Title
Psychophysics and Neurophysiology of Stream Segregation in the Cat

Author
Javier, Lauren

Publication Date
2016

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, IRVINE

Psychophysics and Neurophysiology of Spatial Stream Segregation in the Cat

DISSERTATION

submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in Biological Sciences

by

Lauren Krystal Javier

Dissertation Committee:
Professor John C. Middlebrooks, Chair
Professor Raju Metherate
Professor Virginia M. Richards
Professor Georg F. Striedter

2016

DEDICATION

To

My family and my beloved
For their prayers, love, and support

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

ACKNOWLEDGMENTS

This dissertation work has been years in the making and would not have been possible without the dedicated support and guidance from my mentors, colleagues, family, and friends.

First and foremost, I wish to express my deepest gratitude and appreciation to my mentor, Dr. John Middlebrooks for his continuous support, patience, and immense knowledge. John, thank you for guiding me in my research and in writing this thesis. I could not have imagined a better mentor for my Ph.D. study. I have learned and grown so much since joining the lab. I would also like to thank the members of my dissertation and advancement committees, Dr. Raju Metherate, Dr. Georg Striedter, Dr. Virginia Richards, and Dr. Karina Cramer. Thank you for your helpful feedback and encouragement. You have all been great cheerleaders in helping me believe in myself, especially in times of doubt.

I would like to thank everyone in the Middlebrooks Laboratory. It took a team to orchestrate my dissertation work, and I am grateful for the unwavering support from past and current members of the lab. A special thank you to Elizabeth McGuire for helping with all things cats! We've literally shed blood, sweat, and tears over the past several years and I could not have gotten this far without you. Thank you to my lab "siblings," Justin Yao and Akshat Arneja for your constant encouragement. Justin thank you so much for your help in MATLAB! Zekiye Onsan, thanks so much for being an amazing "lab mom" and for your willingness to help whenever I needed it. Pauline Herbert-Whiting, I cannot thank you enough for helping me with cat training. I'd also like to thank Peter Bremen, Alana Kirby, and Chen-Chung Lee for helping me get my feet wet. In addition to the lab, I'd like to thank ULAR and the UCI veterinarian staff for your expertise and assistance in cat care. I would especially like to thank Dr. Roger

Geertsema, Dr. Claire Lindsell, Annie Ristig, Andrea Linden, and Rommel Medina for ensuring that our cats received the best care.

Thank you to my family and friends for their constant prayers, support, and encouragement over the years. Thank you for keeping me sane and for always believing in me. Mom and dad, I cannot thank you enough for all of your sacrifices. Larielle and Blacky, thank you for the pep talks and for helping me look and feel my best. Jason, you have been with me at the start of this journey. Thank you for inspiring me to always put my best foot forward. I cannot wait for our next big adventure.

Lastly, I am indebted to the cats that have contributed to this research. Although my aim was to teach them an auditory task, they too, have taught me. Thank you for teaching me to be persistent in accomplishing my goals, to be stringent in my time management, and to find joy in the simplest of things. I'd especially like to thank Pumpkin, who taught me that I could love cats too.

These studies were supported by the following grants: NIH RO1 DC000420 and DOE GAANN P200A120165.

CURRICULUM VITAE

Lauren Krystal Javier

EDUCATION

Ph.D., School of Biological Sciences June 2016
Department of Neurobiology and Behavior, University of California, Irvine

M.S., School of Biological Sciences June 2015
Department of Neurobiology and Behavior, University of California, Irvine

Interdisciplinary Neuroscience Program 2009-2010
University of California, Irvine

B.A. Biology December 2008
University of Hawaii, Manoa

RESEARCH EXPERIENCE

Graduate Student Researcher, Ph.D. Candidate 2010-2016
Advisor: Dr. John Middlebrooks
Department of Neurobiology and Behavior, University of California, Irvine

- Psychophysics and neurophysiology of spatial stream segregation in the cat

Undergraduate Student Researcher 2007-2009
Advisors: Dr. Robert Blanchard and Dr. Caroline Blanchard
Minority Access to Research Careers (MARC) Program, University of Hawaii, Manoa

- Anxiolytic effects of a β-adrenoceptor agonist on male CD-1 mice during scent marking behavior.
- c-Fos expression during neural activation in male C57BL/6J mice in scent marking behavior.

Undergraduate Summer Research 2008-2008
Advisor: Dr. John Guzowski
Summer Undergraduate Research Fellowship (SURF) Program, University of California, Irvine

- The effect of cytokines on cognitive function in behavioral, neural circuit, and gene expression.

TEACHING EXPERIENCE

Teaching Assistant; Head TA Fall 2015
Undergraduate Introductory Biology Discussion Course (Bio 93)
School of Biological Sciences, University of California, Irvine

- Designed student-centered learning curriculum to be used by teaching assistants
- Provided pedagogical training to teaching assistants

Undergraduate Research Mentor Summer 2015
School of Biological, University of California, Irvine

- Mentored and supervised an undergraduate research assistant in conducting experiments in animal behavior

Pedagogical Fellows (PF) Program Spring 2014- Winter 2015
Teaching, Learning & Technology Center, University of California, Irvine

- Designed and implemented discipline-specific training for new and experienced teaching assistants
- Provided teaching consultations including analysis, feedback and suggestions to instructors looking to improve their teaching

Teaching Assistant Fall 2014
Undergraduate Introductory Biology Discussion Course (Bio 93)
School of Biological Sciences, University of California, Irvine

- Taught 3 discussion sections of undergraduate students for an introductory biology course
- Designed student-centered learning activities and assignments, and graded student assignments and quizzes

Teaching Assistant Summer 2014
Undergraduate Brain and Behavior Course (Bio 35)
Department of Neurobiology and Behavior, University of California, Irvine

- Graded assignments and exams, proctored exams

Teaching Assistant; Course Coordinator/ Head TA Spring 2013, Winter-Spring 2014
Undergraduate Neurobiology Lab (Bio N113L)
Department of Neurobiology and Behavior, University of California, Irvine

- Trained a group of graduate teaching assistants how to teach the course
- Oversaw all lab course sections which involved providing a lab overview for teaching assistants, assisting with troubleshooting, and ensuring the lab sections ran efficiently
- Coordinated communication between students, teaching assistants, lab leaders, and faculty course director
- Coordinated the construction of the course final examination and grading the exam
- Worked with lecturers from other departments in Biological Sciences to develop assessment methods for student writing assignments
- Developed curriculum for a flipped laboratory classroom, including making lecture and lab demonstration videos.

Guest Lecturer Fall 2013
Undergraduate Brain and Behavior Course (Bio 35)
Department of Neurobiology and Behavior, University of California, Irvine

- Guest lectured to a class of 350 undergraduates on psychomotor stimulants, the dopaminergic system and popular drugs such as nicotine and caffeine.

Teaching Assistant Spring 2012, Winter 2013
Undergraduate Neurobiology Lab (Bio N113L)
Department of Neurobiology and Behavior, University of California, Irvine

- Taught a class of 17 undergraduate students
- Created lectures, quizzes, and contributed to the creation of the final examination
- Set up laboratory equipment and assisted in necessary troubleshooting
- Graded quizzes and written assignments
- Assisted in implementing new lab experiments for the course

Teaching Assistant Fall 2012
Undergraduate Auditory Neuroscience Course (Bio N147/Psych 161H)
Department of Neurobiology and Behavior, University of California, Irvine

- Assisted in examinations which included organizing the exam, proctoring, grading, and meeting with students

Teaching Assistant Winter 2012
Graduate Level Cellular Neuroscience Lab (NBB 207L)
School of Biological Sciences, University of California, Irvine

- Assisted Dr. Ian Parker in a graduate level, cellular neurophysiology laboratory course.
- Assisted in experimental setups including animal dissections and equipment setup for students to conduct neurophysiology experiments
- Developed a Teaching Assistant Manual as a reference for future TAs assisting in the course.

Teaching Intern Spring 2007
Undergraduate Introductory Biology Lab
Department of Biology, University of Hawaii, Manoa

- Undergraduate teaching intern for Dr. Charles Riley who provided training for teaching the course.
- Assisted the Teaching Assistant in running the course by assisting students with questions, grading quizzes and exams, and giving 2 lectures during the semester.

ACADEMIC AND PROFESSIONAL HONORS

Association for Research in Otolaryngology Travel Award ($500) 2015

Pedagogical Fellows (PF) Program ($2000) 2014-2015
University of California, Irvine

- For teaching excellence and contributions to quality undergraduate education

Graduate Assistance in Areas of National Need Program 2013-2015
University of California, Irvine

Graduate Division Award in Neuroscience ($1000) 2015
University of California, Irvine

Minority Access to Research Careers Program 2007-2009
University of Hawaii, Manoa

Associated Students of the University of Hawaii Research Award 2008
University of Hawaii, Manoa

Summer Undergraduate Research Fellowship (SURF) Program 2008
University of California, Irvine

Associated Students of the University of Hawaii Research Award 2008
University of Hawaii, Manoa

Honorable Mention: Best Overall Presentation 2008
SURF Program, University of California Irvine

PUBLICATIONS

L.K. Javier, E.A. McGuire, J.C. Middlebrooks (2016) Spatial stream segregation by cats. *Journal for the Acoustical Society of America* 17: 195-207

K.G. Borelli, D.C. Blanchard, **L.K. Javier,** E.B. Defensor, M.L Brandao, R.J. Blanchard (2009) Neural correlates of scent marking behavior in C57BL/6J mice: detection and recognition of a social stimulus. *Neuroscience* 162: 914-23.

PRESENTATIONS

L. Javier, J.D. Yao., J.C. Middlebrooks. Monkey Business in Cat Auditory Cortex. Association for Research in Otolaryngology 39th Annual Midwinter Meeting; 2016 Feb 20-24; San Diego, CA, ARO, 2016 Pg. 132. Poster number 719

N. Goldberg, **L. Javier**, J. Overman, A. Nicholas. Flipping a Course: Habituation in *C. elegans*. Regional Association for Biology in Laboratory Education (RABLE) Conference, UCI, February 2015

L. Javier, E. Hand, J.C. Middlebrooks. Psychophysical Evaluation of Auditory Spatial Stream Segregation in the Cat. SoCal Hearing Conference, UCI, September 2014

L. K. Javier, E.A. McGuire, J.C. Middlebrooks. 2014. Spatial Stream Segregation in Cat Psychophysics. Association for Research in Otolaryngology 37th Annual Midwinter Meeting; 2014 Feb 22-26; San Diego, CA, ARO, 2014 Pg. 249. Poster number 403

L. Javier, E. Hand, J.C. Middlebrooks. Psychophysical Evaluation of Auditory Spatial Stream Segregation in the Cat. SoCal Hearing Conference, USC, September 2013

K.G Borelli, E. Defensor, **L. Javier**, M.L. Brandao, D.C. Blanchard, R.J, Blanchard. Distribution of Fos Immunoreactivity in the C57BL/6J Mice Brain Associated with Scent Marking Behavior. Society for Neuroscience Conference, Washington D.C., November 2008

L. Javier, T. Cole, G. Lewandowski, J. Guzowski. Systemic Lipopolysaccharide Induces Cytokine Production in the Hippocampus and Impairs Spatial Reversal Learning. SURF Program Symposium, Irvine, CA, August 2008

L. Javier, K.G. Borelli, M.L. Brandao, D.C. Blanchard, R.J. Blanchard. Distribution of Fos Immunoreactivity in the C57BL/6J Mouse Brain Associated with Scent Marking Behavior. Spring Symposium of Manoa Undergraduate Research and Creative Expression, Honolulu, HI, May 2008

L. Javier, H. Arakawa, K. Arakawa, R.J. Blanchard. SR56811A, a β_3-Adrenoceptor Agonist Reduces Scent Marking but not Aggression in the Mouse. ABRCMS Conference, Austin, TX, November 2007

PROFESSIONAL DEVELOPMENT

Course Design Certificate Program, May 2015
University of California, Irvine

- Program on designing a completed curriculum for a course using a backwards, integrated course design beginning with student learning goals and how the class will be taught using active learning.

University Studies Graduate Courses 390A/B/C March 2014-February 2015
University of California, Irvine

- 390A: Coursework to learn about Social Constructivist learning epistemology, active learning strategies, and designing learning-centered curriculum
- 390B: Emphasis on academic job preparation, providing teaching consultations, and developing teaching workshops for the academic year
- 390C: Training in how to conduct peer reviews, consult with course instructors, design rubrics for assessing interviews, and conduct interviews. Selection of new Pedagogical Fellows.

Science, Technology, Education and Math Education September 2013-Present
University of California, Irvine

- Monthly group discussion meetings on higher education student-centered learning techniques

TA Training Workshop: Dealing with Evaluations November 2013
University of California, Irvine
Teaching, Learning & Technology Center

- Workshop on assessing evaluations and developing techniques to improve teaching

Power of Speech Presentation Skills Workshop October 2012
University of California, Irvine

- Interactive workshop that challenges participants to develop communication skills that have an impact on others

SERVICE

Teaching Assistant Professional Development Program Fall 2015
Pedagogical Fellow/ Facilitator

- Designed and facilitated a 1.5-Day program for incoming graduate student teaching assistants. Training included designing learning objectives, incorporating student-centered active learning, learning styles, student diversity, and teaching responsibilities

Foundations of Teaching Program Spring 2015
Micro-Teaching Facilitator

- Program designed by the Graduate Research Center to train graduate students and postdoctoral fellows on research-based teaching strategies to increase learning and maximize student engagement
- Served as a Pedagogical Fellow facilitator and provided feedback on participants' micro-teaching

Brain Bee Fall 2015
Graduate Student Volunteer

- Program designed to stimulate high school students to learn about neuroscience
- Proctored and graded written exams which includes a practical component involving neuroanatomy

Teaching Workshops Fall 2014
Workshop Organizer and Facilitator
Teaching, Learning & Technology Center

- Designed and facilitated two workshops: Effective Lesson Planning and How to Write Better Tests to train graduate students teaching assistants

Teaching Assistant Professional Development Program Fall 2014
Pedagogical Fellow/ Facilitator

- Designed and facilitated a 2-Day program for incoming graduate student teaching assistants. Training included designing learning objectives, incorporating student-centered active learning, learning styles, student diversity, and teaching responsibilities

New Faculty Teaching Academy: How to Make the Most of Your TAs Fall 2014
Pedagogical Fellow

- Participated as a TA representative to answer faculty questions on TA roles and responsibilities

Regional Intel International Science and Engineering Fair Spring 2014
Judge, Qualifying Competition

- Worked with a cohort of judges to select two high school students and 2 alternates from Orange County to compete in the International ISEF competition

Southern California Forum for Diversity in Graduate Education Spring 2013
Graduate Student Volunteer

- Assisted with graduate recruitment under the UCI Graduate Division

- Met with underrepresented undergraduate and Master's students who are interested in pursuing a Ph.D. by sharing graduate school experience, informing them about available programs, etc.
- Assisted with event workshops, being a mediator between speakers and attendees.

Graduate Student Panel Fall 2012

Graduate Student Representative

- Participated as part of a panel of graduate students to assist undergraduates interested in applying to graduate school by answering questions about graduate school education.

ABSTRACT OF THE DISSERTATION

Psychophysics and Neurophysiology of Spatial Stream Segregation in the Cat

By

Lauren K. Javier

Doctor of Philosophy in Biological Sciences

University of California, Irvine, 2016

Professor John C. Middlebrooks, Chair

Listeners have the remarkable ability to disentangle multiple competing sound sequences and organize this mixture into distinct sound sources. A previous study in human listeners has shown that the physical separation between sounds aids in "segregating" between sound sources, whereby sounds located further apart in space are more easily segregated. Furthermore, under anesthetized conditions, animal neurophysiology has been found to parallel conditions in which humans hear one stream or multiple streams. The goal for this dissertation is to evaluate the psychophysics of spatial stream segregation and, in the same species, record neural activity in auditory cortex in the absence of anesthesia. Cats have been used extensively in auditory research due to their well-developed auditory cortex and because they have evolved accurate sound localization ability to support their nocturnal predatory behavior. We developed a novel paradigm testing the spatial resolution of stream segregation in cats to measure psychophysical performance (Chapter 2) and to uncover the spatial cues that are utilized by cats to perform this task. We then implanted chronic electrodes into primary auditory cortex to record single- and multiple-unit neural activity in awake cats (Chapter 3 and Chapter 4). Our findings show that: (1) Cats can segregate streams of broadband sounds with spatial acuity approaching that of

humans. In addition, performance was consistently better for high than for low frequencies which is consistent with previous cat physiological results but contrary to human psychophysics. (2) In the absence of anesthesia, neurons in cat cortex exhibit spectral and temporal properties which are not seen in anesthetized preparations but accord with previous observations in unanesthetized marmosets. (3) Lastly, neurons in auditory cortex of awake cats that are not engaged in an overt auditory task exhibit considerably weaker stream segregation than is observed in anesthetized preparations. Although there is little evidence for stream segregation while cats are not engaged in an auditory task, it may be that segregation is influenced by selective attention or that spatial stream segregation is processed in cortical areas beyond A1. Overall, these findings provide insight into auditory mechanisms underlying stream segregation and neural properties of the unanesthetized cortex.

CHAPTER 1

General Introduction

1.1 Auditory Scene Analysis

In a complex acoustic environment, sounds arising from different sources sum and enter the pinna, or external ear. In such an environment, listeners have the remarkable ability to organize this mixture of competing sounds by "grouping" multiple spectral and temporal components that belong to the same auditory source and "segregating" other components that belong to different sources. As an example, a listener could effectively follow and converse with a speaker of interest in a noisy cocktail party setting by attending to strings of phonemes from that speaker while separating that of other speakers. In addition, this listener can selectively switch his or her attention towards another speaker of interest or to other auditory sources present at the cocktail party such as background music. This phenomenon has been referred to as the cocktail party problem (Cherry, 1953) or auditory scene analysis (ASA) (Bregman, 1990). Sequences of sound originating from multiple sources may be grouped together as a single perception stream, known as stream integration, or perceived as distinct streams, known as stream segregation.

Segregation that occurs when the elements of competing sound sequences overlap in time, refers to concurrent segregation. In contrast to concurrent stream segregation, sequential stream segregation refers to two competing streams that are temporally interleaved in time with no temporal overlap between the elements of the two streams. When both competing sound sequences overlap in spectrum and/or time, this results in masking of the sound sequences

whereby listeners have difficulty segregating between competing streams. In energetic masking, sound sequences overlap in spectral excitation along the basilar membrane and auditory nerve at the same time rendering the distinct sound sequences as inaudible to the listener. In contrast, informational masking is involved with the central auditory system and is quantitatively defined as all masking not pertaining to energetic masking. In cases where interleaved sequences of sound are similar (i.e., target and masker overlap in spectrum but not in time such as a speech signal masked by speech), informational masking could occur whereby listeners are unable to disentangle elements of the target and competing sound sequences (Durlach et al., 2003; Brungart 2005).

Sequential stream segregation has often been demonstrated using repeating ABAB or ABA-ABA stimulus sequences where A and B refer to pure tones of different frequencies (van Noorden, 1975; Bee and Micheyl, 2008; Bregman 1990). In addition to tone-based stream segregation, other factors have been found to facilitate segregation. These include, differences in temporal envelope (Vliegen and Oxenham, 1999), timbre (for example, Wessel, 1979; Iverson, 1995; Singh and Bregman, 1997), component phases of complex tones (Roberts et al., 2002), and the spatial separation (for example, Middlebrooks and Onsan, 2012; Middlebrooks and Bremen, 2013; Yao et al., 2015) between competing sound sources (reviewed by Moore and Gockel, 2002). In this dissertation, we will focus on sequential stream segregation and how the contribution of spatial differences aid in the segregation of competing sound sources.

Although normal hearing listeners can segregate between different sound sources with high fidelity, hearing-impaired individuals with mild-to-moderate hearing loss often report the inability to segregate multiple speakers and the inability to understand speech in a noisy environment (Gatehouse and Nobel, 2004). Thus, understanding the neural correlates and

mechanisms underlying auditory streaming will aid in therapies for improving hearing in complex acoustic environments. This could lead to enhanced diagnosis and treatment for central auditory processing disorders.

1.2 Spatial Hearing as a Factor in Stream Segregation

The ability to localize and segregate various sound sources is important for organizing sounds present in the environment. For animals, the necessity of segregating the presence of prey or avoiding a predator in the presence of other sounds in an acoustically complex environment is paramount to survival. In a cocktail party setting, the ability to segregate a speaker of interest in the presence of other speakers or background noise plays an important role in communication. Locations of sound sources are not mapped topographically in the auditory system. Rather, sound location is computed through converging inputs from the two ears and is further processed and transformed along the ascending auditory system. For sound sources located to the left and/or right of a listener, the difference in path length to each ear results in an interaural difference in sound level (interaural level difference) and an interaural difference in time of sound arrival (interaural time difference) between the two ears. This is commonly referred to as the "duplex" theory of sound localization (Raleigh, 1907). Interaural level difference (ILD) is the dominant spatial cue for high frequency sounds. The wavelength of high frequency sounds is smaller than the size of a listener's head. When sound is presented off to the side, the sound level at the further ear would be at a lower level due to the attenuation of sound waves created by the head shadowing effect. This could provide up to a 35 dB level difference between the two ears (Middlebrooks et al., 1989). At lower frequencies, the sound wave is larger than the size of a person's head rendering ILD negligible. Interaural time difference

(ITD) provides a salient cue for low frequency sounds. For low frequency sounds originating from the side, sound waves reach one ear prior to reaching the other ear. Sensitivity to ITD cues is mostly limited to frequencies below 1500 Hz. In this dissertation, we study these spatial cues in cat psychophysics.

In vertebrates, the location of a sound source is first computed in the brainstem in the superior olivary complex (SOC) through the analysis of acoustic cues such as ITD and ILD. In the SOC, the lateral superior olive (LSO) is the first site of ILD detection and it measures the difference of sound levels presented at the two ears by measuring the balance between afferent inhibition and excitation (Tollin and Yin, 2002; Tollin and Yin, 2005). The medial superior olive (MSO) in the SOC acts as coincidence detectors, with maximal responses when there is simultaneous excitatory input from the ipsilateral and contralateral sides of the brainstem (Goldberg and Brown, 1969; Yin and Chan, 1990; Spitzer and Semple, 1995). This is followed by the synthesis of auditory targets beginning in the inferior colliculus (ICC) (Fujita and Konishi, 1991; McAlpine et al., 1998; McLaughlin et al., 2014). From the ICC, projections are made to the medial geniculate body (MGB) (Calford, 1983; Proctor and Konishi, 1997) and then to auditory cortex (Middlebrooks and Pettigrew, 1981 Jenkins and Masterton, 1982; Middlebrooks and Zook 1983; Jenkins and Merzenich, 1984; Malhotra et. al., 2004; Malhotra et. al., 2008). In A1, the locations of sounds are found to be represented by a distributed code through the coordination of a large population of neurons. Spike rate, first-spike latency and higher-order temporal characteristics are influenced by the location of a sound source in the auditory environment (Middlebrooks et al., 1994; Middlebrooks et al., 1998; Middlebrooks et al., 2002; Stecker and Middlebrooks, 2003).

In the literature, there are conflicting results in which spatial cues either exhibit a weak or robust factor in facilitating stream segregation. These results are dependent on the psychophysical tasks used. For tasks in which listeners are instructed to integrate competing streams (i.e., hear one stream), the strength of segregation is based on the ability of a factor, such as spatial cues, to disrupt integration and thereby perceive segregation. In this case, "obligatory" stream segregation is observed, where no amount of bias of the listener to hear coherence (i.e., integration) could prevent the perception of stream segregation (van Noorden, 1975; Moore and Gockel, 2002; Fishman and Steinschneider, 2010). In tasks requiring listeners to integrate competing streams, spatial cues produce a weak effect in disrupting stream integration (Boehnke and Phillips 2005; Stainsby et al., 2011). Moreover, when competing streams are presented concurrently, whereby sound sequences share a common onset or fundamental frequency (i.e., energetic masking), spatial cues are insufficient in enabling the perception of stream segregation (Broadbent and Ladefoged, 1957; Cutting, 1976; Darwin, 1997; Moore and Gockel, 2002; Shinn-Cunningham, 2005; Moore and Gockel, 2012; Schwartz et al., 2012).

Other studies show that the spatial separation of a target signal and masker has a large effect for tasks that require listeners to segregate between competing sound sources. In this case, spatial cues improve performance on stream segregation tasks. For real-life situations, these tasks are similar to a listener following a conversation at a crowded cocktail party. In such an environment, discriminating between talkers involves a combination of energetic and informational masking. The spatial separation of a target and masker sound produces a large effect for release of informational masking when there is no spectral or temporal overlap between the competing sound sequences (van Noorden, 1975; Bregman, 1990; Hartmann and Johnson, 1991; Kidd, et al., 1994; Kidd et al., 1998; Arborgast et al., 2002). Shinn-Cunningham

(2005) noted in her review that spatial unmasking, or the release from masking due to a spatial separation of a signal with a masker, is largely influenced by spatial attention. This pertains to "voluntary" stream segregation in which the listener has some control over the percept (Bregman, 1990). Spatial attention allows a listener to focus their attention on the target location in the presence of background sound and devote more computational resources towards object selection and formation (Ihlefeld and Shinn-Cunningham, 2008). In addition, if listeners are cued to the location of the target sound, they are easily able to detect a change to that sound (Eramudugolla et al., 2005).

A study by Middlebrooks and Onsan (2012) demonstrated that listeners are able to utilize spatial cues to segregate temporally interleaved sequences of sounds. Sound stimuli were identical in spectra, however, the sound sequences did not have any temporal overlap, eliminating energetic masking. Using a rhythmic masking release (RMR) task (adapted from Sach and Bailey, 2004), listeners were instructed to discriminate between two temporal rhythms in the presence of an interleaved masker that varied in location along azimuth. When the two sound sources were presented from the same location, listeners perceived an indiscriminate sequence of sound and could not identify the temporal rhythm. Only when the sound sources were spatially separated, could listeners identify the target rhythm, which could only be done through stream segregation. This is referred to as spatial release from (informational) masking. Findings show that a spatial separation between the target and masker of ~8° was sufficient to provide robust stream segregation. In addition, Middlebrooks and Onsan found that human listeners performed better under low-band sound conditions, where ITD is the predominant cue, than high-band conditions, where ILD is the predominant spatial cue. Taken together, this study suggests that under conditions in which the listener is focused on a target sound, segregation is

robust when the target sound and masker are spatially separated in azimuth and that ITD cues provide the highest spatial acuity for stream segregation.

Although spatial cues play an important role in facilitating stream segregation and in identifying sound-source locations, a previous study using human psychophysics suggests that sound localization is not necessary for spatial release from masking. In addition to studying the utility of spatial cues in stream segregation, Middlebrooks and Onsan (2012) also studied minimum audible angles (MAA) to measure the acuity of sound localization. If spatial cues are utilized similarly for localization and stream segregation, threshold performance in the MAA task would be similar to performance in the spatial stream segregation task. As previously described in the stream segregation task, human listeners performed well under broadband and low-band conditions, suggesting that ITD was the dominant spatial cue. For the MAA task, stimuli were also presented using broadband, high-band, and low-band conditions. Contrary to performance in the segregation task, human listeners performed similarly well for all passband conditions. This suggests that spatial cues are utilized differently between stream segregation and localization and that localization may not be necessary for spatial release from masking.

Neural correlates of spatial stream segregation have been observed in single neurons in primary auditory cortex of anesthetized cats and rats (Middlebrooks and Bremen, 2013; Yao et al., 2015). Interleaved noise bursts were presented in an alternating ABABAB… pattern, where the A source (i.e., target) was always presented at a fixed location, and the B source (i.e., masker) varied in location from trial to trial. The location of the competing masker sound influenced whether neurons synchronized their responses to the target sound or the masker sound. For example, if the masker was presented contralateral to the recording site, a majority of neurons preferentially synchronized to the masker sound source while only a minority

synchronized to the ipsilateral source. If, however, the masker was presented on the ipsilateral side, these cortical units tended to synchronize to the target sound. A previous study in anesthetized rats suggests that spatial stream segregation arises in two parallel pathways for auditory space processing (Yao et al., 2015). In the tectal pathway, spatial stream segregation arises in the brachium of the inferior colliculus (BIN) and projects to the superior colliculus. In the lemniscal pathway, spatial stream segregation is prominent in a subpopulation of neurons in the ventral division of the medial geniculate body (MGBv) of the thalamus which then projects to primary auditory cortex. In the BIN and MGBv spatial stream segregation is explained by neurons tuned to sound stimuli located in the contralateral hemifield, thereby synchronizing their responses to the most contralateral of the two sound sources. In primary auditory cortex, spatial stream segregation is also largely explained by the sharpening of spatial tuning in those neurons which is further enhanced by a low-pass envelope filter due to forward suppression (Middlebrooks and Bremen, 2013; Yao et al., 2015).

Topics addressed in this dissertation. Previous studies have demonstrated that human listeners and neurons in the auditory cortex in anesthetized cats and rats can segregate sequences of sounds for two competing sources with high spatial acuity (Middlebrooks and Onsan, 2012; Middlebrooks and Bremen, 2013; Yao et al., 2015). The goal for this dissertation is to explore the neurobiological basis of auditory scene analysis by evaluating the psychophysics of spatial stream segregation and, in the same species, recording neural activity in auditory cortex in the absence of anesthesia. Cats have been used extensively in auditory research for their remarkable localization ability as predatory animals and also, for their ability to perform well using objective behavioral methods. In addition, their well-developed auditory cortex is easily accessible on the surface of the brain making it amenable for neurophysiological recordings. We first evaluate the

spatial acuity and putative spatial cues for stream segregation in cat psychophysics (Chapter 2). Here we show that cats exhibit sequential stream segregation with spatial acuity that approaches that of human listeners. Contrary to humans however, cats achieve higher-acuity stream segregation based on high-frequency interaural level differences rather than low-frequency interaural time differences. We then implanted those cats with chronic recording arrays to record single- and multiple-unit neural activity in awake cats not engaged in an auditory task. We first assess the spectral and temporal coding properties of neurons (Chapter 3) followed by studying neural correlates of spatial stream segregation (Chapter 4). In the absence of anesthesia, cortical units exhibited spectral and temporal properties that are rarely seen in anesthetized conditions. We observed a variety of frequency response areas (FRAs) including the typical V-shaped FRAs seen in anesthetized cats and also, level-tolerant I- and O-shape responses. In addition, neurons were sharply tuned and majority of neurons exhibited a sustained response to tonal stimuli. These cortical units showed synchronized responses to both clicks and noise bursts, some at rates $>40\ s^{-1}$. Lastly, we investigated the neural correlates of stream segregation by recording from awake cats (Chapter 4). Interestingly, we observed that neurons phase-locked to both competing streams but showed a slightly stronger response to one sound source. This heightened response however, did not qualify as stream segregation based on our discrimination index criteria, which is contrary to what is seen in anesthetized conditions. It may be that spatial stream segregation is influenced by selective attention or that segregation is processed in cortical areas beyond A1. Taken together, this dissertation provides insight into the auditory mechanisms underlying aspects of spatial stream segregation. It has implications in contributing to solving the complex listening problem and could lead to the enhancement of treating and diagnosing central auditory processing disorders.

CHAPTER 2

Spatial Stream Segregation by Cats

This work has been published in JARO 17(3): 195-207, 2016.

2.1 Summary

Listeners can perceive interleaved sequences of sounds from two or more sources as segregated streams. In humans, physical separation of sound sources is a major factor enabling such stream segregation. Here, we examine spatial stream segregation with a psychophysical measure in domestic cats. Cats depressed a pedal to initiate a target sequence of brief sound bursts in a particular rhythm and then released the pedal when the rhythm changed. The target bursts were interleaved with a competing sequence of bursts that could differ in source location but otherwise were identical to the target bursts. This task was possible only when the sources were heard as segregated streams. When the sound bursts had broad spectra, cats could detect the rhythm change when target and competing sources were separated by as little as 9.4°. Essentially equal levels of performance were observed when frequencies were restricted to a high, 4-to-25-kHz, band in which the principal spatial cues presumably were related to sound levels. When the stimulus band was restricted to 0.4 to 1.6 kHz, leaving interaural time differences as the principal spatial cue, performance was severely degraded. The frequency sensitivity of cats in this task contrasts with that of humans, who show better spatial stream segregation with low- than with high-frequency sounds. Possible explanation for the species difference include the smaller interaural delays available to cats due to smaller sizes of their

heads and the potentially greater sound-level cues available due to the cat's frontally directed pinnae and higher audible frequency range.

2.2 Introduction

In typical auditory environments, listeners show a remarkable ability to isolate sounds of interest amid other competing sounds. This has been referred to as the cocktail party effect (after Cherry, 1953) or auditory scene analysis (Bregman, 1990). One key element of auditory scene analysis is stream segregation, which permits listeners to disentangle multiple temporally interleaved sequences of sounds. An example of stream segregation is that of a listener streaming together sequences of syllables as sentences from one talker while rejecting syllables from one or more other competing talkers. Multiple acoustic features enable stream segregation, including fundamental frequency, temporal envelope, bandwidth, phase, and lateralization (Moore and Gockel, 2002). The present study focuses on the contribution of spatial separation between the sources of the target and distracting sounds.

Previous research in our laboratory has evaluated spatial stream segregation by human listeners, using a non-verbal, objective measure. Listeners were asked to discriminate rhythms of target sequences of broadband noise bursts that were masked by interleaved noise-burst sequences (Middlebrooks and Onsan, 2012). Performance was at chance levels when the target and masker sources were co-located, but improved with increasing target/masker spatial separation. The median rhythmic masking release threshold was 8.1°, which approached those listeners' minimum audible angles for discriminating changes in source locations of single sound bursts. Performance was not significantly different when the noise bursts were band-limited to 0.4 to 1.6 kHz, but thresholds broadened significantly to a median of 15.9° when tested with

bursts band-limited to 4 to 16 kHz. Those results suggest that interaural time differences (ITD) in temporal fine structure were the acoustic cues that provided the highest spatial acuity for humans in that task. A related study examined correlates of spatial stream segregation by neurons in cortical area A1 of anesthetized cats (Middlebrooks and Bremen, 2013). Neurons synchronized preferentially to one or the other of two interleaved sound sequences from spatially separated sources with spatial acuity approaching that of the human listeners in the psychophysical task. Contrary to the expectation from the human results, however, acuity of spatial stream segregation by cat neurons was by most tests finer among neurons tuned to high frequencies than among those tuned to low frequencies.

The purpose of the present study was to evaluate the spatial acuity of stream segregation in cats, thereby providing psychophysical data for the same species in which data from single cortical neurons can be obtained. In particular, we wished to test whether the cat listeners showed finer spatial acuity at low frequencies, like the human listeners, or finer acuity at high frequencies, consistent with the cat cortical physiology. For this study, we modified the task from the two-alternative task employed in the previous human psychophysical study (Middlebrooks and Onsan, 2012) to a hold-release paradigm. Cats depressed a pedal to begin presentation of a standard sound sequence, Rhythm 1, and then released the pedal when the sequence changed to Rhythm 2. The target sounds were interleaved with masker sequences that varied in source location from trial to trial. It was necessary for the cat to segregate target from masker streams in order to detect the change in rhythm and thereby receive a food reward.

The results demonstrate that cats can segregate streams of broadband sounds with spatial acuity nearly as fine as that of humans. Performance was consistently better for high than for low frequencies, consistent with the previous cat physiological results but contrary to the human

psychophysics. Factors contributing to that inter-species difference in frequency dependence could include the narrower interaural time differences provided by the smaller head of the cat as well as potentially greater sound-level cues available due to the cat's frontally directed pinnae and higher audible frequency range.

2.3 Materials and Methods

Animals. All procedures were in accordance with the NIH Animal Welfare Guidelines and with a protocol approved by the Institutional Animal Care and Use Committee at the University of California at Irvine. Six male domestic shorthaired cats (*Felis catus*) were obtained from a breeding colony at the University of California at Davis. No hearing deficits were evident. Ages ranged from 2 to 6 months at the beginning of training and from 8 to 36 months at the time of collection of the reported data. Male cats were used exclusively for this study to facilitate group housing. The cats were neutered to reduce aggressive behavior, making it possible to introduce new animals to the colony. Food was restricted on days that animals were performing the behavioral task (five days a week). On those days, cats received moist food as behavioral reinforcement during training or testing sessions and then were given free access to dry food for up to an hour after the session. On weekends, cats were given free access to dry food for 3 hours per day. Water was freely available in the housing area.

Experimental apparatus. Experiments were conducted in a double-walled sound-attenuating anechoic chamber (Industrial Acoustics; inside dimensions 2.6 x 2.6 x 2.5 m) lined with SONEXone absorbent foam to suppress sound reflections. The chamber contained 13 8.4-cm-diameter two-way loudspeakers positioned on a horizontal circular hoop, 1.2 m in radius, at azimuths of 0 and ±5, 10, 20, 40, 60, 80° relative to the front of the apparatus. The cat was

supported on a raised platform, which was adjusted in height so that the cat's head was centered in the array of loud speakers. A harness restrained the animal to the platform but permitted free movement of the limbs and head. A feeder was mounted on a pneumatic cylinder located on the animal pedestal. The feeder was raised to provide behavioral reinforcement and was lowered during sound presentation. All behavioral sessions were conducted in the dark and were monitored with video using infrared illumination.

Stimulus generation. Stimulus generation and data acquisition used System III hardware from Tucker-Davis Technologies (TDT; Alachua, FL) controlled by custom MATLAB software (The Mathworks; Natick, MA) on a Windows-based computer. Sounds were generated with a 24-bit precision at a sample rate of 97,656 s^{-1}. Loudspeakers were calibrated using a precision ½" microphone (ACO Pacific) that was positioned at the center of the apparatus at the normal location of the animal's head. Golay codes were used as probe sounds (Zhou et al., 1992). The calibration procedure yielded a 1029-tap finite-impulse-response correction filter for each speaker. The filters flattened and equalized the broadband frequency responses of the loudspeakers such that, for each loudspeaker, the standard deviation of the magnitude spectrum across the 0.2–25 kHz calibrated passband was <1 dB. The responses were rolled off by 10 dB from 25 kHz to 40 kHz.

Stimuli consisted of sequences of noise bursts generated in real time by gating a continuous Gaussian noise source generated by the TDT RZ6 digital signal processor. Each noise burst was 20 ms in duration, gated with raised cosine functions with 1-ms rise and fall times. The noise presented from each speaker was filtered with the corresponding speaker correction filter and then was band-pass filtered with fourth-order Butterworth filters to one of the following bands: 0.4-25 kHz for the broadband condition; 0.4-1.6 kHz for the low-band

condition; and 4-25 kHz for the high-band condition. The filter bands were identical between target and masker sounds and throughout each training or testing session. Sounds were presented at 60 dB SPL for all filter conditions. Target sound sequences were presented with two temporal patterns, referred to as Rhythms 1 and 2 (Fig. 2.1A). Each trial began with one to four continuous sequences of Rhythm 1, with the number of sequences varying randomly from trial to trial. The Rhythm-1 sequence was followed without interruption by a 1200-ms Rhythm-2 sequence repeated 1.5 times. Masker sound sequences were interleaved with target sequences (Fig. 2.1B). The masker sequences were exactly complementary to the target sequences such that, when target and masker sources were co-located at 0°, the stimulus was a continuous sequence of undifferentiated noise bursts. The aggregate rates of target and masker noises bursts were 10 s^{-1}, meaning that onsets of target or masker bursts were presented at intervals of 100 ms.

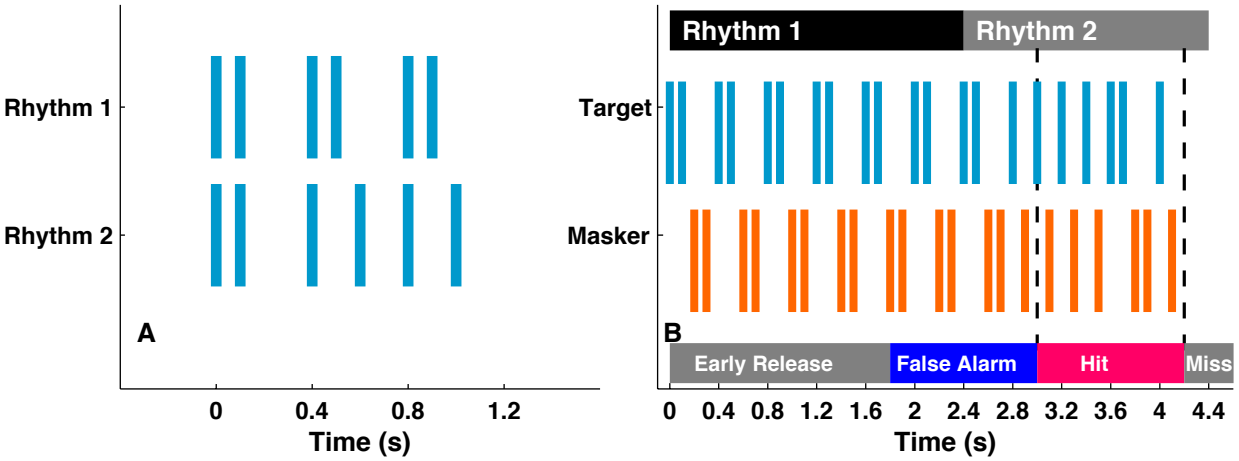

Fig. 2.1. A Two temporal patterns of noise bursts, Rhythm 1 and Rhythm 2. **B** Timing of response windows. This example shows the Hold 2 condition. The target and masker sequences were interleaved in time. The masker pattern was complementary to that of the target. The earliest detectable change from Rhythm 1 to Rhythm 2 occurred 600 ms after the onset of Rhythm 2. That time marked the beginning of a 1200 ms time window in which the cat could release the pedal to score a "hit" and receive a food reward. Releases within 1200 ms prior to the beginning of the hit window were scored as "false alarms". Releases even earlier were scored as "early releases". Releases after the hit window were scored as "misses".

Behavioral task and training. The behavioral task was patterned after the hold-release paradigm described by May and colleagues (1995). Each trial was initiated by an operator who monitored the activity of the cat on a video display. Each trial began with illumination of a green light-emitting-diode (LED) located at 0° azimuth. The green LED signaled the cat to depress a pedal to initiate a sequence of noise bursts, Rhythm 1, from a target source located at 0° azimuth. The target sequence was interleaved with a complementary sequence from a masker source that varied in location from trial to trial. After a variable hold time, the target rhythm changed from Rhythm 1 to Rhythm 2 and the cat was required to release the pedal to receive a food reward. The duration of the sequences of Rhythm 1 varied randomly from trial to trial with equal probability among 1200, 2400, 3600 ms, or 4800 ms in what will be referred to as Hold 1, Hold 2, Hold 3, and Hold 4 trials, respectively. The Hold 4 trials were used only as catch trials for the Hold 3 condition.

Performance on each trial was scored according to the latency of pedal release relative to the time of the first sound burst that differed from Rhythm 1. That pattern change occurred 600 ms after the onset of Rhythm 2 and is indicated in Figure 2B by a vertical dashed line. Starting from that 600 ms time point, cats had a window of 1200 ms in which pedal releases were scored as “hits.” Releases more than 1200 ms prior to the beginning of the hit window were scored as "early release" and were not analyzed further. Releases 1200 to 0 ms prior to the beginning of the hit window were scored as "false alarms.” Releases after the 1200 ms hit window were scored as "misses.” The sounds ceased after the hit window, meaning that later times of pedal releases were of little interest and, therefore were not recorded aside from counting them as misses. Hits were rewarded with delivery of a portion of pureed canned cat food (Purina Friskies). Early releases, false alarms, or misses triggered 4-s time-out periods, signaled by a flashing blue LED,

in which there was no reward and in which no new trial could be initiated. Hold 2, Hold 3, and Hold 4 trials served as catch trials for Hold 1, Hold 2, and Hold 3 trials, respectively. For instance, a hit on a Hold 2 trial was also counted as a correct rejection (i.e., not a false alarm) for a Hold-1 trial. Pedal releases during the hit window during Hold 4 trials were rewarded but were not otherwise scored. The rationale for using Hold *N+1* trials as catch trials for Hold *N* trials, rather than running separate catch trials, was that we obtained a roughly equal number of catch and non-catch trials for each hold time and we potentially could collect both a false-alarm datum and a hit or miss datum on each trial (except for the Hold 4 condition). Time windows in which responses were scored as hits and false alarms were equal in duration, meaning that random pedal releases had roughly equal probability of being scored as hits or false alarms.

Table 2.1 shows the number of trials that were scored for the broadband condition, i.e., excluding early releases, for each Hold time and each cat. The numbers of trials did not vary significantly across Holds 1, 2, and 3 ($\chi^2_{(2)}$=0.087; p=0.96; Friedman Test); the numbers of scored Hold 4 trials were somewhat lower, reflecting a higher probability of early releases in the lengthy time prior to the false-alarm time window in that condition.

Cat	Number of Scored Trials			
	Hold 1	Hold 2	Hold 3	Hold 4
Mu	401	470	426	263
Bo	377	356	358	320
Go	322	268	345	238
Oz	256	340	210	57
Ma	179	378	247	152
St	306	301	301	148

Table 2.1 The number of trials for each Hold number that were included in the data analysis after exclusion of Early Releases. Each column represents a Hold number and each row represents a particular cat.

At the beginning of training for this task, the target sounds were presented without a masker, and Rhythm 2 was presented at a level 10 dB greater than that of Rhythm 1. In that phase, cats were rewarded for detecting the increase in sound level and/or the change in rhythm. The level difference between Rhythm 1 and 2 was decreased as the cat became more proficient at the task. When both rhythms were at the same level, the change from Rhythm 1 to 2 could be detected only on the basis of the temporal pattern. When a cat could detect the rhythm change reliably, a masker was introduced at the ±80° locations. As proficiency improved, the target and masker separations were gradually decreased. Once the animal was proficient with all hold times and varying masker locations, the hold times and masker locations were randomized from trial to trial. After training in the broadband-sound condition, training shifted to the high-band and then low-band conditions (3 cats) or low-band and then high-band conditions (3 cats). Once cats were proficient in all three pass-band conditions, the passband filter conditions were varied every 3-4 days. The reported performance includes a minimum of 10 testing blocks for each passband condition, where each block represents one day of training. Data from enough testing blocks were included for each passband condition for each cat to yield data from ≥20 trials for each target/masker separation.

Each training session lasted for as long as the animal was willing to work, typically around 30 minutes each day. The training periods varied from cat to cat, lasting several months to a year, followed by 3 to 11 months of data collection.

The psychophysical procedure used in the present study, hold-release, differed from the 2-alternative forced-choice procedure used in our previous study of human listeners (Middlebrooks and Onsan, 2012). Also, there were slight differences in the rhythms that were used. The rationales for those differences are considered in the Discussion.

Data analysis. Performance was measured by computing the discrimination index, *d'* (Green and Swets, 1988) for each masker location:

$$d' = z(P_{hit}) - z(P_{false\ alarm})$$

For each masker location, the proportion of hits (P_{hit}) was given by the number of hits divided by the number of hits and misses across Hold 1, 2, and 3 trials, and the proportion of false alarms ($P_{false\ alarm}$) was given by the number of false alarms divided by the number of false alarms, hits, and misses across Hold 2, 3, and 4 trials. P_{hit} and $P_{false\ alarm}$ were transformed to standard deviants (z-scores), and the difference in z-scores gave the discrimination index, *d'*. In some conditions, P_{hit} was 1 or $P_{false\ alarm}$ was 0, meaning that the *z*-score was undefined. In those situations, the proportion of hits or false alarms on *N* trials was expressed as $(N - ½)/N$ or $(½)/N$, respectively (Macmillan and Kaplan, 1985). Values of *d'* for each cat and passband were plotted as a function of masker location. The masker location at which the interpolated plot crossed a criterion of $d' = 1$ or, in a separate computation, $d' = 2$ was used as the rhythmic masking release (RMR) threshold.

The distributions of thresholds were not normally distributed. For that reason, non-parametric statistics were used for comparison of median thresholds between conditions.

2.4 Results

We begin by characterizing observations that were specific to the cats' performance of the hold-release behavioral task. Then, we compare performance among broadband, low-band, and high-band stimulus conditions that were intended to identify the acoustic cues that provide highest spatial acuity for cats.

Task Performance. Cats performed the hold-release task enthusiastically, showing high hit rates for large target/masker separations, declining to chance performance for narrow separations. The positions of cats' heads and pinnae were monitored on the video display. Cats learned early in training to direct their attention toward the green start light and the target sound source, both located at 0° azimuth. In the training trials in which maskers were first introduced, some cats made orienting movements of the head and pinnae toward masker sources at peripheral locations, but that behavior rapidly extinguished. During data collection, cats tended to keep their heads oriented toward the target source at 0° and their mobile pinnae in a fully forward position, seemingly focused on the target.

The histograms in Figure 2.2 show the distributions of latencies to pedal release relative to onsets of the sound sequences; these data are from the broadband condition. Cats Mu and Bo are represented by the left and right columns of panels, respectively. The rows of panels represent the four hold times, i.e., the four durations of Rhythm-1 sequences prior to the change to Rhythm 2. Hold times, and the corresponding hit windows, were varied randomly from trial to trial in order to confound efforts to obtain reinforcement by releasing the pedal at some constant latency. In Figure 2.2, the bars are colored to represent responses that were scored as early releases (green), false alarms (blue), hits (magenta), or misses (white). Misses are grouped in single bars after the hit windows, regardless of how long the cat held the pedal after the offset of sound presentation. The histograms include results from all target/masker separations, including 0°, at which stream segregation was impossible, and ±5°, which proved to be narrower than the thresholds of any of the cats. The trials with those sub-threshold target/masker separations tended to increase the scatter of response latencies among early release, false alarm, and miss windows.

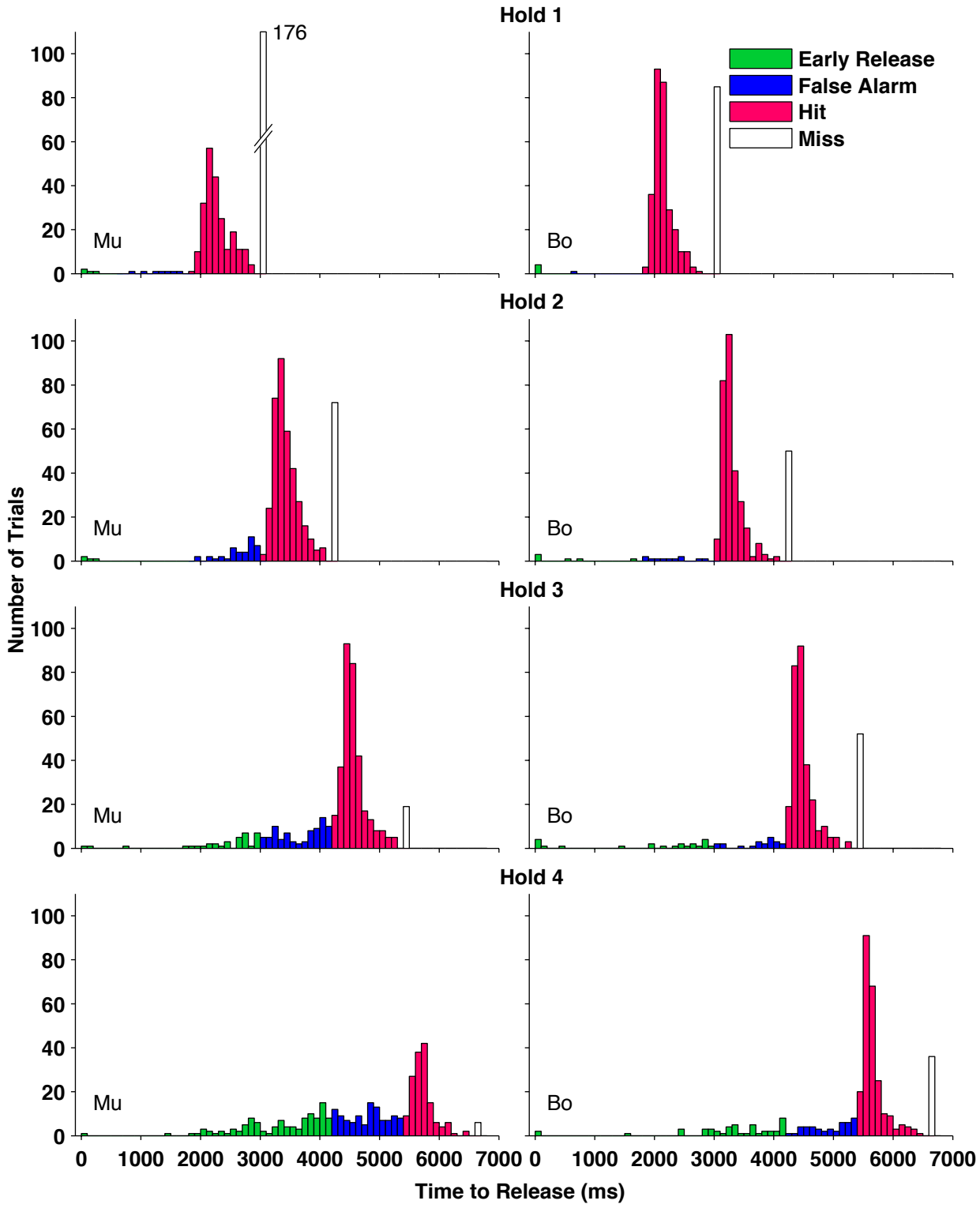

Figure 2.2 Latencies to pedal release for two cats in the broadband condition. Each row of panels represents a different Hold number (i.e., number of repetitions of Rhythm 1). Colors on the bar graph denote scores of early release (green), false alarm (blue), hit (magenta), and miss (white). Columns of panels represent individual cats (Mu and Bo). The 0 ms Pedal Release Time denotes the start of the sound presentation.

In each panel, the numbers of pedal releases were relatively low during presentation of Rhythm 1 (i.e., the hold time), and the numbers of responses increased sharply as the stimulus pattern changed to Rhythm 2, signaling the correct release time. Generally, the hit responses occurred with short latencies relative to the rhythm change, with 81% of hits falling within the first half of the hit window across all cats, target/masker separations, and hold times. The observation that hit responses tended to fall early in the hit window indicates that the cats tended to respond as soon as they detected the increased inter-burst interval that characterized Rhythm 2; that is, the rhythms could be discriminated without listening to the entire rhythm. The numbers of early releases and false alarms increased with increasing hold time. Those increases were seen across all animals tested ($\chi^2_{(2)}$=10.3; p=0.0057; Friedman Test). We attribute the increase in early releases and false alarms with increasing hold times as indicative of the cats' general impatience in waiting for reinforcement.

Latencies to pedal release for hit responses as a function of masker location are shown in Figure 2.3, again in the broadband condition for cat Mu (left) and cat Bo (right). Symbols indicate the pedal release latency on each trial. Data are collapsed across Hold times 1, 2, and 3, and latencies are aligned relative to the time of the rhythm change (i.e., relative to the beginning of the hit window). Numbers of misses, hits, and false alarms are given by the rows of numbers at the top of the figure. At wide target/masker separations (e.g., Masker Locations ±80 and 60°), there were few false alarms, and there were many hits, typically early in the hit window. At narrower separation, the numbers of false alarms increased, the numbers of hits decreased, and pedal releases were later in the hit window. At near-zero separations, pedal releases were scattered fairly randomly throughout the false alarm and hit windows.

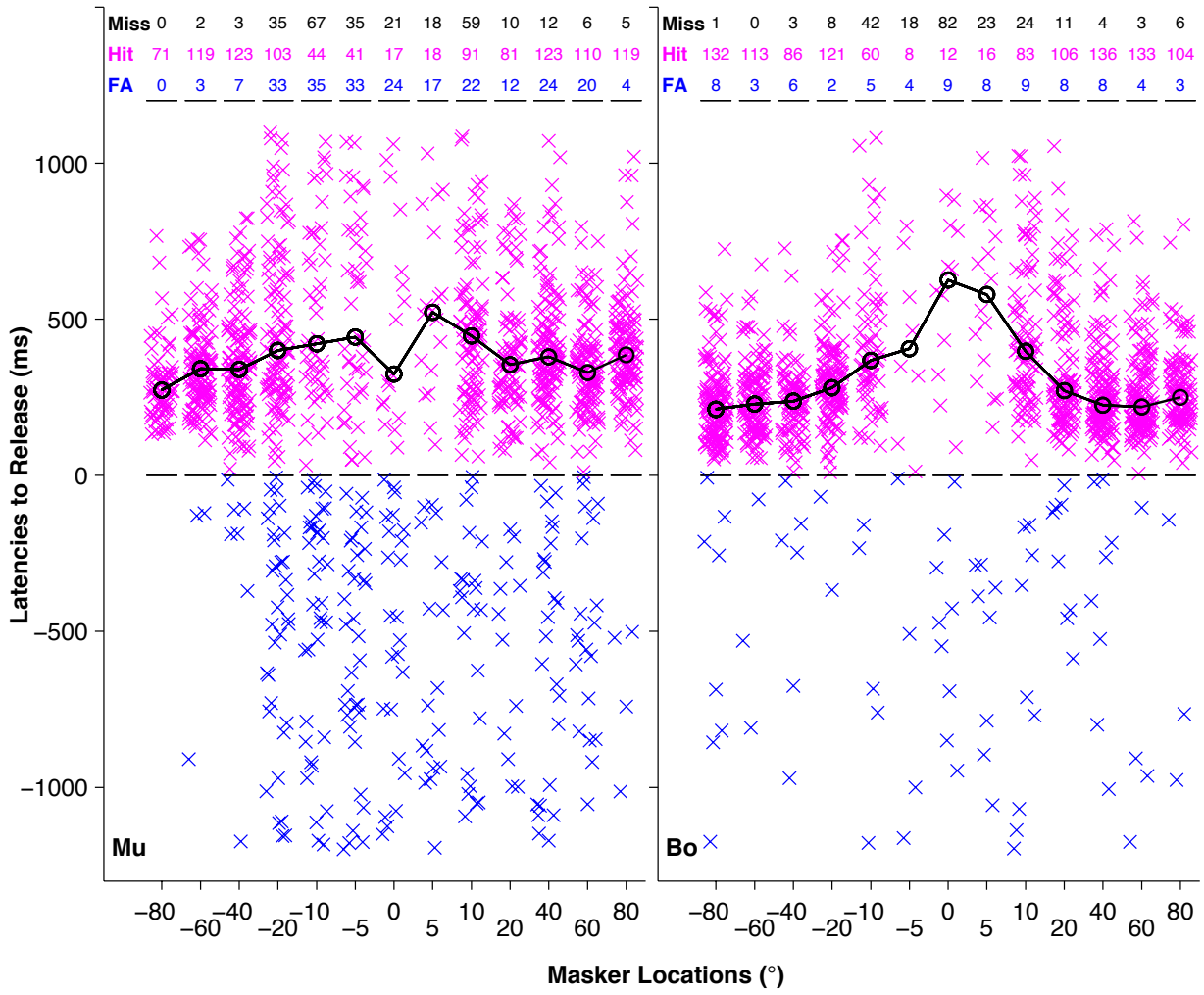

Figure 2.3 Latencies to release as a function of masker location for the same two animals shown in Figure 2 (Mu and Bo). Individual x symbols represent trials that were scored as hits (magenta) or false alarms (blue). The target was always located at 0 degrees. Latency time point of 0 ms represents the start of the hit window. The time point of 1200 ms represents the end of the hit window. Trials below the 0 ms time point are false alarms. Numbers at the top of the figures denote the number of responses for misses, hits, and false alarms (FA) at that respective masker location. The black curves indicate the median latencies of hit responses.

Cats exhibited a range of biases for or against releases of the response pedal. Cat Mu, whose data are shown on the left sides of Figures 2.2 and 2.3, was relatively eager to release the pedal. Compared to Cat Bo (on the right), Cat Mu had a higher false alarm rate at all but the widest target/masker separations. For the widest target/masker separations, shown in Figure 2.3, Cat Mu correctly rejected early pedal releases and released in the hit window. For narrow

separations, however, he apparently was less able to segregate the target and masker sequences and, therefore, was less able to recognize Rhythm 1 during the Hold time. His tendency on such trials was to release early. The sum of his false alarms and hits consistently was higher than the number of his misses, and median latency of his hits in the 0°-masker condition was relatively short. In contrast, Cat Bo shown on the right sides of Figures 2.2 and 2.3 was more conservative. His false alarm rates were low for all target/masker separations. His tendency on the difficult trials with separations ≤5° was to persist in holding the pedal, as indicated by low numbers of false alarms and hits, by large numbers of misses, and by the relatively long median latency for hits in the 0°-masker condition. The *d'* analysis that was used for evaluating performance largely compensated for differences in response bias among cats. That is, values of *d'* in cases of bias toward pedal releases (like Cat Mu), which produced high numbers of false alarms but also high numbers of hits, could be roughly equal to *d'* values in cases of bias against release, which produced lower numbers both of false alarms and hits. The *d'* measures are presented in the next section.

Broadband spatial stream segregation. Figure 2.4 shows the performance for all six cats, where each row of panels represents percentages of hits and false alarm rates and the *d'* for one animal. We first consider data from the broadband condition, indicated by open squares and solid black lines. Hit rates tended to be low at narrow target-masker separations (i.e., 0° and 5° masker locations), ≤50% for most cats. Hit rates increased markedly with increasing target/masker separation, reaching 100% for most cats. The dependence of false alarms on target/masker separation varied somewhat among cats. For the majority of cats (e.g., Cat Mu, top row), false-alarm rates were noticeably higher for narrow separations. For other cats (e.g., Cat Bo), false-alarm rates were largely insensitive to separations. The *d'* values (right column) were

around 0 for near-zero target/masker separations and increased with increasing separations. In the broadband condition, most of the cats reached *d'* around 4, nearly 100% correct, for the widest separations, although performance was not as good for Cats Oz and Ma. The two cats showing differing bias contrasted in the previous section – Mu the eager releaser and Bo the conservative – are represented in the top two rows of Figure 2.4. One can see that their *d'* values in the broadband condition were fairly similar even though Cat Mu's hit and false-alarm rates both were noticeably higher than those of Cat Bo.

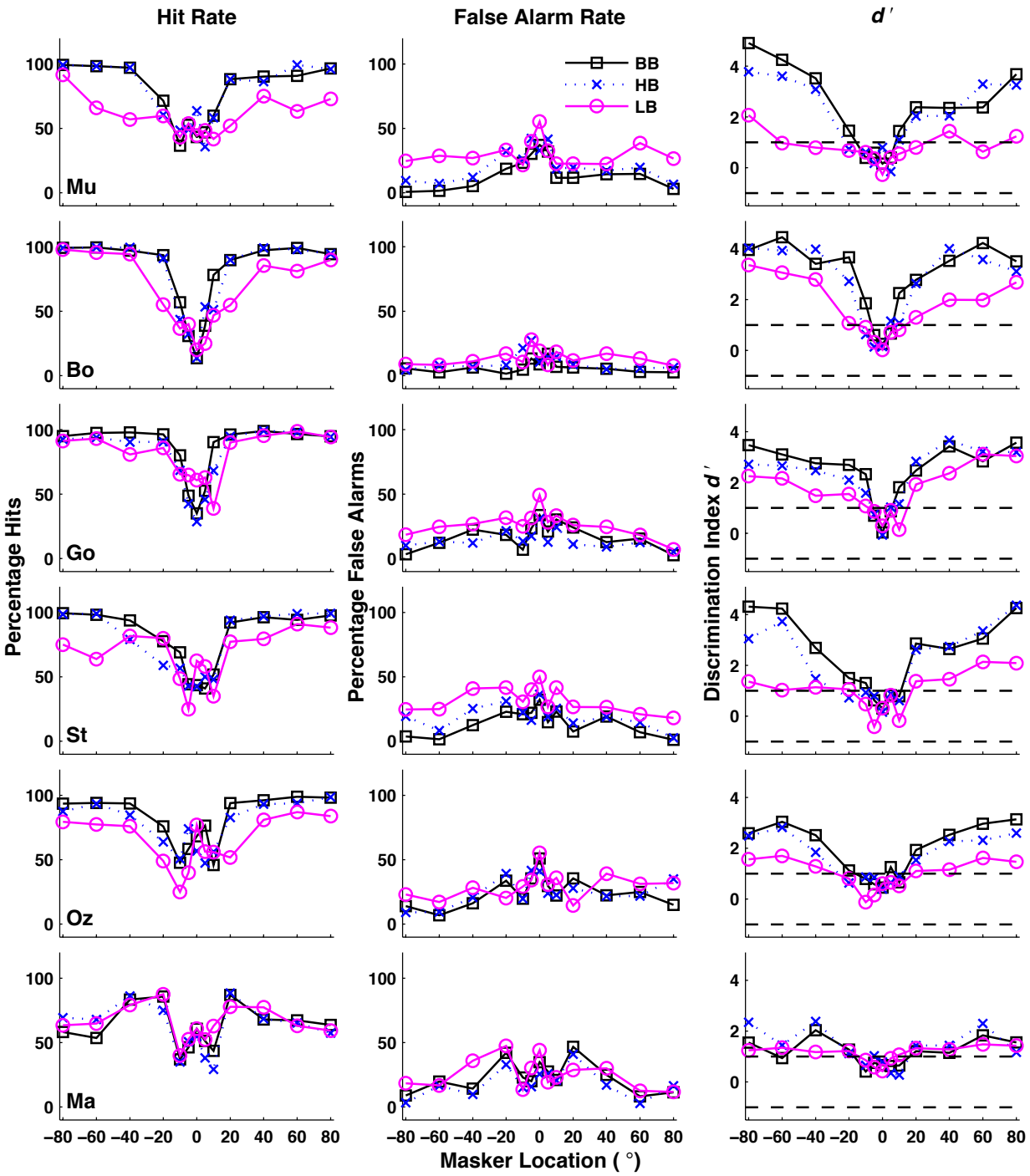

Figure 2.4 Task performance as a function of masker location. The three columns of panels show the proportion of hits, of false alarms, and the discrimination index for each filter condition. Filter conditions are denoted by black solid lines and open squares for broadband (BB), blue dashed lines and x symbols for high-band (HB), and magenta solid lines and open circles for low-band (LB). Each row of panels represents an individual cat.

Rhythmic masking release (RMR) thresholds were given by the narrowest target/masker separation at which *d'* was consistently ≥1 and, in a separate computation, *d'* ≥2. The criterion of ≥1 (Fig. 2.5A, blue) was used to permit comparison with our previous study in humans (Middlebrooks and Onsan, 2012), and the criterion of ≥2 (Fig. 2.5A, magenta) was used to better evaluate the difference in performance in the various passband conditions (presented below). Two RMR thresholds were recorded for each passband condition and *d'* criterion, for maskers to the left and the right of the target. The distributions of RMR thresholds in the broadband condition for all six cats is given in the left-most pair of columns in Figure 2.5B as box plots and with individual symbols. The median broadband RMR threshold was 9.4° for the $d' \geq 1$ criterion and 19.1° for the $d' \geq 2$ criterion.

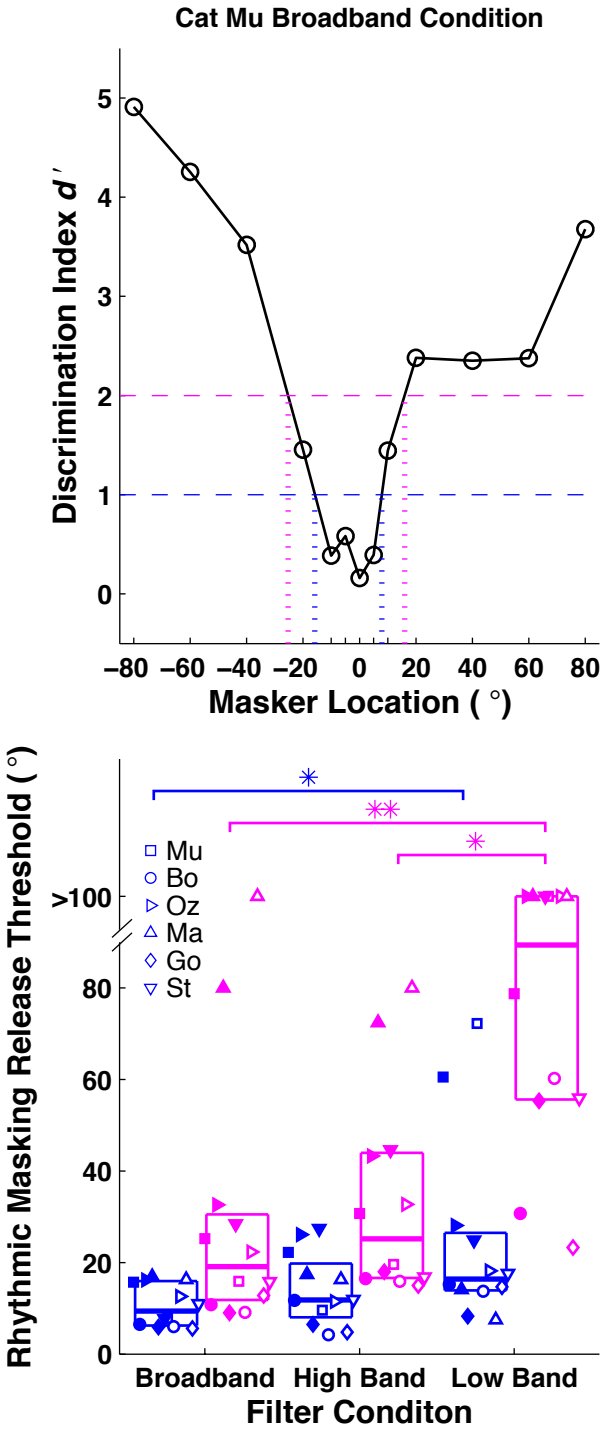

Figure 2.5 A A Estimation of rhythmic masking release (RMR) thresholds. RMR thresholds were given by the masker locations to the left and right of the target at which the interpolated plot crossed a criterion of d' =1 (blue dashed line) and, in a separate computation d'= 2 (magenta dashed line). **B** Distribution of RMR thresholds for each pass-band condition. Data shown in blue reflect the $d' \geq 1$ criterion, and data shown in magenta reflect the $d' \geq 2$ criterion. Each symbol represents a different cat. Filled and open symbols represent RMR thresholds located to the left and right of the target, respectively. A random horizontal offset is added to each symbol to minimize overlap between data points. The horizontal lines of each box represent the 25th, 50th, and 75th percentiles.

Performance in conditions of restricted spatial cues. We tested conditions of limited frequency bandwidth as a means of identifying the major acoustical cues that cats use for spatial stream segregation. The low-band condition used a spectrum limited to 0.4 to 1.6 kHz. We assume that essentially the only useful spatial cue in that frequency band is the ITD in temporal fine structure. The high-band condition used a spectrum of 4.0 to 25 kHz. In that band, the most likely cues involve differences in sound level, both in the form of interaural differences in levels (ILDs) and as differences in target and masker levels at each ear. There might also be some influence of ITDs in high-frequency sound envelopes, although Middlebrooks and Onsan (2012) demonstrated only a weak contribution of envelope ITD to spatial stream segregation by humans.

Spatial stream segregation in the high-band condition was nearly as high as that in the broadband condition. That can be seen for all cats in the *d'* plots on the right column of panels in Figure 4. The blue dotted lines indicating the high-band condition nearly overlie the black solid lines of the broadband condition. In contrast, performance was consistently degraded in the low-pass condition, in which low-frequency ITDs presumably are the spatial cue. In Figure 2.4, the magenta lines indicating the low-band condition generally show lower hit rates, higher false-alarm rates, and lower *d'*.

Distributions of RMR thresholds for the various passbands are shown in Figure 2.5B. Given the criterion of $d' \geq 1$, median threshold values were 9.4° for broadband, 11.8° for high-band, and 16.4° for low-band. Median values varied significantly with passband (Friedman test, $\chi^2_{(2)}=7.8$, $p=0.020$). A *post hoc* analysis with Bonferonni adjustment showed that low-pass thresholds were significantly wider than broadband thresholds ($p<0.05$) but that there was no significant difference between broadband and high-band or between high-band and low-band thresholds ($p>0.05$). The dependence of performance on stimulus passband was greater given a

criterion of $d' \geq 2$. Median threshold values were 19.1° for broadband, 25.2° for high-band, and 89.4° for low-band. Median values varied significantly with passband (Friedman test, $\chi^2_{(2)}=19.3$, $p<0.0001$). The *post hoc* analysis with Bonferonni adjustment showed that low-band thresholds were significantly wider than broadband thresholds ($p<0.01$) and wider than high-band thresholds ($p<0.05$) but, again, that broadband and high-band thresholds were not significantly different ($p>0.05$).

We also compared across passbands the distributions of d' for target/masker separations of 40° (Fig. 2.6). The 40° separation was chosen because that separation tended to produce d' higher than the threshold value of $d'=1$ and lower than asymptotic values for nearly all cats and conditions. Median values of d' at the 40° separation were 2.66 for broadband, 2.42 for high-band, and 1.36 for low-band. The d' values varied significantly with passband (Friedman test, $\chi^2_{(2)}=17.3$, $p=1.7 \times 10^{-4}$). A *post hoc* analysis with Bonferroni correction showed that d' values were significantly higher (i.e., performance was better) for broadband and high-band than for low-band conditions ($p<0.05$) but that there was no significant difference between broadband and high-band conditions ($p>0.05$).

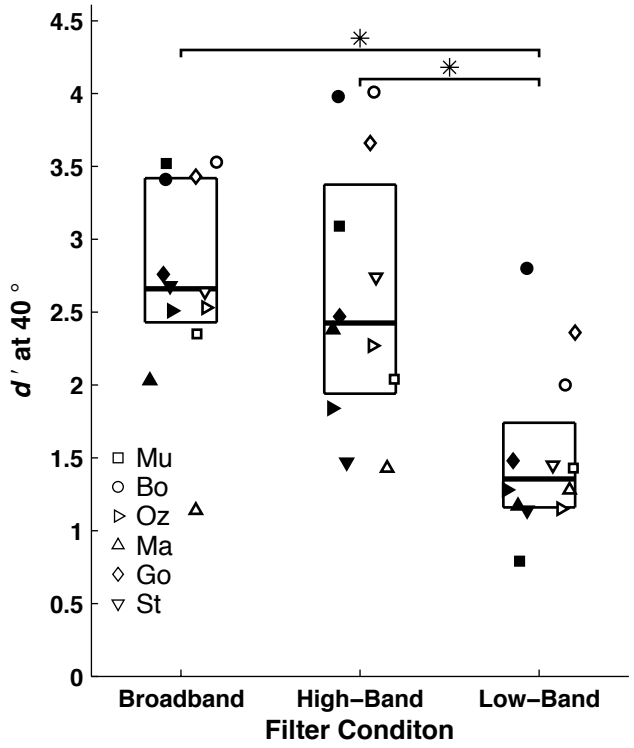

Figure 2.6 Distribution of d' values for 40° target/masker separations for each pass-band condition. Each symbol represents a different cat. Filled and open symbols represent *d'* for maskers located to the left and right of the target, respectively. A random horizontal offset is added to each symbol to minimize overlap between data points. The horizontal lines of each box represent the 25th, 50th, and 75th percentiles

Overall, the results showed little or no impairment of spatial stream segregation by cats when low-frequency ITDs were made unavailable and a severe degradation in performance when high-frequency cues were eliminated. These results make an interesting contrast to the situation in humans, in which performance is substantially better with low-band than with high-band sounds (Middlebrooks and Onsan, 2012).

2.5 Discussion

Cats performed this RMR task reliably, exhibiting spatial stream segregation comparable to that of humans. The median RMR threshold in the broadband condition for 6 cat listeners was 9.4°, only slightly broader than the corresponding median of 8.1° for 7 human listeners

(Middlebrooks and Onsan, 2012). Cats differed from humans in that performance by the cats was better in the high-band than the low-band condition, whereas the opposite was true for humans. Possible reasons for that difference are considered in a later section.

The hold-release task used in the present study differed from the task used in the previous human psychophysical study (Middlebrooks and Onsan, 2012). The human study used a one-interval, 2-alternative design: Rhythm 1 or Rhythm 2 was presented on each trial with equal probability, and the listener responded by pressing one of two keys. Our initial efforts to train cats on the 2-alternative task were unsuccessful, largely because the cats tended to associate one or the other response pedal with the location of the masker source rather than with the stimulus rhythm. The cat and human studies also differed somewhat in that the rhythms in the present cat study were extended from 800 to 1200 ms by adding two bursts to target and masker patterns. The reason for that change was to provide the cats with a longer time window in which to release the pedal during the Rhythm 2 presentation. Informal comparisons by human listeners produced essentially equal RMR thresholds between the two-alternative and hold–release tasks. Finally, the upper limit of the broadband and high-band conditions was extended to 25 kHz to take advantage of the cat's higher audibility range.

Species differences in use of spatial cues. In the present study and the previous human psychophysical study (Middlebrooks and Onsan, 2012), stimulus bandwidths were manipulated to limit the available spatial cues. The low-band condition was intended to minimize use of cues related to sound levels, and the high-band condition was intended to eliminate usable cues from ITDs in temporal fine structure. Cats consistently performed worse in the low-band than in the high-band and broadband conditions in that hit rates were lower, false-alarm rates were higher, *d'* values were lower, and RMR thresholds were broader in the low-band condition. High-band

performance, in contrast, was not significantly different from that in the broadband condition. We take this to mean that in the broadband condition cats relied primarily on high-frequency ILD cues. This contrasts with the previously reported human results in which RMR thresholds in low-band and broadband conditions were not significantly different and thresholds in the high-band condition were significantly broader. Those results suggest that in the broadband condition humans relied primarily on ITD cues.

The superior performance by cats in the high-band condition agrees with expectations based on single-unit recordings from cortical area A1 in anesthetized cats (Middlebrooks and Bremen, 2013). In that study, cortical neurons demonstrated a correlate of spatial stream segregation by synchronizing preferentially to one of two interleaved sequences of broadband noise bursts from sound sources that were separated in location. Neurons that were most sensitive to frequencies greater than 4 kHz tended to show higher *d'* for segregation of sound sequences from alternating sources than did neurons that were most sensitive to lower frequencies. A test of the ability of a linear classifier to discriminate stimulus rhythms based on neural spike patterns also showed good performance among the highest-frequency neurons although, inexplicably, that test also showed good performance among the small sample of units that were most sensitive to frequencies around 500 Hz.

The poorer performance by cats in the low-band condition conflicts with early studies of localization of pure tone stimuli. Casseday and Neff (1973) trained cats to walk to one of two possible pure-tone sources located symmetrically about the frontal midline, and Martin and Webster (1987) used a conditioned-avoidance task in which cats were required to detect a change in the location of a tone source away from the frontal midline. In both of those studies, performance was best for tone frequencies ≤2 kHz, poor at 4 kHz, and improving (Casseday and

Neff, 1973) or irregular (Martin and Webster, 1987) at even higher frequencies. The reason for the difference in frequency dependence between previous and present studies is not obvious, but we note that the stimulus conditions were very different. In the early localization studies, tone bursts were 500 ms in duration and were repeated 5 or more times for each location judgement. Those lengthy sound presentations would have permitted a cat to move its head and ears relative to the sound source during individual sound bursts and thereby obtain dynamic localization cues. In the present study, in contrast, individual sound bursts were only 20 ms in duration. The sequences of such bursts lasted for some seconds, but the task required the cats to segregate successive 20-ms bursts from the two sources in order to evaluate the rhythm conveyed by the sequence from one or the other source. We also note that, in humans, discrimination of the locations of two successive sounds (i.e., a minimum audible angle test) was not a good predictor of the effects of passband on RMR thresholds (Middlebrooks and Onsan, 2012).

Cats in the present study made less effective use of low-frequency spatial cues than do humans (Middlebrooks and Onsan, 2012). That inter-species difference likely can be attributed primarily to differences in the sizes of cat and human heads, resulting in differences in interaural delays. In both species, ITDs vary somewhat with frequency across the 0.4-to-1.6 kHz range of our low-band stimulus. In cats, ITDs at 0.8 kHz are around 100 and 320 μs for sound sources at 15 and 90°, respectively (Roth et al., 1980). In humans, ITDs are ~1.5-to-2 times greater for the same source angles: 140 and 660 μs, respectively, on a human-sized mannequin (Kuhn, 1977). Despite the differences in the ranges of ITDs that cats and humans typically experience, their sensitivity to ITD is similar. Reported just-noticeable differences in ITDs are around 25 μs in cats (Wakeford and Robinson, 1974; Cranford, 1979) and between 9 and 45 μs in human depending on listeners' degree of training (Zwislocki and Feldman, 1956; Klumpp and Eady,

1956; Wright and Fitzgerald, 2001; Middlebrooks et al., 2013). Also, the just-noticeable difference for ITD increases dramatically or becomes immeasurable at tone frequencies greater than ~1.5 kHz in both cats (Wakeford and Robinson, 1974) and humans (Zwislocki and Feldman, 1956; Klumpp and Eady, 1956; Brughera et al., 2013). The cat's smaller head means that, given comparable ITD sensitivity in cats and humans, the displacement of a sound source from the midline needed to achieve a just-noticeable ITD is 1.5-to-2 times larger for a cat than for a human. Scaling of ITDs by a factor of 1.5 to 2 would reduce to some degree the difference in median low-band RMR thresholds between cat (16.4°) and human (5.9°). One can see in Figure 4, however, that a simple scaling of ITD would not bring the psychometric functions for the low-band condition in line with those for broadband and high-band conditions. That is, cats' maximum *d'* levels of performance in the low-band condition rarely reached the high levels of performance attained in the broadband or high-band condition. We conclude that the cat's smaller head size relative to humans almost certainly contributes to the cat's less effective use of low-frequency cues, but that head size cannot entirely account for the inter-species difference.

We considered two other factors that might explain to some degree the cats' relatively poor performance in the low-band condition. One consideration is that detection thresholds by cats are reported to be as much as ~15 dB higher for sounds in the low-frequency compared to the high-frequency bands that were tested (Neff and Hind, 1955; Heffner and Heffner, 1985). We note, however, that our low-band stimuli were ~50 dB above the reported audiograms for cats and, therefore, should have been clearly audible. Also, in pilot studies, we observed that 5-dB increases in the levels of low-band sounds failed to improve performance. The second consideration is simply the observation that cats were less willing to perform the task in the low-band condition. That might be because, for one reason or another, their performance was worse

in that condition so that they received less frequent reinforcement. Alternatively, it might have been that the low-band stimulus was, for some reason, aversive to the cats.

Cats in the present study performed better in the high-band condition (median threshold 11.8°) than they did in the low-band condition (16.4°) and better than humans in the high-band condition (median threshold 15.4°; Middlebrooks and Onsan, 2012). We assume that the principal cues for spatial stream segregation in the high-band condition are related to sound level rather than ITD (Middlebrooks and Green, 1991; Macpherson and Middlebrooks, 2002; Middlebrooks and Onsan, 2012). Other things being equal, we would expect that, at any particular frequency, the cat's smaller head would produce weaker refraction and weaker level-related cues than would the larger human head. Two factors mitigate the possible disadvantage of a smaller head. First, the cat's audible range extends more than an octave higher than that of humans. In the human study, the high-band stimulus cut off at 16 kHz, which is well above the most-sensitive frequency region of listeners. In the cat study, however, the high-band stimulus extended to 25 kHz, which is within the sensitive portion of the cat's behavioral audiogram (Heffner and Heffner, 1985). The higher-frequency hearing by the cat would permit it to benefit from the decrease in wavelengths at higher frequencies, which would result in stronger level cues (Middlebrooks and Pettigrew, 1981; Phillips et al., 1982; Tollin and Koka, 1990a). Second, the directional sensitivity of the cat's external ears could have enhanced spatial level cues around the frontal midline. Previous acoustical measurements have shown that when the cat's ears are oriented frontally, as they were during task performance, the axes of greatest sensitivity are located ~10-40° from the frontal midline, meaning that sensitivity tends to decline fairly steeply across the midline (Middlebrooks and Pettigrew, 1981; Calford and Pettigrew, 1984; Middlebrooks and Knudson, 1987; Musicant et al., 1990; Young et al., 1996; Tollin and Koka,

2009a). The resulting interaural level differences show a particularly steep gradient across the midline in cat at frequencies above ~8 kHz (Middlebrooks and Pettigrew, 1981; Musicant et al., 1990; Tollin and Koka, 2009b). In contrast, the human ear at the highest audible frequencies is focused near the frontal midline, meaning that the spatial gradient of levels in the central ~±15° is relatively flat (Middlebrooks et al., 1989).

There are three mechanisms by which the spatial dependence of sound levels at the ears could support spatial stream segregation. The *first* would be a conventional use of ILDs as spatial cues, resulting in differential representations of locations of target and masker sources. The *second* would be a "better ear" mechanism in which the cat could attend to the ear contralateral to the masker source, thereby optimizing the target-to-masker ratio. The *third* mechanism would be detection of differences in the levels of target and masker sounds at each ear, exploiting the potentially audible rhythms of varying sound levels. Humans can segregate two interleaved sequences of sounds that differ in level by as little as 3 dB, even when the two sources are co-located (Middlebrooks and Onsan, 2012). The human study showed that that target/masker level cues were weaker than the cue given by ILDs. In cats, however, the steeper gradient of high-frequency sound levels around the midline might enhance the contribution of absolute-level cues and, thereby, account for the cat's superior performance in the high-band condition.

2.6 Acknowledgements

The work is supported by National Institute on Deafness and Other Communication Disorders Grant R01 DC000420 (to JCM) and the Graduate Assistance in Areas of National Need Award P200A120165 (to LKJ). We thank Justin Yao, Peter Bremen, Pauline Herbert-Whiting, and Zekiye Onsan for technical and administrative assistance.

CHAPTER 3

Spectral and Temporal Coding Properties in the Auditory Cortex of Awake Cats

3.1 Summary

The sub-cortical auditory pathway exhibits a rich variety of spectral and temporal coding properties, including sharp selectivity for sound frequency and/or level, synchrony of neural firing to relatively high envelope modulation rates, and coding of even higher rates by non-synchronized spike-rate codes. Surprisingly, these coding properties are rarely, if ever seen in the auditory cortex under commonly used anesthetized conditions. We tested the hypothesis that the failure to observe those properties in the auditory cortex is a result of the use of general anesthesia. To test this hypothesis, we recorded from the auditory cortex of unanesthetized cats using chronic recording arrays. These were passive recordings in that the cats were not explicitly engaged in a listening task. Under these conditions, we find spectral and temporal properties not seen in anesthetized cat cortex. These properties include O-shaped frequency response areas that are restricted in frequency and level, sharp frequency tuning bandwidths that are level-invariant, sustained responses to preferred tones, neurons that can phase lock to trains of clicks or noise bursts at rates >40 s^{-1}, and tonic asynchronous firing that increases with increasing stimulus rates. The results accord with previous observations in unanesthetized marmosets suggesting that these properties may be present in other species and are common properties of the normal, unanesthetized, auditory cortex.

3.2 Introduction

In commonly used anesthetized cat preparations, a substantial majority of neurons in auditory cortex exhibit V-shaped frequency response areas (FRAs) that are restricted in bandwidth at low sound level but broaden at higher levels (Heil et al., 1992; Schreiner and Sutter, 1992); some other neurons show multi-peaked FRAs, particularly in dorsal areas of A1 (Middlebrooks and Zook, 1983; Sutter and Schreiner, 1991) or with the use of inhaled anesthetics (Moshitch et al., 2006). Interestingly, in the absence of anesthesia, many neurons in subcortical and cortical areas show FRAs that are restricted in both frequency and sound level. In decerebrate cats for example, neurons in the dorsal cochlear nucleus (DCN) (Young and Brownell, 1976) and inferior colliculus (ICC) (Ramachandran et al., 1999) exhibit FRAs having I- and O-shapes (Type IV neurons observed by Young and Brownell resembled O-shaped neurons). These units differ from V-shaped FRAs in that the frequency tuning of these units does not widen with increasing sound levels and their frequency tuning is narrower (i.e., narrower bandwidth). For O-shaped FRAS, neurons respond maximally at some BL followed by responses declining at higher levels. For I-shaped FRAs, frequency tuning does not change with increasing sound levels. In addition to the DCN and ICC, I- and O-shaped FRA responses have also been recorded in primary auditory cortex of awake marmosets (Sadagopan and Wang, 2008; Bartlett et al., 2011). Moreover, neurons in primary cortex of awake cats and marmosets show sustained responses to preferred tones (cats: Evans and Whitfield, 1964; Chimoto et al., 2002; marmosets: Wang et al., 2005) which is contrary to anesthetized preparations in which neurons respond only to the onsets of sustained stimuli (reviewed by Phillips, 1993).

Throughout the ascending auditory pathway, the ability of neurons to synchronize their responses (i.e., phase lock) to periodic stimuli decreases at successively higher levels (for

review, Langner, 1992). In the ICC for example, neurons can phase lock to amplitude-modulated noise with best modulation frequency for firing rates (BMFs) averaging 30-100 s^{-1} with maximum BMFs as high as 1000 s^{-1} (Langner and Schreiner, 1988). Progressing from the ICC to the thalamus, the maximum rate of phase locking in the ICC is slightly higher than phase locking recorded in the medial geniculate body of the thalamus (MGB). Neurons in the MGB of anesthetized cats are able to synchronize their responses to click rates of ~100 s^{-1} (Rouiller et al., 1981). The ability for neurons to phase lock to periodic stimuli continues to decrease from the thalamus ascending to auditory cortex where neurons in anesthetized cats were shown to phase-lock to clicks or noise bursts at rates ~25 s^{-1} (Schreiner and Urbas, 1988; Lu and Wang, 2000; Eggermont, 1991, Middlebrooks and Bremen, 2013). Under awake conditions however, neurons in marmoset primary auditory cortex demonstrate response synchrony >40 s^{-1} (Lu et al., 2001) which is higher than what is seen under anesthesia. Recordings have also shown a subpopulation of neurons exhibiting nonsynchronous rate coding, in which neurons fire in a sustained manner that increases in spike rate with increasing stimulus rate; Cortical neurons in awake marmosets show nonsynchronous coding up to 330 s^{-1}.

The spectral and temporal response properties observed in unanesthetized marmoset cortex might be unique to the marmoset, representing specialization for reception of the marmoset's sophisticated vocal repertoire. We tested an alternative hypothesis, that failure to observe those properties in the auditory cortex of other species is the result of experimental technique, specifically the use of anesthesia. For this experiment, we recorded from cortical area A1 in awake cats implanted with chronic recording arrays. We recorded under conditions in which the animals were not engaged in an auditory task. Activity was recorded from well isolated single units and from multiples of two or more undifferentiated units. Pure tone stimuli,

varied in frequency and pressure level, were used to measure frequency response areas. We observed spectral and temporal properties not seen in anesthetized cat cortex. These observations include I- and O- shaped FRAs that were restricted in both frequency and sound level, sharp level-invariant frequency tuning bandwidths ≤0.2 octaves, and sustained responses to tone stimuli. In response to trains of clicks or 20-ms noise bursts, that were varied in rate, we observed phase locking to click and pulse rates >40/s, and neural coding of higher stimulus rates by non-synchronized tonic responses that increase in spike rate with increasing stimulus rate.

The results demonstrate that the remarkable spectral and temporal sensitivity observed in awake marmosets is not limited to marmosets. These properties are also present in cats, suggesting that they are common properties of normal, awake, auditory cortex.

3.3 Materials and Methods

Animals. All procedures were in accordance with the NIH Animal Welfare Guidelines and with a protocol approved by the Institutional Animal Care and Use Committee at the University of California at Irvine. Eight domestic shorthaired cats (*Felis catus*), seven males and one female, were obtained from a breeding colony at the University of California at Davis. No hearing deficits were evident. Ages ranged from one to five years at the time of collection of the reported data. Male cats were neutered to reduce aggressive behavior, making it possible to introduce new animals to the colony. Food was restricted on days in which physiological data was collected. During recording sessions, cats received moist food as they sat calmly in the sound booth and then were given free access to dry food for up to an hour after the recordings. On other days, cats were given free access to dry food for three hours a day. Water was freely available in the housing area. These animals had been trained, with varying degrees of success,

in an operant auditory task that required discrimination of temporal patterns of sound bursts. During collection of the present data, however, the operant response lever and food-reward system was inactivated and animals received food without any listening contingency.

Chronic recording arrays. Chronic recording arrays were obtained from Modular Bionics (Santa Ana, CA). The arrays consisted of four parallel shanks, each containing eight recording sites. The shanks were of either equal length at 3 mm or staggered lengths with two shanks at 2.5 mm and two shanks at 3.5 mm. The shanks were arranged on corners of a rectangle, 0.8 μm or 1.0 mm in the medio-lateral dimension and 0.8 μm or 1.5 mm in the rostro-caudal dimension. Along each shank, the recording sites were spaced in 250 μm intervals. The recording sites consisted of the cut ends of 25-μm-diameter platinum-iridium wire, plated with activated iridium. The sites on the array connected with a flexible cable terminating in a percutaneous connector.

Surgical procedures. The chronic recording arrays were implanted under aseptic conditions in an approved surgical suite. We administered dexamethasone (1 mg/kg) to prevent brain edema and buprenorphine (0.1 mg/kg) for analgesia prior to surgery. In addition, we also administered meloxicam (0.2 mg/kg) for inflammation, and ampicillin (10 mg/kg) to prevent bacterial infections. Following surgery, cats were given Clavamox (62.5 mg/kg) as an antibiotic to prevent infection, for 14 days. Anesthesia was induced with an intramuscular injection of ketamine (25 mg/kg) to permit shaving of the cat and insertion of an endotracheal tube. Isoflurane in 100% oxygen was then delivered through the endotracheal tube, with the isoflurane concentration adjusted to maintain an arelexive state. Ophthalmic ointment was instilled in the eyes to prevent corneal drying.

The animals' heads were supported with an orbito-palatal head holder. A craniotomy was performed over the right auditory cortex, which was located on the basis of cortical landmarks. The chronic recording array was implanted in the primary auditory area (A1) using a micromanipulator, with the four shanks approximately orthogonal to the cortical surface. After the array was inserted, the percutaneous cable was led to the midline and the connector was mounted on the skull. A stainless-steel cylinder (1.5 cm in diameter) with a screw-on cap was placed over the connectors to protect it from mechanical damage. The exposed brain surface was covered with silastic (Kwil-Sil, World Precision Instruments) and then with methacrylate cement which was anchored with stainless-steel bone screws. The cylinder was also secured with methacrylate and bone screws. The floor of the cylinder was covered with methacrylate to seal the inside of the cylinder from biological spaces. Following this, the skin was sutured around the cylinder. After approximately two weeks of recovery, we began with physiological recordings for a period of several weeks to several months.

Experimental apparatus. Experiments were conducted in a double-walled sound-attenuating anechoic chamber (Industrial Acoustics; inside dimensions 2.6 x 2.6 x 2.5 m) lined with SONEXone absorbent foam to suppress sound reflections. The chamber contained 13 8.4-cm-diameter two-way loudspeakers positioned on a horizontal circular hoop, 1.2 m in radius, at azimuths of 0 and ±10, 20, 40, 60° relative to the front of the apparatus. Left and right speaker locations are referred to as contralateral (C) and ipsilateral (I), respectively, relative to the electrode placement, which was in the cat's right hemisphere. The cat was supported on a raised platform, which was adjusted in height so that the cat's head was centered in the array of loud speakers. A harness restrained the animal to the platform but permitted free movement of the limbs and head.

Stimulus generation and data acquisition. Stimulus generation and data acquisition used System III hardware from Tucker-Davis Technologies (Alachua, FL) controlled by custom MATLAB software (The Mathworks; Natick, MA) on a Windows-based computer. Sounds were generated with a 24-bit precision at a sample rate of 97,656 s^{-1}. Loudspeakers were calibrated using a precision ½" microphone (ACO Pacific) that was positioned at the center of the apparatus at the normal location of the animal's head. Golay codes were used as probe sounds (Zhou et al., 1992). The calibration procedure yielded a 1029-tap finite-impulse-response correction filter for each speaker. The filters flattened and equalized the broadband frequency responses of the loudspeakers such that, for each loudspeaker, the standard deviation of the magnitude spectrum across the 0.2–25 kHz calibrated passband was <1 dB. The responses were rolled off by 10 dB from 25 to 40 kHz. Tonal stimuli were calibrated using pure-tone bursts in 1/24 octave intervals.

Study of neural responses consisted of measurements of frequency response areas (FRA), and synchrony to click and pulse trains. Frequency response areas were measured using pure tones, 100 ms in duration with 5 ms cosine-squared on and off ramps, at a repetition rate of 0.6/s, presented at the C 40° loudspeaker. To estimate the characteristic frequency (CF), we measured a coarse FRA from 0.5 to 40.0 kHz in 1/3-octave steps in 10 dB steps, from -10 to 60 dB SPL. The frequency response at the lowest threshold (i.e., lowest sound level) was taken as the characteristic frequency (CF). After acquiring a range of CFs from all active recording sites, we measured an FRA at a higher resolution consisting of tones 2 octaves below and 1 octave above the estimated CF range in either 1/6 or 1/12-octave steps, in level steps of 10 dB, typically from -40 to 80 dB SPL. Each combination of frequency and level was repeated 10 times. For the purpose of comparison with a study by Wang and colleagues, we also tested, in 3 animals,

responses to unmodulated tones that were 5 s in duration. These tones were presented only at the level that elicited a maximum response (i.e., BL). Taking the CF from the previous FRAs obtained using 100 ms tones, we presented long-duration tones that were 1/2 octave below and 1/6 octave above that CF. These were then presented in 1/6 octave steps.

Click trains were 1000 ms in duration and presented at a rate of 5 to 320 s^{-1} every 1.5 s. These trains of clicks varied in location between 0 and C 40° loudspeaker locations. Click trains for these locations were repeated 20 times for each presentation rate.

Pulse trains consisted of noise bursts that were generated in real time by gating a continuous Gaussian noise source generated by the TDT RZ6 digital signal processor. Each noise burst was 20 ms in duration and was gated with raised cosine functions with 1-ms rise and fall times presented at 60 dB SPL. Pulse trains were presented at rates of 5 to 40 s^{-1} in increments of 5 s^{-1} steps at a repetition period of 1.5 s, varied in presentation at the 0 and C 40° loudspeaker locations. The pulse train stimuli were repeated 20 times for each stimulus condition.

Experimental procedure. During data collection, cats sat on a raised platform in an anechoic sound attenuating chamber. A harness was used to restrain the animal to the platform but the animal had free movement of limbs and head. During recordings, an experimenter remained in the chamber with the animal, offering pureed canned food to keep the animal calm and quiet, while keeping the animal's head oriented toward the front of the loudspeaker array (~0° azimuth). If the animal became too restless and began moving around the pedestal, the recordings and sound presentation were paused until the animal calmed down. Each recording session was approximately 30 minutes to an hour, or until the animal refused to sit calmly on the pedestal. Data collection and recordings spanned from several days to several weeks, until we completed collecting all data.

Anesthetized Recordings. Two of the animals were also studied in a terminal experiment under anesthetized conditions. Anesthesia was induced with an intramuscular injection of ketamine (25 mg/kg) to permit shaving of the cat and insertion of an endotracheal tube. A 1:1 mixture of Nembutal and Ringer's solution was administrated intravenously to maintain an areflexive state during the physiological recordings. During the recordings, animals were suspended in the center of a sound attenuating booth. The animals' head was supported by a clamp around the metal cylinder that housed the electrode connector. Both pulse and respiratory rates were monitored in half hour intervals. Recordings lasted approximately two to three hours.

Data analysis. Stimulus-evoked neuronal action potentials were identified off-line with a spike sorting procedure that is described in Middlebrooks, 2008. First, a denoising procedure was used to attenuate waveforms that were correlated among all of the recording sites. Following this, the denoised waveforms were then interpolated to a 48.8 kHz time resolution using Fourier interpolation. Candidate action potentials were selected by setting a criterion spike height. Finally, the distributions of spike peak-to-trough amplitudes and times were examined to select single or multiple units. Well isolated single units were characterized by the following criteria: 1) uniform waveform appearance as observed through visual inspection, 2) consistent spike amplitude for the duration of the recording, and 3) inter-spike intervals that were >1 ms (i.e., longer than typical neural refractory periods). In many cases, single units could not be isolated, and we studied multi- unit responses corresponding to spikes of >2 neurons. Unless otherwise stated, all statistical analysis includes both single- and multiple-unit responses.

Spectral coding analysis. For awake cats, consistent responses to tone bursts were obtained from 12 well isolated single units and 42 multi-unit recordings. Frequency sensitivity

was plotted as FRAs consisting of mean spike rate (colors) as a function of sound frequency (horizontal axis) and level (vertical axis). The FRAs were classified as either V-, I-, or O-type following Ramachandran et al., 1999. For V-type FRAs, frequency tuning bandwidths increased with increasing sound level, and spike rates more or less increased with sound levels increasing up to the highest level tested. Units with I-shaped FRAs were level tolerant in that the bandwidth did not increase with increasing sound level; those units showed no marked decline in spike rates with increasing sound levels. Units with O-type FRAs, on the other hand, responded maximally at some BL, with responses declining at higher levels. The shape of the FRAs were based on visual inspection of the FRA. Out of the 54 units, 6 units were classified as O-type units while the remaining 48 were V- or I-type units.

We measured a monotonicity index (MI) to further study FRA shapes based on the neuronal spike rate. The MI was taken as the spike rate at the highest level tested divided by the maximum spike rate across all levels tested. MIs could range from 0, meaning few spikes at the highest level tested which is indicative of O-type units, to 1, meaning that the spike rate at the highest level tested is the same as the maximum spike rate which is indicative of V-type units.

Background rates of firing activity were measured during stimulus-free intervals equal in duration to stimulus intervals. The background rates presumably included spontaneous rates characteristic of each unit as well as neural responses to sounds generated by the cat's licking and other self-generated noises; these background rates tended to be higher than spontaneous rates recorded in head-restrained conditions (Sadagopan and Wang, 2008). Background rates were included in our FRA plots as well as in our measurements for MIs. Frequency tuning bandwidths however, were measured after subtraction of background rates. The maximum spike rate across all frequencies and levels was computed, and a threshold of 20% of that rate was

applied to all frequency/level conditions. The best frequency (BF) was the stimulus frequency yielding the highest mean spike rate across all frequencies and levels. The best level (BL), then, was defined as the sound level yielding the maximum firing rate at the BF. At each level, an area-matched rectangle was fitted to the tuning curve and the width of that rectangle was taken as the bandwidth (BW) at that level.

Analysis of time course of response. To look at the time course of response to tones, we computed a sustained response (SR) index. The SR index was taken as the number of spikes in the second half of the tone duration (60- 110 ms) as a proportion of the spike rate for the entire sound duration (10-110 ms). We had an offset of 10 ms when looking at recordings because neural spikes typically occurred ≥10 ms after the tone onset.

Additional tests were conducted to determine if units would exhibit a sustained response to tones that were 5 s in duration, similar to Wang et al. (2005). We recorded from three animals that we used in the previous measurement to test this and from those animals, we observed 16 multiple units and no single units. For these recordings, we presented tones at the BL and varied the tone frequency based on initial FRA measurements using 100 ms tones. The SR index was calculated as the proportion of spikes from a time range of 110-5010 ms (i.e., omitting the response to tone onset) divided by the response over the entire tone duration with an offset of 10 ms (i.e., 10-5010 ms).

In addition to recording from awake cats not participating in an auditory task, we also recorded from two of these cats under anesthesia. This was for comparison of the spectral coding properties between awake and anesthetized cortical recordings. From these two animals, we obtained 3 single units and 29 multi-unit recordings. We used the same criterion for the awake preparation to measure the MI, BF, BL, BW, and SR index.

Temporal coding analysis. We studied the neural representation of the rates of periodic trains of clicks and pulse trains. Responses to click trains in awake cats were studied at 37 sites, 9 of which yielded single-unit and 28 were multiple-unit responses. For pulse trains, we studied 42 sites. Of these sites, 8 were single units and 34 were multiple-unit recordings. Post-stimulus time histograms (PSTHs) for clicks and pulse trains show actual times in which spikes were elicited. Colors indicate the mean spike rate across time, relative to the stimulus onset.

After Lu et. al. (2001) and Middlebrooks and Snyder (2010), we evaluated coding of click- or pulse-train rates by synchronized neural spikes and by non-synchronized spike rates. Analysis of spike synchrony was based on the response interval from 100 to 1000 ms after stimulus onset, thereby omitting the highly synchronous response to the onset of the stimulus train. Synchrony was computed by treating each spike as a unit vector oriented with respect to the stimulus phase. Unit vectors were summed, then the mean length of the resultant was taken as the vector strength (Goldberg and Brown, 1969). The Rayleigh statistic was calculated using the equation $2nr^2$ (Mardia, 1972), where n is the number of spikes, and r is the vector strength. A Rayleigh-statistic criterion of 13.8 indicated synchronous firing with $p<0.001$. The synchronization boundary was defined as the fastest rate a neuron synchronized significantly to the stimulus train.

Stimulus coding by non-synchronized responses was examined by analysis of spike rates as a function of stimulus rates. Non-synchronized rate coding was characterized by tonic spike rates that were sustained and increased with increasing stimulus rate. The rate boundary for each unit was given by the lowest stimulus rate above which increasing stimulus rates yielded increasing spike rates. Spike-rate coding by high rates was identified by the discharge rate ratio (DRR) which was the highest spike rate to stimulus rates 160 s^{-1} or higher divided by the highest

spike rate across all stimulus rates. Background rates were subtracted prior to calculating the DRR. DRRs could range from negative values up to 1. A DRR of 1, indicates that the highest spike rate to stimulus rate 160 s^{-1} or higher was also the highest spike rate measured. This is indicative of tonic asynchronous response firing at higher presentation rates. Negative DRRs indicate suppression of neural responses by high click rates (i.e., the highest spike rate to stimulus rate 160 s^{-1} or higher was below the firing rate of background activity), indicative of neuronal firing at lower presentation rates.

3.4 Results

We begin by characterizing observations for neural recordings in response to pure tones to study the spectral coding properties in cats. We then characterize the temporal coding properties of neurons in response to trains of clicks and noise bursts.

Spectral coding properties. The analysis for the spectral coding properties include 54 units in awake animals in non-task conditions and 32 units from two of those animals studied under conditions of barbiturate anesthesia. To characterize the spectral response properties in awake cats, we presented 100-ms tones that were varied in both frequency and sound pressure level. We observed a variety of FRAs, which include V-shapes that are commonly observed in anesthetized cat cortex. We also observed level-tolerant I- and O- shapes that are rarely seen in anesthetized animal preparations. Figure 3.1 shows example FRAs of V- (A), I- (B), and O-type (C) responses from single (A) and multiple (B and C) units in awake cats. FRAs are plotted as the mean spike rate, denoted by color, as a function of frequency and sound pressure level. V-shaped responses are characterized as being level intolerant in that the frequency tuning widens with increasing sound levels. In contrast, I- and O-shaped responses show tolerance to

increasing sound levels where I-shaped units exhibit the same bandwidth at all sound levels tested and for O-type units, the frequency tuning narrowed at sound levels increasing above the best level.

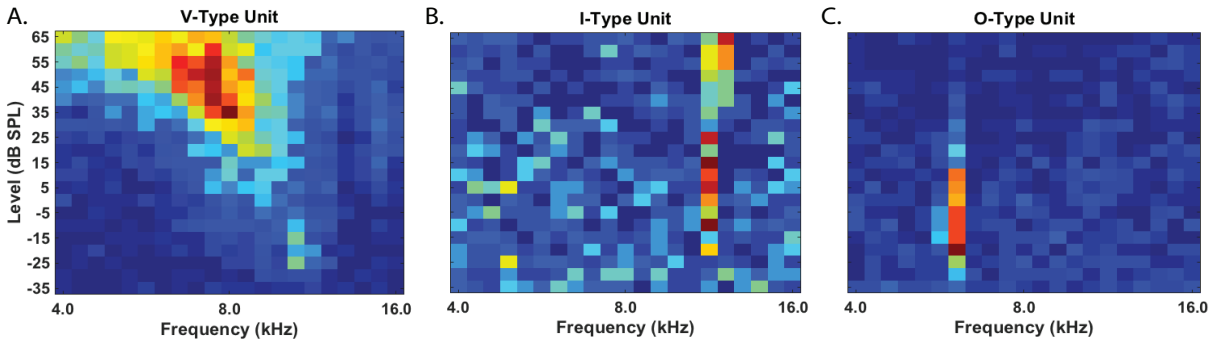

Fig. 3.1. Frequency response areas of single- (A) and multiple (B and C) units from one animal showing V- (A), I- (B), and O-type (C) responses.

To further study FRA types, we measured a monotonicity index (MI) which was defined as the spike rate at the highest level tested divided by the maximum spike rate across all levels tested. This measure determined whether spike rate increased or decreased at the maximum level tested. O-type units for example had low MI values because their response firing declined at higher levels. V-type units on the other hand, typically had higher MI values. The highest MI value of 1, indicates that the spike rate at the maximum level tested is also the maximum spike rate across all levels. An MI of 0, in contrast, would indicate that the spike rate declined from some high value at the best level to zero at the highest level tested. A distribution of MI values in awake cats is shown in Figure 3.2A for O-type (shaded bars) and V/I-type (open bars) units. O- and V/I-type units were identified by visual inspection of FRAs. We did not differentiate between V- and I-type units for this analysis. Median MI values for O-type and V/I-type units are 0.13 and 0.57, respectively. Filled symbols indicate single units. 6 of the 52 units (11%)

were characterized as O-type units, and of those 6, 2 were single units. The majority of units (48/54, 89%) were characterized as V- or I-type units. Of those units, only 10 were single units. We performed a two-sample Kolmogorov-Smirnov test to see whether single- and multiple-units in awake cats showed differences in their MIs and did not find any significant differences (p=0.54). Unlike the awake condition, there were no O-type units in the anesthetized condition. The median MI for units in the anesthetized condition was 0.85. We found a significant difference between MI values for the awake and anesthetized conditions (p=1.20 x 10^{-3}, Kolmogorov-Smirnov).

A few V/I-type units had MI values <0.3. Figure 3.2B and C are examples of two V/I-type FRAs having different MI values. The V-type unit in Figure 3.2B had an MI value of 0.21 while that in Figure 3.2C, had an MI of 1. For comparison, the MI values in Figure 3.1 were 0.60 (A), 0.47 (B) and 0.05 (C).

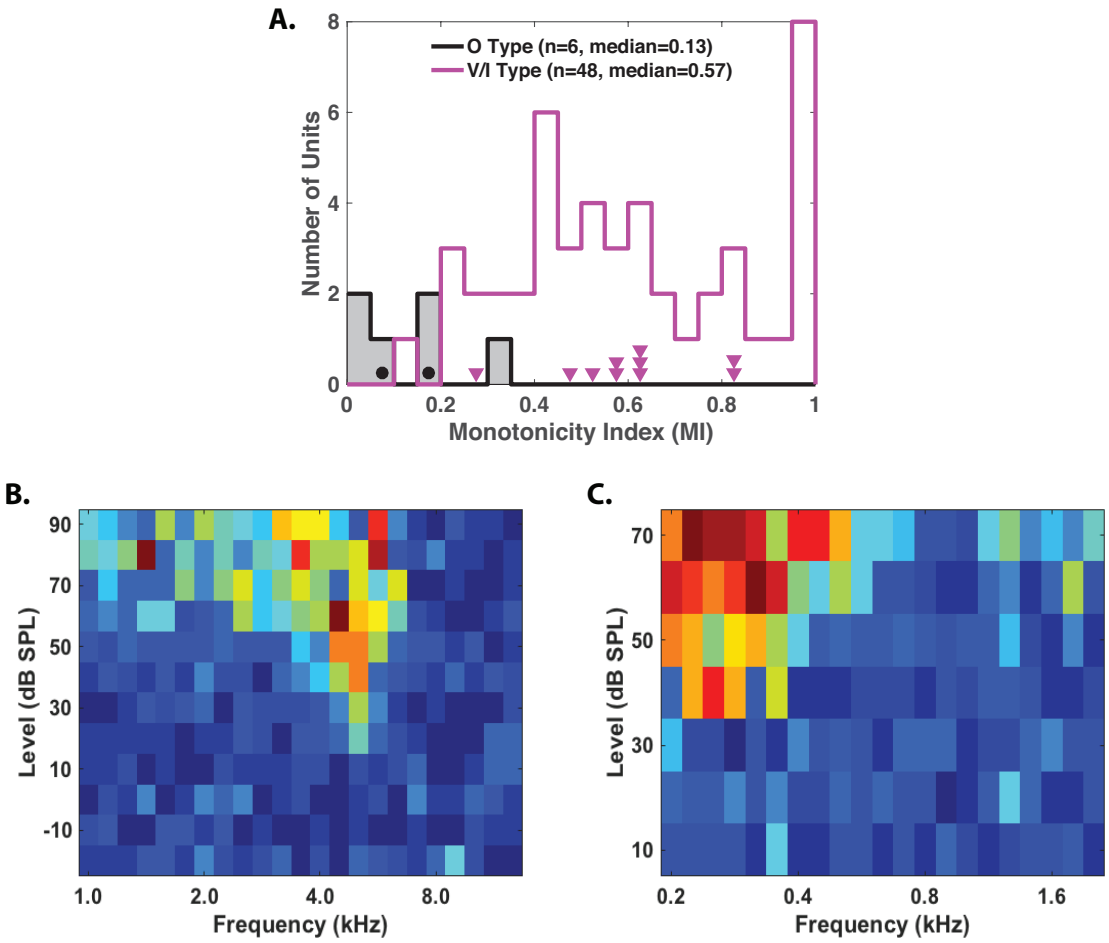

Fig. 3.2. Monotonicity indices (MI). (A) Distribution of MIs for O- (shaded bars) and V/I-type (open bars) units in awake cats. Filled symbols indicate single O- (filled circles) and V/I-type (filled triangle) units. (B and C) FRAs from multiple units with MI values of 0.21 (B) and 1 (C).

We measured frequency tuning BWs to further study whether neurons in awake cortex were sharply tuned compared to neurons in anesthetized cortex. BW was computed at the BL which was computed at the BF. Figure 3.3A plots the BWs for awake (shaded bars) and anesthetized (open bars) animals. Single units are indicated by the filled and open symbols for awake and anesthetized animals, respectively. Median BWs were 0.33 octave for awake conditions and 1.14 octave for anesthetized conditions. The distribution for the awake animals showed a prominent mode below 0.5 octave, encompassing 63% (35/54) of units. For comparison, in the anesthetized condition, only 3% (1/32) of units had a BW <0.5 octave.

Bartlett et al. (2011) defined finely tuned units in the awake marmoset as having BWs ≤0.2 octaves at the BL. Based on that criterion, 20% (11/54) of units were sharply tuned in our awake preparation, and no units met the ≤0.2-octave criterion in the anesthetized preparation. We found a significant difference between awake and anesthetized conditions (p=1.35 x 10^{-10}, Kolmogorov-Smirnov) in that in awake conditions, BWs were narrower compared to anesthetized conditions. We also tested whether single- or multiple-units in awake cortex showed differences in their BW. Single- and multiple-units did not show significant differences in their BWs (p=0.07, Kolmogorov-Smirnov).

In Figure 3.3B, we plotted MI as a function of BW for units in the awake (black circles and magenta triangles) and anesthetized (blue triangles) conditions. Single units are indicated by filled symbols and multiple units are indicated by open symbols. O-type units had MI values <0.4 as well as BWs <0.5 octaves. A few units classified as V/I-type also had low MI and BWs, similar to O-type units. In the anesthetized condition, units characterized as V- and I-type tended to have BWs >0.5 octaves and MIs >0.2.

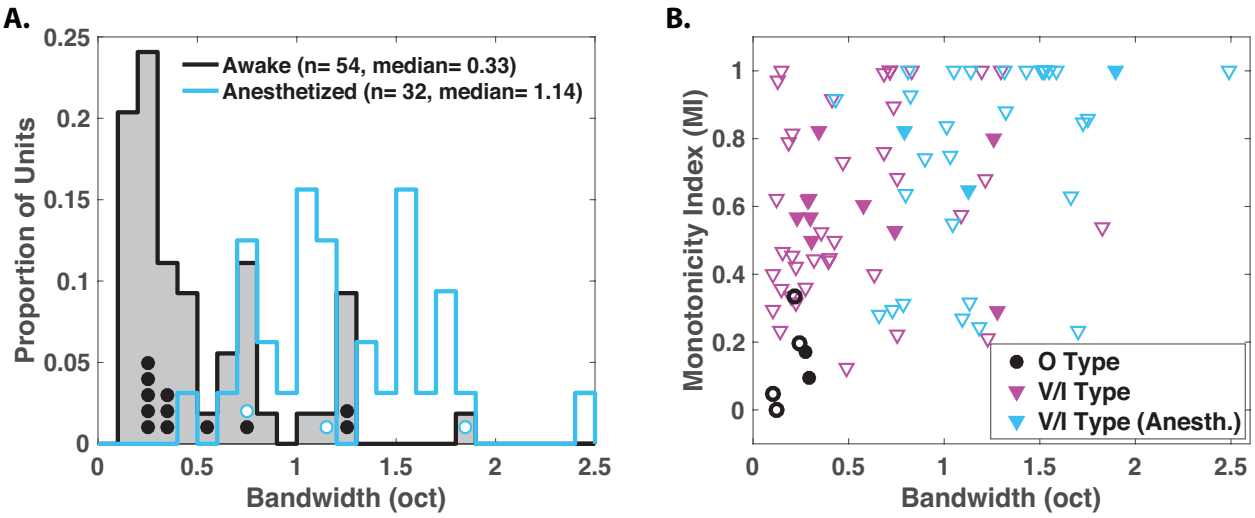

Fig. 3.3. Equivalent rectangular bandwidths measured at the sound level giving the maximum response. The proportion of units separated by bandwidth for awake (shaded bars) and anesthetized (open bars) cats for both single- and multiple- units. Circles indicate single units for awake (black, filled) and anesthetized (blue, open) cats. (B) Scatterplot of MIs as a function of BW. Filled symbols indicate single units and open symbols indicate multiple units. Black circles and magenta triangles are units in awake animals. Blue triangles represent units in anesthetized animals.

We looked at data in a single animal to see whether BF and BW differed under awake and anesthetized conditions. Figure 3.4 shows distributions of BF (A) and BW (B) for this animal. We found that there were no significant differences in BF between the awake and anesthetized conditions (p=0.38, Kolmogorov-Smirnov). There was however, a significant difference between BW in the awake and anesthetized conditions (p=2.47 x 10^{-6}, Kolmogorov-Smirnov). This result coincides with the combined data across all animals as shown previously indicating that the effect we see is not limited across cats but could also be seen in individual cats.

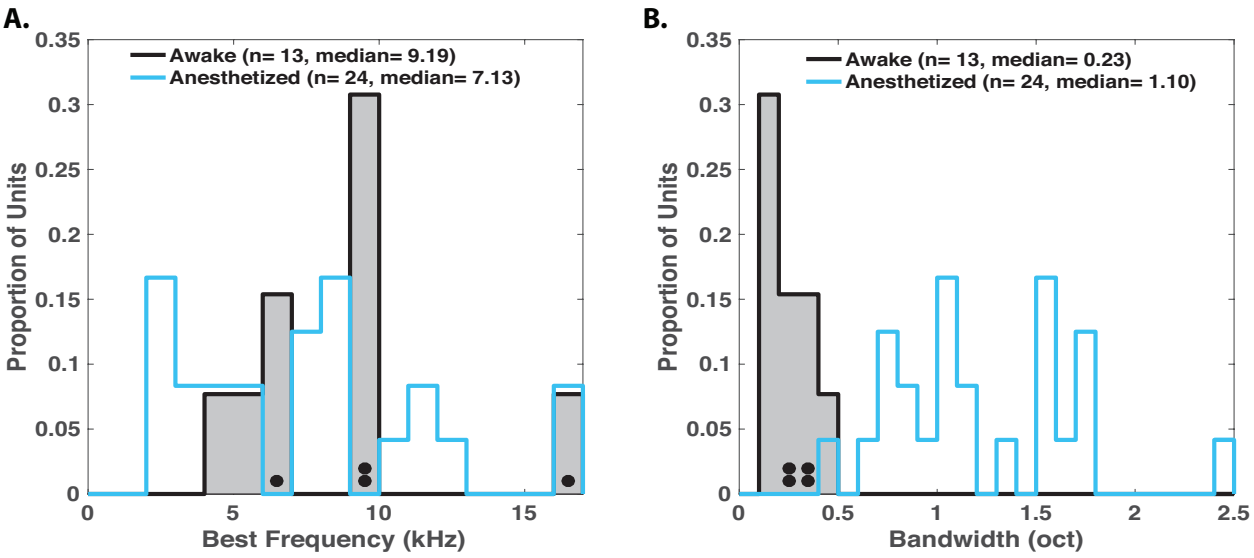

Fig. 3.4. Distributions of BF (A) and BW (B) for a single animal. The proportion of units are separated by awake (shaded bars) and anesthetized (open bars) conditions. Filled circles represent single units in the awake condition. There were no single units in the anesthetized condition.

Time course of responses. Most studies of auditory cortical responses in anesthetized conditions show responses to unmodulated sounds that are largely restricted to the sound onset (reviewed by Phillips, 1993). In contrast, Wang and colleagues (2005) often observed sustained firing in auditory cortex for the entire stimulus duration, especially when neurons were driven by their preferred stimuli. In our awake cat preparation, we often saw sustained firing that lasted throughout or even beyond the 100-ms tone durations when units were presented with preferred frequencies and/or levels. Less-preferred tones only elicited onset responses. Figure 3.5 illustrates examples of response firing of single units showing sustained, onset and offset, and onset only responses. Figure 3.5A and B show responses from the same neuron and differ only in the sound levels that were presented. When the neuron was presented with 100-ms tones at low levels (-30 to -20 dB SPL), the neuron fired throughout the durations of tones in the 10.5 to 12.1 kHz frequency range. If, however, the sound level was increased to 50 to 60 dB SPL, the response broadened to all frequencies tested and the response to the 10.5 to 12.1 kHz frequency

range changed to a strictly onset response. An offset response also appeared at lower frequencies between 8 and 9.2 kHz. Figure 3.5C shows a strong onset response from a different single unit.

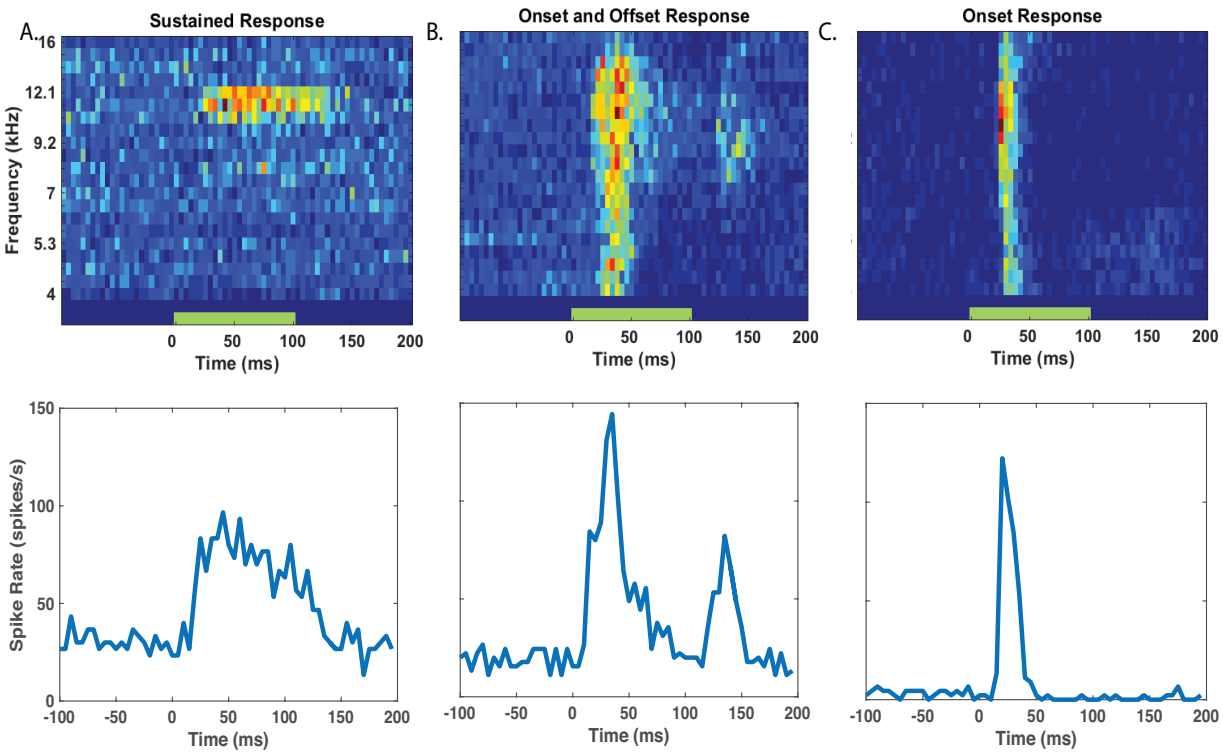

Fig. 3.5. Examples of response firing of single units. In the top row, colors indicate mean spike rates and vertical and horizontal axes represent tone frequency and time re stimulus onset. **A** and **B** are show responses from one single unit to 100-ms tones at low (A: -30 to -20 dB SPL) and high (B: 50 to 60 dB SPL) sound levels. **C** shows the response of a different unit to 50 to 60 dB SPL levels. The bottom panels are PSTHs at a frequency range of 10.5 to 12.1 (A) and 8 to 9.2 (B and C).

To characterize units that exhibited sustained responses, we measured an SR index which is defined as the number of spikes in the second half of the tone duration (60 to 110 ms) as a proportion of the entire recording duration at the BL of the BF. We found that most units exhibiting a sustained response upon inspection of PSTHs had an SR index >0.2. For this reason, we took that as our criterion for classifying neurons that had sustained responses. Of the 54 units for which we studied responses to tones, 76% (41/54) were classified as sustained units and 24% (13/54) were classified as having a phasic (i.e., onset) response at the BL. For V-/I-,

and O- type units, majority of those neurons exhibited sustained responses (35/48 for V-/I- type, and 6/6 for O-type). In anesthetized animals, a number of units showed sustained responses (12/32), but majority of responses were phasic (20/32). Figure 3.6A shows a distribution of SR indices for awake and anesthetized cats. The median SR indices for awake and anesthetized conditions were 0.34 and 0.15, respectively. We performed a two-sample Kolmogorov-Smirnov test to test whether the SR indices differed for awake and anesthetized conditions. Overall, we found a significant difference between awake and anesthetized conditions (p=7.78 x 10^{-4}) in that awake conditions exhibited more sustained responses compared to anesthetized conditions. In figure 3.6B, we plotted a cumulative distribution of SR indices for awake (black line) versus anesthetized (blue line) conditions. Larger SR values reflect more of a sustained response as shown by units in the awake condition.

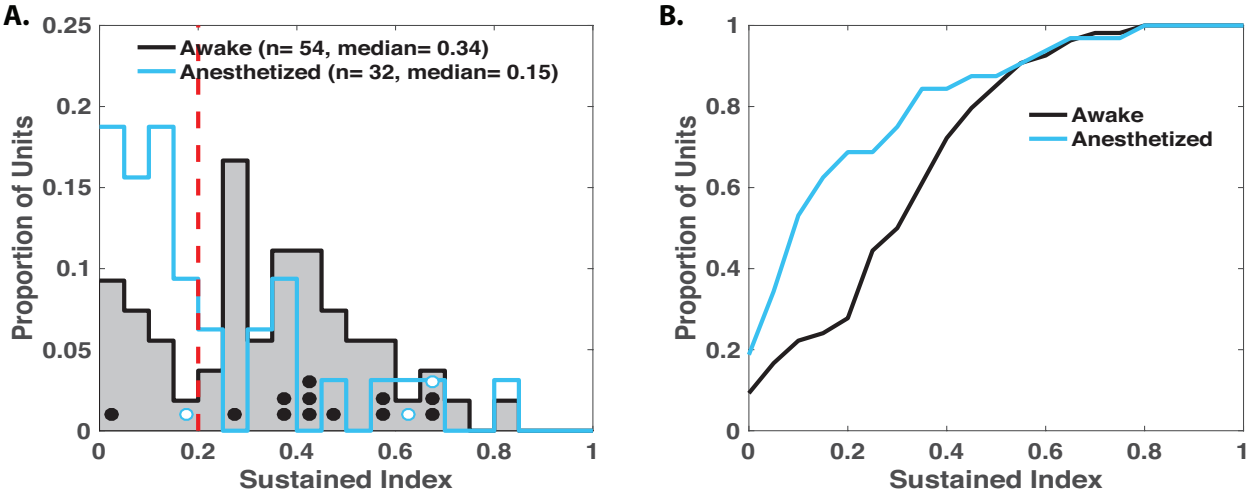

Fig. 3.6. Sustained response (SR) index measured at the level of the BF giving the maximum response. **A** shows the proportion of single- and multiple units separated by SR for awake (shaded bars) and anesthetized (open bars) cats. Single units are also indicated by circles for awake (black, filled) and anesthetized (blue, open) cats. We classified units that had an SR index of ≥0.2 as having a sustained response. The dashed red line indicates an SR index at 0.2. **B** is a cumulative distribution of SR indices for awake (solid black line) and anesthetized (solid blue line).

In addition to presenting 100 ms tones, we also presented 3 animals with tones that were 5 s in duration for comparison to work done by Wang et al. (2005). For that study they observed neurons that fired in a sustained manner for the entire duration of the 5 s tone. For the presentation of long duration tones, we presented tones at the BL (i.e., the level that elicited the maximum response) and varied the tone frequency for each trial. We observed 16 units and all of those were multiple units. Figure 3.7 is an example of a unit exhibiting a sustained response at a frequency range of 4.5 to 5.7 kHz. Figure 3.7A shows the mean spike rate with the first 100 ms of the tone presentation to illustrate the strong onset response across a broad frequency range. The onset response tended to saturate the color scale. Figure 3.7B shows the same data, limited to post-stimulus times 100 to 5000 ms. The line plots show post-stimulus-time response rates for limited ranges of preferred (3.7D) and non-preferred (3.7C and E) frequencies as indicated by the vertical red, magenta, and black lines on the right-hand side of 3.7A and B.

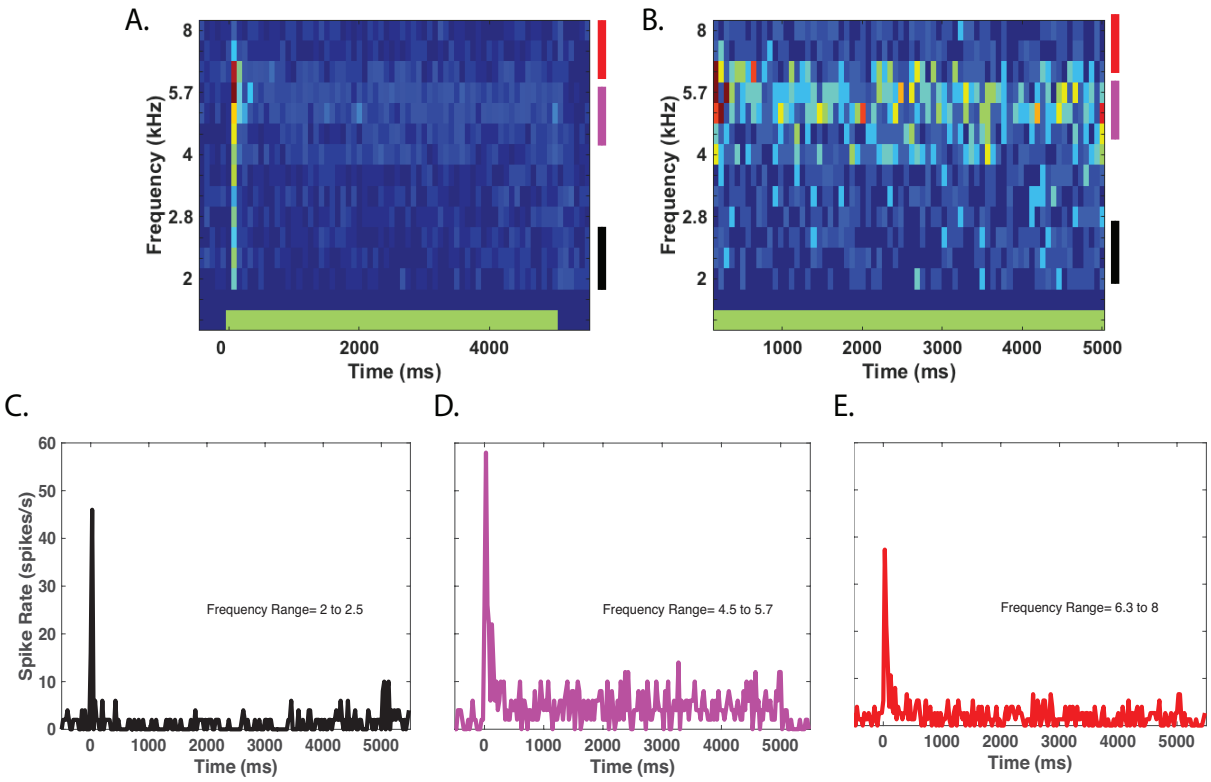

Fig. 3.7. Example of response firing of one multiple unit for 5 s tones showing a sustained response at a frequency range of 4.5 to 5.7 kHz. For **A** and **B**, colors indicate mean spike rates and vertical and horizontal axes represent tone frequency and time re stimulus presentation. **A** and **B** are from the same data but **B** is limited to post-stimulus times of 100 to 5000 ms. Bottom row shows PSTHs of the same neuron but at different frequencies, all at the same level of 50 dB SPL. In **A** and **B**, colored vertical bars on the right-hand side of the panels correspond to the frequency ranges of the post-stimulus-time response rates shown below (**C**, **D**, and **E**).

To characterize units exhibiting sustained responses for 5 s tones, we measured an SR index which was the number of spikes from 110-5010 ms as a proportion of the entire sound duration (10-5010 ms). That number could range from 0, a response only at sound onset, to 1.0, meaning only a late response. The median SR index was 0.87. SR indices were plotted in Figure 3.8.

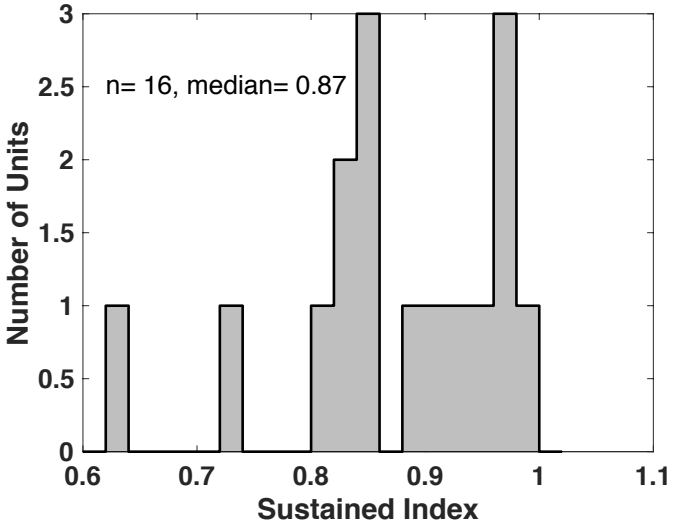

Fig. 3.8. Sustained response (SR) index measured at the BF for long duration tones for multiple units.

In addition to observing sustained and phasic responses when presenting long duration tones, we did find one unit that showed a strong onset response followed by a tonic frequency-specific suppression of background activity. An illustration of this unit is shown in Figure 3.9. At low frequencies, the unit had a strong onset response followed by a suppression of background activity that lasted for the duration of the tone. After the tone presentation, there was an increase of background activity. At high frequencies, background activity was present which does not seem to be affected by the stimulus presentation.

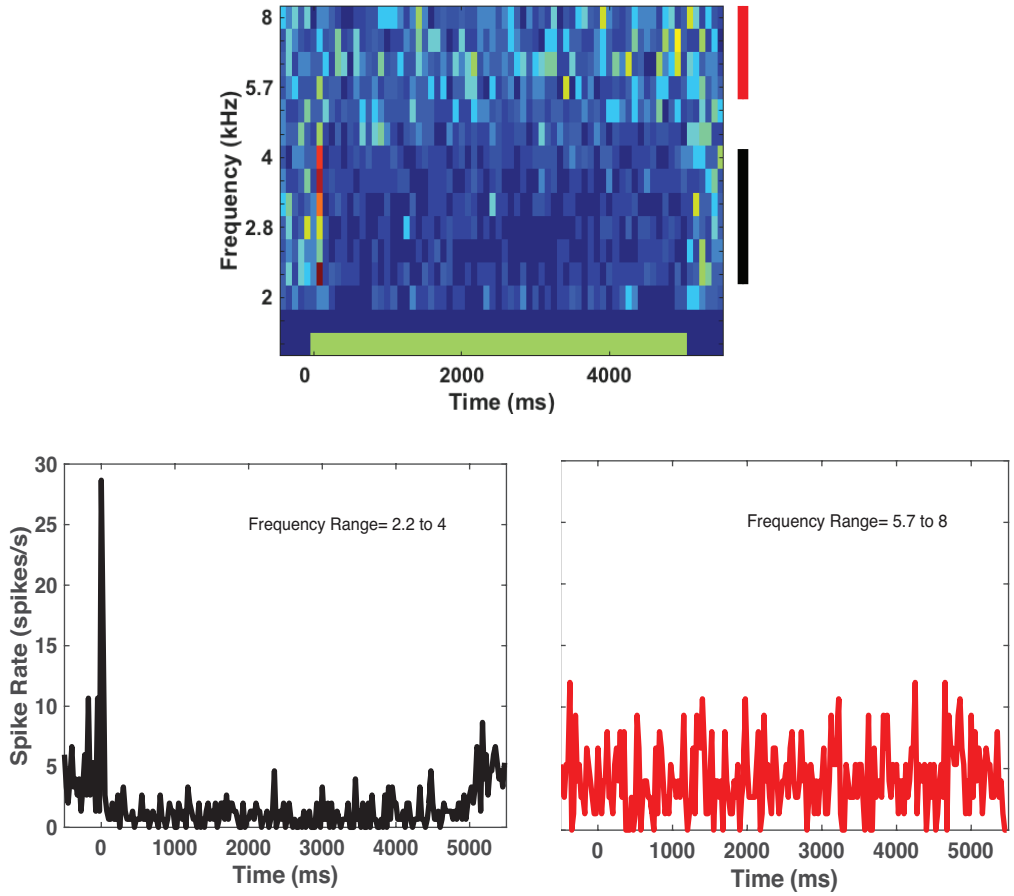

Fig. 3.9. Example of response firing of one multiple unit for 5 s tones showing tonic frequency suppression of background activity. In the top row, colors indicate mean spike rates and vertical and horizontal axes represent tone frequency and time re stimulus onset. Bottom row shows PSTHs of the same neuron but at different frequencies (indicated by the red and black vertical lines on the right-hand side of the top panel), all at the same level of 50 dB SPL.

Temporal coding properties. We measured the ability of neurons to represent the rates of periodic click or pulse trains. The trains of short-duration clicks permitted study at high repetition rates and, also, permitted comparison with published data from the marmoset (Lu et al., 2001). The trains of 20-ms pulses were used to evaluate responses to repeated sounds like those used in our previous psychophysical study (Chapter 2; Javier et al., 2016) and in our study of physiological substrates of stream segregation (Chapter 4); the 20-ms pulses were tested only up to rates of 40 s^{-1}, at which the repetition period of the train is 25 ms. Click and pulse trains all were 1000 ms in duration.

Approximately half of the units demonstrated sustained responses synchronized to trains of clicks or noise bursts, with half of those synchronizing to click rates >40 s^{-1}. Examples of synchronized units are shown in Figure 3.10. The left-most columns are PSTHs for clicks (A) and pulse trains (B), where colors indicate the mean spike rates to individual clicks or noise bursts as a function of time. The vertical and horizontal axes represent the click or pulse presentation rate (s^{-1}) and time re stimulus onset. The horizontal green line at the bottom of the figures represent the duration of the stimulus. Synchronized responses are predominantly seen at lower rates and start to fall off at higher rates. We measured the amount of synchrony in the right-most columns. The solid red line shows the tonic spike rate, excluding the onset response in the first 100 ms and with the background rates subtracted. The dashed red line indicates the spike rate at 0 relative to the background rate. We calculated a Rayleigh Statistic (solid blue line) to characterize synchronous activity. Rayleigh Statistic values that fall above the Rayleigh criterion of 13.8 (dashed blue line) indicates significant synchrony (p<0.001). The stimulus rate at which the Rayleigh Statistic crosses the Rayleigh criterion indicates the synchronization boundary, which is defined by Lu et al. (2001) as the fastest rate in which a neuron shows significant stimulus-synchronized activity. The unit in figure 3.10A synchronized to click trains up to ~67 s^{-1}. The unit in figure. 3.10B synchronized to pulse trains up ~30 s^{-1}. Both of these units showed a decline in tonic spike rates (red lines) with increasing stimulus rate, roughly parallel to the decline in synchrony. The latter observation shows that tonic firing by synchronized units was found only at stimulus rates at which synchrony was present. This contrasts with non-synchronized units, presented below, that responded to high stimulus rates with tonic non-synchronized responses.

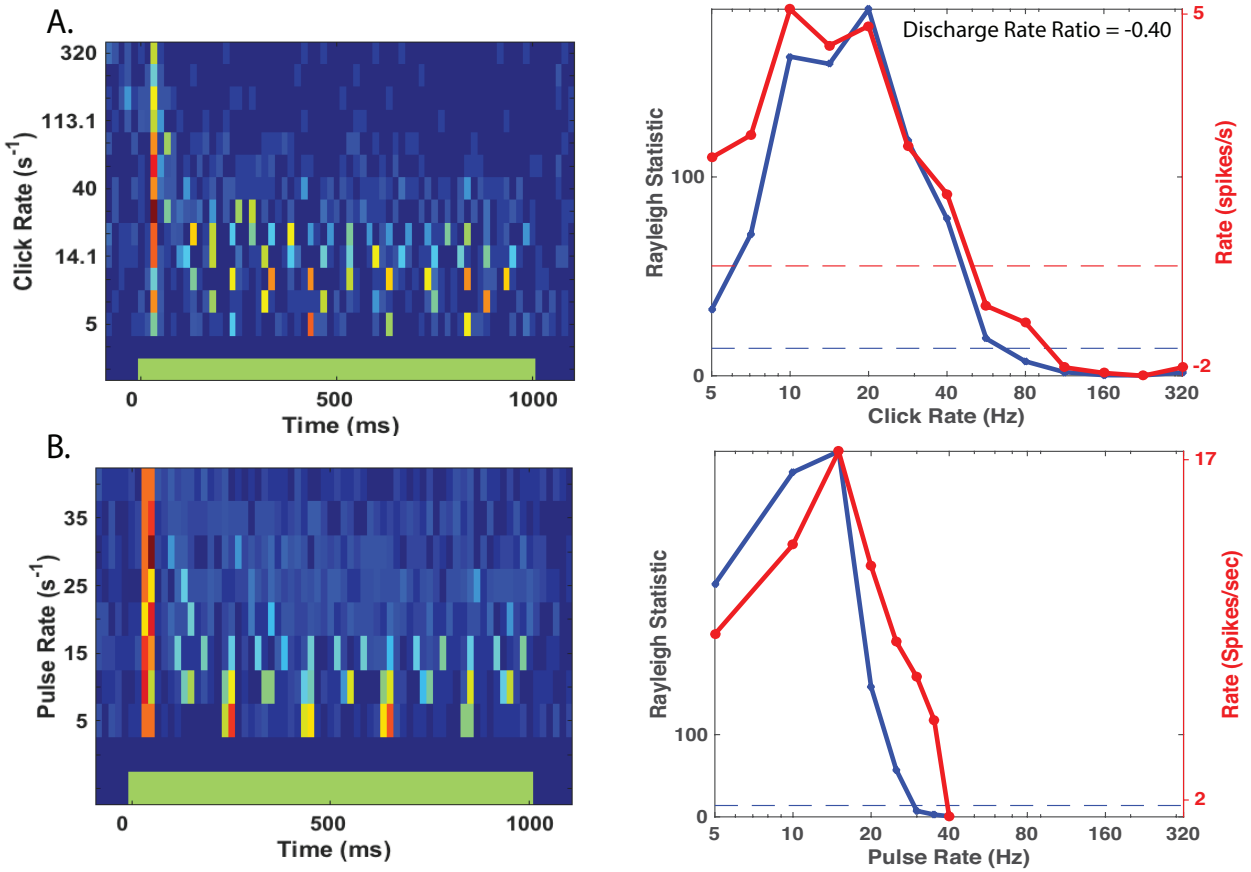

Fig. 3.10. Examples of synchronized units for click (**A**) and pulse (**B**) trains. The left column are PSTHs of neurons synchronized to clicks (top) and pulse (bottom) trains. Colors indicate the mean spike rate. The right column shows the Raleigh Statistic (solid blue line). Rayleigh statistics that fall above the dashed blue line (Rayleigh criterion = 13.8) indicate significant synchrony (p<0.001). The solid red line shows tonic spike rates, excluded the onset response in the first 100 ms and with background rates subtracted. The dashed red line indicates 0 spikes/s re background.

We obtained synchronization boundaries for units that synchronized their responses to click trains (A) and noise bursts (B) and plotted a distribution of these values in Figure 3.11. Shaded bars indicate units with synchronization boundaries and open bars indicate units that did not synchronize their responses to the stimuli. Figure 3.11A shows a distribution of synchronization boundaries for click trains. The median synchronization boundary for clicks was ~40 s^{-1}. Neurons that synchronized to noise bursts, shown in Figure 3.11B, had a median synchronization boundary of ~33 s^{-1}.

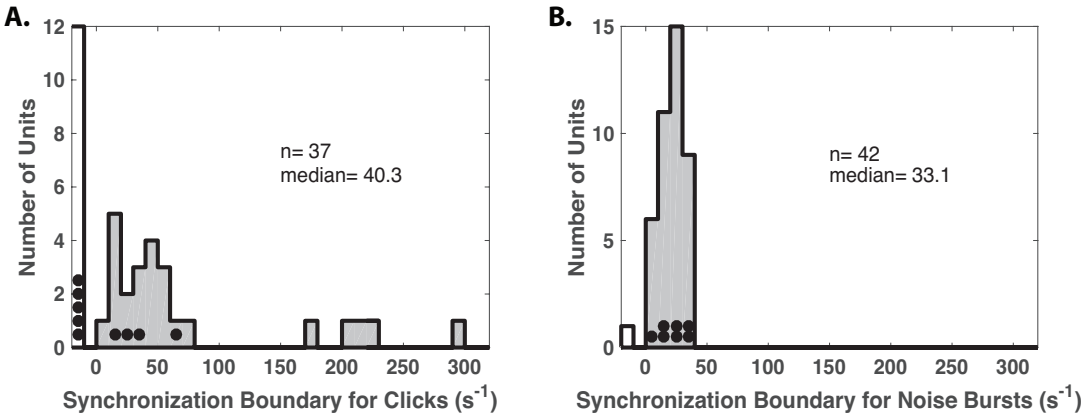

Fig. 3.11. Distribution of synchronization boundaries for click (**A**) and pulse (**B**) trains are shown in shaded bars. The synchronization boundary was defined as the fastest rate in which a neuron showed significant stimulus-synchronized activity (Lu et al., 2001). Open bars represent units that did not phase lock to the stimuli at any tested rate and thus had no synchronization boundary values.

Some units showed minimal sustained responses to clicks at low rates, but fired in a nonsynchronous sustained manner to click rates 50 s^{-1} and greater. In these cases, spike rates increased with increasing click rate with responses peaking at a click rate around 160 s^{-1} and then declining at faster rates. An example of this is shown in Figure 3.12.

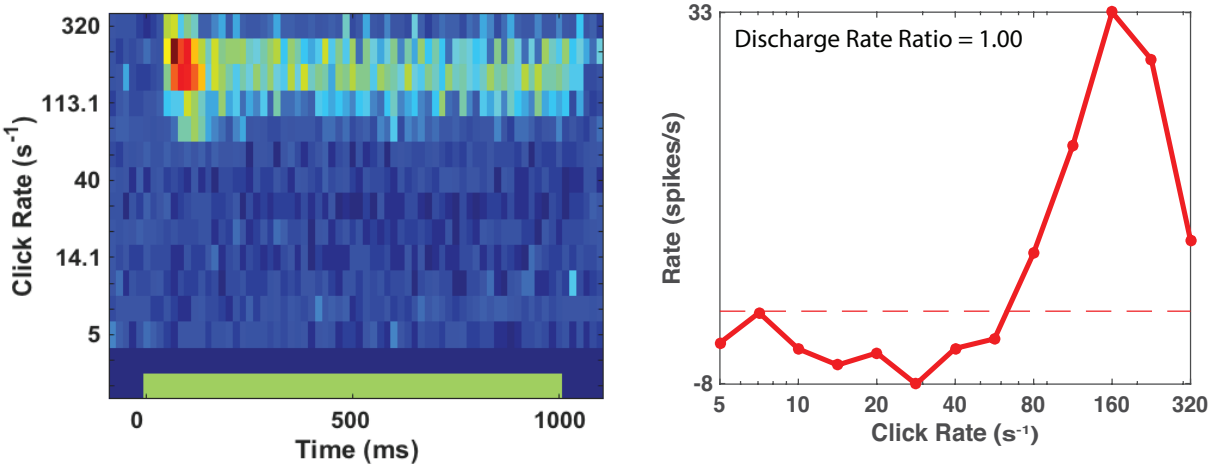

Fig. 3.12. Example of a non-synchronized units for click trains. The left panel is a PSTH of a neuron showing a sustained, nonsynchronous response to trains of clicks. Colors indicate the mean spike rate. The right panel shows the tonic spike rates, excluded the onset response in the first 100 ms and with background rates subtracted. The dashed red line indicates 0 spikes/s re background.

We examined the stimulus coding by non-synchronized responses by looking at the discharge rate ratio (DRR). This was defined as the maximum spike rate to stimulus rates greater than or equal to 160 s^{-1} divided by the maximum spike rates across all stimulus rates. The calculated DRR did not include background firing rates (i.e., background rates were subtracted before computing the ratio). Negative DRRs indicate units that showed neural response firing at low click rates followed by suppression of neural responses at higher rates that were below background activity. An example of this is shown in Figure 3.10A in which the DRR is -0.40. DRRs that approached 1 indicate neurons that responded maximally at higher presentation rates as shown in Figure 3.12 where the DRR is 1. Figure 3.13A shows a distribution of DRRs across both single- and multiple-units for click trains. Filled symbols represent single units. Single units tended to have DRRs approaching that of 1. Although this was the case, there were no significant differences in DRRs between single- and multiple-units (p=0.61, Kolmogorov-Smirnov).

We also looked at the Raleigh Statistic at 10 s^{-1} as a function of the DRR to see if we could parse out the different types of responses (Figure 3.13B). Neurons that had high DRRs approaching 1 had a tendency towards low Rayleigh statistics at 10 s^{-1}. These units tended to fire asynchronously at high rates. One outlier unit that had a DRR of 1 and a Rayleigh statistic of ~200 showed synchronized responses at slow rates and sustained nonsynchronous responses at higher rates. Neurons that had DRRs of ~-2 showed asynchronous firing at low rates. Those units that had DRRs ranging from -1.5 to 0 showed synchronous responses at low rates, indicated by the Rayleigh statistics >13.8 (i.e., criterion for significant response synchrony), followed by suppression at higher rates. The suppression of the neural response at higher rates is lower than that of the background rate, resulting in a negative DRR value. Units that had DRRs from 0 to

<1 showed response firing at higher click rates and often, exhibited phase locking at those higher click rates.

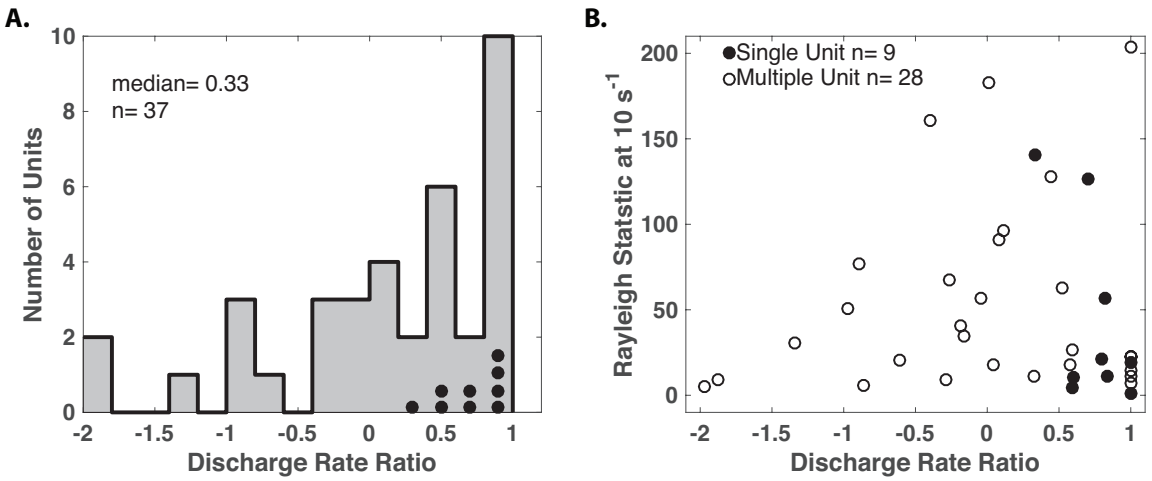

Fig. 3.13. Discharge rate ratios for click trains. The discharge rate ratio was defined as the maximum spike rate to stimulus rates greater than or equal to 160 s^{-1} divided by the maximum spike rates to stimulus rates. **A** Distribution of DRRs across all units. Filled circles represent single units. **B** DRR as a function of Rayleigh Statistics at 10 s^{-1}. Filled symbols represent single units and open symbols represent multiple units.

3.5 Discussion

Neurons in cortical area A1 of awake cats demonstrate remarkable spectral and temporal coding properties that have previously been seen only in awake marmoset cortex. These properties include frequency response areas that are restricted in frequency and level, sharp frequency tuning bandwidths, sustained responses, neurons that can phase lock to rates >40 s^{-1}, and tonic asynchronous firing that increases with increasing stimulus rates. The present results indicate that these properties are not unique to marmoset, but that they apparently can be seen only in unanesthetized animals. We will first discuss spectral coding properties, followed by time course of response and, temporal coding properties.

Spectral coding properties. Studies in the DCN (Young and Brownell, 1976) and ICC (Ramachandran et al., 1999) of decerebrate cats described units with FRAs with neural excitation limited in both frequency and level. The use of decerebrate preparations allow for the characterization of neural response properties without the confounding effects of anesthesia. In the DCN, Young and Brownell (1976) characterized Type IV units, which for the most part, resembled O-type units in the ICC in decerebrate cats (Ramachandran et al., 1999) and in A1 of awake marmosets (Sadagopan and Wang, 2008). Young and Brownell (1976) found that adding a barbiturate caused inhibitory areas of a Type IV neuron to become excitatory, such that frequency tuning became wider with increasing sound level. This changed the neuron's classification from Type IV to Type II/III (i.e., V-shaped). In awake marmosets, Sadagopan and Wang, 2008 classified majority (64%) of A1 neurons as having O-shaped FRAs. In our awake cat preparation, we also observed O-shaped and I-shaped FRAs, however a smaller proportion of our units (11%) were classified as O-type units. For our anesthetized condition, we did not observe any O-type units. Taken together, these studies and ours suggest that the study of O-type and I-type units is possible under awake preparations.

In the current study we found frequency tuning to be narrower in awake cats in comparison to tuning seen in anesthetized cats. The median BW for awake cats was 0.33 octaves while that of two anesthetized animals was 1.14 octaves. The finding of broad tuning in anesthetized conditions is supported by previous studies in anesthetized cat in which BW has been found to be >1 octave (Sutter and Schreiner, 1991; Schreiner and Sutter, 1992; Heil et al., 1992; Moshitch et al., 2006). Studies by Sutter and Schreiner (1991, 1992) measured BW as a function of mediolateral position along isofrequency contours of A1. In the dorsal part of A1, they observed units that exhibited multi-peaked FRAs which contained two or more excitatory

frequency ranges. It is likely that what they referred to as the dorsal part of A1 is now regarded as the dorsal zone (DZ). Middlebrooks and Zook (1983) described area DZ and noted that neurons in this cortical area were more broadly tuned for frequency compared to other areas of A1. In addition, neurons in area DZ have been shown to exhibit multi-peaked frequency tuning (Stecker et al., 2005). In the present experiment, our electrode placements were targeted towards cortical areas ventral to area DZ and we did not see multi-peaked FRAs as reported by Sutter and Schreiner (1991). In a subsequent paper, Schreiner and Sutter (1992) found that single units and units located in the central division of A1 had the narrowest BW. Mean BWs between single- and multiple units in the central division were 0.44 octaves (0.31 octaves for single units), while that in the dorsal and ventral divisions of A1 had mean BWs of 0.93 octaves and 0.68 octaves, respectively. In our study, we found that single and multiple-unit recordings showed no systematic difference in spectral sensitivity, suggesting that BW was similar across units. Our experimental design however, did not permit extensive cortical mapping of A1 like that of Schreiner and Sutter. For that reason, our unit sample might have included recordings from both narrowly and broadly tuned cortical areas. Overall, our findings of fine frequency tuning in awake cats agrees with that of Bartlett et al., (2011) who investigated frequency tuning in awake marmosets. In their findings, 27% of units were sharply tuned with a mean BW of 0.13 octaves. Based on their criteria, we found that 20% of our units were sharply tuned.

Time course of responses. In addition to investigating properties pertaining to FRA shape and tuning, we also looked at the time course of responses to tone stimuli. For both our 100 ms and 5000 ms duration tones, we saw a variety of responses, including sustained responses that lasted for the duration of the stimulus presentation. This agrees with previous studies in which sustained firing has been characterized in awake marmosets (Wang et al., 2005)

as well as in other studies using awake cats (Evans and Whitfield, 1964; Chimoto et al., 2002). We also saw transient onset and offset responses which are predominantly observed in anesthetized conditions. In anesthetized preparations, neurons typically exhibit onset responses regardless of the duration of the stimulus presentation (Eggermont, 1991; reviewed by Phillips, 1993).

Temporal coding properties. Neurons in A1 were found to synchronize their responses to rates >40 s^{-1} which is higher than what is typically seen in anesthetized preparations. Much of the temporal properties presented in this chapter focused on click trains, but we also looked at response phase locking to noise bursts. The animals used for this study were initially trained in a stream segregation task (described in Chapter 2) where they needed to detect a change in rhythm in noise bursts presented at a rate of 10 s^{-1}. We were interested in measuring the maximum rate in which neurons could synchronize to these noise bursts in an awake animal. We focused on click trains for comparison to that of work by Lu et al., (2001) in awake marmosets and also, for the reason that clicks could be presented at faster rates to allow for study of synchronous and nonsynchronous responses. This discussion will first go over response synchrony to noise bursts followed by response synchrony and asynchrony to click trains.

Middlebrooks and Bremen (2013) studied spatial stream segregation in anesthetized cat cortex. They presented trains of noise bursts which were similar to the stimuli presented to these animals in the streaming task. Middlebrooks and Bremen found that most units could synchronize to noise bursts presented at 2.5 and 5 bursts per second (bps) but not many units could synchronize to stimuli that were presented at 10 bps. In the current study neurons were found to synchronize their responses up to the fastest rates presented, which was 40 s^{-1}. The median synchronization boundary was 33.1 s^{-1} which is faster than the rates tested in

anesthetized cats by Middlebrooks and Bremen. Chapter 4 further explores auditory streaming and the ability of cortical neurons to synchronize their responses to competing sound sources in the awake cat.

In response to clicks, the median synchronization boundary was ~40 s^{-1} which was much higher than the 10-25 s^{-1} rates at which sustained responses are lost in anesthetized conditions. Eggermont (1991) studied synchronization of click trains in anesthetized cat primary auditory cortex. For that study, neurons were able to synchronize their responses up to 8 s^{-1} but responses started to fall off at around 16 s^{-1}. Eggermont measured the vector strength for obtaining the limiting rates to which a neuron could synchronize its responses. Based on this measurement, half of the neurons synchronized to rates of ~24 s^{-1}. Similarly, Lu and Wang (2000) observed that neurons in anesthetized cat cortex had a median synchronization boundary of ~25 s^{-1}. Contrary to anesthetized preparations, neurons in auditory cortex of awake cats (Dong et al., 2011) and marmosets (Lu et al., 2001) exhibit response synchrony to higher rates >40 s^{-1} which is similar to our findings in this study.

Interestingly, we also observed units that exhibited tonic asynchronous responses at higher click rates >50 s^{-1}. In addition, the spike rate for these neurons was found to increase with increasing click rates. These responses generally peaked at click rates around 160 s^{-1} and declined at higher rates. This decline differs from what is seen in awake marmoset in which spike rates were shown to increase monotonically to click rates of 333 s^{-1}, the highest rates tested (Lu et al., 2001).

To characterize between synchronous and non-synchronized units in awake marmosets, Lu et al. (2001), looked at the DRR. In their case, the DRR was defined as the maximum spike rates to stimulus rates greater than 200 s^{-1} (we used 160 s^{-1}) divided by the maximum spike rates

to stimulus rates lower than 33 s^{-1}. They found that synchronized units had high Raleigh statistics and low discharge rate ratios that were <1. Non-synchronized units on the other hand, had low Raleigh statistics and high discharge rate ratios. We calculated our DDR values differently in that we subtracted background rates and took the maximum spike rates to stimulus rates greater than 160 s^{-1} divided by the maximum spike rate. Measuring our DRR in this way allowed us to look at responses that exhibited suppression at high presentation rates which were often below that of background activity. Unlike the Lu et al. study, neurons in the current study were not classified predominantly as exhibiting synchronous and nonsynchronous responses. One neuron for example exhibited synchrony at low presentation rates and nonsynchronous activity at high presentation rates. A few neurons also showed nonsynchronous responses at low presentation rates.

Conclusions. The results demonstrate that the spectral and temporal coding properties previously seen in awake marmosets are also present in awake cats. In comparison between cats and marmosets, cats have a wider audibility range (63 Hz- 64 kHz) compared to that of marmosets (125 Hz- 36 kHz) (Heffner and Heffner, 1985; Osmanski and Wang, 2011). Although cats have a larger audibility range, the cat vocal repertoire is not as complex as that of marmosets. The vocal repertoire of marmosets for example, includes twitter calls that consist of repetitive phrases at rates of 6-9 s^{-1} (Wang et al., 1995; Nagarajan et al., 2001). Each phrase contains 2-3 frequency modulation sweeps at different frequencies ranging from 7-8 kHz (Wang et al., 1995; Nagarajan et al., 2001). On the other hand, the cat vocal repertoire consists of at least 6 call types encompassing a broad frequency range (i.e., 1-4 kHz) with fundamental frequencies ranging from 0.5 to 1 kHz (Moelk, 1944; Brown et al., 1978; Farley et al., 1992; Gehr, 2000; Qin et al., 2008). Their vocalizations lack the repetitive structure of marmoset

twitter calls and the amplitude envelope and frequency modulation changes in cat vocalizations are relatively slow. Although these animals differ in the complexity of their vocal repertoires, cortical neurons in both animals exhibit similar spectral and temporal response properties to sound stimuli suggesting that these properties are present only in the awake cortex. To summarize, in regards to spectral coding properties, I- and O- shaped FRAs are observed in the absence of anesthesia. Measurements in frequency tuning show narrower BW in awake compared to anesthetized preparations. Moreover, neurons are found to respond in a sustained manner rather than eliciting only transient onset responses. Neurons also show phase locking to rates >40 s^{-1} which is faster than what is seen under anesthesia. Lastly, neurons are shown to respond in an asynchronous tonic fashion to rates >50 s^{-1} with response firing increasing with increasing presentation rates.

3.6 Acknowledgements

The work is supported by National Institute on Deafness and Other Communication Disorders Grant R01 DC000420 (to JCM) and the Graduate Assistance in Areas of National Need Award P200A120165 (to LKJ). We thank Elizabeth McGuire, Justin Yao, Akshat Arneja, and Zekiye Onsan for technical and administrative assistance.

CHAPTER 4

Spatial Stream Segregation in the Awake Cat Auditory Cortex

4.1 Summary

In a complex acoustic environment stream segregation permits a listener to disentangle multiple competing sequences of sounds. Recent studies have demonstrated correlates of spatial stream segregation in primary auditory cortex in which neurons preferentially synchronize to one of two competing interleaved sound sequences. It has been suggested that the mechanism underlying spatial stream segregation is enhanced by the limited rate in which neurons phase lock to stimuli. This is referred to as forward suppression. We demonstrated in Chapter 3 that neurons in the awake cat cortex could synchronize to noise bursts at rates up to 40 s^{-1} which is higher than what is observed under anesthesia. The ability of neurons in awake cortex to synchronize to faster rates may result in a lack of stream segregation due to the absence of forward suppression seen under anesthesia. We tested the hypothesis that in the awake cortex, cats not engaged in an auditory task, will show minimal segregation as a result of a lack of forward suppression. We recorded from auditory cortex in cats implanted with chronic recording arrays. Interestingly, our findings show that neurons synchronized to both competing streams that were presented at a base rate of 10 s^{-1} and that these neurons showed some preferentiality to synchronize more strongly to one sound source. This preference however, was prominently weaker compared to a previous study in anesthetized cats. It may be that the representation of spatial stream segregation is enhanced when an animal is selectively attending to a target source.

4.2 Introduction

In complex auditory environments listeners can disentangle interleaved sequences of sounds from two or more sources, a phenomenon referred to as "stream segregation" (Bregman, 1990). The difference in the locations of competing sound sources has been shown to be a robust factor in enabling stream segregation. Under anesthesia, neurons in primary auditory cortex of cats (Middlebrooks and Bremen, 2013) and rats (Yao et al., 2015), show spatial stream segregation that reflects what is seen in behavior (Middlebrooks and Onsan, 2012; Javier et al., 2016; Chapter 2). Under conditions in which listeners perceive a single source (i.e., competing sounds are presented at the same location), neurons will non-preferentially synchronize their responses to both competing sources. If, however, the two sources are separated in space such that human listeners perceive segregation, neurons will preferentially synchronize their responses to one of two alternating noise burst sequences.

The mechanism underlying spatial stream segregation in anesthetized cortex has been suggested to be enhanced by forward suppression, in that cortical neurons are limited in the rate at which they could synchronize their responses. When sounds are co-located for example, there is a decrease in neural response due to the inability of neurons to synchronize to the aggregate rate of both sources. When both sources are spatially separated however, the neural response is captured by one source, which is presented at half the base rate of when the sources are co-located, while suppressing the response to the other sound source (i.e., stream segregation) (Middlebrooks and Bremen, 2013; Yao et al., 2015).

In Chapter 3, we have shown that when cats were presented with trains of noise bursts from a single sound source and not engaged in an auditory task, many neurons synchronized their responses up to 40 s^{-1}, the fastest rate tested. The purpose of this study was to investigate if

neurons in the awake cortex could segregate between competing noise bursts even though these neurons have previously been shown to synchronize to rates higher than the burst rate of the streaming stimuli. Streaming stimuli consisted of alternating 20-ms noise bursts. The combined presentation rate of both competing sound sequences was 10 s^{-1}. We recorded from primary auditory cortex in awake cats implanted with chronic recording arrays. We tested the hypothesis that neurons in the auditory cortex of awake cats not overtly engaged in an auditory task would show minimal, if any, stream segregation due to their ability to synchronize to faster stimuli rates. We found that although neurons synchronized to both competing streams, their responses were preferentially stronger to one source over the other. This preferential response however, was not strong enough to elicit stream segregation based on our measurement criteria of d'=1. We also introduce examples of units from cats performing a spatial stream segregation task but due to the scarcity of units, we did not do further data analysis. The results suggest that when an animal is not engaged in an auditory task, neurons exhibit some sensitivity to preferentially synchronize to one sound source. When an animal is on-task and selectively attending to a target sequence, neural responses may be enhanced such that neurons selectively synchronize to the target sequence while suppressing responses to the competing masker sequence, enabling stream segregation. This sharpening of response to the target sequence could then be further processed in higher cortical areas.

4.3 Materials and Methods

All procedures were in accordance with the NIH Animal Welfare Guidelines and with a protocol approved by the Institutional Animal Care and Use Committee at the University of California at Irvine. For most of this study, we recorded from awake cats that were not engaged

in an auditory task, which we will refer to as the "non-task" condition. The same animals that were used in Chapter 3 (seven males, one female) were also used for this study in the non-task condition. For three of those animals, we also collected neurophysiological data while those cats were performing a stream segregation task as described in Chapter 2. We will refer to this condition as "on-task." Surgical procedures to implant chronic recording electrodes are identical to what was described in Chapter 3 and will not be repeated below. For days in which we collected neurophysiological data for the non-task and on-task conditions, food was restricted. Cats received moistened cat food in the non-task condition as they sat calmly in the sound booth, and as reinforcement while they were performing the stream segregation task (on-task condition). On other days in which we did not collect neurophysiological data, cat had free access to dry food for 3 hours a day. Water was freely available in the housing area.

Experimental apparatus. Experiments were conducted in a double-walled sound-attenuating anechoic chamber (Industrial Acoustics; inside dimensions 2.6 x 2.6 x 2.5 m) lined with SONEXone absorbent foam to suppress sound reflections. The chamber contained 13, 8.4-cm-diameter two-way loudspeakers positioned on a horizontal circular hoop, 1.2 m in radius, at azimuths of 0 and ±5, 10, 20, 40, 60, 80° relative to the front of the apparatus. Left and right speaker locations are referred to as contralateral (C) and ipsilateral (I), respectively, relative to the electrode placement which was in the cat's right hemisphere. The cat was supported on a raised platform, which was adjusted in height so that the cat's head was centered in the array of loud speakers. A harness restrained the animal to the platform but permitted free movement of the limbs and head.

Stimulus generation. For the most part, stimulus generation was similar to what is described in Chapter 2. Rather than looking at different passband conditions however, we

focused on only the broadband condition with band-pass noise bursts filtered at 0.4-25 kHz. Noise bursts were 20 ms in duration and was gated with raised cosine functions with 1-ms rise and fall times. Masker noise bursts were interleaved with target noise burst sequences. The masker sequences were complementary to the target sequences, where when co-located, the stimulus was a continuous sequence of undifferentiated noise bursts with aggregate rates of 10 bps. Stimuli were presented at 60 dB SPL. In the non-task condition, stimuli were presented as competing interleaved A and B noise bursts in an alternating ABAB…. pattern where the A source was always presented at 0° in location and the B source varied in location at ± 60, 40, 20, 10, and 0° for each trial. These streaming sequences were repeated 20 times for each stimulus condition.

The on-task condition is the same as what was described in Chapter 2. Target sequences were presented with two temporal patterns, Rhythm 1, AA--AA--AA--, and Rhythm 2, AA--A-A-A-A-, where the dashes represent a gap in which a complementary masker would be presented. Each trial began with the presentation of one to four continuous sequences of Rhythm 1, with the number of Rhythm 1 repetitions varying randomly from trial to trial. This was followed by a presentation of Rhythm 2 which was repeated 1.5 times.

Experimental procedure. The experimental procedure consisted of two conditions: non-task and on-task. The experimental procedure for the non-task condition is similar to what is described in Chapter 3. For this condition, the pneumatic feeder system and response pedal used in the on-task condition was disabled. A harness was used to restrain the animal on the platform but the animal had free movement of limbs and head. During recordings, an experimenter remained in the sound chamber with the animal, offering moistened cat food to keep the animal

calm and quiet and also, to orient the animal's head toward the front of the loudspeaker array (~0° azimuth). Recording sessions lasted for approximately 30 minutes to an hour.

For the on-task condition, cats performed the spatial stream segregation task as described in Chapter 2. The only differences were that cats only performed the task using the broadband condition, and we had a cable attached to the connector of the electrode to record neurophysiology while the cat was performing the task.

Data Analysis. Stimulus-evoked neuronal action potentials were identified off-line with a spike sorting procedure that is described in Middlebrooks (2008) and also, in Chapter 3. Statistical analysis includes the combined single- and multiple-unit responses. To look at auditory streaming in the on-task condition, we only looked at response synchrony to Rhythm 1. We did not look at synchrony to Rhythm 2 because the sound presentation stopped shortly after the presentation of Rhythm 2 which is when the cat detected a change in the temporal pattern. In the non-task condition, we recorded from 86 units. Of those 86 units, 4 were single units and the rest were multiple units. Although responses showed phase-locking to the A and B stimuli, 38 of the 86 units showed slightly stronger responses to one source location. We only had a few units from the on-task condition, and all units were multiple-units. In the on-task condition, only 10 of the 20 units showing auditory evoked response showed stronger responses to either the A or B location. Due to insufficient data for the on-task condition, we focused our data analysis to primarily to the non-task condition.

We measured rate-azimuth functions (RAFs) for competing-sound stimuli similar to Middlebrooks and Bremen (2013) and Yao et al., (2015). RAFs were expressed as the mean number of spikes synchronized to each 20-ms noise burst as a function of loudspeaker location. Spikes tended to fall within ~50 ms after each noise-burst onset. To capture spikes driven by

each noise burst, we counted spikes in the 8-58 ms after stimulus onset. Given this, spikes were attributed to either A or B sources.

We used a procedure based on signal detection theory (Green and Swets, 1966; Macmillan and Creelman, 2005) to measure the discrimination of sound-source locations by trial-by-trial neural spike counts. Similar to Middlebrooks and Bremen (2013) and Yao et al., (2015), we compiled spike counts for all repetitions that were synchronized to the A and B sources. We formed an empirical ROC curve that was based on the trial-by-trial distributions of spike counts elicited on all trials by the two stimuli. The area under the ROC curve gave the probability of correct discrimination of the stimuli. This probability was expressed as a *z*-score and was multiplied by √2, yielding the discrimination index, *d'*. If 100% of the spike rates elicited by one stimulus was greater than any of those elicited by the other stimulus, the area under the ROC curve was 1.0 and the corresponding *z*-score was undefined. Under these conditions, d' was written as ±2.77, corresponding to the 97.5% correct discrimination. The interpolated separation at which the plot of *d'* versus azimuth crossed *d'*= ±1 was taken as the threshold. Positive *d'* values denotes that there were more spikes synchronized to the more-contralateral sound source.

4.4 Results

We presented cats that were not participating in an auditory task with a train of noise bursts in an alternating ABAB stimulus pattern. The A source (i.e., target) was always located at 0° azimuth and the competing, masker source (B-source) varied in location from trial to trial. Synchronized mean spike counts were quantified for the full range of the A and B location combinations. Example multiple-unit rate azimuth functions (RAFs) are shown in Figure 4.1.

The abscissa indicates the B-source location relative to the placement of the recording electrode which was always located in the cat's right hemisphere. The vertical dashed red line indicates the A-source location. Spikes synchronized to the A-source (fixed location; red lines) and B-source (blue lines) varied in response to the location of the B-source. The spike rates per trial are indicated by the left-axis. In some example units, when the A and B sources were separated in azimuth, neurons would preferentially synchronize to either source A or source B. In 4.1A for example, when source B was located contralateral to the recording site, neurons synchronized more strongly to source B. 4.1C shows the opposite response in which neurons preferentially synchronized to source A when source B is located in the contralateral hemifield. On the other hand, in 4.1B, if source B were moved ipsilateral to the recording site, neurons preferentially synchronized to source A. Figure 4.1D shows an example unit in which the neuron synchronized predominantly to source A except for when source B was located at the two furthest contralateral locations. To measure the accuracy to which spike counts discriminated between A and B sources, we quantified a discrimination index, *d'* as indicated by the right y-axis. We took a $d' = \pm 1$ as the threshold for significant discrimination of A and B sources, thus indicating significant stream segregation. The solid black line indicates the discrimination index. The horizontal dashed black lines indicate $d'=0$. Although neurons were shown to preferentially synchronize to one source over the other, the *d'* values often only reached threshold performance at the largest A- and B-source separations. Examples of this are seen in Figure 4.1A and C. Figure 4.1C was rare in that threshold was reached at a ~17° sound source separation.

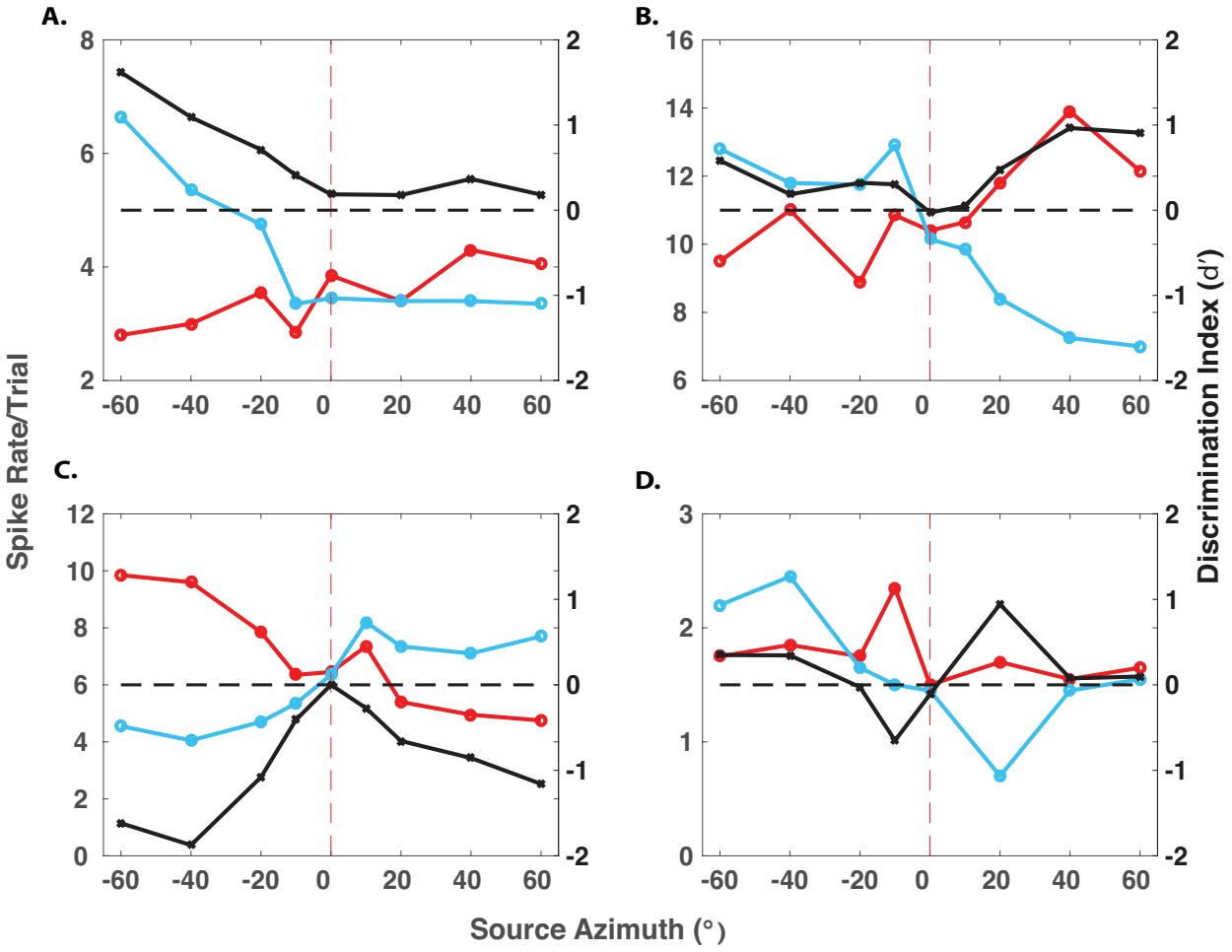

Fig 4.1. Responses to competing source conditions. RAFs showing mean spike rates (left y-axis) as a function of B source location. Red and blue lines indicate responses synchronized to the A and B sources, respectively. B-source locations were as plotted. A-source locations were fixed in the 0° location indicated by the vertical dashed red line. The variation in the responses to the A-source locations reflects the effect of the changing B-source locations. The solid black line indicates the discrimination index (d') of trial-by-trial spike rates elicited by A versus B sources. The dashed black lines indicate *d'*=0.

We plotted histograms of the discrimination index of trial-by-trial spike counts that were synchronized to the A and B sources (Figure 4.2). In these plots, the A-source was always fixed at 0° azimuth and the B source was located either 20° or 60° contralateral (left column, A and C) and ipsilateral (right column, B and D) to the recording site. We chose to look at the 20° B source location for comparison to a previous anesthetized study in cats (Middlebrooks and Bremen, 2013) who looked at stream segregation at this spatial separation. We also looked at

the condition in which the A and B sources were separated by 60° since units tended to exhibit significant stream segregation at the largest sound source separations. The horizontal dashed black lines represent threshold at $d' = \pm 1$. Blue horizontal thick and thin lines indicate the interquartile ranges and 10^{th} to 90^{th} percentile ranges of each distribution, respectively. The Xs indicate medians. The magenta horizontal lines and Xs in Figure 4.2A and B are data from anesthetized cats (Middlebrooks and Bremen, 2013) with the same notation as previously described for the blue horizontal lines. Positive *d'* values indicate units that are more strongly synchronized to the more contralateral sound source. If the B-source is located in the contralateral hemifield (left column) for example, positive values indicate stronger responses to the B-source, while negative *d'* values indicates stronger responses to the A-source. Conversely, if the B-source is located in the ipsilateral hemifield (right column), positive values indicate stronger responses to the A-source and negative *d'* values indicate stronger responses to the B-source location. As shown below, only a few units exhibited stream segregation compared to previous anesthetized cat data. For those few units that showed segregation, most of those units preferentially responded to the A-source (i.e., target) location. Median values and the median absolute deviations are shown for each B location.

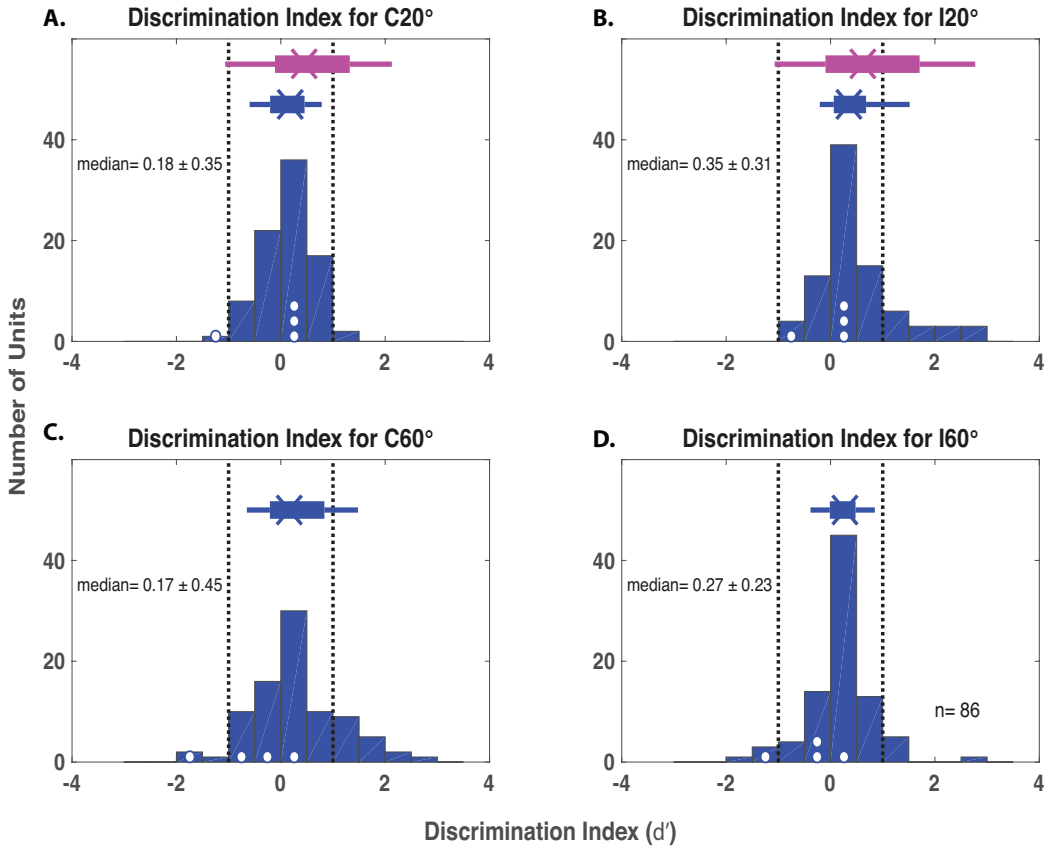

Fig 4.2. Distributions of d' for discrimination of A and B sources. A sources are fixed at 0° and B sources were located either contralateral (A and C) or ipsilateral (B and D) to the recording site. Horizontal thick and thin lines indicate interquartile ranges and 10th and 90th percentile ranges, respectively. The Xs on those lines indicate medians. The blue horizontal lines pertain to the current study and the magenta lines are from anesthetized cat data (Middlebrooks and Bremen, 2013). Positive d' values indicate greater spike rates elicited by the more contralateral sound source. Medians and median absolute deviations are shown in each panel.

We recorded from 20 units while the animals were performing a spatial stream segregation task. Of those 20, only 10 units showed some sensitivity to preferentially synchronize to one of the two competing sound sources. RAFs of multiple units, similarly shown in Figure 4.1, are shown in Figure 4.3. Figure 4.3A shows an example of stronger responses to the B source when the B source is located in the contralateral hemifield and stronger responses to the A source when the B source is located in the ipsilateral hemifield. In Figure 4.3B, this multiple unit preferentially responded to the A source regardless of the B source

location. Most source A and B separations had few trials, with some conditions only having one trial. Due to having few units and few trials for each source A and B separation, we did not conduct further analysis on units recorded in the on-task condition.

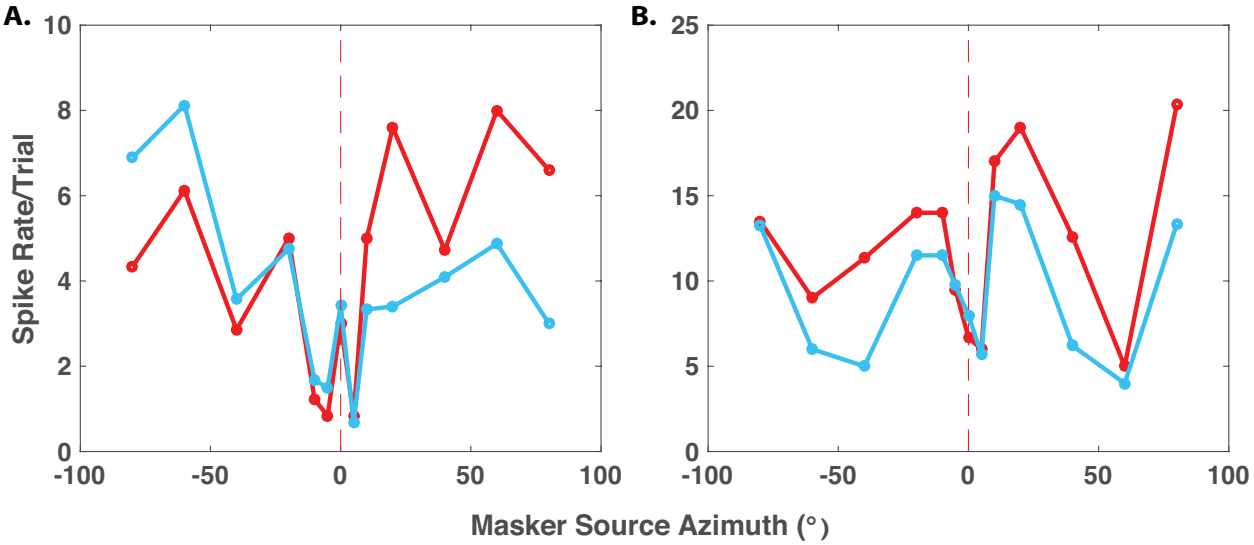

Fig 4.3. RAFs for on-task conditions showing mean spike rates as a function of B source location. Red and blue lines indicate responses synchronized to the A and B sources, respectively. B-source locations were as plotted. A-source locations were fixed in the 0° location indicated by the vertical dashed red line. The variation in the responses to the A-source locations reflects the effect of the changing B-source locations.

4.5 Discussion

Under anesthesia, cortical neurons in A1 are able to segregate between two interleaved competing sound sources (Middlebrooks and Bremen, 2013; Yao et al., 2015). In anesthetized cat recordings for example (Middlebrooks and Bremen, 2013), many cortical neurons showed little sensitivity to the locations of stimuli presented from a single sound source. These neurons responded equally to these sound sequences regardless of where the sound source was located. When a competing sound source was added, neurons preferentially synchronized to one of the two competing sound sources resulting with $d'>1$. Thresholds for segregation was seen by the

minimum separation of competing sound sources tested, which was 10°. These results agree with cat psychophysical experiments in which cats engaged in a spatial stream segregation task could segregate between two sources with separation as little as 9.4° (Javier et al., 2016; Chapter 3).

The robust ability for neurons in anesthetized cortex to segregate between otherwise identical noise-bursts that differed only in their location, was not replicated in the present study. In Chapter 3 we found that neurons could synchronize to noise bursts from a single location at rates up to 40 s^{-1}. In the non-task condition of the current study, when neurons were presented with competing noise bursts that varied in location, neurons synchronized to both the target and masker but showed some sensitivity to preferentially synchronize to one sound source. This however, produced a weak stream segregation effect in that discrimination indices rarely reached threshold and this effect is weaker than what is observed in anesthetized conditions (Middlebrooks and Bremen, 2013; Yao et al, 2015).

Although results for the awake, non-task condition differed from previous anesthetized cat recordings (Middlebrooks and Bremen, 2013), both experimental setups were fairly similar. Both studies utilized alternating noise bursts in which a target sound (source A) was always presented at a fixed location while a second competing masker source (source B) varied in location along azimuth. Similar to our study, Middlebrooks and Bremen presented a condition in which sequences of alternating noise bursts for sources A and B were presented at a base rate of 10 s^{-1}. The current study differed from that of Middlebrooks and Bremen in the duration of individual noise bursts and also, the manner in which the cat was positioned in the sound chamber for recordings. For the current study, the stimuli we used was identical to stimuli used by these animals when they were engaged in an auditory task (i.e., on-task condition) (Javier et

al., 2016; Chapter 2), which consisted of noise bursts that were 20 ms in duration. This differs from Middlebrooks and Bremen in that they used 5 ms noise bursts. We chose to use 20 ms noise bursts for better comparison for when animals were on-task vs. when they were not engaged in a task. It is possible that the duration of the noise bursts contributes to the differences in results, however, under on-task conditions, cats showed behaviorally that they were able to segregate between noise bursts that were 20 ms in duration. Also, studies that looked at stimulus duration as a factor in stream segregation found that longer tone durations more strongly facilitated segregation compared to shorter tone durations (van Noorden, 1975; Beauvois 1998). Even at shorter noise burst durations, Middlebrooks and Bremen (2013) still observed robust stream segregation by cortical neurons.

The current study also differed from Middlebrooks and Bremen in that under anesthesia, cats' heads were fixed in position, whereas in the current study, we had some imprecision in head position in that cats had free movement of their head and limbs. An experimenter offered pureed cat food in order for the cat to sit calmly and also, to orient the cat's head towards the front of the array of loudspeakers. Having the cats eat while recording resulted in background activity that presumably includes sounds generated by the cat's licking and other self-generated noises. For most of the duration of the recording sessions, cats eagerly ate the offered cat food and while it is unclear whether cats paid attention to the noise bursts stimuli, we do see auditory evoked responses to the stimuli. It is possible that having the cats attentively focused on the cat food caused some suppressive effects of any stream segregation in our neural recordings. We did attempt to abstain from offering cat food during recording sessions, but cats refused to sit calmly during those sessions and we could not obtain any reliable recordings. Nevertheless, we

did observe some sensitivity for stream segregation. We think that the lack of robust segregation is due to the lack of attention to the sound stimuli.

The mechanism underlying spatial stream segregation in the anesthetized cortex is suggested to be enhanced by forward suppression. It reflects the inability of cortical neurons to synchronize to faster rates. When for example, both competing sources were presented from the same location, there was a decrease in response synchrony at this higher presentation rate. This is illustrated in anesthetized cat cortex in which there was a decrease in neuronal synchrony to noise bursts presented at 5 and 10 s^{-1} (Middlebrooks and Bremen, 2013). In conditions in which competing sound sources were separated along azimuth, neurons were able to synchronize their responses, but only to one sound source since the presentation rate at a single location was lower than the presentation rate of when both sounds were co-located. Yao and colleagues (2015) quantified the amount of forward suppression in areas of the inferior colliculus (ICC), medial geniculate body (MGB) and A1 of anesthetized rats. In all areas studied, forward suppression was greatest among neurons in A1, in that the ability of neurons to synchronize to trains of noise bursts decreased at increasing presentation rates. In addition, A1 neurons exhibited a sharpening of spatial tuning that could contribute to forward suppression. Yao et al., also investigated whether forward suppression was due to synaptic inhibition within the cortex, or whether it could be attributed to some other mechanism in A1 that limits neurons from synchronizing to higher rates. To study this, they applied GABA antagonists while recoding from neurons in A1. They reasoned that if forward suppression was due to synaptic inhibition, inhibiting GABA receptors would remove cortical inhibition and thus, lead to an increase in the neurons ability to synchronize to higher rates. Administering GABA antagonists did not alleviate forward suppression, suggesting that some other biophysical property prevented neural synchrony. The

authors hypothesize that forward suppression involved in spatial stream segregation is due to synaptic depression at the thalamo-cortical synapse.

Other studies using awake preparations in macaques and European starlings have observed stream segregation in cortical neurons of animals not engaged in any auditory tasks (Fishman et al., 2001; Fishman and Steinschneider, 2004; Bee and Klump, 2004). These studies however, were not looking at the spatial separation of sound sources that enable segregation, but rather, segregation due to frequency separation. Fishman et al., (2001) for example, presented awake macaques with alternating ABAB… streams in which A and B sources differed in frequency. One sound source consisted of tones at the neuron's best frequency (BF) (i.e., A source) while the other sound source consisted of tones not at BF (i.e., B source). When either the A- or B-sources were presented alone, neurons synchronized similarly to both sources. When they were presented in an alternating manner, however, cortical neurons were shown to synchronize to one sound source under conditions in which the frequency separation was large, or if the presentation rate was high. If both A and B-sources had a small frequency separation for example, the neuron responded to both competing sources equally. When the frequency separation was increased, responses to the B tones decreased while synchrony to that of the A source was high. In regards to presentation rate, segregation was seen at the fastest presentation rates of 20 and 40 s^{-1}, but not at the lowest presentation rates of 5-10 s^{-1}. They also found that forward suppression was evident regardless of which sound source started the sequence.

The fact that tone-based segregation was not seen at presentation rates of 5-10 s^{-1} coincides with our study in that we do not see strong spatial stream segregation with our stimuli which are presented at a base rate of 10 s^{-1}. When animals are engaged in an auditory task (Javier et al., 2016; Chapter 2), cats have been shown to spatially segregate between interleaved

noise bursts at base rates of 10 s^{-1} which agrees with anesthetized recordings in cats (Middlebrooks and Bremen, 2013). Since cats show effective stream segregation at these low rates, it raises the question of why we do not see spatial stream segregation in the awake primary auditory cortex. It may be that 1) when an animal is on-task, selective attention enhances the response to the target source or, 2) that the processing of stream segregation in higher cortical areas is more selective than lower cortical areas in its representation of the target sound source.

A previous study, not involving stream segregation, has shown that neuronal selectivity was enhanced when an animal was on-task versus when the animal was off-task. Lee and Middlebrooks (2011, 2013) trained cats to depress a pedal which initiated a presentation of successive broadband noise bursts from varying locations along azimuth. When the cats detected a change in the target, either by a change in elevation or a change from noise bursts to a click train, they released the pedal to receive a food reward. Compared to conditions in which cats were not engaged in the task to when cats were on-task, they found a sharpening of spatial sensitivity when cats were performing the task. The largest effect of neural sensitivity was when cats had to evaluate the location of each stimulus to detect when the target changed in elevation. They observed a suppression of sounds to non-favored locations. The sharpening of tuning to favored source locations may be applicable to the current study in that selective attention may enhance sensitivity to one sound source while suppressing responses to a competing source, enabling stream segregation.

Although selective attention may enhance the representation of spatial stream segregation in A1, the increased selectivity towards one sound source may be processed in higher cortical areas. Studies of human neurophysiological recordings have demonstrated streaming through enhanced synchrony of neural activity to one of two competing speech streams that the listener

was selectively attending to (Ding and Simon, 2012; Mesgarani and Chang, 2012; Golumbic, 2013). Ding and Simon for example, used MEG recordings to study stream segregation. Human listeners were presented narratives of two talkers and they were instructed to attend to one talker over the other. The speech envelopes of both talkers were reconstructed from neural activity by optimization of different MEG sensors across time. They looked at the neural reconstructions of the attended speech and unattended speech and compared that to the envelopes of those speeches. The neural reconstruction of the attended speech was found to be strongly correlated with the envelope of the attended speech. Conversely, there was a weak correlation between the reconstruction of the unattended speech and the attended speech. This modulation by selective attention was localized to the planum temporale, which is outside of primary auditory cortex (i.e., Heschl's gyrus). This is further supported by work done by Golumbic et al., (2013) who found that in low level cortices (i.e., superior temporal gyrus), attention modulates the neuronal representation by enhancing cortical tracking of attended speech streams, although ignored speech is still represented. In higher level cortical areas, the representation becomes more selective towards the attended speech with no detectable tracking of the unattended speech. Interestingly, they also found that brain activity tracked speech streams using low-frequency phase (delta 1-3 Hz and theta 4-7 Hz) fluctuations. In the current study, the presentation rate of individual sound sources of 5 s^{-1} falls into the range of low-frequency brain activity of theta oscillations. It is possible that when an animal is on-task and selectively attending to one sound source, neurons will entrain to these low frequency oscillations whereby the representation of the attended stream will be enhanced. Conversely, the ignored stream will not entrain to these neural oscillations resulting in a suppressed representation of this stream. Based on studies looking at higher cortical areas, it may be that A1 integrates information from areas in the

brainstem and below with some influence by attention. From there, it distributes that information to higher auditory cortical areas for further analysis and further selectivity towards the attended target. This increase in selectivity for the target stream is enhanced by brain activity that tracks the target stream using low-frequency phase fluctuations. In our study, we see representation for both sources with some sensitivity towards the target sound source. Under on-task conditions, attention may modulate the neural representation to strongly favor the target and suppress the response to the masker, exhibiting robust stream segregation. Furthermore, this representation could be sharpened in higher cortical areas.

Overall, this study shows minimal stream segregation when a cat is not engaged in an auditory task. Further studies are needed to parse out the mechanisms underlying spatial stream segregation in primary awake cortex, particularly when a cat is engaged in an auditory task. It might be that spatial stream segregation in on-task conditions will: 1) show a sharpening of spatial tuning in A1 as demonstrated by Lee and Middlebrooks (2011, 2013) and 2) that forward suppression and the entrainment to low frequency oscillations is enhanced, thereby exhibiting robust stream segregation.

4.6 Acknowledgements

The work is supported by National Institute on Deafness and Other Communication Disorders Grant R01 DC000420 (to JCM) and the Graduate Assistance in Areas of National Need Award P200A120165 (to LKJ). We thank Elizabeth McGuire, Justin Yao, Pauline Herbert-Whiting, and Zekiye Onsan for technical and administrative assistance.

CHAPTER 5

Summary and Future Directions

5.1 Summary

In complex auditory scenes, listeners have the exceptional ability of disentangling multiple interleaved sequences of sound and segregating that mixture into distinct sound sources, a phenomenon referred to as "stream segregation." The physical separation of competing sound sources has been shown to be a robust factor in facilitating stream segregation, whereby human listeners can segregate between competing streams with a spatial separation of as little as ~8° (Middlebrooks and Onsan, 2012). Furthermore, neural correlates of spatial stream segregation have been observed in area A1 in anesthetized cats and rats (Middlebrooks and Bremen, 2013; Yao et al., 2015). The goal for this dissertation is to evaluate the psychophysics of spatial stream segregation and within the same species, to conduct neurophysiological recordings in auditory cortex in the absence of anesthesia. Cats are well suited for these experiments as they have been used extensively for sound-localization experiments. In addition, their well-developed cortex is easily accessible on the surface of the brain allowing for neurophysiological recordings. In Chapter 2, I show that cats can segregate between competing streams when sound sources are separated by ~9.4°. I also find that cats rely predominantly on high-frequency interaural level difference cues for stream segregation which is contrary to human listeners who rely more on interaural time difference cues. The superior use of high frequencies by cats, however, agrees with physiological results in anesthetized cats showing stronger stream segregation among cortical neurons having high CFs. In Chapter 3, I observe cortical units in A1 of awake cats that

exhibit spectral and temporal properties that are rarely seen in anesthetized cortex. These properties include frequency response areas that are restricted in frequency and level (i.e., O-shaped units), sharp level-invariant frequency tuning bandwidths, sustained response firing, neurons that can synchronize to clicks and noise bursts at rates >40 s^{-1}, and tonic asynchronous firing that increases with stimulus rates. These findings support the hypothesis that these spectral and temporal response properties are present in the normal, awake auditory cortex. Lastly, in Chapter 4, I find that cortical units do not segregate between competing streams when cats are not engaged in an auditory task, possibly due to a lack of selective attention. Overall, this dissertation highlights cat psychophysics and neural properties in the absence of anesthesia that will pave the way for future studies combining psychophysics and neurophysiology within the same listener.

5.2 Future Directions

More study is needed to investigate neural activity when an animal is engaged in a spatial stream segregation task. This dissertation provides a glimpse of a few cortical recordings while cats are on-task, however, it is insufficient in providing any conclusions on task-dependent modulation of cortical units. A previous study by Lee and Middlebrooks (2011) demonstrated that the spatial sensitivity of neurons was dependent on the behavioral state of the animal where spatial tuning was found to sharpen in A1 when the animal was engaged in a localization task compared to when it was off-task. It may be that on-task conditions will show an improvement of discrimination by neuronal firing rates while the cat is on-task, whereby neurons will preferentially synchronize to the target while suppressing responses to the masker sound source. In addition, while a cat is on-task, further study is needed to evaluate the magnitude of neural

discrimination on individual trials in which the cat does or does not successfully detect the rhythm change and to test whether the neural responses predict psychophysical performance on a trial-by-trial basis.

In addition, all single- and multi-unit studies of cortical mechanisms of stream segregation have been limited to the primary auditory cortex. Further study is needed to investigate processing of spatial stream segregation in cortical structures outside of A1. Human neuroimaging studies suggest that structures outside of A1 are also involved in auditory scene analysis (Gutschalk et al., 2005; Micheyl et al., 2007; Wilson et al., 2007; Ding and Simon, 2012; Mesgarani and Chang, 2012; Golumbic, 2013). Overall, studies outside of area A1 suggest that the role of primary cortex might be to integrate information from the brainstem and below. Information is then further distributed to higher cortical areas for further analysis.

Lastly, future studies could investigate whether the mechanisms underlying spatial stream segregation are the same as those underlying tone-based segregation. The task would be similar to the spatial stream segregation task in Chapter 2, however the competing streams will differ in frequency instead of by location. Previous studies in awake macaques not engaged in an auditory task (Fishman et al., 2001) show tone-based segregation at presentation rates >20 s^{-1}. They did not see segregation at the rate in which our stimuli were presented. Neurophysiological recordings of on-task cats could be compared between spatial-based and tone-based segregation to see if there are any similarities between the two types of segregation. In addition, performance using high-frequency tones versus low-frequency tones in the tone-based segregation task could be compared to the different passband conditions in the spatial segregation task.

Taken together, this dissertation and future work provides insight into the auditory mechanisms underlying spatial stream segregation in cat psychophysics as well as a basis for streaming in awake cortex.

REFERENCES

Arbogast TL, Mason CR, Kidd G Jr. (2002) The effect of spatial separation of informational and energetic masking of speech. *Journal of the Acoustical Society of America* 112: 2086-2098

Bartlett EL, Sadagopan S, Wang (2011) Fine frequency tuning in monkey auditory cortex and thalamus. *Journal of Neurophysiology* 106: 849-859

Beauvois MW (1998) The effect of tone duration on auditory stream formation. *Perception & Psychophysics* 60: 852-861

Bee MA, Klump GM (2004) Primitive auditory stream segregation: A neurophysiological study in the songbird forebrain. *Journal of Neurophysiology* 92: 1088-1104

Bee MA, Micheyl C (2008) The cocktail party problem: what is it? How can it be solved? And why should animal behaviorists study it? *Journal of Comparative Psychology* 122: 235-251

Boehnke SE, Phillips, DP (2005) The relation between auditory temporal interval processing and sequential stream segregation examined with laterality differences. *Perception and Psychophysics* 67: 1088-1101

Bregman AS (1990) Auditory Scene Analysis: The Perceptual Organization of Sound. Cambridge, MA: *MIT Press*

Broadbent DE, Ladefoged (1957) On the fusion of sounds reaching different sense organs. *Journal of the Acoustical Society of America* 29: 708-710

Brughera, A, Dunai, L Hartmann, WM (2013) Human interaural time difference thresholds for sine tones: the high-frequency limit. *Journal of the Acoustical Society of America* 133: 2839-2855

Brungart DS (2005) Informational and Energetic Masking Effects in Multitalker Speech Perception. *In Speech Separation by Humans and Machines* (pp. 261-267). Boston: *Kluwer Academic*

Calford MB (1983) The parcellation of the medial geniculate body of the cat defined by the auditory response properties of single units. *Journal of Neuroscience* 124: 535-547

Calford MB, Pettigrew JD (1984) Frequency dependence of directional amplification at the cat's pinna. *Hearing Research* 14: 13-19

Casseday JH, Neff WD (1973) Localization of pure tones. *Journal of the Acoustical Society of America* 54: 356-372

Cherry CE (1953) Some experiments on the recognition of speech, with one and two ears. *Journal of the Acoustical Society of America* 25: 975-979

Chimoto S, Kitama T, Qin L, Sakayori S, Sato Y (2002) Tonal response patterns of primary auditory cortex neurons in alert cats. *Brain Research* 934: 34-42

Cranford JL (1979) Auditory cortex lesions and interaural intensity and phase-angle discrimination in cats. *Journal of Neurophysiology* 42: 1518-1526

Cutting JE (1976) Auditory and linguistic processes in speech perception: inferences from six fusions in dichotic listening. *Psychological Review* 2:114-140

Darwin, CJ (1997) Auditory grouping. *Trends in Cognitive Sciences* 1: 327-333

Ding N, Simon JZ (2012) Emergence of neural encoding of auditory objects while listening to competing speakers. *Proceedings of the National Academy of Sciences* 109: 11854-11859

Dong C, Qin L, Liu Y, Zhang X, Sato Y (2011) Neural responses in the primary auditory cortex of freely behaving cats while discriminating fast and slow click-trains. *PLoS ONE* 6: 1-11

Eggermont JJ (1991) Rate and synchronization measures of periodicity coding in cat primary auditory cortex. *Hearing Research* 56:153-167

Eramudugolla, R, Irvine DRF, McAnally KI, Martin RL, Mattingley JB (2005) Directed attention eliminates 'change deafness' in complex auditory scenes. *Current Biology* 15: 1108-1113

Evans EF, Whitfield IC (1964) Classification of unit responses in the auditory cortex of the unanesthetized and unrestrained cat. *Journal of Physiology* 171: 476-493

Farley GR, Barlow SM, Netsell R, Chmelka JV (1992) Vocalizations in the cat: behavioral methodology and spectrographic analysis. *Experimental Brain Research* 89: 333-340

Fishman YI, Arezzo JC, Steinschneider M (2004) Auditory stream segregation in monkey auditory cortex: effects of frequency separation, presentation rate, and tone duration. *Journal of the Acoustical Society of America* 116: 1656-1670

Fishman YI, Reser DH, Arezzo JC, Steinschneider M (2001) Neural correlates of auditory stream segregation in primary auditory cortex of the awake monkey. *Hearing Research* 151: 167-187

Fishman YI, Steinschneider M (2010) Formation of Auditory Streams. Oxford, NY: *Oxford University Press.* pp 215-245

Fujita I, Konishi M (1991) The role of GABAergic inhibition processing of interaural time differences in the owl's auditory system. *Journal of Neuroscience* 11: 722-739

Gatehouse S and Noble W (2004) The speech, spatial and qualities of hearing scale (SSQ). *International Journal of Audiology* 43: 85-99

Gehr DD, Komiya H, Eggermont JJ (2000) Neural responses in cat primary auditory cortex to natural and altered species-specific calls. *Hearing Research* 150:27-42

Goldberg JM, Brown PB (1969) Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. *Journal of Neurophysiology* 32: 613-636

Golumbic EMZ, Ding N, Bickel S, Lakatos P, Schevon CA, McKhann GM, Goodman RR, Emerson R, Mehta AD, Simon JZ, Poeppel D, Schroeder CE (2013) Mechanisms underlying selective neuronal tracking of attended speech at a "cocktail party." *Neuron* 77: 980-991

Green DM, Swets, JA (1966) Signal Detection Theory and Psychophysics. New York, NY: *Wiley*

Gutschalk A, Micheyl C, Melcher JR, Rupp A, Scherg M, and Oxenham AJ (2005) Neuromagnetic correlates of streaming in human auditory cortex. *The Journal of Neuroscience* 25: 5382-5388

Harrington IA, Stecker GC, Macpherson EA, Middlebrooks JC (2008) Spatial sensitivity of neurons in the anterior, posterior, and primary fields of cat auditory cortex. *Hearing Research* 240: 22-41

Hartmann WM, Johnson D (1991) Stream segregation and peripheral channeling. *Music Perception* 9: 155-184

Heffner HS, Heffner HE (1985) Hearing range of the domestic cat. *Hearing Research* 19:85-88

Heil P, Rajan R, Irvine DR (1992) Sensitivity of neurons in cat primary auditory cortex to tones and frequency-modulated stimuli. I: Effects of variation of stimulus parameters. *Hearing Research* 63: 108-134

Ihlefeld A, Shinn-Cunningham B (2008) Disentangling the effects of spatial cues on selection and formation of auditory objects. *Journal of the Acoustical Society of America* 124: 2224-2235

Iverson P (1995) Auditory stream segregation by musical timbre: effects of static and dynamic acoustic attributes. *Journal of Experimental Psychology: Human Perception and Performance* 21: 751-763

Javier LK, McGuire EA, Middlebrooks JC (2016) Spatial stream segregation by cats. *Journal of the Association for Research in Otolaryngology 17: 195-207*

Jenkins WM, Masterton RB (1982) Sound localization: effects of unilateral lesions in central auditory system. *Journal of Neurophysiology* 47: 987-1016

Jenkins WM, Merzenich MM (1984) Role of cat primary auditory cortex for sound-localization behavior. *Journal of Neurophysiology* 52: 819-847

Kidd G Jr., Mason CR, Deliwala PS, Woods WS, Colburn HS (1994) Reduced information masking by sound segregation. *Journal of the Acoustical Society of America* 95: 3475-3480

Kidd G Jr., Mason CR, Rohtla RL, Deliwala PS (1998) Release from masking due to spatial separation of sources in the identification of nonspeech auditory patterns. *Journal of the Acoustical Society of America*. 104: 422-431

Klump RG, Eady HR (1956) Some measurements of interaural time difference thresholds. *Journal of the Acoustical Society of America* 28: 859-860

Kuhn GF (1977) Model for the interaural time differences in the azimuthal plane. *Journal of the Acoustical Society of America* 62: 157-167

Langner G (1992) Periodicity coding in the auditory system. *Hearing Research* 60: 115-142

Langner G, Schreiner CE (1988) Periodicity coding in the inferior colliculus of the cat. I. neuronal mechanisms. *Journal of Neurophysiology* 60: 1799-1822

Lee CC, Middlebrooks JC (2011) Auditory cortex spatial sensitivity sharpens during task performance. Nature Neuroscience 14: 108-114

Lee CC, Middlebrooks JC (2013) Specialization for sound localization in fields A1, DZ, and PAF of cat auditory cortex. *Journal of the Association for Research in Otolaryngology* 14: 61-82

Lomber SG, Malhotra S (2008) Double dissociation of "what" and "where" processing in auditory cortex. *Nature Neuroscience* 11: 609-616

Lu T, Wang X (2000) Temporal discharge patterns evoked by rapid sequences of wide- and narrowband clicks in the primary auditory cortex of cat. *Journal of Neurophysiology* 84: 236-24

Lu T, Liang L, Wang X (2001) Temporal and rate representations of time-varying signals in auditory cortex of awake primates. *Nature* 4: 1131-1138

Macmillan NA, Creelman CD (2005) Detection theory: A users guide (2nd ed). New York: *Cambridge University Press*

MacMillan, NA, Kaplan, HL (1985) Detection theory: A user's guide. Mahwah, NJ: *Psychology Press*

Macpherson EA, Middlebrooks JC (2002) Listener weighting cues for lateral angle: The duplex theory of sound localization revisited. *Journal of the Acoustical Society of America* 111: 2219-2236

Malhotra S, Hall AJ, Lomber SG (2004) Cortical control of sound localization in the cat: unilateral cooling deactivation of 19 cerebral areas. *Journal of Neurophysiology* 92: 1625-1643

Malhotra S, Stecker GC, Middlebrooks JC, Lomber SG (2008) Sound localization deficits during reversible deactivation of primary auditory cortex and/or the dorsal zone. *Journal of Neurophysiology* 99:1628-1642

Mardia KV (1972) Statistics of Directional Data. New York: *Academic Press*

Martin RL, Webster WR (1987) The auditory spatial acuity of the domestic cat in the interaural horizontal and median vertical planes. *Hearing Research* 30: 239-252

May BJ, Huang AY, Aleszczyk CM, Hienz RD (1995) Design and Conduct of Sensory Experiments for Domestic Cats. Methods in Comparative Psychoacoustics. Edited by Klump GM, Basel: *Birkhauser Verlag*

McAlpine D, Jiang D, Shackleton TM, Palmer AR (1998) Convergent input from brainstem coincidence detectors onto delay-sensitive neurons in the inferior colliculus. *Journal of Neuroscience* 18: 6026-6039

McLaughlin M, Franken TP, van der Heijden M, Joris PX (2014) The interaural time difference pathway: a comparison of spectral bandwidth and correlation sensitivity at three anatomical levels. *Journal for the Association for Research in Otolaryngology* 15: 203-18

Mesgarani N, Chang EF (2012) Selective cortical representation of attended speaker in multi-talker speech perception. *Nature* 485: 233-236

Micheyl C, Carylon RP, Gutschalk A, Melcher JR, Oxenham AJ, Rauschecker JP, Tian B, Wilson EC (2007) The role of auditory cortex in the formation of auditory streams. *Hearing Research* 229: 116-131

Middlebrooks JC (2008) Auditory cortex phase locking to amplitude-modulated cochlear implant pulse trains. *Journal of Neurophysiology* 100: 76-91

Middlebrooks JC, Bremen P (2013) Spatial stream segregation by auditory cortical neurons. *Journal of Neuroscience* 33: 10986-11001

Middlebrooks JC, Clock AE, Xu L, Green DM (1994) A panoramic code for sound location by cortical neurons. *Science* 264: 842-844

Middlebrooks JC, Dykes RW, Merzenich MM (1980) Binaural response-specific bands in the primary auditory cortex (A1) of the cat: topographical organization orthogonal to isofrequency contours. *Brain Research* 181: 31-48

Middlebrooks JC, Green DM (1991) Sound localization by human listeners. *Annual Review of Psychology* 42: 135-159

Middlebrooks JC, Knudson EL (1987) Changes in external ear position modify the spatial tuning of auditory units in the cat's superior colliculus. *Journal of Neurophysiology* 57: 672-687

Middlebrooks JC, Makous JC, Green DM (1989) Directional sensitivity of sound-pressure levels in the human ear canal. *Journal of the Acoustical Society of America* 86: 89-108

Middlebrooks JC, Nick HS, Subramony SH, Advincula J, Rosales RL, Lee LV, Ashizawa T, Waters MF (2013) Mutations in the Kv3.3 voltage-gated potassium channel causing spinocerebellar ataxia 13 disrupts sound-location mechanisms. *PLoS One* 8: e76749

Middlebrooks JC, Onsan ZA (2012) Stream segregation with high spatial acuity. *Journal of the Acoustical Society of America* 132: 3896-3911

Middlebrooks JC, Pettigrew JD (1981) Functional classes of neurons in primary auditory cortex of the cat distinguished by sensitivity to sound location. *Journal of Neuroscience* 1: 107-120

Middlebrooks JC, Snyder RL (2010) Selective electrical stimulation of the auditory nerve activates a pathway specialized for high temporal acuity. *Journal of Neuroscience* 30: 1937-1946

Middlebrooks JC, Xu L, Eddins AC, Green DM (1998) Codes for sound-source location in nontonotopic auditory cortex. *Journal of Neurophysiology* 80: 863-881

Middlebrooks JC, Xu L, Furukawa S, Macpherson EA (2002) Cortical neurons that localize sounds. *The Neuroscientist* 8: 73-83

Middlebrooks JC, Zook JM (1983) Intrinsic organization of the cat's medial geniculate body identified by projections to binaural response-specific bands in the primary auditory cortex. *Journal of Neuroscience* 3: 203-224

Moelk M (1944) Vocalizing in the house cat; A phonetic and functional study. *Journal of Psychology* 57: 184-205

Moore BC, Gockel H (2002) Factors influencing sequential stream segregation. *Acta Acustica* 88: 320-332

Moore BJC, Gockel HE (2012) Properties of auditory stream formation. *Philosophical Transactions of The Royal Society* 367: 919-931

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. *Journal of Neurophysiology* 96: 3756-3769

Musicant AD, Chan JC, Hind JE (1990) Direction-dependent spectral properties of cat external ear: new data and cross-species comparisons. *Journal of the Acoustical Society of America* 87: 757-781

Nagarajan SS, Cheung SW, Bedenbaugh P, Beitel RE, Schreiner CE, Merzenich MM (2001) Representation of spectral and temporal envelope of twitter vocalizations in common marmoset primary auditory cortex. *Journal of Neurophysiology* 87: 1723-1737

Neff WD, Hind JE (1955) Auditory thresholds of the cat. *Journal of the Acoustical Society of America* 27: 480-483

Osmanski MS, Wang X (2011) Measurement of absolute auditory thresholds in the common marmoset (*Callithrix jacchus*). *Hearing Research* 277: 127-133

Phillips DP (1993) Neural representation of stimulus times in the primary auditory cortex. *Annals of the New York Academy of Sciences* 682: 104-118

Phillips DP, Calford MB, Pettigrew JD, Aitkin LM, Semple MN (1982) Directionality of sound pressure transformation at the cat's pinna. *Hearing Research* 8:13-28

Proctor L, Konishi M (1997) Representation of sound localization cues in the auditory thalamus of the barn owl. *Proceedings of the National Academy of Sciences* 94: 10421- 10425

Qin L, Wang JY, Sato Y (2008) Representation of cat meows and human vowels in the primary auditory cortex of awake cats. *Journal of Neurophysiology* 99: 2305-2319

Ramachandran R, Davis KA, May BJ (1999) Single-unit responses in the inferior colliculus of decerebrate cats I. Classification based on frequency response maps. *Journal of Neurophysiology* 82: 152-163

Rayleigh L, (1907) XII. On our perception of sound direction. *Philosophical Magazine* 13: 214-32

Roberts B, Glasberg BR, More BCJ (2002) Primitive stream segregation of tone sequences without differences in fundamental frequency or passband. *Journal of the Acoustical Society of America* 112: 2074- 2085

Roth GL, Kochhar RK, Hind JE (1980) Interuaral time differences: implications regarding the neurophysiology of sound localization. *Journal of the Acoustical Society of America* 68: 1643-1651

Rouiller E, de Ribaupierre Y, Toros-Morel A, de Ribaupierre F(1981) Neural coding of repetitive clicks in the medial geniculate body of cat. *Hearing Research* 5: 81-100

Sach AJ, Bailey PJ (2004) Some characteristics of auditory spatial attention revealed using rhythmic masking release. *Perception & Psychophysics* 66: 1379-1387

Sadagopan S, Wang X (2008) Level invariant representation of sounds by populations of neurons in primary auditory cortex. *Journal of Neuroscience* 28: 3415-3426

Schreiner CE, Sutter ML (1992) Topography of excitatory bandwidth in cat primary auditory cortex: single-neuron versus multiple-neuron recordings. *Journal of Neurophysiology* 68: 1487-1502

Schreiner CE, Urbas JV (1988) Representation of amplitude modulation in the auditory cortex of the cat. II. Comparison between cortical fields. *Hearing Research* 32: 49-64

Schwartz A, McDermott JH, Shinn-Cunningham B (2012) Spatial cues alone produce inaccurate sound segregation: The effect of interaural time differences. *Journal of the Acoustical Society of America* 132: 357-368

Shinn-Cunningham BC (2005) Influences of spatial cues on grouping and understanding sound. *In Proceedings of Forum Acusticum*, Budapest, Hungary

Singh PG, Bregman AS (1997) The influence of different timbre attributes on the perceptual segregation of complex-tone sequences. *Journal of the Acoustical Society of America* 102: 1943- 1952

Spitzer MW, Semple MN (1995) Neurons sensitive to interaural phase disparity in gerbil superior olive: diverse monaural and temporal response properties. *Journal of Neurophysiology* 73: 1668-1689

Stainsby TH, Fullgrabe C, Flanagan HJ, Waldman SK, Moore BCJ (2011) Sequential streaming due to manipulation of interaural time differences. *Journal of the Acoustical Society of America* 130: 904-914

Stecker GC, Harrington IA, Macpherson EA, Middlebrooks JC (2005) Spatial sensitivity in the dorsal zone (area DZ) of cat auditory cortex. *Journal of Neurophysiology* 94: 1267-1280

Stecker GC, Mickey BJ, Macpherson EA, Middlebrooks JC (2003) Spatial sensitivity in field PAF of cat auditory cortex. *Journal of Neurophysiology* 89: 2889-2903

Stecker GC, Middlebrooks JC (2003) Distributed coding of sound locations in the auditory cortex. *Biolgoical Cybernetics* 89: 341-349

Sutter ML, Schreiner CE (1991) Physiology and topography of neurons with multipeaked tuning curves in cat primary auditory cortex. *Journal of Neurophysiology* 65: 1207-1226

Tollin DJ, Koka K (2009) Postnatal development of sound pressure transformations by the head and pinnae of the cat: Monaural characteristics. *Journal of the Acoustical Society of America* 125: 980-994

Tollin DJ, Yin TCT (2002) The coding of spatial location by single units in the lateral superior olive of the cat. I. Spatial receptive fields in azimuth. *Journal of Neuroscience* 22: 1454-1467

Tollin DJ, Yin TCT (2005) Interaural phase and level difference sensitivity in low-frequency neurons in the lateral superior olive. *Journal of Neuroscience* 26: 10648-10657

van Noorden LPAS (1975) Temporal coherence in the perception of tone sequences. *Ph.D. Diss., Tech Hogeschool, Eindhoven, The Netherlands, Pub. by The Institute of Perception Research, Eindhoven, The Netherlands*

Vliegan J, Oxenham AJ (1999) Sequential stream segregation in the absence of spectral cues. *Journal of the Acoustical Society of America* 105: 339-346

Wakeford, OS, Robinson, DE (1974) Lateralization of tonal stimuli by the cat. *Journal of the Acoustical Society of America* 55: 649-652

Wang X, Merzenich MM, Beitel R, Schreiner CE (1995) Representation of species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. *Journal of Neurophysiology* 74:2685-2706

Wang X, Lu T, Snider RK, Liang L (2005) Sustained firing in auditory cortex evoked by preferred stimuli. *Nature* 435: 341-346

Wessel DL (1979) Timbre space as a musical control structure. *Computer Music Journal* 3: 45-52

Wilson EC, Melcher JR, Micheyl C, Gutschalk A, Oxenham AJ (2007) Cortical fMRI activation to sequences of tones alternating in frequency: Relationship to perceived rate and streaming. *Journal of Neurophysiology* 97: 2230- 2238

Wright BA, Fitzgerald MB (2001) Different patterns of human discrimination learning for two interaural cues to sound-source location. *PNAS* 98: 12307-12312

Yao JD, Bremen P, Middlebrooks JC (2015) Emergence of spatial stream segregation in the ascending auditory pathway. *The Journal of Neuroscience* 35: 16199-16212

Young ED, Brownell WE (1976) Responses to tones and noise of single cells in dorsal cochlear nucleus of unanesthetized cats. *Journal of Neurophysiology* 39: 282-300

Young ED, Rice JJ, Tong SC (1996) Effects of pinna position on head-related transfer functions in the cat. *Journal of the Acoustical Society of America* 99: 3064-3076

Zhou B, Green DM, Middlebrooks JC (1992) Characterization of external ear impulse responses using Golay codes. *Journal of the Acoustical Society of America* 92: 1169-1171

Zwislocki J, Feldman RS (1956) Just noticeable differences in dichotic phase. *Journal of the Acoustical Society of America* 28: 860-864