month after experience (Shepard, Lin, et al., 2015), despite mouse mothers still behaviorally responding to infant cries (Lin et al., 2013)), functionally relevant, long-term excitatory plasticity does emerge in core auditory cortex in other ways. For example, in the first demonstration of coding differences between maternal and non-maternal animals, multiunit auditory cortical responses in anesthetized mothers, but not pup-naïve virgins, were found to be robustly entrained by the ~5 Hz temporal rhythm of natural bouts of pup USVs (R. C. Liu et al., 2006). Moreover, the call-elicited firing rates of

multiunits themselves, irrespective of how the units' frequency tuning contributes to the tonotopic map, also appear to convey more information in mothers than in virgins for detecting pup USVs and discriminating between them (Liu et al., 2007).

Although informative, these early studies were conducted in anesthetized rather than awake mice, and utilized multiunit recordings, which is a somewhat coarse measure of neural activity. Multiunit recordings pool responses of many individual neurons, such that changes happening within only specific subpopulations could be obscured. To overcome this, methods of segregating neuronal subtypes could be applied. In particular, a computational model of auditory sensitivity to the amplitude envelope of a sound (Neubauer et al., 2008) can successfully carve out subsets of putative inhibitory interneurons and pyramidal neurons in auditory cortex whose spiking either encodes for the acoustics of a sound's onset, or not (Lin et al., 2010). In essence, these so-called well- and poorly-predicted neurons, respectively, have feedforward response latencies to pup USVs that either can or cannot be easily predicted from their pure tone responses.

Such a classification scheme has revealed a fairly specific form of excitatory plasticity for pup USVs within a subset of putative pyramidal neurons in core auditory cortex, recorded from passively listening, awake, head-fixed animals (Shepard, Lin, et al., 2015). The putative pyramidal neuron subset shows more delayed, and occasionally sustained, responses to USVs that are poorly-predicted by pure tone responses. The physiological characteristics of putative pyramidal neurons on average showed a longer peak-to-peak distance (thick-spiking cells) and late onset of response (Lin et al., 2010). Interestingly, in this subset, there was a significantly enhanced evoked firing rate response to the collection of pup compared to matched adult USVs in maternal animals, but not in non-maternal animals (Shepard, Lin, et al., 2015). This indicates an improved discrimination of acoustic features that separate pup from adult USVs in animals that

have learned the behavioral relevance of the vocalizations. However, the results of this study did not separate On or Offset responses, and looked exclusively within the core auditory cortex, and is unable to precisely pinpoint the acoustic features that maternal animals are using to discriminate the two sound categories. The work in this thesis aims to address each of these points, as well as forming a more generalized view of sound processing in the auditory cortex.

## **1.6 Summary and Significance**

Overall, researchers have approached the study of the auditory cortex from many different directions, probing responses using stimuli of varying acoustic parameters. In the study of ethologically or behaviorally relevant sound cues, utilizing sounds that more closely represent natural sounds may be an effective method of dissecting the encoding of sounds by the auditory cortex (Bennur et al., 2013). However, simply utilizing natural sounds without precisely describing their acoustic parameters sacrifices the quantitative accuracy and control over the acoustic stimulus parameters afforded by sounds such as pure tones or frequency modulation sweeps and gratings. Given the importance and prevalence of frequency modulation in natural sounds, we sought to utilize a parameterizable model of frequency modulation that could be fit to natural vocalizations, such that the frequency modulations present within natural sounds and neural tuning to those modulations could be quantitatively described.

Given the emergence of Offset responses as a potentially important mechanism by which frequency modulations are encoded (Sollini et al., 2018), we sought to systematically investigate On and Offset responses to parameterized variations in frequency modulation. Additionally, due to the preponderance of evidence showing late Onset or Offset responses present in the secondary auditory cortex (Bizley et al., 2005; Joachimsthaler et al., 2014; Schulze et al., 1997; Wallace et al., 2000), in this work I will

be comparing responses to frequency modulated tones in primary and secondary auditory cortex. Finally, animals are capable of learning the meaning of natural sounds, in which frequency modulation is a distinguishing factor between different sound categories. How the representation of frequency modulation across auditory cortex changes between animals that are either naïve to or have learned the meaning of a sound category will be addressed in this work.

### **Experience Dependent Coding of Intonations by Offsets in Mouse Auditory Cortex**

#### **2.1. Background**

Across many types of natural sounds, the frequency trajectory can be important to its meaning. In human speech, for example, pitch modulation conveys emotional and linguistic meaning, including in tonal languages (Bachorowski et al., 1995; Cutler et al., 1997; Lehiste, 1970). In many other species, including rodents, monkeys, dogs, birds, and even some marine animals, the temporal features of a vocalization including its frequency trajectory vary by context, and can express the intention as well as identity of the vocalizer (Brudzynski, 2007; Fischer et al., 1995; Klump et al., 1984; Taylor et al., 2009; Watwood et al., 2004). Indeed, intonation-sensitive areas have been reported in the human auditory cortex using subdural electrophysiology (Tang et al., 2017).

Furthermore, the ability to process and recognize pitch trajectory in vocalizations has been shown to be shaped by experience (Bijeljac-Babic et al., 2012; Karen Mattock et al., 2006; K. Mattock et al., 2008), and successful human adult learning of pitch modulation in words has been associated with activity in left as opposed to right auditory cortex (Wong et al., 2007). However, the temporal or spatial resolution of methods used in human studies such as fMRI, EEG, and ECoG are coarse compared to the rate at which neurons in the auditory system are responding to the sounds, which themselves may have rapid frequency modulation.

Previous studies of frequency trajectory encoding at the single neuron level utilized stimuli with linear or logarithmic unidirectional frequency modulation (Mendelson et al., 1985; Nelken et al., 2000), in which sweep direction selectivity and sweep velocity preference have been observed in primary auditory cortex. Others investigated periodic sinusoidal frequency modulation on the scale of seconds (Gaese et al., 1995; Whitfield

et al., 1965), in which phase locking to the modulation frequency is observed in primary auditory cortex. From these studies alone, it is unclear how their results would translate to how neurons would respond to natural sounds with more complex frequency modulation. A more recent study looked at both frequency modulation sweep responses and natural vocalization responses (Carruthers et al., 2013), demonstrating that primary auditory cortical preference for original vocalizations over temporally compressed, dilated, or reversed vocalizations can be explained with a combination of their best frequency (BF) as well as their responses to linear FM sweeps. However, in the study there was still a sizable population of neurons, about half, for which responses could not be explained using their model based on linear FM alone. This may not be surprising given most vocalizations contain more than linear FM.

Directional selectivity for linear frequency modulation within auditory cortical neurons is purported to arise from differences in receptive fields within a given neuron's On and Offset responses (Sollini et al., 2018). This result points to the importance of segregating evoked responses that occur at the On or Offset of sound playback, particularly when considering responses to frequency modulation. All these studies provide insight into frequency modulation encoding of linear or sinusoidal modulation alone in the primary auditory cortex (A1). Less is known about frequency modulation encoding in secondary auditory cortex (A2), and how experience modifies responses in both primary and secondary auditory regions. Natural vocalization frequency trajectories often contain more than just linear or sinusoidal modulation alone. By utilizing a model that combines the two, natural vocalizations can be better modelled.

In this study, we model mouse USVs using a parameterized sinusoidal frequency modulation (sFM) model. Mice can learn to recognize pup ultrasonic vocalizations (or USVs) after maternal experience, and will display preferential approach behavior to pup USVs over neutral sounds (Ehret, 2005; Ehret et al., 1987; Lin et al., 2013). We

conducted head-fixed, awake single unit electrophysiology and presented sFM with systematically varied parameters to probe plasticity in frequency trajectory sensitivity across natural experience and auditory region. Here, we find tuning to frequency modulation amplitudes less than  $1/8^{\text{th}}$  octave, a much finer degree of FM than uncovered in previous FM studies. We also observe an enhancement in the prevalence of Offset responses to natural USVs after maternal experience in A2. In maternal A2 units, a bias emerges in both On and Offset responses that favors vocalizations that have pup-like sFM parameters. This bias is furthermore explained by a shift in tuning of maternal A2 towards frequency trajectory parameter values that are more characteristic of the pup USV category. This work furthers our understanding of how auditory cortex attunes to features in acoustic space, demonstrating that offset tuning to frequency trajectory in secondary auditory cortex plays a role in natural sound category learning.

## **2.2 Materials and Methods**

### **2.2.1 Experimental Details**

#### *2.2.1.1 Experimental model - mouse*

Female wildtype CBA/CaJ mice (RRID:IMSR\_JAX:000654) between 12-18 weeks of age were used in this study. Mice had varying levels of pup experience, with the non-maternal animal group consisting of pup-naïve females, pup-naïve females that have only been passively exposed to pup USVs without social interaction, and females that had acted as cocarers with a mother and her litter but are at the time of post-weaning, during which they no longer display preferential pup USV phonotaxis (Lin et al., 2013). The maternal animal group consisted of post-weaning mothers (P21) as well as females acting as cocarers in a cage with a mother and her litter for 5-7 days. Animals were socially housed in single-sex cages until breeding age on a 10h light/ 10h dark reverse light cycle with ad libitum access to food and water; mice were moved to individual housing during experiments. Mothers included in the study had experience with only one litter before experiments. For mice in the maternal group, pup retrieval was conducted on postnatal days P5-7, in which pups were scattered in the home cage and animals were given 5 minutes to retrieve pups back to the nest. Pup scattering was repeated for a total of three times. Only animals that performed pup retrieval successfully were included in the maternal animal group. All animal procedures used in this study were approved by the Emory University Institutional Animal Care and Use Committee (IACUC).

#### *2.2.1.2 Sinusoidally Frequency Modulated (sFM) USV Modeling*

Mouse pup and adult USVs are complex, single-frequency whistles that have naturally variable frequency trajectories. In order to capture and parameterize the

various frequency trajectories of mouse USVs, we use a parameterized sinusoidally frequency modulated (sFM) model fit to the frequency trajectory of each call. The parameterized model contains a total of six parameters: duration ( $dur$ ), onset frequency ( $f_0$ ), sFM amplitude ( $A_{fm}$ ), sFM frequency ( $f_{fm}$ ), sFM phase ( $\varphi$ ), and linear FM slope ( $f_{slope}$ ). Parameters were fit minimizing mean squared error between the model and call, and the original amplitude modulation of each natural USV was used for each fit. Our entire call library of pup ( $n=62394$ ) and adult ( $n=11247$ ) USVs were fit to generate a call distribution in six-parameter space. For playback during neural recording, we used a curated natural USV stimulus set containing 18 pup USVs and 18 adult USVs that are matched for various acoustic dimensions including duration, onset frequency, and degree of frequency modulation (FM) at onset (Shepard, Lin, et al., 2015).

### *2.2.1.3 Sound Stimulus Playback*

USVs ( $n=36$ ) plus a silent trial were played with up to 50 trials per stimulus, randomly interleaved. For a subset of cells, to compare natural USV responses to sFM model fit USV responses, the original set of 36 USVs plus 36 sFM model versions of each USV were played back for 25 trials per stimulus, randomly interleaved.

For each neuron that showed an evoked response to USVs, we played an additional sFM stimulus optimized around the call that elicited the best response to assess tuning for sFM parameters. This stimulus contained sFM exemplars with  $f_0$ ,  $f_{slope}$ ,  $f_{fm}$ ,  $\varphi$ , and  $dur$  equal to the call eliciting the best response, with  $A_{fm}$  varied in 19 logarithmic steps across the range of natural  $A_{fm}$  values in USVs ( $A_{fm} = 180.78 * \exp(0.1051*n)$ ;  $n = [2:2:38]$ ), plus the original best-response call's  $A_{fm}$ , for a library of 20 different stimuli. A total of 30 trials per stimulus were collected. For this stimulus, if the unit's half-max spike rate was below its spontaneous firing rate, in which evoked responses were not significantly greater than spontaneous activity, units were excluded

from further analysis. Following playback of the  $A_{fm}$  tuning stimulus, a control stimulus with spectrally-matched noise to the 20  $A_{fm}$  tuning stimuli was played back and responses collected. Within the spectrally matched noise control stimulus, for each corresponding  $A_{fm}$  tuning stimulus ( $n=20$ ), 3 instances of randomly generated white noise, each with different random seeding, were first generated and then filtered based on the spectral content of the original  $A_{fm}$  stimulus, for a total of 60 stimuli. Each of the 60 stimuli were played for a total of 10 trials per stimulus, such that total presentation time approximately matches that of the experimental  $A_{fm}$  tuning stimulus.

For each individual unit, a frequency tuning curve was also obtained by playing a pure tone tuning stimulus (60ms duration, 6 sound intensities from 15 to 65dB SPL at 30 frequencies log-spaced 5-80kHz, repeated 15 times each, presented in pseudorandom order). For neurons that showed a best frequency, an additional sFM stimulus centered at the best frequency of the neuron was played. This stimulus had  $f_0$  equal to the best frequency, dur of 60ms,  $f_{slope}$  and  $\phi$  equal to 0,  $f_{fm}$  varied in logarithmic steps along natural USV ranges (15, 20, 28, 39, 53, 73, 100, 137 Hz), and  $A_{fm}$  varied as a fraction of best frequency in logarithmic steps following a logarithmic fit to  $A_{fm}$  natural USV distribution (fraction of BF: 0.004, 0.008, 0.016, 0.030, 0.055, 0.104, 0.194). Stimulus was played at a sound intensity level that elicited the best response at the neuron's best frequency. A total of 25 trials per stimulus were collected. For each of these stimuli, a corresponding bandwidth matched noise stimulus was played, with noise generated realtime via RpvdsEx software (TDT).

For all stimuli described, trial length was 600ms, with stimulus playback beginning at 200ms.

#### *2.2.1.4 Single Unit Awake Electrophysiology*

At 12-18 weeks of age, mice were moved to individual housing and headpost attachment with small hole craniotomy surgery was conducted (Shepard, Lin, et al., 2015). Briefly, animals were anesthetized with isoflurane (2-5%, delivered with oxygen) and buprenorphine (0.1mg/kg) was administered as an analgesic. Animals underwent aseptic surgery to stereotaxically define a recording grid over the left auditory cortex, as the left auditory cortex is putatively associated with vocalization processing, particularly in the maternal paradigm (Geissler et al., 2004). The skull is exposed, and the left temporal muscle is deflected to permit access to the bone overlying auditory cortex. Using sterile tattoo ink applied to a stiff wire mounted on a stereotaxic manipulator, a grid of dots designating drill points over the auditory cortex is made. Dots ~100 $\mu$ m in diameter were drawn on the skull in three rows (1.5, 2.0, and 2.5mm below bregma) and five columns (spanning 50-90% of the distance between bregma and lambda, in 10% steps). Following recording grid placement, dental cement was used to secure an inverted flat-head screw on the midline equidistant from bregma and lambda. The animal is recovered in the home cage upon a heating pad and is administered saline subcutaneously for fluid replacement.

The day before a recording, the animal is reanesthetized with isoflurane and holes (~150 $\mu$ m in diameter) are hand drilled on one or more grid points with an insect needle held by a pin vise, by manually holding the insect needle and rotating the needle in a drilling motion to create a hole. A hole for a ground wire is also drilled in the left frontal cortex. On the following day, the animal is allowed to recover, and is acclimated to a foam-lined cylindrical (~3cm diameter) restraint device, where the animal is handled and placed into the device for 15 minutes before being returned to the home cage. On the day of recording, the animal is handled for 10 minutes and then placed in the same restraint device, which secures the body while leaving the head exposed. The implanted head post is secured to a post mounted on a vibration-isolation table, while the capsule

containing the body is suspended from rubber bands to reduce torque force on the headpost. Recordings typically lasted 2-4h, and excessive movement or signs of stress signaled the end of an experiment.

Electrophysiological activity in the auditory cortex is recorded (sample rate 24414.0625/s, at which single units can be detected) with single 6 M $\Omega$  tungsten electrodes (FHC), filtered at >300Hz and <3 or 6kHz. Using a hydraulic Microdrive (FHC), the electrode is driven orthogonally into auditory cortex to an initial depth of ~200 $\mu$ m. The electrode is then advanced in 5 $\mu$ m steps until an SU is detected. SU isolation is based on the absence of spikes during the absolute refractory period (1ms), and on cluster analysis of various spike features (e.g., first vs second peak amplitudes, peak-peak times). In some cases, multiple SUs are recorded at one location and could be extracted by clustering based on spike features.

## **2.2.2 Data Analysis**

### *2.2.2.1 Analysis of $A_{fm}$ and $f_{fm}$ Tuning around Best Frequency*

For analysis of whether units show preference for sFM stimuli over pure tone, an 8x8  $A_{fm}$  and  $f_{fm}$  tuning stimulus was played back to neurons as described in sound stimulus playback. Data used in this analysis included n=48 units recorded from n=24 animals. Unit responses were divided into the On response (during 60ms window of stimulus playback) and the Offset response (window defined starting after stimulus playback ends until the end of the raster response). Out of n=48 units, n=41 exhibited an On response and n=40 exhibited an Offset response. Absolute spike rates were calculated over the On and Offset window for each stimulus to generate a response heat map. To determine whether response to a specific combination of  $A_{fm}$  and  $f_{fm}$  significantly differed from pure tone, the spike rates from all pure tone trials were compared to spike rates from all trials with the corresponding parameter values using the nonparametric

Wilcoxon Rank Sum test. To account for multiple testing, the Bonferroni correction was applied, with the alpha level taken as  $\alpha < (0.05/64)$  or 0.00078. The unit was considered to prefer sFM over pure tone if there is at least one set of sFM parameters (where  $A_{fm}$  is non-zero) with significantly greater response than pure tone for either its On or Offset response.

#### *2.2.2.2 Pure Tone Integration Model for Prediction of $A_{fm}$ Stimulus Spike Rates*

In order to predict the  $A_{fm}$  stimulus spike rates from a unit's pure tone tuning curve, the power spectral density of each of the  $A_{fm}$  stimuli were calculated. The amount of power in each frequency bin of the power spectrum was multiplied by corresponding frequency bin's spike rate in the unit's pure tone tuning curve, normalized such that the evoked response from the best frequency pure tone stimulus for both the  $A_{fm}$  tuning stimulus and the pure tone tuning stimulus are matching. As the number of frequencies played back during pure tone tuning is fewer than the number of frequency bins in the power spectrum, the pure tone tuning curve was interpolated linearly across missing frequency bins.

#### *2.2.2.3 sFM Matched Noise and Linear FM Analysis*

For comparison of responses to sFM and to matched bandwidth noise stimuli, responses were calculated during the On, Offset, and the entire evoked response window, and normalized by subtracting the spontaneous rate. A nonparametric paired Wilcoxon Signed Rank Test was performed comparing the sFM and paired noise normalized spike rate. In this analysis, a total of  $n=280$  paired sFM and noise stimuli were played to  $n=40$  units in  $n=21$  animals. For comparing sFM to linear FM, we also calculated the responses during the On, Offset, and entire evoked response time windows, normalizing by subtracting the spontaneous rate, followed by performing

paired Wilcoxon Signed Rank Test. For the linear FM control, a total of n=35 paired sFM and IFM stimuli were played to n=20 units from n=13 animals. All these units showed an excited response to the stimuli so they were considered for analysis.

#### *2.2.2.4 Spike Density Plot*

A spike density plot was generated to visualize the overall population spiking activity across n=36 natural mouse USVs, sorted by increasing USV duration. Plots are generated using a modified version of the scatterplot density MATLAB function `dscatter.m`, freely available online (Eilers et al., 2004). Default smoothing settings were used, in which smoothing is conducted over 20 bins, with the time axis divided into 0.1ms bins (6000 bins across a 600ms period). For visualization purposes, the y axis divided into 100 bins per stimulus (3600 bins for 36 stimuli), and randomized jitter [0-1 bins] was applied to each individual spike along the y axis. A total of n=136 units are included, pooling across region and animal group.

#### *2.2.2.5 On and Offset Prevalence Analysis*

For analysis of prevalence of on and offset responses to pup USVs, the shortest calls in the stimulus library on the order of 12-15ms were excluded from analysis as On and Offset responses to these calls could not be distinguished and may be overlapping. On and Offset responses were classified by two independent investigators (KC and DK). Data used in this analysis comes from n=3264 calls (n=971 show USV-excited responses) played to n=136 units (USV-excited) from n=55 animals. Units were classified as exhibiting On or Offset responses, in which a response above spontaneous *during* call playback for any of 36 USVs would signify the unit has an On response, and a response *after* call playback has ended for any of 36 USVs would signify the unit has an Offset response. On and Offset responses did not necessarily have to be from the

same call within the unit for the unit to be classified as having an On / Offset response. Separately, responses were also classified on a per call basis, in which only trials associated with a single USV were included for classification of On / Offset response, and was repeated for all 36 USVs played to each SU. Again, responses were classified by two independent investigators (KC and DK).

#### 2.2.2.6 Generalized Linear Modeling

With the ability to describe the frequency trajectory of our USV library with six sFM parameters, we are interested in which of these features are encoded by the On and Offset responses to calls. Generalized linear modeling, allows us to see which sFM parameters significantly explain the On and Offset response spike rates. To model spike data, we propose to utilize GLMs with a Poisson distribution (Dayan et al., 2001) and a Log link function. The spike rate response to each call was calculated as the firing rate within the corresponding On or Offset time window, normalized by subtracting the spontaneous firing rate. Due to the nature of the Poisson distribution, negative firing rates (which are firing rates below the spontaneous rate) was removed from analysis. In addition, due to the difficulty of distinguishing On and Offset responses in short calls (<30ms), short calls were excluded from this analysis. The six parameters in the sFM fit to the call will be used as the individual model parameters. To test and account for overdispersion (as the Poisson distribution describing the distribution of spike likelihood may not necessarily be equal to 1), the overdispersion test can be conducted and accounted for within the standard error as well as p-values (if dispersion >1). Correction is conducted by multiplying the standard error by a factor  $\sigma$ , defined as  $\sigma = \sqrt{\frac{Pearson \chi^2}{DF}}$ . Additionally, false discovery rate (FDR) correction was applied to the parameter estimate p-values. An alpha level  $\alpha < 0.05$  was

taken as significance for corrected p-values. GLM analysis was performed in R (<https://www.r-project.org/>) or in JMP Pro 13 (SAS).

#### *2.2.2.7 Logistic Regression Modeling for Classification of sFM Parameter Combinations as Pup-like vs Adult-like*

A nominal logistic regression was performed using our library of 10,353 adult and 57,989 pup ultrasonic calls, where calls shorter than 4ms duration with minimum frequencies less than 45kHz were excluded. A nominal logistic regression model was fit using the six sFM parameters (A, F, Phi, b, m, dur) to best predict “pup” (pup-like) and “adult” (adult-like) labels. The model followed the format:  $\text{Score} = \text{logit}(\beta_0 + \beta_1 A_{fm} + \beta_2 f_{fm} + \beta_3 \varphi + \beta_4 f_0 + \beta_5 f_{\text{slope}} + \beta_6 \text{dur})$ , where each of six beta coefficients were fit to maximize prediction accuracy. An ROC curve was constructed using the resulting model, and a threshold was selected that maximizes the sensitivity and minimizes 1-specificity. Significance of the ROC was assessed using a bootstrap analysis (N=1000). All analyses were performed using JMP Pro 13 (SAS).

#### *2.2.2.8 $A_{fm}$ Tuning Analysis*

As  $A_{fm}$  is one of the parameters capable of separating the pup and adult USV categories, we sought to investigate whether tuning in this parameter space changes as a result of learning the meaning of pup USVs. Tuning curves for  $A_{fm}$  generated by the stimulus for which 19 steps of  $A_{fm}$  were played back with the remaining sFM parameters centered around the values of a USV that elicits the best response from the unit (from a survey of n=36 USVs in a curated stimulus set). If a unit shows responses to both pup and adult USVs, two  $A_{fm}$  tuning stimuli were presented, with one centered around the best pup USV and the other around the best adult USV. Data in this analysis comes

from n=25 calls played to n=18 units from n=15 animals. For this analysis, only the Offset response was considered, as very few calls elicited On responses from these units (n=5/25 of calls played to n=4/18 units from n=3/15 animals showed On responses, all of which were from core auditory cortex). Responses with a half max spike rate below its spontaneous rate were excluded due to responses not being significantly greater than spontaneous activity. To determine the best  $A_{fm}$  frequency and the tuning bandwidth, a Gaussian fit ( $a \cdot e^{-\left(\frac{x-b}{c}\right)^2} + d$ ) was placed over the spike rates to determine the peak (Gaussian mean,  $b$ ) and bandwidth (Gaussian standard deviation or std,  $c$ ) of the unit's  $A_{fm}$  tuning. Fit parameter initialization [lower bound, upper bound] were as follows:  $a = 1$  [0,Infinity],  $b = A_{fm}(\text{max spike rate})$  [0,10000],  $c = A_{fm}$  bandwidth (halfmax to halfmax width) [0,10000],  $d = 0$  [0,Inf]. Reported best  $A_{fm}$  values are obtained from the Gaussian fit tuning curves. Note that when using the  $A_{fm}$  value without a Gaussian fit, the results remain the same as reported. Results were divided by animal experience group (maternal or nonmaternal) as well as auditory cortical region (Core or A2), and group comparisons were conducted with a nonparametric Wilcoxon Kruskal-Wallis Test followed by the post-hoc Tukey Kramer HSD method. A  $p < 0.05$  was taken as significant. Note when conducting analysis by call, by unit, or by animal, results remain significant.

## 2.3 Results

### 2.3.1 SUs can show tuning to $A_{fm}$

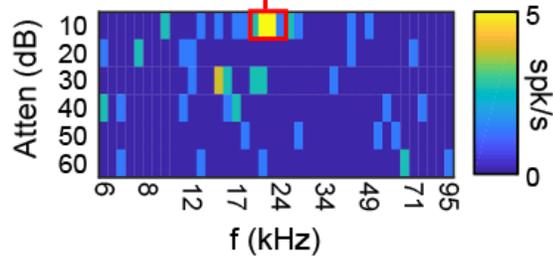
In order to better understand frequency trajectory encoding in the auditory cortex, we sought to quantitatively describe frequency trajectories of natural sounds utilizing a sinusoidally frequency-modulated (sFM) model with six parameters:  $A_{fm}$  (Amplitude of frequency modulation, Hz),  $f_{fm}$  (Frequency of frequency modulation, Hz),  $\phi$  (Phase, radians),  $f_o$  (Onset frequency, Hz),  $f_{slope}$  (Slope of linear frequency modulation component, Hz/sec), and  $dur$  (duration, ms) (**Fig 2.1A**). Using this model, we assessed the response of isolated single units to variation in one of the sFM parameters,  $A_{fm}$ . We first determined a unit's best frequency (BF) by playing pure tone tuning stimuli (**Fig 2.1B**), and then starting from the unit's BF, introduced small amounts of frequency modulation in the range of 0 – 1 octave centered linearly around the BF. All of the example unit's responses to each of the 60ms sFM stimuli were well above spontaneous activity, including the pure tone response. In addition, we find that the unit exhibits tuning in  $A_{fm}$ , and responds best to a fine frequency modulation (1/10<sup>th</sup> octave) compared to a pure tone or larger frequency modulations (**Fig 2.1C**). Using a pure tone integration model that attempts to predict the spike rate to the  $A_{fm}$  tuning stimulus based on the unit's pure tone response, we find that the responses are significantly above what would be predicted by pure tones alone. This unit was also tested with a longer duration pure tone of 120ms, and we found that the responses are slightly delayed and are still significantly above the pure tone integration model. However, these results do not eliminate the possibility of static inhibitory sidebands causing the results we see. To account for this possibility, we introduced variation in a parameter in which does not change the overall bandwidth of the sound, in this case the frequency of the frequency modulation,  $f_{fm}$ .

# Example Single Unit (SU 2981)

A)

## Pure Tone (PT) Tuning

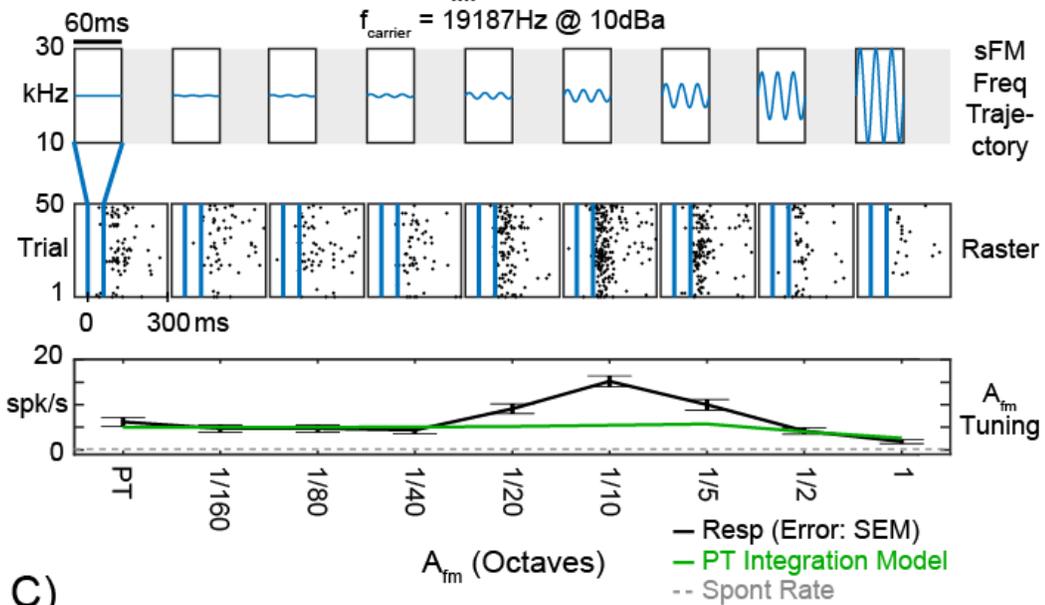
BF: 19187Hz



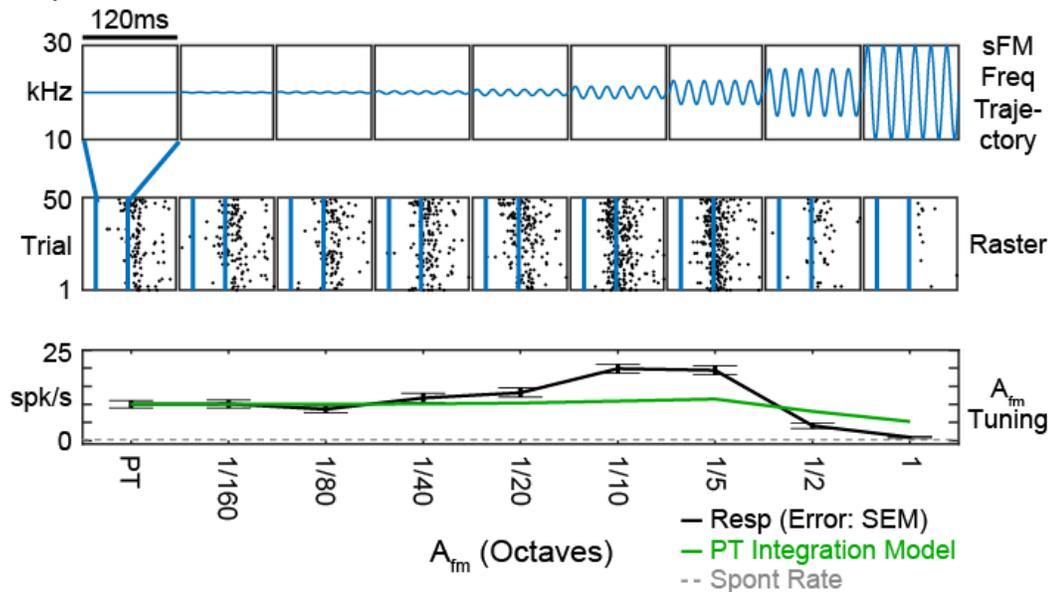
B)

## $A_{fm}$ Tuning

$f_{carrier} = 19187\text{Hz} @ 10\text{dBa}$



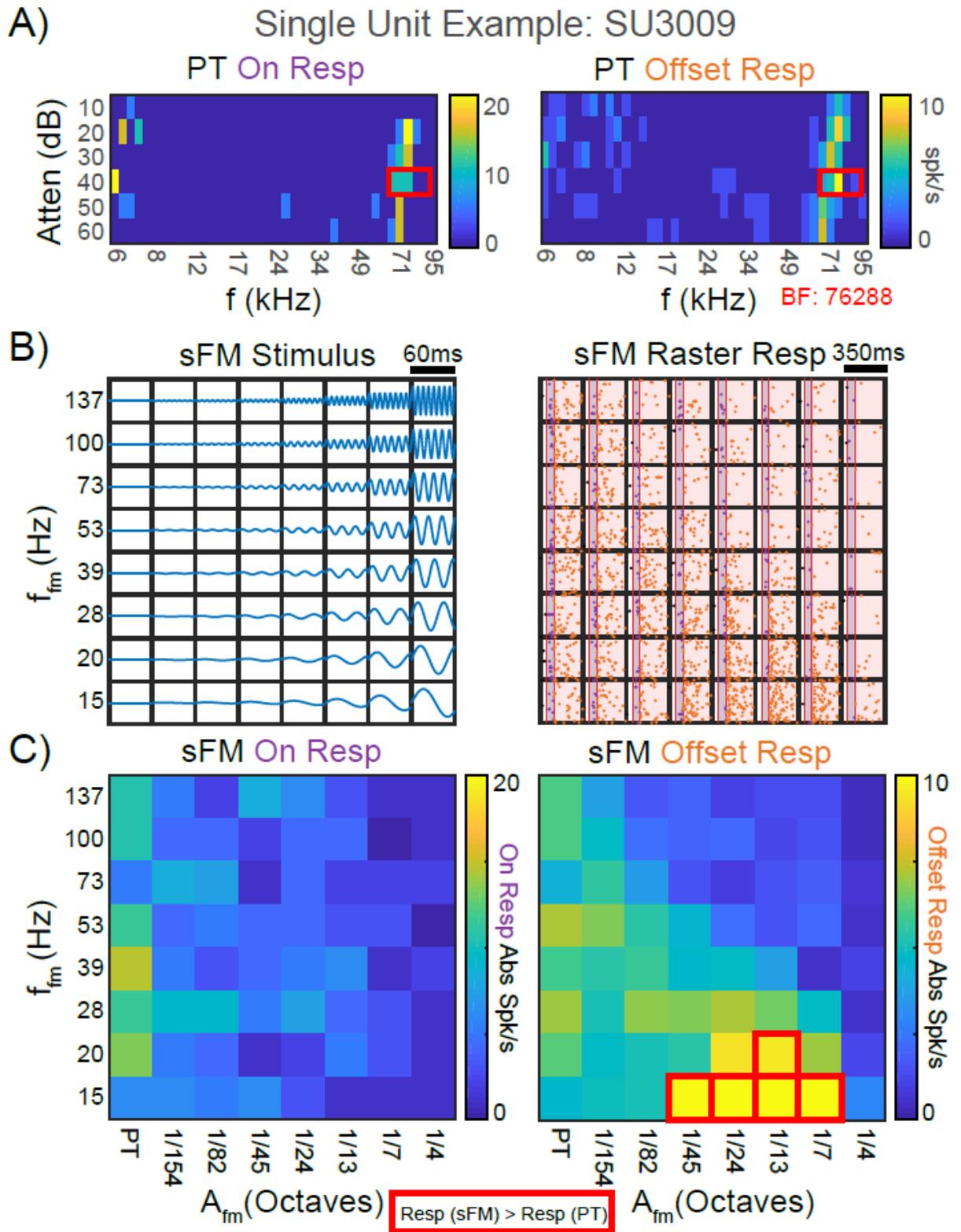
C)



**Figure 2.1. sFM model and sample single unit  $A_{fm}$  tuning.** **A)** sFM model with six parameters, depicting formula and each of the model's six parameters in red. **B)** Pure tone tuning response area across six decibel levels of attenuation (dBa) and 40 logarithmic frequency steps spanning 6 – 96kHz, with 60ms tones. The unit's absolute spike rate is represented by the heatmap color, with hotter colors representing higher spike rate. The neuron's best responding area is indicated in red, with a best frequency (BF) of 19187 Hz. **C)**  $A_{fm}$  tuning stimulus using the unit's BF as the carrier frequency of the  $A_{fm}$  signal. **C: Top row:** Cartoon stimulus spectrograms, with all other parameters fixed at:  $f_{fm} = 50\text{Hz}$ ,  $\varphi = 0$ ,  $f_0 = \text{BF}$  (19187Hz),  $f_{\text{slope}} = 0 \text{ Hz/s}$ ,  $\text{dur} = 60\text{ms}$ .  $A_{fm}$  is varied in logarithmic steps from 0 to 1 octave. **C: Middle row:** Rasters, with stimulus length depicted by the vertical blue bars, and individual spikes depicted with black dots. A total of fifty trials for each stimulus were presented. **C: Bottom row:** Response of the neuron in black with error bars representing standard error of the mean. Spontaneous rate is represented by the dotted gray line. A pure tone (PT) integration model that predicts the spike rate response to the  $A_{fm}$  tuning stimulus using the unit's pure tone tuning curve is depicted in green. **D)** Similar plot to C, instead with the  $A_{fm}$  tuning stimulus using a duration of 120ms, and all other parameters identical to C. The pure tone (left most) stimulus in this shows that latency of neural responses can shift to follow the offset of stimulus playback. The stimuli with more modulation moving rightwards show that units can also spike earlier, as if integrating a sufficient degree of the frequency trajectory is can elicit a response.

### **2.3.2 SUs can show tuning to $A_{fm}$ and $f_{fm}$ specifically in their Offset response**

We looked at tuning to sFM while varying two sFM parameters,  $A_{fm}$  and  $f_{fm}$ . In a sample single unit, after acquiring a pure tone tuning response to determine its BF (**Fig 2.2A**), we then presented a stimulus that varied  $A_{fm}$  and  $F_{fm}$  in 8 steps each around the BF at the sound level that elicited the best response from the unit. Parameters were chosen such that  $f_{fm}$  ranged from 15Hz – 137Hz, and  $A_{fm}$  ranged from 0 – 1/4<sup>th</sup> octaves; sound duration was fixed at 60ms, with an 8x8 stimulus grid (**Fig 2.2B**). In this different example unit, we found responses both during sound playback and after sound playback ended, which were separately analyzed as the On and Offset response respectively. We found that this unit's On response was strongest for 0  $A_{fm}$  (pure tone). However, the unit's Offset response preferred modulation by small, nonzero amounts of  $A_{fm}$  at low  $f_{fm}$  values (**Fig 2.2C**). We additionally sampled a larger population of neurons that exhibit On or Offset responses to assess their sFM responses as a whole.



**Figure 2.2.  $A_{fm}$  and  $f_{fm}$  tuning in a sample single unit. A)** Pure tone tuning response areas for the On response (left) and Offset response (right) of sample unit SU3009. The

unit's absolute spike rate is represented by the heatmap color, and the neuron's best responding area for their Offset response is indicated in red with BF = 76288 Hz. **B)** Left: Cartoon stimulus spectrograms of the  $A_{fm} \times f_{fm}$  tuning stimulus, each at 60ms duration. Right: Raster responses of the single unit, with the purple shaded area representing the On response window, and the orange shaded area representing the Offset response window. Dots represent individual spikes. **C)** Heatmap representing the spike rate of the unit within the On response window (left) and Offset response window (right). Red boxes indicate stimuli that had a significantly higher spike rate compared to the pure tone (PT) spike rate ( $p < 0.05$ , Wilcoxon Rank Sum Bonferroni Corrected).

### ***2.3.3 SU Offset responses are more likely to prefer sFM stimuli over pure tone compared to On responses***

We assessed On and Offset response sFM tuning around BF across a population of  $n=48$  neurons, of which  $n=41$  exhibited an On response, and  $n=40$  exhibited an Offset response. There was variability in how units responded to the  $8 \times 8$  sFM stimulus in both On and Offset responses, with some showing no apparent pattern, while others showing preference for different parts of  $A_{fm}$  and  $f_{fm}$  stimulus space (**Fig 2.3A**). Comparing units overall, the Offset response is more likely than the On response to show preference for sFM over pure tone, although a small subset of On responses can also show preference for sFM over pure tone (Bonferroni-corrected Wilcoxon Rank Sum  $p < 0.05$ , **Fig 2.3B**).

To demonstrate that sFM responses are sensitive to the frequency trajectory as it unfolds over time, rather than simply the spectral power contained in the stimulus, we recorded responses to separate stimuli containing narrowband noise that is spectrally matched to the  $8 \times 8$   $A_{fm}$  and  $f_{fm}$  tuning stimulus. Units respond preferentially to sFM stimuli compared to matched bandwidth noise stimuli ( $p < 0.01$  Paired Wilcoxon) when considering the Offset response (**Fig 2.3C**). This data includes  $n=280$  sFM – noise pairs presented to  $n=40$  units in  $n=21$  animals. This preference is also significant when the On response is considered. Note that these results hold regardless of what type of spike

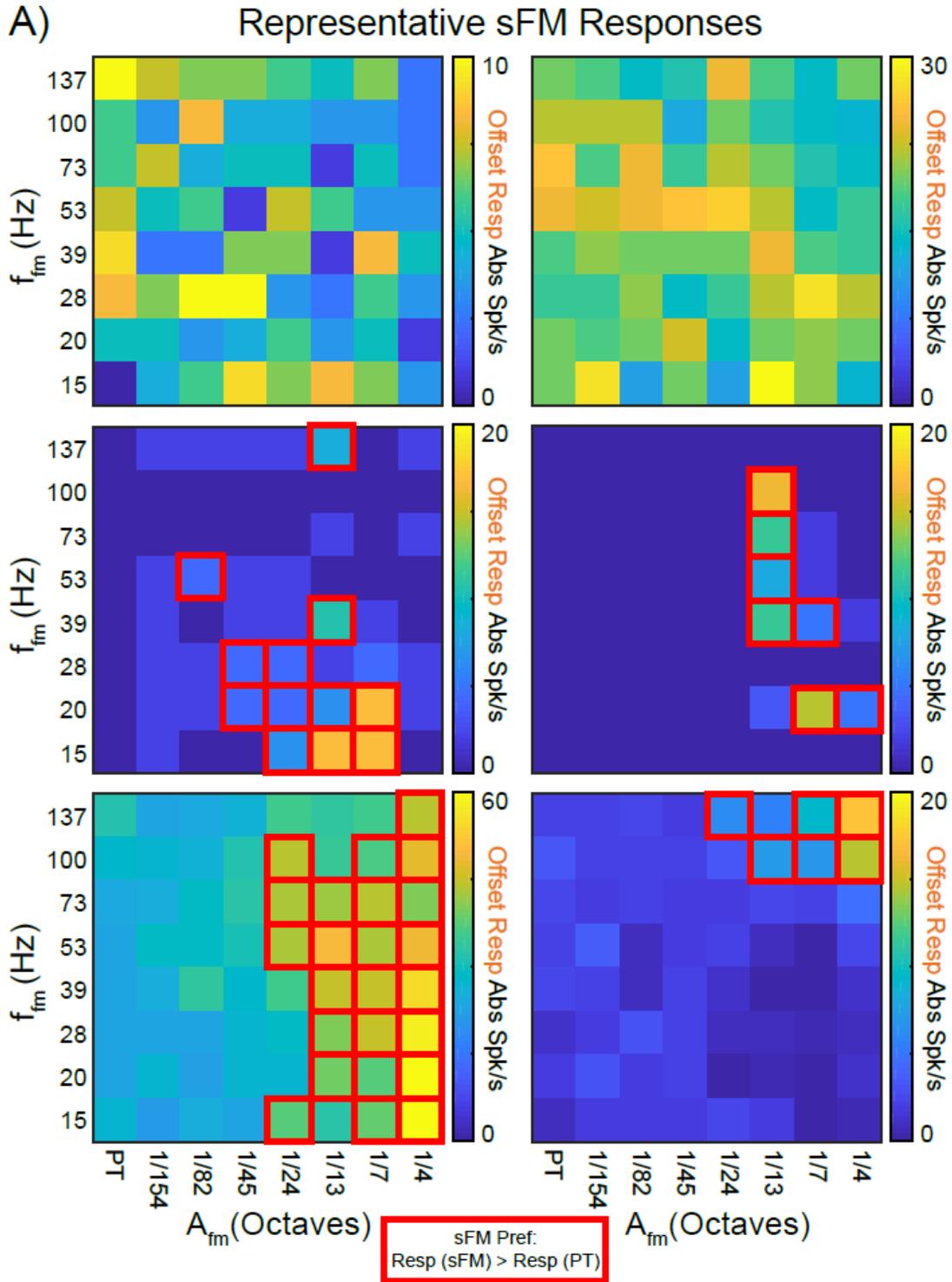
rate calculation is done (absolute spike rate, normalizing by subtracting the spontaneous rate, or normalizing by dividing the spontaneous rate).

As an additional control to ensure that responses to sFM were specifically to the frequency trajectory of the sound, linear frequency modulated (IFM) stimuli containing the same frequencies as the sFM stimuli were also played to a subset of units. This control addresses the possibility that responses to matched bandwidth noise may be suppressed due to the noise simultaneously hitting a unit's inhibitory sideband, which may decrease the unit's response strength, as has been demonstrated in the past using two-tone stimuli (Young et al., 1976). Presumably, sFM stimuli with matched spectral power would also be hitting these sidebands, although it would not be doing so simultaneously as would be occurring in the noise stimulus. As a result, a pure tone with IFM can serve as a secondary control for trajectory rather than just spectral sensitivity. When comparing Offset responses to sFM vs IFM, we find that sFM stimuli were preferred over matched IFM stimuli (pairwise signed rank Wilcoxon  $p < 0.01$ , **Fig 2.3C**). Note that On responses to sFM vs IFM are not significantly different (not shown).

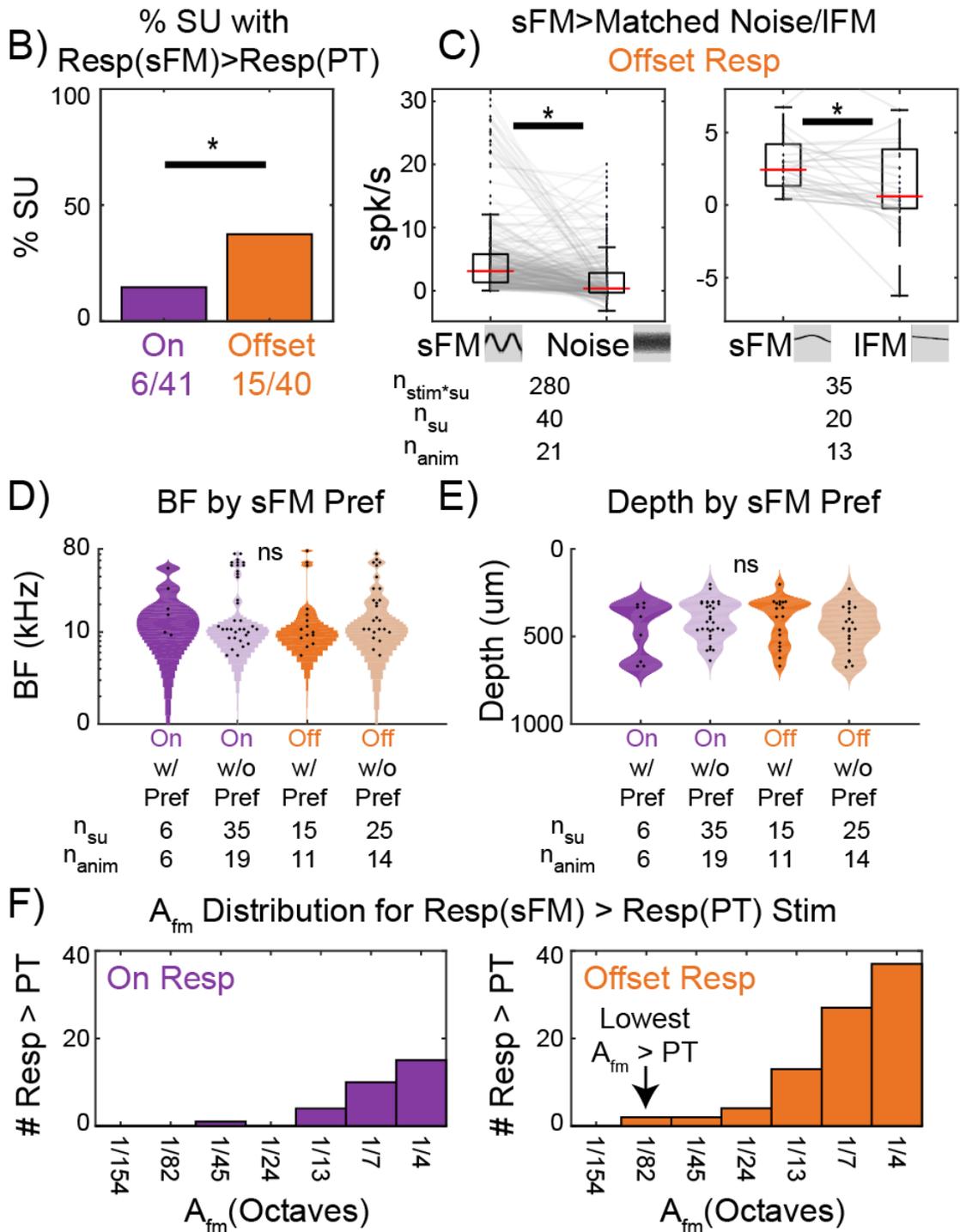
The distribution of best frequencies between units that either show or do not show sFM preference in their On or Offset responses does not significantly differ (**Fig 2.3D**). Units can show preference for sFM across the entire mouse hearing range from 6-80kHz. Similarly, the distribution of cortical depths between units with or without sFM preference does not differ (**Fig 2.3E**). Units showing preference for sFM span the entire range of cortical depths that were sampled (200-700 $\mu$ m).

These results demonstrate that auditory cortical neurons can show tuning to frequency trajectory parameters across the mouse's hearing frequency, and that the responses cannot be explained by static spectral tuning. These types of frequency trajectories are prevalent in natural vocalizations, and sensitivity to these trajectories may change as the meanings of vocalizations are learned. To investigate frequency

trajectory responses in natural sounds, we focus on a natural vocalization learning paradigm using mouse ultrasonic vocalizations (USVs).



## Single Unit Population Data



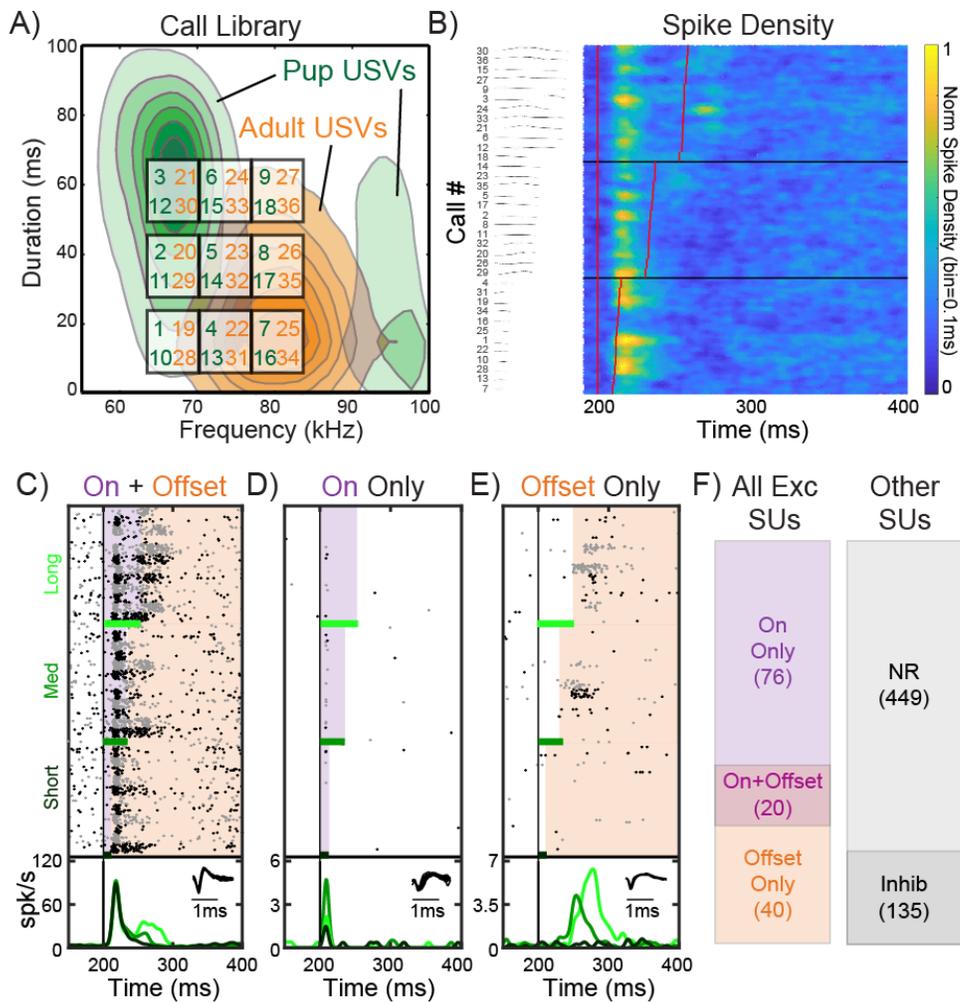
**Figure 2.3. Population sFM Tuning Data. A)** Offset response  $A_{\text{fm}} \times f_{\text{fm}}$  heatmaps across six different sample units. Red boxes indicate stimuli that had a significantly higher spike rate compared to the pure tone (PT) spike rate ( $p < 0.05$ , Wilcoxon Rank Sum Bonferroni Corrected). **B)** Proportion of units with an On response that prefer sFM

stimuli over pure tone compared to Offset responses that prefer sFM stimuli over pure tone ( $p < 0.05$ , Wilcoxon Rank Sum). **C**) Control stimulus set depicting comparison of sFM stimuli responses to matched bandwidth noise stimuli (left) as well as matched linear FM stimuli (right). Offset responses are depicted (\*  $p < 0.05$ , Wilcoxon Signed Rank paired test). **D**) Distribution of best frequencies (BFs) of units that either had a preference for sFM over pure tone or not in their On or Offset response. No differences are seen overall in BFs across group (NS, Kruskal Wallis). **E**) Distribution of unit recording depths of units. No differences are seen in unit depth (NS, Kruskal Wallis). **F**) Distribution of  $A_{fm}$  values for stimuli that elicited a significantly greater response than pure tone for both On responses (left) and Offset responses (right), based on statistical test done in (A).

### **2.3.4 SUs show Heterogeneity in timing of responses to USVs**

Mice can naturally learn the behavioral relevance of pup USVs with maternal experience (Ehret et al., 1987; Lin et al., 2013). We recorded responses to a library of  $n=36$  curated pup and adult USVs with matched onset frequency and duration properties (**Fig 2.4A**). Auditory cortical population-averaged responses to such natural mouse USVs are shown in **Fig 2.4B**, pooling across region and maternal experience. Maternal animals included mothers and co-caring females co-housed with a mother and her dam, while nonmaternal animals included pup-naïve females and females that have been passively exposed to the sound of pup calls without social interaction with pups (Yoked mice, see **Appendix C**). A total of  $n=36$  USVs of varying frequency trajectories were presented to units across region and experience,  $n=136$  of which were call-excited, and their pooled spike density plot was visualized (**Fig 2.4B**). The presence of Offset responses is most pronounced for the longest USV durations, while the On and Offset responses were difficult to separate for USVs with the shortest duration. As a result, for all subsequent analysis of On and Offset responses, the shortest group of USVs was excluded. Individual units can also respond either at the On and/or Offset of USV playback. Units that were call excited were further sub-classified into three categories: 1) On+Offset responding, in which at least one of 36 calls played elicited an On response, and at least one call elicited an Offset response, as rated by two independent investigators (**Fig 2.4C**); 2) On only, in which at least one of 36 calls played elicited an

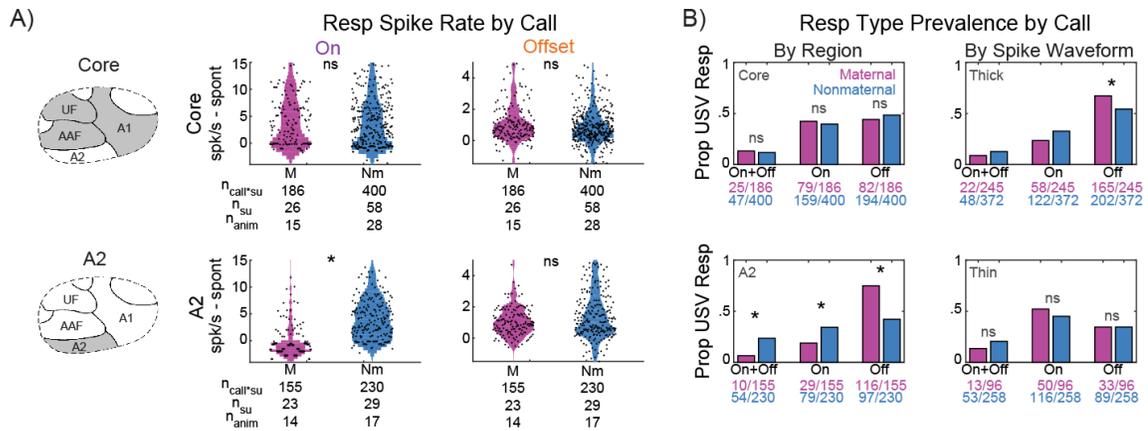
On response while none of the calls elicited an Offset response (**Fig 2.4D**); and 3) Offset only, in which at least one of 36 calls played elicited an Offset response while none of the calls elicited an On response (**Fig 2.4E**). All n=136 units were classified into one of these three groups, and we found most units show only an On response (n=76), while there was a subset that shows both (n=20) or only Offset responses (n=40). These are pooled numbers of call-responsive cells found across all auditory fields (A1, AAF, UF, A2). To reduce selection bias against neurons with lower spontaneous firing rates or highly selective responses, recordings were completed on all units that could be isolated. A fraction of neurons were not responsive to any USVs (NR n=449), while some were inhibited by USVs (Inhib n=135) (**Fig 2.4F**).



**Figure 2.4. USV On and Offset Responses across units. A)** Mouse USV playback library of a curated set of n=36 calls overlaid on a probability distribution of pup (green) and adult (orange) calls. **B)** Spike density plot depicting overall population responses to USVs across all call-excited single units recorded (n=136, n=55 animals), sorted by increasing USV length. USV length is represented by the vertical red line, with playback starting at 200ms. Call spectrograms are depicted on the left. **C)** Representative single unit showing both On and Offset responses. Rasters alternate between black and gray to delineate when the call # changes. Purple shaded area represents the On response window, while the Orange shaded area represents the Offset response window. A Peristimulus Time Histogram (PSTH) pooling responses for the corresponding call length is depicted under the rasters, with an inset representing the spike waveform. **D)** Similar to B, a representative single unit showing only an On response. **E)** Similar to B, a representative single unit showing only an Offset response. **F)** Overall distribution of recorded single units and their response characteristics. Units that did not respond (NR, non-responsive, or Inhibited) were not included in subsequent analyses.

### **2.3.5 Strength of On Responses to USVs decreases with experience**

We then assessed the spike rate of On and Offset responses across auditory region (Core: A1, AAF, Ultrasound Field UF, versus A2) and across maternal experience. Spike rates were calculated on a per-call basis and normalized by subtracting the unit's spontaneous firing rate. We found that in Core, no changes across experience are seen in either On or Offset spike rates (**Fig 2.5A**, top). In A2, maternal animals show a significantly decreased On response firing rate ( $p < 0.0001$ , Bonferroni corrected Wilcoxon Rank Sum: Maternal n=155, Mean $\pm$ SE = 0.45 $\pm$ 0.69 spk/s; Nonmaternal n=230, Mean $\pm$ SE = 7.61 $\pm$ 0.56 spk/s), while no changes are seen in evoked Offset spike rates (**Fig 2.5A**, bottom). When analyses are performed on a per-unit basis rather than a per-call basis, A2 still shows significantly decreased On response spike rate in maternal animals ( $p < 0.01$ , Wilcoxon Rank Sum; Maternal n=23, Mean $\pm$ SE= 0.07 $\pm$ 1.04; Nonmaternal n=29, Mean $\pm$ SE= 3.02 $\pm$ 0.92). On a per-animal basis, the comparison is trending (NS  $p = 0.14$ , Wilcoxon Rank Sum; Maternal n=8, Mean $\pm$ SE=0.48 $\pm$ 1.64; Nonmaternal n=5, Mean $\pm$ SE=1.96 $\pm$ 1.64).



**Figure 2.5. On and Offset Responses to USVs across Region and Experience. A)** Response spike rates on a per-call basis divided by auditory cortical region: Core (A1, AAF, UF) and A2, and by On (left) and Offset (right) response. Spike rates are normalized by subtracting the unit's spontaneous rate. Magenta: maternal animals (pup-experienced); Cyan: nonmaternal animals (pup-naïve). \*  $p < 0.0001$ , Wilcoxon Rank Sum Bonferroni Corrected. **B)** Prevalence of each type of response on a per-call basis, divided by auditory region (left), or by spike waveform (right). Spike waveforms were divided by Thick (peak to peak distance  $> 0.35$ ms) and thin (peak to peak distance  $< 0.35$ ms). Magenta: maternal animals (pup-experienced); Cyan: nonmaternal animals (pup-naïve). \*  $p < 0.0001$ , Fisher's Exact Test.

### 2.3.6 Prevalence of Offset Response to Pup USVs increases with experience

Changes may also be happening in the proportion of On or Offset responses being elicited, rather than in the spike rate of the response itself. When looking at how often either On only, Offset only, or On+Offset responses are observed among responses to individual calls, we find that in Core, no differences are seen after maternal experience (**Fig 2.5B**, top left). In A2 though, we find that the prevalence of Offset responses significantly increases ( $p < 0.0001$ , Bonferroni corrected Fisher's Exact), while On-only and On+Offset responses significantly decrease ( $p < 0.0001$ , Fisher's Exact) (**Fig 2.5B**, bottom left). Comparing on a per-unit basis, A2 neurons that are Offset-only responding are still significantly more prevalent ( $p < 0.005$ , Bonferroni corrected Fisher's Exact, data not shown; A2 Offset only: Maternal  $n=13$  of 23; Nonmaternal  $n=5$  of 29). On a per-animal basis, analyzing prevalence of On and Offset responses does not yield

significant results as any given animal contains both On and Offset responding units, and would always be classified as both an On+Offset responding animal.

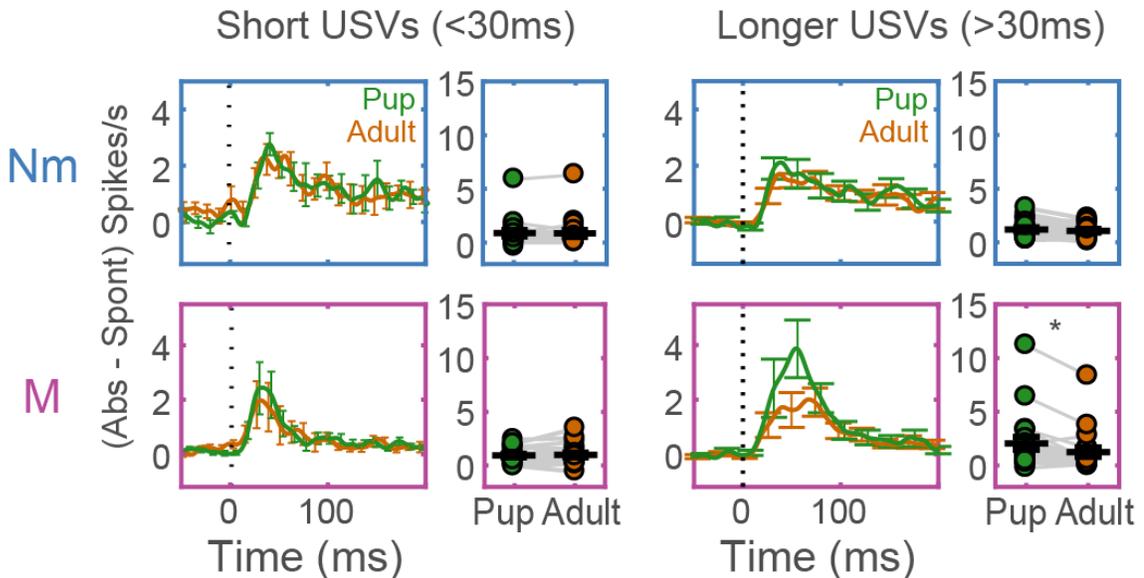
Responses were also analyzed based on spike waveform, as the spike waveform can be used to infer the neuronal subtype, and has been previously found to be predictive of response characteristics (Lin et al., 2010; Shepard, Lin, et al., 2015; Tsunada et al., 2012). In maternal Thick-spiking (peak-to-peak width > 0.35ms) units, significantly more Offset only responses were observed ( $p < 0.01$ , Bonferroni corrected Fisher's Exact) (**Fig 2.5B**, top right). No changes were seen in the prevalence of On only or On+Offset responses. Responses from Thin-spiking (peak-to-peak width < 0.35ms) units show no changes in prevalence of any three types of responses (**Fig 2.5B**, bottom right). On a per-unit basis, Thick-spiking units trend towards showing more Offset only responses, but comparisons do not survive multiple testing correction (ns, Bonferroni corrected Fisher's Exact; Thick-spiking Offset-only Units, Maternal  $n=18$  of 38; Nonmaternal  $n=16$  of 61).

Combined with the previous results for spike rate, average changes in On and Offset responses after maternal experience are more prominent at a population level in A2 than in Core auditory cortex. This does not exclude the possibility of changes in the Core happening to a smaller subset of neurons, however. In fact, a physiologically defined subset of units in maternal Core auditory cortex can discriminate different USV categories (Shepard, Lin, et al., 2015).

### ***2.3.7 Maternal SUs are capable of discriminating between USV categories with matched onset properties for longer calls only***

Natural mouse USVs from different categories such as pup vs. male adult USVs can overlap in acoustic space. Pup and adult USV spectrograms have differences in frequency trajectory, which may be a feature neurons are using to discriminate between

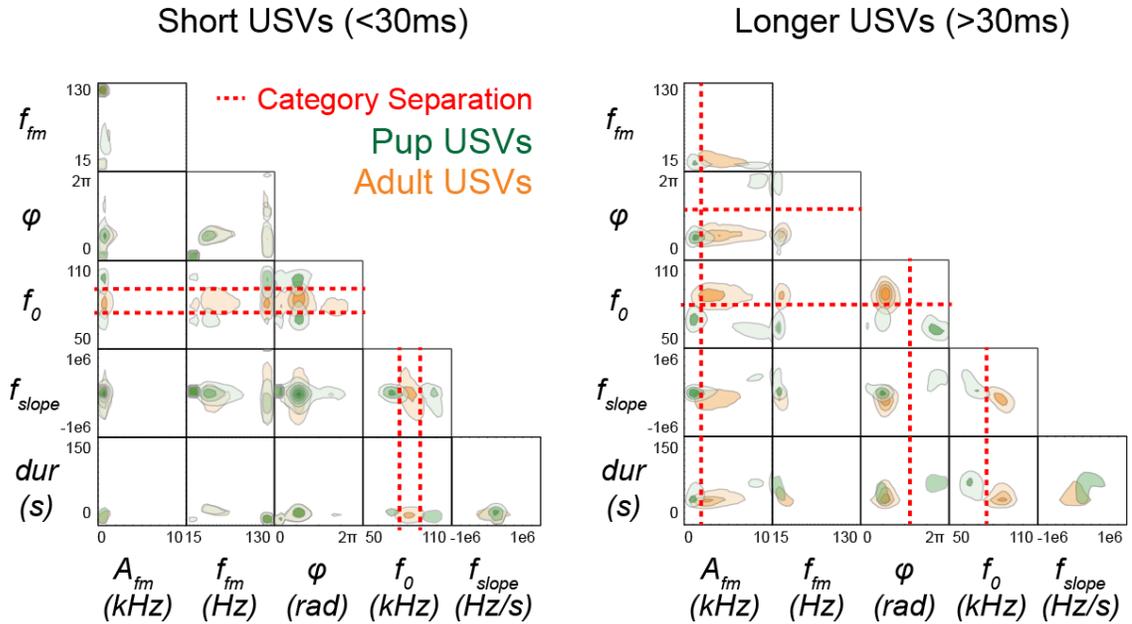
calls with overlapping onset acoustic properties. A subset of putative pyramidal neurons in Maternal animals' Core auditory cortex can discriminate between pup and adult calls with matched onset acoustic properties, as our previous studies have shown (Shepard, Lin, et al., 2015). Our call library consists of calls in three distinct clusters of call durations: ~13ms (short), ~35ms (medium), and ~60ms (long) (**Fig 2.4A**). Here, we extend our earlier findings with new units, and focus only on analyzing medium and long pup and adult USVs, which are the ones that can be discriminated by maternal SUs (**Fig 2.6**). Short pup and adult USVs that have matched onset frequency do not show response differences in these neurons, presumably because short USVs do not have as much time for acoustic trajectories to diverge so that they could be used to discriminate between different call categories. This leads to the question of what and how the acoustic information beyond the onset of the sound is used for discrimination by neurons in the auditory cortex. Response onset, whether at the sound onset or delayed from sound onset, may be important for discriminating call categories.



**Figure 2.6: Maternal SUs discriminate USVs with matched onset properties for long calls.** PSTHs of Putative Pyramidal Neuron responses (M: Maternal n=20; Nm: Nonmaternal n=19). Neurons are only capable of discriminating longer duration calls

(>30ms, which includes both medium and long USV clusters, Right), and not short duration calls (<30ms, Left).

Maternal animals' ability to discriminate longer, but not shorter, pup and adult USVs matched in onset frequency/modulation may be simply due to shorter USVs being impossible to acoustically separate from one another. To illustrate this, each call in our library was fit to a sFM function, and values for each of the six parameters were generated. To determine whether pup and adult USV categories can be separated in this parameter space, we plotted probability density clouds for pup and adult USVs within the six-parameter space to visualize whether there is separation between these two call categories (**Fig 2.7**). We note that when all calls are included regardless of their length, the parameter  $f_0$  (onset frequency) is best able to separate the two categories, as might be expected since adult calls tend to be higher in frequency than pup calls (Liu et al., 2003). However, since there is considerable acoustic overlap between the frequencies in pup and adult call categories, we also considered separability for cases in which a pup and adult USV have matched  $f_0$ . When we plot only short USVs ( $\leq 30$ ms) or only long USVs ( $> 30$ ms), we find that specifically for longer calls, the sFM parameter  $A_{fm}$  (amplitude of frequency modulation) shows category separation (**Fig 2.7**). From the distribution of the acoustics of pup and adult USVs,  $A_{fm}$ , as well as to some degree  $f_{slope}$  and  $\phi$  of the USV frequency trajectory, can be used to discriminate the two categories for longer calls. In shorter USVs, the only parameter that visually separates the call categories is  $f_0$  (onset frequency). As the frequency trajectory of short calls does not separate in sFM parameter space, for analysis of frequency trajectory, we considered only longer ( $> 30$ ms) calls.



**Figure 2.7: sFM six parameter distribution of pup and adult USVs.** sFM parameter space distribution of pup (green) and adult (orange) USVs. Left: Calls < 30ms in duration. Right: Calls > 30ms in duration. Short calls only show separation in  $f_0$  (onset frequency, which is only an onset property). Long calls begin to show separation in  $A_{fm}$  and  $\varphi$  (frequency trajectory properties). Red lines show where category separation can be seen.

### 2.3.8 Acoustic Features Encoded by On and Offset Responses via Generalized Linear Modeling

We investigated whether responses that occur at different times relative to sound playback encode different features of the sound. We divided responses into those occurring during sound playback or after playback ends, which were referred to as On or Offset responses respectively. To first determine what acoustic features are generally encoded by On and Offset responses regardless of animal experience, we utilized a Generalized Linear Model (GLM) that uses the six sFM parameters as input and attempts to predict the spike rate of the On or Offset response. Spike rates used in the model were normalized by the unit's spontaneous rate, and GLMs used a Poisson distribution with a Log link function. Only excited (non-negative) responses were included in the analysis.

When looking at only On responses, or the brief, phasic responses occurring during the sound waveform, we found that the only parameter that is significant is  $f_0$ , or frequency at the onset of the call (**Table 2.1**). All other parameters are nonsignificant. Given this is the predominantly phasic response when the sound is first turned on, encoding properties of the sound that are apparent at the onset is something we might expect.

**Table 2.1: GLM Parameter Estimates when fitting sFM parameters of USVs to spike rate responses to USVs** (Poisson distribution, log link function). Left: Parameter estimates when sFM parameters are used to predict transient (<20ms, n=219) On responses. Right: Parameter estimates when sFM parameters are used to predict Offset responses (300ms window after stimulus Offset, n=318). Note this analysis includes all call excited units across all auditory cortical regions and animal groups, including units whose data had been previously published. The Intercept term's significance indicates that if all coefficients for each of the other parameters are equal to zero, that the predicted spike rate is still non zero.

#### Transient On response

Term	Estimate	P > ChiSq
<b>Intercept</b>	<b>6.3098</b>	<b>&lt;0.0001</b>
Dur	-0.0021	0.6792
$A_{fm}$	3.616e-5	0.0596
$f_{fm}$	-0.0013	0.6817
$\varphi$	-0.0005	0.3409
<b><math>f_0</math></b>	<b>-4.684e-5</b>	<b>&lt;0.0001</b>
$f_{slope}$	-3.073e-7	0.0959

#### Offset Response

Term	Estimate	P > ChiSq
<b>Intercept</b>	<b>2.6056</b>	<b>&lt;0.0001</b>
<b>Dur</b>	<b>0.0063</b>	<b>0.0194</b>
<b><math>A_{fm}</math></b>	<b>4.9873e-5</b>	<b>&lt;0.0001</b>
$f_{fm}$	0.0021	0.2196
$\varphi$	-0.0003	0.2837
<b><math>f_0</math></b>	<b>-3.368e-5</b>	<b>&lt;0.0001</b>
<b><math>f_{slope}</math></b>	<b>-6.194e-7</b>	<b>&lt;0.0001</b>

Looking at what sFM parameters significantly explain Offset response spike rate, we find that several parameters are significant, including onset frequency ( $f_0$ ),  $A_{fm}$ ,  $\varphi$ ,  $f_{slope}$ ,  $dur$  (**Table 2.1**). From this, we observe that the Offset response encodes some of the sFM parameters that describe the acoustics of the frequency trajectory of the call beyond just the onset properties. Moreover, the parameters  $A_{fm}$ ,  $\varphi$ , and  $f_{slope}$  are parameters in which longer pup and adult USV categories show separation in sFM parameter space (**Fig 2.7**), and would be acoustic features that are useful for discriminating the call categories. The Offset response specifically shows sensitivity to these acoustic parameters, lending to the importance of the Offset response in discriminating sounds in which the frequency trajectory is an important cue for identifying the meaning of a sound. However, the presence of differences in what parameters the Offset response is sensitive to does not give a full picture of how exactly the responses or receptive field of the neurons are changing with experience. One way to look at this was to determine whether the Offset response is better able to respond to one USV category versus another with experience.

### ***2.3.9 Maternal SUs respond to the same number of pup USVs on average as nonmaternal SUs***

In order to determine whether maternal units have an increased general responsiveness to pup USVs, for a given unit, we can quantify the total number of pup USVs that a unit exhibited an On or an Offset response to. We find that there is no significant difference in the number of pup USVs maternal units respond to compared to nonmaternal units, whether we are considering the On or Offset response (Wilcoxon rank sum test NS). Similarly, no differences are seen in the number of adult USVs that maternal or nonmaternal units respond to in the On or Offset response. No differences were seen in the distribution of number of pup USVs units were responsive to (K-S test NS). Results

hold regardless of whether region (Core or A2) or spike waveform (thick or thin) are considered. Overall, it appears the number of calls that units will respond to on average does not change across experience. This is not entirely unexpected in the context of previous studies that have found no frequency representation expansion of the ultrasound frequency range with maternal experience (Shepard, Lin, et al., 2015). Rather than excitatory plasticity at the population level, a finer excitatory plasticity occurs within a smaller subset of putative pyramidal neurons in the core auditory cortex, where this subpopulation shows the ability to discriminate between USV categories even when onset acoustic properties (such as frequency and frequency modulation) were matched. However, in this study's subset of units, few putative pyramidal neurons were recorded from so this was not observed in our units.

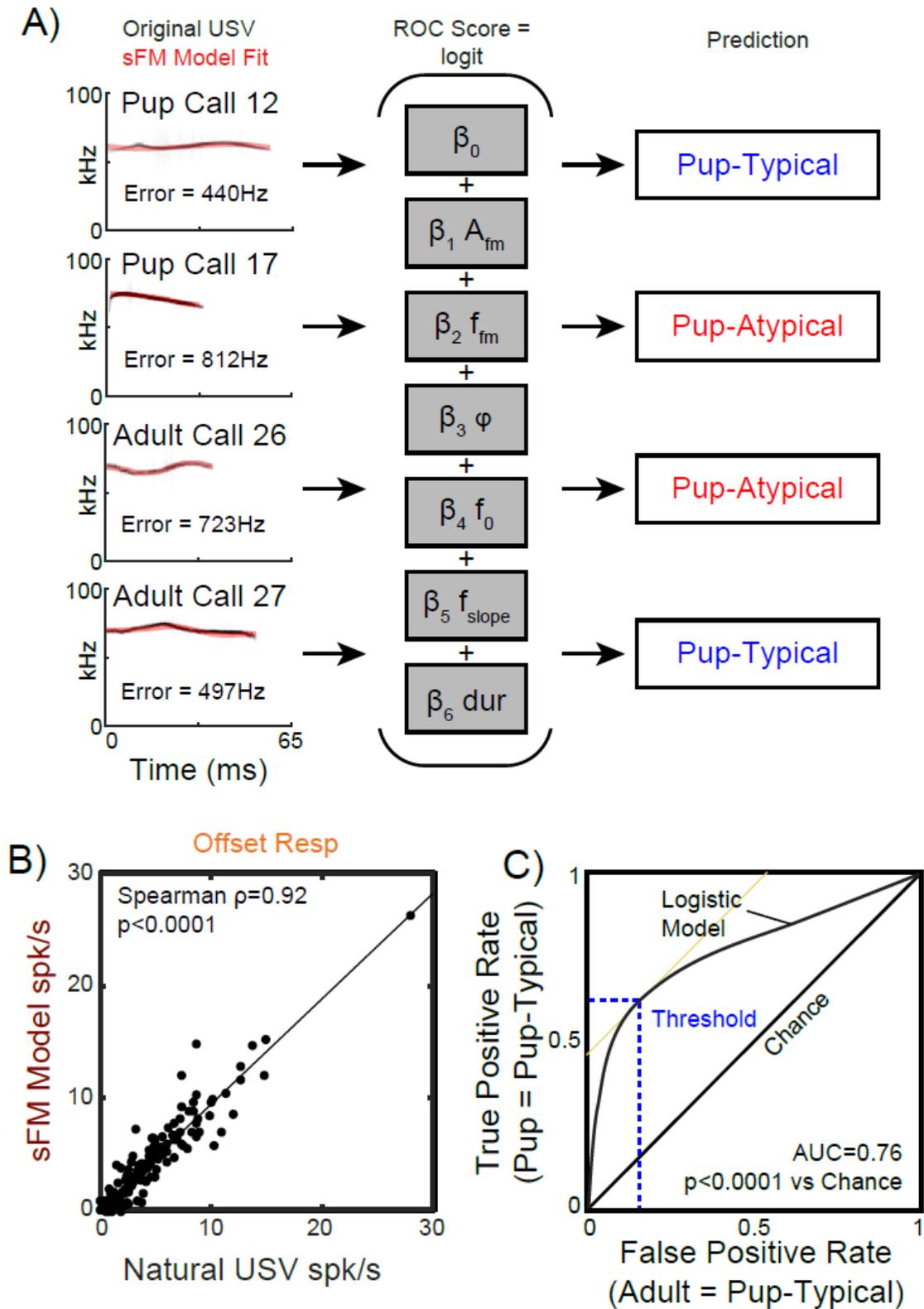
#### ***2.3.10 Population of Maternal and Nonmaternal SUs equally likely to respond at On or Offset of all USVs***

Rather than looking at the number of calls a given unit responds to, we can look at the proportion of maternal or nonmaternal units that respond to each individual call. For each of  $n=24$  medium/long calls in our USV stimulus library, the proportion of units that have an On or Offset response to the call was compared between maternal and nonmaternal units. No differences are seen between the proportion of maternal and nonmaternal units that respond to any of the 24 calls with either an On or Offset response (Bonferroni corrected Fisher's exact test). In addition, no differences are seen when breaking units down by region (Core or A2). This is surprising given the enhanced prevalence of Offset responses seen in maternal animals in A2. However, in this analysis, all calls are pooled and considered equally when they may in fact have different acoustic parameter distributions that may make some of them more pup-like or adult-like. Rather than showing a difference on an individual call basis, there is still the possibility that A2

maternal units, on average, will respond to certain more stereotypical calls that fall within a specific acoustic space (such as what would be a space that is “pup-like”).

### ***2.3.11 Sinusoidally Frequency Modulated (sFM) Tones can be used as a model of Natural Vocalization Frequency Trajectories***

In order to define what acoustic parameter distributions would be pup-like or adult-like, we used the sFM model to parameterize the frequency trajectories of mouse USVs. We fit our USV call library of 57,929 pup USVs and 10,353 adult USVs to the six parameter model and show that error for each fit is relatively small, on the order of <1 kHz, which is small compared to the vocalizations themselves that reside in the range of 60-80 kHz (**Fig 2.8A**, left). We also show that Offset spiking responses elicited by the natural USV and the paired sFM model USV are significantly correlated ( $p < 0.0001$ , Spearman  $\rho = 0.92$ , **Fig 2.8B**). Note that even when including the full response rather than only the Offset response, natural USV and sFM model USV spike rates are significantly correlated (data not shown). These results demonstrate that our sFM model elicits similar neural responses to the natural calls, and can be used to model mouse USVs.



**Figure 2.8. sFM works as a model for USVs and can be used to discriminate call categories using a Logistic Regression model. A) Left: Spectrograms of sample pup**

and adult calls with original USVs in black and sFM model fits in red. For each call, a set of six parameters is generated, which is then used to attempt to predict call type using a nominal logistic regression model (pup: pup-like; adult: adult-like). **B)** Comparison of original, natural USV Offset spike rates to matched sFM model spike rates, showing high correlation (Spearman rho=0.92, p<0.0001). **C)** Performance of the nominal logistic regression model via ROC analysis in correctly classifying pup calls as pup-like (true positive) and in incorrectly classifying adult calls as pup-like (false positive). Threshold represents the cutoff at which the true positive rate is maximized and the false positive rate is minimized. The model performs above chance (AUC = 0.76, p<0.0001 via Bootstrap analysis, N=1000).

### **2.3.12 sFM Parameters can be used to classify call types using a Logistic**

#### **Regression Model**

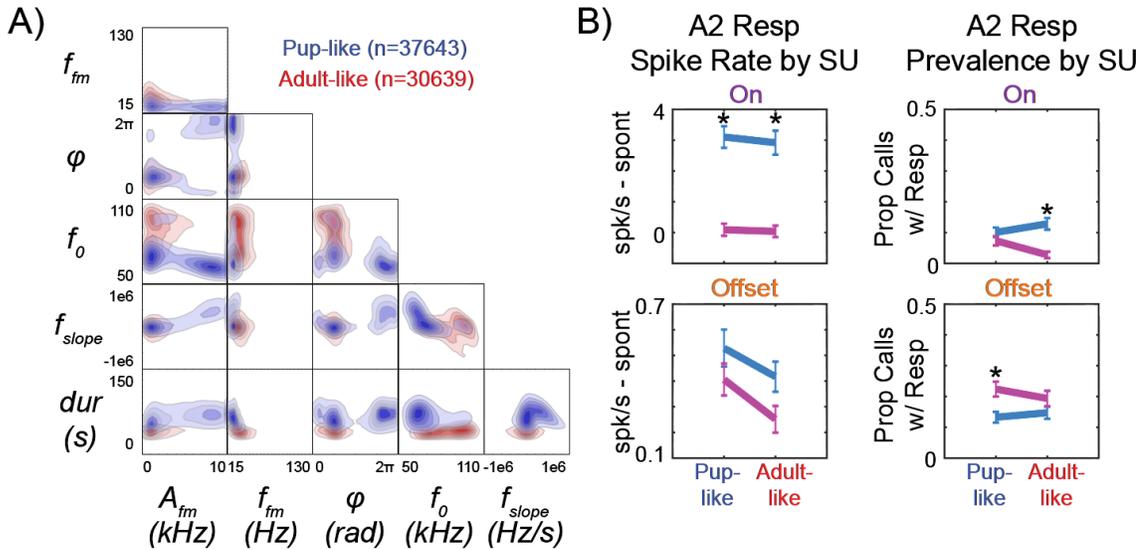
We then wanted to determine whether the two call types (pup and adult) were separable in our parameter space, and whether the combination of a call's sFM parameters could correctly predict which category (pup or adult) it came from. To do this, we used a nominal logistic regression model that attempts to classify each call as a pup-like or an adult-like based on the six sFM parameters (**Fig 2.8A**, right; Model parameters:  $\beta_{Afm} = -1.6e-5$ ;  $\beta_{Ffm} = -0.0024$ ;  $\beta_{\phi} = 0.08$ ;  $\beta_{f0} = -9.43e-6$ ;  $\beta_{fslope} = -2.10e-8$ ;  $\beta_{dur} = 0.034$ ; Model  $\chi^2 = 7.16e3$ , p<0.0001). The call classification model performs significantly above chance using receiver operating curve (ROC) analysis (Area Under Curve AUC=0.75842, 95% Confidence Interval CI: 0.75414 – 0.76242, Bootstrap N=1000, p<0.0001, **Fig 2.8C**). The best performing score threshold is 0.8686, where a score > 0.8686 indicates the call would be classified as having a pup-like set of parameters. The model's sensitivity (true positive rate) is 61.99%, and 1-specificity (false positive rate) is 16.82%. Based on these results, frequency trajectory as described by the six sFM model parameters can be used to distinguish between pup and adult calls above chance performance with an ideal observer.

We can then look at how the calls within our stimulus set of n=36 calls would be classified by this ideal observer. The calls within this set are specifically sampled such

that pup and adult calls are matched in onset acoustic parameters, and sampled evenly across duration and onset frequency, which led to many of them falling outside of the typical acoustic space for pup calls, so a lower classification accuracy would be expected. Of  $n=18$  pup calls, 7 pup calls were classified as pup-like (all 6 short duration pup calls were not classified by the regression as pup-like). Of  $n=18$  adult calls, 6 adult calls were classified as pup-like. The overall distribution of pup-like and adult-like calls in six parameter space was visualized (**Fig 2.9A**). The calls classified as pup-like tend to have greater duration and lower  $f_{fm}$ .

### ***2.3.13 A2 On-only and Offset-only responses are suppressed in Maternal units for Adult-like USVs***

Now that we are able to classify calls by whether they have pup-like or adult-like frequency trajectories, we looked back at our prior results of decreased On response spike rate and increased Offset response prevalence in A2 of maternal animals. Here, we asked whether the changes were happening specifically for calls that have a pup-like or adult-like frequency trajectory. We find that when splitting calls by pup-like or adult-like, the decreased On response spike rate occurs universally for both types of calls (Fisher's Exact Bonferroni corrected  $p < 0.05$ ; **Fig 2.9B**, left). We find that prevalence of On responses *decreases* specifically for adult-like calls, while the prevalence of Offset response *increases* specifically for pup-like calls in maternal animals ( $p < 0.05$ ; **Fig 2.9B**, right). This shows that in maternal A2, units become more likely to show Offset responses specifically to calls that have a frequency trajectory that is more pup-like.

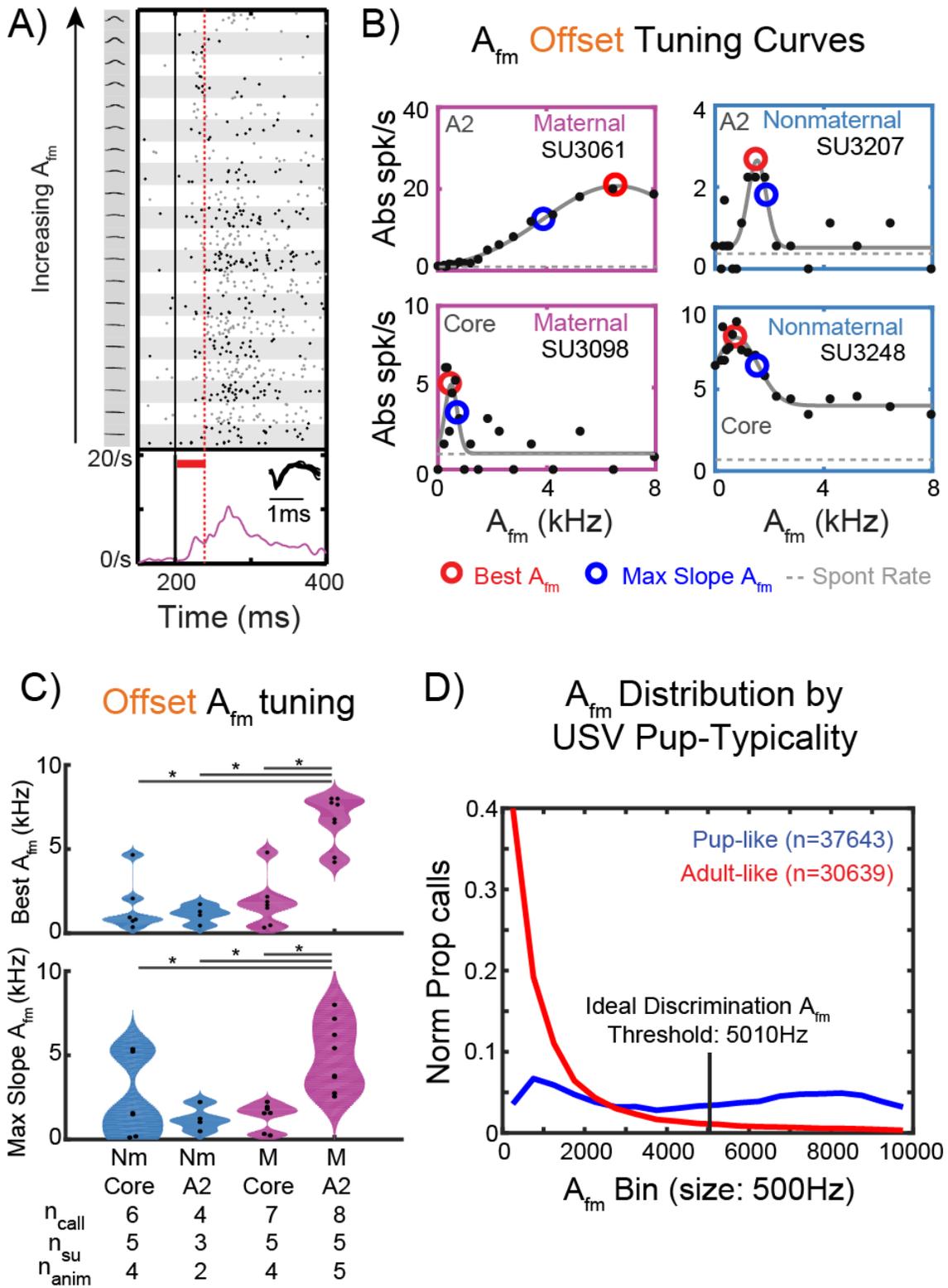


**Figure 2.9. sFM Parameter Distribution and Unit Responses to USVs classified by ROC. A)** Distribution of all USVs across the six sFM parameters, with pup-like calls depicted in blue and adult-like calls depicted in red. Some degree of separation between categories is seen in  $A_{fm}$ , duration, phase ( $\varphi$ ), and onset frequency ( $f_0$ ). **B) Left:** Single unit spike rate responses to USVs during the On response window (top) and Offset response window (bottom), separated by USVs that are pup-like and adult-like. **Right:** Prevalence of On responses (top) or Offset responses (bottom) to USVs that are either pup-like or adult-like. \* $p < 0.05$  Fisher's Exact Test Bonferroni Corrected.

### 2.3.14 A2 Offset responses in Maternal Units are tuned to higher $A_{fm}$

We next asked whether maternal A2 units' Offset responses change in how they are frequency trajectory parameters. Given that units can show tuning to  $A_{fm}$  for general stimuli, and that  $A_{fm}$  is one of the parameters that appears to separate pup-like and adult-like calls (**Fig 2.9A**), we sought to evaluate whether  $A_{fm}$  tuning changes for maternal A2 Offset responses. We selected the best call for each SU during recording using a curated set of  $n=36$  USVs, and designed a stimulus with systematically varied  $A_{fm}$  in 19 steps around the best call, while all other five sFM parameters were fixed to those equal to the best call (**Fig 2.10A**). From the Offset response of the neuron, we generated an  $A_{fm}$  tuning curve, where the best  $A_{fm}$  as well as the point at which the tuning curve exhibits the greatest slope (Max slope) can be calculated (**Fig 2.10B**). The Max slope  $A_{fm}$  value was of interest as it represents the theoretical point at which a

neuron's  $A_{fm}$  discrimination is the greatest, and a small change in  $A_{fm}$  leads to the largest change in firing rate. In evaluating the best  $A_{fm}$  across experience and region, we find that Offset responses in maternal (M) A2 show a preference for higher  $A_{fm}$  values than nonmaternal (Nm) A2 (on a per-call basis, M A2  $n=8$ ; M Core  $n=7$ ; Nm A2  $n=4$ ; Nm Core  $n=6$ ;  $p<0.0001$  Tukey-Kramer Honestly Significant Difference (HSD) significant for [M A2] vs [Nm Core, Nm A2, and M Core]; **Fig 2.10C**, top). For the Max Slope  $A_{fm}$  value, we also find that M A2 units have a significantly greater Max Slope  $A_{fm}$  compared to other groups ( $p<0.0001$  Tukey-Kramer HSD significant for [M A2] vs [Nm Core, Nm A2, and M Core]; **Fig 2.10C**, bottom). All results still hold if done by unit or by animal, taking the average best  $A_{fm}$  value per unit or per animal. On responses are not significantly different across groups (data not shown). In this cohort, many of the units did not exhibit an On response to the presented USV stimuli (8 of 25 show an On response). Note that this result is only seen when looking at best  $A_{fm}$  around natural calls. When comparing the best  $A_{fm}$  for stimuli centered around a unit's BF (spanning both low and high frequencies), we see no differences across animal groups or region. These results demonstrate an upward shift in  $A_{fm}$  tuning of Offset responses that is specific to USVs in maternal A2 units.



**Figure 2.10.  $A_{fm}$  tuning around USVs changes with experience.** **A)**  $A_{fm}$  tuning stimulus around call played to a sample single unit. Stimulus consists of 19 steps of  $A_{fm}$  from 0 to 8000Hz spanning the range of natural USV  $A_{fm}$  values. The other sFM

parameters are fixed at values matching the USV that elicits the best response from the given unit. Small spectrogram depictions of the stimulus are shown on the left. A PSTH shows pooled spike rate responses underneath the raster, with the red line denoting the end of the stimulus. **B**)  $A_{fm}$  tuning curves from samples units' Offset responses. Two units from A2 (top) and two from Core (bottom) are depicted, with a gaussian fit over the tuning curve and the peak denoting the unit's best preferred  $A_{fm}$  value, and the triangle depicting the value at which the maximum slope in the unit's response is located. Spontaneous rates are shown with a dotted gray line. **C**) Overall population data showing the best  $A_{fm}$  (top) and Max Slope  $A_{fm}$  (bottom) divided by animal experience group (M = maternal, Nm = nonmaternal) as well as auditory region. \* $p < 0.05$  Tukey Kramer's HSD. **D**) Histogram showing the  $A_{fm}$  distribution of calls in either the pup-like (blue) or adult-like category (red), as a normalized proportion of calls (over total number of calls in the category). The ideal observer discrimination threshold for pup-like vs adult-like, as determined by ROC, is shown in black at  $A_{fm} = 5010\text{Hz}$ .

### **2.3.15 Pup-like Calls have Higher $A_{fm}$ Values**

Given that  $A_{fm}$  tuning in maternal A2 neurons are significantly higher than for other groups, we asked how these changes are reflective the acoustics of the pup call category. We looked at how  $A_{fm}$  correlates with the pup-likelihood of a given call (as measured by the logistic regression score). We find that higher  $A_{fm}$  values are correlated with higher pup-likelihood score (Spearman rho = 0.6766;  $p < 0.0001$ ). The distribution of  $A_{fm}$  across pup-like calls shows an almost bimodal distribution and has a larger proportion of pup-like calls with high  $A_{fm}$  values (**Fig 2.10D**). We also calculated the theoretical best  $A_{fm}$  threshold to discriminate pup from adult USVs by performing a nominal logistic regression using only  $A_{fm}$  as a parameter. We found the threshold that best separates pup and adult calls is  $A_{fm} = 5009.7\text{ Hz}$  (AUC 0.61574  $p < 0.0001$  True Positive = 38.8%; False Positive = 13.15%, **Fig 2.10D**). Interestingly, the mean Max Slope  $A_{fm}$  value of the M A2 group is 4950Hz, which is very close to the ideal pup vs adult USV discrimination threshold (compared to Nm A2's Max Slope  $A_{fm}$  of 1245Hz, **Fig 2.10C**, bottom). These results show that with experience, A2 exhibits a shift in frequency trajectory parameter ( $A_{fm}$ ) tuning that reflects acoustics that are stereotypical of the sound category that was learned.

## 2.4 Discussion

Our results demonstrate that auditory cortical neurons can show tuning to subtle sinusoidal frequency modulations that cannot be explained by the spectral power in the signal. We utilized a sinusoidally frequency modulated (sFM) tone with six parameters to describe the sound's frequency trajectory. We specifically chose to investigate variations in amplitude ( $A_{fm}$ ) and frequency ( $f_{fm}$ ) of frequency modulation as studies have previously explored the encoding of linear frequency modulation in the auditory cortex (Gaese et al., 1995; Mendelson et al., 1985; Rauschecker et al., 2000), where preference for up or downward linear frequency modulation sweeps have been documented. In the context of natural vocalizations, frequency trajectories can undergo more complex frequency modulations that are not entirely captured by linear modulations alone, so we selected an sFM model that better fits natural frequency modulations. Our results showed that neurons can prefer sinusoidal FM over linear FM spanning the same frequencies, further emphasizing the importance of being able to capture more complex frequency trajectories in the study of auditory cortical encoding of frequency modulation.

Recent studies have also begun to highlight the importance of segregating On and Offset, with studies demonstrating that On and Offset responses arise from different synaptic sources (Scholl et al., 2010). Across studies, it is frequently reported that On and Offset responses of the same neuron can be tuned to distinct frequencies (Qin et al., 2007; Scholl et al., 2010; Sollini et al., 2018). Offset responses have been found to be important for gap detection (Anderson et al., 2016), and the ability to perform gap detection plays an important role in speech processing (Weible et al., 2014). In addition, a late responding subset of thick-spiking, putative pyramidal neurons has been previously described to acquire the ability to discriminate between call categories after maternal experience in the mouse (Shepard, Lin, et al., 2015), pointing to later responses playing a potentially important role in vocalization discrimination. In this study,

we found that the prevalence of Offset responses to calls increases specifically in thick-spiking units and not in thin-spiking units across auditory cortical areas. We also show that Offset responses play an important role in frequency trajectory encoding, and that plasticity in Offset responses occurs over the course of natural vocalization learning. Specifically, we found that the tuning of Offset responses to frequency trajectory parameters shifts towards those that more closely resemble the frequency trajectories in the learned vocalization category.

In our study, we also investigate the role of secondary auditory cortex (A2) in frequency trajectory encoding. We observed plasticity in Offset response tuning to  $A_{fm}$  on a population level in secondary auditory cortex of maternally experienced animals and not in primary auditory regions. In prior literature, little is known about the response properties of the secondary auditory cortex A2, with most auditory cortical studies on response properties focusing on the primary auditory areas. A2 has been reported to show less tonotopic organization compared to primary auditory regions (Guo et al., 2012; Stiebler et al., 1997), although a frequency gradient within A2 has been reported using  $Ca^{2+}$  imaging when using sinusoidally amplitude modulated tones rather than pure tones (Issa et al., 2014). A2 has been thought to play a role in higher-order sound processing as well as for segregation of sound objects (Geissler et al., 2004; Joachimsthaler et al., 2014). This study sheds light on the response properties of A2, demonstrating that A2 Offset responses are tuned to frequency modulation parameters, and that this tuning is plastic with experience.

This study is the first to demonstrate Offset responses can show tuning to sinusoidal frequency modulation (sFM), and that with experience in a natural vocalization learning paradigm, frequency trajectory tuning in the Offset response can shift to better match the characteristics of the learned vocalization category. The changes are specifically observed on the population level in the ventral secondary

auditory cortex (A2) rather than primary or core auditory areas, and changes are predominantly seen in thick-spiking neurons. Our results suggest that sFM may be an acoustic feature that neurons in auditory cortex attune to, and that segregating On and Offset responses are important in studying auditory cortical plasticity.

## Chapter 3

### Conclusions

Our current understanding of the encoding of sound by the auditory cortex is shaped by the library of stimuli that have been used to probe auditory cortical responses, as well as the details of the experimental preparation during study, such as type of anesthesia and recording methodology. How the details of the experiment are selected are important to the types of conclusions that can be gleaned from the study. For example, the type of stimulus selected can lead to very different conclusions regarding the receptive field of a given neuron. Certain types of anesthesia can eliminate the presence of Offset responses, leading to very different neural activity than what one may observe in an awake subject. Additionally, response properties of neurons depend heavily on what region of the auditory cortex they are sampled from, as different areas of auditory cortex are involved in different kinds and levels of processing.

The work in this thesis specifically examines the mouse core and secondary auditory cortex in the awake state with single-unit electrophysiology. The findings demonstrate the importance of Offset responses in frequency trajectory encoding, as well as the presence of Offset response plasticity in the ventral secondary auditory field during the learning of natural sounds.

#### ***3.1 Frequency Trajectory Encoding in the Auditory Cortex***

Historically, pure tones have been used to probe auditory cortical responses, and have led to important insights about the response properties of neurons in the auditory cortex. However, work using more complex stimuli such as two-tone stimuli have made it clear that responses observed to pure tones cannot alone be used to predict responses to all other types of stimuli. The auditory system has adapted over years for the

processing of sounds that are encountered naturally in the environment, oftentimes with complex temporal structure such as varying envelope or frequency modulation. One of the less well-understood aspects of auditory cortical encoding is how frequency trajectory is represented in the auditory cortex.

In this study, we used stimuli whose frequency trajectory was mathematically described using a parameterized model of a sine wave. The frequency trajectory of a sound is useful for indicating emphasis, intention, or word meaning in tonal languages (Lieberman et al., 1956; Liberman et al., 1957; Miller et al., 1979; Stevens et al., 1974). Studies have not extensively explored encoding of frequency trajectory, even though complex frequency trajectories are often a feature of natural sound. Complex frequency trajectories are multidimensional, and their complexity may contribute to why it has not been studied in the past.

Using the parameterized model, by specifically varying the amplitude or frequency of frequency modulation, we demonstrated that a fraction of neurons can show tuning to these parameters, and that their response properties are not explainable by their spectral content. Preference for temporal modulation in stimuli for auditory cortex over pure tones has been observed across species (Hall et al., 2002; Hart et al., 2003; Liang et al., 2002), although previous studies did not look at tuning across frequency modulation depth. We also demonstrated that neurons can respond better to sFM compared to linear frequency modulation, which further supports that neurons are responding to the frequency trajectory traversed by the stimulus rather than the spectral content. Using a generalized linear model, we showed that On responses encode primarily onset frequency, whereas Offset responses can encode various other sFM parameters that describe the frequency trajectory. This work is the first to investigate frequency trajectory coding by Offset responses. As we are interested in studying

frequency trajectory encoding for natural sounds, we then utilized the sFM stimulus in a natural adulthood learning paradigm.

### ***3.2 The Mouse Model for Experience Dependent Plasticity***

This work utilized a natural paradigm in which mouse mothers learn the behavioral significance of pup ultrasonic vocalizations (USVs), which are frequency-modulated whistles. Pups isolated from the nest emit vocalizations that mothers learn to approach to retrieve the pup to the nest. Pup-naïve female mice do not behaviorally respond to these vocalizations, but can learn to do so through approximately 5-7 days of housing with a mother and pups. The mouse USVs are primarily single frequency whistles (Liu et al., 2003), which allowed us to recreate and parameterize USVs using the sFM model. Our work demonstrated that it is possible to use sFM tones as a model of natural mouse USVs, and that neural responses to an sFM model and natural USV, when envelopes are matched, are comparable. Previous research within our lab has shown that maternal animals can discriminate between pup and adult USVs that have matched onset acoustic properties (Shepard, Lin, et al., 2015), which led us to believe that perhaps how the sound changes temporally, such as in its frequency trajectory, allows animals to discriminate sound category. Indeed, single units in maternal animals are capable of distinguishing pup and adult USVs for long calls only. This led us to examine changes in the Offset response for maternally experienced animals.

### ***3.3 Offset Responses in the Auditory Cortex***

Reporting of Offset responses in auditory cortex has been inconsistent (Heil, 1997a, 1997b; Hind, 1953; Mendelson et al., 1997; Merzenich et al., 1975; Phillips et al., 1990; Phillips et al., 1994; Schreiner et al., 1992; Sutter et al., 1995); (Moshitch et al., 2006; Volkov et al., 1991), mainly due to different recording conditions such as

anesthesia (Qin et al., 2007). Work in awake subjects has made it clear that the Offset response is an informative physiological response about the stimulus, whereas the Offset response was not observed in earlier barbiturate-anesthetized electrophysiology studies (Heil, 1997a, 1997b; Hind, 1953; Mendelson et al., 1997; Merzenich et al., 1975; Phillips et al., 1990; Phillips et al., 1994; Schreiner et al., 1992; Sutter et al., 1995). More recently Offset responses have been studied separately from responses at the Onset of sound playback, and studies have observed that individual neurons can show an On and Offset response that is tuned to different areas of frequency space (Fishman et al., 2009; Qin et al., 2007; Tian et al., 2013). Others have extended this work to demonstrate that the different receptive fields in On and Offset responses contributing to a unit's upward or downward directional preferences for linear frequency modulated sweeps (Sollini et al., 2018).

In this work, we demonstrated that Offset responses are present in response to both pure tones and to natural vocalizations, and that Offset responses are more likely to show a preference for frequency modulated stimuli over pure tones than On responses. This illustrates the importance of the Offset response in sound encoding, and further differentiates the two response types beyond the previously reported differences in frequency tuning.

Most importantly, we observed adult plasticity in Offset responses. With maternal experience, we found an enhanced prevalence of Offset responses and decreased On response strength in the adult auditory cortex. Furthermore, the Offset response's tuning to sFM parameters in maternal animals shifts towards values that are characteristic of pup USVs. It is important to note that these changes were confined to the secondary auditory cortex, and were not observed in what is commonly referred to as core areas (A1, AAF, UF). However, this does not exclude the possibility of changes occurring in a small, defined subpopulation of neurons in core auditory cortex as previously reported,

which may be obscured when looking at responses across the entire population of core neurons (Shepard, Lin, et al., 2015).

### ***3.4 The Role of Secondary Auditory Cortex in Frequency Trajectory Encoding***

The secondary auditory cortex has historically been demonstrated to be poorly responsive to pure tones, preferring stimuli with more complex features such as broadband noise (Kaas et al., 1999; Kikuchi et al., 2010; Rauschecker et al., 1995). In the human, the secondary auditory cortex is thought to be important for processing complex sounds such as speech (Geschwind, 1972; Kimura, 1961; Zatorre et al., 2002) and music (Patterson et al., 2002). Most electrophysiological mouse studies have concentrated on the neural responses in A1, with a select few that look at A2 some for the purpose of mapping the entire mouse auditory cortex (Bandyopadhyay et al., 2010; Guo et al., 2012; Rothschild et al., 2010; Stiebler et al., 1997). In the case of electrophysiological mapping studies, tonotopy has not been observed in A2. However, higher resolution methods including intrinsic imaging and Ca<sup>2+</sup> imaging have demonstrated tonotopic gradients in A2 (Issa et al., 2014; Kubota et al., 2008; Tsukano et al., 2015; Tsukano et al., 2016; Tsukano et al., 2017). It should be noted that for these imaging studies, the pure tones presented also included amplitude or frequency modulations during their presentation. It remains that A2 is often reported to be poorly driven by pure tones, and preferentially responds to complex stimuli (Kaas et al., 1999; Kikuchi et al., 2010; Rauschecker et al., 1995).

In our study, we observed plasticity specifically in the Offset response of A2 to behaviorally relevant calls. Specifically, in maternal animals, there was an enhanced prevalence of Offset responses on a population level, and the tuning of the Offset response to  $A_{fm}$  shifted towards values that are more characteristic of the pup USV category. The enhanced prevalence of Offset responses in A2 improves the

representation of pup USVs in maternal AC by increasing the number of neurons responding to the sound, and biasing their responses towards frequency trajectories that are likely in pup USVs. This illustrates an important role for plasticity in ventral auditory Offset responses in the processing and categorization of natural frequency modulated sounds, and contributes to our understanding on the role of Offset responses and the nature by which they can be altered through experience.

### **3.5 Future Directions**

Much remains to be uncovered in the exploration of frequency modulation encoding by the auditory cortex. Beyond the exploration of tuning in other parameters that were not used in this study (such as the phase or duration), there remains questions regarding the mechanism of the Offset plasticity observed. How do more Offset responses arise in the secondary auditory cortex? Some possibilities include increased or decreased inhibition of On or Offset responses respectively by interneuronal populations. It has been shown that at least one subtype of interneuron, the Parvalbumin-expressing subtype, is sparse in the secondary auditory cortex (Cruikshank et al., 2001), and has been verified in the mouse strain this work is conducted in (See **Appendix D.2**). Targeting other types of interneurons, such as Calbindin, Somatostatin, or Vasoactive Peptide expressing interneurons may be useful in identifying what subtypes contribute to On or Offset responses. For example, viral genetic manipulation or transgenic breeding allows for expression of an inhibitor opsin in a target cell-type, such that only that cell population is suppressed during the presentation of the correct light wavelength. Auditory cortical recordings conducted with or without inhibition across auditory cortical areas can help discern whether a cell type contributes generally across the auditory cortex to On or Offset responses, helping to further elucidate the mechanism by which these responses are generated. Cell type specific recording can be

made possible using the technique of Photoidentification of Neuronal Populations (See **Appendix H**).

Additionally, there is the question of how behaviorally important the Offset response is. Maternally experienced mice will respond to pup USV playback with preferential phonotaxis or approach behavior. Based on the results of this work, the Offset response in A2 becomes biased towards responding to pup USV-like frequency trajectories, although it is unclear whether the Offset response is itself necessary for the mouse to behaviorally respond. To answer the question of whether Offset response activity is necessary for the expression of preferential phonotaxis, an experiment can be conducted with inhibition of A2 that is locked to the timing of the Offset of sound playback. The precise inhibition timing can be achieved through optogenetic inhibition, and successful inhibition can be confirmed with simultaneous electrophysiology such that On responses can still be confirmed while Offset responses are shown to be suppressed. The behavioral expression of phonotaxis in a T-maze or W-maze setup (Lin et al., 2013) can then be assessed. As a separate experimental group, inhibition of only the On response in A2 during stimulus playback can be conducted to compare behavioral responses. This experiment will allow us to determine how important the On or Offset responses in A2 are to the expression of maternal approach behavior.

## Appendix A

### List of Single Units used in Experiments

#### A.1 List of Units in On and Offset Prevalence Analysis

A total of n=137 call-excited units were included in On and Offset prevalence analysis, and their unit IDs are listed below.

**Table A.1. List of Units in On and Offset Prevalence Analysis.** Single unit **SUnitIDs** based on the MasterDB.xls file are listed in the first column. **Animal IDs** are listed, where each unique animal ID corresponds to a unique animal. **Animal Group** is listed by experience type, where Parents are post-weaning late mothers, EarlyCocare are cocarers at the P5-7 time point, Naïve are completely pup-naïve females, and Yoked are animals that are housed in a divided cage on the side opposite from a mother and her litter. All animals in this list are females. **Region** shows the auditory cortical area that the unit comes from: Primary Auditory Cortex (A1), Anterior Auditory Field (AAF), border between A1 and AAF (AAFA1), Ultrasound Field (UF), Secondary Auditory Cortex (A2). **SpkType** denotes the spike waveform: Thick, peak-to-peak width > 0.35ms; thin, peak-to-peak width < 0.35ms.

<b>SUnitID</b>	<b>AnimalID</b>	<b>Group</b>	<b>Region</b>	<b>SpkType</b>
2329	E713072302A	Yoked	UF	Thick
2333	E713073103A	Yoked	A1	Thick
2338	E713081403A	Yoked	AAFA1	Thick
2344	E713081905A	Yoked	UF	Thick
2350	E713101102A	EarlyCocare	UF	thin
2371	E914012305B	Naïve	A2	thin
2374	E914012305B	Naïve	AAF	Thick
2388	E914013104B	Yoked	A1	Thick
2393	E914021504A	Yoked	AAF	thin
2394	E914021504A	Yoked	AAF	thin
2399	E914021904A	Yoked	A2	Thick
2400	E914021904E	Yoked	A2	thin
2409	E914021904E	Yoked	AAFA1	thin
2411	E914021904E	Yoked	UF	Thick
2412	E914021904E	Yoked	UF	thin
2413	E914021904E	Yoked	UF	Thick
2414	E914013101A	Parent	A2	Thick
2418	E914031205A	Yoked	A2	Thick
2434	E914031804A	Yoked	A2	Thick
2435	E914031804A	Yoked	AAFA1	Thick

**Table A.1** continued

2442	E914031704A	Yoked	AAF	thin
2458	E914032004B	Yoked	UF	thin
2463	E914032504A	Yoked	AAF	Thick
2466	E914032504A	Yoked	A2	Thick
2467	E1014060904A	Yoked	A2	Thick
2468	E1014060904A	Yoked	A2	Thick
2486	E1014061104A	Yoked	A2	thin
2495	E1014061104A	Yoked	AAF	thin
2499	E1014061104A	Yoked	AAF	Thick
2505	E114070305A	Yoked	A2	Thick
2509	E114070305A	Yoked	A2	thin
2512	E114070305A	Yoked	A2	Thick
2524	E114070305A	Yoked	AAFA1	thin
2535	E114070905A	Yoked	A2	thin
2545	E114070905A	Yoked	AAF	Thick
2549	E114072904A	Yoked	A2	Thick
2550	E114072904A	Yoked	A2	Thick
2555	E114082904A	Yoked	AAF	thin
2560	E114082904A	Yoked	AAFA1	Thick
2579	E114090104A	Naïve	AAF	Thick
2580	E114090104A	Naïve	AAF	Thick
2581	E114090104A	Naïve	UF	Thick
2584	E114090104A	Naïve	UF	thin
2592	E114090104A	Naïve	AAFA1	Thick
2593	E114090104A	Naïve	AAFA1	Thick
2603	E214101001A	Naïve	AAF	thin
2605	E214101001A	Naïve	AAF	Thick
2607	E214101001A	Naïve	AAF	Thick
2631	E1114100702A	EarlyCocare	AAF	Thick
2647	E214101002A	Naïve	A2	Thick
2654	E214101002A	Naïve	UF	Thick
2655	E214101002A	Naïve	UF	thin
2656	E214101002A	Naïve	UF	Thick
2661	E214092701A	Parent	A2	Thick
2665	E214092701A	Parent	A2	Thick
2666	E214092701A	Parent	A2	Thick
2669	E214092701A	Parent	A2	Thick
2679	E214092704A	Naïve	AAF	Thick
2682	E214092704A	Naïve	UF	Thick
2683	E214092704A	Naïve	UF	thin

**Table A.1** continued

2689	E214092704A	Naïve	AAFA1	Thick
2691	E214111804A	Yoked	A2	Thick
2692	E214111804A	Yoked	A2	Thick
2698	E214111804A	Yoked	UF	Thick
2701	E214111804A	Yoked	UF	Thick
2705	E214111804A	Yoked	UF	Thick
2706	E214111804A	Yoked	UF	thin
2715	E214110902A	Parent	AAF	Thick
2719	E214110902A	Parent	AAFA1	Thick
2726	E214121604A	Yoked	AAF	Thick
2736	E315010201B	Parent	AAF	Thick
2739	E315010201B	Parent	AAF	thin
2750	E315010201B	Parent	UF	Thick
2752	E315010201B	Parent	UF	Thick
2753	E315010201B	Parent	UF	Thick
2754	E315010201B	Parent	UF	thin
2755	E315010201B	Parent	UF	Thick
2757	E315011904A	Yoked	A2	thin
2760	E315011904A	Yoked	A2	thin
2770	E315011904A	Yoked	UF	Thick
2798	E115020904A	Yoked	AAF	Thick
2799	E115020904A	Yoked	AAF	Thick
2800	E115020904A	Yoked	AAF	thin
2801	E115020904A	Yoked	AAF	thin
2807	E115020904A	Yoked	UF	thin
2848	E315011905A	Parent	AAFA1	thin
2849	E315031302A	EarlyCocare	A2	Thick
2851	E315031302A	EarlyCocare	A2	Thick
2852	E315031302A	EarlyCocare	A2	Thick
2858	E315031302A	EarlyCocare	UF	thin
2870	E315022104A	Naïve	AAF	Thick
2871	E315022104A	Naïve	AAF	Thick
2872	E315022104A	Naïve	AAF	Thick
2879	E315022104A	Naïve	UF	Thick
2881	E315022104A	Naïve	UF	thin
2908	E415051203A	Parent	A2	Thick
2910	E415051203A	Parent	UF	Thick
2916	E415060502A	EarlyCocare	A2	thin
2923	E215051602A	Parent	UF	Thick
2929	E515062502A	EarlyCocare	A2	Thick

**Table A.1** continued

2933	E515062502A	EarlyCocare	UF	Thick
2956	E215092602A	Naïve	A2	thin
2960	E215092602A	Naïve	AAFA1	Thick
2965	E215092602A	Naïve	UF	Thick
2967	E215101604A	Naïve	A2	Thick
2970	E215101604A	Naïve	A2	Thick
2971	E215101604A	Naïve	A2	Thick
2972	E215101604A	Naïve	A2	Thick
2977	E215101604A	Naïve	UF	Thick
2991	E715120702B	EarlyCocare	AAF	Thick
2995	E715120702B	EarlyCocare	A2	thin
2996	E715120702B	EarlyCocare	UF	thin
3003	E315112601B	Parent	A2	Thick
3007	E315112601B	Parent	UF	thin
3008	E315112601B	Parent	UF	Thick
3058	E716011401A	Parent	A2	Thick
3061	E716011401A	Parent	A2	Thick
3062	E716011401A	Parent	A2	Thick
3063	E716011401A	Parent	UF	Thick
3065	E716011401A	Parent	UF	thin
3097	E216021401A	Parent	AAF	Thick
3112	E316022501A	Parent	A2	Thick
3126	E516030601A	Parent	A2	Thick
3137	E316032101A	Parent	UF	Thick
3138	E316030901A	Parent	A2	Thick
3139	E316030901A	Parent	A2	Thick
3141	E316030901A	Parent	A2	Thick
3144	E316052904A	Naïve	A2	Thick
3148	E316052904A	Naïve	A2	Thick
3150	E316052904A	Naïve	A2	Thick
3151	E316052904A	Naïve	A2	Thick
3156	E316052904A	Naïve	UF	Thick
3164	E216042902A	Parent	A2	Thick
3170	E416052102A	Parent	A2	Thick
3190	E516063001A	Parent	UF	Thick
3191	E516063001A	Parent	UF	thin

## A.2 List of Units in $A_{fm}$ $f_{fm}$ 8x8 Tuning around Pure Tone

These units were used for the 8x8 tuning around pure tone analysis.

**Table A.2. List of Units in  $A_{fm}$   $f_{fm}$  8x8 Tuning around Pure Tone.** Single unit table columns are as described for **Table A.1**.

SUnitID	AnimalID	Group	Region	SpkType
2993	E715120702B	EarlyCocare	AAF	Thick
2994	E715120702B	EarlyCocare	A2	Thick
2999	E715120702B	EarlyCocare	UF	Thick
3065	E716011401A	Parent	UF	thin
3133	E316032101A	Parent	AAF	thin
3139	E316030901A	Parent	A2	Thick
3148	E316052904A	Naïve	A2	Thick
3166	E216042902A	Parent	UF	Thick
3170	E416052102A	Parent	A2	Thick
3171	E416052102A	Parent	AAF	Thick
3228	E617033003A	Parent	AAF	Thick
3229	E617033003A	Parent	AAF	Thick
3236	E817052604A	Naïve	UF	Thick
3256	E717080205A	Naïve	UF	Thick
3257	E717080205A	Naïve	UF	Thick
3132	E316032101A	Parent	AAF	Thick
3160	E316052904A	Naïve	A1	Thick
3164	E216042902A	Parent	A2	Thick
3007	E315112601B	Parent	UF	thin
3122	E316022501A	Parent	AAF	Thick
3159	E316052904A	Naïve	A1	Thick
3241	E617042302A	Parent	UF	Thick
2995	E715120702B	EarlyCocare	A2	thin
3116	E316022501A	Parent	AAF	Thick
3130	E516030601A	Parent	UF	Thick
3131	E316032101A	Parent	AAF	Thick
3144	E316052904A	Naïve	A2	Thick
3149	E316052904A	Naïve	A2	Thick
3191	E516063001A	Parent	UF	thin
3193	E416080102A	Naïve	AAF	Thick
3210	E516090702A	EarlyCocare	AAF	Thick
3251	E717070405A	Naïve	AAF	Thick
3064	E716011401A	Parent	UF	Thick
3128	E516030601A	Parent	AAFA1	Thick

**Table A.2** continued

3142	E316042504A	Parent	A2	Thick
3151	E316052904A	Naïve	A2	Thick
3195	E416080102A	Naïve	AAF	Thick
3024	E215122601A	Parent	A2	Thick
3119	E316022501A	Parent	UF	Thick
3127	E516030601A	Parent	AAFA1	Thick
3098	E216021401A	Parent	AAF	Thick
3106	E216031304A	Yoked	A2	Thick
3110	E216031304A	Yoked	UF	Thick
3126	E516030601A	Parent	A2	Thick
3189	E516063001A	Parent	UF	Thick
3212	E616120204A	Yoked	A2	Thick
3239	E817052604A	Naïve	AAFA1	Thick
3261	E917080704A	Naïve	UF	Thick

**A.3 List of Units in  $A_{fm}$  Tuning around USV**

These units were used for the  $A_{fm}$  tuning around best responding USV.

**Table A.3. List of Units in  $A_{fm}$  Tuning around USV.** Single unit table columns are as described for **Table A.1**.

SUnitID	AnimalID	Group	Region	SpkType
3144	E316052904A	Naïve	A2	Thick
3151	E316052904A	Naïve	A2	Thick
3207	E516090304A	Yoked	A2	Thick
3207	E516090304A	Yoked	A2	Thick
3256	E717080205A	Naïve	UF	Thick
3248	E917062005A	Naïve	UF	Thick
3248	E917062005A	Naïve	UF	Thick
3249	E917062005A	Naïve	UF	Thick
3274	E917083101A	Naïve	UF	Thick
3272	E917083102A	Naïve	UF	Thick
3164	E216042902A	Parent	A2	Thick
3164	E216042902A	Parent	A2	Thick
3139	E316030901A	Parent	A2	Thick
3139	E316030901A	Parent	A2	Thick
3170	E416052102A	Parent	A2	Thick

**Table A.3** continued

3126	E516030601A	Parent	A2	Thick
3061	E716011401A	Parent	A2	Thick
3061	E716011401A	Parent	A2	Thick
3098	E216021401A	Parent	AAF	Thick
3098	E216021401A	Parent	AAF	Thick
3121	E316022501A	Parent	AAF	Thick
3137	E316032101A	Parent	UF	Thick
3137	E316032101A	Parent	UF	Thick
3231	E617040205A	Parent	UF	Thick
3065	E716011401A	Parent	UF	thin

## Appendix B

### Auditory Stimuli

Auditory stimuli utilized during electrophysiological experiments vary by the experiment conducted as well as the single neuron's responses to stimuli. Below is a list of stimuli used in the work included in this thesis.

#### ***B.1 Pure Tone Tuning Stimulus***

For all single units recorded, a pure tone tuning stimulus was presented to obtain a tuning curve for the unit. Briefly, 40 pure tones logarithmically spaced from 6kHz to 95kHz were presented at a total of six sound intensity levels from 10 – 60 dBa attenuation. Tones were played in pseudorandom order from loudest to softest intensity, with each trial 600ms, and the tone beginning 200ms into the trial and lasting for 60ms. Each individual tone and intensity combination was repeated 5 times for a total of 1200 stimuli per run. A table summarizing the pure tone tuning stimulus presented during recordings is available (**Table B.1**).

**Table B.1.** Pure tone stimulus parameters listing each of 40 pure tone frequencies and attenuation levels. This stimulus was used to generate a frequency response area for each single unit.

Pure Tone Stimulus Parameters		
Dur (ms)	Frequency (Hz)	Atten (dBa)
60	5580	10, 20, 30, 40, 50, 60
60	6001	10, 20, 30, 40, 50, 60
60	6453	10, 20, 30, 40, 50, 60
60	6939	10, 20, 30, 40, 50, 60
60	7462	10, 20, 30, 40, 50, 60
60	8024	10, 20, 30, 40, 50, 60
60	8629	10, 20, 30, 40, 50, 60
60	9279	10, 20, 30, 40, 50, 60
60	9978	10, 20, 30, 40, 50, 60
60	10730	10, 20, 30, 40, 50, 60
60	11539	10, 20, 30, 40, 50, 60
60	12408	10, 20, 30, 40, 50, 60
60	13343	10, 20, 30, 40, 50, 60
60	14349	10, 20, 30, 40, 50, 60
60	15430	10, 20, 30, 40, 50, 60
60	16592	10, 20, 30, 40, 50, 60

**Table B.1** continued

60	17843	10, 20, 30, 40, 50, 60
60	19187	10, 20, 30, 40, 50, 60
60	20633	10, 20, 30, 40, 50, 60
60	22188	10, 20, 30, 40, 50, 60
60	23859	10, 20, 30, 40, 50, 60
60	25657	10, 20, 30, 40, 50, 60
60	27590	10, 20, 30, 40, 50, 60
60	29669	10, 20, 30, 40, 50, 60
60	31905	10, 20, 30, 40, 50, 60
60	34309	10, 20, 30, 40, 50, 60
60	36894	10, 20, 30, 40, 50, 60
60	39674	10, 20, 30, 40, 50, 60
60	42664	10, 20, 30, 40, 50, 60
60	45878	10, 20, 30, 40, 50, 60
60	49335	10, 20, 30, 40, 50, 60
60	53053	10, 20, 30, 40, 50, 60
60	57050	10, 20, 30, 40, 50, 60
60	61349	10, 20, 30, 40, 50, 60
60	65971	10, 20, 30, 40, 50, 60
60	70942	10, 20, 30, 40, 50, 60
60	76288	10, 20, 30, 40, 50, 60
60	82036	10, 20, 30, 40, 50, 60
60	88218	10, 20, 30, 40, 50, 60
60	94865	10, 20, 30, 40, 50, 60

**B.2 Natural Mouse Ultrasonic Vocalization Stimulus**

As long as units are held after the pure tone tuning stimulus completes, a library of natural mouse ultrasonic vocalizations (USVs) was presented. The library consists of n=18 pup USVs and n=18 adult USVs that are matched for basic acoustic properties such as duration, onset frequency, and onset frequency modulation. Calls are played in a pseudorandom order, for a total of n=50 trials each, and randomly interspersed with n=50 silent trials. The library is as described previously (Shepard, Lin, et al., 2015). Each of these calls were fit to an sFM model and six parameters were generated. The parameters for each of the n=36 calls, along with basic acoustic parameter values for each of the calls are listed (**Table B.2**).

**Table B.2.** Natural USVs from the n=36 curated library presented to all SUs. Yellow highlighted columns represent sFM fits to each of the individual calls.

Natural USV Library sFM Parameters and Attributes

Call #	Calltype	Onset Frequency (Hz)	Onset FM Category	dur [ms]	Afm [Hz]	ffm [cyc/sec]	f0 [Hz]	fslope [Hz/sec]	Phi [deg]	fm_start[Hz]
1	Pup	64523	Flat	12	959	90	68141	-412728	96	67187

**Table B.2** continued

2	Pup	66267	Flat	35	997	28	67325	-52399	89	66328
3	Pup	67139	Flat	57	-6784	26	62590	417804	126	68088
4	Pup	73242	Flat	15	416	64	73195	12332	102	72788
5	Pup	70626	Flat	36	-530	43	71043	-155824	165	71179
6	Pup	73242	Flat	55	616	15	75481	-156263	67	74914
7	Pup	79346	Flat	10	-210	121	79196	-61498	2	79202
8	Pup	76730	Flat	34	-573	36	76528	-2901	7	76597
9	Pup	77602	Flat	57	285	27	78810	-147283	49	78596
10	Pup	68882	Sweep	12	686	91	69541	219646	125	68981
11	Pup	68882	Sweep	34	1853	28	71358	49240	94	69510
12	Pup	68011	Sweep	55	-1178	24	70356	41178	45	71194
13	Pup	74986	Sweep	11	1294	94	76137	116240	103	74875
14	Pup	74114	Sweep	38	1758	26	77560	-115348	81	75827
15	Pup	72370	Sweep	56	740	34	75309	-137999	41	74822
16	Pup	79346	Sweep	13	341	73	80287	73726	98	79949
17	Pup	72370	Sweep	35	929	27	83769	-260572	63	82938
18	Pup	77602	Sweep	51	1117	28	79329	-188008	41	78602
19	Adult	69754	Flat	14	-506	73	68960	227800	107	69443
20	Adult	67139	Flat	33	-2171	42	64337	703069	80	66476
21	Adult	68011	Flat	55	6696	21	78557	-102542	104	72064
22	Adult	74114	Flat	12	38	68	74436	-131925	95	74398
23	Adult	74114	Flat	36	1243	30	75365	-183514	146	74675
24	Adult	71498	Flat	47	8622	22	78383	-51143	104	70011
25	Adult	78474	Flat	12	595	81	78931	-204479	96	78339
26	Adult	78474	Flat	32	-2697	41	74809	134112	98	77476
27	Adult	79346	Flat	58	2051	21	82483	-68327	93	80436
28	Adult	69754	Sweep	11	593	88	71132	-67697	77	70554
29	Adult	66267	Sweep	32	3119	33	67695	655368	131	65336
30	Adult	68011	Sweep	59	9381	18	82152	-146645	90	72772
31	Adult	72370	Sweep	14	2157	76	73984	148491	104	71889
32	Adult	71498	Sweep	33	3461	35	73665	338016	151	71979
33	Adult	71498	Sweep	55	3473	25	77882	12875	102	74482
34	Adult	79346	Sweep	14	1885	75	80891	37383	103	79058
35	Adult	78474	Sweep	37	4144	27	84947	-68047	89	80804
36	Adult	81089	Sweep	59	4921	22	88776	-235936	106	84051

### **B.3 $A_{fm}$ and $f_{fm}$ Tuning Stimulus (8x8)**

If a neuron showed excited responses to pure tones and a best frequency could be calculated, then a stimulus that varies  $A_{fm}$  and  $f_{fm}$  in 8 logarithmic steps each spanning the range of frequency modulation parameter values seen in mouse USVs was used, for a total of  $n=64$  stimuli. For this stimulus, the best frequency was taken as the carrier frequency of the sFM stimuli, and the sound intensity level that elicited the strongest response from neurons was used for all stimuli. A total of  $n=25$  trials per stimulus were played. The stimulus parameters are listed in a table for convenience (Table B.3).

**Table B.3.** Values of sFM parameters for 8x8  $A_{fm}$  and  $f_{fm}$  tuning stimulus are around best frequency. Atten (dBa) is taken as the sound level that elicited the best response from the neuron, any value of the six possible played during the pure tone tuning curve: 10, 20, 30, 40, 50, 60dBa. The  $f_0$  value, which is the carrier frequency of the sFM stimulus, is taken to be the best frequency of the neuron as determined from its pure tone tuning curve. The  $f_{fm}$  values are fixed to be eight values of 15, 28, 39, 53, 73, 100, or 137Hz/s regardless of the carrier frequency. The  $A_{fm}$  values are dependent upon the carrier frequency, and are calculated as a fraction of the carrier frequency of 0, 0.004, 0.008, 0.016, 0.030, 0.055, 0.104, 0.194, which correspond approximately to 0, 1/154, 1/82, 1/45, 1/24, 1/13, 1/7, and 1/4 octaves.

8x8  $A_{fm}$  x  $f_{fm}$  Tuning Stimulus around BF

Stim	Atten (dBa)	Dur (ms)	$A_{fm}$ (Fraction of $f_0$ )	$f_{fm}$ (Hz/s)	$f_0$ (Hz)
1	10-60	60	0	15	BF
2	10-60	60	0.004	15	BF
3	10-60	60	0.008	15	BF
4	10-60	60	0.016	15	BF
5	10-60	60	0.030	15	BF
6	10-60	60	0.055	15	BF
7	10-60	60	0.104	15	BF
8	10-60	60	0.194	15	BF
9	10-60	60	0	20	BF
10	10-60	60	0.004	20	BF
11	10-60	60	0.008	20	BF
12	10-60	60	0.016	20	BF
13	10-60	60	0.030	20	BF
14	10-60	60	0.055	20	BF
15	10-60	60	0.104	20	BF
16	10-60	60	0.194	20	BF
17	10-60	60	0	28	BF
18	10-60	60	0.004	28	BF
19	10-60	60	0.008	28	BF
20	10-60	60	0.016	28	BF
21	10-60	60	0.030	28	BF
22	10-60	60	0.055	28	BF
23	10-60	60	0.104	28	BF
24	10-60	60	0.194	28	BF
25	10-60	60	0	39	BF
26	10-60	60	0.004	39	BF
27	10-60	60	0.008	39	BF
28	10-60	60	0.016	39	BF
29	10-60	60	0.030	39	BF
30	10-60	60	0.055	39	BF
31	10-60	60	0.104	39	BF
32	10-60	60	0.194	39	BF
33	10-60	60	0	53	BF
34	10-60	60	0.004	53	BF
35	10-60	60	0.008	53	BF
36	10-60	60	0.016	53	BF
37	10-60	60	0.030	53	BF
38	10-60	60	0.055	53	BF
39	10-60	60	0.104	53	BF
40	10-60	60	0.194	53	BF
41	10-60	60	0	73	BF
42	10-60	60	0.004	73	BF
43	10-60	60	0.008	73	BF
44	10-60	60	0.016	73	BF
45	10-60	60	0.030	73	BF
46	10-60	60	0.055	73	BF
47	10-60	60	0.104	73	BF
48	10-60	60	0.194	73	BF
49	10-60	60	0	100	BF
50	10-60	60	0.004	100	BF
51	10-60	60	0.008	100	BF
52	10-60	60	0.016	100	BF
53	10-60	60	0.030	100	BF
54	10-60	60	0.055	100	BF
55	10-60	60	0.104	100	BF

**Table B.3** continued

56	10-60	60	0.194	100	BF
57	10-60	60	0	137	BF
58	10-60	60	0.004	137	BF
59	10-60	60	0.008	137	BF
60	10-60	60	0.016	137	BF
61	10-60	60	0.030	137	BF
62	10-60	60	0.055	137	BF
63	10-60	60	0.104	137	BF
64	10-60	60	0.194	137	BF

**B.4  $A_{fm}$  and  $f_{fm}$  Tuning Stimulus Matched Bandwidth Noise Control**

For units to which the 8x8  $A_{fm}$  x  $f_{fm}$  stimulus was presented, a subset of units which were still well isolated and did not drift from the electrode after the stimulus completed were then presented with a control stimulus. This stimulus contains n=8 stimuli that are narrowband noise spanning the frequencies equal to the bandwidth of  $A_{fm}$  for the corresponding 8x8 tuning stimulus set. The sound level, center frequency, duration, and noise bandwidth are all matching that of the n=8 steps of  $A_{fm}$  presented in the 8x8 stimulus. No values for  $f_{fm}$  in this stimulus are available, as the stimulus is noise rather than a frequency modulated tone. For this stimulus, n=25 presentations of each stimulus in a pseudorandom order was conducted. Stimulus parameters are summarized in a table (**Table B.4**).

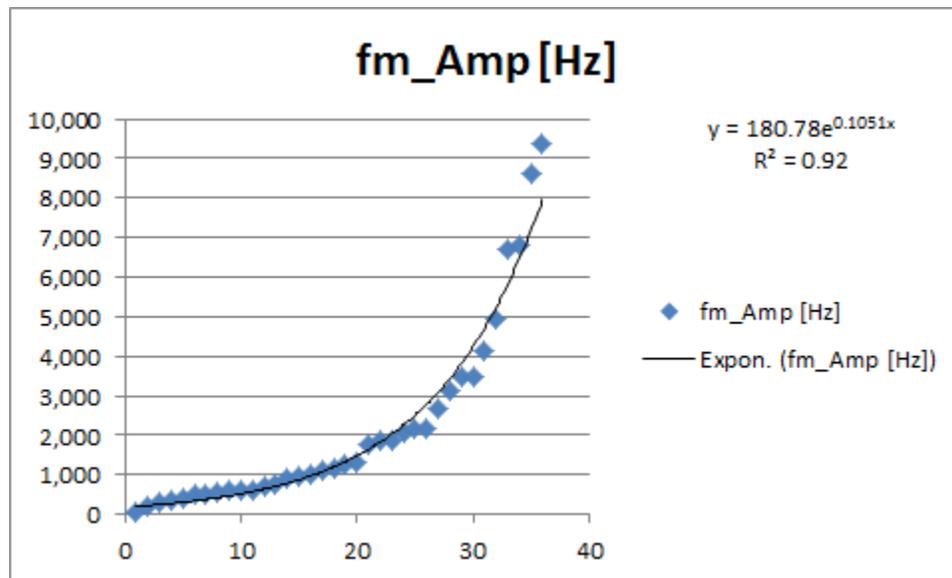
**Table B.4.** Stimulus parameters for the matched bandwidth control stimulus used in tandem with the 8x8  $A_{fm}$  and  $f_{fm}$  tuning stimulus around BF. Noise bandwidths are matched to the  $A_{fm}$  values used in the 8x8 tuning stimulus.

**Matched Bandwidth Noise Control Stimulus**

LineNo	Atten (dBa)	Dur (ms)	Noise_BW (Hz)	f0 (Hz)
1	10-60	60	0	BF
2	10-60	60	0.004	BF
3	10-60	60	0.008	BF
4	10-60	60	0.016	BF
5	10-60	60	0.03	BF
6	10-60	60	0.055	BF
7	10-60	60	0.104	BF
8	10-60	60	0.194	BF

### B.5 $A_{fm}$ Tuning Stimulus around USV

For neurons that showed excited responses to the n=36 library of USVs, an additional  $A_{fm}$  tuning stimulus was presented with sFM parameters centered around the USV that elicited the best response from the neuron via real time analysis. For each call, an  $A_{fm}$  tuning stimulus was constructed ahead of time such that depending on which call a unit responds to, any of the  $A_{fm}$  tuning stimuli could be used on demand. The stimulus consisted of n=20 stimuli, with n=19  $A_{fm}$  logarithmic steps spanning the range found within the library of n=36 calls. Briefly, to obtain individual  $A_{fm}$  steps, the  $A_{fm}$  values for the n=36 calls were plotted along increasing order. An exponential fit function was generated to describe how  $A_{fm}$  varies across the curated call library (**Fig B.1**), and 19 evenly spaced steps across the exponential function were taken as the  $A_{fm}$  values for the tuning stimulus.



**Fig B.1.** Distribution of  $A_{fm}$  values in our curated call library of n=36 calls.  $A_{fm}$  (y-axis) and call number (x-axis), sorted by  $A_{fm}$  showing the exponential distribution of  $A_{fm}$  values. The fit function depicted in the upper right was used to determine the 19 steps of  $A_{fm}$  used in the tuning curve (x:[2:2:38]).

For each stimulus set, the 20<sup>th</sup> stimulus is the original USV around which the  $A_{fm}$  tuning stimulus is centered. For this stimulus, a total of  $n=30$  trials per stimulus were collected. A table summarizing the stimulus parameters is below (**Table B.5**).

**Table B.5.** Parameters in the  $A_{fm}$  tuning stimulus around best USV. Note that duration,  $f_{fm}$ ,  $f_0$ ,  $f_{slope}$ , and Phi values are all equal to the parameter values of the sFM fit to the best USV.

Stim#	dur [ms]	fm_Amp [Hz]	fm_Fre [cyc/sec]	fm_offset [Hz]	fm_slope [Hz/sec]	fm_Phi [deg]
1	dur (BestUSV)	0	ffm (Best USV)	f0 (Best USV)	fslope (Best USV)	Phi (Best USV)
2	dur (BestUSV)	223	ffm (best USV)	f0 (best USV)	fslope (best USV)	Phi (best USV)
3	dur (BestUSV)	275	ffm (best USV)	f0 (best USV)	fslope (best USV)	Phi (best USV)
4	dur (BestUSV)	340	ffm (best USV)	f0 (best USV)	fslope (best USV)	Phi (best USV)
5	dur (BestUSV)	419	ffm (best USV)	f0 (best USV)	fslope (best USV)	Phi (best USV)
6	dur (BestUSV)	517	ffm (best USV)	f0 (best USV)	fslope (best USV)	Phi (best USV)
7	dur (BestUSV)	638	ffm (best USV)	f0 (best USV)	fslope (best USV)	Phi (best USV)
8	dur (BestUSV)	787	ffm (best USV)	f0 (best USV)	fslope (best USV)	Phi (best USV)
9	dur (BestUSV)	972	ffm (best USV)	f0 (best USV)	fslope (best USV)	Phi (best USV)
10	dur (BestUSV)	1199	ffm (best USV)	f0 (best USV)	fslope (best USV)	Phi (best USV)
11	dur (BestUSV)	1479	ffm (best USV)	f0 (best USV)	fslope (best USV)	Phi (best USV)
12	dur (BestUSV)	1825	ffm (best USV)	f0 (best USV)	fslope (best USV)	Phi (best USV)
13	dur (BestUSV)	2252	ffm (best USV)	f0 (best USV)	fslope (best USV)	Phi (best USV)
14	dur (BestUSV)	2779	ffm (best USV)	f0 (best USV)	fslope (best USV)	Phi (best USV)
15	dur (BestUSV)	3429	ffm (best USV)	f0 (best USV)	fslope (best USV)	Phi (best USV)
16	dur (BestUSV)	4231	ffm (best USV)	f0 (best USV)	fslope (best USV)	Phi (best USV)
17	dur (BestUSV)	5221	ffm (best USV)	f0 (best USV)	fslope (best USV)	Phi (best USV)
18	dur (BestUSV)	6443	ffm (best USV)	f0 (best USV)	fslope (best USV)	Phi (best USV)
19	dur (BestUSV)	7950	ffm (best USV)	f0 (best USV)	fslope (best USV)	Phi (best USV)
20	dur (BestUSV)	Afm (BestUSV)	ffm (best USV)	f0 (best USV)	fslope (best USV)	Phi (best USV)

### ***B.6 $A_{fm}$ Matched Bandwidth Noise Stimulus around USV***

The  $A_{fm}$  tuning stimuli may have more complex spectral content than the  $A_{fm} \times f_{fm}$  tuning stimulus centered around BF, given nonzero linear slope values, such that the spectral power of the stimulus is not restricted only within the  $A_{fm}$  bandwidth. As a result, to generate a proper matched bandwidth noise stimulus for this set, the power spectrum of each of the  $n=20$  generated stimuli in the  $A_{fm}$  tuning set was constructed, and a noise stimulus filtered such that it contained the same power spectrum was constructed. To avoid the possibility of patterns within a single randomly generated noise seed, we constructed three exemplars of noise with different random seeds, and matching power spectra. This means for each single stimulus in the  $A_{fm}$  tuning set, three matched bandwidth noise stimuli with matching power spectra, each with different noise randomization were used. The root mean squared values for the noise were matched to the corresponding original stimuli as well. The matched bandwidth noise for each  $A_{fm}$  tuning around best USV set totaled  $n=60$  stimuli per set. Each individual noise exemplar was presented for a total of  $n=10$  trials during playback.

## Appendix C

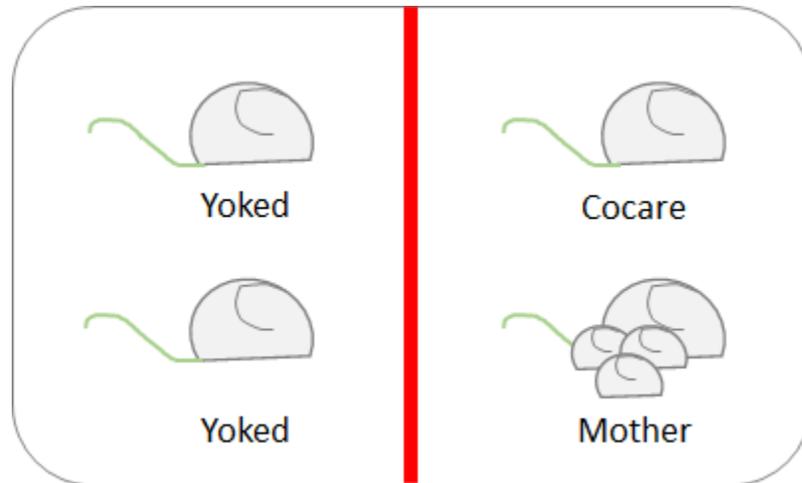
### Yoked Cage Design

Animals display the ability to learn and display novel behavioral responses to new sound cues in a social situation, which can be important to enhance survival. For example, in the mouse maternal paradigm, pup-naïve mice can acquire a preference for approaching pup USVs over a neutral sound when housed for at least 5-7 days with a mother and her litter (Lin et al., 2013). At this early time point, these early cocarers would be able to help retrieve isolated pups that have been separated from the nest. This preference is present acutely but disappears by the time of post-weaning (21 days after birth in late cocarers). We find that both mothers and early cocarers exhibit similar physiological responses to pup USV playback in the auditory cortex, different from completely pup-naïve mice and late cocarers (Shepard, Lin, et al., 2015). However, differences that are seen when comparing pup-naïve mice to early cocarers may arise simply from exposure to pup USVs that early cocarers have had, or perhaps also require direct social interaction with pups. As it currently stands, the two are conflated. We sought to create a social experience control of sound-yoked mice in which pup-naïve mice would be passively exposed to the sounds of a mother and her litter, but would not be allowed to socially interact with the mother and pups.

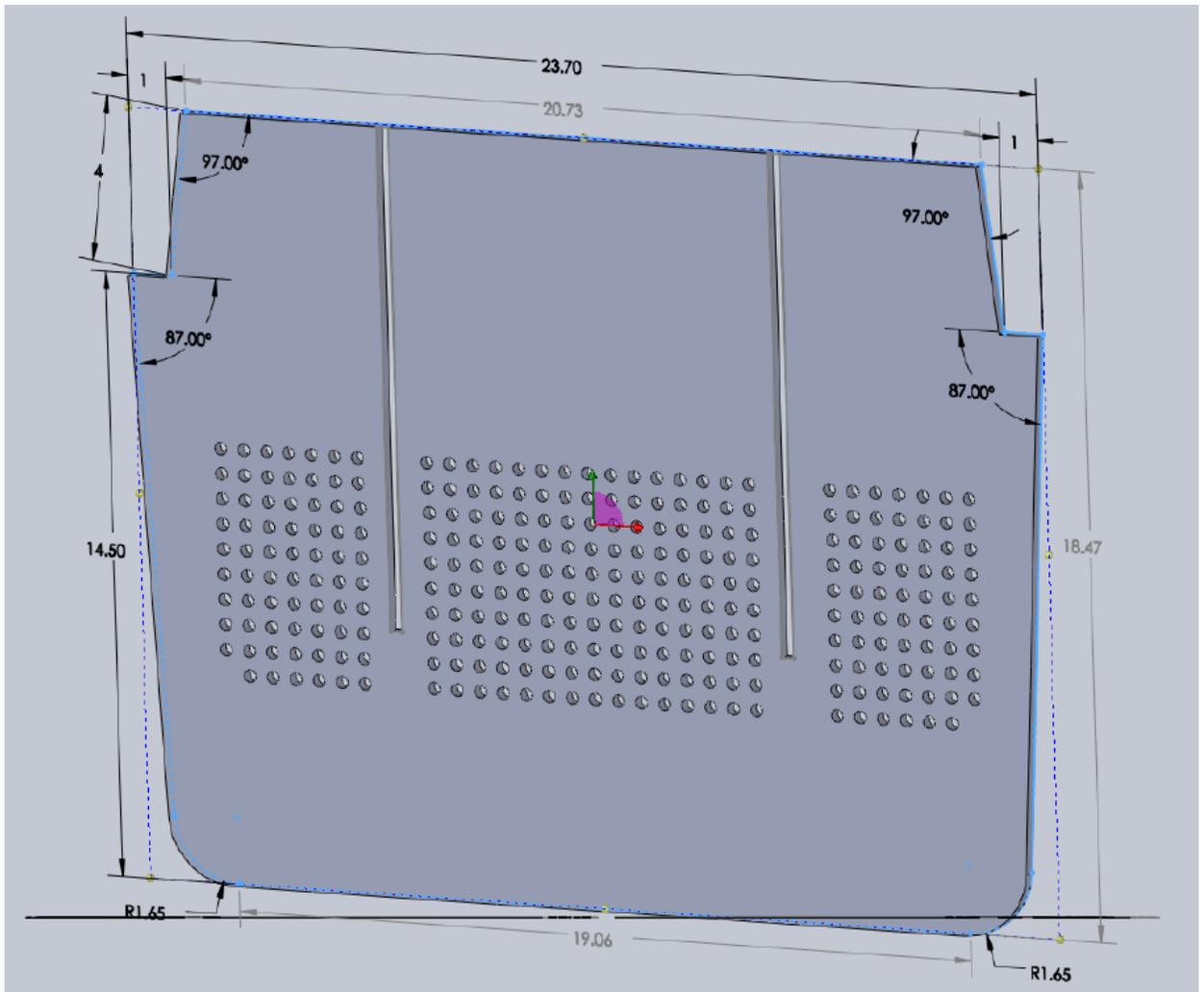
To setup the yoked mice in a cage with a mother and litter without allowing interaction, we created cage dividers with holes that would allow sensory cues to cross between sides without allowing animals to touch and interact. Cage dividers were sized to fit a rat cage, such that enough floor space was given to the mother and litter in addition to the yoked mice in each compartment. Custom-milled dividers were generated (Sabic Polymershapes) using red, opaque colored polycarbonate sheets of 0.3cm

thickness, drafted in Solid Works 2013. Additional divider specifications are listed in the divider specification images (**Fig C.2, C.3, C.4**).

In experiments, a total of four females (siblings) would be used, two of which are housed on the yoked side, one of which is mated and gives birth to a litter (mother), and one of which acts as a cocarer and is housed with the mother (**Fig C.1**).



**Fig C.1. Cartoon schematic of yoked cage setup.** Two pup-naïve females are housed on the left side, and passively experience the auditory and other sensory stimuli from the right side. A cocarer female is housed with a mother and her litter on the right side. Mice are taken for electrophysiological or histology studies at P5-7 when maternal sensitization of the cocarer is expressed. All four adult female mice are siblings.



**Fig C.2. Cage divider specifications, outer** – measurements of outer length in centimeters. Additional cage divider specifications can be found in the folder: Yoking\Cage Divider Design\Cage Divider Specifications revision.pdf, with the Solidworks file in the same directory named: Cage Divider (widthslot).SLDPRT

## Appendix D

### Immunohistochemical Staining

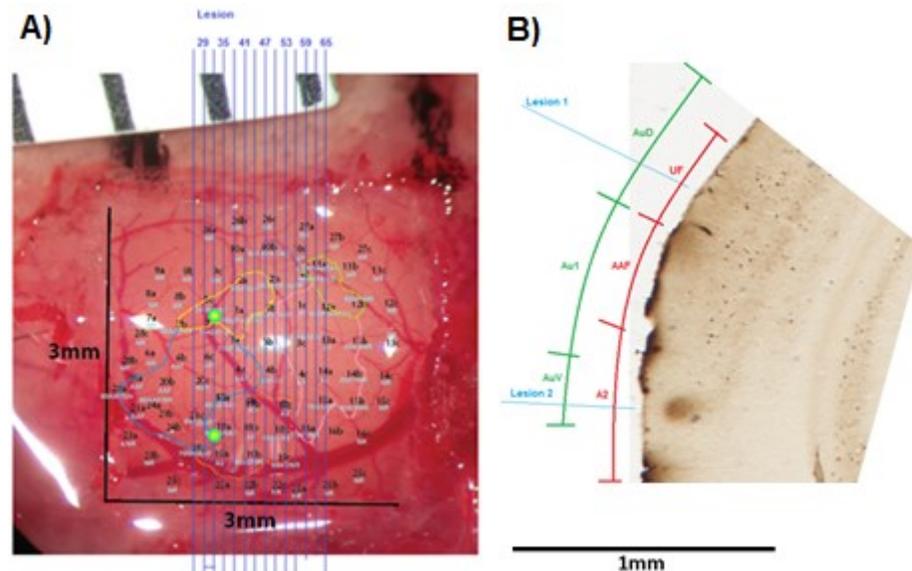
The mouse auditory cortex contains a heterogeneous population of neuronal subtypes. Various interneuron subtypes present within the cortex are thought to play an important role in regulating the excitatory-inhibitory balance, which can shape neural plasticity with experience (Froemke, 2015; Wehr et al., 2003). Inhibitory plasticity has been observed in the lateral band of the mouse auditory cortex after maternal experience, which functionally can enhance the signal to ratio of behaviorally relevant ultrasonic vocalizations by suppressing lower frequency sounds (Lin et al., 2010). Whether this is due to greater activity of putative interneuronal subtypes was of interest.

As a first step, we sought to quantify the relative density of different interneuron subtypes, namely Parvalbumin (PV) and Calbindin (CB)-expressing interneurons. These two neuronal types have been found in the auditory cortex of a different mouse strain (Cruikshank et al., 2001), and we wished to confirm their distribution within the CBA/CaJ strain that we use. Moreover, the results from the prior study utilize the mouse atlas to delineate primary and secondary auditory areas. The correspondence of these histologically defined regions (Au1, AuD, AuV) to areas delineated based on physiological response properties as observed by Stiebler (A1, AAF, UF, dorsoposterior field DP, A2) (Stiebler et al., 1997) has never been systematically quantified.

#### ***D.1 Alignment of Atlas and Physiologically Defined Auditory Regions***

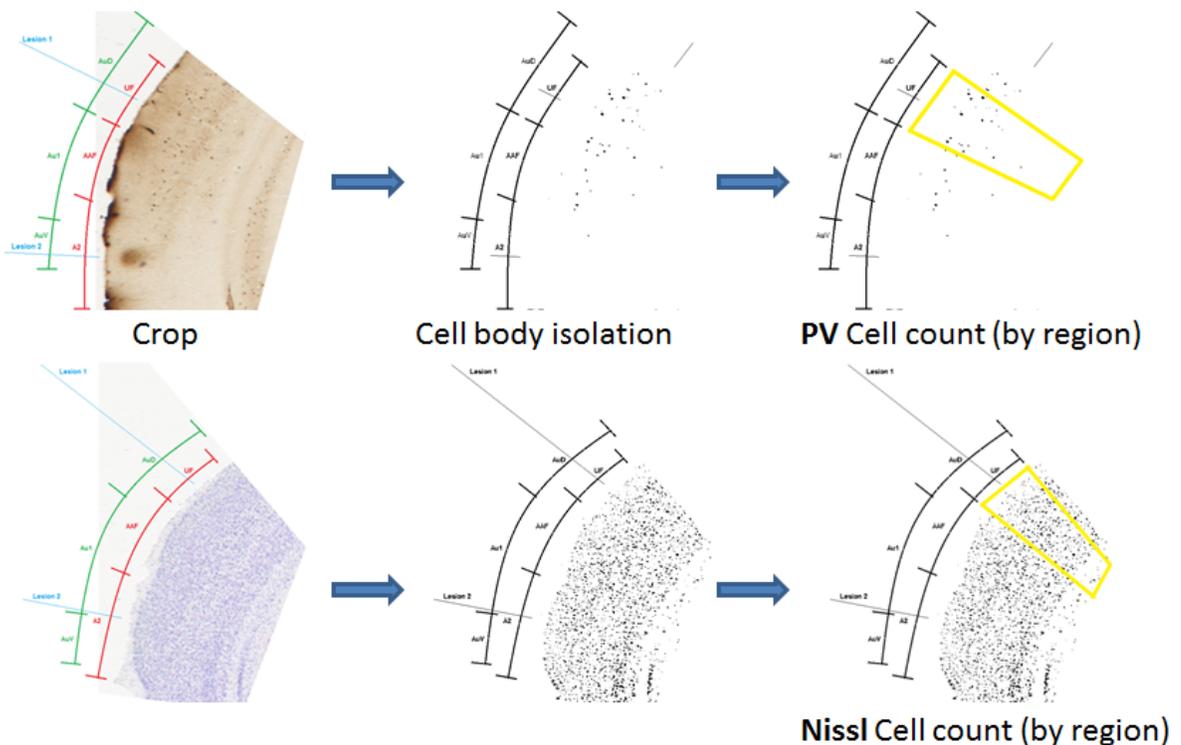
In order to delineate the boundaries of physiologically defined auditory cortical regions on a coronal brain section, before sacrifice, mice underwent surgical craniotomy and tonotopic mapping of their entire left auditory cortex as previously described (Shepard, Liles, et al., 2015). Briefly, after craniotomy, a headpost is affixed to the skull

and multiunit recordings were taken across left auditory cortex using 4 MOhm 3x1 tungsten electrode arrays (FHC) with 305 um interelectrode spacing. Electrodes were driven to layer IV (approximately 400um depth) and pure tones were presented to the right ear across frequencies from 6 to 95kHz. Tuning curves were generated for each electrode penetration, and auditory cortical regions were identified by clear frequency tuning as well as the timing of the PSTH peak. A peak of <40ms from sound onset was taken as core auditory cortex, while latencies >40ms were considered noncore areas. The spatial coordinates of each recording site were tracked using a high-resolution photo of the brain surface. Two lesions were made after recording was complete, with one in the core auditory area and one in the noncore auditory area (**Figure D.1**).



**Figure D.1: Craniotomy Image with Electrode Penetrations and Histological section.** **A)** Cortical surface with electrode penetrations are shown with black lettering. Green dots denote the placement of electric lesions. Vertical blue lines denote the relative position of stained coronal sections. Colored lines show the relative placement of each individual auditory area as defined through electrophysiological mapping. **B)** Coronal section stained for parvalbumin with the lesion positions demarcated with blue lines. Red lines indicate electrophysiological auditory cortical regions. Green lines indicate mouse atlas regions obtained by overlaying the corresponding mouse atlas coronal image over the whole-brain image of the stained tissue. The relative sizing of the scale bar on the cortical surface map and the coronal brain section were used to delineate the boundaries of individual auditory brain regions.

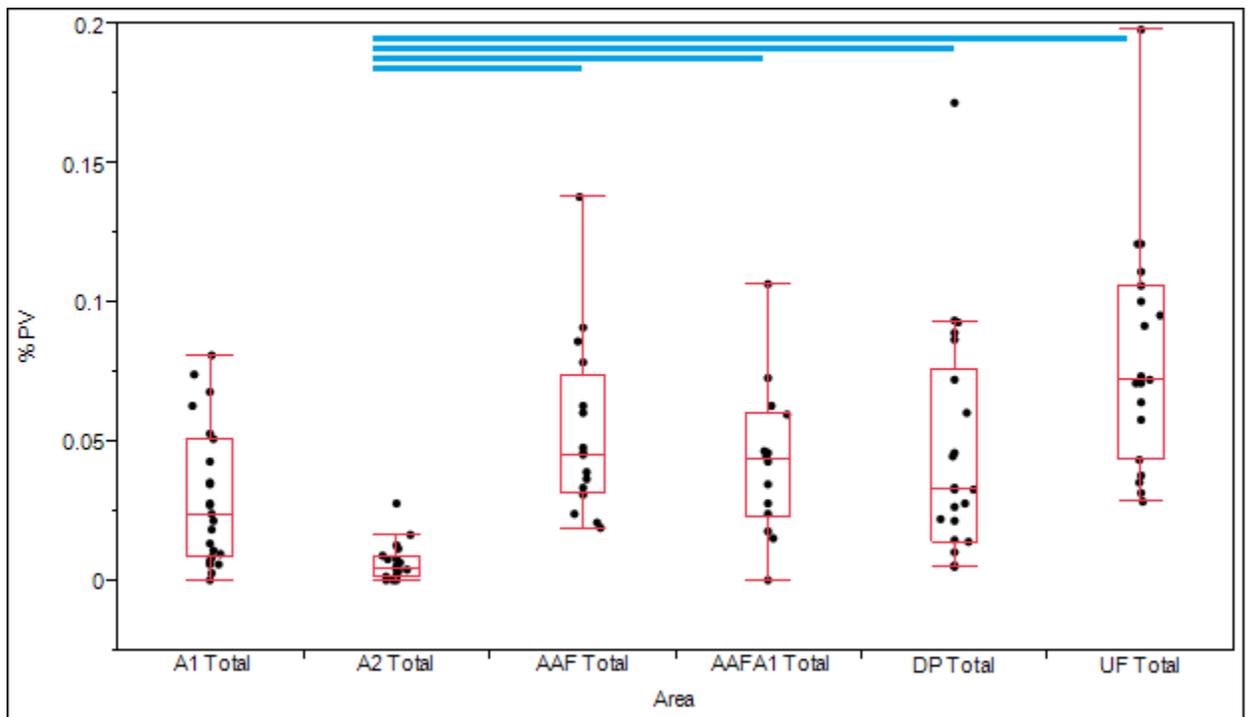
For each animal, the brain was perfused with 4% paraformaldehyde, and tissue was suspended overnight in paraformaldehyde before being transferred to a solution of 30% sucrose. After 24 hours in sucrose, the brain is then sectioned into 40um thick sections on a microtome, and all sections spanning the auditory cortex were saved. An alternating series of three stains were conducted on each brain for a total of n=4 animals, with every third section stained for PV, CB, and Nissl staining respectively. Nissl staining was used to obtain a total cell body count in the relative area, while cell counts for bodies stained for PV and CB across the same areas were used to calculate a percentage of PV / CB-stained cells per section (**Fig D.2**).



**Figure D.2: Cell Counting and Quantification methods.** For each section, auditory cortical regions as denoted by histology (green) as well as physiology (red) are demarcated, and lesions are located by looking through all tissue sections and using surface and hippocampal structures as landmarks. For calculation of % PV or CB stained cells, each PV-stained section has a paired nearest Nissl-stained section. The number of cell bodies expressing PV/CB or stained with the Nissl stain is quantified in ImageJ on a per-region basis, per-section. The number of PV or CB cells is then divided by the total number of Nissl-stained cells to obtain an approximate percentage of PV/CB cells within each auditory area on a given section.

## D.2 Parvalbumin Expression in Secondary Auditory Cortex

Overall, we find that PV expression is significantly lower in the ventral secondary auditory cortex A2 (**Fig D.3**), which agrees with previously reported findings in a different mouse strain (Cruikshank et al., 2001). However, we do not find any significant differences in the expression of CB across auditory cortical areas (not shown). We did find that CB expression is relatively confined to the more shallow cortical layers (2/3), which agrees with previously reported findings in Cruikshank's work.



**Figure D.3: Parvalbumin expression across Physiologically Defined Auditory Cortical Regions.** The percentage of PV expressing cells, as calculated by the total PV expressing cell count divided by the total Nissl cell count for the corresponding section and auditory region are shown. Each dot represents a section, with multiple sections per animal shown. Blue lines indicate groups that are statistically different from one another using the non-parametric post-hoc Steel Dwass test. A2 is significantly lower in PV expression compared to core areas AAF, DP, UF, and the high frequency reversal area between AAF and A1.

Interestingly, A2 does not show a very high preponderance of PV interneurons, despite being one of the most common types of interneurons in the cortex (up to 40% of

the total population) (Butt et al., 2005). PV neurons are capable of shaping cortical responses of pyramidal neurons in the visual cortex (Atallah et al., 2012), and plasticity in PV networks in the hippocampus as well as in the barrel cortex play an important role in learning (Donato et al., 2013; Nowicka et al., 2009). However, as A2 PV expression is relatively sparse, changes in A2 and the plasticity mechanisms involved likely are independent from the PV interneuron subtype.

Taken with the results regarding frequency trajectory tuning in maternal A2, the changes we see in A2 neural Offset responses to  $A_{fm}$  tuning are likely from a mechanism that does not involve the PV interneuron subtype. Although other interneuronal subtypes may be involved, such as CB, or others that we did not directly investigate such as somatostatin (SST) and vasoactive intestinal peptide (VIP), the observation that there were no significant changes in On or Offset response spike rate or prevalence in the thin-spiking (putative interneuron) subpopulation means that physiological changes in the interneuronal subtypes is not easily visible on the population level.

### D.3 Parvalbumin Staining Protocol

Parvalbumin IHC Protocol:

#### Day 1

1. Rinse 4 X 5 min in 0.5M Tris, pH7.5  
For 1L of 0.5M Tris, add:
  - 63.5g Trizma-Hydrochloride
  - 11.8g Trizma Base
  - Add to 1L of ddH<sub>2</sub>O
2. Incubate sections in blocking buffer for 1h @ RT  
For 30mL of blocking buffer, add:
  - 28.5mL 0.5M Tris
  - 0.5% Triton-X (1.5mL)
  - 5% NHS (7 drops, Vectastain ABC Kit, PK-4002)
3. Blot, then Incubate at RT with light agitation overnight in primary antibody:  
For 30mL of primary antibody, add:
  - 30mL 0.5M Tris
  - 7.5 ul PV antibody (Swant Mouse Monoclonal Ab PV235, 1:4000 dilution)  
**Note: Calbindin Staining followed the same protocol, with the exception of using a different antibody: 7.5 ul CB antibody (Sigma Aldrich C9848 CB-955, 1:4000 dilution)**
  - 2% NHS (3 drops)
  - 0.1% Triton-X (300ul)

#### Day 2

4. Blot then rinse 4 X 5 min in 0.5M Tris.
5. Incubate in biotinylated secondary antibody in 0.5M Tris with 2% NHS for 1h @ RT  
For 30mL of secondary antibody, add:
  - 30mL 0.5M Tris
  - 2% NHS (3 drops)
  - Vectastain Anti-Mouse IgG (3 drops)
6. Make ABC reagent 10 minutes before secondary antibody incubation is completed.  
For 30mL of ABC reagent, add:
  - 30mL 0.5M Tris
  - Reagent A (6 drops)
  - Reagent B (6 drops)
7. Blot and rinse 4 X 5 min in 0.5M Tris.
8. Incubate in ABC reagent for 1-2h @ RT (1 hour is usually sufficient)
9. Make DAB reagent 10 minutes before ABC incubation is completed. **NOTE: Prepare DAB under the hood. DAB reagent is highly carcinogenic and should be handled with caution. Gloves that come into contact with DAB should be placed into a biohazard bag and disposed of as biohazardous waste.**  
For 30mL of DAB reagent, add the following under the hood:
  - 29mL 0.5M Tris
  - 24 drops DAB (Vector Laboratories, SK-4100)
  - 12 drops Nickel Solution (SK-4100)
  - 12 drops H<sub>2</sub>O<sub>2</sub> (SK-4100) – Use caution, can cause burns
10. Rinse 4 X 5 min in 0.5M Tris.
11. Move the orbital shaking platform under the hood, and incubate in DAB watching closely (5-15min). Tissue sections should turn brown.
12. When staining reaches a desired appearance, Rinse 4 X 5 min in 0.5M Tris buffer.
  - All DAB trays and glassware that touch DAB need to be bleached with 5% bleach, and then placed in a DAB waste bottle located in the hood. Afterwards, glassware should be thoroughly rinsed in tap water followed by dH<sub>2</sub>O.
13. Mount from 0.5M Tris Buffer onto gelatinized slides; air dry overnight and coverslip.

## **Appendix E**

### **Human Psychophysics**

#### ***E.1 Introduction***

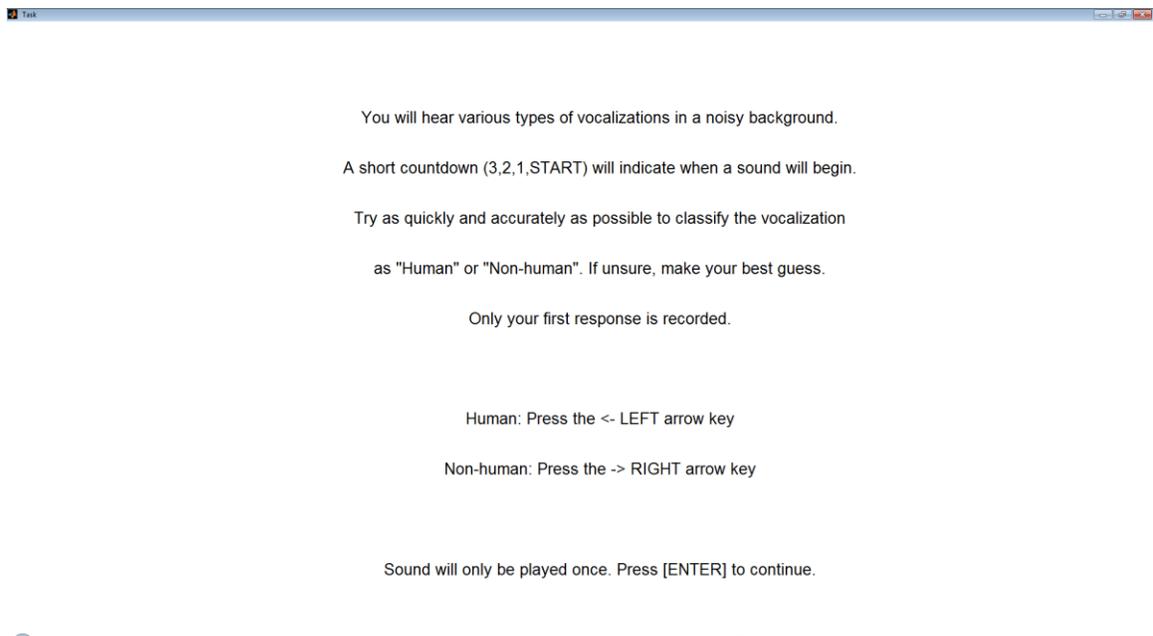
We sought to verify findings from electrophysiological studies in the mouse auditory cortex in humans. We found that in mice, inhibitory plasticity is present that suppresses background acoustic features to enhance the relative signal to noise ratio (SNR) of a behaviorally relevant vocalization. Specifically, in mouse mothers that have learned the meaning of pup ultrasonic vocalizations, we find increased inhibition of auditory areas tuned to lower frequencies that are not present in the ultrasonic pup calls (Galindo-Leon et al., 2009).

In humans, it is known that infant cries are a recognizable and behaviorally relevant auditory stimulus that can elicit parenting behavior and induce physiological responses such as the maternal letdown reflex (Mead et al., 1967). We investigated whether the process of learning to recognize vocalizations led to any perceptual changes in how these vocalizations are processed. Previous studies have shown changes in the processing of familiar sounds by showing that subjective loudness of a background noise is decreased for sentences read in familiar voices compared to novel voices (Goldinger et al., 1999). This mechanism may be similar to the increased suppression of off-target acoustic responses seen in mice. In order to assess detectability of infant cries, vocalizations can be played at varying levels of attenuation against background noise with varying acoustical properties. We will conduct psychophysical testing to determine the subjective background noise loudness as well as assessing whether the subject identifies the vocalization as familiar.

#### ***E.2 Experimental design***

We obtained a library of vocalizations from the OxVoc database, consisting of 173 non-verbal sounds from infants, adults, and animals expressing a range of emotional states (Parsons et al., 2014). We assessed infant cry and adult human vocalization detectability and discriminability against a white noise background of varying SNR of +18dB, +3dB, and -3dB. The task itself was designed with two phases. The first phase asked subjects to classify sounds as “human” or “non-human” and consisted of half of the sounds in the database. The second phase asked subjects to rate the loudness of the background noise on a scale of 1 (soft) to 5 (loud), while at the same time reporting whether the sound is familiar (was heard in the first phase) or unfamiliar (novel). Subjects were given separate prompts for task 1 (**Fig E.1A**), and task 2 (**Fig E.1B**), and were given break time between two tasks for as long as the subject chose. A pilot set of n=10 subjects were recruited, in which all were healthy adult participants who did not have children. Of subjects, n=7 were female, and n=3 were male.

## A) Task #1



# START

Human: <- [LEFT] Arrow Key

Non-human: -> [RIGHT] Arrow Key

## B) Task #2

### Switching Tasks!

This next set of tasks will be a familiarity rating task. Similar vocalizations will be played with background noise of varying loudness.

You will be asked to classify vocalizations as familiar (in first task) or unfamiliar, and to rate the BACKGROUND noise on a scale of: 1 (softest) to 5 (loudest).

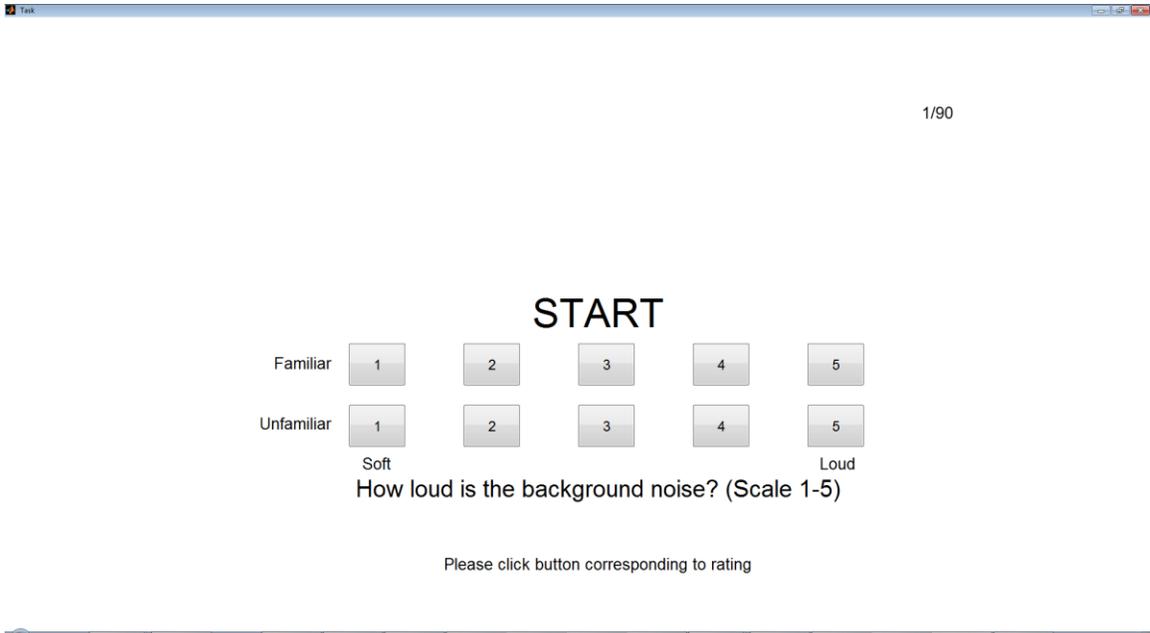
An example of what you will use to respond is below:

Familiar	<input type="button" value="1"/>	<input type="button" value="2"/>	<input type="button" value="3"/>	<input type="button" value="4"/>	<input type="button" value="5"/>
Unfamiliar	<input type="button" value="1"/>	<input type="button" value="2"/>	<input type="button" value="3"/>	<input type="button" value="4"/>	<input type="button" value="5"/>
	Soft				Loud

**How loud is the background noise?**

Use the mouse to click the corresponding button on screen to make your choice.

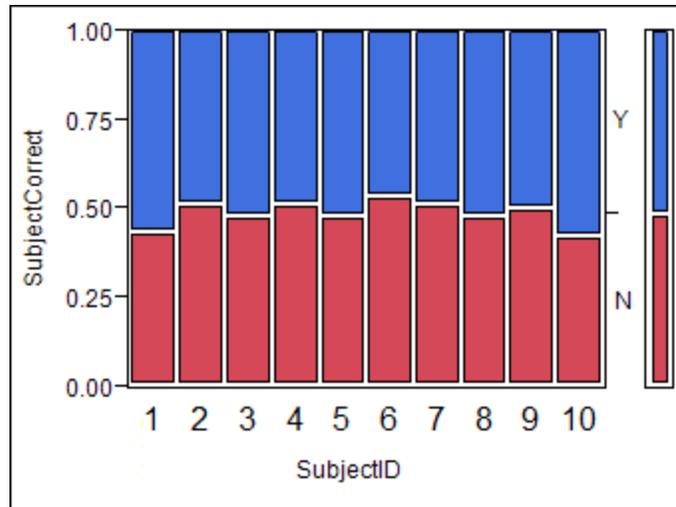
Sound will only be played once. Press [ENTER] to continue.



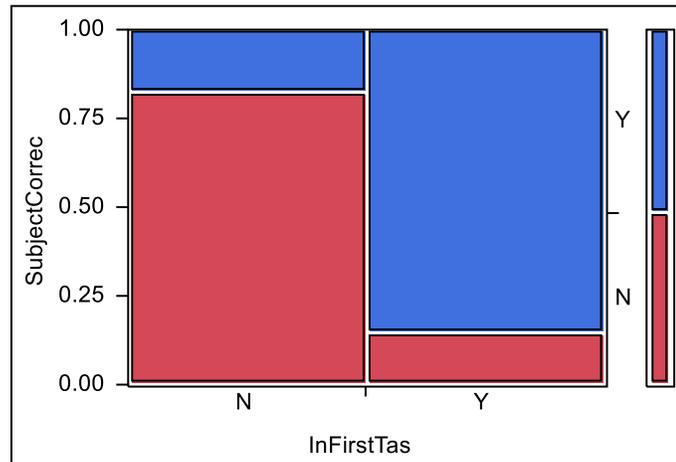
**Figure E.1.** Sample images of the task administered to subjects. **A)** Task #1, in which subjects were asked to classify a library of non-speech vocalizations as human or non-human with a corresponding left or right arrow keystroke. **B)** Task #2, in which subjects were asked to rate the loudness of background noise on a scale from 1 – 5, as well as reporting whether the sound was familiar (presented in the first task) or unfamiliar.

### ***E.3 Results***

Overall, we found that subjects had trouble correctly classifying sounds correctly as familiar during the second phase if they were presented in the first phase (**Fig E.2**). Overall, the proportion of correct responses did not differ significantly from chance performance. We found that subjects' poor performance were attributed to a high number of false positives, in which subjects would rate a novel sound in the second phase as "Familiar" (**Fig E.3**).



**Fig E.2.** Proportion of Vocalizations Correctly Classified as Familiar if presented in Phase 1. Subjects in this pilot of  $n=10$  did not differ significantly from chance (0.5) performance.

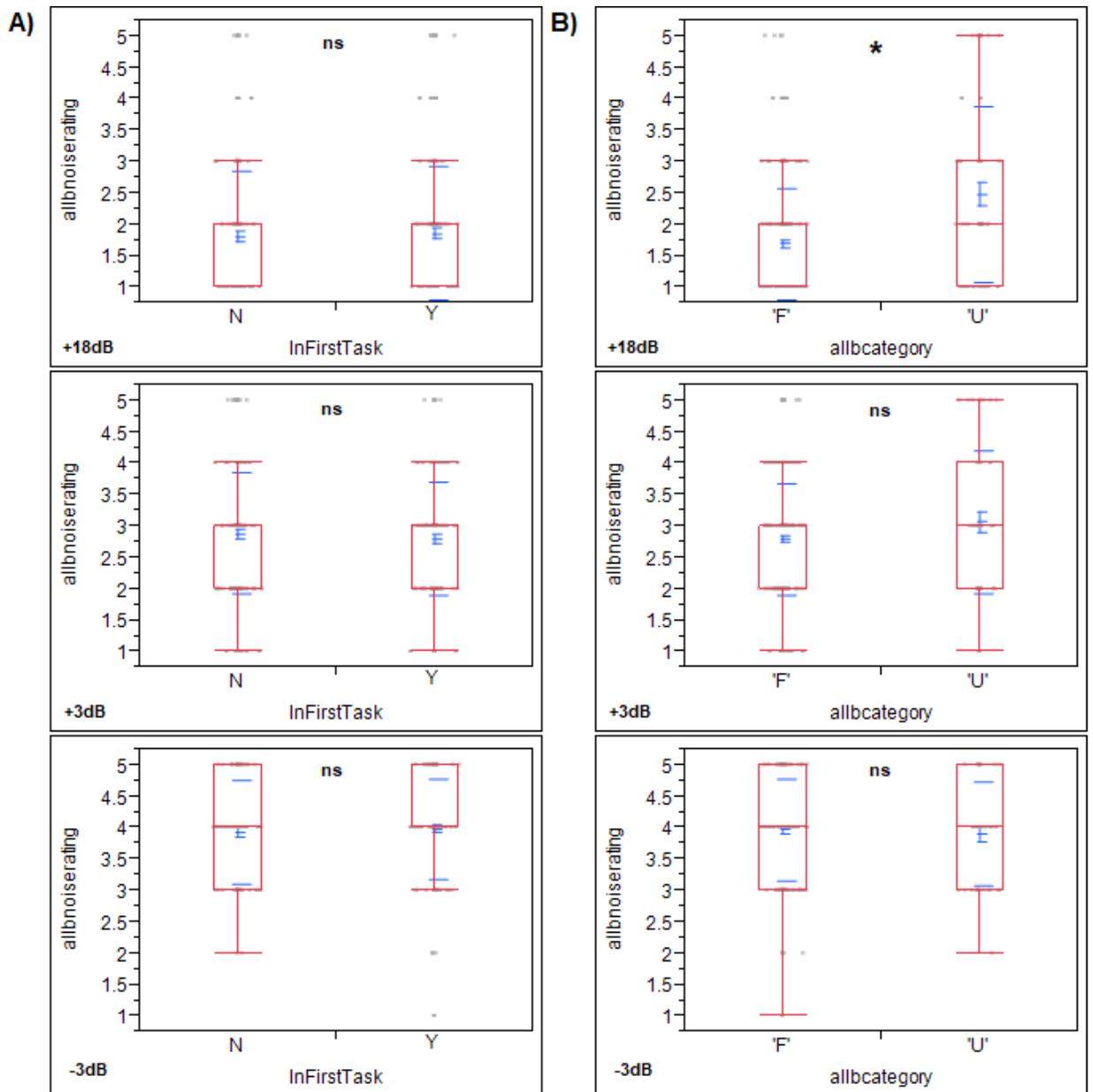


**Fig E.3.** High Proportion of Vocalizations Correctly Classified as Familiar. Proportion of correct classifications are high if the sound was presented in the first task (Y), meaning the number of true positives was high. The proportion of correct classifications was very low if the sound was NOT presented in the first task (N), which means the number of false positives was also high. Subjects were likely to call vocalizations familiar regardless of whether they were in the first task or not.

We then examined whether the subjective noise ratings differed based on whether a vocalization was presented in the first phase or not. We found that regardless of the SNR level (+18dB, +3dB, or -3dB), loudness ratings did not significantly differ between sounds presented in the first phase or compared to novel sounds (**Fig E.4A**). However, if the subject classified the sound as “Familiar”, for the softest noise level (+18dB SNR),

subjects would tend to rate “Familiar” sounds as having lower background noise than “Unfamiliar” sounds (**Fig E.4B**).

Overall, we find that there is an effect of the subject’s own reported familiarity with the sound rather than the ground truth of whether the sound was previously presented to the subject before. This suggests there may be top-down influences on the perceived background noise level, and that these effects are only pronounced when the background noise is relatively quiet.



**Fig E.4.** Subjects' Noise Loudness Ratings on a scale of 1 to 5. **A)** Ratings divided by whether the sound was presented in the first phase of the task, split by the vocalization SNR (+18dB: Top; +3dB: Middle; -3dB: Bottom). All comparisons are not significant. **B)** Ratings divided by whether the subject classified the vocalization as Familiar (F) or Unfamiliar (U), split by vocalization SNR (+18dB: Top; +3dB: Middle; -3dB: Bottom). For the softest noise level, noise loudness ratings are significantly lower for familiar rated vocalizations ( $p < 0.0001$ , Wilcoxon Rank Sum).

## Appendix F

### Stimulus Specific Adaptation (SSA)

The phenomenon of stimulus-specific adaptation (SSA) is when the spike rate of a sensory neuron to a stimulus tends to decrease if the stimulus is presented repeatedly (Ulanovsky et al., 2003). This has been conducted with pure tone stimuli, in which tones of a certain frequency are presented with high probability, while a deviant tone that has a small likelihood of being presented is interspersed randomly amongst common frequency presentations. Reports have shown that auditory cortical neurons in A1 will show increased activity for the deviant tone presentations (Taaseh et al., 2011) under anesthesia.

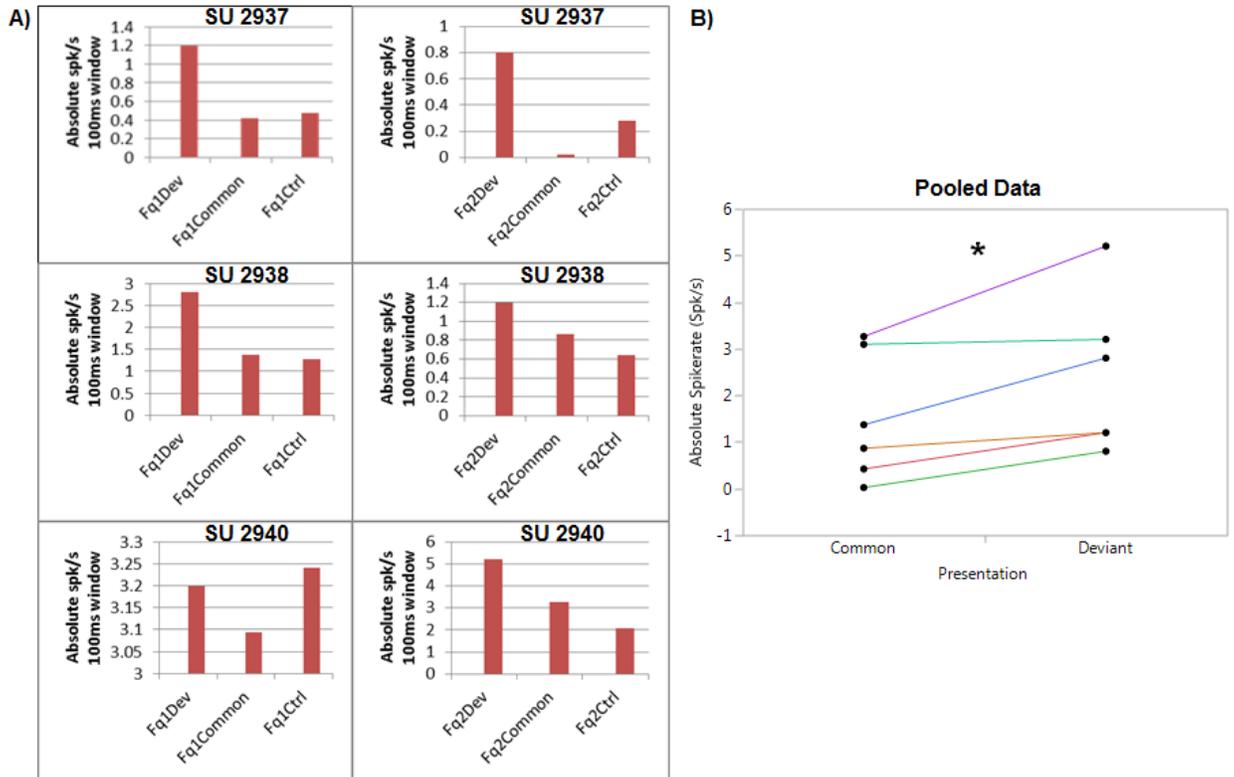
We sought to replicate these findings in the awake head-fixed paradigm, as well as extend these findings to the possibility of SSA generalizing across sounds that fall in the same category. This will be referred to as category-specific adaptation (CSA), or adaptation to sounds falling within the same category.

#### ***F.1 Verifying SSA in Awake Auditory Cortex***

First, to replicate SSA in our CBA/CaJ mouse strain, briefly, headpost attachment and small hole craniotomy was conducted over the auditory cortex. The mouse was given 24 hours of recovery time before the SSA recording session. During recording, the awake mouse was placed into a restraint and head-fixed in an anechoic chamber with a speaker positioned 11cm from its right ear. A single 6M $\Omega$  tungsten electrode was inserted into the auditory cortex in either a Core area (Ultrasound Field UF) or a Noncore area (Secondary Auditory Cortex A2). For each isolated single unit, a tuning stimulus was presented with 60ms tones logarithmically spaced in 40 steps from 6kHz to 95kHz. A tuning curve was generated to determine the neuron's best frequency at 20dB SPL.

If the neuron is responsive to tones, then a second SSA stimulus set was presented. The SSA stimulus set consists of two pure tones (freq 1 and freq 2) with frequencies close to the neuron's best frequency. Three stimulus configurations were presented: 1) Freq 1/Freq 2 presentation ratio of 25/475, where Freq 1 is Deviant, 2) Freq 1/Freq 2 ratio of 475/25, where Freq 2 is Deviant, 3) Freq 1/Freq 2 ratio of 250/250, where both tones are equally likely. In all cases, tones were 60ms in duration, with an inter-stimulus interval of 300ms. Absolute spike rates to each of the stimuli were calculated and a peristimulus time histogram (PSTH) was generated for each set.

In this pilot test to validate the presence of SSA, we collected n=1 UF neuron and n=2 A2 neurons from one animal. Out of the three neurons recorded, all neurons showed SSA, or an increased absolute spike rate to the deviant compared to the common tone (**Fig F.1**). We were able to confirm that we observe SSA in auditory cortical neurons in the awake mouse auditory cortex.



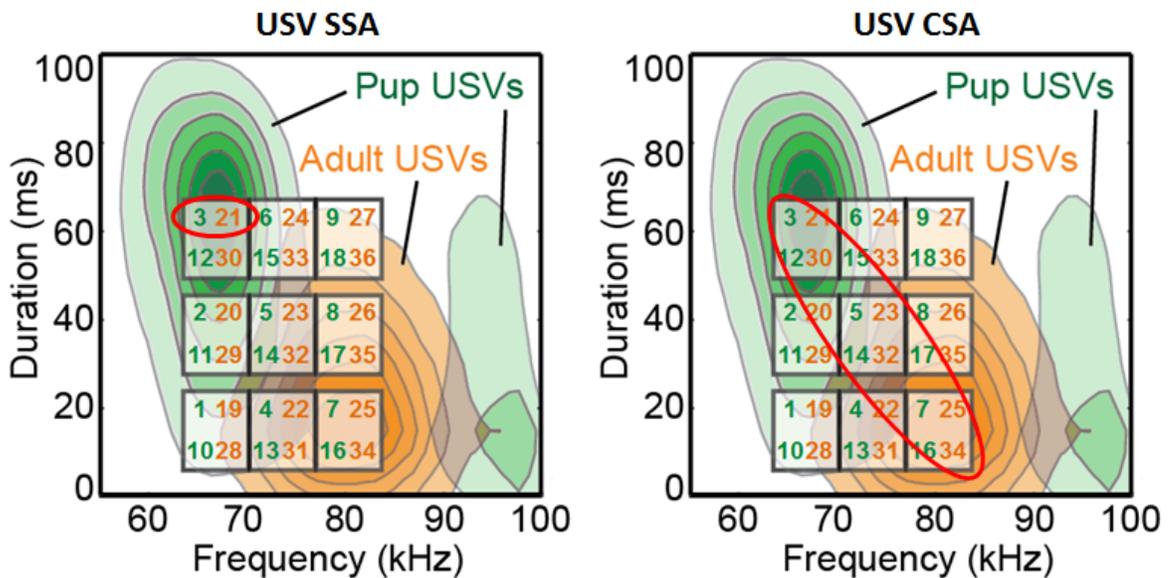
**Fig F.1.** Units spike more when a tone is deviant compared to when it is common. **A)** Three units' individual absolute evoked spike rates for both Freq1 (Fq1) and Freq2 (Fq2) are shown across the three conditions of when the tone is Deviant (Dev), Common (Common), or equally presented (Ctrl). **B)** Pooling all data, when a tone is a deviant, the absolute spike rate is significantly higher compared to when the tone is common across the six tones presented ( $p < 0.05$  paired Wilcoxon signed rank).

## F.2 Investigating the Presence of Category Specific Adaptation (CSA)

For this set of experiments, the same surgical procedures as was performed for SSA were conducted. Two different types of stimulus were tested. First, we designed a stimulus that more closely resembled SSA, although instead of using pure tones, mouse USVs were used. A pup call (Call #3) and adult call (Call #21), which are matched for onset frequency and duration were used for the stimulus (**Fig F.2**). Three stimuli were constructed: 1) Pup#3 / Adult#21 presentation ratio of 25/475, where pup call #3 is Deviant, 2) Pup#3 / Adult#21 presentation ratio of 475/25, where adult call #21 is

Deviant, 3) Pup#3 / Adult#21 presentation ratio of 250/250, where both calls are equally likely.

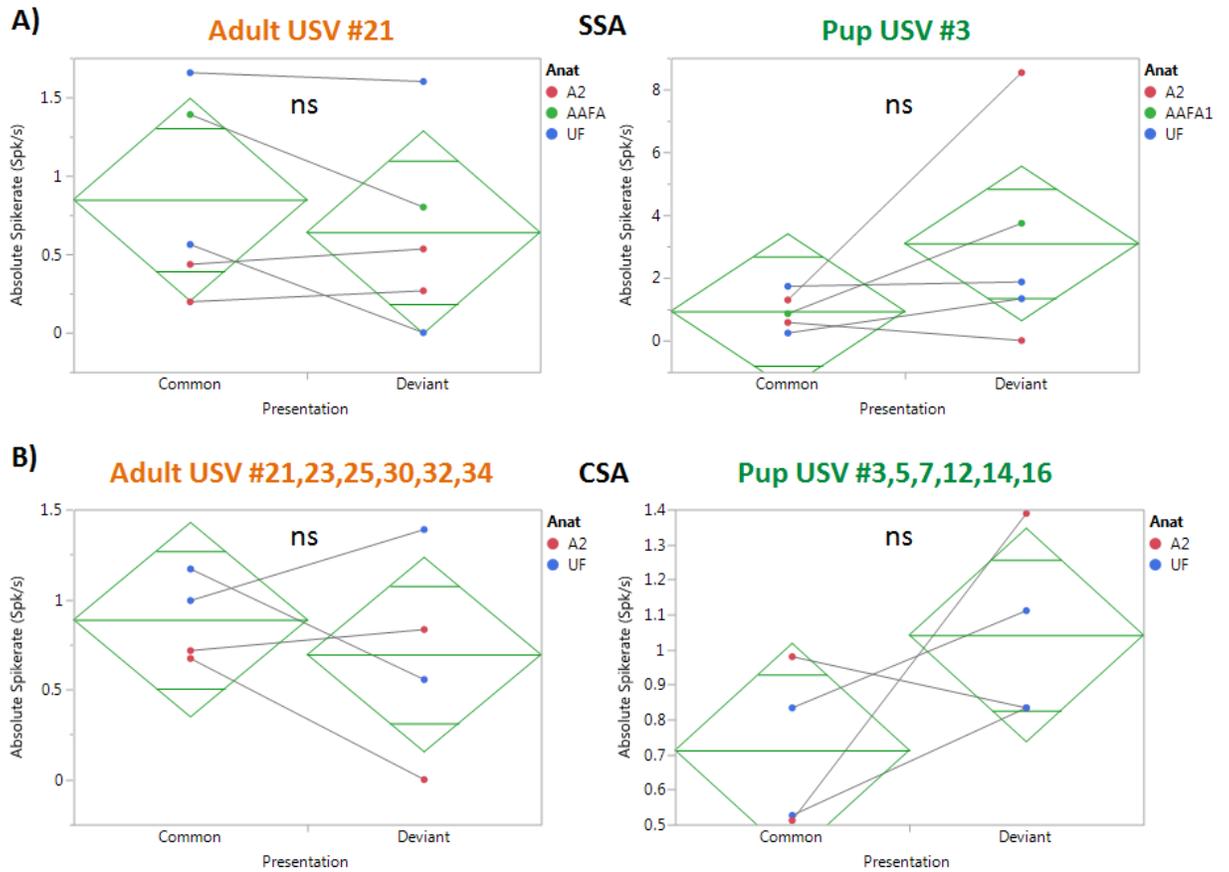
A second CSA stimulus set was also presented, in which a set of 6 pup calls (#3, 5, 7, 12, 14, 16) comprised the pup call category, and a set of 6 adult calls (#21, 23, 25, 30, 32, 34) comprised the adult call category. Here, we constructed three stimuli sets in a similar fashion to the SSA stimulus: 1) Pup/Adult presentation ratio of 24/456, where pup calls are deviant, 2) Pup/Adult presentation ratio of 456/24, where adult calls are deviant, 3) The entire library of n=36 calls was presented at equal ratio of 900/900 pup and adult calls. Each individual deviant call was presented 4 times ( $4 \times 6 = 24$ ), while each common call was presented 76 times ( $76 \times 6 = 456$ ), totaling 480 stimuli for the first two sets.



**Fig F.2.** Mouse USVs Selected for stimulus generation. Left: An SSA stimulus was constructed using a pup USV (#3) and adult USV (#21) instead of using pure tones. The calls are matched for onset frequency and duration, and are approx. ~60ms in duration. Right: A category-specific adaptation (CSA) stimulus was constructed using a set of six pup calls (3, 5, 7, 12, 14, 16) and six adult calls (21, 23, 25, 30, 32, 34). The entire category of calls would either be “deviant” or “common”, and are also equally presented in a control stimulus.

We collected n=5 units (AAFA1=1, UF=2, A2=2) from two pup-naïve animals. We found that for the SSA stimulus, we do not see any SSA for pup or adult calls whether they are the common or deviant sound, although the result is trending towards higher spike rates seen in pup calls (**Fig F.3A**). When using the CSA stimulus, we also do not see any significant differences between when a category is deviant or common, although again the result is trending for pup calls and not adult calls (**Fig F.3B**). Note that one of the units that SSA was conducted on was lost shortly after SSA recording was completed, and was not included in the CSA analysis.

Overall, it appears that the SSA observed for pure tones in auditory cortex does not generalize to natural USVs, either when they are presented as a deviant single exemplar or when an entire category of calls is presented as deviant in our pilot study.



**Fig F.3.** No SSA or CSA seen for Mouse USVs. **A)** Stimulus Specific Adaptation using USVs: When only a single Adult USV exemplar (#21) and Pup USV exemplar (#3) are presented, we see no significant changes in the neural response to these calls when they are presented as a deviant or common stimulus (ns, paired Wilcoxon signed rank). Results are trending for the pup USVs ( $p=0.18$ ) but not significant. **B)** Category Specific Adaptation using USVs: When six exemplars of Adult USVs and six exemplars of Pup USVs are presented, we see no significant changes in the neural response to these calls when they are presented as deviant or common categories (ns, paired Wilcoxon signed rank). Results are again trending for pup USVs but not significant ( $p=0.25$ ).

## Appendix G

### Voltage Sensitive Dye and Protein Imaging

Voltage sensitive dyes (VSDs) offer a method of looking at membrane potentials across an entire cortical surface. When looking at specific cell subtypes in the auditory cortex, the sparse distribution of certain cell types renders conventional electrophysiological techniques such as single/multiunit, local field potential, or patch clamp recording impractical. This is due to limitations on the number and density of electrode penetrations to maintain tissue viability (Moore et al., 2013; Wang et al., 2012). VSD imaging allows for real-time optical imaging of activity of an entire cortical neuronal population (Blasdel et al., 1986). Signals can be isolated to a specific neuron subset by conditionally expressing a voltage-sensitive fluorescent protein (VSFP) under type-specific promoters (Akemann et al., 2010).

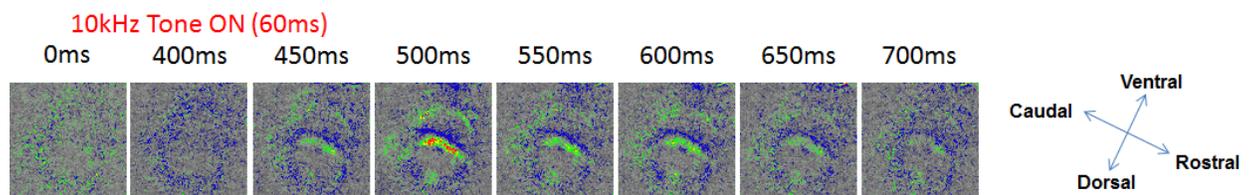
#### ***G.1 Voltage Sensitive Dye Imaging***

We initially utilized a VSD, RH-1691 (Slovin et al., 2002), which is a blue dye whose absorption spectrum is farther away from that of hemoglobin, and in principal would result in less blood flow associated artifacts present in the signal. CBA/CaJ mice were anesthetized and craniotomy over the left auditory cortex was conducted, a headpost was attached, and a well of dental cement was constructed around the craniotomy to hold the dye. The cortical surface was kept moist with saline, and RH-1691 was administered with a saline rinse over the course of 2 hours. Afterwards dye was drained and the well was filled with agar, and covered with a coverslip. From then, imaging was conducted.

In initial runs, n=4 animals had their craniotomy conducted under isoflurane anesthesia, and after recovery from isoflurane, urethane (500-1500mg/kg initial dose;

500mg/kg maintenance) was administered subcutaneously. Only one out of four animals survived this anesthesia regimen. For the next n=3 animals, all procedures including the craniotomy were done exclusively with urethane (1500mg-2000mg/kg initial dose; 500mg/kg maintenance) administered subcutaneously. None of these animals survived to imaging. For the final n=5 animals, procedures were also done exclusively with urethane (1500-2000mg/kg; 500mg/kg maintenance), this time administered intraperitoneally. Three out of five animals survived to imaging for these runs. As a result, for best results in animal survival when conducting urethane anesthesia, urethane should be administered intraperitoneally (IP) and should not be preceded by isoflurane anesthesia.

We were able to visualize cortical activity after pure tone presentation using VSD. A 10kHz tone was presented for 60ms and captured with a MiCAM Ultima camera in single camera mode with a red halogen light source. The camera lens was mounted at an angle relatively perpendicular to the left auditory cortical brain surface, and a speaker (Pioneer) was positioned 11cm away from the mouse's right ear. Activity in the cortex was observed at its most intense fluorescence approximately 100ms after tone playback (Fig G.1).



**Fig G.1.** Auditory cortical Voltage Sensitive RH-1691 Dye imaging. Tone playback begins at 400ms, with a 10kHz tone presented for 60ms. Two hotspots of activity are seen in the auditory cortex, likely representing the two low-frequency tuned Core auditory fields, A1 and AAF.

However, VSD RH-1691 would bleach rather quickly over the course of 20-30 minutes, and the signal to noise ratio would suffer near the end of recording sessions. Given that a full tuning curve can take approximately 10 minutes to collect, usage of the RH-1691 dye was limited for our application.

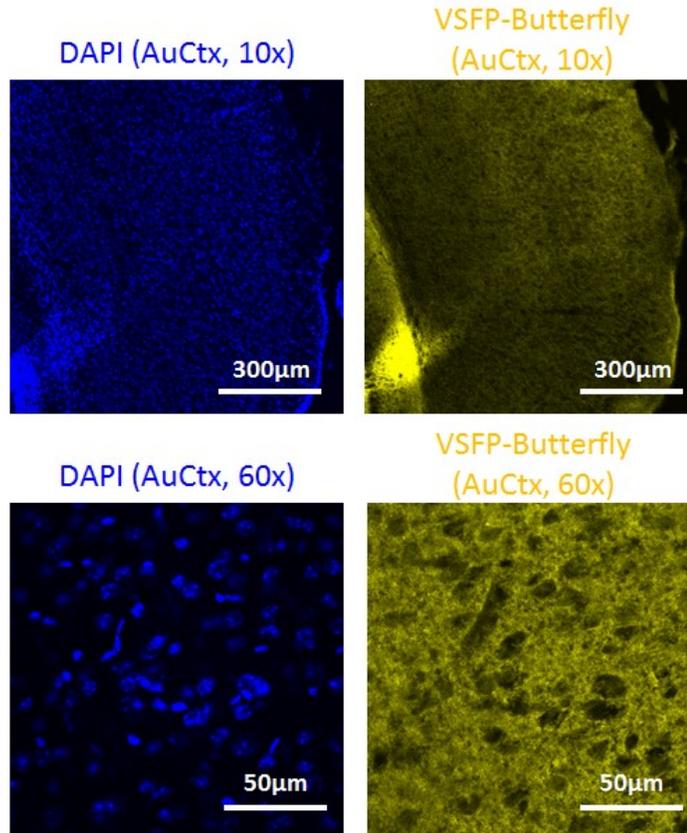
## ***G.2 Voltage Sensitive Fluorescent Protein Imaging***

An alternative to VSDs are genetically expressed voltage-sensitive fluorescent proteins (VSFP) (Akemann et al., 2010), also known as Genetically Encoded Voltage Indicators (GEVI) (Platasa et al., 2018). These proteins do not require a dye recirculation step as the voltage indicator is transgenically expressed by the mouse strain, and bleaching occurs at a slower rate. A transgenic strain on an Ai78 background expressing VSFP Butterfly 1.2 (VSFPB) (Akemann et al., 2012; Akemann et al., 2013) under the control of a pyramidal neuronal marker, EMX1 (Chan et al., 2001), was utilized for imaging studies. VSFPB contains two fluorescent reporters mKate2 (red) and mCitrine (yellow), which produce opposing fluorescent signals as voltage changes. The ratio of these two signals can be then calculated to approximate the voltage of an imaged area. This ratiometric normalization helps to reduce the heart pulsation artifact present in the signal (Akemann et al., 2012).

For experiments, briefly, mice underwent headpost attachment and a skull-thinning technique over the auditory cortex under isoflurane anesthesia, during which a glass coverslip is secured with cyanoacrylate glue over the thinned, smoothed skull atop auditory cortex. The mouse is allowed to recover for 24 hours after surgery. On the following day, the mouse is placed in a restraint while awake, and the headpost is secured under the imaging setup. A speaker is positioned 11cm from the right ear of the mouse, and a MiCAM Ultima camera is angled towards the auditory cortex. Due to the dual fluorescent nature of VSFPB, the camera is run in dual camera mode, with two filter

cubes (em:542, 594, di 580, and 506 di). A 100x100 pixel image is collected, with trial stimulus playback and camera recording synced using a trigger signal generated by LabView. Data collection was conducted at a rate of 200Hz, with 512 frames captured over the course of 2.56 seconds. The MiCAM Ultima was set to collect data under the DIF (CDS) mode, in which the first image is used as the baseline fluorescence, and the delta fluorescence for subsequent images is calculated based on the change in fluorescence from the baseline image.

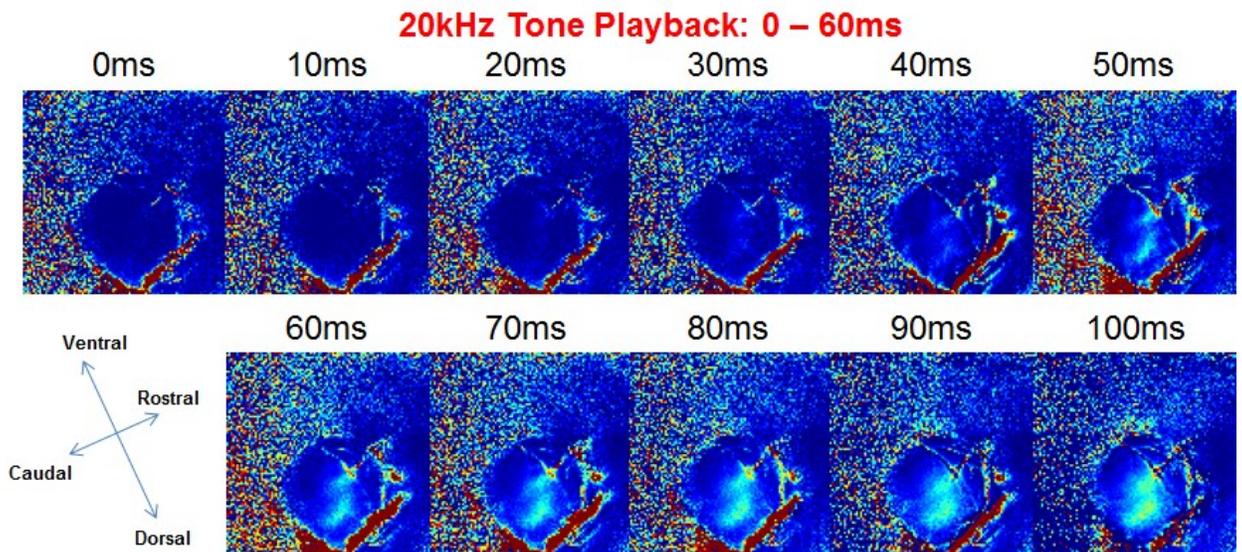
Surgical preparation for imaging was conducted on a total of n=12 animals. We confirmed the expression of VSFPB by collecting brain tissue from one of the experimental animals. Briefly, the animal was perfused with 4% paraformaldehyde, and brain tissue was then transferred to 30% sucrose the next day. The brain was then sectioned on a microtome into 40um thickness sections, before mounting, DAPI staining, and imaging. Expression of VSFPB is seen throughout the brain outside of the DAPI-stained nuclei and in the neuropil (**Fig G.2**).



**Fig G.2.** Histological images of VSFPB expression in the mouse auditory cortex. Sections are 40µm thickness. DAPI stain (Blue) visualizes the location of cell nuclei, while the inherent fluorescence of VSFP (yellow: mCitrine) shows expression of VSFPB throughout the entire auditory cortex, restricted mostly outside of the cell nucleus and throughout the neuropil.

Note that for the majority of animals (n=11 of 12), VSFP imaging data could not be used due to one of multiple reasons: 1) The craniotomy coverslip window was clouded and the cortical surface could not be visualized (n=1); 2) The headpost was dislodged before imaging began (n=3); 3) The data was collected and saved in an incorrect format (n=1); 4) We discovered the playback of the sound and camera recording time were not synced and had varying start times for earlier trials (n=6), so these data could not be used.

In the latest VSFPB trial, we were able to observe changes in fluorescence locked to sound playback. Playback of a pure tone 60ms in length directed at the right ear during left cortical VSFPB imaging was conducted. An average of 5 trials was taken, with the imaging camera focused at 300um depth under the cortical surface, and the sound intensity at +6dB SNR. Sample images taken over the course of a 20kHz tone playback show two distinct activation foci, likely representing the two low frequency primary auditory cortical areas, A1 and AAF (**Fig G.3**). This preliminary data shows that VSFPB can be used to image auditory cortical activity in awake, head-fixed mice.



**Fig G.3.** Time course images of voltage sensitive fluorescent protein Butterfly 1.2 in the auditory cortex during 20kHz pure tone playback at +6dB SNR and 60ms duration. Averaged over 5 trials at a sound level of +6dB SNR.

## Appendix H

### Photoidentification of Neuronal Populations (PINP) and Viral Vectors

The mouse model provides many opportunities for more precise dissection of neural circuitry using genetic tools. Many transgenic strains expressing various proteins under different neuronal cell type markers are available from vendors. Viral vectors can also be used to induce expression of specific proteins of interest, without requiring breeding. A genetic tool that has gained a great deal of traction in the neuroscience field is Optogenetics, where the expression of light-sensitive ion channels can enable precise optogenetic activation of specific cell subtypes (Boyden, 2015; Deisseroth, 2015; Kim et al., 2017; Mei et al., 2012; Zhang et al., 2010). Briefly, a viral vector, which is usually an inactivated viral particle that can introduce genetic material into a cell is packaged with a plasmid, or custom-made ring of genetic material. The sequence of the plasmid will determine the type of genes as well as the conditions under which they will be expressed. Optogenetic proteins such as Channelrhodopsins can be made to be expressed only under certain gene promoters, such that very specific neuronal subtypes will express the protein and thus be light sensitive.

One application of optogenetics in neuroscience for electrophysiological study is Photoidentification of Neuronal Populations (PINP). PINP requires a mouse to be expressing a type of light-sensitive ion channel under control of a cell type-specific promoter, which allows light-driven activity of neurons to determine what neuronal subtype is being recorded (Lima et al., 2009; Moore et al., 2014). Previously, electrophysiological recordings are made blind to the type of neuron that is being recorded from, and the genetic profile of a recorded neuron is completely unknown. At most, inferences can be made about the identity of a cell based on the spiking characteristics of the cell, such as the shape and peak-to-peak width of a cell's

waveform (Bartho et al., 2004; Mountcastle et al., 1969), with larger peak-to-peak widths (“thick-spiking”) more likely to be attributed to pyramidal cells, and smaller widths (“thin-spiking”) attributed to interneurons. However, even within pyramidal and interneurons, there are a large variety of subtypes, and each can play a different role in the processing of information in the brain (Markram et al., 2004; Molyneaux et al., 2007). Thus, a technique that can more precisely identify the neuron subtype during recording, such as PINP, is desirable.

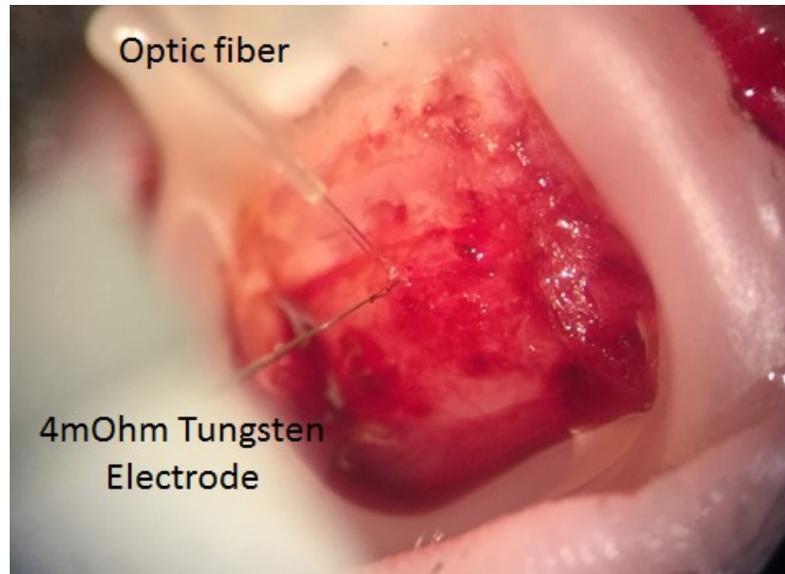
### ***H.1 PINP***

Neurons that express channelrhodopsin characteristically will fire upon presentation of the correct light wavelength. For the most common channelrhodopsin, ChR2, a blue light of 473nm wavelength is generally used. Light delivery can either be conducted via a combined electrode and fiber optic, or “optrode”. Alternatively, when an electrode is inserted, an optic fiber is hovered over the exposed cortical surface and light is delivered. In our application, we used the latter, where an optic fiber is hovered above the auditory cortical surface with an electrode inserted into the surface. To avoid heat damage on tissue, the total light intensity should not exceed a power of 10mW. Calibration of light intensity was conducted such that light intensity delivered by a laser (Cobolt, 473nm) at a distance of 1mm from a 300µm diameter optic fiber (Thorlabs Part No. M69L01) ranges from 2 – 8mW (**Table H.1**).

Target mW	AMP (V) – Cobolt 473nm
2	0.573708
3	0.582086
4	0.590465
5	0.598844
6	0.607222
7	0.615601
8	0.62398

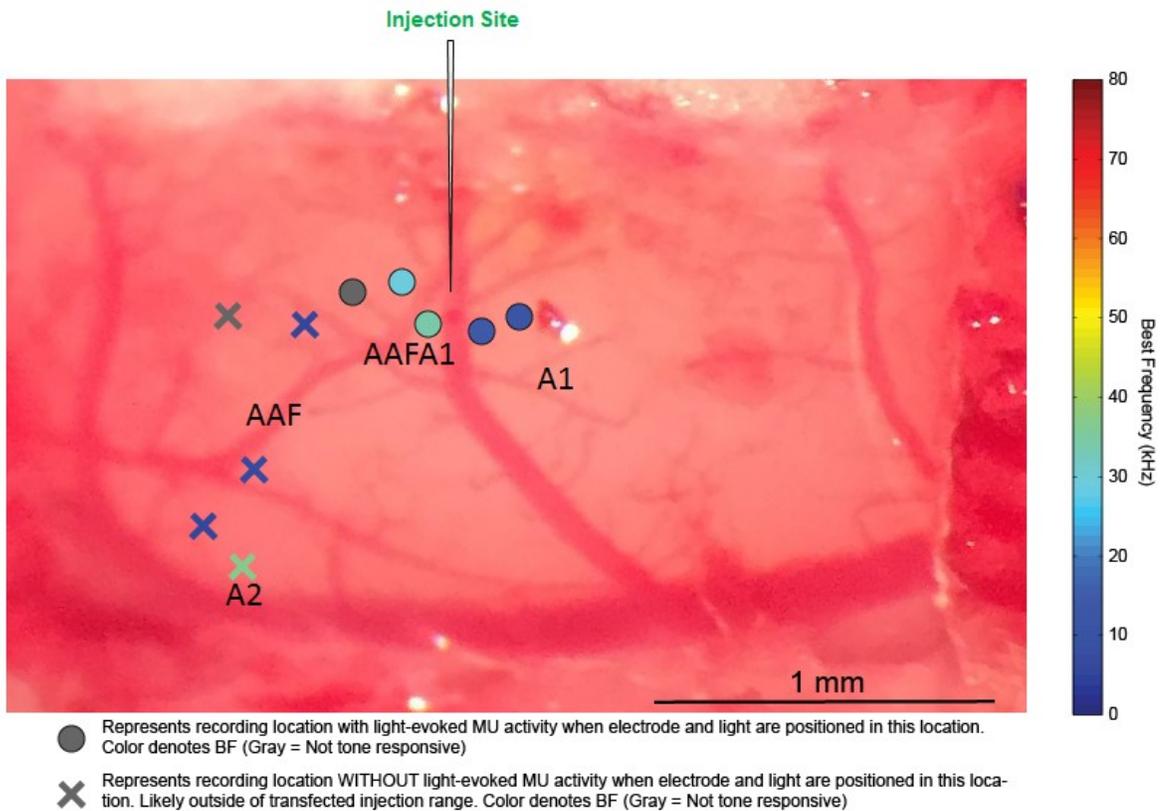
**Table H.1.** Laser power calibration information using the Cobolt laser system. Light intensity as measured using a light diode at 1mm distance from the fiber tip, and corresponding voltage (V) values delivered to the laser system are listed.

To test whether PINP neurons could be electrophysiology isolated in auditory cortex, CBA/CaJ were injected with 32.2 – 64.4nl of adeno-associated virus rAAV5-aCamKIIa-hChR2-EYFP at a stereotaxic position of Bregma X – 3.5mm; Bregma Y – 4.0mm; Bregma Z [cortical surface] – 0.5mm, corresponding approximately to the high frequency reversal region AAFA1 in the left auditory cortex. Injections were performed using a pressure injector (Nanoject II, Drummond, Cat 3-000-204) at a rate of 0.5nl/s (**Appendix H.4**). Mice were then allowed to recover and given 3-7 weeks to allow the protein to be expressed. On the PINP procedure day, mice were anesthetized with isoflurane and a craniotomy was performed over left auditory cortex. A headpost (inverted screw) was affixed to the skull with dental cement, and the mouse was placed in an anechoic chamber with a speaker (Emit) positioned 11cm from the right ear of the mouse. A 4 M $\Omega$  tungsten electrode (FHC) was inserted into the cortical surface with an initial depth of ~200 $\mu$ m using a microdrive and an optical fiber is positioned over the penetration (**Fig H.1**). Light playback was used as a search stimulus as the electrode was advanced slowly at 5-10 $\mu$ m/s to locate PINP neurons.

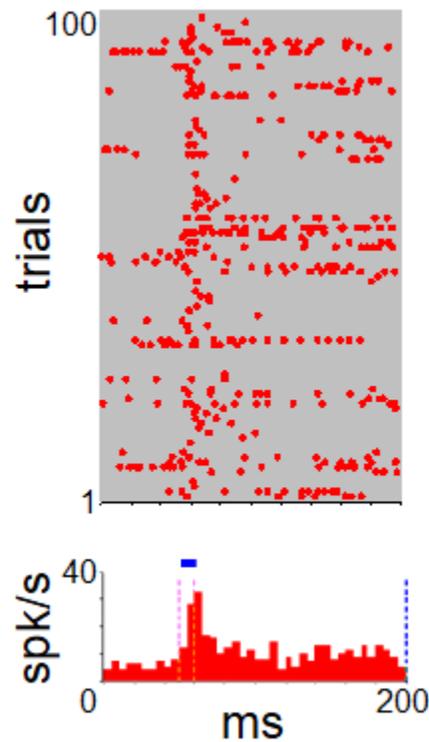


**Fig H.1.** Image of PINP over Mouse Left Auditory Cortex Craniotomy, with tungsten electrode inserted and optic fiber positioned over the surface of the auditory cortex penetration site.

A)



**B)**

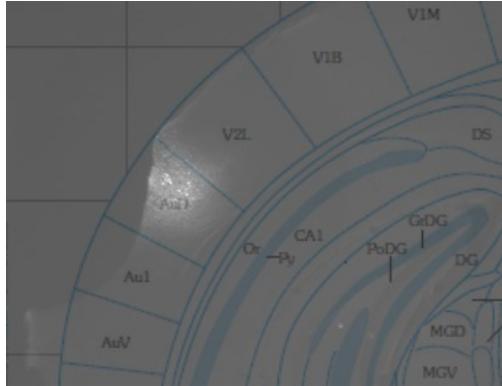


**Figure H.2.** Sample light-driven responses. **A)** Cortical surface image depicting electrode penetration sites. Each circle or X symbol indicates an electrode recording, with the symbol color denoting the best frequency of the multiunit (MU) recorded at 400um depth (which approximately correlates with the Thalamic input layer). **B)** Sample raster of a multiunit recorded from a penetration with a 10ms duration light pulse (blue bar) delivered 50ms into the trial. Multiunit activity is light-evoked. Given that the ChR2 is expressed under EMX1, which is a pyramidal neuronal marker, the activation of these cells may be activating other nearby neurons, which results in a slightly broader and prolonged response that is less tightly locked to light delivery.

Multiunit (MU) activity was recorded across penetration sites, and their relative positioning was kept track of via a high resolution cortical surface image using vascular structure as landmarks. Light responsive regions of auditory cortex are observed close to the injection site, whereas regions more distal to the injection site did not show light-evoked activity (**Fig H.2**).

Histology was also performed to confirm that expression of the channelrhodopsin protein was in the left auditory cortex as expected. For histology, briefly, the mouse was perfused with 4% paraformaldehyde and the brain was extracted and submerged in

paraformaldehyde overnight. The brain tissue was then transferred to a 30% sucrose solution for 24 hours before the brain was sectioned on a microtome. Auditory cortical sections were collected at 40um thickness, and mounted onto a glass microscope slide before coverslipping using a DAPI stain and imaging. Section images were overlaid with the mouse atlas to confirm infection location (**Fig H.3**).



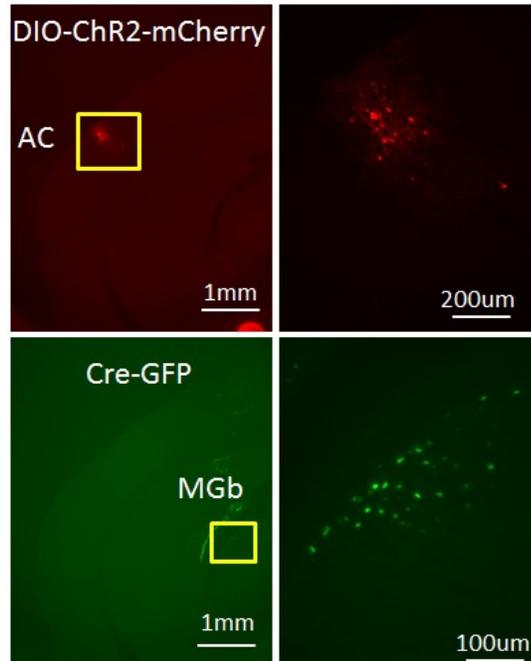
**Figure H.3.** Histological section showing yellow fluorescent protein (EYFP) expression at ChR2 injection site. Injection site overlaps with AuD, dorsal auditory cortex.

For identification of single neurons using PINP, neurons that are expressing ChR2 should spike reliably to light presentation for 70-100% of trials with a latency of 2-5ms (Lima et al., 2009). Cells that are monosynaptically connected to expressing cells but do not themselves express ChR2 also will spike but with a lower reliability of 0-40% of trials with a longer latency of 8+ms and with less tightly phase-locked spike timing to light presentation. Our usage of the EMX1 promoter targets excitatory pyramidal neurons, and made it particularly difficult to separate single ChR2-expressing units out from multiple neighboring neurons that are spiking due to the excitation of the EMX1-expressing. As a result, for the most part, multiunit recordings were yielded from these PINP trials.

## ***H.2 PINP with Transsynaptic Tracing***

A recent paper had demonstrated anterograde activity in adeno-associated virus serotype 1 (AAV1), opening the possibility of using the virus for identification or recording from neurons receiving projections from specified brain regions (Zingg et al., 2017). Using a combination of AAV1-Cre in an upstream brain location, and AAV-dflox-(Gene marker)-Chr2 in a downstream brain region, PINP can be conducted to identify neurons that receive projections from other regions.

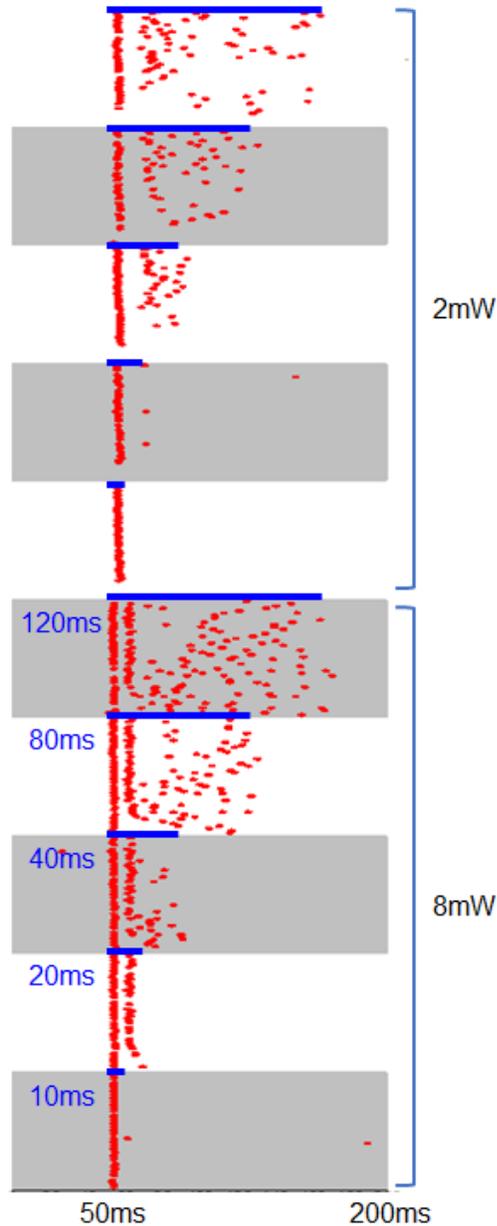
In a pilot study to validate the transsynaptic activity of AAV1 and the ability to record from a projection-receiving neuron, we conducted viral injections in CBA/CaJ mice. Two injections were made: 64.4nl of AAV1-Cre-GFP was injected into the left auditory thalamus, the medial geniculate body (MGB), and 64.4nl of AAV1-DIO-ChR2-mCherry was injected into the left auditory cortex (AC) and an incubation period of 4 weeks was used to allow the proteins to be expressed. Using this combination of viruses, auditory cortical neurons should only fluoresce with mCherry if they are receiving projections from the auditory thalamus. After four weeks, perfusion was conducted on the animal, and the brain was sectioned into 40um sections encompassing MGB and AC, and fluorescent imaging of brain sections was conducted. Sections were aligned with the mouse atlas (Allen Brain Atlas 3<sup>rd</sup> ed.) to confirm injection coordinates and spread of expression. We were able to confirm the transsynaptic action of AAV1-Cre, as well as the correct injection placements in MGB and AC (**Fig H.4**).



**Figure H.4.** Histological Confirmation of AAV Transsynaptic activity. Top left: AAV1-DIO-ChR2-mCherry red fluorescent expression in auditory cortex (AC); scale bar on bottom right. Top right: Corresponding zoomed in image of the region highlighted with a yellow box, depicting staining of cell bodies and neuropil. Bottom left: AAV1-Cre-GFP green fluorescent expression in the medial geniculate body of the thalamus (MGb). Bottom right: Zoomed in image of the bottom left image's region highlighted with a yellow box, showing staining of cell bodies in the MGB.

We then sought to confirm that we could electrophysiologically record from and identify the ChR2-expressing neurons in the auditory cortex using light stimulation. In a separate animal cohort, the same injections of AAV1-Cre-GFP into left MGB and AAV1-DIO-ChR2-mCherry into left AC. After 3-4 weeks, animals underwent headpost attachment and small-hole craniotomy, and were given 24 hours of recovery post-surgery. On subsequent days, fully awake head-fixed single unit recordings were done in a similar fashion to the earlier PINP pilot studies, except rather than a full craniotomy, electrodes were directed into small holes drilled into the skull with a fiber optic pointed at the hole and delivering a 473nm laser (Cobolt).

Neurons expressing ChR2 in the auditory cortex were identified by delivering a 10ms blue light pulse as a single tungsten electrode (6M $\Omega$ ) was advanced into the auditory cortex. Neurons that exhibited a low 2-5ms spike latency with >70% trial consistency were considered as ChR2-expressing. Across three animals, one ChR2-expressing neuron was recorded (**Fig H.5**).



**Figure H.5.** Raster responses of an AC neuron receiving projections from MGB as identified by PINP. Blue bar represents laser on time. Laser power was set at two levels of 2mW or 8mW, with varying laser pulse lengths. High first spike time consistency

across trials and presence of spiking across >70% of trials indicated that this was a ChR2-expressing neuron. Note higher laser power can cause ChR2-expressing neurons to fire in doubles, and longer laser pulse durations induce scattered spiking throughout laser presentation time.

## REFERENCES

- Aertsen, A. M., & Johannesma, P. I. (1981). The spectro-temporal receptive field. A functional characteristic of auditory neurons. *Biol Cybern*, *42*(2), 133-143.
- Aitkin, L. M., & Webster, W. R. (1971). Tonotopic organization in the medial geniculate body of the cat. *Brain Res*, *26*(2), 402-405.
- Akemann, W., Mutoh, H., Perron, A., Park, Y. K., Iwamoto, Y., & Knopfel, T. (2012). Imaging neural circuit dynamics with a voltage-sensitive fluorescent protein. *J Neurophysiol*, *108*(8), 2323-2337. doi:10.1152/jn.00452.2012
- Akemann, W., Mutoh, H., Perron, A., Rossier, J., & Knopfel, T. (2010). Imaging brain electric signals with genetically targeted voltage-sensitive fluorescent proteins. *Nat Methods*, *7*(8), 643-649. doi:10.1038/nmeth.1479
- Akemann, W., Sasaki, M., Mutoh, H., Imamura, T., Honkura, N., & Knopfel, T. (2013). Two-photon voltage imaging using a genetically encoded voltage indicator. *Sci Rep*, *3*, 2231. doi:10.1038/srep02231
- Anderson, L. A., & Linden, J. F. (2016). Mind the Gap: Two Dissociable Mechanisms of Temporal Processing in the Auditory System. *J Neurosci*, *36*(6), 1977-1995. doi:10.1523/JNEUROSCI.1652-15.2016
- Atallah, B. V., Bruns, W., Carandini, M., & Scanziani, M. (2012). Parvalbumin-expressing interneurons linearly transform cortical responses to visual stimuli. *Neuron*, *73*(1), 159-170. doi:10.1016/j.neuron.2011.12.013
- Atencio, C. A., Blake, D. T., Strata, F., Cheung, S. W., Merzenich, M. M., & Schreiner, C. E. (2007). Frequency-modulation encoding in the primary auditory cortex of the awake owl monkey. *J Neurophysiol*, *98*(4), 2182-2195. doi:10.1152/jn.00394.2007

- Bachorowski, J.-A., & Owren, M. J. (1995). Vocal Expression of Emotion: Acoustic Properties of Speech Are Associated With Emotional Intensity and Context. *Psychological Science*, 6(4), 219-224. doi:10.1111/j.1467-9280.1995.tb00596.x
- Bandyopadhyay, S., Shamma, S. A., & Kanold, P. O. (2010). Dichotomy of functional organization in the mouse auditory cortex. *Nat Neurosci*, 13(3), 361-368. doi:10.1038/nn.2490
- Bartho, P., Hirase, H., Monconduit, L., Zugaro, M., Harris, K. D., & Buzsaki, G. (2004). Characterization of neocortical principal cells and interneurons by network interactions and extracellular features. *J Neurophysiol*, 92(1), 600-608. doi:10.1152/jn.01170.2003
- Bennur, S., Tsunada, J., Cohen, Y. E., & Liu, R. C. (2013). Understanding the neurophysiological basis of auditory abilities for social communication: a perspective on the value of ethological paradigms. *Hear Res*, 305, 3-9. doi:10.1016/j.heares.2013.08.008
- Bijeljac-Babic, R., Serres, J., Hohle, B., & Nazzi, T. (2012). Effect of bilingualism on lexical stress pattern discrimination in French-learning infants. *PLoS One*, 7(2), e30843. doi:10.1371/journal.pone.0030843
- Bizley, J. K., Nodal, F. R., Nelken, I., & King, A. J. (2005). Functional organization of ferret auditory cortex. *Cereb Cortex*, 15(10), 1637-1653. doi:10.1093/cercor/bhi042
- Blasdel, G. G., & Salama, G. (1986). Voltage-sensitive dyes reveal a modular organization in monkey striate cortex. *Nature*, 321(6070), 579-585. doi:10.1038/321579a0
- Boyden, E. S. (2015). Optogenetics and the future of neuroscience. *Nat Neurosci*, 18(9), 1200-1201. doi:10.1038/nn.4094

- Britt, R., & Starr, A. (1976). Synaptic events and discharge patterns of cochlear nucleus cells. II. Frequency-modulated tones. *J Neurophysiol*, *39*(1), 179-194.  
doi:10.1152/jn.1976.39.1.179
- Brown, T. A., & Harrison, R. V. (2009). Responses of neurons in chinchilla auditory cortex to frequency-modulated tones. *J Neurophysiol*, *101*(4), 2017-2029.  
doi:10.1152/jn.90931.2008
- Brudzynski, S. M. (2007). Ultrasonic calls of rats as indicator variables of negative or positive states: acetylcholine-dopamine interaction and acoustic coding. *Behav Brain Res*, *182*(2), 261-273. doi:10.1016/j.bbr.2007.03.004
- Brugge, J. F., & Merzenich, M. M. (1973). Responses of neurons in auditory cortex of the macaque monkey to monaural and binaural stimulation. *J Neurophysiol*, *36*(6), 1138-1158. doi:10.1152/jn.1973.36.6.1138
- Butt, S. J., Fuccillo, M., Nery, S., Noctor, S., Kriegstein, A., Corbin, J. G., & Fishell, G. (2005). The temporal and spatial origins of cortical interneurons predict their physiological subtype. *Neuron*, *48*(4), 591-604. doi:10.1016/j.neuron.2005.09.034
- Carrasco, A., & Lomber, S. G. (2011). Neuronal activation times to simple, complex, and natural sounds in cat primary and nonprimary auditory cortex. *J Neurophysiol*, *106*(3), 1166-1178. doi:10.1152/jn.00940.2010
- Carruthers, I. M., Natan, R. G., & Geffen, M. N. (2013). Encoding of ultrasonic vocalizations in the auditory cortex. *J Neurophysiol*, *109*(7), 1912-1927.  
doi:10.1152/jn.00483.2012
- Chan, C. H., Godinho, L. N., Thomaidou, D., Tan, S. S., Gulisano, M., & Parnavelas, J. G. (2001). Emx1 is a marker for pyramidal neurons of the cerebral cortex. *Cereb Cortex*, *11*(12), 1191-1198.

- Chevillet, M., Riesenhuber, M., & Rauschecker, J. P. (2011). Functional correlates of the anterolateral processing hierarchy in human auditory cortex. *J Neurosci*, *31*(25), 9345-9352. doi:10.1523/JNEUROSCI.1448-11.2011
- Chimoto, S., Kitama, T., Qin, L., Sakayori, S., & Sato, Y. (2002). Tonal response patterns of primary auditory cortex neurons in alert cats. *Brain Res*, *934*(1), 34-42.
- Christensen-Dalsgaard, J., Brandt, C., Wilson, M., Wahlberg, M., & Madsen, P. T. (2011). Hearing in the African lungfish (*Protopterus annectens*): pre-adaptation to pressure hearing in tetrapods? *Biol Lett*, *7*(1), 139-141. doi:10.1098/rsbl.2010.0636
- Creutzfeldt, O., Hellweg, F. C., & Schreiner, C. (1980). Thalamocortical transformation of responses to complex auditory stimuli. *Exp Brain Res*, *39*(1), 87-104.
- Cruikshank, S. J., Killackey, H. P., & Metherate, R. (2001). Parvalbumin and calbindin are differentially distributed within primary and secondary subregions of the mouse auditory forebrain. *Neuroscience*, *105*(3), 553-569.
- Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: a literature review. *Lang Speech*, *40* ( Pt 2), 141-201. doi:10.1177/002383099704000203
- Dayan, P., & Abbott, L. F. (2001). *Theoretical neuroscience* (Vol. 806): Cambridge, MA: MIT Press.
- deCharms, R. C., Blake, D. T., & Merzenich, M. M. (1998). Optimizing sound features for cortical neurons. *Science*, *280*(5368), 1439-1443.
- deCharms, R. C., & Merzenich, M. M. (1996). Primary cortical representation of sounds by the coordination of action-potential timing. *Nature*, *381*(6583), 610-613. doi:10.1038/381610a0

- Deisseroth, K. (2015). Optogenetics: 10 years of microbial opsins in neuroscience. *Nat Neurosci*, 18(9), 1213-1225. doi:10.1038/nn.4091
- Depireux, D. A., Simon, J. Z., Klein, D. J., & Shamma, S. A. (2001). Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex. *J Neurophysiol*, 85(3), 1220-1234. doi:10.1152/jn.2001.85.3.1220
- Donato, F., Rompani, S. B., & Caroni, P. (2013). Parvalbumin-expressing basket-cell network plasticity induced by experience regulates adult learning. *Nature*, 504(7479), 272-276. doi:10.1038/nature12866
- Eggermont, J. J. (1991). Rate and synchronization measures of periodicity coding in cat primary auditory cortex. *Hear Res*, 56(1-2), 153-167.
- Eggermont, J. J. (1994). Temporal modulation transfer functions for AM and FM stimuli in cat auditory cortex. Effects of carrier type, modulating waveform and intensity. *Hear Res*, 74(1-2), 51-66.
- Ehret, G. (2005). Infant Rodent Ultrasounds – A Gate to the Understanding of Sound Communication. *Behavior Genetics*, 35(1), 19-29. doi:10.1007/s10519-004-0853-8
- Ehret, G., & Haack, B. (1984). Motivation and Arousal Influence Sound-induced Maternal Pup-retrieving Behavior in Lactating House Mice. *Zeitschrift für Tierpsychologie*, 65(1), 25-39. doi:doi:10.1111/j.1439-0310.1984.tb00370.x
- Ehret, G., Koch, M., Haack, B., & Markl, H. (1987). Sex and parental experience determine the onset of an instinctive behavior in mice. *Naturwissenschaften*, 74(1), 47.
- Eilers, P. H., & Goeman, J. J. (2004). Enhancing scatterplots with smoothed densities. *Bioinformatics*, 20(5), 623-628. doi:10.1093/bioinformatics/btg454
- Evans, E. F., & Whitfield, I. C. (1964). Classification of Unit Responses in the Auditory Cortex of the Unanaesthetized and Unrestrained Cat. *J Physiol*, 171, 476-493.

- Fischer, J., Hammerschmidt, K., & Todt, D. (1995). Factors Affecting Acoustic Variation in Barbary-macaque (*Macaca sylvanus*) Disturbance Calls. *Ethology*, *101*(1), 51-66. doi:doi:10.1111/j.1439-0310.1995.tb00345.x
- Fishman, Y. I., & Steinschneider, M. (2009). Temporally dynamic frequency tuning of population responses in monkey primary auditory cortex. *Hear Res*, *254*(1-2), 64-76. doi:10.1016/j.heares.2009.04.010
- Flores, E. N., Duggan, A., Madathany, T., Hogan, A. K., Marquez, F. G., Kumar, G., . . . Garcia-Anoveros, J. (2015). A non-canonical pathway from cochlea to brain signals tissue-damaging noise. *Curr Biol*, *25*(5), 606-612. doi:10.1016/j.cub.2015.01.009
- Frisina, R. D., Smith, R. L., & Chamberlain, S. C. (1990). Encoding of amplitude modulation in the gerbil cochlear nucleus: I. A hierarchy of enhancement. *Hear Res*, *44*(2-3), 99-122.
- Froemke, R. C. (2015). Plasticity of cortical excitatory-inhibitory balance. *Annu Rev Neurosci*, *38*, 195-219. doi:10.1146/annurev-neuro-071714-034002
- Gaese, B. H., & Ostwald, J. (1995). Temporal coding of amplitude and frequency modulation in the rat auditory cortex. *Eur J Neurosci*, *7*(3), 438-450.
- Galambos, R. (1956). XC Some Recent Experiments on the Neurophysiology of Hearing. *Annals of Otology, Rhinology & Laryngology*, *65*(4), 1053-1059. doi:10.1177/000348945606500418
- Galindo-Leon, E. E., Lin, F. G., & Liu, R. C. (2009). Inhibitory plasticity in a lateral band improves cortical detection of natural vocalizations. *Neuron*, *62*(5), 705-716. doi:10.1016/j.neuron.2009.05.001
- Geissler, D. B., & Ehret, G. (2002). Time-critical integration of formants for perception of communication calls in mice. *Proc Natl Acad Sci U S A*, *99*(13), 9021-9025. doi:10.1073/pnas.122606499

- Geissler, D. B., & Ehret, G. (2004). Auditory perception vs. recognition: representation of complex communication sounds in the mouse auditory cortical fields. *Eur J Neurosci*, *19*(4), 1027-1040.
- Geschwind, N. (1972). Language and the brain. *Sci Am*, *226*(4), 76-83.
- Godey, B., Atencio, C. A., Bonham, B. H., Schreiner, C. E., & Cheung, S. W. (2005). Functional organization of squirrel monkey primary auditory cortex: responses to frequency-modulation sweeps. *J Neurophysiol*, *94*(2), 1299-1311.  
doi:10.1152/jn.00950.2004
- Goldinger, S. D., Kleider, H. M., & Shelley, E. (1999). The marriage of perception and memory: creating two-way illusions with words and voices. *Mem Cognit*, *27*(2), 328-338.
- Graven, S. N., & Browne, J. V. (2008). Auditory Development in the Fetus and Infant. *Newborn and Infant Nursing Reviews*, *8*(4), 187-193.  
doi:<https://doi.org/10.1053/j.nainr.2008.10.010>
- Guo, W., Chambers, A. R., Darrow, K. N., Hancock, K. E., Shinn-Cunningham, B. G., & Polley, D. B. (2012). Robustness of cortical topography across fields, laminae, anesthetic states, and neurophysiological signal types. *J Neurosci*, *32*(27), 9159-9172. doi:10.1523/JNEUROSCI.0065-12.2012
- Hall, D. A., Johnsrude, I. S., Haggard, M. P., Palmer, A. R., Akeroyd, M. A., & Summerfield, A. Q. (2002). Spectral and temporal processing in human auditory cortex. *Cereb Cortex*, *12*(2), 140-149.
- Han, Y. K., Köver, H., Insanally, M. N., Semerdjian, J. H., & Bao, S. (2007). Early experience impairs perceptual discrimination. *Nature Neuroscience*, *10*, 1191.  
doi:10.1038/nn1941  
<https://www.nature.com/articles/nn1941#supplementary-information>

- Hari, R., Pelizzone, M., Makela, J. P., Hallstrom, J., Leinonen, L., & Lounasmaa, O. V. (1987). Neuromagnetic responses of the human auditory cortex to on- and offsets of noise bursts. *Audiology*, *26*(1), 31-43.
- Harms, M. P., Guinan, J. J., Jr., Sigalovsky, I. S., & Melcher, J. R. (2005). Short-term sound temporal envelope characteristics determine multisecond time patterns of activity in human auditory cortex as shown by fMRI. *J Neurophysiol*, *93*(1), 210-222. doi:10.1152/jn.00712.2004
- Hart, H. C., Palmer, A. R., & Hall, D. A. (2003). Amplitude and frequency-modulated stimuli activate common regions of human auditory cortex. *Cereb Cortex*, *13*(7), 773-781.
- Heil, P. (1997a). Auditory cortical onset responses revisited. I. First-spike timing. *J Neurophysiol*, *77*(5), 2616-2641. doi:10.1152/jn.1997.77.5.2616
- Heil, P. (1997b). Auditory cortical onset responses revisited. II. Response strength. *J Neurophysiol*, *77*(5), 2642-2660. doi:10.1152/jn.1997.77.5.2642
- Heil, P., Rajan, R., & Irvine, D. R. (1992). Sensitivity of neurons in cat primary auditory cortex to tones and frequency-modulated stimuli. II: Organization of response properties along the 'isofrequency' dimension. *Hear Res*, *63*(1-2), 135-156.
- Hind, J. E. (1953). An electrophysiological determination of tonotopic organization in auditory cortex of cat. *J Neurophysiol*, *16*(5), 475-489. doi:10.1152/jn.1953.16.5.475
- Issa, J. B., Haeffele, B. D., Agarwal, A., Bergles, D. E., Young, E. D., & Yue, D. T. (2014). Multiscale optical Ca<sup>2+</sup> imaging of tonal organization in mouse auditory cortex. *Neuron*, *83*(4), 944-959. doi:10.1016/j.neuron.2014.07.009
- Joachimsthaler, B., Uhlmann, M., Miller, F., Ehret, G., & Kurt, S. (2014). Quantitative analysis of neuronal response properties in primary and higher-order auditory

- cortical fields of awake house mice (*Mus musculus*). *Eur J Neurosci*, 39(6), 904-918. doi:10.1111/ejn.12478
- Johnson, D. H. (1980). The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. *J Acoust Soc Am*, 68(4), 1115-1122.
- Joris, P. X., & Yin, T. C. (1992). Responses to amplitude-modulated tones in the auditory nerve of the cat. *J Acoust Soc Am*, 91(1), 215-232.
- Kaas, J. H., Hackett, T. A., & Tramo, M. J. (1999). Auditory processing in primate cerebral cortex. *Curr Opin Neurobiol*, 9(2), 164-170.
- Kikuchi, Y., Horwitz, B., & Mishkin, M. (2010). Hierarchical auditory processing directed rostrally along the monkey's supratemporal plane. *J Neurosci*, 30(39), 13021-13030. doi:10.1523/JNEUROSCI.2267-10.2010
- Kim, C. K., Adhikari, A., & Deisseroth, K. (2017). Integration of optogenetics with complementary methodologies in systems neuroscience. *Nat Rev Neurosci*, 18(4), 222-235. doi:10.1038/nrn.2017.15
- Kimura, D. (1961). Some effects of temporal-lobe damage on auditory perception. *Can J Psychol*, 15, 156-165.
- Klump, G. M., & Shalter, M. D. (1984). Acoustic Behaviour of Birds and Mammals in the Predator Context; I. Factors Affecting the Structure of Alarm Signals. II. The Functional Significance and Evolution of Alarm Signals. *Zeitschrift für Tierpsychologie*, 66(3), 189-226. doi:doi:10.1111/j.1439-0310.1984.tb01365.x
- Kowalski, N., Depireux, D. A., & Shamma, S. A. (1996). Analysis of dynamic spectra in ferret primary auditory cortex. I. Characteristics of single-unit responses to moving ripple spectra. *J Neurophysiol*, 76(5), 3503-3523. doi:10.1152/jn.1996.76.5.3503
- Kubota, Y., Kamatani, D., Tsukano, H., Ohshima, S., Takahashi, K., Hishida, R., . . . Shibuki, K. (2008). Transcranial photo-inactivation of neural activities in the

mouse auditory cortex. *Neurosci Res*, 60(4), 422-430.

doi:10.1016/j.neures.2007.12.013

Lehiste, I. (1970). *Suprasegmentals* (Cambridge, Mass.): MIT Press.

Liang, L., Lu, T., & Wang, X. (2002). Neural representations of sinusoidal amplitude and frequency modulations in the primary auditory cortex of awake primates. *J Neurophysiol*, 87(5), 2237-2261. doi:10.1152/jn.2002.87.5.2237

Liberman, A. M., Delattre, P. C., Gerstman, L. J., & Cooper, F. S. (1956). Tempo of frequency change as a cue for distinguishing classes of speech sounds. *J Exp Psychol*, 52(2), 127-137.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *J Exp Psychol*, 54(5), 358-368.

Lima, S. Q., Hromadka, T., Znamenskiy, P., & Zador, A. M. (2009). PINP: a new method of tagging neuronal populations for identification during in vivo electrophysiological recording. *PLoS One*, 4(7), e6099.  
doi:10.1371/journal.pone.0006099

Lin, F. G., Galindo-Leon, E. E., Ivanova, T. N., Mappus, R. C., & Liu, R. C. (2013). A role for maternal physiological state in preserving auditory cortical plasticity for salient infant calls. *Neuroscience*, 247, 102-116.  
doi:10.1016/j.neuroscience.2013.05.020

Lin, F. G., & Liu, R. C. (2010). Subset of thin spike cortical neurons preserve the peripheral encoding of stimulus onsets. *J Neurophysiol*, 104(6), 3588-3599.  
doi:10.1152/jn.00295.2010

Liu, L. F., Palmer, A. R., & Wallace, M. N. (2006). Phase-locked responses to pure tones in the inferior colliculus. *J Neurophysiol*, 95(3), 1926-1935.  
doi:10.1152/jn.00497.2005

- Liu, R. C., Linden, J. F., & Schreiner, C. E. (2006). Improved cortical entrainment to infant communication calls in mothers compared with virgin mice. *Eur J Neurosci*, 23(11), 3087-3097. doi:10.1111/j.1460-9568.2006.04840.x
- Liu, R. C., Miller, K. D., Merzenich, M. M., & Schreiner, C. E. (2003). Acoustic variability and distinguishability among mouse ultrasound vocalizations. *J Acoust Soc Am*, 114(6 Pt 1), 3412-3422.
- Liu, R. C., & Schreiner, C. E. (2007). Auditory cortical detection and discrimination correlates with communicative significance. *PLoS Biol*, 5(7), e173. doi:10.1371/journal.pbio.0050173
- Lu, T., Liang, L., & Wang, X. (2001). Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. *Nat Neurosci*, 4(11), 1131-1138. doi:10.1038/nn737
- Markram, H., Toledo-Rodriguez, M., Wang, Y., Gupta, A., Silberberg, G., & Wu, C. (2004). Interneurons of the neocortical inhibitory system. *Nat Rev Neurosci*, 5(10), 793-807. doi:10.1038/nrn1519
- Mattock, K., & Burnham, D. (2006). Chinese and English Infants' Tone Perception: Evidence for Perceptual Reorganization. *Infancy*, 10(3), 241-265. doi:doi:10.1207/s15327078in1003\_3
- Mattock, K., Molnar, M., Polka, L., & Burnham, D. (2008). The developmental course of lexical tone perception in the first year of life. *Cognition*, 106(3), 1367-1381. doi:10.1016/j.cognition.2007.07.002
- Maunsell, J. H., & Gibson, J. R. (1992). Visual response latencies in striate cortex of the macaque monkey. *J Neurophysiol*, 68(4), 1332-1344. doi:10.1152/jn.1992.68.4.1332
- Mead, M., & Newton, N. (1967). Cultural patterning of perinatal behavior.

- Mei, Y., & Zhang, F. (2012). Molecular tools and approaches for optogenetics. *Biol Psychiatry*, 71(12), 1033-1038. doi:10.1016/j.biopsych.2012.02.019
- Mendelson, J. R., & Cynader, M. S. (1985). Sensitivity of cat primary auditory cortex (AI) neurons to the direction and rate of frequency modulation. *Brain Res*, 327(1-2), 331-335.
- Mendelson, J. R., Schreiner, C. E., & Sutter, M. L. (1997). Functional topography of cat primary auditory cortex: response latencies. *J Comp Physiol A*, 181(6), 615-633.
- Merzenich, M. M., Knight, P. L., & Roth, G. L. (1975). Representation of cochlea within primary auditory cortex in the cat. *J Neurophysiol*, 38(2), 231-249.  
doi:10.1152/jn.1975.38.2.231
- Miller, J. L., & Liberman, A. M. (1979). Some effects of later-occurring information on the perception of stop consonant and semivowel. *Percept Psychophys*, 25(6), 457-465.
- Moller, A. R., & Jannetta, P. J. (1982). Auditory evoked potentials recorded intracranially from the brain stem in man. *Exp Neurol*, 78(1), 144-157.
- Molyneaux, B. J., Arlotta, P., Menezes, J. R., & Macklis, J. D. (2007). Neuronal subtype specification in the cerebral cortex. *Nat Rev Neurosci*, 8(6), 427-437.  
doi:10.1038/nrn2151
- Moore, A. K., & Wehr, M. (2013). Parvalbumin-expressing inhibitory interneurons in auditory cortex are well-tuned for frequency. *J Neurosci*, 33(34), 13713-13723.  
doi:10.1523/JNEUROSCI.0663-13.2013
- Moore, A. K., & Wehr, M. (2014). A guide to in vivo single-unit recording from optogenetically identified cortical inhibitory interneurons. *J Vis Exp*(93), e51757.  
doi:10.3791/51757

- Moshitch, D., Las, L., Ulanovsky, N., Bar-Yosef, O., & Nelken, I. (2006). Responses of neurons in primary auditory cortex (A1) to pure tones in the halothane-anesthetized cat. *J Neurophysiol*, *95*(6), 3756-3769. doi:10.1152/jn.00822.2005
- Mountcastle, V. B., Talbot, W. H., Sakata, H., & Hyvarinen, J. (1969). Cortical neuronal mechanisms in flutter-vibration studied in unanesthetized monkeys. Neuronal periodicity and frequency discrimination. *J Neurophysiol*, *32*(3), 452-484. doi:10.1152/jn.1969.32.3.452
- Nelken, I., Prut, Y., Vaadia, E., & Abeles, M. (1994). Population responses to multifrequency sounds in the cat auditory cortex: one- and two-parameter families of sounds. *Hear Res*, *72*(1-2), 206-222.
- Nelken, I., & Versnel, H. (2000). Responses to linear and logarithmic frequency-modulated sweeps in ferret primary auditory cortex. *Eur J Neurosci*, *12*(2), 549-562.
- Neubauer, H., & Heil, P. (2008). A physiological model for the stimulus dependence of first-spike latency of auditory-nerve fibers. *Brain Res*, *1220*, 208-223. doi:10.1016/j.brainres.2007.08.081
- Nowicka, D., Soulsby, S., Skangiel-Kramaska, J., & Glazewski, S. (2009). Parvalbumin-containing neurons, perineuronal nets and experience-dependent plasticity in murine barrel cortex. *Eur J Neurosci*, *30*(11), 2053-2063. doi:10.1111/j.1460-9568.2009.06996.x
- Okada, T., Honda, M., Okamoto, J., & Sadato, N. (2004). Activation of the primary and association auditory cortex by the transition of sound intensity: a new method for functional examination of the auditory cortex in humans. *Neurosci Lett*, *359*(1-2), 119-123. doi:10.1016/j.neulet.2004.02.018
- Palmer, A. R. (1982). Encoding of rapid amplitude fluctuations by Cochlear-nerve fibres in the guinea-pig. *Arch Otorhinolaryngol*, *236*(2), 197-202.

- Parsons, C. E., Young, K. S., Craske, M. G., Stein, A. L., & Kringelbach, M. L. (2014). Introducing the Oxford Vocal (OxVoc) Sounds database: a validated set of non-acted affective sounds from human infants, adults, and domestic animals. *Front Psychol*, 5, 562. doi:10.3389/fpsyg.2014.00562
- Patterson, R. D., Uppenkamp, S., Johnsrude, I. S., & Griffiths, T. D. (2002). The processing of temporal pitch and melody information in auditory cortex. *Neuron*, 36(4), 767-776.
- Perrodin, C., Kayser, C., Logothetis, N. K., & Petkov, C. I. (2011). Voice cells in the primate temporal lobe. *Curr Biol*, 21(16), 1408-1415. doi:10.1016/j.cub.2011.07.028
- Pfingst, B. E., & O'Connor, T. A. (1981). Characteristics of neurons in auditory cortex of monkeys performing a simple auditory task. *J Neurophysiol*, 45(1), 16-34. doi:10.1152/jn.1981.45.1.16
- Phillips, D. P., & Hall, S. E. (1990). Response timing constraints on the cortical representation of sound time structure. *J Acoust Soc Am*, 88(3), 1403-1411.
- Phillips, D. P., Hall, S. E., & Boehnke, S. E. (2002). Central auditory onset responses, and temporal asymmetries in auditory perception. *Hear Res*, 167(1-2), 192-205.
- Phillips, D. P., Semple, M. N., Calford, M. B., & Kitzes, L. M. (1994). Level-dependent representation of stimulus frequency in cat primary auditory cortex. *Exp Brain Res*, 102(2), 210-226.
- Platasa, J., & Pieribone, V. A. (2018). Genetically encoded fluorescent voltage indicators: are we there yet? *Curr Opin Neurobiol*, 50, 146-153. doi:10.1016/j.conb.2018.02.006
- Qin, L., Chimoto, S., Sakai, M., Wang, J., & Sato, Y. (2007). Comparison between offset and onset responses of primary auditory cortex ON-OFF neurons in awake cats. *J Neurophysiol*, 97(5), 3421-3431. doi:10.1152/jn.00184.2007

- Rauschecker, J. P., & Tian, B. (2000). Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proc Natl Acad Sci U S A*, *97*(22), 11800-11806.  
doi:10.1073/pnas.97.22.11800
- Rauschecker, J. P., Tian, B., & Hauser, M. (1995). Processing of complex sounds in the macaque nonprimary auditory cortex. *Science*, *268*(5207), 111-114.
- Reale, R. A., & Imig, T. J. (1980). Tonotopic organization in auditory cortex of the cat. *J Comp Neurol*, *192*(2), 265-291. doi:10.1002/cne.901920207
- Recanzone, G. H. (2000). Response profiles of auditory cortical neurons to tones and noise in behaving macaque monkeys. *Hear Res*, *150*(1-2), 104-118.
- Recanzone, G. H., Schreiner, C. E., & 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.
- Reed, A., Riley, J., Carraway, R., Carrasco, A., Perez, C., Jakkamsetti, V., & Kilgard, M. P. (2011). Cortical map plasticity improves learning but is not necessary for improved performance. *Neuron*, *70*(1), 121-131.  
doi:10.1016/j.neuron.2011.02.038
- Romand, R., & Ehret, G. (1990). Development of tonotopy in the inferior colliculus. I. Electrophysiological mapping in house mice. *Brain Res Dev Brain Res*, *54*(2), 221-234.
- Rothschild, G., Nelken, I., & Mizrahi, A. (2010). Functional organization and population dynamics in the mouse primary auditory cortex. *Nat Neurosci*, *13*(3), 353-360.  
doi:10.1038/nn.2484
- Scholl, B., Gao, X., & Wehr, M. (2010). Nonoverlapping sets of synapses drive on responses and off responses in auditory cortex. *Neuron*, *65*(3), 412-421.  
doi:10.1016/j.neuron.2010.01.020

- Schreiner, C. E., & Sutter, M. L. (1992). Topography of excitatory bandwidth in cat primary auditory cortex: single-neuron versus multiple-neuron recordings. *J Neurophysiol*, *68*(5), 1487-1502. doi:10.1152/jn.1992.68.5.1487
- Schreiner, C. E., & Urbas, J. V. (1988). Representation of amplitude modulation in the auditory cortex of the cat. II. Comparison between cortical fields. *Hear Res*, *32*(1), 49-63.
- Schulze, H., Ohl, F. W., Heil, P., & Scheich, H. (1997). Field-specific responses in the auditory cortex of the unanaesthetized Mongolian gerbil to tones and slow frequency modulations. *J Comp Physiol A*, *181*(6), 573-589.
- Shamma, S. A., Fleshman, J. W., Wiser, P. R., & Versnel, H. (1993). Organization of response areas in ferret primary auditory cortex. *J Neurophysiol*, *69*(2), 367-383. doi:10.1152/jn.1993.69.2.367
- Shannon-Hartman, S., Wong, D., & Maekawa, M. (1992). Processing of pure-tone and FM stimuli in the auditory cortex of the FM bat, *Myotis lucifugus*. *Hear Res*, *61*(1-2), 179-188.
- Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, *270*(5234), 303-304.
- Shepard, K. N., Kilgard, M. P., & Liu, R. C. (2013). Experience-Dependent Plasticity and Auditory Cortex. In Y. E. Cohen, A. N. Popper, & R. R. Fay (Eds.), *Neural Correlates of Auditory Cognition* (pp. 293-327). New York, NY: Springer New York.
- Shepard, K. N., Liles, L. C., Weinshenker, D., & Liu, R. C. (2015). Norepinephrine is necessary for experience-dependent plasticity in the developing mouse auditory cortex. *J Neurosci*, *35*(6), 2432-2437. doi:10.1523/JNEUROSCI.0532-14.2015
- Shepard, K. N., Lin, F. G., Zhao, C. L., Chong, K. K., & Liu, R. C. (2015). Behavioral relevance helps untangle natural vocal categories in a specific subset of core

- auditory cortical pyramidal neurons. *J Neurosci*, 35(6), 2636-2645.  
doi:10.1523/JNEUROSCI.3803-14.2015
- Sinex, D. G., & Geisler, C. D. (1981). Auditory-nerve fiber responses to frequency-modulated tones. *Hear Res*, 4(2), 127-148.
- Slovin, H., Arieli, A., Hildesheim, R., & Grinvald, A. (2002). Long-term voltage-sensitive dye imaging reveals cortical dynamics in behaving monkeys. *J Neurophysiol*, 88(6), 3421-3438. doi:10.1152/jn.00194.2002
- Sollini, J., Chapuis, G. A., Clopath, C., & Chadderton, P. (2018). ON-OFF receptive fields in auditory cortex diverge during development and contribute to directional sweep selectivity. *Nat Commun*, 9(1), 2084. doi:10.1038/s41467-018-04548-3
- Stevens, K. N., & Klatt, D. H. (1974). Role of formant transitions in the voiced-voiceless distinction for stops. *J Acoust Soc Am*, 55(3), 653-659.
- Stickney, G. S., Nie, K., & Zeng, F. G. (2005). Contribution of frequency modulation to speech recognition in noise. *J Acoust Soc Am*, 118(4), 2412-2420.
- 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. *J Comp Physiol A*, 181(6), 559-571.
- Suga, N. (1964). Recovery Cycles and Responses to Frequency Modulated Tone Pulses in Auditory Neurones of Echo-Locating Bats. *J Physiol*, 175, 50-80.
- Sutter, M. L., & Schreiner, C. E. (1995). Topography of intensity tuning in cat primary auditory cortex: single-neuron versus multiple-neuron recordings. *J Neurophysiol*, 73(1), 190-204. doi:10.1152/jn.1995.73.1.190
- Taaseh, N., Yaron, A., & Nelken, I. (2011). Stimulus-specific adaptation and deviance detection in the rat auditory cortex. *PLoS One*, 6(8), e23369.  
doi:10.1371/journal.pone.0023369

- Takahashi, H., Nakao, M., & Kaga, K. (2004). Cortical mapping of auditory-evoked offset responses in rats. *Neuroreport*, *15*(10), 1565-1569.
- Tallal, P., Miller, S. L., Bedi, G., Byma, G., Wang, X., Nagarajan, S. S., . . . Merzenich, M. M. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science*, *271*(5245), 81-84.
- Tang, C., Hamilton, L. S., & Chang, E. F. (2017). Intonational speech prosody encoding in the human auditory cortex. *Science*, *357*(6353), 797-801.  
doi:10.1126/science.aam8577
- Taylor, A. M., Reby, D., & McComb, K. (2009). Context-Related Variation in the Vocal Growling Behaviour of the Domestic Dog (*Canis familiaris*). *Ethology*, *115*(10), 905-915. doi:doi:10.1111/j.1439-0310.2009.01681.x
- Theunissen, F. E., Sen, K., & Doupe, A. J. (2000). Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *J Neurosci*, *20*(6), 2315-2331.
- Thornton, A. R. (1976). Electrophysiological studies of the auditory system. *Audiology*, *15*(1), 23-38.
- Tian, B., Kusmirek, P., & Rauschecker, J. P. (2013). Analogues of simple and complex cells in rhesus monkey auditory cortex. *Proc Natl Acad Sci U S A*, *110*(19), 7892-7897. doi:10.1073/pnas.1221062110
- Tian, B., & Rauschecker, J. P. (2004). Processing of frequency-modulated sounds in the lateral auditory belt cortex of the rhesus monkey. *J Neurophysiol*, *92*(5), 2993-3013. doi:10.1152/jn.00472.2003
- Tsukano, H., Horie, M., Bo, T., Uchimura, A., Hishida, R., Kudoh, M., . . . Shibuki, K. (2015). Delineation of a frequency-organized region isolated from the mouse primary auditory cortex. *J Neurophysiol*, *113*(7), 2900-2920.  
doi:10.1152/jn.00932.2014

- Tsukano, H., Horie, M., Hishida, R., Takahashi, K., Takebayashi, H., & Shibuki, K. (2016). Quantitative map of multiple auditory cortical regions with a stereotaxic fine-scale atlas of the mouse brain. *Sci Rep*, *6*, 22315. doi:10.1038/srep22315
- Tsukano, H., Horie, M., Ohga, S., Takahashi, K., Kubota, Y., Hishida, R., . . . Shibuki, K. (2017). Reconsidering Tonotopic Maps in the Auditory Cortex and Lemniscal Auditory Thalamus in Mice. *Front Neural Circuits*, *11*, 14. doi:10.3389/fncir.2017.00014
- Tsunada, J., Lee, J. H., & Cohen, Y. E. (2012). Differential representation of auditory categories between cell classes in primate auditory cortex. *J Physiol*, *590*(13), 3129-3139. doi:10.1113/jphysiol.2012.232892
- Ulanovsky, N., Las, L., & Nelken, I. (2003). Processing of low-probability sounds by cortical neurons. *Nat Neurosci*, *6*(4), 391-398. doi:10.1038/nn1032
- Volkov, I. O., & Galazjuk, A. V. (1991). Formation of spike response to sound tones in cat auditory cortex neurons: interaction of excitatory and inhibitory effects. *Neuroscience*, *43*(2-3), 307-321.
- Wallace, M. N., Rutkowski, R. G., & Palmer, A. R. (2000). Identification and localisation of auditory areas in guinea pig cortex. *Exp Brain Res*, *132*(4), 445-456.
- Wang, X., Lu, T., Snider, R. K., & Liang, L. (2005). Sustained firing in auditory cortex evoked by preferred stimuli. *Nature*, *435*(7040), 341-346. doi:10.1038/nature03565
- Wang, X., & Walker, K. M. (2012). Neural mechanisms for the abstraction and use of pitch information in auditory cortex. *J Neurosci*, *32*(39), 13339-13342. doi:10.1523/JNEUROSCI.3814-12.2012
- Watwood, S. L., Tyack, P. L., & Wells, R. S. (2004). Whistle sharing in paired male bottlenose dolphins, *Tursiops truncatus*. *Behavioral Ecology and Sociobiology*, *55*(6), 531-543. doi:10.1007/s00265-003-0724-y

- Wehr, M., & Zador, A. M. (2003). Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. *Nature*, *426*(6965), 442-446.  
doi:10.1038/nature02116
- Weible, A. P., Moore, A. K., Liu, C., DeBlander, L., Wu, H., Kentros, C., & Wehr, M. (2014). Perceptual gap detection is mediated by gap termination responses in auditory cortex. *Curr Biol*, *24*(13), 1447-1455. doi:10.1016/j.cub.2014.05.031
- Weinberger, N. M. (2004). Specific long-term memory traces in primary auditory cortex. *Nat Rev Neurosci*, *5*(4), 279-290. doi:10.1038/nrn1366
- Wessinger, C. M., VanMeter, J., Tian, B., Van Lare, J., Pekar, J., & Rauschecker, J. P. (2001). Hierarchical organization of the human auditory cortex revealed by functional magnetic resonance imaging. *J Cogn Neurosci*, *13*(1), 1-7.
- Whitfield, I. C. (1956). Electrophysiology of the central auditory pathway. *Br Med Bull*, *12*(2), 105-109.
- Whitfield, I. C., & Evans, E. F. (1965). Responses of Auditory Cortical Neurons to Stimuli of Changing Frequency. *J Neurophysiol*, *28*, 655-672.  
doi:10.1152/jn.1965.28.4.655
- Wong, P. C., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat Neurosci*, *10*(4), 420-422. doi:10.1038/nn1872
- Young, E. D., & Brownell, W. E. (1976). Responses to tones and noise of single cells in dorsal cochlear nucleus of unanesthetized cats. *J Neurophysiol*, *39*(2), 282-300.  
doi:10.1152/jn.1976.39.2.282
- Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends Cogn Sci*, *6*(1), 37-46.

- Zeng, F. G., Nie, K., Stickney, G. S., Kong, Y. Y., Vongphoe, M., Bhargave, A., . . . Cao, K. (2005). Speech recognition with amplitude and frequency modulations. *Proc Natl Acad Sci U S A*, *102*(7), 2293-2298. doi:10.1073/pnas.0406460102
- Zhang, F., Gradinaru, V., Adamantidis, A. R., Durand, R., Airan, R. D., de Lecea, L., & Deisseroth, K. (2010). Optogenetic interrogation of neural circuits: technology for probing mammalian brain structures. *Nat Protoc*, *5*(3), 439-456. doi:10.1038/nprot.2009.226
- Zhao, L., & Zhaoping, L. (2011). Understanding auditory spectro-temporal receptive fields and their changes with input statistics by efficient coding principles. *PLoS Comput Biol*, *7*(8), e1002123. doi:10.1371/journal.pcbi.1002123
- Zingg, B., Chou, X. L., Zhang, Z. G., Mesik, L., Liang, F., Tao, H. W., & Zhang, L. I. (2017). AAV-Mediated Anterograde Transsynaptic Tagging: Mapping Corticocollicular Input-Defined Neural Pathways for Defense Behaviors. *Neuron*, *93*(1), 33-47. doi:10.1016/j.neuron.2016.11.045