Copyright Warning & Restrictions

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material.

Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be "used for any purpose other than private study, scholarship, or research." If a, user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use" that user may be liable for copyright infringement,

This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve violation of copyright law.

Please Note: The author retains the copyright while the New Jersey Institute of Technology reserves the right to distribute this thesis or dissertation

Printing note: If you do not wish to print this page, then select "Pages from: first page # to: last page #" on the print dialog screen

NJIT New Jersey's Science & Technology University

The Van Houten library has removed some of the personal information and all signatures from the approval page and biographical sketches of theses and dissertations in order to protect the identity of NJIT graduates and faculty.

ABSTRACT

TOWARDS UNDERSTANDING THE ROLE OF CENTRAL PROCESSING IN RELEASE FROM MASKING

by
Nima Alamatsaz

People with normal hearing have the ability to listen to a desired target sound while filtering out unwanted sounds in the background. However, most patients with hearing impairment struggle in noisy environments, a perceptual deficit which current hearing aids and cochlear implants cannot resolve. Even though peripheral dysfunction of the ears undoubtedly contribute to this deficit, surmounting evidence has implicated central processing in the inability to detect sounds in background noise. Therefore, it is essential to better understand the underlying neural mechanisms by which target sounds are dissociated from competing maskers. This research focuses on two phenomena that help suppress background sounds: 1) dip-listening, and 2) directional hearing.

When background noise fluctuates slowly over time, both humans and animals can listen in the dips of the noise envelope to detect target sound, a phenomenon referred to as dip-listening. Detection of target sound is facilitated by a central neuronal mechanism called envelope locking suppression. At both positive and negative signal-to-noise ratios (SNRs), the presence of target energy can suppress the strength by which neurons in auditory cortex track background sound, at least in anesthetized animals. However, in humans and animals, most of the perceptual advantage gained by listening in the dips of fluctuating noise emerges when a target is softer than the background sound. This raises the possibility that SNR shapes the reliance on different processing strategies, a hypothesis tested here in awake behaving animals. Neural activity of Mongolian gerbils is measured by chronic implantation of silicon probes in the core auditory cortex. Using appetitive conditioning, gerbils

detect target tones in the presence of temporally fluctuating amplitude-modulated background noise, called *masker*. Using rate- vs. timing-based decoding strategies, analysis of single-unit activity show that both mechanisms can be used for detecting tones at positive SNR. However, only temporal decoding provides an SNR-invariant readout strategy that is viable at both positive and negative SNRs.

In addition to dip-listening, spatial cues can facilitate the dissociation of target sounds from background noise. Specifically, an important cue for computing sound direction is the time difference in arrival of acoustic energy reaching each ear, called interaural time difference (ITD). ITDs allow localization of low frequency sounds from left to right inside the listener's head, also called sound lateralization. Models of sound localization commonly assume that sound lateralization from interaural time differences is level invariant. Here, two prevalent theories of sound localization are observed to make opposing predictions. The labelled-line model encodes location through tuned representations of spatial location and predicts that perceived direction is level invariant. In contrast, the hemispheric-difference model encodes location through spike-rate and predicts that perceived direction becomes medially biased at low sound levels. In this research, through behavioral experiments on sound lateralization, the computation of sound location with ITDs is tested. Four groups of normally hearing listeners lateralize sounds based on ITDs as a function of sound intensity, exposure hemisphere, and stimulus history. Stimuli consists of low-frequency band-limited white noise. Statistical analysis, which partial out overall differences between listeners, is inconsistent with the place-coding scheme of sound localization, and supports the hypothesis that human sound localization is instead encoded through a population rate-code.

TOWARDS UNDERSTANDING THE ROLE OF CENTRAL PROCESSING IN RELEASE FROM MASKING

by
Nima Alamatsaz

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology and
Rutgers Biomedical and Health Sciences–Newark
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Biomedical Engineering

August 2021

Copyright © 2021 by Nima Alamatsaz

ALL RIGHTS RESERVED

APPROVAL PAGE

TOWARDS UNDERSTANDING THE ROLE OF CENTRAL PROCESSING IN RELEASE FROM MASKING

Nima Alamatsaz

Dr. Antje Ihlefeld, Dissertation Advisor — Date
Assistant Professor of Biomedical Engineering, New Jersey Institute of Technology

---

Dr. Sergei Adamovich, Committee Member — Date
Professor of Biomedical Engineering, New Jersey Institute of Technology

---

Dr. Mesut Sahin, Committee Member — Date
Professor of Biomedical Engineering, New Jersey Institute of Technology

---

Dr. Robin L. Davis, Committee Member — Date
Professor of Cell Biology and Neuroscience, Rutgers University–New Brunswick

---

Dr. Catherine E. Myers, Committee Member — Date
Professor of Pharmacology, Physiology and Neuroscience,
Rutgers Biomedical and Health Sciences–Newark

BIOGRAPHICAL SKETCH

Author: Nima Alamatsaz

Degree: Doctor of Philosophy

Date: August 2021

Undergraduate and Graduate Education:

- Doctor of Philosophy in Biomedical Engineering,
  New Jersey Institute of Technology & Rutgers Biomedical and Health Sciences, Newark, NJ, 2021

- Bachelor of Science in Electrical Engineering,
  University of Isfahan, Isfahan, Iran, 2013

Major: Biomedical Engineering

Presentations and Publications:

Min Zhang, Nima Alamatsaz, and Antje Ihlefeld, "Hemodynamic responses link individual differences in informational masking to the vicinity of superior temporal gyrus." *Frontiers in Neuroscience*, 15:677, 2021.

Nima Alamatsaz and Antje Ihlefeld, "Teaching electronic circuit fundamentals via remote laboratory curriculum." *Biomedical engineering education*, 1(1):105-109, 2021.

Nima Alamatsaz and Antje Ihlefeld, "Remote laboratory curriculum for electrical fundamentals of biomedical engineering." *Zenodo*, 2020.

Nima Alamatsaz and Antje Ihlefeld, "EARS: Electrophysiology auditory recording system (v1.0)." *Zenodo*, 2019.

Antje Ihlefeld, Nima Alamatsaz, and Robert M Shapley, "Population rate-coding predicts correctly that human sound localization depends on sound intensity." *eLife*, 8:e47027, 2019.

Nima Alamatsaz and Antje Ihlefeld, "Leveraging adaptation to study perceptual weighting of interaural time differences." *International Congress on Acoustics (ICA): Integrating 4th EAA Euroregio*, 23(1):8253-8256, 2019.

Antje Ihlefeld, Nima Alamatsaz, and Robert M Shapley, "Circling back on theories of sound localization." *The Journal of the Acoustical Society of America*, 145(3):1759, 2019.

Nima Alamatsaz and Antje Ihlefeld, "The role of central processing in modulation masking release." *The Journal of the Acoustical Society of America*, 144(3):1900, 2018.

Antje Ihlefeld, Nima Alamatsaz, and Robert M Shapley, "Human sound localization depends on sound Intensity: implications for sensory coding." *bioRxiv*, 378505, 2018.

Antje Ihlefeld, Matthew Ning, Sahil S Chaubal, and Nima Alamatsaz, "Neural correlates of modulation masking release: the role of sound deprivation." *The Journal of the Acoustical Society of America*, 141(5):3894, 2017.

Nima Alamatsaz, Antje Ihlefeld, and Robert M Shapley, "The effect of sound intensity on lateralization with interaural time differences." *The Journal of the Acoustical Society of America*, 141(5):3639, 2017.

I would like to dedicate this dissertation to my dearest parents, ***Dr. Mohammad Hossein Alamatsaz*** *and* ***Mrs. Nahid Meamaripour****, whom I can't begin to describe how utterly unable would I have been to take even the smallest of steps without them always being at my side, while being physically far away.*

ACKNOWLEDGMENT

It is a genuine pleasure to express my gratitude to my dissertation adviser, **Dr. Antje Ihlefeld**, for her all her teachings, constant help and support throughout this long journey.

I owe a deep sense of gratitude to my dissertation committee, **Drs. Sergei Adamovich**, **Mesut Sahin**, **Robin Davis**, and **Catherine Myers**, for having had the opportunity to benefit from their classes, learn from their lectures, present my progress to them, and for providing me with their valuable inputs on every step of the way.

I hereby thank National Institutes of Health (NIH) and New Jersey Health Foundation (NJHF) for supporting this research by funds from NIH-NIDCD R01 DC014008-01, NIH-NIDCD R01 DC019126-01, and NJHF PC 24-18.

I would also like to thank **Drs. Dan Sanes**, **Brad Buran**, **Todd Mowery**, **Merri Rosen**, **Justin Yao**, **Kristina Penikis**, and **Pelin Avcu** for many helpful comments and extensive guidance on the data acquisition and interpretation.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

Chapter **Page**

TABLE OF CONTENTS
(Continued)

Chapter **Page**

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES
(Continued)

Figure **Page**

LIST OF FIGURES
(Continued)

Figure **Page**

LIST OF FIGURES
(Continued)

Figure **Page**

LIST OF FIGURES
(Continued)

Figure **Page**

LIST OF FIGURES
(Continued)

Figure **Page**

CHAPTER 1

INTRODUCTION

A major hurdle of patients with conductive hearing-loss (CHL) is the ability to understand speech in noisy environments. Many studies have attempted to characterize this deficiency and associate it with different regions along the auditory pathway, more specifically the periphery auditory system [75, 85], and the central nervous system (CNS) processing [31, 53, 57], however, the full extent of influence of CHL on CNS and its underlying neural mechanisms is still under investigation. A necessary first step for understanding the limitations in the hearing impaired is to study normal auditory function. Normal-hearing listeners (NH) utilize a combination of mechanisms to resolve target sounds in presence of a background noise (masker). Here we focus on two of these mechanisms: a specific case of energetic masking referred to as modulation masking release, and an attentional cue called spatial release from masking.

1.1 Modulation Masking Release

It is well documented that for tone detection in background noise, normally-hearing (NH) listeners have better behavioral thresholds when that noise is temporally modulated as compared to being stationary, a perceptual phenomenon referred to as modulation masking release (MMR). However, hearing impaired listeners often do not show a dramatic difference in performance across these two tasks. Behavioral evidence from Mongolian gerbils (Meriones unguiculatus) with conductive hearing loss (CHL) supports the idea that sound deprivation alone can reduce MMR. MMR cannot be alone explained by the mechanical processes of the ear [127]. Due to similarity of low-frequency hearing range of gerbils to that of humans [125], a series of electrophysiological investigations were undertaken on gerbils that lead to

detection of neural correlates of MMR in inferior colliculus, medial geniculate body and primary auditory cortex A1 [73]. Moreover, a recently published behavioral study has demonstrated that gerbils exposed to early onset conductive hearing loss have reduced MMR activity [59].

In previous studies, it has been reported that attention modulates neural activity of various regions along the auditory pathway [16]. Hence, to measure neural correlates of MMR in A1, the subject needs to be actively engaged in a behavioral task. Moreover, several acoustic limitations dictate the head of the subject to be fixated in a predetermined position to be able to reliably present sound stimuli. By bearing these considerations in mind, an experimental paradigm has to behaviorally engage gerbils to respond to acoustic stimuli as the neural activity is recorded via an implanted microelectrode array.

1.2 Interaural Time Differences

In a different mechanism called spatial release from masking, it has been observed that spatial segregation of the sources of competing sounds improves detection thresholds [77]. A wide range of species uses the time delay in propagation of sound waves from one ear to the other (referred to as interaural time difference or ITD) as a cue for localizing the direction of a sound source in the azimuthal plane, a phenomenon called sound lateralization. The specific neural mechanisms by which humans can accomplish this feat are still incompletely understood. Two competing computational models, called the labelled-line model [65, 29] and the hemispheric-difference model [126, 82], can predict aspects of human lateralization performance. These two models differ in the assumptions they make about the underlying neural code [58].

In the mammalian ascending auditory pathway, the first neural processing stage where ITDs are encoded, on the timescale of microseconds, is the medial superior olive (MSO). Here, temporally precise binaural inputs converge, causing changes

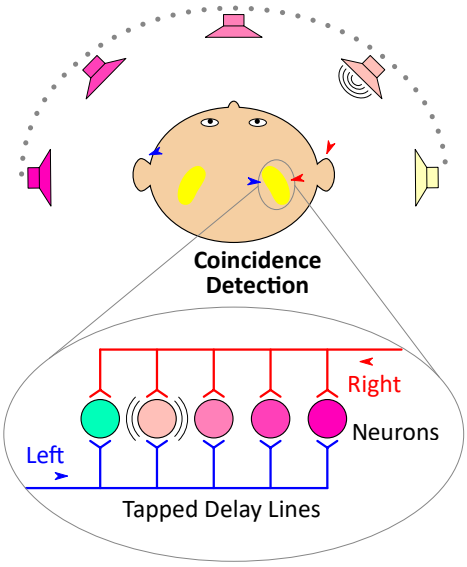

Figure 1.1 Coincidence detection model for processing of interaural time differences. The construct of tapped delay lines renders each neuron sensitive to a specific time delay between the right and left auditory signals.

in MSO output firing rate as a function of ITD [36, 136, 119, 25] with ITD-rate output functions resembling a band-limited cross-correlation of the acoustic inputs to each ear [6]. A prominent neural model for sound localization, originally proposed by Jeffress, consists of neurons that are sensitive to the binaural synchronicity of neural inputs from each ear [65]. In this model, coincidence detector neurons receive inputs from both ears through labelled delay lines, making each neuron maximally sensitive to a specific magnitude of ITD (Figure 1.1). This model is computationally equivalent to calculating the normalized cross-correlation of sounds reaching both ears and versions of it have successfully been applied in many engineering applications predicting human localization performance and/or human ability to ignore competing background masking sound [29, 22, 46, 122, 13].

Several studies support the existence of Jeffress-like neural place code mechanisms in the avian brain [18, 92]. However, a growing literature indicates that the Jeffress model cannot fully explain mammalian sensitivity to ITD. Neurons in the mammalian MSO are assumed to receive two excitatory input from the ipsilateral and contralateral

auditory pathways, each with a specific delay as a result of different axonal properties (such as their length, thickness and also myelination patterns) of their post-synaptic neurons originating from the anterio-ventral cochlear nucleus (AVCN) [44]. Aside from these excitatory pathways, a set of inhibitory neurons from the medial nucleus of the trapezoid body (MNTB) also form synaptic connections to the MSO neurons. Evidence from Mongolian gerbils shows that precise glycinergic inhibitory inputs play a crucial role in processing ITDs [12, 95], however their role has been controversial [43]. Using a population rate code and incorporating these precise inhibitory pathways, a non-normalized coincidence detection model of rate-based ITD coding in the MSO proposes that sound localization can be modeled by calculating differences in firing rates across hemispheres [82]. This model generally predicts that the locus of the highest information content of neuronal responses is at the steepest slopes of rate-ITD curves, where firing rates change most strongly, consistent with the observation that the peak ITDs of rate-ITD curves often fall outside the physiologically plausible range [82, 44]. Further support of a rate-based neural population code stems from the findings that mammalian sound localization can adapt to stimulus history [99], consistent with a disynaptic feedback loop through the MSO and the superior periolivary nucleus that normalizes the MSO output gain [120].

The labelled-line model detects temporal coincidence of the signals reaching both ears and relies on place-coding: neurons are maximally sensitive to a best ITD; whereas the hemispheric-difference model compares firing rates across hemispheres and favors a population spike-rate neural code. Both types of model can predict human sound localization across a range of conditions. They both predict that the ability to lateralize sound should degrade with decreasing sound intensity, whereas their predictions differ qualitatively. The Jeffress-type labelled-line model predicts that neural discharge will become increasingly less sound-evoked and thus more random as sound intensity decreases. As a result, this model predicts that listeners

will on average hear sounds from similar locations at low versus high sound levels, except that the response variability should increase with decreasing sound levels. In contrast, rate-based models, including hemispheric-difference models, predict that as sound intensity decreases, firing rates will decrease. Because perceived laterality in this family of models is computed via the difference across rate-ITD functions, this means that the maximal discharge rate will be below peak value, causing the location percept to become increasingly medially biased.

Table 1.1 Conflicting Evidence in the Literature Regarding the Effect of Sound Level on ITD-based Lateralization

Study	Stimulus	Duration [ms]	Level [dB SPL]	Hemisphere	N	Bias
Bekesy 1960	Click	n/a	10, 30, 40, 50	n/a	10	Most lateral, few none
Guttman 1962	Click: 0.06-13 kHz	0.1	16,20, 36	n/a	3	Medial
Teas 1962	Click: high-pitched	0.5	10, 20, 30, 40	Right	2	Medial
Mickunas 1963	Click	0.2	30, 60	Right	3	None
Babkoff 1966	Click	0.5	15, 25, 35, 45	Bilateral	2	Medial
Sabin et al. 2006	Noise: 0.5-16 kHz	250	0, 5, 10, 15, 20, 30, 40, 50, 60	Bilateral	6	Medial

In the literature there has been insufficient evidence on the effect of stimulus level on ITD-invoked lateralization (Table 1.1). In an experiment performed on two subjects, Teas [124] claims as sounds become softer, there is a significant tendency in perceiving the sound direction to be more medial (towards the nose line) as opposed to the louder sounds being heard more lateral. Babkoff et al. [4] also measured the Hornbostel-Wertheimer constant on two subjects that yielded the same medial bias at 10 dB SL sounds. A conflicting experiment undertaken by Bekesy [129] demonstrated a bias towards the opposite directions (i.e., lateral side) with sounds at near sensation threshold. Meanwhile, Mickunas [84] did not observe any level effects after testing three subjects on relatively loud 30 and 60 dB SL stimuli. All of the mentioned studies were undertaken on limited number of subjects and the results are not considered to

be statistically significant. An argument that can be made about the later experiment is that the stimuli were not as close to the sensation threshold as to invoke any level dependent effects. A more recent study reported a similar bias at 5 dB SL in a sound localization task performed on ten blind folded subjects while presented with stimuli through an array of loud speakers [111].

1.3 Structure of this Dissertation

In Chapter 2, a closed-loop system is developed for synchronized control of animal behavior and recording neuronal electrophysiology. Gerbils are tested in a behavioral paradigm that trains them to perform a nose poke and wait for a target sound. When and if the target sound is presented, the gerbil can advance towards a lick spout and receive water as reward. In cases where the target sound is not presented, the gerbil should withdraw from the nose poke and make another attempt. As the behavioral task in is progress, electrophysiological activity of the auditory cortex will be recorded using a chronically implanted microelectrode array. The recorded signal is further processed, visualized and stored for offline analysis. Chapter 3 presents the results of testing gerbils with target sounds that have lower or higher energy than the modulated background sound. The hypothesis that signal-to-noise ratio (SNR) shapes how listening in fluctuating background sound operates is tested by assessing different neural rate and temporal mechanisms in the primary auditory cortex of gerbils.

Chapter 4 examines the effect of binaural sound level on lateralization with low-frequency ITD. For sounds near sensation threshold, the hemispheric-difference model predicts that with decreasing sound intensity, location percepts will become increasingly biased towards a focal point. This is contrary to the level-invariance prediction of place-coding. This confound is tested in two experiments: an ITD-ITD pointer-target matching task and an ITD identification task. In both

experiments, normal-hearing listeners judge the lateralization of band-limited noise tokens. Chapter 5 expands on the findings of the previous chapter and reveals how ITD lateralization adapts to stimulus history. In the labelled-line model, neurons only respond to limited spatial directions. Therefore, preceding sound should only lead to adaptation in neurons sensitive to the preceding sound's direction. In contrast, the hemispheric-difference model represents space through spike-rate distances relative to a perceptual anchor, and hence, predicts that perceived location should adapt towards a focal point. Here, in four experiments, trained and naïve subjects are either unilaterally or bilaterally tested on an ITD identification task.

Finally, Chapter 6 summarizes the findings of these experiments, and discusses their implications on future research.

CHAPTER 2

ELECTROPHYSIOLOGY AUDITORY RECORDING SYSTEM

Electrophysiology Auditory Recording System is a software and hardware bundle developed for synchronized rodent auditory behavioral assessment with simultaneous real-time wireless recording of cortical electrophysiology to study hearing in situations with background sound. The Electrophysiology Auditory Recording System software (referred to as EARS software, or EARS for short) supports data acquisition with up to two simultaneous sound sources, a target and a masker. The masker sound is played continuously throughout each recording session. The target sound is triggered through animal behavior and played back at a fixed phase delay relative to the masker. The signal-to-noise ratio of target and masker can be adjusted online. General digital/analog input/output operations are facilitated through two PCIe data acquisition (DAQ) cards (PCIe-6321 and PCIe-6341, National Instruments Corporation). A single 16-bit analog channel is utilized at 100 kS/s to output auditory stimuli to a sound amplifier and then into a loudspeaker. 15 analog input channels are employed for collecting physiological signals from a bioamplifier at a sampling frequency of 31.25 kS/s per channel. Custom electronic circuits are designed to drive infrared (IR) emitter diodes and also to condition the output signal of their paired IR photodiodes to make them compatible with the DAQ card's digital input channels. An additional digital channel is used to control the testing booth's lights. All input/output channels are synchronized by sharing the sampling clock pulse of the two DAQ cards via an RTSI bus cable. A syringe pump (NE-1000 Programmable Single Syringe Pump, New Era Pump Systems, Inc.) is interfaced through a USB-RS232 emulator using its own factory defined protocols and set of commands, to allow delivering water rewards in an appetitive behavioral paradigm.

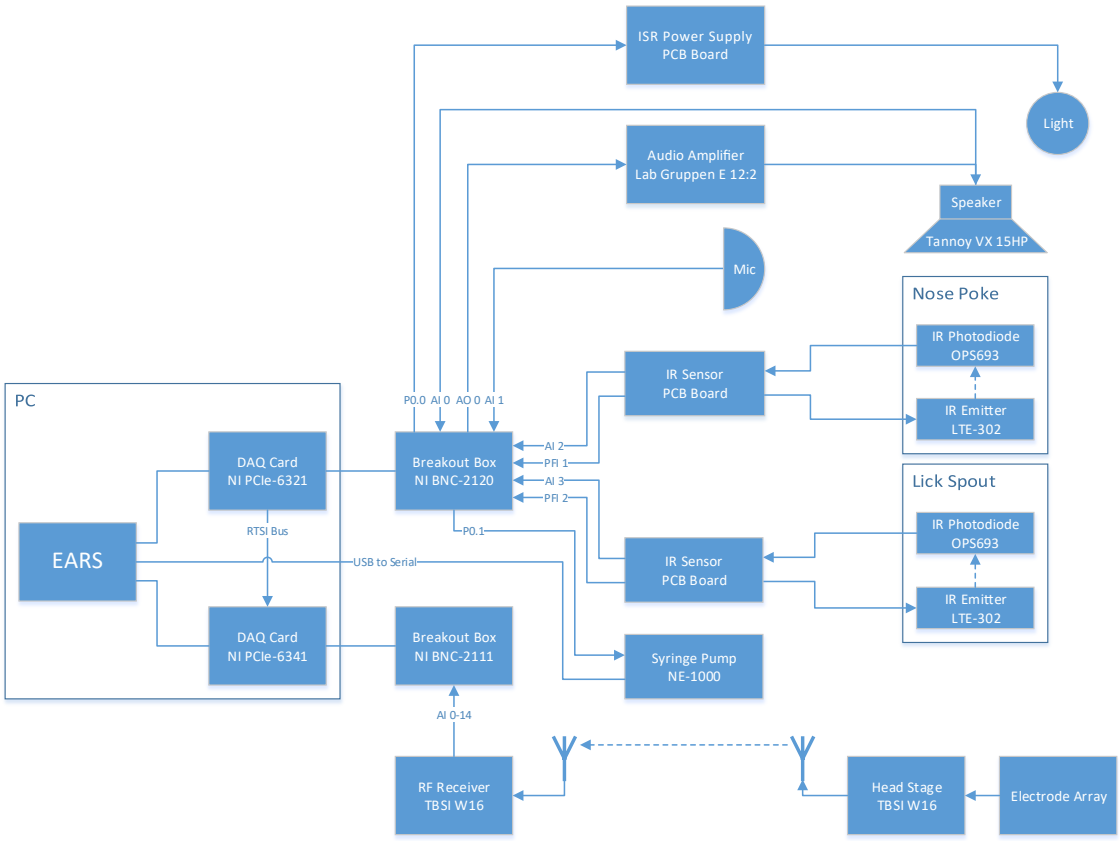

Figure 2.1 The full hardware schematic of the closed-loop recording system illustrates the association among the aforementioned components. The system is comprised of the EARS software running an a PC, which control the entire system through two DAQ cards. These DAQs allow outputting signals to control the experimental setup, play acoustic stimulus, switch the light on/off, and also enables receiving signals about contact with nose-poke or lick-spout sensors, recording of neural activity, and recording the microphone. The neural activity are picked up by an implanted electrode array attached to wireless head stage transmitter. IR emmiter and photodiodes are interfaced using custom-designed electronic boards, and the testing booth's light is switched by a custom-designed power supply.

2.1 Electronic Circuits

The hardware components that constitute the recording system may or may not be able to directly communicate with each other. Most parts are easily made compatible by using an appropriate cable (such as 68-Pin VHDCI for DAQs), or an adapter (USB-RS232 emulator for the syringe pump). Although, not all connections are as easy as plug-and-play, and require custom electronics to fine tune and condition the signals between different devices.

For example, IR emitter diodes require a certain amount of current to switch on and radiate an IR beam. Providing less current than required results in a weaker beam that will not be detected by the paired photodiode, whereas providing more current than a certain limit would lead to the component burning out. A similar principle applies to IR photodiodes; they require a bias circuitry specified by the manufacturer to operate. Moreover, the output of this circuitry is not necessarily compatible with the readout device (here, a DAQ) and its voltage levels should be adjusted.

The following sections describe the custom electronic circuitry and their printed circuit boards (PCBs) that were designed to fit the recording system's use case. These boards, as they were, along with all the wires interconnecting them to other devices, would have been difficult to maintain and manage. Hence, a modular enclosure (referred to as the electronic box, see Figure 2.2) was designed to contain and organize the circuit boards. The electronic box was designed in a planar layout and was laser cut[1] out of transparent acrylic sheets. The box could fit multiple movable shelves that allowed addition of PCBs with various sizes using plastic screws and bolts. The front and back panels had cuts and engravings for operable switches, knobs, LED indicators, and BNC connectors.

[1]For laser cutting see: https://www.ponoko.com

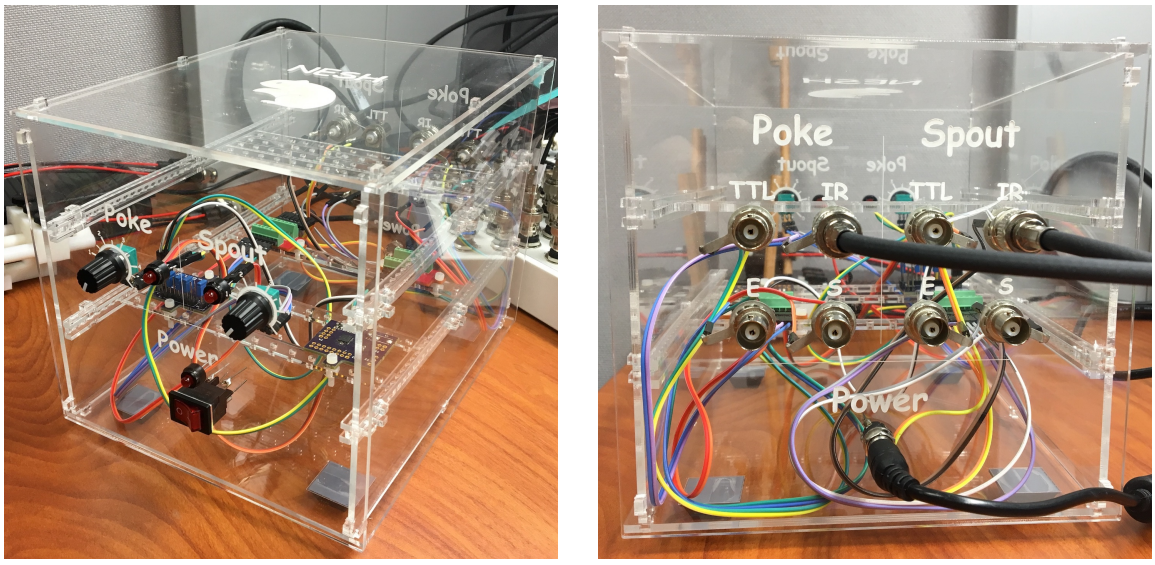

Figure 2.2 The electronic box, a modular enclosure for custom-designed electronic circuitry, laser cut and made out of transparent acrylic sheets.

2.1.1 ISR Power Supply

Electronic circuitry are often comprised of various smaller components that each have their own manufacturer defined power supply level for operation. This could particularly introduce a challenge in designing a modular system with unknown requirements for future components. To allow providing power to various components, an adjustable power supply was designed using the TPS54232 chip (Texas Instruments, Inc.). This chip is an integrated switching regulator (ISR) capable of providing up to 2 A of current. The switching technology used in the chip emanates minimal heat and has very low power loss. The designed circuitry (Figure 2.3) is activated with an input voltage in the range of 3.5-28 V, and can be programmed to output any voltage between 0.8 V and 25 V (in steps of 0.1 V). The output is adjusted by placing jumpers on the P4 pin-headers. Voltage values indicated by the jumpers will be summed as V_{adj} and will set the output to:

$$V_{out} = V_{adj} + 0.8. \tag{2.1}$$

In the current setup of the recording system, the main use of this power supply is to provide the IR sensor circuitry with a low noise 5 V supply. In order to yield a 5 V power supply, three jumpers should be place on the 3.2, 0.8 and 0.2 pin-headers:

$$V_{adj} = 3.2 + 0.8 + 0.2 = 4.2V,$$
$$V_{out} = V_{adj} + 0.8 = 4.2 + 0.8 = 5V. \tag{2.2}$$

The main repository for this ISR power supply is located on GitHub[2].

Another instance where this custom circuitry aided design of the recording system was for rapid switching of the testing booth's lights. Commercially available devices for switching power to the lights use electromechanical relays. When changing from on to off or vice versa, these relays introduce a loud acoustic noise which interferes with our auditory experiment. The noise often startles the gerbils and disrupts their behavior. More over, this interference would be reflected in the recorded neural activity and leads to contamination of all analysis. Another issue with these types of relays is that they connect to the main power cord, and hence, they would be switching the entire adapter that delivers power to the lights. Most adapters contain electrical transformers and are relatively slow when turning on. The animal training process is extremely time sensitive and such delays would completely impede forming behavioral associations.

The solution to this problem was quite simple. The ISR power supply boards have an enable switch (EN) which can either be manually activated with jumpers, or automatically by receiving a digital signal from the DAQ. The facilitate this feet, another ISR power supply board was placed in the electronic box. The extra board was configured to output 11.4 V for powering 2 LED light strips attached to the ceiling of the animal testing booth.

[2]ISR power supply design files: https://github.com/nalamat/supply-isr-adj-single

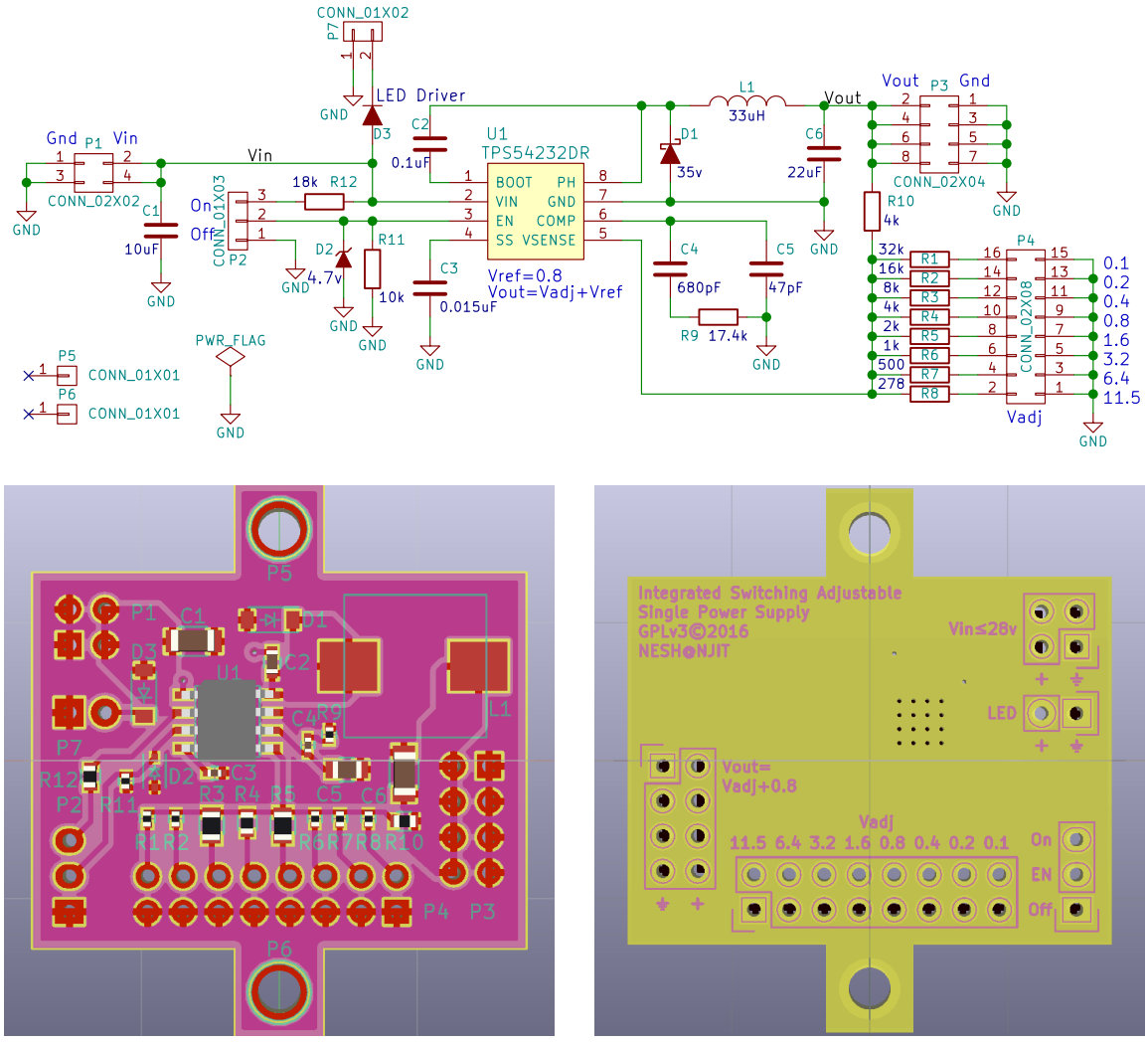

Figure 2.3 Circuit schematic (top) and PCB layout (bottom left and right) for the integrated switching regulator (ISR) power supply with an adjustable output voltage.

2.1.2 LDO Power Supply

ISR power supplies generally excel at power efficiency and are often the go-to type of regulator for most applications including consumer devices. However, they are not ideal for use cases where it is necessary to have an accurate and stable supply voltage without noise. An immediate example of this in neurophysiology would be when using instrumentation amplifiers. These devices usually have high amplification gains and small fluctuations in their power supply could greatly impact the quality of their output signal. For having a high quality and low noise power supply, a low dropout (LDO) power supply was designed based on the TPS7A470X chip (Texas Instruments, Inc.). This power supply has an input voltage range of 3-36 V, and provides a single output voltage that is adjustable from 1.4 to 20.5 V (in 0.1 V steps) via the P3 pin-headers. The output voltage is determined as:

$$V_{out} = V_{adj} + 1.4. \tag{2.3}$$

The main repository for this LDO power supply is located on GitHub[3].

2.1.3 IR Sensing

A key point of control in a closed-loop system is sensing. In the designed recording system, sensing is achieved via IR sensors that detect contact with the nose-poke or lick-spout. There are two parts to IR sensing; first, an IR emitter diode (less accurately referred to as IR LED) has to generate a sufficiently powerful infrared beam. Second, an IR photodiode must be placed directly in front of the emitter at a close enough proximity. When there is no contact, the IR beam will be detected by the photodiode and electrical current starts flowing through it. On the contrary, when an animal approaches the nose-poke or lick-spout, the beam is intercepted and the flow of current through the photodiode stops.

[3]LDO power supply design files: https://github.com/nalamat/supply-ldo-adj-single

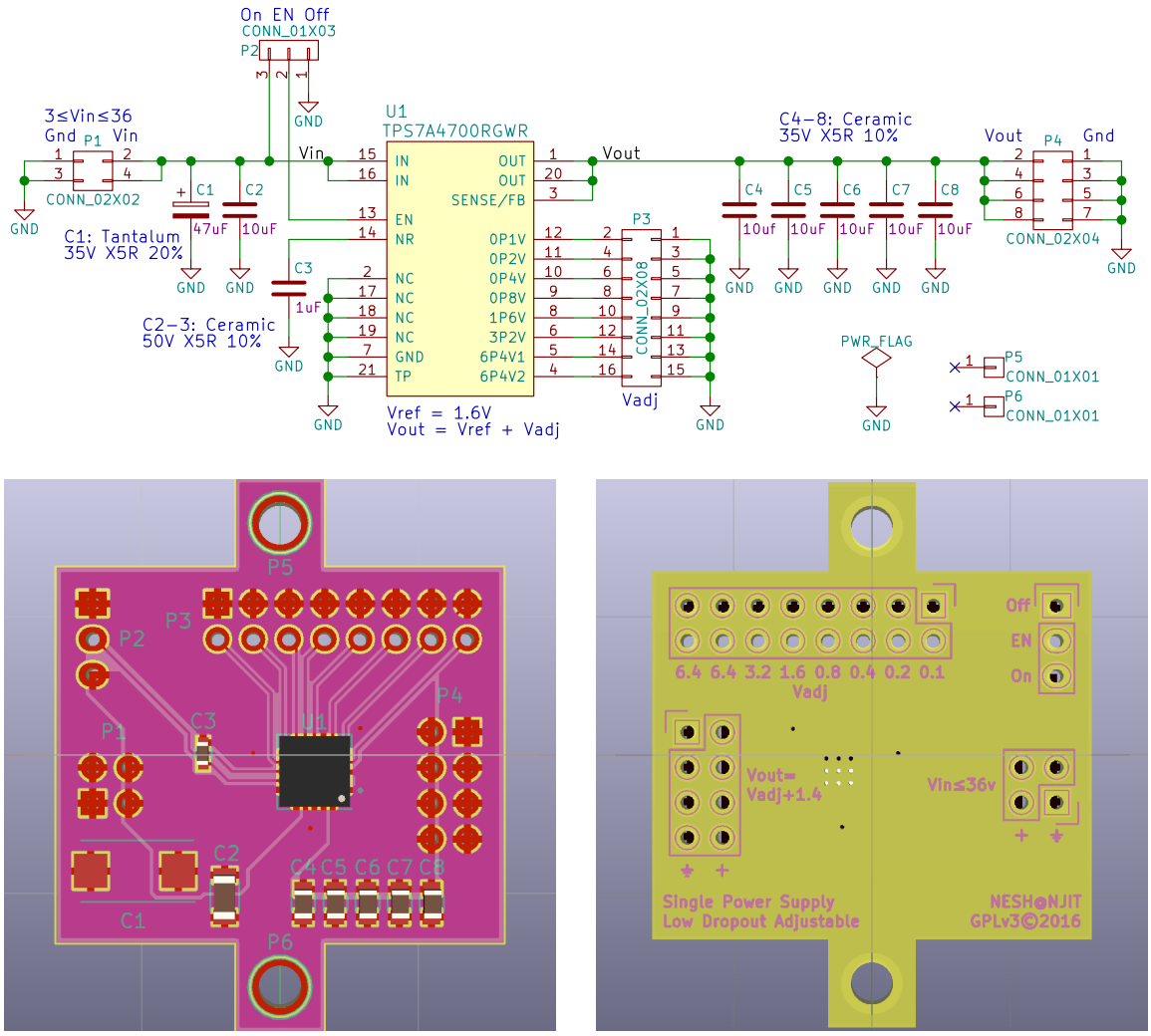

Figure 2.4 Circuit schematic (top) and PCB layout (bottom left and right) for the low dropout (LDO) power supply with an adjustable output voltage.

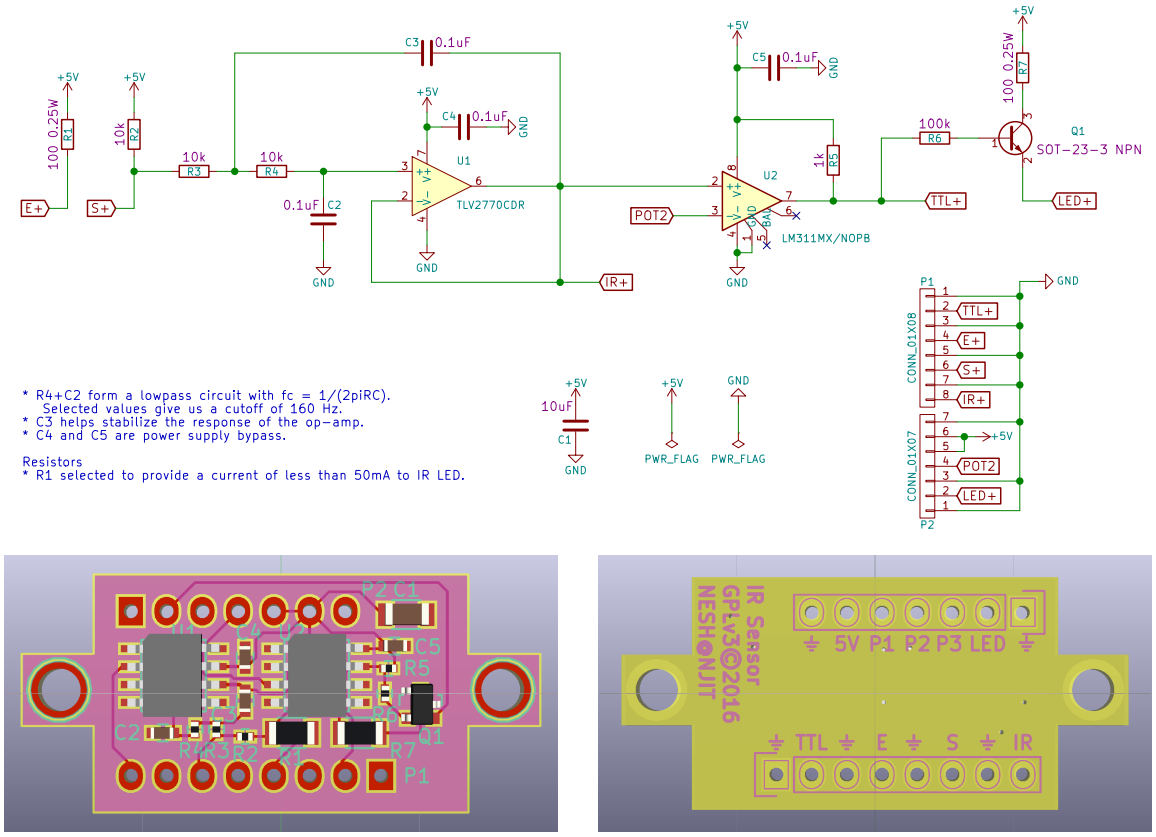

Figure 2.5 Circuit schematic (top) and PCB layout (bottom left and right) for the IR sensor board. This circuit can activate a pair of IR emitter diode and photodiode.

To facilitate IR sensing, the circuit board shown in Figure 2.5 was adapted from a similar design[4] using KiCad EDA. This circuit board activates an IR emitter diode (LTE-302) connected to the `E` pin, and continuously reads the current flowing through an IR photodiode (OPS693) connected to the `S` pin. The value read from the photodiode is transformed and amplified into a voltage between 0 V and 5 V, and made available via the `IR` pin. This readout is an analog signal and thus the circuit has to further threshold it in order to decide whether a contact has occurred or not. The cut off threshold is manually adjusted by a potentiometer connected to the `POT2` pin. Eventually, the result of this thresholding is transferred to the `TTL` pin as a digital signal with values of either 0 V or 5 V.

[4]IR sensor design files by Brad Buran: https://github.com/bburan/circuit-IR-sensor

2.2 Mechanical Design

2.2.1 Nose Poke

During the behavioral task, gerbil learn to poke their nose inside a cylindrical shape to initiate a trial. This cylinder was designed as a 3D model in SketchUp (Trimble, Inc.) and printed using a 3D printer (Flashforge, Zhejiang Flashforge 3D Technology Co., Ltd.). The 3D printer used acrylonitrile butadiene styrene (ABS) or polylactic acid (PLA) filaments in an additive process to materialize the computer designs. Both these materials are bio-compatible and safe to use for animal experiments. The nose-poke (shown in Figure 2.6) had two pair of slots (four slots in total) incorporated around the body that allowed insertion of IR sensors without exposing the wires to the animal. Each slot pair had small holes facing inside the nose-poke to allow emanation of the IR beam from one side to the other. The slot pairs were placed at different depths to provide the experimenter with flexibility of using shallower contact sensing in earlier stages of training, or deeper/longer contact during advanced stages of testing and physiology recording. The sensors were held in place using small plastic screws (nylon) on the sides, and the nose-poke itself was secured to the cage bars using larger screw and 3D printed bracket.

2.2.2 Lick Spout

The lick-spout show in Figure 2.7 was also designed and printed as a 3D model using ABS or PLA filaments. Compared to the nose-poke, the lick-spout required more flexibility for spout and sensor placement; hence, in addition to the main chassis, it was designed with two sliding parts. The first part is a long beam at the center which has a narrowing length-long hole. This hole was inserted with a water tube from the rear to provide gerbils with a flow of water. The other two lateral beams had slots for holding IR sensors, and were interconnected to ensure the sensors were always directly facing each other. On top of the chassis, three thumb screws held the sliding beams

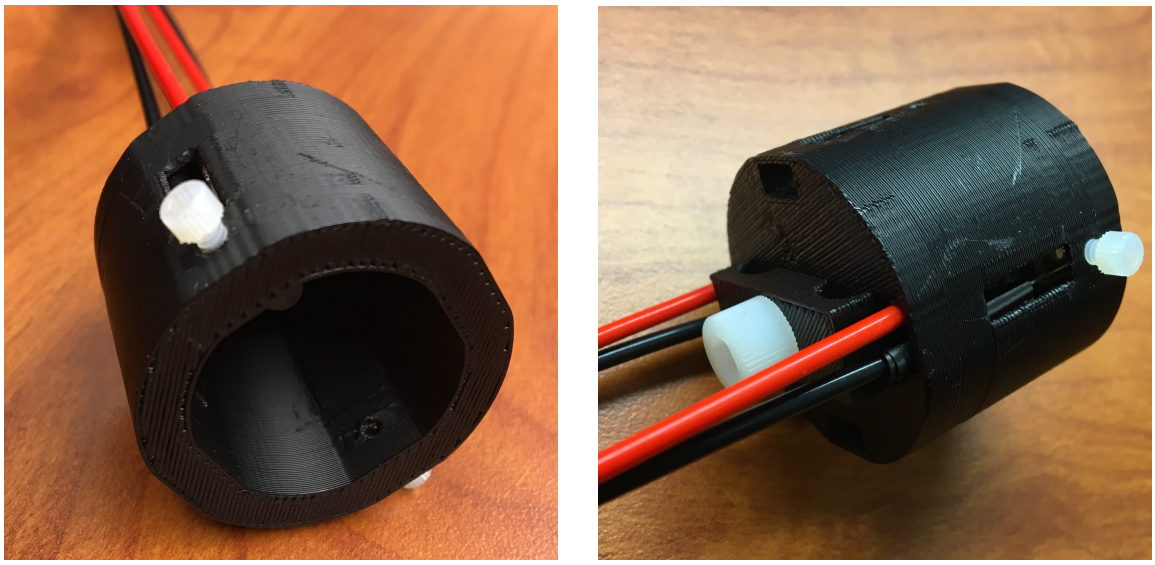

Figure 2.6 Cylindrical nose-poke designed as a 3D model and printed out of plastic filaments. IR emitter diode and photodiode pairs are inserted through slots from behind the nose-poke and are held in place by small screws. The nose-poke itself is secured to cage bars using a bigger screw and a bracket piece that was also 3D printed.

in place. During early stages of training the experimenter has to ensure that gerbils discover the lick-spout, and thus, has to advance the sensor and spout beams inside the cage. Later on, when gerbils fully understand the functionality of the lick-spout but become impatient with NoGo trials, the beams are retracted behind the cage bars. Doing so allows licking of the spout for receipt water while prevents excessive chewing and destruction of the apparatus.

2.2.3 Testing Cage

Gerbils were trained and tested inside a custom cage made mostly out of plastic. In choosing the proper material for these experiments multiple considerations had to be made: 1) the material needed to be bio-compatible and safe for contact with animals, 2) sensitization is often a concern for surfaces that come in contact with animals, and materials such as wood or fabric are often out of question, 3) gerbils have a stereo-typical chewing behavior that in the long-term is destructive to most types of

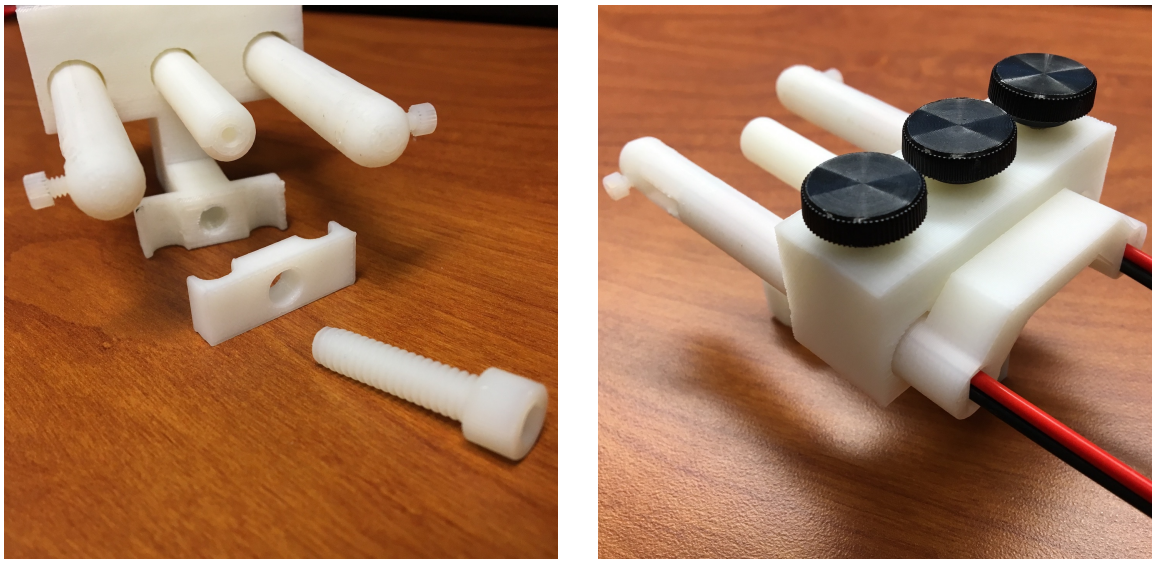

Figure 2.7 Multi-piece lick-spout designed as a 3D model and printed out of plastic filaments. Two long slots on the back allow insertion of IR emitter diode and photodiode in two opposing beams. A hole in the back is used to attach a tube of water to the central beam. Three screws on the top hold the sliding beams in place. Similar to nose-poke, a 3D printed bracket piece is used to secure the entire lick-spout against cage bars.

materials, 4) very hard surfaces have excessive reverberation and could potentially confound an acoustic experiment, 5) most metals could distort the quality of radio signals and interfere with wireless transmission of neural recordings.

Based on these considerations, high density poly ethylene (HDPE) plastic was chosen for the cage frame due its relatively high strength and bio-compatibility. Although for cage bars, fiberglass-epoxy (brand name Garolite) was used that can endure any amount of chewing. The bars were 1/4" in diameter with 3/8" space in between each one (the space between the frame and last bars on each side was not exactly 3/8"). A plastic mesh was placed on the bottom of the cage (see Figure 2.8) to allow disposal of feces to the bottom tray. The cage was secured at the center of an HDPE sheet, and the sheet was held on top of a set of polyvinyl chloride (PVC) pipes forming a table base. The rest of the testing apparatus such as nose-poke, lick-spout and recording antennae were attached to the cage bars or frame.

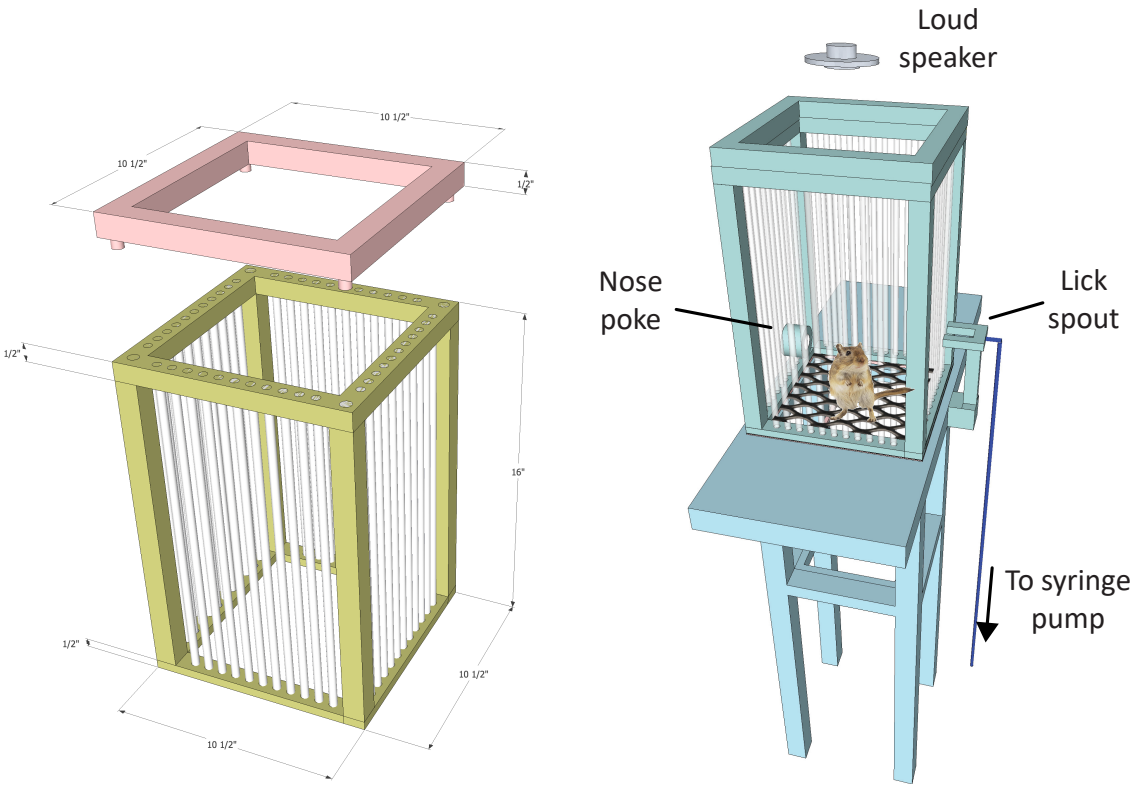

Figure 2.8 Testing cage specifications (on the left) and schematic representation of all mechanical components that are attached to the cage (on the right). The cage frame is made out of HDPE and held together using plastic pins and screws. For cage bars, Garolite was used due to its endurance against chewing. The entire cage is placed on an HDPE sheet held by a table base made from PVC pipes. Nose-poke and lick-spout are secured against cage bars, and a loudspeaker is placed directly on top of the cage.

2.3 EARS Software

The recording system's main point of control is the EARS software [1] that was developed as an open-source package[5] using the Python programming language. EARS provides a platform for running auditory experiments on awake and freely behaving rodents. This software incorporate many other open-source libraries[6], and was inspired by and partially rewritten from NeuroBehavior[7].

EARS allows playback of two simultaneous acoustic stimuli, a continuous background sound and an on-demand target sound, with microsecond precision timing. At the same time, signals from the nose-poke and the lick-spout are monitored for instantaneous decision making. All of the parameters and timings of the experimental paradigm along with the characteristics of acoustic stimuli could be adjusted by the experimenter in real-time.

The EARS software facilitates incremental storage of behavioral and electrophysiological data in HDF5 file format using the PyTables library. In addition to data storage, the NumPy and SciPy libraries are utilized to process the acquired data online. The results are then visualized with customized plotting widgets implemented using the PyQtGraph package. All these tasks are performed in parallel and with consideration of asynchronous programming confounds and thread-safety measures.

The main entry point to the software is the `load.py` script which can be executed from command line as follows:

```
$ python load.py
```

Optionally, if no particular hardware (DAQ, pump, etc.) is connected, a simulation mode is available for development and testing purposes:

[5]EARS source code: https://github.com/nalamat/ears
[6]SciPy, NumPy, Pandas, PyTables, PyQt, PyQtGraph, PyOpenGL, and PyDAQmx
[7]NeuroBehavior source code by Brad Buran: https://github.com/bburan/NeuroBehavior

```
$ python load.py --sim
```

First the setup window opens for general experiment settings. Depending on the selected recording type, either the behavior or both behavior and physiology windows will open for viewing and controlling the experiment session.

2.3.1 Setup and Calibration

The setup window is the first dialog that is shown after opening EARS (see Figure 2.9 Left and `setup.py`[8]). Here, the experimenter can adjust general settings of an experiment session. The default values are always populated from the last application run, and thus, for most training sessions, only the `Subject ID` and `Paradigm file` must be selected. Subject ID is an internal name chosen for gerbils at the time of delivery or when juveniles are being weened. This ID is often concatenation of the original cage ID and the side of the body where shaving marks are located. The paradigm file holds all the parameters of an experiment, and is generally named with either the subject ID or the cage ID. `Experiment mode` determines the stage of training (spout/poke/target training, Go/NoGo) or the task engagement mode (passive, all other options are active). When gerbils are fully trained, implanted, and ready to be recorded, `Recording` must be set to `Physiology` to activate visualization and storage of neural recordings. `Data file` specifies the location where all data will be stored. This field is by default automatically generated from other settings; however, by unchecking the box next to it, a custom file name can be manually entered. The `Rove` checklist allows selection of all parameters that are required to be randomly roved throughout different trials. When selecting a pre-existing `Paradigm file`, `Rove` becomes deactivated since the rove parameters are already

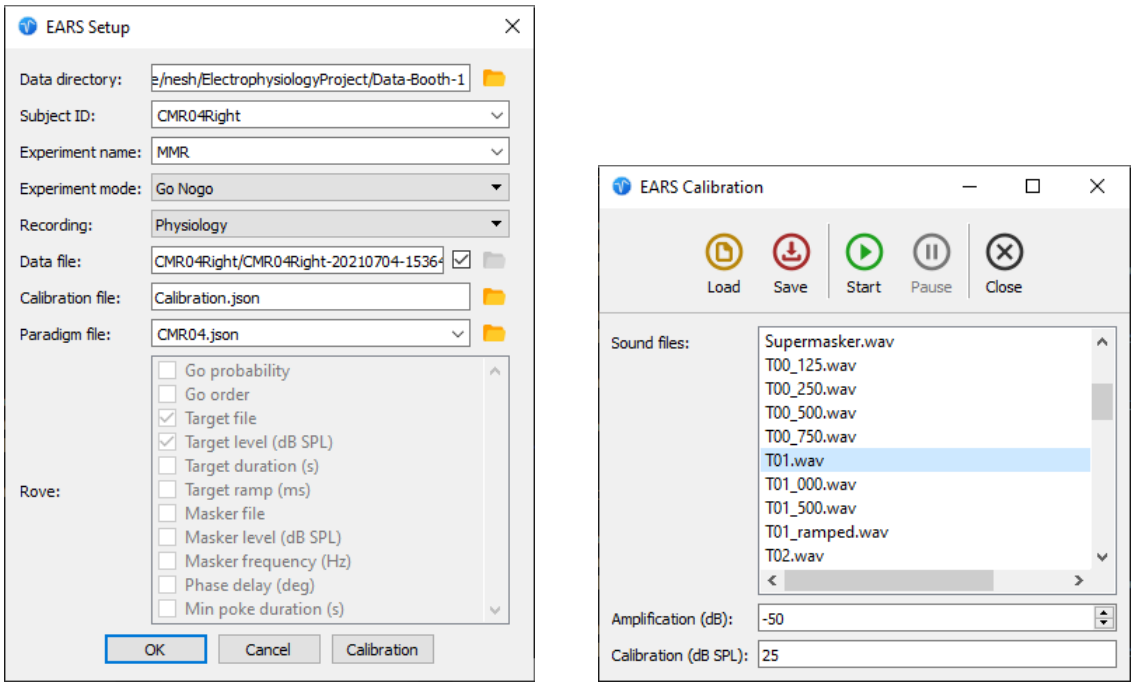

Figure 2.9 **Left** The setup window is the first dialog shown to users after opening EARS, providing options for setting up an experiment session. **Right** The calibration window allows playing and adjusting sound level of experimental stimuli based on readings from a sound level meter.

specified by the paradigm. To setup a completely new paradigm and specify the desired rove parameters, `Paradigm file` has to be cleared.

The setup window also has an option for selecting a `Calibration file` for loading calibration data of acoustic stimuli. Although, if there are no calibration data available, or whenever there has been a change in the acoustic setup of the testing chamber, all stimuli need to be (re-)calibrated. This process can be performed in the calibration window that lists all available sound files in the `/stim` directory (see Figure 2.9 Right and `calibration.py`[9]). For calibration, a sound file must be selected, played, and has its level adjusted using the `Amplification` field. It's best to start with a low amplification and gradually increase it until a comfortably audible sound level is achieved. Then, the read measurement from a sound level meter can be entered in the `Calibration` field. The current version of EARS only stores one

[8]Setup window: https://github.com/nalamat/ears/blob/master/setup.py

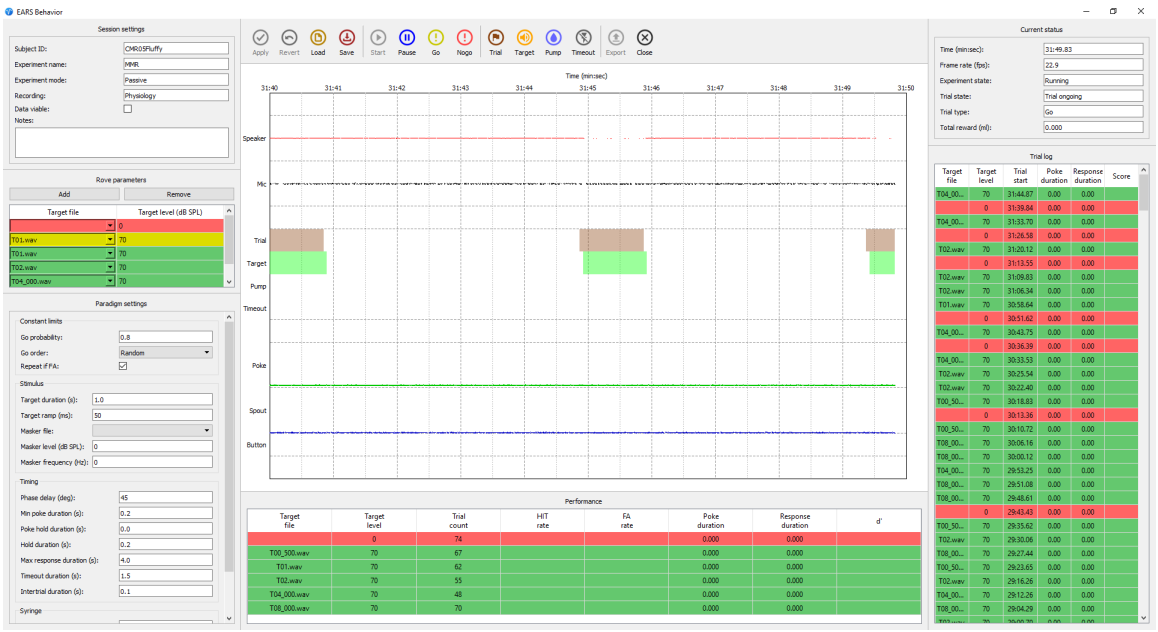

Figure 2.10 The behavior window allows controlling all aspects of the behavioral paradigm, as well as visualizing all events occurring in real-time.

calibration point per sound file; hence, it is recommended to calibrate all sounds at the same dB SPL.

2.3.2 Behavior and Physiology

After adjusting general settings of the session, the behavior window is shown (see Figure 2.10 and `behavior.py`[10]). This dialog has multiple sections for controlling and monitoring an experiment in real-time.

On the top left corner, some of the session settings that were selected in the setup window are shown. Below session settings, a multi-column list allows specifying values for all of the rove parameters. Each green row represents a set of parameters that will be randomly selected for each Go trial. Values from the red row are only used for NoGo trials, and the yellow row is used for Go Remind trials. New sets of values for Go trials can be added or removed.

[9]Calibration window: https://github.com/nalamat/ears/blob/master/calibration.py
[10]Behavior window: https://github.com/nalamat/ears/blob/master/behavior.py

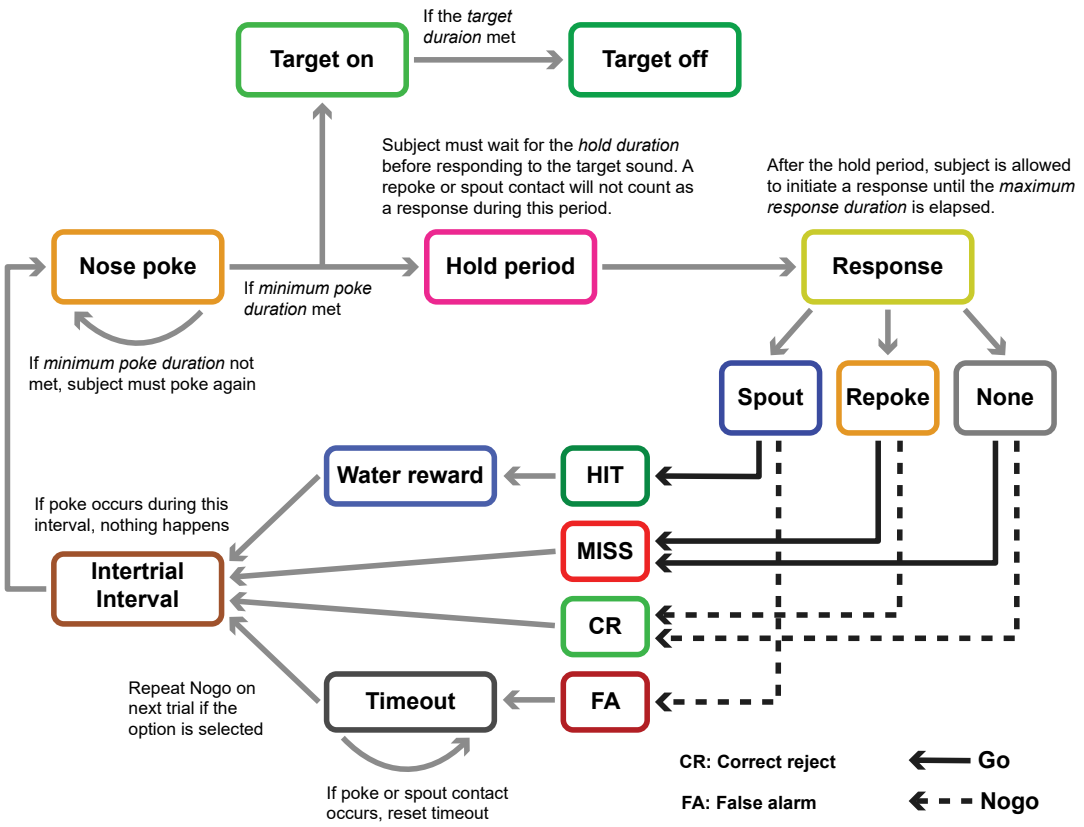

Figure 2.11 Flowchart of the appetitive Go/NoGo behavioral paradigm.

On the bottom left side of the behavior window, all paradigm parameters are listed except the ones used for rove. These parameters allow fine tuning of the behavioral paradigm as shown in Figure 2.11. The most important ones are as follows:

- `Go probability`: The probability that the next trial would be randomly selected as Go instead of NoGo (modeled with the Bernoulli distribution). This random selection only occurs if the next trial is not forced to be a Go Remind, NoGo Remind, or NoGo Repeat.
- `Repeat if FA`: If checked, after every False Alarm, the trial type is forced to be NoGo Repeat.
- `Target file`: The sound file to be played after successful initiation of a trial. Note that in Figure 2.10 `Target file` is selected as a rove parameter and is not shown under paradigm settings.
- `Target level`: The loudness of the target sound in dB SPL. Calibration data is used to accurately scale and play the sound at the specified level. Note that in Figure 2.10, `Target level` is selected as a rove parameter, and therefore, is not shown under paradigm settings.
- `Target duration`: The duration of the target sound in seconds. If the specified duration exceeds the length of the target file, the waveform is repeated without any gaps.
- `Masker file`: The sound file to be played continuously throughout the entire experiment session. The waveform is looped and repeated without any gaps.
- `Masker level`: The loudness of the masker sound in dB SPL, scaled according to the calibration data.
- `Masker frequency`: This value indicates the frequency of the masker in Hz, and is only used for insertion of target at a particular phase of a modulated masker. For stationary maskers, set to zero.
- `Phase delay`: Phase delay in degrees is used in combination with masker frequency to calculate the exact insertion time of target at a particular phase of a modulated masker.
- `Min poke duration`: In seconds, the amount time a nose-poke should be held until its acknowledged by the software for trial initiation (See Figure 2.11).
- `Max response duration`: Determines how long in seconds does the animal have to respond after successful initiation of a trial to either perform another nose-poke or approach the lick-spout. If this duration is elapsed, a response of `None` is recorded.

- `Timeout duration`: The duration in seconds for which the lights are turned off after a False Alarm.
- `Pump rate`: Adjusts the rate of water flow by the syringe pump in milliliters per minute. Typical values are 1-2 mL/min.
- `Reward volume`: The volume of water in µL given as reward after a successful spout contact (Hit) in a Go trial.

The bar at the top of behavior window provides controls for saving or loading paradigm parameters, starting or pausing the experiment, and forcing next trial type to Go or NoGo Remind. Furthermore, the user can intervene in the predefined paradigm flow by manually initiating certain actions such as trial start, target stimulus, pump activation, and timeout. The central pain contains visualizations for all software or animal generated signals to help monitor the training process and also the functionality of the system. Finally, the trial log and performance sections on the right and bottom in order demonstrate trial by trial and statistical summary of the subject's behavioral performance throughout the session. More information about certain fields such as Hit rate, FA rate and d′ is given in Chapter 3.

While configuring an experiment session in the setup window, if `Recording` mode is set to `Physiology`, the physiology window will be shown (see (Figure 2.12 and `physiology.py`[11]). This dialog visualizes the neural activity that are wirelessly transmitted from the brain implant. On the left side of the physiology window, there are options for adjusting the pre-processing of these signals, such as: the lower and higher cutoff frequencies of the applied band-pass filter, amplification scale, visibility of traces from each electrode, and also whether each trace is included in the grand mean subtraction.

[11]Physiology window: https://github.com/nalamat/ears/blob/master/physiology.py

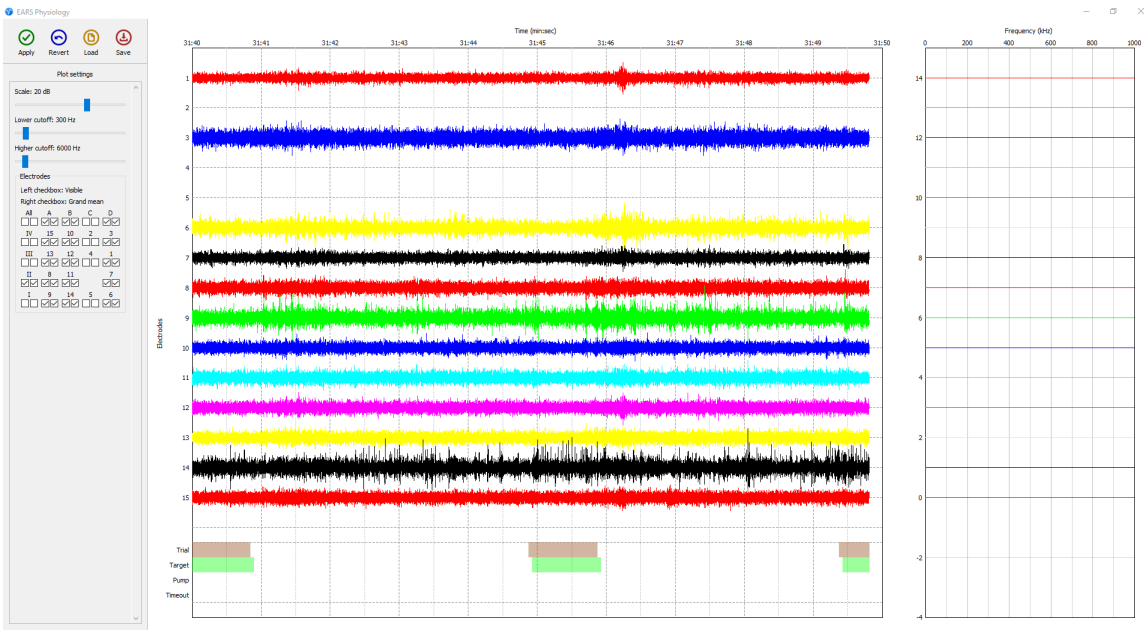

Figure 2.12 The physiology window shows the pre-processed neural activity in real-time.

2.3.3 EasyDAQmx: Data Acquisition Interface

National Instruments DAQ cards (NI-DAQs) offer a cost effective means of recording time series in the physiology of animals and humans. However, a roadblock to widespread adoption of NI-DAQs in many research labs is that the NI-DAQ hardware's drivers, as well as the existing NI-DAQmx programming interface require expert training in order to be used for sophisticated experimental design. To overcome these challenges, EasyDAQmx was developed as an open-source library in the Python programming language and made available for academic use. The main repository for this library is located on GitHub[12].

The library has the following main functionalities: 1. digital input, 2. digital output, 3. analog input, and 4. analog output.

EasyDAQmx relies on the existing NI-DAQmx programming interface. Having the capability to both sample and output various modes of data, the NI-DAQ hardware functions as a hub between a computer and data acquisition hardware.

[12]EasyDAQmx source code: https://github.com/nalamat/easydaqmx

This allows researchers to build a methodical data acquisition system, for almost any research study or experiment.

NI-DAQmx is a free to download application programming interface (API) which allows users to take full advantage of their NI-DAQ card. This driver gives users the ability to build applications without manually programming the data collection and output hardware. Additionally, it allows for the synchronization of multiple hardware components for near-instantaneous communication between them in the system.

Another open-source library called PyDAQmx allows users to utilize the NI-DAQ hardware using Python. Providing a transition from the C++ standard NI-DAQmx interface to a Python-based interface, PyDAQmx allows for a more dynamic approach in setting up and conducting experiments with NI-DAQs. This package is readily accessible from the Python Package Index (PyPI) through the Python Package Installer (pip) tool.

Based on PyDAQmx, EasyDAQmx simplifies the existing functionalities in order to provide a more user-friendly programming interface. Primarily this is accomplished through the creation of several classes that concentrate on how the user utilizes an NI-DAQ card. For example, when using a card to sample an input, one would create an instance of the `AnalogInput` or `DigitalInput` classes, and when providing an output signal, one would create an instance of the `AnalogOutput` or `DigitalOutput` classes. That being said, EasyDAQmx is primarily broken down into six classes four of which are critical for use, and as such will be discussed here. It is important to note that for each NI-DAQ card in use a separate instance of an input or output class must be defined in order to prevent overlap among channels/lines, however, any number of channels or lines can be created per an input/output class.

The `DigitalInput` class, as the name suggests, initializes the digital inputs for a NI-DAQ card. The class fundamentally works by using a callback function whenever

a `DigitalInput` function is called to determine whether a line is on or off. That being said, the class takes one required parameter `line`, which defines the location of the line to be read (i.e., `'/dev1/ai0'` or `'/dev2/ai0:15'`). The class also takes six optional parameters, one of which will be discussed here. The `debounce` parameter allows for the adjustment of noise filtering on a digital line with options for zero seconds, 5.12 milliseconds, 10.24 microseconds, and 160 nanoseconds. This allows for the optimization of any data acquisition system using digital lines preventing false-positive signals on a line by line basis.

```
digitalInput = daq.DigitalInput('/dev1/port0/line0', debounce=5.12e-3)
digitalInput.start()

value = digitalInput.read()    # Returns a single or an array of integers

digitalInput.stop()
```

```
def edgeDetected(task, lineName, edge):
    print('{edge.title()} edge detected on `{lineName}` digital line'

digitalInput = daq.DigitalInput('/dev1/port0/line0:1',
    edgeDetected=edgeDetected, lineNames=['poke', 'spout'])
digitalInput.start()
```

The `DigitalOutput` class, initializes the digital outputs for a NI-DAQ card and allows for the creation of several output lines. The class has one required parameter line, functioning identically to its use in the `DigitalInput` class. The class also has two optional parameters: `name` and `initialState`, each functioning intuitively as named. One key functionality of the `DigitalOutput` class, that improves upon previous usage of python for the NI-DAQ, is that it allows the user to easily change the output value of lines using the `write()` function.

```
digitalOutput = DigitalOutput('/dev1/port1/line0')
digitalOutput.start()

digitalOutput.write(1)    # Output logical high

digitalOutput.stop()
```

The `AnalogInput` class initializes the analog inputs for a NI-DAQ card, taking three required parameters: `line`, `fs`, and `samples`. The `line` parameter is used identically to its use in the `DigitalInput` class, the `fs` parameter is used to set the desired analog sampling frequency, and the `samples` parameter refers to the number of samples needed to be read from a line. In particular, when defining the `samples` parameter, it can be defined for finite sampling given an integer, or it can be defined for infinite sampling given `numpy.inf`. However, it should be considered that if given `numpy.inf` the user will have to define a callback function. This is done in order to process data in chunks as they are being made available instead of reading all the data at once, improving processing efficiency.

```
analogInput = AnalogInput('/dev1/ai0', fs=1000, samples=1000)
analogInput.start()
data = analogInput.read()
```

```
def process(task, newData):
    store(newData)
    filteredData = filter(newData)
    plot(filteredData)

analogInput = AnalogInput('/dev1/ai0', fs=1000, samples=np.inf,
    dataAqcuired=process)
analogInput.start()
```

The `AnalogInput` class also takes 10 other optional parameters, three of which will be expanded on here. The `timebaseSrc` parameter allows for the definition of a sample clock and can be initialized using the sample clock in any NI-DAQ card in the data acquisition system, or even an external clock. This allows for synchronous sampling across multiple NI-DAQ cards, if needed, and overcomes the time difference in sample clocks across hardware. Additionally, the `timebaseRate` parameter allows for the specification of rate in Hz of an identified sample clock in the aforementioned parameter, when using an external clock. It should be noted that, in order to use the time base functionalities of EasyDAQmx an RTSI cable is required to share and maintain sample clock signals. Lastly, the `startTrigger` parameter and functionality allows for the instantaneous start of an analog input channel upon the rising edge of a, currently active, digital line. This also enables the synchronous start of multiple tasks as the `startTrigger` for one task can be another task, successfully overcoming the issue of working across multiple NI-DAQ cards or other sample clock issues.

```
analogInput1 = daq.AnalogInput('/dev1/ai0:1', fs=1000, samples=1000,
    timebaseSrc='/dev1/20MHzTimebase')

analogInput2 = daq.AnalogInput('/dev2/ai1:15', fs=1000, samples=1000,
    timebaseRate=20e6, timebaseSrc='/dev1/20MHzTimebase',
    startTrigger='/dev1/ai/StartTrigger')

analogInput1.start()    # analogInput2 will also start synchronously
```

The `AnalogOutput` class initializes the analog outputs for a NI-DAQ card. The class consists of three required parameters `lines`, `fs`, and `samples`, used identically as in the `AnalogInput` class, and 11 additional parameters, four of which are critical to use. Three of these additional parameter `timebaseSrc`, `timebaseRate`,

and `startTrigger` are used identically to their use in the `AnalogInput` class; the `dataNeeded` parameter, however, is unique to the `AnalogOutput` class. The `dataNeeded` parameter functions to maximize the processing capability of a data acquisition system, by controlling the amount of data being processed by a task. When a task commences, the task is given a portion of the data being collected to be analyzed, but once the task is almost complete the `dataNeeded` function is called. This function then prompts the user to provide more data to the task in order for the task to continue, preventing task error and maximizing the efficiency of the data acquisition system

```
data = np.random.randn(1000)
analogOutput = daq.AnalogOutput('/dev1/ao0', fs=1000, samples=len(data))
analogOutput.write(data)
analogOutput.start()
```

```
def dataNeeded(task, nsWritten, nsNeeded):
    newData = np.random.randn(nsNeeded)
    return newData

analogOutput = daq.AnalogOutput('/dev1/ao0', fs=1000, samples=np.inf,
    dataNeeded=dataNeeded)
analogOutput.start()
```

In summation, EasyDAQmx simplifies the PyDAQmx programming interface allowing researchers to more readily use Python with NI-DAQ cards. This is done by reforming the existing, relatively-complex Python functions more intuitively and prioritizing ease of use and practicality over extensive modularity.

2.3.4 Pypeline: Online Stream Processing

The Pypeline[13] module provides a generic, easy to use, and extendable object-oriented framework for online processing of data streams in Python, which is particularly suited for handling multi-channel electrophysiology signals in the EARS software. The main repository for this module is located on GitHub[14].

Inspired by the dplyr package in R, Pypeline allows definition of stream processing stages as `Node`s that can be connected to each other using `>>`, the shift operator. Alternatively, `|` or the shell pipe operator can be used.

```
daqs.physiologyInput \
    >> pypeline.LFilter(fl=300, fh=6e3, n=6) \
    >> self.physiologyPlot
```

Here, a key difference with dplyr is that `>>` only declares the connections in the pipeline, but no actual processing of data occurs at this statement. All `Node`s inherit the `write()` method that when called, passes new data into the node for processing. The processed data will further be passed to the connected nodes downstream.

Normally data is processed synchronously in the pipeline, meaning the execution of the code goes on halt until all downstream nodes are done with their tasks. Although, a `Thread` node can be inserted into the pipeline to allow asynchronous processing of the data. This could prove useful in situations that the data acquisition thread must not be blocked for too long or when implementing an interactive graphical user interface (GUI). Note that in the current implementation of `Thread`, due to limitations of Python's global interpreter lock (GIL), multithreaded code does not actually run in parallel, but asynchronously. If necessary, the `wait()` method can

[13] *Pypeline* was named by combining the words *python* and *pipeline*, following the convention of naming Python packages by either starting with *py*, or ending with it.
[14] Pypeline source code: https://github.com/nalamat/pypeline

be called on the root node to block execution until an asynchronous pipeline is done processing.

```
daqs.physiologyInput \
    >> pypeline.Thread() \
    >> pypeline.LFilter(fl=300, fh=6e3, n=6) \
    >> pypeline.GrandAverage() \
    >> pypeline.DownsampleMinMax(ds=32) \
    >> self.physiologyPlot
```

Other than linear pipelines, it is possible to connect nodes to multiple branches. This is done by applying `>>` between a node and an iterator of nodes. If the left hand side of `>>` has a single node, its output data will be passed to each of the nodes in the iterator in the order of appearance. In the following example, the same signal is filtered at different frequency bands and then passed on to different plots.

```
daqs.physiologyInput \
    >> pypeline.Thread() \
    >> (pypeline.LFilter(fl=None, fh=300 , n=6) >> self.physiologyPlotLow,
        pypeline.LFilter(fl=300 , fh=6e3 , n=6) >> self.physiologyPlotMid,
        pypeline.LFilter(fl=6e3 , fh=None, n=6) >> self.physiologyPlotHigh)
```

If instead of passing the same data to all downstream nodes, a splitting behavior is required, use the `Split` node between the source and the list of nodes. The code below passes spikes detected from each channel of the recorded signal to separate plots:

```
daqs.physiologyInput \
    >> pypeline.Thread() \
    >> pypeline.LFilter(fl=300 , fh=6e3 , n=6) \
    >> pypeline.SpikeDetector() \
```

```
5       >> pipeline.Split() \
6       >> (self.spikePlot1, self.spikePlot2, self.spikePlot3)
```

2.3.5 glPlotLib: GPU Accelerated Plotting

The Python programming language is enriched with many data visualization libraries that are suitable for various applications. However, when it comes to visualizing neural recordings in real-time, the current tools fail to keep up with the large number of samples across multiple channels that are being continuously acquired. In the initial version of the EARS software, PyQtGraph was employed to create optimized widgets that could partially overcome the limitations of these tools (see `plotting.py`[15]). One downside of choosing this—or any other—approach that depends on the central processing unit (CPU) for graphical processing is that it utilizes the majority of the system's processing power, thus, impeding performance of other tasks. Alongside data visualization, EARS also has to perform heavy input and output operations and insufficient system resources for these tasks often leads to fatal crash of the software.

Most modern computers come with a graphical processing unit (GPU) already packaged with their CPU, or contain a discrete GPU. GPUs are made of many smaller processing units that are capable of performing in parallel, hence, outperforming traditional CPUs in specific tasks such as signal processing and graphical computation. To further improve the plotting widgets with GPU acceleration and allow advanced visualization of detected spikes and response time histograms, the PyOpenGL library was utilized. PyOpenGL has the ability to program GPU cores to run a custom set of instruction in the GLSL language. By taking advantage of this feature, a new set of GPU accelerated plotting widgets are defined within the glPlotLib module. The main repository for this module is located on GitHub[16].

[15]Plotting widgets source code: https://github.com/nalamat/ears/blob/master/plotting.py
[16]glPlotLib source code: https://github.com/nalamat/glplotlib

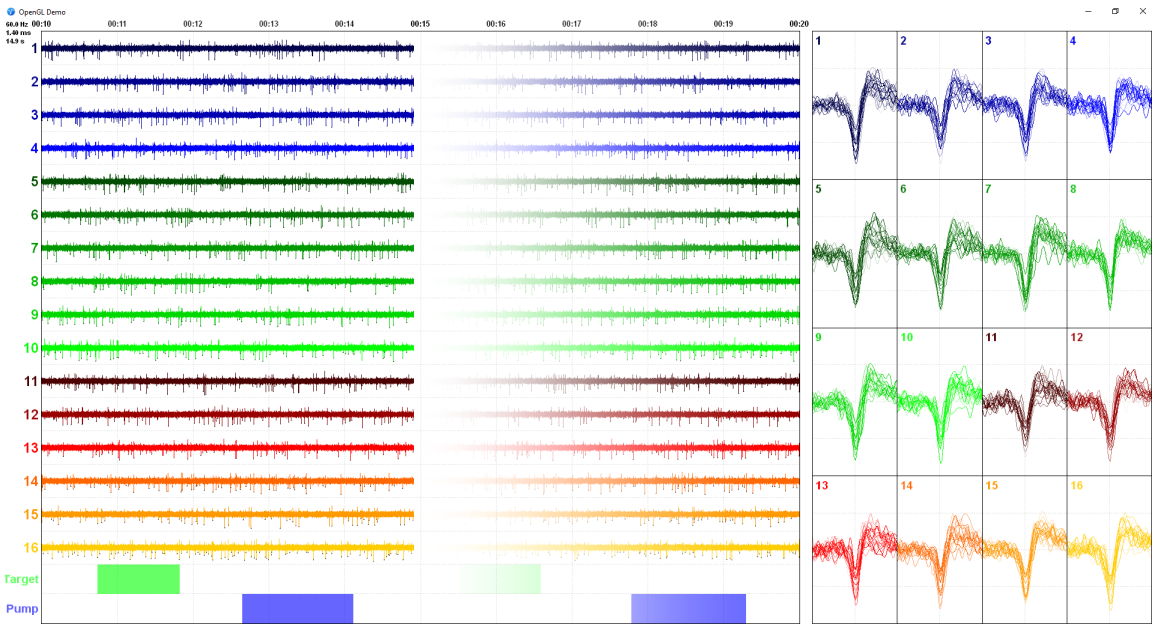

Figure 2.13 GPU accelerated plotting using glPlotLib. The main section of the screen on the left shows the acquired traces of all electrode channels in a wrap-around display similar to an oscilloscope. The horizontal axis represents time over an adjustable time window (currently set to 10 seconds, see labels on the top). Rectangular markings on the bottom show behavioral events, such as presentation of the target stimulus, that are most likely to invoke changes in the neural activity. Circular annotations on neural traces indicate detection of spike events. Enlarged snippets of these events per each electrode channel are also displayed on the left side of the screen.

In Figure 2.13 an oscilloscope like display shows the recorded traces of all electrode channels. With CPU plotting, this display was limited to at most 16 channels. However, GPU accelerated plotting can simultaneously show more than 128 channels, along with annotation of detected spikes and a separate enlarged display of spike snippets. Furthermore, a histogram of the spiking activity relative to the target stimulus onset provides experimenters with more insight regarding the neural activity. This information allow quick decision making during an ongoing experiment without the need for offline processing of the recorded traces.

CHAPTER 3

CORTICAL RELIANCE ON TEMPORAL CODING AT NEGATIVE SIGNAL-TO-NOISE RATIO

Everyday environments often contain multiple concurrent sound sources that fluctuate over time. Normally hearing listeners can benefit from high signal-to-noise ratios (SNRs) in energetic dips of temporally fluctuating background sound, a phenomenon called dip-listening. Specialized mechanisms of dip-listening exist across the entire auditory pathway. An unresolved issue regarding cortical mechanisms of dip-listening is how target perception remains invariant to SNR. The current work tests the hypothesis that at negative SNRs, neuronal readout mechanisms need to increasingly rely on decoding strategies based on temporal spike patterns, as opposed to spike count. Recordings from chronically implanted electrode arrays in core auditory cortex of trained and awake Mongolian gerbils that are engaged in a tone detection task in 10 Hz amplitude-modulated background sound reveal that rate-based decoding is not SNR-invariant, whereas temporal coding is informative at both negative and positive SNRs.

3.1 Background

A signature problem for most hearing-impaired individuals is their reduced ability to dissociate target from background sound. When an auditory target and a background sound source coincide, the background sound can energetically mask the target by swamping or occluding the target's cochlear representation, a problem that even ideal rehabilitative listening devices cannot solve. However, because natural sounds inherently fluctuate over time, they rarely overlap continuously [23]. To segregate competing sound sources into perceived auditory objects, both humans and animals can exploit these temporal fluctuations [80, 9], a phenomenon called "dip

listening." Identifying coding strategies for dip-listening wherein behavioral ability to suppress temporally fluctuating background sound co-varies with target-evoked neuronal activation may thus hold a key for defining treatment targets for hearing loss.

Neurophysiological studies in animal models have greatly advanced our mechanistic understanding of dip listening. Specialized neuronal rate- and temporal coding mechanisms enhance target representations by leveraging high-and-short-term signal-to-noise energy ratios (SNRs) across multiple processing stages, including cochlea [87], cochlear nucleus [104], inferior colliculus [73] and auditory cortex (ACx, [113]). In addition, dip listening can be behaviorally quantified by comparing detection, discrimination or identification thresholds between temporally fluctuating vs. steady maskers. The difference in thresholds, called modulation masking release (MMR), is refractory to training and varies with SNR. At positive SNRs, where target energy dominates the acoustic mixture, MMR is not observed across a range of tasks. Even a coding strategy based on net change in long-term acoustic energy may suffice to detect a target sound. Although, at negative SNRs, where dip-listening can dissociate the target from the acoustic mixture if the masking envelope fluctuates moderately (4-32 Hz, peaking around 10 Hz), MMR increases with decreasing SNR. This suggests that SNR shapes the demands on reliance on short-term temporal processing, thereby modulating how listening in fluctuating background sound operates, a hypothesis tested here. Alternatively, SNR in isolation may not be the major factor limiting dip listening, and task difficulty, which often covaries with SNR, may drive the neuronal coding strategy for listening in fluctuating background sound.

To test this hypothesis, we here study ACx because this is where auditory objects emerge, making ACx a promising target while searching for SNR invariance. Moreover, a mechanism thought to underlie dip listening, envelope locking suppression, reduces the fidelity by which ACx neurons track background sound when a target

occurs, at both positive and negative SNRs, at least in anesthetized animals [89, 73]. Whereas, a mechanistic understanding of how SNR shapes the ACx neuronal code is complicated by the fact that prior neurophysiological work assessing dip listening often uses untrained animals or animals that could not detect target sound at negative SNRs [117, 20]. Here we are interested in SNR invariance of auditory cortical responses in awake gerbil during dip-listening. Gerbils have low-frequency sensitivity and MMR of comparable magnitude as humans [34, 59], making them a suitable model for studying SNR invariance during dip listening.

Using appetitive psychometric testing and chronically implanted recording electrodes, we simultaneously quantify behavioral sensitivity and ACx single-unit responses. Awake animals either actively detect a tone in modulated masking noise, at three SNRs (-10, 0 and 10 dB), or passively hear the same sounds without task engagement. Behavioral accuracy is comparable across SNRs, controlling for task difficulty. In experiment 1, we contrast the effect of SNR on rate vs. temporal coding strategies in sound-detecting animals. We previously discovered that a measure of similarity between the sound-evoked cortical responses when target sound is present vs. absent, mutual information, is smaller for behaviorally relevant sound as compared to sound of no behavioral significance, hinting that behavioral relevance may increase reliance on temporal as opposed to rate coding [83]. To elucidate the role of behavioral relevance, using the same sounds and similar operant conditioning as in experiment 1, except for giving response reward irrespective of decision, in experiment 2, we test non-sound-detecting control animals. Our results support the interpretation that temporal coding can directly serve to detect target sound at across all tested SNRs, whereas the neuronal information carried in rate coding needs a more nuanced readout strategy with differing decision criteria at positive vs. negative SNRs.

3.2 Materials and Methods

All experimental protocols were approved by the Rutgers University Institutional Animal Care and Use Committee.

3.2.1 Housing

Animals were group housed unless warranted by veterinary exemption for post-surgical recovery. Environmental enrichment was provided with soft nesting materials and toys for chewing. The gerbils had unrestricted access to a nutritionally complete diet through food pellets along with unrestricted water prior to training and during water breaks, as well as controlled water during testing. Life-brand cereal was given as treats after each training session. For chronically implanted animals, dietary supplements such as diet gel, hydrogel, and sunflower seeds were provided before and after the surgical procedure. Daily checkups ensured animal health and safety. The vivarium temperature (65-75°F) and humidity ($< 50\%$) were logged and maintained, and the vivarium lights were automatically switched throughout the day to regulate sleep-wake cycle of the animals.

3.2.2 Behavioral Testing

For experiment 1, using an appetitive Go/NoGo paradigm with controlled water access (Figure 3.1A), six adult Mongolian gerbils (*Meriones unguiculatus*) were trained and tested on a tone detection task while fluctuating noise, called *masker*, continuously played in the background throughout each session [59]. Specifically, gerbils detected whether or not a 1 kHz target tone (1 s duration, 50 ms cosine-squared rise/fall, 40-60 dB SPL) was present in 10 Hz rectangularly amplitude-modulated and band-limited background noise (50% duty cycle, 10 ms cosine-squared rise/fall ramp on each rectangular burst, 50 dB SPL, 2/3 octave bandwidth, pseudorandom frozen noise, centered at the target frequency of 1 kHz; Figure 3.1B).

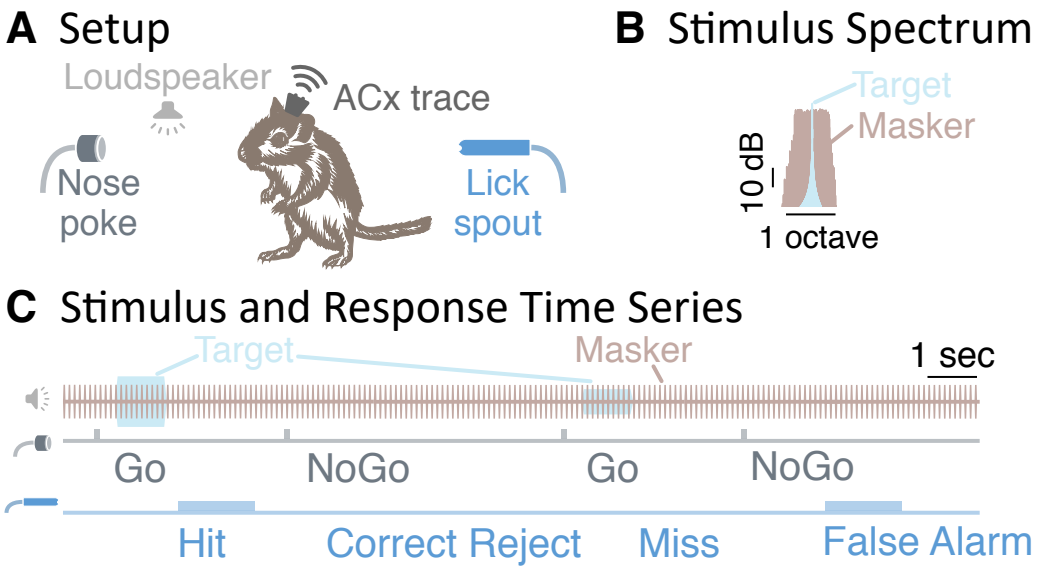

Figure 3.1 Testing apparatus and behavioral design. **A** The test setup included a loudspeaker above the test area, a nose poke and a lick spout. In addition, for chronically implanted animals, a wireless system recorded the cortical traces. **B** The background sound (brown), consisting of 10 Hz amplitude modulated noise, was continuously present. On Go trials, an target sound (blue) was additionally played, consisting of a 1 kHz tone and randomly chosen from -10, 0 or 10 dB SNR. **C** The gerbil triggers a new trial by breaking a light beam inside the nose poke, and could obtain water reward through the lick spout. A loudspeaker, mounted above the test area played the The gerbil then respond to the trial condition either by licking the water spout, or by withholding a response through waiting or by poking the nose poke once more. Depending on the stimulus condition, this response resulted in either a Hit, a Correct Reject, a Miss or a False Alarm.

Gerbils were placed inside a testing cage with a nose poke to initiate or abort a trial and a lick spout to access water (Figure 3.1A). The entire setup was inside a radio frequency shielded sound-attenuating booth with sound-absorption walls (booth dimensions: 8' x 10' x 7'). A loudspeaker was mounted on the booth ceiling, approximately 1 m above the center of the cage. Gerbils were initially trained to associate the 1 kHz target tone with water, then learned to use a nose poke to initiate a target tone (*Go* trial) and finally were taught that on some trials, despite nose poking, no target tone would play and no water reward would be given (*NoGo* trial). Specifically, to initiate a trial, gerbils were trained to hold their nose inside the nose poke for 200 ms and wait for a target stimulus to potentially occur for a potential water reward from the lick spout.

On Go trials, a target was presented in the continuous background sound, whereas on NoGo trials, no target was played and thus only the background could be heard (compare light blue target vs. light brown masking sound traces in Figure 3.1C). On Go trials, gerbils were then supposed to approach a water spout within a 4.2 s period after poke onset for a water reward. Successful contact with the spout during Go trials was scored as HIT. Other responses (either a re-poke in the nose poke or no poke response) were scored as MISS. On NoGo trials, gerbils were supposed to withhold any lick spout response. A withheld lick spout response or a re-poke response to the nose poke were each scored as CORRECT REJECT. However, if gerbils approached the water spout during a NoGo trial, no water was released, a 1-1.5 sec mandatory timeout was given and the response was scored as FALSE ALARM (compare light blue / brown sound, grey nose poke and blue lick spout traces in Figure 3.1C). To discourage guessing, each FALSE ALARM trial was immediately followed by another NoGo trial for a maximum of 15 sequential NoGo trials. For each session, the rate of hit responses (HR) and the rate of false alarms (FAR) was then

used to calculate the behavioral sensitivity, called d′, according to Equation (3.1):

$$d' = \sqrt{2}\, erf^{-1}\left(2\ \mathrm{HR} - 1\right) - \sqrt{2}\, erf^{-1}\left(2\ \mathrm{FAR} - 1\right), \tag{3.1}$$

where erf^{-1} is the inverse error function. Test sessions typically lasted 60 minutes, but varied in duration, depending on the animal's satiety and willingness to perform the task. Averaged across sessions and animals, animals typically performed approximately 108 Go trials and 62 NoGo trials per session. To computationally eliminate the possibility of infinite d′, both (HR) and (FAR) were conservatively bracketed via thresholding such that any observed rate below 0.05 was assumed to equal 0.05, and any rate exceeding 0.95 was assumed to equal 0.95. Those thresholds were chosen assuming one guessed trial per SNR per session.

For both training and testing, the masker was fixed throughout each session. To ensure a fixed phase delay between target and masker, after initiation of a trial via nose poke, the target onset was delayed until the next 45° phase of the masker occurred. As a result, the target onset could occur between 250 to 350 ms after onset of a nose poke. This time delay arose as the sum of the 200 ms hold duration for the nose poke, a fixed 50 ms delay imposed due to limitations of the recording system for updating the acoustic output, and a 0 to 100 ms phase delay, depending on the phase when the animal initiated the nose poke during the 100 ms masker cycle.

All animals were initially trained with a masker level of 20 dB SPL and a target level of 75 dB SPL (55 dB SNR), until they reached criterion performance (FAR below 30% and d′ above 1.5). Across sequential sessions, the masker was then gradually raised in steps of 10 dB, to a final masker level of 50 dB SPL (25 dB SNR). In the final training stage, target sounds at additional, softer intensities were gradually added in 5 or 10 dB steps across sessions, resulting in a total of three SNRs. The experimenter's decisions to increase the masker or decrease the target levels was guided by whether or not the animal reached criterion performance at all tested SNRs

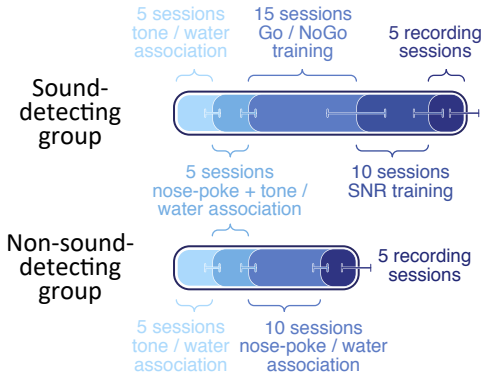

Figure 3.2 Average number of sessions per each training and testing stage shown as shaded progress bars for each of the two gerbil groups. The lower and upper bounds of session counts are indicated with error bars.

and how confident the animal appeared in the task. All gerbils completed training within 30 sessions or less. At the conclusion of training, all gerbils were able to reliably detect tones at -10 dB, 0 and 10 dB SNR.

For testing, the masking noise was continuously present in the background at 50 dB SPL. Three target intensities (40, 50 and 60 dB SPL) were randomly interleaved from trial to trial, resulting in SNRs of -10 dB, 0 dB and 10 dB. In addition NoGo trials, without target energy, were randomly interleaved (SNR $= -\infty$ dB) After animals completed at least three sessions with d′ of 1.9 or better at all SNRs, four of these sound-detecting gerbils advanced to the neural recording stage.

The two other sound-detecting gerbils were advanced to a separate behavioral test, to measure the minimum tone duration needed to perform the tone detection task. These two gerbils were behaviorally tested on tone detection at 0 dB SNR (50 dB SPL target level and 50 dB SPL masker level) as a function of target tone duration, across 13 sessions. Specifically, in each test session, seven different target tone durations (50, 100, 200, 400, 600, 800, 1000 ms) were randomly interleaved from trial to trial and the gerbil's behavioral responses were recorded.

To control for the effects of arousal in the neural recordings, two additional gerbils were trained for experiment 2, using a non-sound-detecting Go/NoGo

paradigm. Similar to the sound-detecting gerbils in experiment 1, these two gerbils were initially trained to associate the target tone with water, and then to trigger target tones by using the nose poke. Unlike the sound-detecting gerbils, the non-sound-detecting gerbils then moved to a different training stage where water reward was only contingent on initiating a trial via nose poke and on reaching the water spout on time. Therefore, both Go trials (when target and background sound was played) and NoGo trials (when only the background sound was played) could result in water rewards. Once these two control animals reached a minimum 95% (HR) and (FAR), indicating that they did not behaviorally discriminate between Go and NoGo before approaching the lick spout, they were advanced to the neural recording stage.

3.2.3 Surgical Procedure

For the neural recording stage, we chronically implanted recording electrodes into left ACx [16, 17, 134]. Specifically, gerbils were initially anesthetized with 4% isoflurane, 1.5 mg/kg ketoprofen, 0.35 mg/kg dexamethasone, and continuously given 1-3% isoflurane to maintain sedation. Using stereotactic coordinates, with the medial-rostral corner at 5 mm lateral and 5 mm rostral to the λ landmark, a craniotomy was performed (approximate extent 1mm x 2mm). Using a 25 degree approach angle relative to the vertical axis, the targeted implantation site was then ascertained at 4.6 lateral and 3.4 mm rostral to λ. At the targeted site, a durotomy was performed before a silicon microelectrode array with 16 channels (A4x4-4mm-200-200-1250-H16-21mm, NeuroNexus Technologies, Inc.) was lowered into the brain at an initial insertion depth of 1.3-1.5 mm from the surface of the brain. A custom-made microdrive held the array in place inside a recording chamber, enabling post surgical advancement of electrodes, deeper into the brain tissue. The

head-post along with the microdrive and implanted array were fixed on the skull using 4-6 bone screws and two layers of dental acrylic.

3.2.4 Recording System

Trained gerbils were tested while cortical potentials were simultaneously recorded, amplified and transmitted wirelessly to a receiver positioned approximately 1 m from the cage (W16, Triangle BioSystems International). All input/output channels were synchronized at 100 kS/s and 16 bits by sharing the sampling clock pulse of two data acquisition cards (DAQs, PCIe-6321 and PCIe-6341, National Instruments Corporation) via a Real-Time System Integration bus cable. Specifically, custom written software, called Electrophysiology Auditory Recording System (EARS), synchronously controlled auditory stimuli delivery and recorded both behavioral and physiological responses [1]. EARS communicated with the loudspeaker, nose poke, licks spout, W16 and a personal computer for data storage and analysis via the two DAQs. In addition, EARS interfaced with a syringe pump (NE-1000 Programmable Single Syringe Pump, New Era Pump Systems, Inc.) via a USB-RS232 emulator.

Audio Delivery Auditory stimuli were generated in EARS, D/A converted and preamplified (E 12:2, Lab.gruppen) before being sent to the loudspeaker (DX25TG59-04 tweeter, Tymphany HK Ltd). Sound calibration was initially performed by playing 18 bits long maximum length sequence (MLS) twice in a row, recording the response with a sound level meter placed in the center of the cage (Brüel & Kjær 2250), and inverting the cross-correlation between the second portion of the MLS and microphone recording to generate a pre-amplification audio filter that flattened the speaker response. Periodic re-calibration verified acoustic integrity of the recording system with +/-2 dB precision from 0.5 to 8 kHz.

Data Acquisition Furthermore, custom printed circuit boards, connected to nose poke and lick spout, drove infrared emitter diodes (LTE-302, LITE-ON Technology Corporation) and their paired photosensors (OPS693, TT Electronics Plc), and conditioned these optical channel responses before sending them to their appropriate DAQ digital input channels. In addition, 15 analog input channels collected cortical potentials from the wireless receiver at a sampling frequency of 31.25 kS/s per channel.

3.2.5 Behavioral-Cortical Assessment

On each recording day, cortical activity in animals was recorded twice, once during active task engagement mode, and later while the animal passively heard similar stimuli in a randomly different order of presentation. At the end of each recording session, electrodes were manually advanced by turning an advancing screw on the custom-made microdrive by approximately 40 μm.

3.2.6 Analysis

Raw recorded cortical traces along with their associated behavioral events were analyzed offline in MATLAB (R2017a, The MathWorks Inc), by subtracting from each recording channel the grand mean across all channels and band-pass filtering this bias-corrected trace (zero-phase Butterworth, 6th order, 300-6000 Hz). Next, the resulting band-limited trace were time-windowed around the response time interval. Specifically, response time intervals were defined as starting at -1 s before potential target onset and ending +1 s after potential target offset, with the next 45° phase of the masker after the nose poke onset denoted as 0 s. In other words, 0 s marks the time point when the acoustic target onset occurred during Go trials, or the time when the target would have started during NoGo trials. Next, spike events were detected as negative peaks exceeding a threshold of $4.8\sigma_n$, where σ_n was the estimated standard

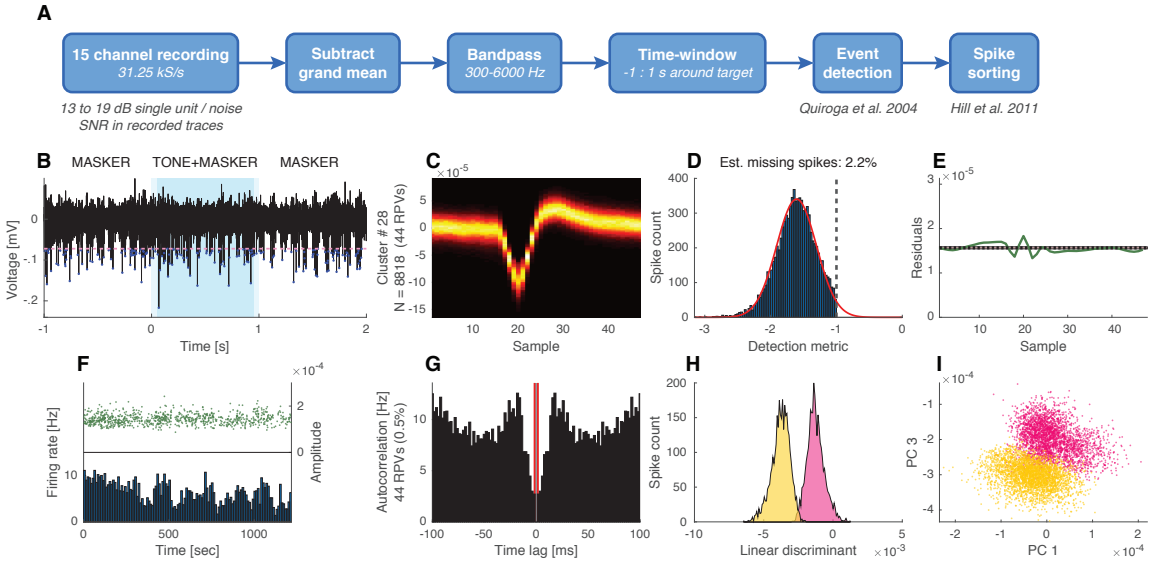

Figure 3.3 **A** Recorded trace analysis pipeline. **B** A sample of a time-windowed recorded trace. Detection threshold is shown with a grey dashed line, and detected events are marked with blue points. **C-G** Metrics of a representative cluster after automatic spike sorting and visual inspection. **H-I** Separation of two sample clusters with linear discriminant and in principle component space.

deviation of the background noise [105]. All events with amplitudes exceeding $30\sigma_n$ were rejected as artifacts.

The extracted event waveforms were further processed into putative units with an automatic spike sorting algorithm, using Principal Component Analysis (PCA) and k-means clustering algorithm (UltraMegaSort 2000, [52]). Visual inspection of each putative unit then verified that the shape of the unit's waveform conformed to a time series typical for an action potential, that the firing rate of this unit was stable throughout the session, that there were infrequent refractory period violations (less than 1%), and that each putative unit was separated in at least one principal component space plot or with the best linear discriminant.

Statistical Comparisons Units that fulfilled all criteria were labelled as single-unit and further analyzed using seven rate and temporal coding metrics defined below.

To assess whether rate vs. temporal response codes varied with SNR and across active and passive listening conditions, repeated measures or mixed design analysis of variance (rANOVA or mANOVA) was implemented with the *rstatix* package version 0.7.0 [68] in R version 4.0.3 [106]. A significance level of $\alpha = 0.001$ was consistently used throughout the analysis. Whenever the sphericity assumption for within-subjects factors was violated (based on Mauchly's test), Greenhouse-Geisser correction was applied to the results. In addition, behavioral psychometric functions of (HR) were estimated using generalized linear models (GLM, [106]), assuming binomial response distributions (link function: logit), and weighing HR proportionally to the trial count of each stimulus condition.

Response Time Histograms For each unit and SNR during Go trials, as well as for NoGo trials, response time histograms (RTHs) were calculated with 50 ms resolution, by first binning spike events into 10 ms time-windows and then convolving the resulting response probability densities with a 50 ms rectangular kernel. This resolution was chosen because the slowest time-window to capture envelope-related acoustic responses is 50 ms, corresponding to the Nyquist rate of the 10 Hz masking envelope.

Firing Rate For each single-unit and SNR, firing rates were calculated by counting the number of spike events during incremental time-windows in steps of 50 ms starting at the target onset, and dividing it by the length of the window in seconds. The across-trial average and standard deviation of the resulting firing rates were then used to calculate the separation between Go and NoGo responses. Specifically, for each SNR, neurometric rate z-scores were calculated as a function of time, according to Equation (3.2):

$$z = \frac{\mu_{\mathrm{Go}} - \mu_{\mathrm{NoGo}}}{\frac{1}{2}\left(\sigma_{\mathrm{Go}} + \sigma_{\mathrm{NoGo}}\right)}. \tag{3.2}$$

Specifically, for the rate z-score, μ_{Go} and σ_{Go} were the average firing rate and standard deviation of firing across Go trials for each target SNR, and μ_{NoGo} and σ_{NoGo} were the average and standard deviation of firing rate for the NoGo trials.

Power Spectral Density To estimate the spectral density of the single-unit responses, point process multi-taper spectrum (MTS) analysis was derived from the spike event times and pooled across trials [10]. Sparse single-unit events impede reliable MTS estimates at the single-trial level, a limitation solved here by sampling, for each SNR, half of the trials randomly with replacement, across 20 repetitions, using Equation (3.2).

Vector Strength To calculate the strength by which single-units followed the masker envelope for each individual trial, using Equation (3.3), we calculated the vector strength (VS) at 10 Hz.

$$\begin{aligned} \theta_i &= 2\pi \left(t_i f_{\text{base}} - \lfloor t_i f_{\text{base}} \rfloor \right), \\ \text{VS} &= \frac{1}{n} \sqrt{\left(\sum_{i=1}^{n} \cos\theta_i \right)^2 + \left(\sum_{i=1}^{n} \sin\theta_i \right)^2}, \end{aligned} \tag{3.3}$$

where t_i is the spike time relative to target onset, f_{base} is the base frequency at which VS is calculated for (ranging from 1 to 20 Hz in steps of 1 Hz), θ_i is the phase of each spike, and n is the total number of spikes (across trials, per target level). Using Equation (3.4) to approximate the p-value for Rayleigh's test for uniformity [137], across all NoGo trials and for each SNR during the Go trials, we then ascertained whether the VS to the masking envelope differed from zero.

$$p = \exp\left[\sqrt{1 + 4n + 4n^2 \left(1 - \text{VS}^2\right)} - (1 + 2n) \right]. \tag{3.4}$$

Target-Evoked Decorrelation Response To quantify how much the temporal response pattern of single-unit responses changed when a target tone was added to

the background sound vs. when just the background sound was present, for each SNR and single-unit, we compared RTHs across Go vs. NoGo trials. Specifically, Pearson's correlation coefficient ρ, between Go and NoGo RTHs were calculated in running time-windows of 300 ms duration, sliding across the full duration of the response interval from -1 s to 2 s, with a step size of 10 ms. The 300 ms duration was chosen conservatively based on the results of the behavioral tone duration thresholds (see Figure 3.5). A high ρ indicates a high trial-to-trial similarity in the single-unit RTH between the masker-only response during NoGo trials and the target-and-masker-combined response during Go trials. In addition, target-evoked decorrelation responses were derived by calculating the ρ z-score between the masker-only RTH immediately preceding the target onset vs. the early response during the response time-window (Equation (3.5), where μ_{onset} and σ_{onset} were the average and standard deviation of ρ during the first 300 ms of target response, and μ_{poke} and σ_{poke} were the average and standard deviation of ρ within 300 ms of the nose poke events).

$$z_{\text{decorrelation}} = \frac{\mu_{\text{onset}} - \mu_{\text{poke}}}{\frac{1}{2}\left(\sigma_{\text{onset}} + \sigma_{\text{poke}}\right)}. \tag{3.5}$$

Sustained Response To quantify how separable the neural responses were between Go and NoGo trials, we analyzed the sustained-response interval, excluding onset and offset responses. Specifically, from each unit's response 100 ms after tone onset to 100 ms before tone offset we calculated seven metrics. We call the difference in sustained rates between Go and NoGo in z-units the $z_{sustained}$.

Mutual Information Under the assumption that during the sustained response interval neural firing follows a stationary independent identically distributed Poisson process, we estimated λ parameters of the underlying Poisson distributions. λ was used to approximate mutual information [83] as an agnostic measure for comparing response similarity. Mutual information ranges from 0 to 1, with higher values

indicating low similarity between the Go and NoGo response, and hence, higher transfer of information over the target-evoked response channel. Conversely, lower mutual information can be interpreted as sparsity of the neural code in conveying information.

Similarity Index To quantify the overall similarity of the sound-evoked responses to the mixture of target and background sound vs. just the background sound, for each single-unit, a similarity index was calculated using a metric defined in prior work [73]. Specifically, for each SNR, the time series of each unit's Go and NoGo RTHs during the sustained period were plotted against each other and the resulting scatter plots were fitted using linear regression. The slope of these regression lines is called similarity index [73]. A similarity index of 1 indicates that when a target occurs, the response to the mixture of target and background sound is similar to the response to just the background sound, whereas a zero similarity index indicates that the presence of target energy strongly changes how a unit responds.

3.3 Results

3.3.1 Sound-Detecting Gerbils

Across all SNRs, the four implanted sound-detecting gerbils could reliably detect the target tone (solid line at or above d'=1.9 in Figure 3.4), an ability that subtly improved with increasing SNR [rANOVA: $F(1.46, 26.21) = 11.543$, $p < 0.001$, $\eta^2_G = 0.18$], consistent with our prior work [59]. To reach performance of d'= 1.9 or better, the target tone needed to be at least 380 ms long, as estimated from tone detection thresholds at 0 dB SNR as a function of tone duration in two additional, non-implanted gerbils (see Figure 3.5 and results of GLM fit in Table 3.1). Analysis revealed that behavioral sensitivity asymptotes at approximately 500-600 ms.

A total of 151 single-units in the freely moving, awake and sound-detecting gerbils showed target-evoked responses in the presence of masking noise, as evaluated

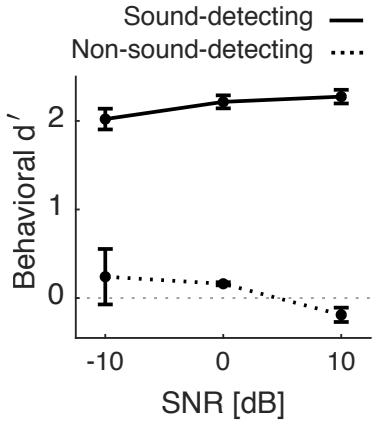

Figure 3.4 Average behavioral performance of each group of gerbils across the recorded sessions with active engagement (test n=4, control n=2, sessions=5 each on average). Error bars show one SEM.

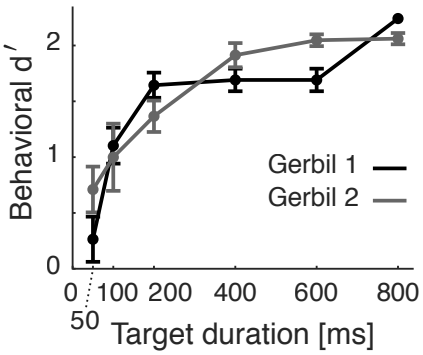

Figure 3.5 Behavioral curves at 0 dB SNR as a function of tone duration for non-implanted gerbils (n=2, sessions=13 each on average). Error bars show one SEM.

Table 3.1 GLM Results for Behavioral Performance vs. Tone Duration

	Description	Estimate	Std. Error	z-value	p-value	
β_0	Intercept	-0.429	0.047	-9.131	< 0.001	***
β_1	Tone duration	9.041	0.376	24.026	< 0.001	***

Model: HR $= \text{logit}^{-1}\left(\beta_0 + \beta_1 \times \text{Tone duration}\right)$.

Deegrees of freedom: 143.

Significance codes: '***' $p < 0.001$, '**' $p < 0.01$, '*' $p < 0.05$, '.' $p < 0.1$, ' ' $p \geq 0.1$.

offline by visually comparing Go vs. NoGo RTHs. Therefore, no units were excluded from further statistical analysis. First-spike response latencies, shown in Figure 3.6, were consistent with those typically observed in primary auditory cortex of gerbil [110]. During NoGo trials, when only the masking sound was played, RTHs robustly phase-locked to the masker envelope, both in the active and in the passive conditions (light red and light grey lines in Figure 3.7 A and B track the 10 Hz acoustic envelope of the masker). In the passive conditions, overall firing rate did not appreciably vary across time. However, in the active conditions, the firing rate was modulated by the animal's behavior, increasing by 21.8% for 200 ms (or two masker cycles) after the onset of the nose poke ([paired $t(151) = 6.2$, $p < 0.001$]; note how the black lines are steady prior to the tone onset in Figure 3.7A, whereas the red lines rise above baseline after the nose-poke in Figure 3.7B), followed by suppression until target tone onset.

In addition, during Go trials, across-unit average RTHs showed target-evoked onset and offset responses (see darker lines around the blue-shaded target time-windows in Figure 3.7 A and B). Onset enhancement occurred at all SNRs. Offset enhancement was pronounced at 10 dB SNR, but weak or absent at 0 and -10 dB SNR. During the sustained portion of the target sound, from 100 to 950 ms, suppression

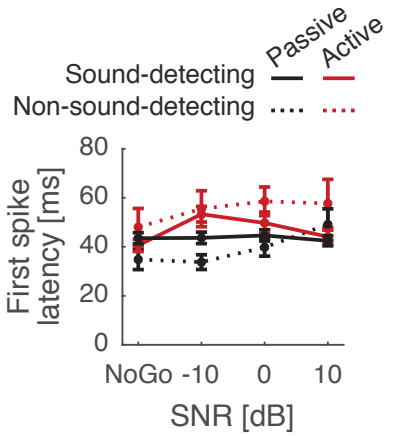

Figure 3.6 Average first-spike latency of all single-units. Error bars show one SEM.

occurred at 10 dB SNR, vs. enhancement at 0 dB and -10 dB SNR, relative to the response to the masker alone (dark lines fall below zero in Figure 3.7 C and D, whereas the lighter lines stay positive throughout the sustained time interval). Despite these non-monotonic changes in sustained firing rate between Go and NoGo trials, power spectral density analysis confirmed the strength by which the single-units tracked the masker *monotonically* decreased with increasing SNR (Figure 3.7 I and J show sharp peaks near 10 Hz, peak height increasing with decreasing SNR).

Figure 3.8A shows the average sustained firing rates vs. vector strengths at 10 Hz for all single-units, with large symbols denoting units with vector strengths that were significantly greater than 0, per Rayleigh's z-test. Linear regression analysis reveals that at -10 dB SNR and during NoGo trials, units that tended to phase-lock more strongly to the envelope of the masker also tended to have lower firing rates in the passive conditions (notice the near absence of data points with high firing rate and high vector strength in Figure 3.8A), accounting for approximately 15% of the variance in the data. In contrast, in the active conditions, and in the passive conditions at 0 dB and 10 dB SNR, average sustained firing rates and vector strengths were not significantly correlated. Figure 3.8B shows for both active and passive conditions, vector strength decreased monotonically with increasing SNR [rANOVA

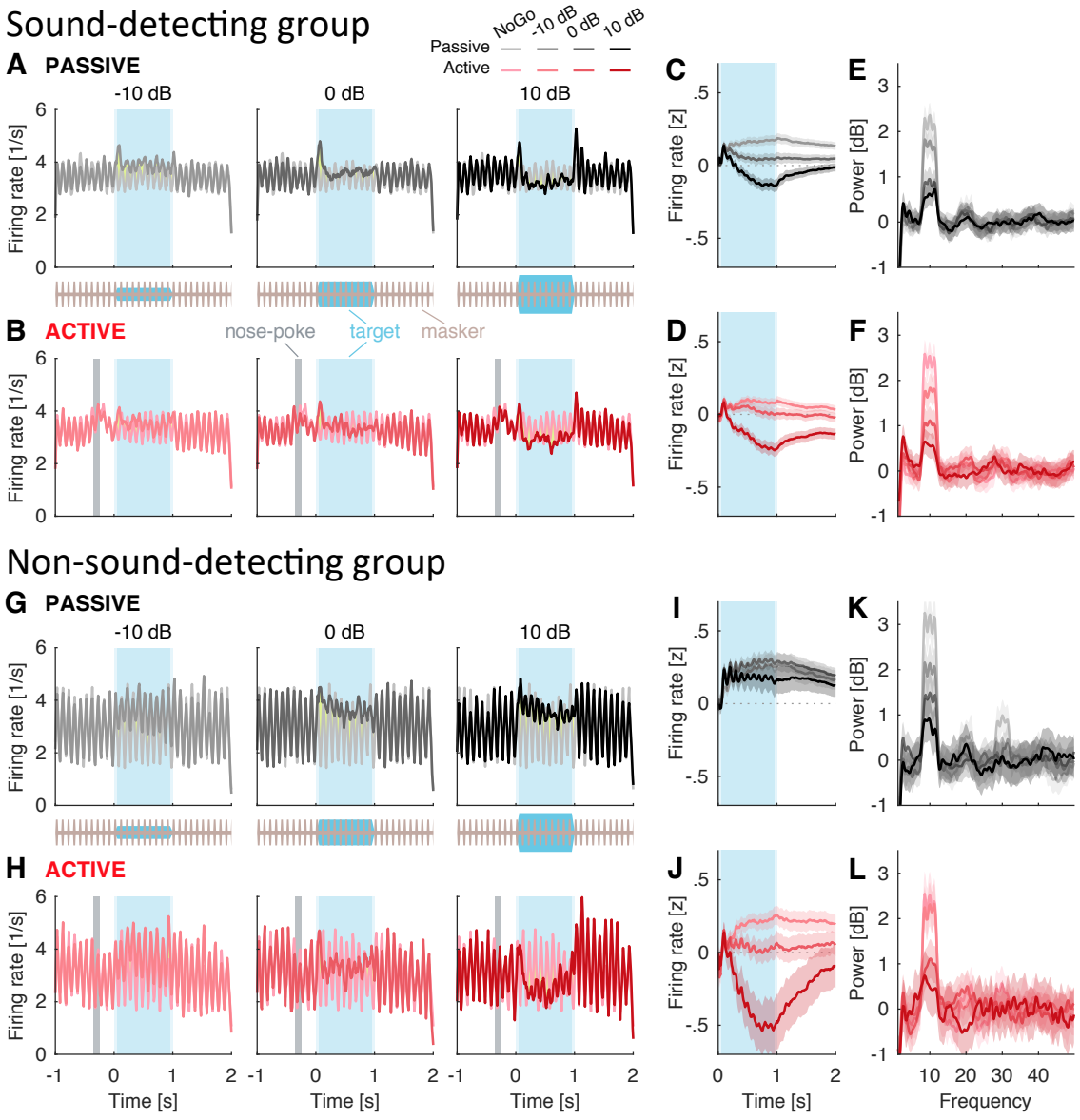

Figure 3.7 **A,B,G,H** Response time histogram for sound-detecting (n=4, units=151) and non-sound-detecting gerbils (n=2, units=56). **C,D,I,J** Firing rate z-score of the neural response as a function of time, calculated in incremental windows relative to target onset. **E,F,K,L** Power of spectral density of the neural activity calculated with MTS at different frequencies. Ribbons indicate one standard error of the mean (SEM).

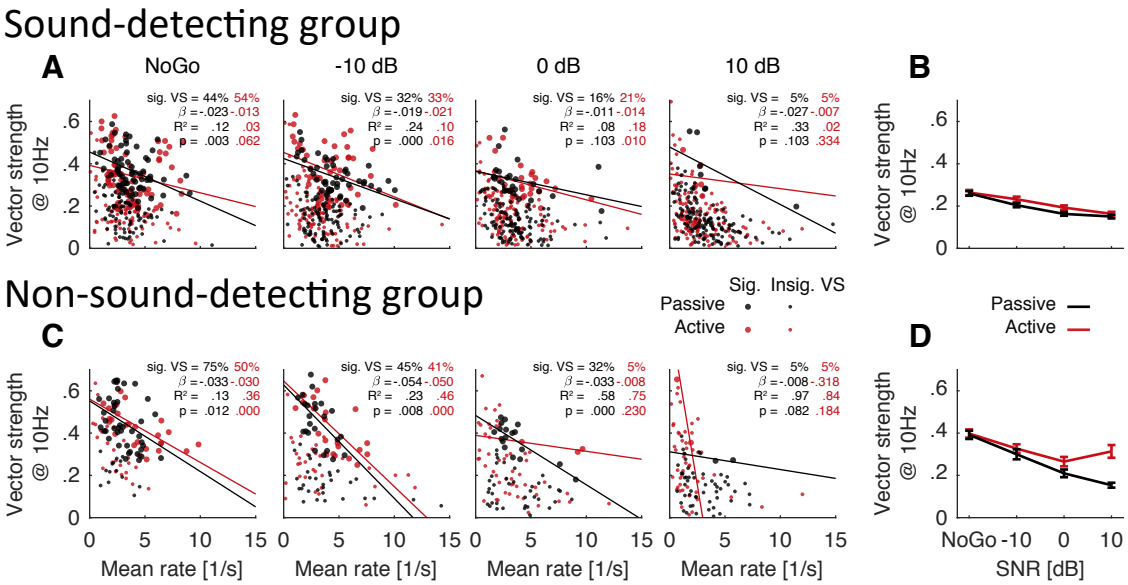

Figure 3.8 **A,C** Vector strength vs. mean firing rate of all units during the sustained response period. Units with a statistically significant vector strength are indicated with larger points. Percentage of these units can be seen at the corner of each plot, along with the details of regression fits over these points. **B,D** Vector strength during the sustained response period as a function of SNR. Here only phasic units are included which have at least one significant vector strength across all SNRs and task engagement conditions. Points mark the average, and error bars show one SEM.

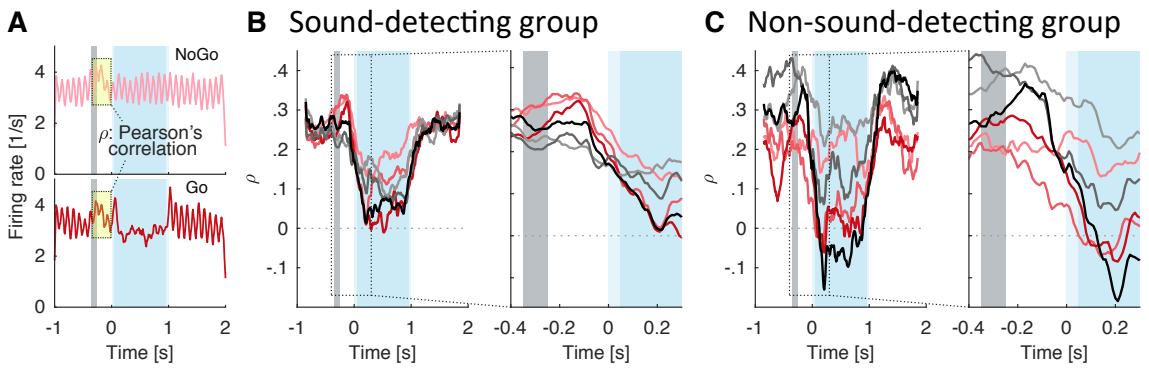

Figure 3.9 **A** Schematic calculation of Pearson's correlation ρ of neural responses for a 300 ms time-window starting at nose-poke. Firing rates at matching time points are correlated between each Go condition and the NoGo. **B,C** Running Pearson's correlation of neural responses during Go trials in reference to NoGo. Colors follow the same code as Figure 3.7.

F(2.4, 350.2) = 53.2, p < 0.001, η^2_G = 0.112], with no appreciable differences in vector strengths between active vs. passive listening [rANOVA F(1, 145) = 5.5, p = 0.020, η^2_G = 0.004].

Next we wondered whether target-evoked decorrelation ρ could be a reliable cue for detecting target sound at both positive and negative SNRs. This metric is a similarity measures that captures whether temporal information changes when target sound occurs, relative to just having background sound. Across both listening conditions and all SNRs, the target onset sharply reduces ρ (Figure 3.9B, left panel), showing that in the absence of target energy, single-unit responses are self-similar across trials. Note that this analysis includes all single-units. Even single-units with low to absent 10 Hz vector strength, i.e., units that do not appear to track the masker envelope, robustly and repeatably show this trend (see Supplemental Information). In addition, in the active conditions, the sharp reduction in ρ is preceded by a small but consistent increase in ρ (see time-expanded plot in Figure 3.9B, right panel), suggesting that after the animal initiates as trial via the nose-poke, the single-unit responses become more sharply tuned to the background sound only, before the

addition of target sound decorrelates the temporal structure in their responses relative to just the presence of background sound.

3.3.2 Non-Sound-Detecting Gerbils

The passive conditions were always recorded following active testing on the same day, after the animal had reached satiety. One striking observation from the sound-detecting gerbils above is the high similarity in the single-unit responses between active vs. passive conditions. This raises the possibility that in the nominally passive listening conditions, these highly trained were, in fact, not ignoring the sounds reaching their ears, even though they were not actively seeking rewards. As a control for this caveat, we next tested gerbils that had been conditioned to expect a water reward simply for initiating the nose poke, even when no tone was playing, i.e., both Go and NoGo trials were always rewarded. Note that while normally-hearing gerbils can be trained to reliably detect target sound the current task even at -20 dB SNR over a dozen training sessions, naive gerbils typically cannot perform the task at negative SNRs [59].

In this non-sound-detecting control group, we isolated 56 target-responsive single-units. Visual inspection of the RTHs reveals target-evoked onset or offset responses at 10 dB SNR, but much reduced or absent onset/offset responses at 0 and -10 dB SNR (Figure 3.7 G and H). In the passive conditions, the presence of target energy consistently increased the firing rate as compared to background-sound-only (all line falls above zero in Figure 3.7I). This response pattern was qualitatively different from the sound-detecting group (compare Figure 3.7 C vs. I). Not only were the firing rate differentials between Go and NoGo consistently positive, but they also fluctuated at 10 Hz (note how all lines are positive in Figure 3.7I, and fluctuate at 10 Hz). In the active conditions, the difference in firing rates between Go and NoGo sustained responses decreased with increasing SNR, a difference that interacted with

SNR similar to results from the sound-detecting gerbils (compare Figure 3.7 J vs. B). Specifically, Go firing rates were suppressed relative to NoGo firing rates at 10 dB SNR (dark red line falls below zero in Figure 3.7D), but enhanced relative to NoGo firing rates at 0 and -10 dB SNR (lighter red lines fall above zero in Figure 3.7J). Of note, unlike in the sound-detecting gerbils, firing rates did not increase after the nose poke onset and even slightly decreased by 2.6% [paired $t(56) = -3.6$, $p < 0.001$)].

Power spectral density was tuned to 10 Hz, and the 10 Hz peak height decreased with increasing SNR, consistent with the idea of tracking the temporal envelope of the masker (Figure 3.7 K and L). In the passive conditions, 10 Hz vector strength and mean firing rate co-varied significantly, except at 10 dB SNR, with lower rate units more likely to significantly phase-lock to the masker envelope (Figure 3.8C). In the active conditions, this trend was also apparent, but only for NoGo and -10 dB SNR cases, where the non-sound-discriminating gerbils presumably could not hear the target. Similar to the sound-detecting gerbils, here, across-unit average vector strength decreased with increasing SNR in the passive conditions (Figure 3.8D). On the other hand, in the active conditions, vector strength plateaued between 0 and 10 dB SNR. Overall vector strength was approximately 0.1 higher in the non-sound detecting as compared to the sound-detecting group (compare vertical ranges across Figure 3.8 B and D, [mANOVA $F(1, 200) = 68.2$, $p < 0.001$, $\eta^2_G = 0.103$]).

Target-evoked response correlations ρ were considerably more variable across conditions, as compared to ρ in the sound-detecting gerbils (compare vertical spread of curves in Figure 3.9 B vs. C). However, even in the non-sound-detecting gerbils, ρ consistently decreased in the presence of target sound. Of note, unlike for the sound-detecting gerbils, nose poke initiation was not followed by increased ρ (red lines do not increase to the right of the grey bar in the time-dilated curve of Figure 3.9C, right panel).

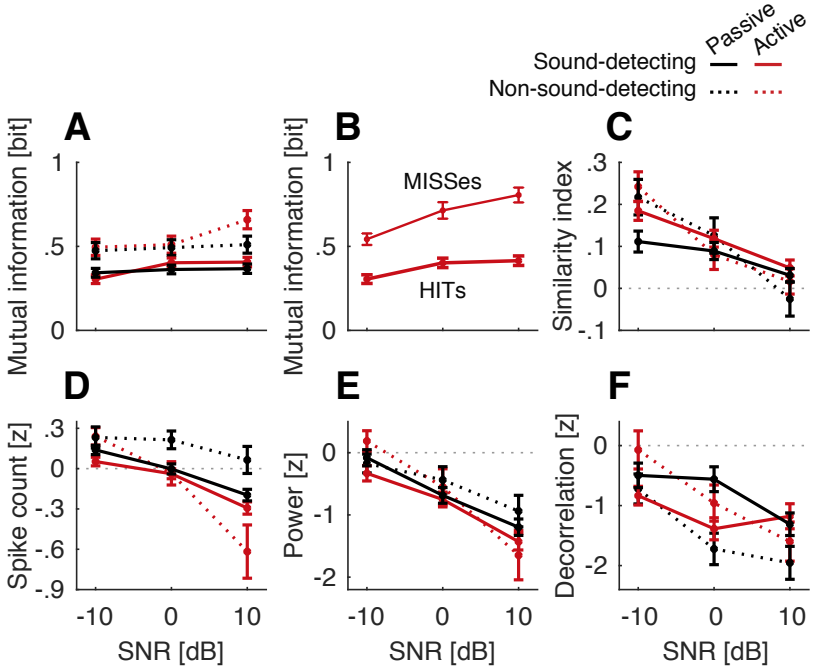

Figure 3.10 Average of all neurometric rate and temporal measures that were calculated relative to NoGo for each gerbil group and task engagement condition. These metrics were derived for the sustained response period. **A** Mutual information. **B** Mutual information for the sound-detecting gerbils during active task engagement, separated by HIT and MISS trials. **C** Similarity index calculated as slope of a linear regression fit. **D** Spike count z-score. **E** Power spectral density z-score at 10 Hz derived with MTS. **F** Target-evoked decorrelation response z-score. Error bars show one SEM.

3.3.3 Neurometric Analysis

To test our core hypothesis that SNR shapes reliance on short-term temporal processing, we next analyzed how rate vs. temporal coding strategies varied with SNR by comparing NoGo vs. Go responses (Figure 3.10). Mutual information between NoGo and Go spike probability distributions, shown in Figure 3.10A, did not appreciably vary with SNR [mANOVA $F(1.9, 368.7) = 6.9$, $p = 0.001$, ${\eta^2}_G = 0.009$], consistent with the observation that gerbils can be trained to hear the target at all SNRs. Sound-discriminating gerbils had much reduced mutual information, as compared to non-sound discriminating gerbils [mANOVA $F(1, 191) = 31.6$, $p < 0.001$, ${\eta^2}_G = 0.044$]. There was no overall effect of active vs. passive task engagement [mANOVA $F(1, 191) = 0.9$, $p = 0.356$, ${\eta^2}_G < 0.001$]. Meanwhile, mANOVA revealed a crossover interaction of SNR and engagement [mANOVA $F(2, 382) = 3.1$, $p = 0.048$, ${\eta^2}_G = 0.004$], consistent with the observation that for both groups of gerbils, the mutual information was somewhat higher during active vs. passive listening at +10 dB SNR.

The association of reduced mutual information and better task performance was also born out when looking at only the Go trials in the active conditions. Figure 3.10B shows a subset of the data in Figure 3.10A, specifically: the active sound-detecting data separated into HITs, i.e., Go trials when sound-detecting animals correctly detected the target, vs. MISSes, i.e., trials when sound-detecting animals fail to detect a target during a Go trial. The mutual information between NoGo and Go spike distributions during HITs was lower as compared to MISSes [rANOVA $F(1, 22) = 45.5$, $p < 0.001$, ${\eta^2}_G = 0.268$]. The overall effect of SNR remained significant [rANOVA $F(2, 44) = 7.8$, $p = 0.001$, ${\eta^2}_G = 0.091$]. The interaction between mutual information score and SNR was significant, consistent with the observation that at +10 dB, the difference between MISSes and HITs was greatest overall [rANOVA $F(2, 44) = 3.5$, $p = 0.039$, ${\eta^2}_G = 0.037$].

Therefore, the current results show, somewhat paradoxically, that when gerbils were better able to hear the tone, the Go and NoGo responses were more similar to each other as compared to when tone detection performance was worse (HITs vs. MISSes or Sound-Detecting vs. Non-Sound-Detecting gerbils). However, note also that the mutual information metric is agnostic to the readout mechanism employed by the central nervous system. The results are consistent with the interpretation that task training prunes information in the neuronal code, removing neuronal discharge patterns that could be informative because they vary consistently with the presence of the target but that are presumably not helpful to the sensory decoding mechanisms that the animal uses.

Figure 3.10C shows the similarity index, i.e., the slope of the regression fit linking Go target-background-mixture responses with NoGo responses to the background sound alone [73]. Here, the similarity index did not differ appreciably across groups [mANOVA $F(1, 200) = 0.1$, $p = 0.769$, ${\eta^2}_G < 0.001$], or active/passive listening [mANOVA $F(1, 200) = 0.8$, $p = 0.371$, ${\eta^2}_G < 0.001$], but decreased with increasing SNR across all conditions [mANOVA $F(1.8, 366.9) = 45.8$, $p < 0.001$, ${\eta^2}_G = 0.061$]. This is consistent with the interpretation that the Go - vs. NoGo-evoked responses become increasingly similar with decreasing SNR. The sign of the similarity index is positive across all conditions, suggesting that an SNR-invariant decision metric based on overall similarity could underlie target detection.

The similarity index combines both rate cues and temporal information. To disentangle how these cues drive behavioral task performance across positive and negative SNRs, we next separately analyzed putative decoding mechanism relying on spike count vs. spike timing. In both groups, here, the Go-NoGo differences in spike count decreased with increasing SNR (Figure 3.10D; mANOVA $F(1.6, 321.1) = 126.1$, $p < 0.001$, ${\eta^2}_G = 0.123$), and the magnitude of the Go-NoGo difference was greater in the active than in the passive conditions [mANOVA $F(1, 196) = 35.9$, $p < 0.001$, ${\eta^2}_G$

= 0.033]. There was no main effect of group [mANOVA F(1, 196) = 3.0, p = 0.087, η^2_G = 0.007]. Across all configurations, except for the passive non-sound-detecting configuration, the spike count curve as a function of SNR crossed zero, showing that at -10 dB SNR, more spikes occurred during Go than NoGo trials, whereas at 10 dB, NoGo trials had a higher spike count than Go trials (In Figure 3.10D, all lines except for dashed black line cross zero). Therefore, a putative neuronal readout relying on Go-NoGo spike count distances would need to use different decision strategies at positive vs. negative SNR, detecting target sound via a decrease in sustained firing at positive SNR, as compared to increased sustained firing at negative SNRs.

In contrast to the spike count, putative temporal readouts varied monotonically with SNR, without crossing zero. Specifically, both the 10 Hz modulation power (Figure 3.10E) and the ρ-index (Figure 3.10F) decreased with increasing SNR [Power: mANOVA F(1.7, 300.7) = 73.4, p $<$ 0.001, η^2_G = 0.066; ρ: mANOVA F(2, 406) = 16.1, p $<$ 0.001, η^2_G = 0.023], with negative z-distances between NoGo vs. Go at all SNRs. There was no main effect of group [Power: mANOVA F(1, 173) = 0.5, p = 0.499, η^2_G = 0.001; ρ: mANOVA F(1, 203) = 1.2, p = 0.275, η^2_G = 0.001]. Analyzing the 10-Hz vector-strength across groups reveals that the vector strength was higher in the non-sound detecting as compared to the sound detecting group ([mANOVA F(1, 200) = 68.2, p $<$ 0.001, η^2_G = 0.103]; vector strength in Figure 3.8 D was overall higher than in Figure 3.8 B), with a main effect of SNR [mANOVA F(2.5, 505.5) = 79.3, p $<$ 0.001, η^2_G = 0.121] and main effect of task engagement [mANOVA F(1, 200) = 28.6, p $<$ 0.001, η^2_G = 0.018]. The Go-NoGo difference in 10 Hz power was greater in the active than passive conditions [mANOVA F(1, 173) = 7.2, p = 0.008, η^2_G = 0.009], but ρ did not differ appreciably between active and passive listening [mANOVA F(1, 203) = .3, p = 0.598, η^2_G $<$ 0.001].

Together, these results suggest that both rate and temporal cues could be used at the positive SNR. In contrast, at 0 dB SNR, spike count does not appreciably

differ between Go and NoGo trials [paired t(413) = -2.6, adjusted p = 0.031, Bonferroni-adjusted for multiple comparisons]. Moreover, at the negative SNR, only decision metrics relying on temporal coding can operate with a decision criterion that is consistent with that at the positive SNR.

3.4 Discussion

Moderate background sound disrupt auditory clarity in people with hearing loss, a challenge that no current clinical treatment approach overcomes for the majority of individuals. Prior behavioral work establishes that both peripheral mechanisms and central auditory processing contribute to dip listening. Evidence supporting central mechanisms include the comodulation masking release phenomenon, where added masker energy can paradoxically improve target detection and identification performance, an effect obliterated by backward-masking, another central phenomenon [47, 24]. Moreover, after an episode of temporary conductive hearing loss from otitis media, normally hearing children experience reduced ability to listen in the dips up to six months [41]. Consistent with this, gerbils with chronic juvenile-onset sound deprivation have raised detection thresholds in modulated noise despite not showing widened cochlear filters as compared to normally hearing controls [59, 135]. Even a theoretically ideal hearing aid could not compensate for peripheral dysfunction. In contrast, central processes could be targeted by rehabilitative technology. To better understand how hearing loss disrupts auditory clarity in situations with background sound and how auditory clarity could be restored, we therefore need to define central mechanisms for hearing in background sound. The current study intends to contribute towards this goal by examining SNR invariance in auditory cortex.

To elucidate the factors causing dip listening to operate across a range of negative and positive SNRs, this study sought to test the hypothesis that listening at negative SNRs is mediated by a stronger reliance on temporal cues, as compared to

positive SNRs. Evidence supporting this hypothesis comes from behavioral studies in human listeners showing that an individual's ability to resolve temporal fine structure information predicts their ability to suppress temporally fluctuating background sound [80, 56]. We here recorded auditory cortical responses in normally hearing, awake and freely moving adult gerbils, while these gerbils either actively detected on a target tone in 10 Hz modulated noise or passively heard the same sounds. Gerbils performed with comparable behavioral sensitivity across three SNRs that were either negative, balanced or positive. In the following sections, we discuss how SNR shapes the potency of rate vs. temporal neuronal coding cues for solving this task.

Using the conceptual framework of classic decision theory, a tacit premise of prior work looking at the effect of SNR on tone detection in modulated noise is that the separation between target and masker along an internal decision variable reduces with decreasing SNR. Although, this assumption had not explicitly been tested, raising the possibility that SNR shapes decision variables differently depending on whether they rely on rate or temporal coding.

3.4.1 Envelope Locking Suppression

In anesthetized untrained cats, phase-locking to a moderately slow amplitude-modulated masker envelope is suppressed 75 ms after onset of a target tone [73]. This phenomenon, referred to as hypersensitive locking suppression, emerges at the medial geniculate body, sculpting response pattern in primary auditory cortex across a range of positive to negative SNRs, and even for target tones below the quiet threshold of the neurons [73].

The existence of envelope locking suppression was confirmed in single units in the primary auditory cortex of anesthetized and untrained rats [50] as well as mice, albeit with low prevalence across recorded sites [117]. Specifically, in rat, at 0 dB SNR, average neuronal responses to tones in unmodulated masking noise did not change

much relative to responses to unmodulated masking noise in isolation, but showed suppressed envelope following responses when the masking noise was modulated [50]. Furthermore, in mouse, for SNRs ranging from -10 to 20 dB, single units responded more weakly to tones in modulated narrow-band noise when this narrow-band noise was presented in isolation, vs. with an incoherently modulated flanking noise vs. with a coherently flanking noise, and showed overall stronger envelope locking suppression in the presence of forward-masking fringes [117]. Moreover, in mouse, with decreasing SNR the firing rate in most units increased, and increased most steeply in the presence of coherently modulated background sound [117]. In contrast to these mammalian responses, neuronal responses in the primary auditory cortex homologue in birds, area L2 of the avian brain, lack locking suppression [72, 91], and have previously been compared to more strongly resemble inferior colliculus responses in cat vs. auditory cortex [73].

We here confirm envelope locking disruption at negative and positive SNRs, both quantitatively through the similarity index as well as qualitatively through visual inspection of the response time histograms. However, unlike in the prior work in anesthetized animals, in the current design with awake animals, overall neuronal discharge rates are enhanced, as opposed to suppressed, at negative SNRs. This phenomenon was observed in all animals that were engaged in a behavioral task, even when those animals were not engaged in sound detection. However, non-sound-detecting animals that were passively exposed to the same sounds showed enhanced neuronal discharge rates across all SNRs. As a result, spike count differences between Go and NoGo trials tend to be negative at positive SNRs, and positive at negative SNRs, an observation that is inconsistent with the interpretation that rate coding is SNR-invariant. In contrast, neurometric measures that are based on temporal patterns, including the similarity index and the decorrelation index ρ, do not change

signs with decreasing SNR, making them viable candidate metrics for the behaviorally observed SNR-invariant detection of target sound.

Previously, in trained rhesus monkeys that were engaged in a tone detection task with amplitude-modulated background noise, primary auditory cortex recordings showed that single-neuron responses were insufficient and pooling of spiking activity across the population of single neurons was necessary to predict behavioral performance from neuronal discharge rates [20]. The monkeys were tested across a range from -5 to 20 dB SNR. Although, performance at 0 and -5 dB SNR was at chance, suggesting that the monkeys could not detect the target sound at the two lowest SNRs [20]. Population responses showed enhancement followed by suppression after the target onset, a phenomenon confirmed by the current results for 0 and 10 dB SNR (Figure 3.10). Our current work supports these findings but shows that at negative SNRs and at the population level, spike count is not a viable metric for tone detection.

In addition, prior work on zebra finches finds evidence for SNR-invariant target feature detectors in primary auditory cortex. When listening for a conspecific target vocalization while a chorus of different conspecific vocalizations was concurrently present in the background, some single auditory cortical neurons selectively encoded the target sound in a manner that is invariant to the background, when those target sounds are behaviorally recognizable [113]. However, the zebra finches in this study could not perform the task at the only tested negative SNR of -15 dB [113]. It is therefore unclear whether the background-invariant encoding of target sound generalizes to negative SNRs. Here, at the single-unit level, only in the sound-detecting gerbils did we observe six highly specialized units with responses in background sound that closely resembled the responses to the target in quiet, across all SNRs (see Supplements FR+/VS+), hinting that the previously proposed selective target encoding mechanism may generalize to the mammalian auditory cortex.

3.4.2 Task Engagement

Others have reported and we confirm that task engagement shaped the anticipatory neural response, prior to the potential target onset [16]. Specifically, in sound-detecting animals, the average discharge rate during NoGos did not vary much between active vs. passive listening, except that after the animal initiated a trial via the nose poke, the majority of units showed a characteristic increase in firing rate, followed by a decrease. Whereas, in non-sound-detecting animals, firing rate did not increase and even modestly declined after nose poking triggered a trial. This suggests that the increase in firing rate immediately after nose-poking is an anticipatory auditory-specific response to support the ability to hear out the target, but only when a gerbil is trying to hear target sound.

3.4.3 Task Training

Overall, sound-detecting animals had higher non sound-evoked discharge rates as compared to non-sound-detecting animals. We observed that the mutual information between sustained NoGo and Go responses was higher for non-sound-detecting than for sound-detecting animals, and higher when sound-detecting animals missed a Go trial vs. when they responded correctly, consistent with the interpretation that sensory task training reduced the amount of shared information between Go and NoGo responses. This reduction in mutual information was not born out in the other neurometric analyses, where differentiation in neural responses between Go and NoGo stimuli appeared comparable for sound-detecting gerbils vs. non-sound detecting gerbils, for both rate and temporal coding metrics. Together, these results strongly suggest that much of the information that a classifier may initially learn to interpret from the sensory cortical responses in non-sound detecting gerbils was not, in fact, utilized by the highly trained animals and hint that learned behavioral ability to detect the target tone was mediated by readout metrics that are not captured by

simple rate-decoding or envelope locking suppression, e.g., via emergence of highly specialized template-matching units.

3.4.4 Clinical Implications

Most hearing-impaired individuals whose hearing has been restored via hearing aids or cochlear implants find it difficult to dissociate target sound that they are trying to hear from background sound, a phenomenon called masking. Moreover, even individuals with comparable peripheral audiometric thresholds can vastly differ in how well they can identify masked speech, a phenomenon that is thought to arise from central processing deficits [79, 118, 3, 55, 78]. An extensive literature demonstrates that the ability to listen in the dips is disrupted in individuals with hearing loss and in cochlear implant users [90, 61], suggesting that dip-listening mechanisms could be a key to restoring auditory clarity when background sound for people with hearing loss. Provision of visual lip-reading cues or reduction in target set size can restore a hearing impaired person's ability to identify target speech at negative SNRs [8, 7]. This phenomenon can be interpreted to mean that auditory dip-listening mechanisms in hearing impaired individuals function normally, provided that the overall *peripheral* SNR is low enough for dip-listening benefits to be effective. However, an alternative interpretation of this work is that provision of temporal cues via lipreading or increased stimulus redundancy via smaller set sizes substitutes malfunctioning auditory temporal cues, filling in *central* access to temporal information that is needed to listen in the dips. We currently lack physiological data to disambiguate these possibilities. The current work demonstrates that both rate and temporal cues can be effective at positive SNRs, but that reliance on temporal information is needed for SNR-invariant hearing at negative SNRs. Moreover, we previously demonstrated that gerbils raised with sound deprivation have reduced dip-listening abilities despite the fact that their peripheral tuning appears

intact [59, 135], suggesting that central processes play a key role in dip-listening. Future work is needed to understand whether the temporal information needed for dip-listening at negative SNRs can be augmented through other sensory modalities.

3.5 Supplemental Information

3.5.1 Single-Unit Response Types

In order to identify different patterns of single-unit activity, we categorized units as phasic or tonic based on their ability to follow the envelope of the masker. Pooling across Go and NoGo trials during either passive or active listening, units that had at least one significant vector strength (Rayleigh's $p < 0.001$, see Equation (3.4) for details) at the 10 Hz modulation rate of the masker were selected as phasic. The remaining units that were not able to phase-lock to the masker at any of the stimulus or task engagement conditions were categorized as tonic.

Among phasic units, we observed four phenotypes based on changes in firing rate and vector strength in response to the target stimulus during active engagement. This change was measured by comparing average activity of the unit across all trials (Go and NoGo) prior to tone onset, with the response of the unit to 10 dB SNR tones. Single-units that increased in both firing rate and vector strength were labeled as FR+/VS+, units that only had an enhanced firing rate but not vector strength as FR+/VS-, units that suppressed their overall firing rate but were better able to phase-lock to the masker as FR-/VS+, and units that decreased in both firing rate and vector strength were labeled FR-/VS-. These four phenotypes cover all quadrants of a rate/temporal response space.

3.5.2 Decoding Sensory Information by Single-Unit Response Type

Table 3.2 shows unit counts for all 4 phasic phenotypes and also tonics that were detected in each gerbil group. Results show that more than 90% of single-units in

both gerbil groups had the ability to follow the envelope of the masker. Among them, the most common phenotype was FR-/VS- with a preference for suppressing both their overall firing rate, and also their phase-locking to the masker in response to the target stimulus.

Table 3.2 Detected Unit Counts per Phenotype for Each of the Gerbil Groups

Gerbil group	FR-/VS+	FR+/VS+	FR-/VS-	FR+/VS-	Tonic	Total
Sound-detecting	28	6	75	28	14	151
Non-sound-detecting	18	1	30	4	3	56

In all types of phasic and tonic units, sensory information at the population level could predict behavioral sensitivity (Figure 3.13). For both FR-/VS+ and FR-/VS-, units at 0 and 10 dB SNR average firing rates decreased in Go vs NoGo trials, consistent with envelope locking suppression (Figure 3.13 A and C). However, at -10 dB SNR, firing rate increased during Go trials as compared to NoGo, a phenomenon not predicted by envelope locking suppression. In quiet, FR-/VS+ units showed suppression of non-sound evoked discharge at the highest sound level, vs. an increase in firing rate at the lowest sound level.

The FR+/VS+ units responded to the tone-masker mixture with a sustained enhancement in firing rate (Figure 3.11B). Interestingly, this activity pattern was consistent across all SNRs in both active and passive tasks, and closely resembled their response to 60 dB SPL tones in quiet. The spike count z-scores in Figure 3.13B also reflect this SNR-invariant response strategy, which would be ideal for a masker-agnostic readout of the target stimulus. Sound-detecting animals reveal 6 units out of 151 that follow this pattern.

3.5.3 Discussion of Single-Unit Response Types

The temporal pattern of envelope locking suppression varies non-linearly across units and SNRs. For instance, while a target sound is present, the majority of units, FR-/VS- units, show decreased vector strength in following the masker envelope, a weak target onset response that does not vary appreciably across SNR, and increased sustained firing with decreasing SNR. In contrast, FR-/VS+ units show increased vector strength to the masker envelope, increased target onset responses, and decreased target sustained responses with decreasing SNR. However, for both types of units, decorrelation between NoGo and Go is a robust, SNR-invariant cue.

When the target energy dominated the acoustic mixture, at 10 dB SNR, the majority of single-units during active listening decreased their firing rate (FR-) as compared to when only masker energy was present, in both sound-detecting (103/151) and non-sound-detecting (48/56) animals. Of those rate-suppressing units, 73% in sound-detecting and 63% in non-sound-detecting animals also decreased the vector strength (VS-) by which they tracked the masker. More so, these nominally rate-suppressing units typically showed a small but consistent *enhancement* in firing rate at -10 dB SNR, rejecting the idea of an SNR-invariant rate-based decision criterion from these units. Intriguingly, unlike in the active modes, and unlike in sound-detecting animals, in the passive responses of non-sound-detecting animals, firing rate consistently increases after target onset. Rate-based neurometric sensitivity in rate-enhancing (FR+) units of non-sound-detecting and passively listening animals did not vary appreciably was positive at all SNRs.

In contrast, across all rate-suppressing units, in sound-detecting animals, decorrelation d', a template model based on the magnitude of the decorrelation of the signal reaching the ears prior vs. during tone onset, is directly proportional to behavioral sensitivity and does not vary appreciably with SNR. In non-sound-detecting animals, the decorrelation d' analysis predicts an ability to detect the target at 10 and 0 dB,

but not -10 dB. This prediction mirrors the learning curves in the behavioral data demonstrating that at the beginning of training, gerbils cannot detect a tone at -10 dB SNR. Indeed, it often takes more than 20 training sessions until gerbils are reliably able to signal the presence of a tone at -10 dB SNR.

When responding to a target in quiet, during passive listening, the RTHs of the populations of rate-suppressing units tended to display strong onset responses followed by sustained suppression. In contrast, rate-enhancing unit responded more slowly to target sound and in sustained patterns. FR+/VS+ units also showed offset responses. This hints that these units rebound after inhibition in the presence of target sound [123]. Intriguingly, in the presence of target sound, FR+/VS+ units increase the vector strength by which they track the masker. But, despite this increased envelope locking, decorrelation-based neurometric sensitivity decreases with increasing SNR, and is small or absent at 0 and 10 dB SNR. In contrast, rate-based neurometric sensitivity appears to be roughly SNR-invariant, at least over the range of tested SNRs. Indeed, unlike in the other three unit types, the RTH patterns of FR+/VS+ unit are comparable across SNRs, but differ from their responses to the same targets in the absence of background sound.

Together, these findings show that at the level of core auditory cortex, at least two readout mechanisms exist. The majority of units change their firing rate when target sound is added. In non-sound-detecting and passively listening animals, firing rate increased at all SNR. In sound-detecting animals or in tasked-engaged non-sound-detecting animals, firing rate depended non-monotonically on SNR. Although, decorrelation in the envelope locking response of these units, either through a decrease or an increase in envelope locking, varies consistently with a change in behavioral sensitivity and is robust to SNR.

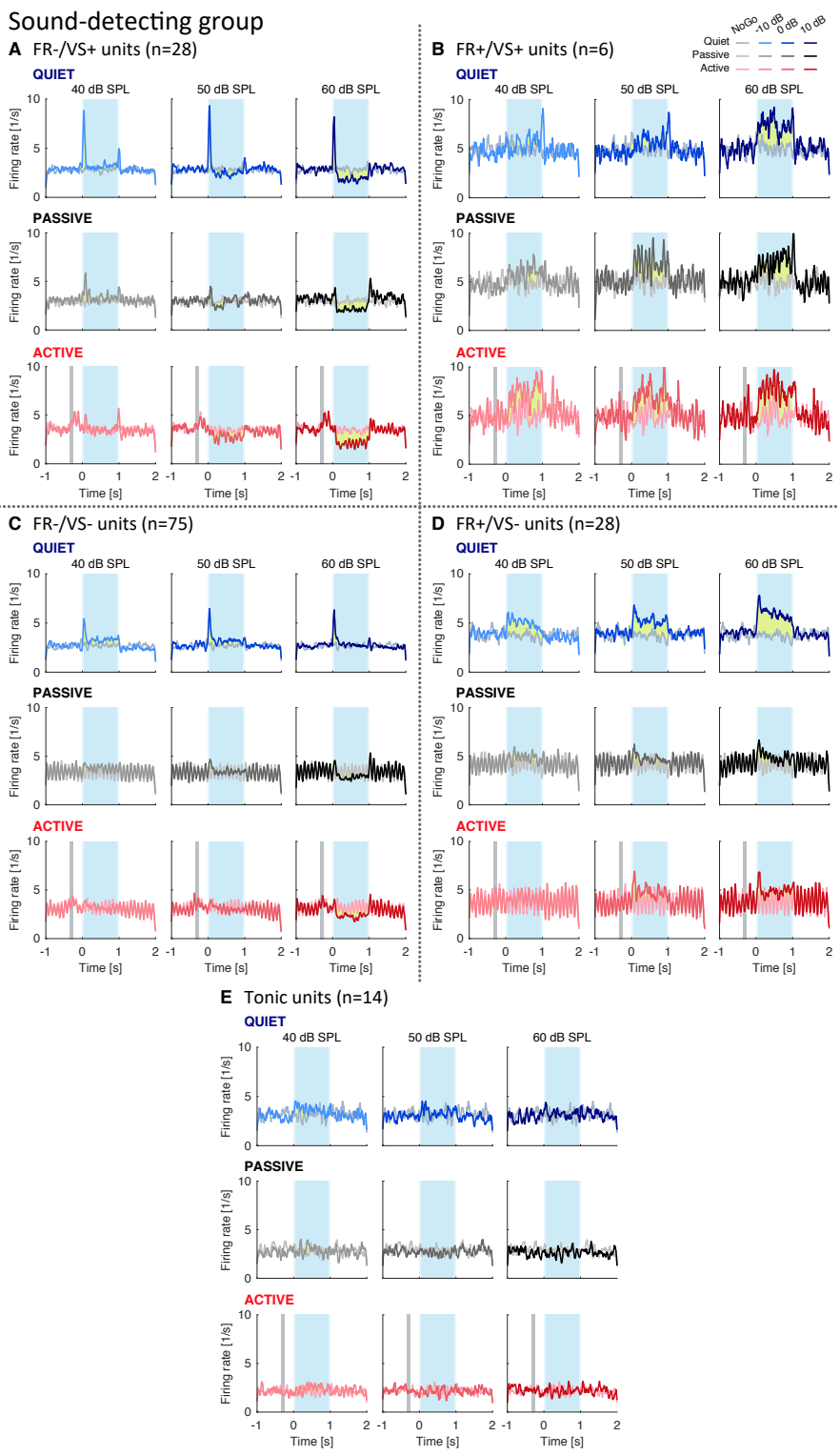

Figure 3.11 RTH of different unit phenotypes for sound-detecting gerbils. In addition to the passive and active engagement conditions with modulated masker (shades of black and red, same as Figure 3.7), here, unit responses to pure tone in quiet (i.e., in absence of background sound) are also illustrated as shades of blue.

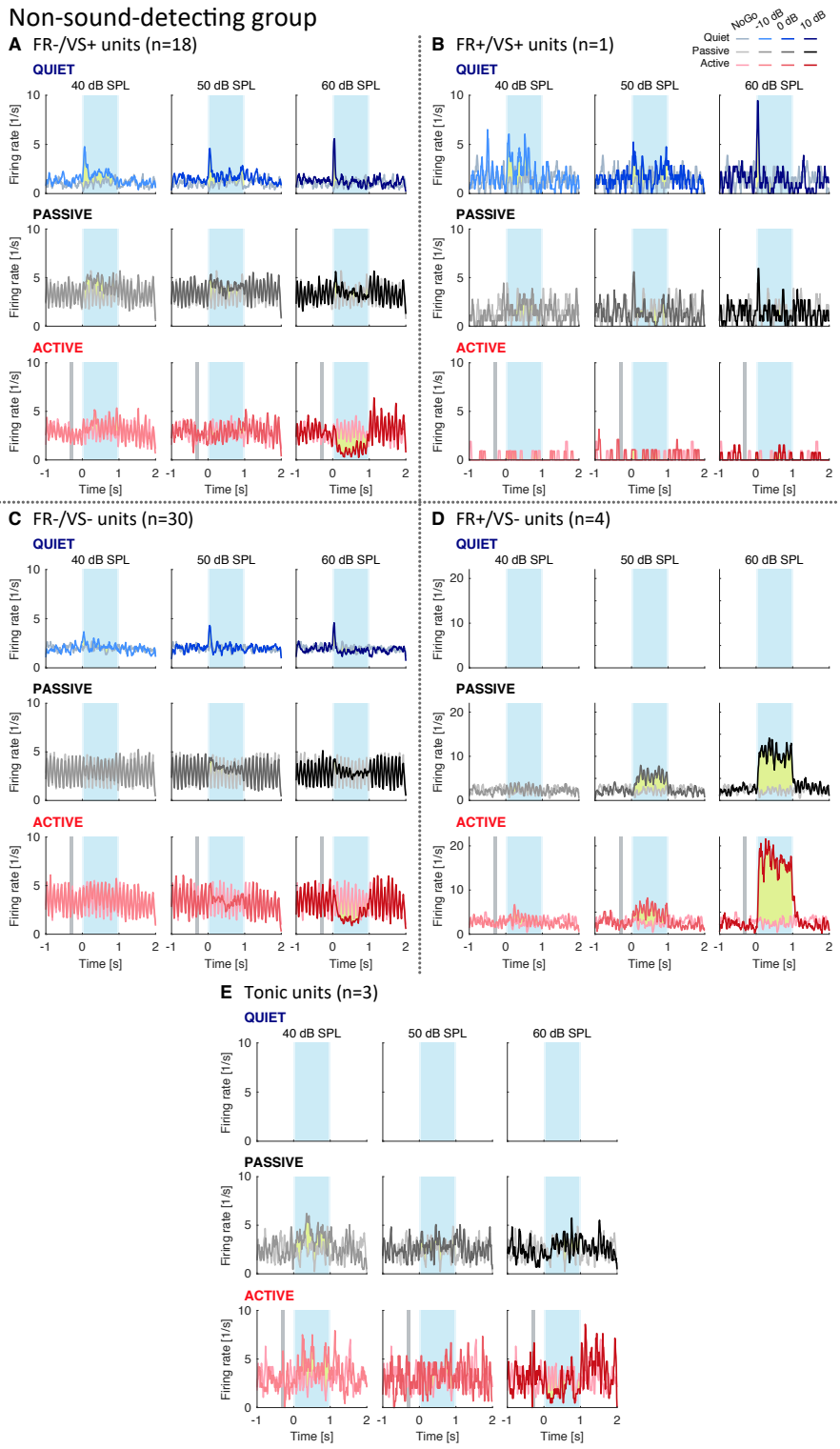

Figure 3.12 RTH of different unit phenotypes (similar to Figure 3.11) for non-sound-detecting gerbils. Note: FR+/VS- and tonic units have missing recordings for the quiet condition.

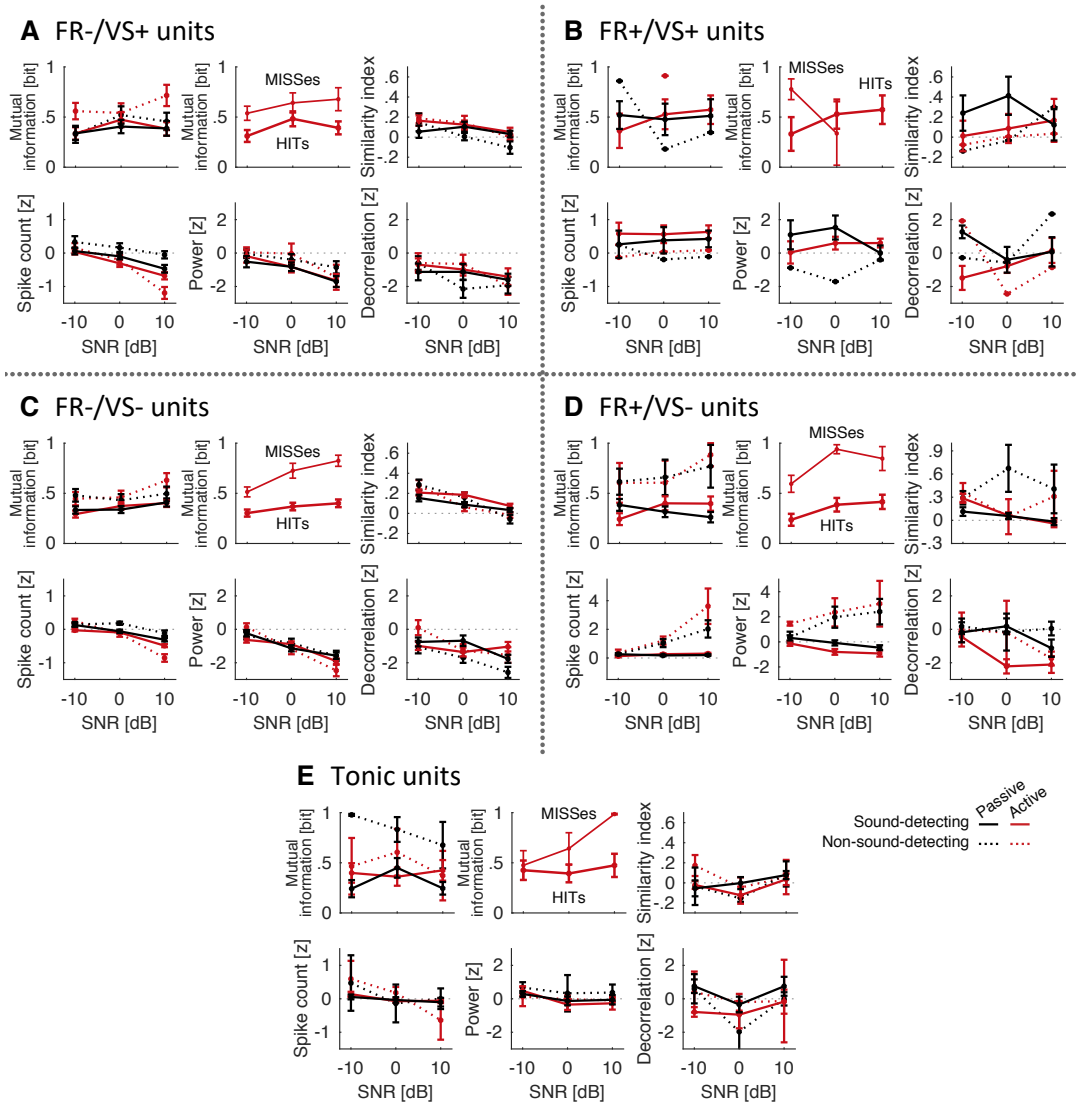

Figure 3.13 Average of neurometric rate and temporal metrics (similar to Figure 3.10) broken down per each of the five unit phenotypes. For unit counts per phenotype for each gerbil group, refer to Table 3.2.

CHAPTER 4

LEVEL DEPENDENCE OF HUMAN SOUND LATERALIZATION WITH INTERAURAL TIME DIFFERENCES

Human sound localization is an important computation performed by the brain. Models of sound localization commonly assume that sound lateralization from interaural time differences is level invariant. Here, we observe that two prevalent theories of sound localization make opposing predictions. The labelled-line model encodes location through tuned representations of spatial location and predicts that perceived direction is level invariant. In contrast, the hemispheric-difference model encodes location through spike-rate and predicts that perceived direction becomes medially biased at low sound levels. Behavioral experiments find that softer sounds are perceived closer to midline than louder sounds, favoring rate-coding models of human sound localization. Analogously, visual depth perception, which is based on interocular disparity, depends on the contrast of the target. The similar results in hearing and vision suggest that the brain may use a canonical computation of location: encoding perceived location through population spike rate relative to baseline.

4.1 Introduction

A fundamental question of human perception is how we perceive target locations in space. Through our eyes and skin, the activation patterns of sensory organs provide rich spatial cues. However, for other sensory dimensions, including sound localization and visual depth perception, spatial locations must be computed by the brain. For instance, interaural time differences (ITDs) of the sounds reaching the ears allow listeners to localize sound in the horizontal plane. In the ascending mammalian auditory pathway, the first neural processing stage where ITDs are encoded, on the timescale of microseconds, is the medial superior olive (MSO). Here, temporally

precise binaural inputs converge, and their ITDs are converted to neural firing rate [35, 136, 119, 96, 25]. The shape of the MSO output firing rate curves as a function of ITD resembles that of a cross-correlation operation on the inputs to each ear [6]. How this information is interpreted downstream of the MSO has led to the development of conflicting theories on the neural mechanisms of sound localization in humans. One prominent neural model for sound localization, originally proposed by Jeffress, consists of a labelled line of coincidence detector neurons that are sensitive to the binaural synchronicity of neural inputs from each ear [65], with each neuron maximally sensitive to a specific magnitude of ITD (Figure 4.1A). This labelled-line model is computationally equivalent to a neural place-code based on bandlimited cross-correlations of the sounds reaching both ears [29]. Several studies support the existence of labelled-line neural place-code mechanisms in the avian brain [18, 92], and versions of it have successfully been applied in many engineering applications predicting human localization performance [30, 46, 122, 13, 48].

A growing literature proposes an alternative to the labelled-line model to explain mammalian sensitivity to ITD [74]. One reason for an alternative is that two excitatory inputs should suffice to implement the labelled-line model, but evidence from experiments on Mongolian gerbils shows that in addition to bilateral excitatory inputs, sharply tuned bilateral inhibitory inputs to the MSO play a crucial role in processing ITDs [12]. Moreover, to date no labelled-line type neurons encoding auditory space have been discovered in a mammalian species. Indeed, using a population rate-code, several studies proposed that mammalian sound localization can be modeled based on differences in firing rates across the two populations of neurons that are tuned to opposing hemispheres (Figure 4.1B; [126, 82, 26]). Rate-based models generally predict that neuronal responses carry most information at the steepest slopes of neural-discharge-rate versus ITD curves, where neural discharge changes most strongly [121], consistent with the observation that the peak ITDs of

rate-ITD curves often fall outside the physiologically plausible range ([82, 44]; but see also [66]). In addition, some authors have suggested that how mammalian sound localization adapts to stimulus history further supports a rate-based neural population code, as assessed behaviorally or via magnetoencephalography [99, 120, 112].

It is unknown which of the two competing models, broadly characterized as labelled-line versus rate-code model, describes human sound localization better. Here, we observe that the two different models predict different dependencies of sound localization on sound intensity. By combining behavioral data on sound intensity dependence in normal-hearing listeners with numerical predictions of human sound lateralization from both models, we attempt to disentangle whether human auditory perception is based on a place-code, akin to the labelled-line model, or whether it is instead more closely described by a population rate-code.

An extensive physiology literature characterizes labelled-line versus population-rate type neurons and suggests that, at least from the perspective of evolution, birds and mammals use different neural mechanisms to calculate sound direction (review: [42]). Thus, we searched the avian and mammalian physiology literature and identified two studies that characterized labelled-line versus population rate-code neurons at low sound levels and as a function of both sound level and ITD [97, 139]. Both Peña and colleagues (1996) and Zwiers and colleagues (2004) report neural firing rate in response to acoustic noise stimuli and are thus suitable for predicting each model's sensitivity to the acoustic noises we tested in the current study. Here, we ran a meta-analysis, reconstructing simulated neurons with response characteristics from each of the two studies and using maximum likelihood estimation to predict source laterality from these previous findings.

4.2 Model Predictions

To predict how lateralization depends on sound intensity from the responses of labelled-line neurons, we estimated neural firing rates from previous recordings in the nucleus laminaris in barn owl [97]. To estimate lateralization's dependence on level based on a population rate-code, we used previous recordings from the inferior colliculus of rhesus macaque monkey and calculated hemispheric differences in firing rate [139]. The labelled-line neurons predicted that, as sound intensity decreases, perceived source laterality would converge towards similar means for low versus high sound intensities, with increased response variability at decreasing sound intensities (Figure 4.1C). In contrast, the hemispheric-difference model predicted that as sound intensity decreases to near threshold levels, perceived laterality would become increasingly biased toward the midline reference (Figure 4.1D, e.g., note the shallower slope and thus compressed laterality percepts for red versus blue curves). For lower sound intensities, predicted source direction is biased towards midline (compare red and orange versus blue or yellow). For higher sound intensities, predicted source direction is intensity invariant (blue on top of yellow line). At higher overall sound intensities, both models predicted that lateralization would be intensity invariant (see insets in Figure 4.1C versus D). Therefore, analyzing how sound intensity affects perceived sound direction near sensation threshold offers an opportunity to disentangle whether our human auditory system relies on a place-based or rate-based population code for localizing sound based on ITD.

A listener's ability to discriminate ITD can vary with sound intensity [27]. However, it is difficult to interpret previous findings linking sensitivity to ITD and a listener's judgement of sound source direction as a function of sound intensity. Some reported decreased perceived source laterality near sensation threshold [124, 111], but others reported weak or no level effects on perceived lateralization [129, 84, 49, 81, 62, 128, 15, 33]. Several factors complicate the interpretation of these previous findings

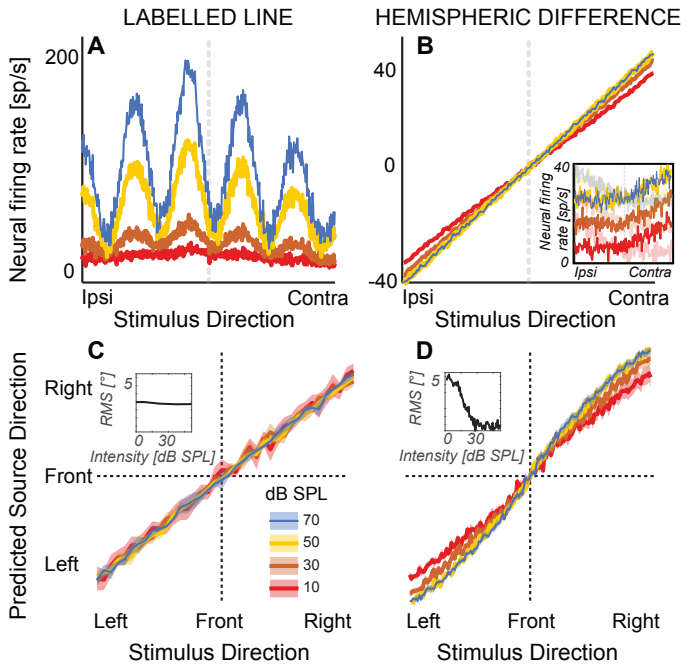

Figure 4.1 **A** Firing rate of a simulated *nucleus laminaris* neuron with a preferred ITD of 375 µs, as a function of source ITD. The model predicts source laterality based on the locus of the peak of the firing rate function. **B** Hemispheric differences in firing rates, averaged across all 81 simulated *inferior colliculus* units. Rate models assume that source laterality is proportional to firing rate, causing ambiguities at the lowest sound intensities. Inset: Reconstructed responses of an *inferior colliculus* unit. The unit predominantly responds contralaterally to the direction of sound (high-contrast traces). The hemispheric difference model subtracts this activity from the average rate on the ipsilateral side (example shown with low-contrast traces). **C** Mean population response using labelled-line coding across a range of ITDs and sound intensities. Inset: The root-mean square (RMS) difference relative to estimated angle at 80 dB SPL does not change with sound intensity, predicting that sound laterality is intensity invariant. **D** Mean population response using hemispheric-difference coding. Inset: RMS difference relative to estimated angle at 80 dB SPL decreases with increasing sound intensity, predicting that sound laterality is not intensity invariant. Ribbons show one standard error of the mean across 100 simulated responses. Sound intensity is denoted by color (see color key in the figure).

in the context of the current hypothesis. For instance, assuming an approximately 30 dB dynamic range of rate-level function either at the MSO or downstream in the binaural pathway (e.g., *medial superior olive*: [35]; *inferior colliculus*: [139]), for stimuli at higher sensation levels (SL) where the rate-level functions saturate, both the labelled-line and the hemispheric difference model predict level invariance. This could explain how studies that tested for sound intensity effects over a range of high intensities did not see an effect. Moreover, when presented in the free field, sounds also contain interaural level differences and spectral cues, in addition to ITD. For low-frequency sound, listeners rely dominantly on ITD when judging lateral source angle. In contrast, for broadband sound, listeners integrate across all three types of spatial cue [131, 60]. Unlike ITDs, interaural level differences and overall sound intensity both decrease with increasing source distance, raising the possibility that for stimuli with high-frequency content, listeners judged softer sounds to be more medial because they interpreted them to be farther away than louder sounds. Further, at low sound intensities, the sound-direction-related notches of the spectral cues at high-frequencies should have been less audible than at higher sound intensities, increasing stimulus ambiguity. A resulting increase in response variability may have obscured the effect of level on ITD coding. Finally, some historic studies used only two or three listeners, suggesting that they may have been statistically underpowered. Thus, the literature provides insufficient evidence on how ITD-based lateralization varies with level near sensation threshold.

4.3 Human Perception

Here, we contrasted two competing hypotheses toward the goal of disentangling whether ITD-based human sound localization relies on a labelled-line versus a population rate-place neural code. The labelled-line code hypothesis predicted that the mean perceived direction based on ITD would be intensity invariant, even at

intensities close to SL. Using a psychophysical paradigm, we studied lateralization based on ITD as a function of sound intensity in a group of ten normally hearing listeners (experiment 1). Stimuli consisted of low-frequency noise tokens that were bandlimited to cover most of the frequency range where humans can discriminate ITD [14] (here, corner frequencies from 300 to 1200 Hz, shown in Figure 4.2A). In each one-interval trial, listeners had to indicate perceived laterality across a range of ITDs from -375 to 375 μs. Lateralization was measured as function of SL. To examine how sound intensity affects perceived ITD coding of source direction, we modelled perceived laterality with a nonlinear mixed effect model (NLME) that included fixed effects of ITD and sound intensity as well as a random effect of listener.

Figure 4.2B depicts lateralization performance with spectrally flat noise at two sound intensities for a representative listener (TCW). Figure 4.2C shows raw data (circles) and NMLE fits (lines) across all listeners. Error bars show one standard error of the mean across listeners, and shaded ribbons indicate one standard error of the mean fit across listeners. This model predicts 80.6% of the variance in the measured responses and is deemed an appropriate fit of the data. Table 4.1 lists all NLME parameters. In this table *Laterality:sound intensity* refers to the NLME weight attributed to acoustic sound intensity of the auditory target. In contrast, *Laterality:audibility* captures the NLME weight attributed to pure tone audiometric thresholds based on the listeners' perceptual abilities. The NLME model results show that perceived laterality scores increased with increasing ITD, as expected. With decreasing sound intensity, percepts were increasingly biased towards midline (compare order of colored lines, magnified in the inset of Figure 4.2C). These trends were supported by the NLME model, which revealed significant effects of ITD ($p < 0.001$; α_{x1}) and sound intensity ($p < 0.001$; α_{y1}) on the maximal extent of laterality, confirming the predicted trend from the hemispheric difference model and rejecting our null hypothesis. Average pure tone audiometric thresholds affect

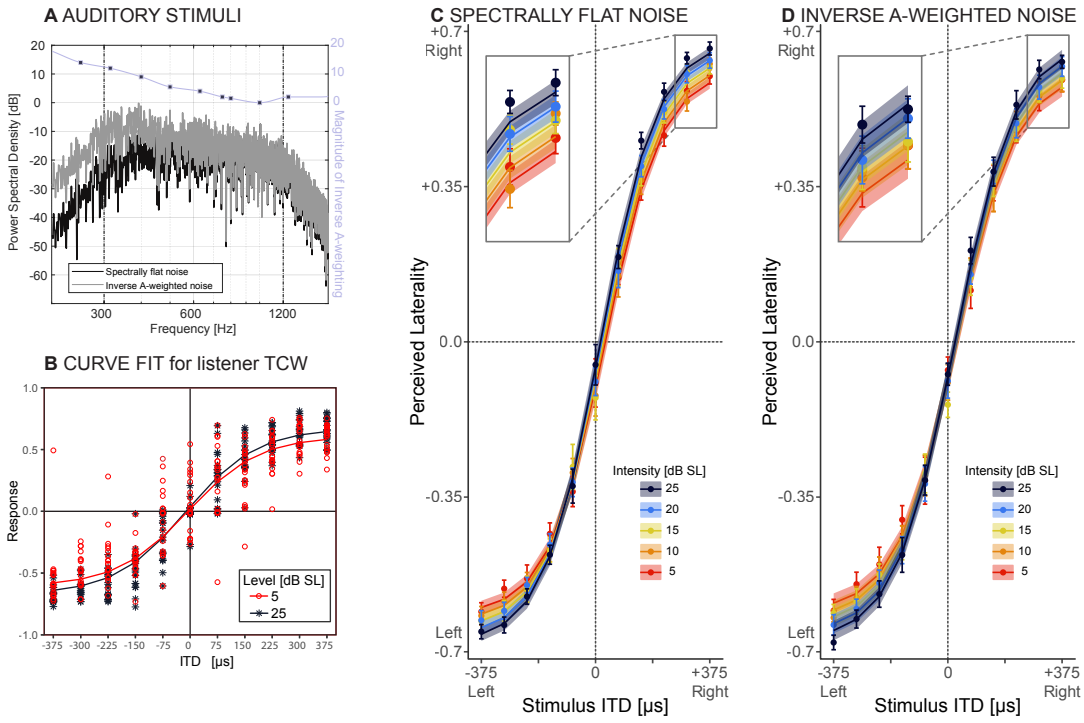

Figure 4.2 Behavioral experiment. **A** Stimuli: spectrally flat noise, used in experiment 1 (dark grey) versus A-weighted noise, tested as a control for audibility in experiment 2 (light grey). The purple line shows the magnitude of the zero-phase inverse A-weighting filter. **B** Responses from one representative listener (TCW) across two sound intensities and the corresponding NLME fits for these data. Perceived laterality as a function of ITD for **C** spectrally flat noise (experiment 1) or **D** A-weighted noise (experiment 2). Error bars, where large enough to be visible, show one standard error of the mean across listeners. Colors denote sound intensity. Insets illustrate magnified section of the plots. Circles show raw data, lines and ribbons show NLME fits and one standard of the mean.

perceived laterality, albeit mildly ($p < 0.001$; α_{y2} = -0.06), consistent with a slight leftward response bias. Sound intensity did not significantly affect the slope of the psychometric functions ($p = 0.14$; α_{x2}).

Table 4.1 Non-Linear Mixed Effects Model for Flat-Spectrum Noise Condition

	Description	Estimate	Std. Error	t-value	p-value	
α_{x0}	*Intercept: ITD*	0.06	0.04	1.58	0.11	
α_{x1}	*Slope: ITD*	2.45	0.05	46.15	<0.001	***
α_{x2}	*Slope: Sound Intensity*	0.02	0.01	1.47	0.14	
α_{y1}	*Laterality: Sound Intensity*	0.05	0.01	7.59	<0.001	***
α_{y2}	*Laterality: Audibility*	0.01	0.01	4.86	<0.001	***

Note: 10986 degrees of freedom.

Significance codes: '***' $p < 0.001$, '**' $p < 0.01$, '*' $p < 0.05$, '.' $p < 0.1$, ' ' $p \geq 0.1$.

In a second experiment, we examined whether these results were robust to the spectral details of the stimuli. A caveat of testing spectrally flat noise at low sound intensities is that parts of the spectrum may be inaudible, and this may contribute to the slight but significant effect of audibility on laterality (α_{y2}). Therefore, the results of experiment 1 could potentially be confounded by the fact that the bandwidth of the audible portion of the noise tokens decreased with decreasing sound intensity. Alternatively, the effect of absolute pure tone detection thresholds that we observe in our normal-hearing listeners may reflect differences in neural function beyond audibility. As a control for perceived stimulus bandwidth, the same listeners were tested again, using inverse A-weighted noises (experiment 2). Inverse A-weighting boosts sound energy at each frequency in rough proportion to the human threshold. Resulting inverse A-weighted sensitivity thus achieves nearly constant sensation level across frequency. All of the original ten listeners from experiment 1 completed

experiment 2. Methods were similar as in the first experiment, except that the stimuli consisted of inversely A-weighted noise (compare magnitude spectra in Figure 4.2A). The data and NLME model fits for the second experiment are shown in Figure 4.2D (color key identical to Figure 4.2C), and coefficients are listed in Table 4.2. This second model accounts for 80.4% of the variance in the data, closely fitting the measured responses. All NLME coefficients are significant ($p<0.001$ for α_{x1}; α_{x2}; α_{y1} and α_{y2}). Intercept coefficient (α_{x0} estimate = -0.60, SE = 0.03, $p < 0.001$) revealed a slight leftward response bias, consistent with a slight narrowband interaural level difference in our stimuli due to precision limits of our test system. The fact that α_{x2} is significant shows that when all noise portions are approximately equally audible, as here, with inverse A-weighted noise, both perceived laterality and the slope linking the change in laterality to ITD decrease with decreasing sound intensity. This is consistent with the interpretation that by controlling for audibility across-frequency, the sensitivity of the task to sound level increases, revealing a medial bias effect not only for the most lateral but also for more medial source angles.

Table 4.2 Non-Linear Mixed Effects Model for Inverse A-Weighted Noise Condition

	Description	Estimate	Std. Error	t-value	p-value	
α_{x0}	*Intercept: ITD*	-0.60	0.03	-19.28	<0.001	***
α_{x1}	*Slope: ITD*	2.57	0.06	46.26	<0.001	***
α_{x2}	*Slope: Sound intensity*	0.06	0.01	4.98	<0.001	***
α_{y1}	*Laterality: Sound intensity*	0.04	0.01	7.10	<0.001	***
α_{y2}	*Laterality: Audibility*	0.01	0.01	3.30	<0.001	***

Note: 10986 degrees of freedom.

Significance codes: '***' $p < 0.001$, '**' $p < 0.01$, '*' $p < 0.05$, '.' $p < 0.1$, ' ' $p \geq 0.1$.

Thus, the results confirm the effect of biasing perceived laterality toward midline with decreasing sound intensity. Therefore, for both spectrally flat noise and A-weighted noise, statistical analyses, which partialed out overall differences between listeners, are inconsistent with a labelled-line model of human sound localization.

4.4 Experimental Model

4.4.1 Subject Details

Twelve naïve normal-hearing listeners (ages 18-27, five females) were enrolled in this study and paid for their time. Their audiometric thresholds, as assessed via a calibrated GSI 39 Auto Tymp device (Grason-Stadler), were 25 dB hearing level or better at octave frequencies from 250 to 8000 Hz, and did not differ by more than 10 dB across ears at each octave frequency. This study was approved by and all testing was administered according to the guidelines of the Institutional Review Board of the New Jersey Institute of Technology, protocol F217-14. All listeners gave written informed consent both to participate in the study and to publish the results with confidential listener identity.

4.4.2 Method Details

Listeners were seated in a double-walled sound-attenuating booth (Industrial Acoustics Company) with a noise floor of 20.0 dB SPL (wideband LAFeq). Stimuli were digitally generated in Matlab R2016b (The MathWorks, Inc.), D/A converted through an external sound card (Emotiva Stealth DC-1) at a sampling frequency of 192 kHz, with a resolution of 24 bits per sample, and presented to the listener through ER-2 insert earphones (Etymotic Research Inc.). The equipment was calibrated using an acoustic mannequin (KEMAR model, G.R.A.S. Sound and Vibration) with a precision of less than $\pm$ 5 μs ITD and less than $\pm$1 dB interaural level difference. Foam eartips were inserted following guidelines provided by Etymotic Research to

encourage equal representation of sounds to both ears and minimize interaural leakage. Each session lasted approximately 60 minutes. Listeners kept the insert earphones placed inside their ears throughout testing. Insert earphones were replaced by the experimenter after each break. Throughout this study, to generate stimuli, tokens of uniformly distributed white noise were generated and bandpassed using a zero-phase Butterworth filter with 36 dB/octave frequency roll-off, and 3 dB down points at 300 and 1200 Hz. Each noise token was 1 s in duration, including 10 ms long squared cosine ramps at the onset and offset.

4.4.3 Sensation Level Measurements

At the beginning of each session, and, as a re-test control, mid-way through each session, each individual listener's SL was measured for the type of sound that was later on used for training and testing, via one run of adaptive tracking. On each one-interval trial of each track, a new noise token was generated and presented diotically. Trials were spaced randomly in time (uniform distribution, inter-token intervals from 3 to 5.5 s). Listeners pressed a button when they heard a sound. No response feedback was given.

On each trial, a response was scored a "hit" if a listener responded with a button push before the onset of the subsequent trial, and a "miss" if the listener did not respond during the interval. If a listener's response changed from hit to miss or from miss to hit across sequential trials, this was interpreted as a response reversal. Using one-up-one-down adaptive tracking, the noise intensity was increased or decreased after each reversal, with a step size of 5 dB (decreasing) or 2.5 dB (increasing). Each listener completed ten adaptive-track reversals, with SL threshold equaling the median of the final six reversals. Each SL was used as reference intensity for the subsequent 30 minutes of testing. If detection thresholds changed between initial test and re-test control by more than 5 dB, this indicated that an insert earphone moved,

and the experimenter replaced the earphones. Thresholds generally did not change by more than 5 dB.

4.4.4 Training

To train listeners on consistently reporting their perception of ITD, using adaptive tracking, listeners matched the perceived laterality of a variable-ITD pointer to that of a fixed-ITD target. Target token intensity was set relative to the listener's own diotic sensation threshold, at 10 or 25 dB SL, and presented with 0 dB interaural level difference. The pointer intensity was fixed at 25 dB SL. Target ITDs spanned the range from 375 to 375 μs, in 75 μs steps. Target ITDs and SLs were randomly interleaved across runs, but held fixed throughout each adaptive run. In each two-interval trial of a run, the pointer token was presented in the first, and the target token in the second interval. The start ITD of the pointer token at the beginning of each run equaled 0 μs. Using a hand-held controller (Xbox 360 wireless controller for Windows, Microsoft Corp.), listeners adjusted the ITD of the pointer token. Specifically, listeners pushed the directional keys (D-pad) either to the left or right in order to move and match the pointer direction with that of the target sound. When a listener indicated a left- or right-ward response, the pointer ITD was decreased or increased. Initial ITD step size equaled 100 μs, then 50 $\pm$ 5 μs (uniformly distributed) after the first reversal. By the end of the second reversal, ITD step size was reduced to 25 $\pm$ 5 μs (uniformly distributed) and remained the same for all of the following reversals. Listeners were instructed to "home in" on the target by moving the pointer initially to a position more lateral than the target, then more medial than the target with the goal of centering on the target. No response feedback was provided. A run was completed after a listener had completed a total of five adaptive-tracking reversals. For each target ITD, the matched pointer ITD was estimated by averaging the pointer ITDs of the final two reversals. Each listener

performed three sessions of training: In the first session only a subset of target ITDs were presented (-375, -150, 0, 150 and 375 μs), whereas the two following sessions included all of the eleven ITDs. Per training session, each ITD was presented once at 10 and 25 dB SL, for a total of 54 adaptive tracking runs across all training sessions. To familiarize listeners with the experimental task (described below), at the end of second and third sessions of training listeners performed an additional five blocks of the experimental testing task, without response feedback. These task training data were not used for statistical analysis.

To assess whether listeners could reliably report their lateralization percepts, training performance was evaluated for each listener by calculating the Pearson correlation coefficient between target ITD and matched pointer ITD in the final training session. Criterion correlation equaled 0.9 (N=11 ITDs, significance level=0.01, power=0.95). Ten listeners reached criterion, suggesting that they were able to consistently report where they perceived the sounds based on ITD. Two of the originally recruited twelve listeners failed to reach training criterion (R^2 = <0.84, 0.87>) and were excluded from testing.

4.4.5 Testing

Using the method of fixed stimuli, we tested lateralization in two experiments. Except for the stimuli, which consisted of spectrally flat noise tokens in experiment 1 and A-weighted noise tokens in experiment 2, the methods were similar across the two experiments. Noise tokens were generated from a statistically similar noise distribution as those presented during both SL measurements and training (see Overall Design). A touchscreen monitor (Dell P2314T) displayed the response interface at about 40 cm distance from the listener. Using a precise touch stylus (MEKO Active Fine Point Stylus 1.5 mm Tip), listeners indicated perceived laterality of noise in a one-interval task. Noise tokens were presented at 5, 10, 15, 20, and 25 dB

SL. ITDs varied randomly from trial to trial, in 75 μs steps spanning the range from -375 μs to $\pm$ 375 μs. On each trial, a new token of noise was generated. Each listener performed 20 blocks of 55 trials each (11 ITDs at each of the 5 sound intensities), with SL measured both before the first and the eleventh block. ITDs and sound intensity were randomly interleaved from trial to trial such that each combination of ITD and sound intensity was presented once before all of them were repeated in a different random order.

4.5 Analysis

4.5.1 Models

We estimate the combined effects of ITD and sound intensity on predicted source laterality both in avian labelled-line type units and in binaurally sensitive units of a mammalian auditory system. The sound intensities where we expect to see an effect of overall sound level fall below 30 dB SPL, because only in this range would most auditory neurons fire below saturation, allowing us to disambiguate labelled-line versus hemispheric rate-difference coding. However, scant data exist for either type of unit at sound pressure levels below 30 dB SPL. We identified two prior studies that have measured neural discharge rate as a function of ITD at these very low sound intensities. Both studies used noise as acoustic stimuli, and the neural response statistics they report are thus suitable for estimating what type of information would be available to either type of coding mechanism with the type of noise stimuli that human listeners lateralized in the behavioral experiments here.

One study in barn owl shows that the output functions of *nucleus laminaris* neurons can be modeled through interaural cross-correlation functions, even at very low sound intensities [97]. That study reports Pearson correlation coefficients between the neural response function of *nucleus laminaris* units at 50 dB SPL versus all other tested sound levels. To reconstruct the spatial information realistically available from

the output of labelled-line neurons, across both a range of -375 to 375 μs ITD in 20 μs steps, we first constructed biologically plausible interaural cross-correlation functions at 50 dB SPL and then added internal noise to the resulting curves to mimic the Pearson correlation coefficients reported by Peña and colleagues (1996). Our model predictions pertain to sound intensities spanning the range from 10 to 70 dB SPL, similar to previous work [97]. Due to overall scarcity of available data at low dB SPL, here we use firing rate characteristics for unit # 0123795-530.02 [97] with a nominal best frequency of 1 kHz. To generate the acoustic inputs to the labelled-line model, we initially generated a Gaussian noise token, duplicated it and introduced a variable ITD, spanning a range from -375 to 375 μs, with 20 μs step size and 0 dB interaural level difference. To simulate ITD information available after cochlear processing, we then processed both noises with a 1/3-octave wide bandpass filter with 24 dB/octave frequency roll-off, followed by half-wave rectification and low-pass filtering at 1500 Hz. We then simulated internal noise by adding uniformly distributed dichotic noise tokens with mean spontaneous firing rates of 5% of the root mean square value of the signal, resulting in left (L) and right (R) inputs to the binaural cross-correlation neurons, called $x_L(t)$ and $x_R(t)$. To establish 50 dB SPL reference functions, at each simulated ITD, we then calculated the binaural cross-correlation function $cc(\tau)$ of $x_L(t)$ and $x_R(t)$, as follows:

$$cc(\tau) = 300 + (450 - 300)\frac{\int_{-\infty}^{+\infty} x_L(t)x_R(t+\tau)dt)}{max|\int_{-\infty}^{+\infty} x_L(t)x_R(t+\tau)dt|}, \tag{4.1}$$

with τ signifying the best ITD of each neuron, and extrema scaled such that $cc(\tau)$ spans a range from 300 to 450 spikes/sec, approximating *nucleus laminaris* firing rates at 50 dB SPL [97]. To simulate non-sound driven neural discharge, we then added uniformly distributed random noise $\widehat{cc}_{ref}(\tau) = cc(\tau) + U(0, \mu)$, with a mean discharge of μ=5 spikes/sec, [97]. The resulting signal is our reference cross-

correlation function at 50 dB SPL, called $cc_{ref}(\tau)$, shown in Figure 4.1A as yellow bold line for a representative simulated neuron.

For each sound level and ITD, we then statistically reconstructed a family of interaural cross-correlation functions that match the originally reported functions [97]. Specifically, we added scaled dichotic uniformly distributed noise tokens $n_L(t) \leftarrow U(0, \mu)$ and $n_R(t) \leftarrow U(0, \mu)$ to the $x_L(t)$ and $x_R(t)$, such that the monaural inputs to the binaural cross-correlation functions equal $\widehat{x}_{L,R}(t) = \alpha x_{L,R}(t) + \sqrt{1-\alpha^2} n_{L,R}(t)$. The resulting cross-correlation function for each sound level and ITD is then $\widehat{cc}(\tau) = (x_L \star x_R)(\tau)$, shown for a representative neuron in Figure 4.1A as blue, brown and red lines corresponding to 70, 30 and 10 dB SPL. We then searched through the space of scaling coefficients α until the Pearson correlation coefficient between $cc_{ref}(\tau)$ and $\widehat{cc}(\tau)$ matched the coefficients originally reported by Peña and colleagues (1996) with a precision error of less than 10%.

To estimate predicted sound laterality as a function of sound intensity for these simulated labelled-line neurons, at each intensity, we then identified the τ where $\widehat{cc}(\tau) = max(\widehat{cc}(\tau))$. For each sound level and ITD, we calculated predicted sound laterality in 100 repetitions of these simulations. Figure 4.1C shows mean estimated laterality across these 100 simulations, with ribbons showing 1 standard error of the mean across simulations.

To estimate source laterality based on rate-coding, we assayed the mammalian auditory system, where one previous study reports firing statistics for 81 *inferior colliculus* units in rhesus macaque as a function of ITD and over a wide range of sound intensities, including very low sound intensities [139]. From the previously published linear regression parameters, we initially reconstruct linear regression functions linking ITD, sound intensity and firing rate [139]. However, while linear regression fits afford statisticial convenience, they cannot fully capture the sigmoidally shaped firing rate functions in mammalian *inferior colliculus* units. Therefore, we

multiplied the original linear reconstructions with sigmoid functions. Specifically, consistent with prior literature, each simulated sigmoidal output function saturates over a 30 dB dynamic range, has linear growth over the physiologically plausible range of contralateral ITDs, has a threshold between uniformly distributed between 0 and 10 dB SPL, and a spontaneous non-sound-evoked discharge of between 2 and 10 spikes/second (e.g., [107]).

The inset of Figure 4.1B shows a representative simulated *inferior colliculus* unit (color denotes sound intensity, dark shading shows contralateral responses), whereas Figure 4.1B shows the differences in firing rates for contra minus ipsi-lateral simulated firing rates, averaged across all 81 simulated *inferior colliculus* units. From these resulting differences in contra versus ipsi firing rates we calculated, collapsed across sound intensities from 0 to 80 dB SPL, the probability density of the firing rate for each *inferior colliculus* unit as a function of source ITD. Assuming an ideal observer, we then classified the sound azimuth as a function of sound intensity via maximum likelihood estimation. To calculate the mean and variance of predicted ITD as a function of sound intensity, we then ran a bootstrapping analysis, sampling with replacement 100 times. Figure 4.1D shows the across-simulation average predicted source laterality, with ribbons showing one standard error of the mean across simulations.

4.5.2 Quantification and Statistical Analysis

Growth curve analysis was used to analyze perceived laterality scores as a function of ITD and sound intensity. For each of the two noise conditions, the perceived laterality scores were fitted with an NLME model. The model included fixed effects α and random effects β. Equation (4.2) describes a sigmoidal function linking ITD to perceived laterality, with a score from left (-1) to right (1). The effect of sound intensity on the maximal extent of lateralization is α_{y1}. To factor out across-listener

differences in absolute hearing thresholds, for each listener, we calculated the pure tone average (PTA) detection threshold in Quiet, averaged across ears, and across 500 and 1000 Hz. Weight α_{y2} models the contribution of PTA. The slope terms are α_{x1} for perceived laterality changes attributed to ITD, and α_{x2} for laterality-ITD slopes attributed to sound intensity. Our stimuli were initially calibrated to have a broadband interaural level difference of 0 dB. However, because the transfer function of our sound card was not perfectly flat across frequency, fluctuations of ±1 dB interaural level difference occurred across frequency, on the same order of magnitude as the minimal threshold for human interaural level difference discrimination [32]. Thus, parameter α_{x0} factors out central response bias from the lateralization scores. Random effects of individual differences across listeners were used to model both the maximal extent of lateralization, $\beta_{y0,listener}$, and the perceived midline, $\beta_{x0,listener}$, centering the sigmoid:

$$response \sim (\alpha_{y2} \times PTA + \alpha_{y1} \times intensity + \beta_{y0,listener}) \times \left(\frac{1}{1 + e^{-[\alpha_{x2} \times intensity + \alpha_{x1} \times (ITD - \alpha_{x0} - \beta_{x0,listener})]}} - 0.5 \right). \quad (4.2)$$

To better conform with the assumptions of the NLME model, prior to fitting, ITD and sound intensity parameters were scaled by subtracting the mean stimulus value, and dividing by the standard deviation of stimulus parameters, resulting in distributions of stimulus parameters with zero-mean and a variance of one. Laterality scores were then fitted using these normalized parameters, with the nlme package, programmed in RStudio 1.1 for Windows (RStudio Inc., Boston, MA, USA).

4.5.3 Data and Software Availability

All data and analysis code are available at Dryad: doi:10.5061/dryad.t8c381f.

4.6 Discussion

Population rate-coding to compute sensory dimension may not be unique to the auditory system. Similar to sound localization based on the comparison of signals from the two ears (Figure 4.3A), visual depth is computed in the cerebral cortex based on signals from the two eyes (Figure 4.3B; [100, 94, 93]). Specifically, in both primary V1 and extrastriate V3a cortex of rhesus macaque monkeys, three types of neurons are thought to encode binocular disparity. "Tuned-excitatory" neurons respond best to zero spatial disparity between the two eyes, whereas "near cells" responds more vigorously when an object approaches, increasing crossed disparity between the eyes [94]. Finally, "far cells" fire more vigorously as uncrossed disparity increases. In V1, the most frequently encountered type of binocular neurons are of the tuned-excitatory type. However, in V3a the large majority of neurons is stereo-specific [101] and most neurons are either near or far cells. Functional magnetic resonance imaging experiments on human stereoscopic vision found that unlike V1 activity, the activity in cortical area V3a predicts behavioral performance on tasks involving stereoscopic depth [5]. These observations lead us to propose that in order to compute perceptual space from sensory input, the central nervous system has evolved a canonical computation that is common to different sensory modalities. Specifically, we propose that near and far cells encode visual distance from the fixation plane in a way similar to how inferior colliculus neurons encode auditory azimuthal angle away from midline reference: firing rate increases monotonically with distance from perceptual reference anchor or fixation.

We observe that in both the auditory and the visual system, the same cells that are tuned to binaural ITD or binocular disparity also have intensity-response functions. A rate-code based on a population of these cells should cause ambiguities when stimulated below the saturation firing rate, either at low sound intensity or at low contrast (Figure 4.3C). Thus, based on the analogies between the stereo-depth

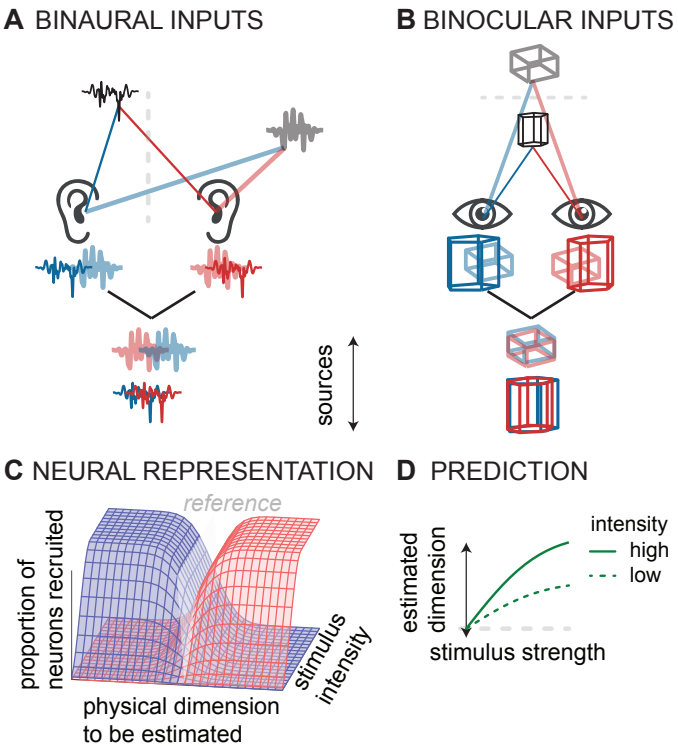

Figure 4.3 **A** Computing sound direction requires analysis of the binaural difference between the signals reaching the left and right ear. **B** Estimating visual depth hinges on analysis of the binocular disparity between the signals reaching left and right eye. **C** For both hearing and vision, the proportion of the neural population that is stimulated (in the *inferior colliculus* or *V3*) depends both on the physical dimension to be estimated (source laterality or source distance) and the intensity of the stimulus (sound intensity or visual contrast). For hearing and vision, ambiguity in this putative neural code predicts **D** biased responses at low stimulus intensities (sound intensity or contrast).

computation and the azimuth-ITD computation, we hypothesized that low visual contrast might affect the computation of depth in a manner analogous to the effect of low sound levels in sound localization—there might be a bias to lower perceived depth at lower contrast (Figure 4.3D). Indeed, one study found such an effect, but only in some observers [21]. A confounding factor in that earlier study is that perceived depth is a complicated neural computation, not only dependent on stereoscopic disparity but also on monocular cues including contrast [93]. Several studies on depth perception indicate that low contrast is interpreted by the brain as a cue for distance; lower contrast targets are perceived farther away [114, 109]. However, experiments that controlled for low contrast bias demonstrated that low contrast causes perceived depth to shrink, both for near and far deviations from baseline [19]. Thus, there is a link between population rate-coding and stimulus intensity in perceived visual depth as in perceived auditory azimuth, two perceptual spatial dimensions computed by the brain.

To illustrate how rate-based decoding of target location varies with sound intensity, we here chose a rate-coding model that compares firing rates across two populations of neurons, tuned to opposite hemifields. This read-out is a direct realization of the original canonical rate-based model for ITD decoding [126]. Alternative rate-code readouts exist (for a recent summary of binaural models, see [28]). Most of these rate-code models rely on subtractive comparisons between populations of neurons that are tuned to opposite hemifields, inherently sharing ambiguous readouts at low super threshold sound intensities. In contrast, divisive comparisons between ipsi- and contralaterally tuned neural populations are less likely to predict the observed behavioral bias due to stimulus intensity [39]. Future work will need to delineate how specific implementations of rate-based readouts shape the intensity-induced bias of sound localization. Moreover, it has been suggested that depending on perceptual task, the mammalian brain could combine place-

and rate-codes [103, 37]. For instance, the mammalian auditory pathway may convert place- into rate-codes and vice versa [40, 103]. Even though, downstream from the inferior colliculus, rate-coding seems to be maintained, at least in the superior colliculus of rhesus macaque [130, 74]. Moreover, our psychophysical and computational results suggest that for sound localization based on ITD at low sound levels, cortical maps do not play a role. There is good evidence for spatial map-like signals in higher order auditory cortical fields when interaural level differences are present, at medium to high sound levels [51]. How these ILD-based cortical maps influence sound localization behavior is yet to be determined.

An additional factor restricting rate-based readouts is that auditory cortex units display non-linear rate-intensity functions. For instance, excitatory-excitatory (EE) cells in auditory cortex that are tuned to sound locations near midline are also often tuned for sound intensity [115, 102, 138, 108, 51]. This intensity tuning may complicate rate-based decoding at higher sound intensities. Although, it is not apparent at the very low sound intensities needed to explain the perceptual bias observed here. There are additional fascinating findings in the neurophysiological literature regarding frequency and intensity tuning, and interesting correlations between non-monotonicity in the azimuthal and intensity dimensions [132], but a detailed discussion of these points is beyond the scope of the present behavioral-computational study.

In summary, unlike predictions from a rate-code neuronal readout, labelled-line coding predicts that sound localization is intensity invariant. Our experimental results show that for low frequency noise, where ITDs are the dominant localization cue, and at low sound intensities, sound lateralization based on ITD is not intensity invariant; it becomes increasingly medially biased with decreasing SL. The observed localization bias is overall small in magnitude, showing that the brain can robustly localize based on ITD across a large range of sound intensities. However, this bias is of theoretical

importance as it confirms the prediction of a subtractive rate-based neuronal readout. Moreover, our auditory finding parallels a phenomenon of visual fixation bias when calculating visual distance from binocular disparity at low contrast. This casts doubt on the idea that the neural mechanism of ITD-based sound localization and binocular disparity-based visual distance estimation are based on place-based coding. Instead, our perceptual data on auditory localization together with previously published data on visual distance perception are parsimonious with the idea that a population rate-code underlies the brain's computation of location.

CHAPTER 5

LEVERAGING ADAPTATION TO STUDY PERCEPTUAL WEIGHTING OF INTERAURAL TIME DIFFERENCES

An important question in auditory cognition is how we perceive the location of an object in space. Converging evidence from animal models and humans suggests that when judging sound direction, the central nervous system weighs the anticipated reliability of binaural cues. Here, we used short-term adaptation to bias normal-hearing listeners towards source direction favoring either the left or the right frontal quadrant. Listeners rated perceived laterality of tokens of band-pass filtered noise (300 Hz – 1200 Hz) with interaural time differences that were randomly selected from a uniform distribution spanning either -375 to 0 μs or 0 to 375 μs. Using non-linear mixed effects modeling of behavioral laterality reports, we tested how exposure to source quadrant affects how listeners weigh the reliability of interaural time differences. The cue reliability hypothesis predicts that perceived direction should be skewed, such that unreliable frontal source angles are more affected by short-term adaptation than the more reliable lateral source angles. Alternatively, short-term adaptation may affect all source angles equally, predicting an overall shift in perceived direction. Results show that frontal angles are more strongly affected by short-term adaptation than lateral angles, supporting the cue reliability hypothesis.

5.1 Introduction

A wide range of species relies on sound localization for both navigation and auditory scene analysis. For low-frequency sound, interaural time differences (ITDs) are the dominant cue for determining the direction of a source in the horizontal plane, a phenomenon called sound lateralization. The mechanisms by which the nervous system maps ITD into perceived sound direction are incompletely understood.

In anechoic spaces, a given source angle in the horizontal plane typically gives rise to the same ITD, across a wide range of source distances. Although, in everyday environments where background sound and reverberant energy are often present, ITDs are much less reliable indicators of source direction. This raises the possibility that when estimating source direction, a listener's interpretation of ITD changes depending on the context of the listening environment. Indeed, previous work shows plasticity in perceived sound direction across a wide range of conditions, including: 1) Prolonged exposure to constant interaural delays [64], 2) Modifications to shape of the pinnae [54], 3) Long-term procedural learning [133], 4) Chronic unilateral ear plugging [71], 5) Short-term adaptation to stimulus history [120], 6) In presence of preceding distractors [70].

However, we have an incomplete understanding of short-term adaptation of sound lateralization based on the overall uncertainty of ITDs. The steepest point of the psychometric function linking ITD and perceived laterality is midline (the frontal direction). Thus, a small change in ITD around 0 μs baseline ITD causes a much larger change in perceived laterality than a small change in ITD around 375 μs ITD. Thus, the cue reliability hypothesis predicts that perceived direction should be skewed, such that unreliable frontal source angles are more affected by short-term adaptation than the more reliable lateral source angles. Alternatively, short-term adaptation may affect all source angles equally, predicting an overall shift in perceived direction. Leveraging short-term plasticity based on stimulus context, we set out to test the cue reliability hypothesis in human psychophysics experiments. We varied context by blocking stimuli such that they would predominantly arise from the left, from the right or from both sides of the head. Normal-hearing listeners judged lateralization of band-limited noise tokens as a function of interaural time difference and context.

5.2 Methods

Overall methods were similar to previous work (for detailed description see: [58]). We tested 34 naïve normal-hearing listeners, 16 females, age 20-30. Audiometric thresholds were 20 dB HL or better at all octave frequencies and did not differ by more than 10 dB across ears at each octave frequency. We divided our listeners into four groups. Using a target matching task without response feedback, we initially trained three groups to be reliable reporters. Listeners trained and tested on the frontal 1) left hemisphere (LEFT HEMI, N=8) or, 2) right hemisphere (RIGHT HEMI, N=8) or, 3) both hemispheres (BOTH HEMIS, N=10). A fourth control group, NAÏVE RIGHT HEMI (N=8) was tested on the right frontal hemisphere without training.

Listeners were seated in a double-walled sound-attenuating booth. Sounds were presented through ER-2 earphones. Stimulus consisted of a band-limited random noise token, generated independently on each trial (300 to 1200 Hz, 1 sec duration, 10 ms cos-squared onset/offset ramps, 192 kHz sampling frequency, 24 bit resolution, 0 dB ILD). ITDs covered the range of 0, $\pm$75, $\pm$150, $\pm$225, $\pm$300 or $\pm$375 μs. Sound levels varied from 10 to 40 dB sensation level (SL).

Prior to testing all sounds were calibrated using KEMAR manikin and B&K 2250 with 2-cc coupler. This study was approved by the NJIT Institutional Review Board.

All listeners performed three training sessions (except the naive group) and then a single session of testing. Each group were presented with their associated range of ITDs, i.e., RIGHT HEMI: positive ITDs only (including 0), LEFT HEMI: negative ITDs only (including 0), BOTH HEMI: full range of ITDs.

To analyze the listeners' responses, a non-linear mixed-effects model (NLME) was developed that captured the sigmoidal sensitivity of ITD lateralization (for details, see [58]). The level-1 model formula shown in Equation (5.1) is a shifted logistic curve (adapted from [76]). Here, fixed effects are designated with β as their

coefficient, and random effects with b. β_3 models the perceived laterality of the sounds. Larger values for β_3 are interpreted as a wider lateralization range with all ITDs being perceived further away from the midline. β_2 captures the slope of the sigmoidal curve, that is the rate at which increasing ITDs map to more lateral directions. Larger slopes do not influence the midline anchor, nor they change the lateralization range, but they push the perceived direction of smaller ITDs laterally. β_1 is the ITD bias that models the shift to the stimulus ITD that was heard by the listener. For example, the perceived shift in the ITD of the current trial due to ITD of the previous trial was modeled by this coefficient. β_0 captures the laterality bias, which is a constant shift in the perceived direction for all ITDs. Since other coefficients also contribute to fitting changes in direction, this bias is mainly observed as a shift in the perceived midline for stimuli with a 0 μs ITD.

$$\text{Response} = (\beta_3 + b_3)\left(\frac{1}{1 + e^{-\beta_2(ITD+\beta_1+b_1)}} - 0.5 + \beta_0\right) + \varepsilon. \tag{5.1}$$

The level-2 model formulas for the fixed effects (coefficients β_0 - β_3) are linear sums of the intercept, the main effects, and the interactions associated with each coefficient (Equation (5.2)). "Hemi" is an encoded variable indicating the training hemisphere for each listener group (-1 for left, 0 for both, and +1 for right). All variables except "Hemi" were normalized as z-scores. "ΔITD_{-1}" is the difference between the current trial's ITD and the 1-back ITD. This value was positive for trials which had an ITD larger than (to the right of) their previous trial. "Block" indicates the block number throughout a testing session which a trial was presented in.

$$\begin{aligned}
\beta_0 &= \beta_{00} + \beta_{01}\text{Hemi} + \beta_{02}\text{Intensity} + \beta_{03}\text{Hemi} \times \text{Intensity},\\
\beta_1 &= \beta_{10} + \beta_{11}\Delta\text{ITD}_{-1} + \beta_{12}\text{Intensity} + \beta_{13}\text{Intensity} \times \Delta\text{ITD}_{-1},\\
\beta_2 &= \beta_{20},\\
\beta_3 &= \beta_{30} + \beta_{31}\text{Intensity} + \beta_{32}\text{Block}.
\end{aligned} \tag{5.2}$$

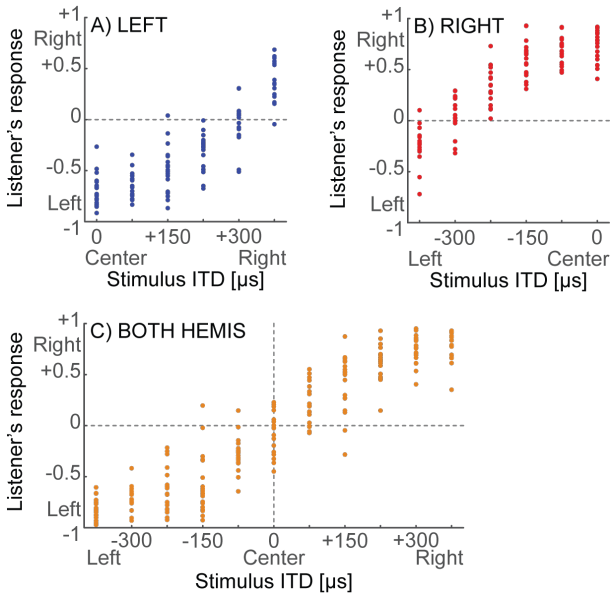

Figure 5.1 Raw lateralization scores from one representative listener.

5.3 Results

Figure 5.1 shows the results of a representative listener. Results follow the expected sigmoidal trend. Figure 5.2A shows the responses of all 34 participants in the lateralization task. Circles show the raw across-listener average for each corresponding ITD and listener group, with error bars showing one standard error of the mean across listeners. Lines are the fitted model value, with ribbons indicating the predicted standard error of the mean fit. Figure 5.2B shows the slopes of these fits as a function of ITD.

Analysis of listener responses with NLME (see Table 5.1) revealed a main effect of training hemisphere (β_{01}, $p < 0.001$) on the laterality bias. Groups that were trained and tested unilaterally on one hemisphere had a bias in their perceived midline towards the opposite hemisphere. However, the bilaterally tested listeners had no discernible bias (β_{00}, $p = 0.787$). This hemisphere-dependant laterality bias increased during trials that had a higher sound intensity (β_{03}, $p < 0.001$). For instance, listener's that were trained and tested on the right hemisphere perceived stimuli with a 0 μs

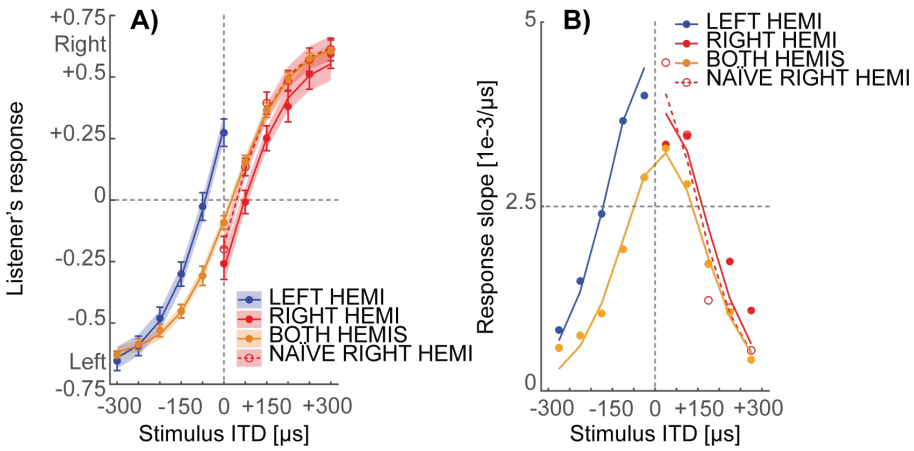

Figure 5.2 Fitted lateralization curves across all four context groups.

ITD to be coming from the left rather than the center. The bias was enhanced for louder sounds and the perceived directions was shifted even more to the left.

For ITD bias, the results show a main effect of difference with 1-back ITD (β_{11}, p < 0.001). Overall, listeners were impacted by the ITD of the previous trial's stimulus, and their responses were consistently biased towards the direction of the 1-back ITD. This bias was less pronounced during trials with a louder stimulus (β_{13}, p < 0.001). Finally, the laterality of listener responses was reduced throughout a testing session (β_{32}, p < 0.001), and all groups had a lower range of responses towards the end of a session.

5.4 Discussion

Current results support the idea that short-term plasticity of ITD is more pronounced in front than at the sides. At the least reliable cue point, 0 μs, results show robust response biases in all unilaterally tested listeners towards the opposite side, with a larger effect for the trained group versus the naïve. In contrast, lateralization at 375 μs base ITD is not appreciably affected by context. Moreover, slope of lateralization curves increases for unilateral groups compared to the bilateral, hinting at an expansion in the perceived space and increase in lateralization sensitivity. Thus,

despite the fact that humans are much more sensitive to frontal ITD than lateral ITD, as evidenced by lower just-noticable differences in ITD discrimination [14], frontal ITDs are also more susceptible to short-term adaptation than lateral ITDs, supporting the cue reliability hypothesis.

Table 5.1 NLME Results for All Bilaterally and Unilaterally Trained Groups

	Description	Estimate	Std. Error	t-value	p-value	
	Laterality Bias					
β_{00}	Intercept	0.0016	0.0058	0.27	0.787	
β_{01}	Hemisphere	-0.1353	0.0122	-11.14	<0.001	***
β_{02}	Sound Intensity	-0.0101	0.0030	-3.37	<0.001	***
β_{03}	Interaction	-0.0082	0.0020	-4.17	<0.001	***
	ITD Bias					
β_{10}	Intercept	-0.0259	0.0363	-0.71	0.476	
β_{11}	Difference with 1-Back ITD	-0.0242	0.0068	-3.53	<0.001	***
β_{12}	Sound Intensity	0.0765	0.0090	8.49	<0.001	***
β_{13}	Interaction	0.0260	0.0062	4.23	<0.001	***
	Slope					
β_{20}	Intercept	2.0647	0.0617	33.48	<0.001	***
	Laterality					
β_{30}	Intercept	1.7907	0.1185	15.11	<0.001	***
β_{31}	Sound Intensity	-0.0118	0.0118	-1.00	0.315	
β_{32}	Block Number	-0.0114	0.0017	-6.82	<0.001	***

Note: 8723 degrees of freedom.

Significance codes: '***' $p < 0.001$, '**' $p < 0.01$, '*' $p < 0.05$, '.' $p < 0.1$, ' ' $p \geq 0.1$.

CHAPTER 6

SUMMARY

This dissertation established an experimental framework for training and testing of awake and freely behaving rodents in auditory tasks. A multi-faceted system was developed for simultaneous behavioral and electrophysiological assessment, that was used to investigate neural mechanisms of hearing at negative SNRs under interference with a modulated background. Two groups of gerbils were tested, one naïve to the task, and another fully trained to detect a target tone stimulus. Computational techniques were then utilized to analyze different neural coding mechanisms that could give rise to dip listening. The results demonstrated that both rate and temporal cues can enable hearing at positive SNRs, but for hearing at negative SNRs, an SNR-invariant readout pertains dependence on temporal information. Future work is needed to understand whether the temporal information needed for dip-listening at negative SNRs can be augmented through other sensory modalities.

In human experiments, it was revealed that at sound levels near the sensation threshold, lateralization with ITDs becomes less potent. Analysis of behavioral data with NLME, a powerful statistical tool, showed a medial bias in responses to similar ITDs at very soft levels. This supports the hypothesis that MSO neurons do not act as mere cross-correlation calculators and in fact receive level dependent inputs that influences the measured ITD decision threshold. Furthermore, additional experiments with four groups of listeners showed robust response biases in all unilaterally-tested listeners towards the opposite side, with a larger effect for the trained groups versus the naïve. In contrast, bilaterally trained listeners did not display any response bias. Results are consistent with the conclusion that human sound localization relies on population spike-rate coding as opposed to a spatially tuned map.

APPENDIX

TEACHING ELECTRONIC CIRCUIT FUNDAMENTALS VIA REMOTE LABORATORY CURRICULUM

A.1 Challenge Statement

The course "Electrical Fundamentals" (EF) is a core requirement for all undergraduate students in our biomedical engineering program. The curriculum introduces general principles of device development for electronics-based bioinstrumentation, comprehensively covering foundational acquisition concepts for bioelectric signals. Laboratory-based learning modules provide hands-on experience with circuit fundamentals. We here introduce our 6-hour lab curriculum for distance learning, which we developed in response to the COVID-19 outbreak.

The learning outcomes for the lab curriculum are twofold. The first learning outcome enables students to design and analyze basic electric circuits of resistors, capacitors and operational amplifiers (op-amps). The second learning outcome teaches students to interpret signal characteristics using core bioinstrumentation equipment, including oscilloscope, function generator and multimeter. Moreover, at this stage in the EF students' curriculum, the concept of filters is still in progress. Therefore, the lab curriculum builds up and deepens the theoretical concept of low- and high-pass filters through practical experimentation.

Traditionally, our EF labs are taught in studio-style on campus where groups of students learn to use oscilloscopes, function generators, multimeters and breadboards. Here, the challenge is to replace these precision instruments with affordable, reliable and portable components that students could safely and robustly assemble at home. Moreover, the material needs to be engaging to nurture student interest. Furthermore, some of our students have limited access to internet. To enable accessibility to all students, labs need to accommodate asynchronous learning. Finally, due to initial

uncertainty during the COVID-19 outbreak, the lab curriculum needs to be designed such that single students can complete lab exercises independently, as opposed to working in groups.

A.2 Novel Initiative

The remote lab curriculum is taught synchronously in three two-hour sessions. A ratio of 10 students per teacher is desirable for effective real-time trouble-shooting. Therefore, for our large undergraduate cohort, we recruit "peer mentors," in addition to the EF instructor and teaching assistants. Peer mentors are outstanding undergraduate students who have previously taken the EF class, who are being mentored by the principal instructor, and who are now helping their peers with hands-on circuit troubleshooting.

Lessons utilize the Arduino Uno, a versatile tool that can be robustly employed in undergraduate engineering education ranging from foundational instruction to project-based learning and student competitions (e.g., [88, 63, 45]). Here, the Arduino replaces three traditional pieces of laboratory-learning equipment, 1) the multimeter, 2) the function generator and 3) the multichannel oscilloscope [2]. We provide students with a customized Arduino kit, including breadboard and all circuit components. In addition, each student needs access to a laptop, internet and webcam.

For optimal learning outcomes, the current curriculum encourages synchronous learning. However, lab instructions are self-contained, enabling students to follow them independently should poor internet connectivity necessitate asynchronous learning. At the beginning of each lab session, the instructor demonstrates the following principles via synchronous teaching: how to assemble the circuit, how to upload code to the Arduino chip, how to use SerialPlotter to view and interpret time series data read in by Arduino, and how to further analyze these time series data in Excel and Matlab.

Following initial instruction, students then independently work through the lab instructions. Work is self-paced, but students are encouraged to finish within the two-hour time frame of the lab. To facilitate timely completion and to avoid having students get stuck on their own, the lab instructions provide milestones, frequently prompting students to submit photos of their circuit as well as numerical values or plotted time series, via private messaging to the instructional team. As problems arise, teachers then trouble-shoot with the students via synchronous video chat. Concepts taught by each milestone are completed before a new concept is introduced, reducing cognitive load and helping learning and retention (e.g., [67]).

The curriculum consists of three take-home laboratory exercises. Lab 1 instructs students on component tolerances [69] and reviews basic circuit knowledge. Students learn how to use the Arduino as a multimeter to measure voltages, currents, and verify resistances of parallel-serial resistor circuits. In Lab 2, the Arduino functions both as function generator and multi-channel oscilloscope. Students learn about charge and discharge behavior of resistor-capacitor (RC) circuits and the time constant tau, laying the foundation for the concept of low-pass filtering. Lab 3 leverages sonification to teach active amplification via operational amplifiers (op-amp) circuits. Op-amps are a challenging new concept for the majority of EF students. To motivate students, here, the output of the circuit is fed into a loudspeaker, allowing students to experience the tuning of RC circuits by hearing how the circuit filters sound, as well as listen to how loud music sounds when it is processed by op-amp circuits. All three lab exercises, as well as all components needed, are comprehensively explained in the laboratory instructions (see Supplemental Materials).

To illustrate the overall concept of Arduino-based learning of circuit fundamentals, we next explain Lab 2 in detail. The learning goal of Lab 2 is to understand charging and discharging of resistor-capacitor (RC) circuits. Students learn to use the voltage traces over R or C to estimate and verify the time constant τ, while also

building up intuition about temporal and spectral filtering properties of the circuit. To illustrate temporal filtering properties of a serial RC circuit, students record both the generated input pulses and the voltage output of the circuit over R or C . Students then plot the charging and discharging curve, take fiduciary markers and estimate the decay constant τ. Students discover that for sufficiently slow pulse rates, the capacitor has enough time to charge and discharge. However, this behavior breaks down as the pulse rate increases. To estimate the maximum pulse rate that the C can robustly follow, students increase pulse rate on the function generator, measuring the peak to peak voltage of the capacitor, $V_{C_{pp}}$, on the oscilloscope. Students then plot $V_{C_{pp}}$ as a function of pulse rate and observe how the $V_{C_{pp}}$ amplitude decreases with increasing pulse rate.

Traditionally, studio-style instruction in Lab 2 uses a function generator to send voltage traces as inputs to the circuit, consisting of pulses of variable repetition rate. Students then trace the circuit with a multichannel oscilloscope. Here, the Arduino functions both as multichannel oscilloscope and simplified function generator, using the same sampling period for both, with a precision of ± 50 microseconds. Specifically, using our custom code for Lab 2, the Arduino delivers a train of rectangular pulses (0 or 5 V, with 50% duty cycle and 100% modulation depth, from Arduino pin A1). An output voltage, across R or C, is simultaneously read in by the Arduino, sample by sample (from pin A0). The voltage traces of both the generated pulse train and the recorded measurements are then sent to the student's laptop via serial port. SerialPlotter, which is part of Arduino IDE on the laptop, then plots both voltage traces sample by sample, generating the type of trace that students would traditionally observe on a multi-channel oscilloscope. The numerical values of these voltage traces can then be copy-pasted into Excel or Matlab where students can use them to run detailed analysis, including estimating peak-to-peak voltages and time constants of the circuit.

Prior to uploading the code onto the Arduino chip, students can modify four parameters in the code: 1) the sampling period, 2) the total number of samples that the Arduino records and sends out (this limits the number of samples plotted, preventing the plot from rolling), 3) the pulse period (this adjusts the pulse widths of the output pulse train), and 4) the initial state of the capacitor (charged vs discharged). On their personal computer or laptop at home, every student opens the Arduino IDE, inspects the customized code and programs it via the "upload" IDE function onto the Arduino chip. To prevent lack of coding skills from interfering with the learning goals of the remote lab curriculum, students do not need to code. However, they are encouraged to inspect the Arduino code and special breakout sessions are offered to explain the code to interested students.

A.3 Reflection

The new remote training curriculum was implemented during Spring and Summer 2020 with two different instructors and 70 students. All students were able to complete all sections of all labs. Due to the ad hoc nature of this learning intervention, no formal learning outcome data were collected. Instead, student satisfaction was gathered from informal feedback, suggesting that the remote lab curriculum resulted in self-reported skill gains and overall high satisfaction.

A number of students commented that it felt empowering to build and trace circuits at home, with affordable materials, as compared to using expensive high-fidelity devices at the university. Consistent with prior work on enriched learning, the inclusion of a sonification experience via their own self-made loudspeaker to validate circuit functionality was particularly rewarding to many students [116]. Going forward, we will routinely use Arduino-based learning as a first step before introducing the professional oscilloscope and function generator devices.

Of note, at the onset of the pandemic, not all of our students had access to webcams via their laptops. Instead of live circuit trouble-shooting via webcam, these students emailed us photographs of their circuits instead. However, we discovered that, aside from being an inefficient use of instructor time, email slows down the learning momentum [98]. We now mandate webcam-enabled laptops for our course.

Our undergraduate cohort includes a high proportion of underrepresented minority students, providing the kind of environment where utilization of peer mentors is a particularly promising method for nurturing BME undergraduate students in their professional and personal development [86]. Here, peer mentors allowed for effective trouble-shooting during labs. Moreover, peer mentors made students more comfortable to ask basic questions as well as help maintain a sense of community during social isolation due to the COVID-19 crisis. Furthermore, the peer mentors were able to deepen their knowledge of electronics by learning through teaching their peers. Both the students in class and the peer mentors reflected very positively on this experience.

In summary, when student access was limited during COVID-19, a novel Arduino-based intervention enabled remote teaching while satisfying original learning outcomes of the EF course. Students were able to measure key parameters of serial-parallel resistors, RC-circuits and op-amp circuits. Experiential learning via a custom Arduino kit demonstrated to students how they can build simple electronic circuits at home, while maintaining accessibility of EF teaching to all students. A lab format of initial instruction followed by prompting for "fill-in-the-blanks" milestone answers proved effective for remote learning with synchronous trouble-shooting of circuits. The use of sonification and peer mentoring further enriched student motivation. Informal course evaluations suggest that many students benefited from the remote lab curriculum. Future work will need to formally test teaching effectiveness [38, 11].

A.4 Supplemental Materials

All lab instructions and software tools are available on GitHub[1] [2].

[1]Lab instructions and Arduino source code: https://github.com/nalamat/njit-bme-301

REFERENCES

[1] Nima Alamatsaz and Antje Ihlefeld. EARS: Electrophysiology auditory recording system (v1.0). *Zenodo*, June 2019.

[2] Nima Alamatsaz and Antje Ihlefeld. Remote laboratory curriculum for electrical fundamentals of biomedical engineering. *Zenodo*, July 2020.

[3] Elizabeth S Anderson, David A Nelson, Heather Kreft, Peggy B Nelson, and Andrew J Oxenham. Comparing spatial tuning curves, spectral ripple resolution, and speech perception in cochlear implant users. *The Journal of the Acoustical Society of America*, 130(1):364–375, 2011.

[4] Harvey Babkoff and Samuel Sutton. End point of lateralization for dichotic clicks. *The Journal of the Acoustical Society of America*, 39(1):87–102, 1966.

[5] Benjamin T Backus, David J Fleet, Andrew J Parker, and David J Heeger. Human cortical activity correlates with stereoscopic depth perception. *Journal of Neurophysiology*, 86(4):2054–2068, 2001.

[6] Ranjan Batra and Tom CT Yin. Cross correlation by neurons of the medial superior olive: a reexamination. *Journal of the Association for Research in Otolaryngology*, 5(3):238–252, 2004.

[7] Joshua GW Bernstein and Douglas S Brungart. Effects of spectral smearing and temporal fine-structure distortion on the fluctuating-masker benefit for speech at a fixed signal-to-noise ratio. *The Journal of the Acoustical Society of America*, 130(1):473–488, 2011.

[8] Joshua GW Bernstein and Ken W Grant. Auditory and auditory-visual intelligibility of speech in fluctuating maskers for normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 125(5):3358–3372, 2009.

[9] Peter A Bohlen, Margit E Dylla, Courtney Timms, and Ramnarayan Ramachandran. Detection of modulated tones in modulated noise by non-human primates. *Jaro-journal of The Association for Research in Otolaryngology*, 2014.

[10] Hemant Bokil, Peter Andrews, Jayant E Kulkarni, Samar Mehta, and Partha P Mitra. Chronux: a platform for analyzing neural signals. *Journal of Neuroscience Methods*, 192(1):146–151, 2010.

[11] Maura Borrego and Charles Henderson. Increasing the use of evidence-based teaching in stem higher education: A comparison of eight change strategies. *Journal of Engineering Education*, 103(2):220–252, 2014.

[12] Antje Brand, Oliver Behrend, Torsten Marquardt, David McAlpine, and Benedikt Grothe. Precise inhibition is essential for microsecond interaural time difference coding. *Nature*, 417(6888):543, 2002.

[13] Jeroen Breebaart, Steven Van De Par, and Armin Kohlrausch. Binaural processing model based on contralateral inhibition. i. model structure. *The Journal of the Acoustical Society of America*, 110(2):1074–1088, 2001.

[14] Andrew Brughera, Larisa Dunai, and William M Hartmann. Human interaural time difference thresholds for sine tones: The high-frequency limit. *The Journal of the Acoustical Society of America*, 133(5):2839–2855, 2013.

[15] Douglas S Brungart and Brian D Simpson. Effects of temporal fine structure on the localization of broadband sounds: potential implications for the design of spatial audio displays. In *Proceedings of the 14 International Conference on Auditory Display, Paris, France June 24 - 27, 2008*, pages 1–5. International Community for Auditory Display, 2008.

[16] Bradley N Buran, Gardiner von Trapp, and Dan H Sanes. Behaviorally gated reduction of spontaneous discharge can improve detection thresholds in auditory cortex. *Journal of Neuroscience*, 34(11):4076–4081, 2014.

[17] Melissa L Caras and Dan H Sanes. Top-down modulation of sensory cortex gates perceptual learning. *Proceedings of the National Academy of Sciences*, 114(37):9972–9977, 2017.

[18] Catherine E Carr and Masakazu Konishi. Axonal delay lines for time measurement in the owl's brainstem. *Proceedings of the National Academy of Sciences*, 85(21):8311–8315, 1988.

[19] Pei-Yin Chen, Chien-Chung Chen, and Christopher W Tyler. The perceived depth from disparity as function of luminance contrast. *Journal of Vision*, 16(11):20–20, 2016.

[20] Kate L Christison-Lagay, Sharath Bennur, and Yale E Cohen. Contribution of spiking activity in the primary auditory cortex to detection in noise. *Journal of Neurophysiology*, 2017.

[21] Patricia M Cisarik and Ronald S Harwerth. The effects of interocular correlation and contrast on stereoscopic depth magnitude estimation. *Optometry and Vision Science*, 85(3):164–173, 2008.

[22] Harry Steven Colburn and NI Durlach. Time-intensity relations in binaural unmasking. *The Journal of the Acoustical Society of America*, 38(1):93–103, 1965.

[23] Martin Cooke. A glimpsing model of speech perception in noise. *The Journal of the Acoustical Society of America*, 119(3):1562–1573, 2006.

[24] Torsten Dau, Stephan Ewert, and Andrew J Oxenham. Auditory stream formation affects comodulation masking release retroactively. *The Journal of the Acoustical Society of America*, 125(4):2182–2188, 2009.

[25] Mitchell L Day and Malcolm N Semple. Frequency-dependent interaural delays in the medial superior olive: implications for interaural cochlear delays. *American Journal of Physiology-Heart and Circulatory Physiology*, 2011.

[26] Sasha Devore, Antje Ihlefeld, Kenneth Hancock, Barbara Shinn-Cunningham, and Bertrand Delgutte. Accurate sound localization in reverberant environments is mediated by robust encoding of spatial cues in the auditory midbrain. *Neuron*, 62(1):123–134, 2009.

[27] Mathias Dietz, Leslie R Bernstein, Constantine Trahiotis, Stephan D Ewert, and Volker Hohmann. The effect of overall level on sensitivity to interaural differences of time and level at high frequencies. *The Journal of the Acoustical Society of America*, 134(1):494–502, 2013.

[28] Mathias Dietz, Jean-Hugues Lestang, Piotr Majdak, Richard M Stern, Torsten Marquardt, Stephan D Ewert, William M Hartmann, and Dan FM Goodman. A framework for testing and comparing binaural models. *Hearing research*, 360:92–106, 2018.

[29] Robert H Domnitz and H Steven Colburn. Lateral position and interaural discrimination. *The Journal of the Acoustical Society of America*, 61(6):1586–1598, 1977.

[30] Nathaniel I Durlach. Equalization and cancellation theory of binaural masking-level differences. *The Journal of the Acoustical Society of America*, 35(8):1206–1218, 1963.

[31] Jos J Eggermont. Acquired hearing loss and brain plasticity. *Hearing Research*, 343:176–190, 2017.

[32] Tom Francart and Jan Wouters. Perception of across-frequency interaural level differences. *The Journal of the Acoustical Society of America*, 122(5):2826–2831, 2007.

[33] Yan Gai, Janet L Ruhland, Tom CT Yin, and Daniel J Tollin. Behavioral and modeling studies of sound localization in cats: effects of stimulus level and duration. *American Journal of Physiology-Heart and Circulatory Physiology*, 2013.

[34] Otto Gleich, Malte C Kittel, Georg M Klump, and Jürgen Strutz. Temporal integration in the gerbil: the effects of age, hearing loss and temporally unmodulated and modulated speech-like masker noises. *Hearing Research*, 224(1-2):101–114, 2007.

[35] Jay M Goldberg and Paul B Brown. Functional organization of the dog superior olivary complex: an anatomical and electrophysiological study. *Journal of Neurophysiology*, 31(4):639–656, 1968.

[36] Jay M Goldberg and Paul B Brown. Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. *Journal of Neurophysiology*, 32(4):613–636, 1969.

[37] Dan FM Goodman, Victor Benichoux, and Romain Brette. Decoding neural responses to temporal cues for sound localization. *eLife*, 2:e01312, 2013.

[38] James E Groccia and William Buskist. Need for evidence-based teaching. *New Directions for Teaching and Learning*, 2011(128):5–11, 2011.

[39] Jennifer M Groh. Converting neural signals from place codes to rate codes. *Biological Cybernetics*, 85(3):159–165, 2001.

[40] Jennifer M Groh, Kristin A Kelly, and Abigail M Underhill. A monotonic code for sound azimuth in primate inferior colliculus. *Journal of Cognitive Neuroscience*, 15(8):1217–1231, 2003.

[41] John H. Grose, Joseph W. Hall, Emily Buss, and Debora R. Hatch. Detection of spectrally complex signals in comodulated maskers: Effect of temporal fringe. *Journal of the Acoustical Society of America*, 2005.

[42] Benedikt Grothe. New roles for synaptic inhibition in sound localization. *Nature Reviews Neuroscience*, 4(7):540, 2003.

[43] Benedikt Grothe and Michael Pecka. The natural history of sound localization in mammals–a story of neuronal inhibition. *Frontiers in Neural Circuits*, 8:116, 2014.

[44] Benedikt Grothe, Michael Pecka, and David McAlpine. Mechanisms of sound localization in mammals. *Physiological Reviews*, 90(3):983–1012, 2010.

[45] Radhika Grover, Shoba Krishnan, Terry Shoup, and Maryam Khanbaghi. A competition-based approach for undergraduate mechatronics education using the arduino platform. In *Fourth Interdisciplinary Engineering Design Education Conference*, pages 78–83. IEEE, 2014.

[46] Ervin R Hafter. Quantitative evaluation of a lateralization model of masking-level differences. *The Journal of the Acoustical Society of America*, 50(4B):1116–1122, 1971.

[47] Joseph W Hall and John H Grose. Effect of otitis media with effusion on comodulation masking release in children. *Journal of Speech, Language, and Hearing Research*, 37(6):1441–1449, 1994.

[48] William M Hartmann, Brad Rakerd, and Aaron Koller. Binaural coherence in rooms. *Acta Acustica United with Acustica*, 91(3):451–462, 2005.

[49] William Morris Hartmann and Brad Rakerd. Auditory spectral discrimination and the localization of clicks in the sagittal plane. *The Journal of the Acoustical Society of America*, 94(4):2083–2092, 1993.

[50] Itai Hershenhoren and Israel Nelken. Detection of tones masked by fluctuating noise in rat auditory cortex. *Cerebral Cortex*, 2016.

[51] Nathan C Higgins, Douglas A Storace, Monty A Escabí, and Heather L Read. Specialization of binaural responses in ventral auditory cortices. *Journal of Neuroscience*, 30(43):14522–14532, 2010.

[52] Daniel N Hill, Samar B Mehta, and David Kleinfeld. Quality metrics to accompany spike sorting of extracellular signals. *Journal of Neuroscience*, 31(24):8699–8705, 2011.

[53] Sally E Hind, Rachel Haines-Bazrafshan, Claire L Benton, Will Brassington, Beverley Towle, and David R Moore. Prevalence of clinical referrals having hearing thresholds within normal limits. *International Journal of Audiology*, 50(10):708–716, 2011.

[54] Paul M Hofman, Jos GA Van Riswick, and A John Van Opstal. Relearning sound localization with new ears. *Nature Neuroscience*, 1(5):417–421, 1998.

[55] Laura K Holden, Charles C Finley, Jill B Firszt, Timothy A Holden, Christine Brenner, Lisa G Potts, Brenda D Gotter, Sallie S Vanderhoof, Karen Mispagel, Gitry Heydebrand, et al. Factors affecting open-set word recognition in adults with cochlear implants. *Ear and Hearing*, 34(3):342, 2013.

[56] Kathryn Hopkins and Brian CJ Moore. The contribution of temporal fine structure to the intelligibility of speech in steady and modulated noise. *The Journal of the Acoustical Society of America*, 125(1):442–446, 2009.

[57] Larry E Humes. Do 'auditory processing'tests measure auditory processing in the elderly? *Ear and Hearing*, 26(2):109–119, 2005.

[58] Antje Ihlefeld, Nima Alamatsaz, and Robert M Shapley. Human sound localization depends on sound intensity: Implications for sensory coding. *bioRxiv*, page 378505, 2018.

[59] Antje Ihlefeld, Yi-Wen Chen, and Dan H Sanes. Developmental conductive hearing loss reduces modulation masking release. *Trends in Hearing*, 20:2331216516676255, 2016.

[60] Antje Ihlefeld and Barbara G Shinn-Cunningham. Effect of source spectrum on sound localization in an everyday reverberant room. *The Journal of the Acoustical Society of America*, 130(1):324–333, 2011.

[61] Antje Ihlefeld, Barbara G. Shinn-Cunningham, and Robert P. Carlyon. Comodulation masking release in speech identification with real and simulated cochlear-implant hearing. *Journal of the Acoustical Society of America*, 2012.

[62] Jinro Inoue. Effects of stimulus intensity on sound localization in the horizontal and upper-hemispheric median plane. *Journal of University of Occupational and Environmental Health*, 23(2):127–138, 2001.

[63] Peter Jamieson and Jeff Herdtner. More missing the boat—arduino, raspberry pi, and small prototyping boards and engineering education needs them. In *2015 IEEE Frontiers in Education Conference (FIE)*, pages 1–6, 2015.

[64] Amin R Javer and Dietrich W Schwarz. Plasticity in human directional hearing. *The Journal of Otolaryngology*, 24(2):111–117, 1995.

[65] Lloyd A Jeffress. A place theory of sound localization. *Journal of Comparative and Physiological Psychology*, 41(1):35, 1948.

[66] Philip X Joris, Bram Van de Sande, Dries H Louage, and Marcel Van der Heijden. Binaural and cochlear disparities. *Proceedings of the National Academy of Sciences*, 103(34):12917–12922, 2006.

[67] Slava Kalyuga. Cognitive load theory: How many types of load does it really need? *Educational Psychology Review*, 23(1):1–19, 2011.

[68] Alboukadel Kassambara. *rstatix: Pipe-Friendly Framework for Basic Statistical Tests*, 2020. R package version 0.7.0.

[69] Michael F Kavanaugh. Including the effects of component tolerances in the teaching of courses in introductory circuit design. *IEEE Transactions on Education*, 38(4):361–364, 1995.

[70] Norbert Kopčo, Gabriela Andrejkova, Virginia Best, and Barbara Shinn-Cunningham. Streaming and sound localization with a preceding distractor. *The Journal of the Acoustical Society of America*, 141(4):EL331–EL337, 2017.

[71] Daniel P Kumpik, Oliver Kacelnik, and Andrew J King. Adaptive reweighting of auditory localization cues in response to chronic unilateral earplugging in humans. *Journal of Neuroscience*, 30(14):4883–4894, 2010.

[72] Ulrike Langemann and Georg M. Klump. Signal detection in amplitude-modulated maskers. i. behavioural auditory thresholds in a songbird. *European Journal of Neuroscience*, 2001.

[73] Liora Las, Edward A Stern, and Israel Nelken. Representation of tone in fluctuating maskers in the ascending auditory system. *The Journal of Neuroscience*, 2005.

[74] Jungah Lee and Jennifer M Groh. Different stimuli, different spatial codes: a visual map and an auditory rate code for oculomotor space in the primate superior colliculus. *PloS one*, 9(1):e85017, 2014.

[75] M Charles Liberman, Leslie D Liberman, and Stéphane F Maison. Chronic conductive hearing loss leads to cochlear degeneration. *PloS one*, 10(11):e0142341, 2015.

[76] Mary J Lindstrom and Douglas M Bates. Nonlinear mixed effects models for repeated measures data. *Biometrics*, pages 673–687, 1990.

[77] Ruth Y Litovsky. Spatial release from masking. *Acoustics Today*, 8(2):18–25, 2012.

[78] Ruth Y Litovsky, Matthew J Goupell, Shelly Godar, Tina Grieco-Calub, Gary L Jones, Soha N Garadat, Smita Agrawal, Alan Kan, Ann Todd, Christi Hess, et al. Studies on bilateral cochlear implants at the university of wisconsin's binaural hearing and speech laboratory. *Journal of the American Academy of Audiology*, 23(6):476–494, 2012.

[79] Enrique A Lopez-Poveda, Peter T Johannesen, Patricia Pérez-González, José L Blanco, Sridhar Kalluri, and Brent Edwards. Predictors of hearing-aid outcomes. *Trends in Hearing*, 21:2331216517730526, 2017.

[80] Christian Lorenzi, Gaëtan Gilbert, Héloïse Carn, Stéphane Garnier, and Brian CJ Moore. Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy of Sciences*, 103(49):18866–18869, 2006.

[81] Ewan A MacPherson and John C Middlebrooks. Localization of brief sounds: effects of level and background noise. *The Journal of the Acoustical Society of America*, 108(4):1834–1849, 2000.

[82] David McAlpine and Benedikt Grothe. Sound localization and delay lines–do mammals fit the model? *Trends in Neurosciences*, 26(7):347–350, 2003.

[83] Ketan Mehta, Jörg Kliewer, and Antje Ihlefeld. Quantifying neuronal information flow in response to frequency and intensity changes in the auditory cortex. In *2018 52nd Asilomar Conference on Signals, Systems, and Computers*, pages 1367–1371. IEEE, 2018.

[84] Juozas Mickunas Jr. Interaural time delay and apparent direction of clicks. *The Journal of the Acoustical Society of America*, 35(5):788–788, 1963.

[85] Brian CJ Moore. *Cochlear hearing loss: physiological, psychological and technical issues*. John Wiley & Sons, 2007.

[86] Danielle X Morales, Amy E Wagler, and Angelica Monarrez. Build peer mentor training model: developing a structured peer-to-peer mentoring training for biomedical undergraduate researchers. *Understanding Interventions Journal*, 11(1), 2020.

[87] John B Mott, Lynn P McDonald, and Donal G Sinex. Neural correlates of psychophysical release from masking. *The Journal of the Acoustical Society of America*, 88(6):2682–2691, 1990.

[88] Antonio A Moya. An arduino experiment to study charge–voltage relationships in capacitors. *Physics Education*, 54(1):015005, 2018.

[89] Israel Nelken, Yaron Rotman, and Omer Bar Yosef. Responses of auditory-cortex neurons to structural features of natural sounds. *Nature*, 397(6715):154–157, 1999.

[90] Peggy B Nelson and Su-Hyun Jin. Factors affecting speech understanding in gated interference: Cochlear implant users and normal-hearing listeners. *The Journal of the Acoustical Society of America*, 115(5):2286–2294, 2004.

[91] Andreas Nieder and Georg M. Klump. Signal detection in amplitude-modulated maskers. ii. processing in the songbird's auditory forebrain. *European Journal of Neuroscience*, 2001.

[92] Edwin M Overholt, Edwin W Rubel, and Richard L Hyson. A circuit for coding interaural time differences in the chick brainstem. *Journal of Neuroscience*, 12(5):1698–1708, 1992.

[93] Andrew J Parker. Vision in our three-dimensional world. *Philosophical Transactions of the Royal Society*, 371:20150251, 2016.

[94] Andrew J Parker and Bruce G Cumming. Cortical mechanisms of binocular stereoscopic vision. In *Progress in Brain Research*, volume 134, pages 205–216. Elsevier, 2001.

[95] Michael Pecka, Antje Brand, Oliver Behrend, and Benedikt Grothe. Interaural time difference processing in the mammalian medial superior olive: the role of glycinergic inhibition. *Journal of Neuroscience*, 28(27):6914–6925, 2008.

[96] Michael Pecka, Ida Siveke, Benedikt Grothe, and Nicholas A Lesica. Enhancement of itd coding within the initial stages of the auditory pathway. *Journal of Neurophysiology*, 2009.

[97] Jose Luis Peña, Svenja Viete, Yehuda Albeck, and Masakazu Konishi. Tolerance to sound intensity of binaural coincidence detection in the nucleus laminaris of the owl. *Journal of Neuroscience*, 16(21):7046–7054, 1996.

[98] Geoff Petty. *Evidence-based teaching*. Nelson Thornes, 2009.

[99] Dennis P Phillips and Susan E Hall. Psychophysical evidence for adaptation of central auditory processors for interaural differences in time and level. *Hearing Research*, 202(1-2):188–199, 2005.

[100] Gian F Poggio. Mechanisms of stereopsis in monkey visual cortex. *Cerebral Cortex*, 5(3):193–204, 1995.

[101] Gian F Poggio, Francisco Gonzalez, and F Krause. Stereoscopic mechanisms in monkey visual cortex: binocular correlation and disparity selectivity. *Journal of Neuroscience*, 8(12):4531–4550, 1988.

[102] George D Pollak, R Michael Burger, Thomas J Park, Achim Klug, and Eric E Bauer. Roles of inhibition for transforming binaural properties in the brainstem auditory system. *Hearing Research*, 168(1-2):60–78, 2002.

[103] Kristin Kelly Porter and Jennifer M Groh. The "other" transformation required for visual–auditory integration: representational format. *Progress in Brain Research*, 155:313–323, 2006.

[104] Daniel Pressnitzer, Ray Meddis, Ray Meddis, Roel Delahaye, and Ian M. Winter. Physiological correlates of comodulation masking release in the mammalian ventral cochlear nucleus. *The Journal of Neuroscience*, 2001.

[105] R Quian Quiroga, Zoltan Nadasdy, and Yoram Ben-Shaul. Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. *Neural Computation*, 16(8):1661–1687, 2004.

[106] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2020. version 4.0.3.

[107] Ramnarayan Ramachandran, Kevin A Davis, and Bradford J May. Single-unit responses in the inferior colliculus of decerebrate cats i. classification based on frequency response maps. *Journal of Neurophysiology*, 82(1):152–163, 1999.

[108] Khaleel A Razak and Zoltan M Fuzessery. Gaba shapes a systematic map of binaural sensitivity in the auditory cortex. *Journal of Neurophysiology*, 104(1):517–528, 2010.

[109] Ann Marie Rohaly and Hugh R Wilson. The effects of contrast on perceived depth and depth discrimination. *Vision Research*, 39(1):9–18, 1999.

[110] Merri J Rosen, Emma C Sarro, Jack B Kelly, and Dan H Sanes. Diminished behavioral and neural sensitivity to sound modulation is associated with moderate developmental hearing loss. *PLoS One*, 7(7):e41514, 2012.

[111] Andrew T Sabin, Ewan A Macpherson, and John C Middlebrooks. Human sound localization at near-threshold levels. *Hearing Research*, 199(1-2):124–134, 2005.

[112] Nelli H Salminen, Hannu Tiitinen, Santeri Yrttiaho, and Patrick JC May. The neural code for interaural time difference in human auditory cortex. *The Journal of the Acoustical Society of America*, 127(2):EL60–EL65, 2010.

[113] David M Schneider and Sarah MN Woolley. Sparse and background-invariant coding of vocalizations in auditory scenes. *Neuron*, 2013.

[114] Clifton M Schor and Peter A Howarth. Suprathreshold stereo-depth matches as a function of contrast and spatial frequency. *Perception*, 15(3):249–258, 1986.

[115] Malcolm N Semple and Leonard M Kitzes. Focal selectivity for binaural sound pressure level in cat primary auditory cortex: two-way intensity network tuning. *Journal of Neurophysiology*, 69(2):462–473, 1993.

[116] Sudarshan Sivaramakrishnan and Alexander Ganago. Teaching strategy focused on sensory perception, students' interest and enjoyment: Successful application in electrical engineering (ee) lab for non-ee majors. In *2013 IEEE Frontiers in Education Conference (FIE)*, pages 296–302. IEEE, 2013.

[117] Joseph Sollini and Paul Chadderton. Comodulation enhances signal detection via priming of auditory cortical circuits. *The Journal of Neuroscience*, 2016.

[118] Pamela Souza, Kathryn Arehart, Tim Schoof, Melinda Anderson, Dorina Strori, and Lauren Balmert. Understanding variability in individual response to hearing aid signal processing in wearable hearing aids. *Ear and Hearing*, 40(6):1280–1292, 2019.

[119] MW Spitzer and MN Semple. Neurons sensitive to interaural phase disparity in gerbil superior olive: diverse monaural and temporal response properties. *Journal of Neurophysiology*, 73(4):1668–1690, 1995.

[120] Annette Stange, Michael H Myoga, Andrea Lingner, Marc C Ford, Olga Alexandrova, Felix Felmy, Michael Pecka, Ida Siveke, and Benedikt Grothe. Adaptation in sound localization: from gaba b receptor–mediated synaptic modulation to perception. *Nature Neuroscience*, 16(12):1840, 2013.

[121] G Christopher Stecker, Ian A Harrington, and John C Middlebrooks. Location coding by opponent neural populations in the auditory cortex. *PLoS Biology*, 3(3):e78, 2005.

[122] Richard M Stern and Constantine Trahiotis. Models of binaural interaction. *Handbook of Perception and Cognition*, 6:347–386, 1995.

[123] Hirokazu Takahashi, Masayuki Nakao, and Kimitaka Kaga. Cortical mapping of auditory-evoked offset responses in rats. *Neuroreport*, 15(10):1565–1569, 2004.

[124] Donald C Teas. Lateralization of acoustic transients. *The Journal of the Acoustical Society of America*, 34(9B):1460–1465, 1962.

[125] Hardy Thomas, Jochen Tillein, Peter Heil, and Henning Scheich. Functional organization of auditory cortex in the mongolian gerbil (meriones unguiculatus). i. electrophysiological mapping of frequency representation and distinction of fields. *European Journal of Neuroscience*, 5(7):882–897, 1993.

[126] Willem A van Bergeijk. Variation on a theme of bekesy: a model of binaural interaction. *The Journal of the Acoustical Society of America*, 34(9B):1431–1437, 1962.

[127] Jesko L Verhey, Daniel Pressnitzer, and Ian M Winter. The psychophysics and physiology of comodulation masking release. *Experimental Brain Research*, 153(4):405–417, 2003.

[128] Joyce Vliegen and A John Van Opstal. The influence of duration and level on human sound localization. *The Journal of the Acoustical Society of America*, 115(4):1705–1713, 2004.

[129] Gerog Von Békésy and Ernest G Wever. *Experiments in hearing*. New York: McGraw-Hill, 1960.

[130] Uri Werner-Reiss and Jennifer M Groh. A rate code for sound azimuth in monkey auditory cortex: implications for human neuroimaging studies. *Journal of Neuroscience*, 28(14):3747–3758, 2008.

[131] Frederic L Wightman and Doris J Kistler. The dominant role of low-frequency interaural time differences in sound localization. *The Journal of the Acoustical Society of America*, 91(3):1648–1661, 1992.

[132] Timothy M Woods, Steve E Lopez, James H Long, Joanne E Rahman, and Gregg H Recanzone. Effects of stimulus azimuth and intensity on the single-neuron activity in the auditory cortex of the alert macaque monkey. *Journal of Neurophysiology*, 96(6):3323–3337, 2006.

[133] Beverly A Wright and Matthew B Fitzgerald. Different patterns of human discrimination learning for two interaural cues to sound-source location. *Proceedings of the National Academy of Sciences*, 98(21):12307–12312, 2001.

[134] Justin D Yao and Dan H Sanes. Developmental deprivation-induced perceptual and cortical processing deficits in awake-behaving animals. *eLife*, 7:e33891, 2018.

[135] Yi Ye, Antje Ihlefeld, and Merri J Rosen. Conductive hearing loss during development does not appreciably alter the sharpness of cochlear tuning. *Scientific Reports*, 11(1):1–11, 2021.

[136] Tom C Yin and Joseph C Chan. Interaural time sensitivity in medial superior olive of cat. *Journal of Neurophysiology*, 64(2):465–488, 1990.

[137] Jerrold H Zar. *Biostatistical analysis*. Pearson Education India, 1999.

[138] Jiping Zhang, Kyle T Nakamoto, and Leonard M Kitzes. Binaural interaction revisited in the cat primary auditory cortex. *Journal of Neurophysiology*, 91(1):101–117, 2004.

[139] Marcel P Zwiers, Huib Versnel, and A John Van Opstal. Involvement of monkey inferior colliculus in spatial hearing. *Journal of Neuroscience*, 24(17):4145–4156, 2004.