The Impact of Sensitivity to Fast Spectrotemporal Chirps on Speech Coding in the Mammalian Inferior Colliculus: Physiology and Modeling

by

Paul W. Mitchell

Submitted in Partial Fulfillment of the

Requirements for the Degree

Doctor of Philosophy

Supervised by Professor Laurel H. Carney

Department of Biomedical Engineering

Arts, Sciences and Engineering

Edmund A. Hajim School of Engineering and Applied Sciences

University of Rochester

Rochester, New York

2024

Table of Contents

Biographical Sketch

Paul Mitchell graduated from the University of Texas at Dallas in 2017 with a Bachelor of Science degree in Biomedical Engineering. He began his doctoral studies in the Department of Biomedical Engineering at the University of Rochester in 2017, and received a Master of Science degree in Biomedical Engineering in 2019. Paul was awarded an NIH F31 Predoctoral Fellowship in 2021. Since 2018, Paul has pursued a thesis focused on understanding the nature and implications of fast frequency chirp sensitivity in the auditory midbrain, under the supervision of Dr. Laurel H. Carney.

The following publications were a result of work conducted during doctoral study:

Journal Publications:

Mitchell, P. W., Henry, K. S., & Carney, L. H. (2023). Sensitivity to direction and velocity of fast frequency chirps in the inferior colliculus of awake rabbit. Hearing research, 440, 108915.

Mitchell, P. W. & Carney, L. H. (2024, in press). A Computational Model of Auditory Chirp-Velocity Sensitivity and Amplitude-Modulation Tuning in Inferior Colliculus Neurons. The Journal of Computational Neuroscience.

Acknowledgements

First and foremost, I would like to thank my advisor, Dr. Laurel Carney. I started working with Laurel during a period of personal uncertainty towards my future. Without her example and guidance, I would certainly not be completing my PhD now.

I would also like to acknowledge all current and past committee members, including Dr. Ross Maddox, Dr. Zhiyao Duan, Dr. Mark Bocko, and especially Dr. Ken Henry, who provided advice throughout my PhD.

Thanks to all current and past Carney lab graduate students and post-docs, Dr. Langchen Fan, Dr. Braden Maxwell, Dr. Afagh Farhadi, Johanna Fritzinger, Dr. Daniel Guest, and Dr. Swapna Agarwalla, as well as members of the Henry lab, Yingxuan Wang and Leslie Gonzales, for the years of camaraderie and support. Also, special acknowledgement to Doug Schwarz and Kris Abrams for their assistance and friendship throughout my PhD.

I would like to thank my wonderful girlfriend Ellen Kneeskern for being an irrepressible ray of optimism and joy even in the most difficult stretches of graduate school. I am so happy we can be a context for one other.

Finally, I would like to thank my parents, Jerry and Wei Mitchell, and my brothers, Stephen and David. Your lifelong belief in me has made all of this possible.

Abstract

Recordings of auditory midbrain neurons in the inferior colliculus (IC) have revealed large rate differences in response to fast frequency chirps of different directions or velocities. This observation has implications for IC responses to perceptually-important sounds such as speech: neural response rates may be substantially impacted by chirp cues in a way not predicted by more commonly studied response properties, such as characteristic frequency (CF) or modulation transfer function (MTF) shape. In vowels, the phase transitions associated with resonances of the vocal tract suggest that chirp cues may arise near formant frequencies. Thus, it is important to account for the effect of chirp cues in IC speech responses.

To characterize chirp-sensitive neurons, recordings were made of IC units. A novel stimulus was developed to characterize IC neuron velocity sensitivity in an aperiodic context using rate-velocity functions (RVFs). RVFs were shown to be independent from CF or MTF shape. The majority of IC neurons displayed direction-selectivity to chirps for at least one velocity—this selectivity was most often observed at lower chirp velocities compared to higher velocities.

IC chirp sensitivity was incorporated into a computational model that retained tuning to frequency and AM. The mechanism of chirp sensitivity in the IC was hypothesized to originate in octopus cells of the posteroventral cochlear nucleus, which display a similar diversity of chirp sensitivity to the IC. Octopus cell chirp sensitivity was modeled using sequence detection, whereby correct arrival order of auditory-nerve inputs is required to elicit a response. Octopus-cell output was an inhibitory input the IC,

alongside an inhibition to produce band-enhanced MTFs. The model was capable of simulating neurons with physiologically valid CFs and RVFs.

The impact of chirp sensitivity on vowel coding was assessed with a combination of physiology and modeling strategies. IC responses to vowel stimuli were classified based on average rate and spike timing. Overall classification accuracy was found to be correlated with directionless velocity-sensitivity and high-velocity direction bias. Additionally, individual neurons' vowel responses were examined in the chirp model, but the relationship of model neuron RVFs and vowel classification remained unclear.

Contributors and Funding Sources

This work was supported by a dissertation committee consisting of Professor Laurel H. Carney (advisor) of the Departments of Biomedical Engineering and Neuroscience, Professor Kenneth S. Henry of the Departments of Otolaryngology, Neuroscience, and Biomedical Engineering, and Professors Zhiyao Duan and Mark Bocko of the Department of Electrical and Computer Engineering. Professor Ross K. Maddox of the University of Michigan also contributed to all committee meetings except the defense. All work conducted for the dissertation was completed by the student independently, under the supervision of Prof. Laurel H. Carney. Work was supported by the National Institute of Health and National Institute on Deafness and Other Communication Disorders under the grant numbers NIH R01-DC001641, NIH R01-DC010813, and NIH F31-DC019816.

List of Tables

List of Figures

Chapter 1: Introduction

The auditory system is a network of neural pathways that accomplishes the task of converting sound into useful information about the world. Broadly, the auditory system can be divided into two parts. The peripheral auditory system transduces sound into a signal understandable by the brain, ultimately a series of electric impulses in time. The central auditory system, the part included within the brain, processes and integrates signals received from the periphery, eventually resulting in perception (Yost, 2013). For humans, an essential purpose of the auditory system is deciphering complex sound stimuli like speech into meaning.

Responses of the auditory system are indescribable by linear system—that is, a system in which outputs are linearly related to inputs. In particular, nonlinear aspects of the periphery strongly impact signals in the central auditory system. One example is the active mechanism of the cochlea by outer hair cells (OHCs), which lends the mechanical response of the basilar membrane higher frequency specificity than could be achieved with an inactive mechanism alone (Kim, 1986), and serves to selectively amplify signal gain across frequencies (Ashmore, 2008; Hudspeth, 2014). It is important to note that these nonlinearities are not undesirable or negligible, but essential in the sense that they shape and define an emergent code of sound features that is used by subsequent areas in the auditory pathway. For instance, loss of the active processes of the cochlea (through hearing loss) is implicated with a substantial decrease in the ability to perceive speech, especially in background noise (Oxenham et al., 2003), presumably due to a degradation in signal contrasts achieved by the normal compressive properties of the

cochlea (Carney, 2018; Carney, 2024). Therefore, the nonlinear properties of the auditory system crucially impact how the brain processes complex sounds like speech.

However, peripheral nonlinearities often clash with historically leading theories of complex sound coding. In the auditory nerve (AN), the two most prominent theories are coding based on rate and temporal fine structure (TFS) (Sachs and Young, 1979; Young and Sachs, 1979), both of which theorize a representation of spectral peaks contained in speech, either based on the firing rate of tonotopically organized AN fibers (Cedolin and Delgutte, 2005) or the precise timing of responses of those fibers (Sachs et al., 1988). Both of these theories are challenged by properties of peripheral nonlinearities. Firing rates of AN fibers saturate with increasing stimulus level (Costalupes et al., 1984). Specifics of this rate-saturation differ between fibers depending on their spontaneous rates (Liberman, 1978). Additionally, AN fibers can be organized into high-, medium-, and low-spontaneous-rate groups (HSR, MSR, and LSR, respectively). HSR fibers, comprising the majority, saturate at low sound pressure levels (SPLs), about 20-30 dB SPL (Bharadwaj et al., 2014; Huet et al., 2016), as compared to conversational speech levels, about 55 – 66 dB (Olsen, 1998). Although MSR and LSR fibers have extended dynamic ranges compared to HSR fibers (Yates, 1990), the majority of AN fibers (HSRs) saturating at levels at which conversation is normally conducted is incongruous with rate-based coding. Furthermore, TFS code degrades in the AN both above 1 kHz frequencies (Johnson, 1980), and at conversational speech levels (Delgutte and Kiang, 1984). Therefore, a comprehensive code for complex sounds like speech at the periphery based on rate or TFS is unlikely, motivating the need to examine the central auditory system for coding mechanisms that might take advantage of peripheral nonlinearities instead of being hindered by them.

The inferior colliculus (IC) is a nearly obligatory point of convergence in the ascending auditory pathway, receiving diverse excitatory and inhibitory input from all lower regions. The IC is sensitive to sound features—notably including frequency, binaural cues, and amplitude modulations (AM) (Aitkin and Reynolds, 1975; Stiebler and Ehret, 1985; Langner and Schreiner, 1988)—but importantly, these sensitivities are diverse between individual IC neurons, and the same neuron can exhibit sensitivity to multiple sound features. IC neurons are tuned toward the specific modulation frequencies of AM sounds, commonly displaying enhanced or suppressed response rates relative to unmodulated sounds (Kim et al., 2020). These modulation frequencies also coincide with the range of speech pitch (Bishop and Keating, 2012). AM tuning in the IC has been theorized as the crux of an alternative method of coding speech based on neural fluctuations of AN fibers (Carney, 2018; Carney, 2024). In this theory, AN fibers tuned to spectral peaks saturate due to a combination of cochlear compression and transduction saturation of the inner-hair-cell (IHC), another peripheral nonlinearity. Meanwhile, AN fibers tuned away from spectral peaks display comparatively higher amplitude fluctuations in their neural response patterns. In the IC, AM-sensitive neurons convert the presence or absence of neural fluctuations into a rate code, and form a spectral profile that encodes the location of peaks across frequency (Carney, 2018). Such a representation of speech would additionally be robust at conversational speech levels and even in background noise. (Carney, 2018). Neural fluctuation coding of speech serves as an example whereby feature sensitivities of the IC can be leveraged to construct a representation of complex sounds resistant to the pitfalls of coding schemes limited to the periphery.

In this thesis, we describe efforts to update and expand our understanding of complex sound representation in the IC to include a relatively new and under-explored sensitivity to the direction and velocity of fast frequency chirps. This work began by summarizing response patterns of IC neurons to chirps, continued with constructing a model capable of chirp sensitivity, and ended with an analysis of the impact of chirp-sensitivity on speech coding.

1.1 Characterizing IC sensitivity to frequency chirps

Much of the inciting evidence for this work stems from a series of preliminary experiments conducted in 2016-2017 consisting of extracellular, single-unit recordings of the responses of IC neurons to Schroeder-harmonic complexes (this work would also eventually culminate in Steenken et al., 2022). In this work, Schroeder-harmonic complexes (henceforth SCHR stimuli) were used as a way to examine IC responses to fast frequency sweeps.

SCHR stimuli were invented by Schroeder (1970) to introduce a formula to calculate the necessary phase of a harmonic complex resulting in a maximally flat stimulus envelope. The stimulus has a flat magnitude spectrum and phase spectrum defined by the equation:

$$\theta_n = C\pi n(n-1)/N$$

which describes θ_n, the phase of the nth harmonic, as a function of harmonic number n and the total number of harmonics N (Schroeder, 1970). Additionally, C represents a duty-cycle factor that also determines the sign of the phase function. These spectra are depicted in Figure 1.1.

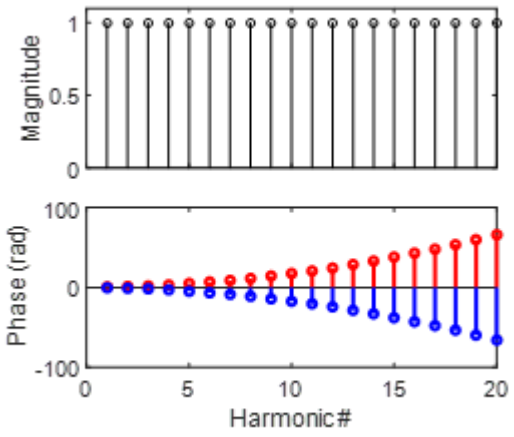

Figure 1.1 – Magnitude and phase spectra of the SCHR stimulus (positive-SCHR in red, negative-SCHR in blue).

SCHR stimuli are defined by a linear frequency sweep extending from the fundamental frequency (F0) to the highest harmonic during each stimulus period. The velocity of this sweep can be controlled by modulating F0—for instance, doubling F0 from 50 Hz to 100 Hz doubles the velocity, as the sweep travels a comparable frequency range over half the time. Depending on the sign of C, the frequency sweep contained in the SCHR stimulus either increases over time (for C = -1) or decreases (for C = 1). The resulting waveform and spectrogram depicting the within-period frequency sweep is depicted in Figure 1.2. Note that SCHR stimuli are typically assigned a name associated with direction of their phase functions, not the direction of their frequency sweep. Thus, +SCHR stimulus contains a negative frequency sweep, and vice versa. To eliminate ambiguity, the remainder of this document will refer to +SCHR stimuli as “downward SCHR” and -SCHR as “upward SCHR”—additionally, all figures depicting SCHR stimuli or related stimuli and responses will use the color red to refer to downward SCHR, and blue to refer to upward SCHR (including Fig. 1.1 and 1.2).

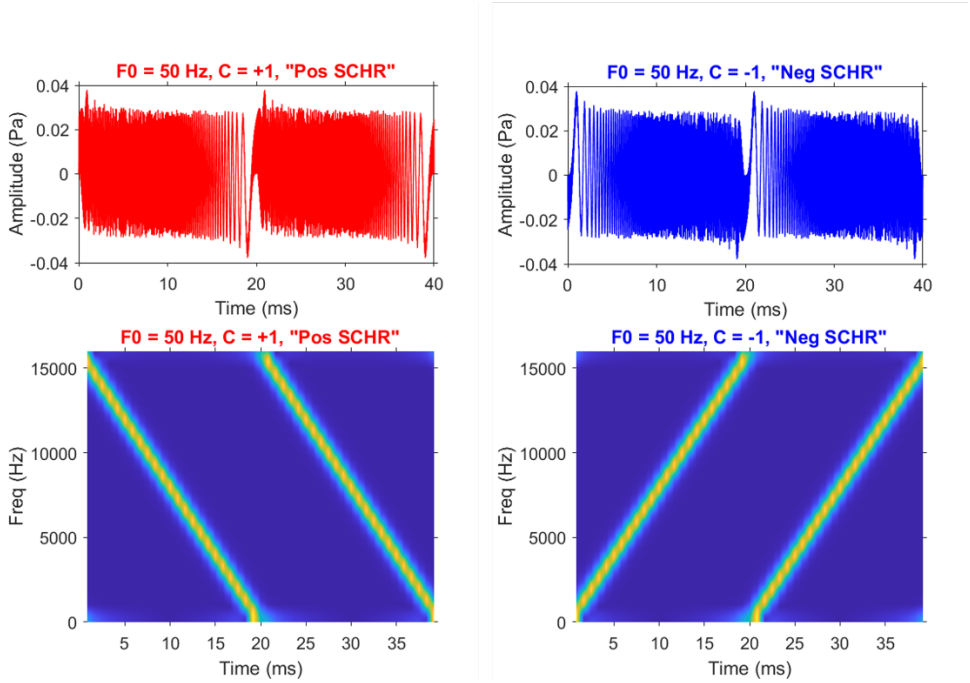

Figure 1.2 – SCHR stimulus waveforms and spectrograms. Top: SCHR waveform with F0 = 50 Hz, C = +1 (red, left) and C = -1 (blue, right). Bottom: Spectrogram illustrating negative frequency sweep in +SCHR (left), and positive frequency sweep in -SCHR (right).

In the preliminary experiments, the response rates of IC neurons varied substantially for stimuli that differed only in the direction or velocity of frequency sweep (Figure 1.3). Furthermore, in the cases that the stimuli differed only in direction of sweep (i.e., phase direction), these stimuli were time-reversed versions of one another, suggesting the existence a nonlinearity in responses of the IC to SCHR sweeps. Finally, neurons expressed a diversity of responses to SCHR stimuli, rarely responding uniformly by sweep direction, instead responding differently with velocity.

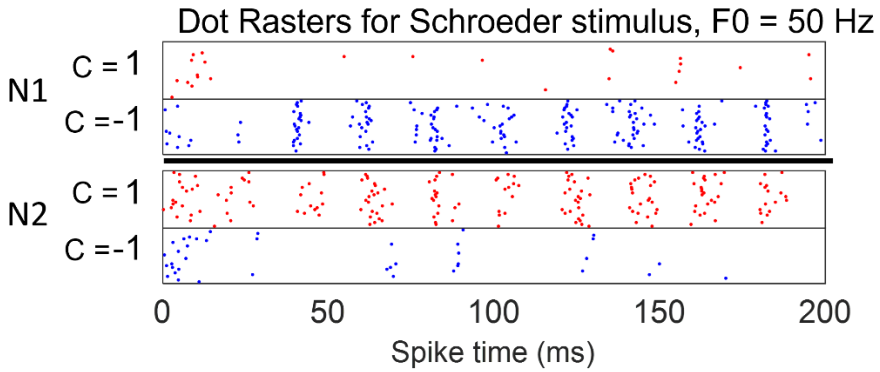

Figure 1.3 – Dot rasters depicting responses of two IC neurons to SCHR stimuli of 50 Hz and C = ±1. Within each row, stimulus repetition number is indicated by vertical location of dot, and time by horizontal location. Top: neuron 1 (N1) displaying direction bias towards -SCHR (upward sweep). Bottom: neuron 2 (N2) displaying direction bias towards +SCHR (downward sweep).

It is important to note that the frequency sweep represented by the SCHR stimulus is distinct from many conventionally considered frequency modulation (FM) stimuli, in the context of the auditory system. Primarily, the speed of the SCHR sweep is much faster than spectrotemporal changes in other signals, such as formant transitions in vowels (Liberman and Mattingly, 1989), which occur over the course of several pitch-periods and represent frequency-shifts in the hundreds of Hz. For comparison, SCHR sweeps travel from F0 to the highest harmonic (in this work, 16 kHz) over the fundamental period (here, 2.5 to 20 ms), placing the spectrotemporal velocities in the realm of bat echolocation sweep velocities (Fuzessery, 1994; Gordon and O'Neill, 1998). To distinguish SCHR sweeps from slower FM, we refer to them and similar stimuli as chirps for the remainder of this document.

Additionally, the responses in Fig. 1.3 were recorded in rabbits, a species that does not produce specialized vocalizations containing chirps (Awan et al., 2014), unlike bats. On the contrary, rabbits are considered hearing generalists, with high-frequency

audiometric thresholds comparable to similarly-sized mammals, and low-frequency thresholds similar to humans (Heffner and Masterson, 1980). Altogether, this data suggests that IC neurons are sensitive to the direction and velocity of frequency chirps in SCHR stimuli.

For the implications of this preliminary data to be interrogated further, it was necessary to fully define the nature of supposed chirp sensitivity and its relationship to other IC feature sensitivities. Notably, SCHR stimuli, as harmonic complexes, are highly periodic and thus represent a strong AM cue to an individual, IC neuron. In the first chapter of this work, we tested the hypothesis that the observed sensitivity of IC neurons to SCHR chirps represented a general chirp velocity sensitivity that would manifest in aperiodic contexts. We aimed to use physiological methods to characterize the trends of chirp sensitivity with other known feature sensitivities of the IC, in particular frequency and AM tuning.

One crucial outcome of this work was the development of the aperiodic chirp stimulus—this stimulus was inspired by techniques such as modulation transfer functions (MTFs), which describe a neuron's response rate to modulated stimuli as a function of modulation frequency. Similarly, the objective of the aperiodic chirp stimulus was to describe neural response rates to chirps as a function of chirp velocity, completely removed from the periodic context of SCHR stimuli, in a summary figure called a rate-velocity function (RVF).

Aperiodic chirp stimuli are designed to be SCHR-like—they are comprised of single-cycles of SCHR stimuli (i.e., individual chirps) presented in random order and with random silent-interval spacing to ensure aperiodicity. Similarity to the SCHR stimulus is

important because it allows RVFs and MTFs to be used as orthogonal feature vectors, representing chirp velocity and periodicity tuning, respectively, upon which SCHR responses can be predicted. One finding of this analysis was that, while the majority of SCHR response variance was predicted by neurons' MTFs, the RVF contributed a high percentage of variance explained for a subset of the population comprising the neurons with direction bias in their SCHR responses.

Another outcome of this work was the use of receiver operating characteristic (ROC; Egan, 1975) analysis to quantify the direction selectivity of a neuron to particular pairs of equal-speed (absolute velocity) chirps. Neurons were labeled as significantly biased towards a direction if the ROC area under the curve (AUC) surpassed the threshold (70.7%) equivalent to that of a two-down, one-up psychophysical procedure (Levitt, 1971)—the direction that AUC passed this threshold (i.e., above 70.7% or below 29.3%) determined the direction of bias. Using this definition of direction bias, we reached the conclusion that the vast majority of IC neurons (>90%) were direction biased for at least one equal-speed chirp pair. Additionally, the prevalence of direction biased neurons did not appear to vary with characteristic frequency (CF) or MTF shape. Rather, the defining feature of chirp sensitivity among IC neurons tended to be diversity.

Finally, principal component analysis (PCA) on the entire population of RVFs revealed three dimensions along which the RVFs most differed—average rate (PC1), direction-independent velocity sensitivity (PC2) and low-velocity direction bias (PC3). In particular, PC2 and PC3 pointed to a natural divide in the way neurons responded to chirp-velocities at 2 kHz/ms. PC2 indicated that neurons tended to have high response-rates to low chirp velocities (< 2 kHz) and low response-rates to high chirp velocities (> 2

kHz), or vice versa. PC3 reflected the chirp-direction bias to low-velocity chirps (< 2 kHz).

Overall, the completion of this study provided evidence that chirp direction and velocity sensitivity was its own feature sensitivity independent of periodicity cues, and additionally showed the physiological ubiquity of chirp sensitivity. Furthermore, it introduced several analysis tools such as aperiodic chirp stimuli, RVFs, and ROC direction bias, that would be employed in future work (Mitchell et al., 2023).

1.2 Incorporating chirp sensitivity into a computational model of the IC

With chirp direction and velocity sensitivity established as a common characteristic among IC neurons, the need for a computational model to simulate responses to chirps became clear. Such a model has two utilities: first, design of a plausible computational model allows the testing of candidate mechanisms for chirp sensitivity, for which the contributing physiological sources are still unclear. Second, such a model would allow the nature of chirp-sensitivity in the model neuron to be experimentally altered directly, permitting further study on the impact of chirp sensitivity on responses to complex sounds, like speech.

Bat literature has identified several potential mechanisms whereby chirp direction sensitivity might arise. The most basic model depends on asymmetric sideband inhibition alongside excitation (Rees and Langner, 2005)—the differently-timed excitatory and inhibitory inputs combine to form a spectrotemporal filter. For example, chirps that traverse the inhibitory region prior to excitation do not result in a response due to the coincidence of excitation and slower-acting inhibition (Gordon and O'Neill,

1998; Portfors, 2018). Other mechanisms have been hypothesized to impact chirp sensitivity, including early on-CF inhibition, which may act to reject chirps traversing the excitatory region too slowly (Fuzessery et al., 2006; Fuzessery et al., 2011), and wide-band facilitation, which requires multiple cross-frequency inputs with differing delays to sum to overcome blanket inhibition (Fuzessery and Hall, 1996). Another theory implicates inhibition originating from the superior paraolivary complex (SPON) as a source of non-directional velocity tuning (Pollak et al., 2011). Finally, another hypothesis proposes that cells with high input resistances and long time constants may become chirp direction sensitive through asymmetry in input magnitudes rather than timing (Gittelman et al., 2009).

The physiological mechanism that informed the model presented here arises in the octopus cells of the posteroventral cochlear nucleus (PVCN) (Oertel et al., 2000). In particular, the model was heavily influenced by the work of Lu et al. (2022), which presented a theory of sequence detection to explain the strong sensitivity to chirp direction and velocity observed in octopus cells. Octopus cells have wide dendritic fields that receive AN inputs from many different frequencies (Osen, 1969). Octopus cells are also characterized by low-voltage-activated potassium (KL) channels, known for slow recovery times (Bal and Oertel, 2001). Acknowledging that inputs of different frequencies may also have different amplitudes, a potential mechanism to select for chirp direction emerges: a suprathreshold excitatory post-synaptic potential (EPSP) in isolation will trigger a response before hyperpolarization associated with KL channels prevents subsequent response. If a subthreshold EPSP arrived before the suprathreshold EPSP, the hyperpolarization initiated by the subthreshold input would coincide with the suprathreshold EPSP, preventing an action potential. If we imagine these two inputs as

having different frequencies, this mechanism would select for a particular sequence of cross-frequency inputs, and could explain chirp direction selectivity (Lu et al., 2022).

Furthermore, octopus cells project to the contralateral ventral nucleus of the lateral lemniscus (VNLL) (Adams, 1997; Vater et al., 1997), which is known to provide inhibition to the IC. The combination of mechanism and well-studied projection to the IC makes octopus cell sequence detection a promising focus for the computational model.

In designing the model, we employed a strategy pioneered by Krips and Furst (2009), which describes model cells as coincidence detectors (CD) that combine any number of excitatory or inhibitory inputs with their own parameters determining timing and strength. Krips and Furst's work comes from a family of models originated by Siebert (1965), which broadly use statistical decision theory to estimate psychophysical thresholds as a function of stimulus parameters—this method has previously been used to examine discrimination of tone frequency, sound level, and interaural time and level differences (Siebert, 1970; Colburn, 1973; Heinz et al., 2001a,b). Krips and Furst innovated upon this tradition by showing that if the statistics of CD outputs conformed to nonhomogeneous Poisson processes (NHPPs), those outputs could be used to drive inputs to a subsequent model cell, effectively extending the model from the periphery into the central nervous system (Krips and Furst, 2009). Finally, and of chief importance to the work presented here, Krips and Furst's method represents a highly flexible strategy in which numerous and varied model architectures can be theorized.

In the third chapter of this work, we tested the hypothesis that RVFs recorded from physiological units could be simulated by an IC model receiving input from an

octopus cell with sequence detection. Crucially, we aimed for the model to explain chirp sensitivity alongside other IC feature sensitivities, chiefly AM tuning.

Within the Krips and Furst framework, sequence detection was implemented in the model's first stage, wherein an octopus cell was simulated with two excitatory inputs, representing AN inputs of different magnitudes, and two inhibitory inputs, representing the long-term hyperpolarization related to KL channels. The model's first stage consisted of an IC cell receiving on-CF excitation alongside delayed on-CF inhibition, and leading octopus-cell inhibition. The two inhibitory processes controlled AM tuning and chirp sensitivity, respectively. Delayed same-frequency inhibition and excitation (SFIE) (Nelson and Carney, 2004) is a strategy for creating band-enhanced (BE) MTFs, characterized by elevated rates to a band of modulation frequencies relative to unmodulated.

In the main outcome of this work, we show that parameters can be proposed for model neurons that mimic the range of CFs seen in physiological recordings of chirp sensitive neurons. Across this CF range, RVFs of either upward or downward direction bias can be proposed, all with BE MTFs, and all within the same model cell. In summary, using sequence-selective octopus-cell inhibition as the basis of chirp sensitivity in the IC, we proposed a model that can accurately portray multiple IC sound-feature sensitivities in the same model neuron (Mitchell and Carney, 2024). These qualities make the presented model a useful tool for examining responses of the IC to complex sounds.

1.3 Assessing the impact of chirp sensitivity on speech coding

Vowels make up a significant portion of available phonemes in language, and make strong contributions to overall speech intelligibility and linguistic information

(Ladefoged and Broadbent, 1957; Kewley-Port et al., 2007). Vowel identity is defined by spectral peaks known as formants (Fant, 1971). One proposed code of formant frequency in the IC is based on contrasts in the profile of neural fluctuations of AN responses to spectral peaks and valleys (Carney, 2018; Carney, 2024)—neural fluctuation code is resistant to rate saturation of the AN and background noise.

Through the sensitivity to chirp direction and velocity described in earlier chapters, it is possible that chirp-cues in vowels serve to strengthen formant representation in the IC by providing a secondary code of formant frequencies. Consider SCHR stimuli—the chirps contained within SCHR harmonic complexes are the result of phase-differences between harmonic components. In vowels, vocal tract resonances shape phase spectra, centered around formant frequencies, implying that formant frequencies coincide with phase inflections. Congruently, group-delay functions tracking time delay across frequencies indicate that group-delay peaks around formant frequencies, resulting in spectrotemporal fluctuations at formant frequencies (Bozkurt et al., 2006; Rajan et al., 2013).

In the final part of these studies, we aimed to assess the impact on vowel coding of chirp sensitivity. We tested the hypothesis that vowel phase properties would amount to a cue usable by chirp-sensitive IC neurons, and that chirp sensitivity would impact coding of vowels in both physiological and modeling methods.

One outcome of this study was the development of a chirp-identification strategy based on component decomposition. Given an input vowel, the power spectrum was used to approximate component frequencies—then parameter optimization was utilized to fit a sinusoid to these estimated components, matching phase and magnitude. Finally,

a synthetic version of the original vowel could be generated, with clearly defined phase between harmonic components. These synthetic vowels revealed the presence of time delays centered around formant frequencies, as well as inflections spanning between these frequencies of a similar velocity to chirps used in prior studies (i.e., aperiodic chirp stimulus).

Steady-state vowel tokens from the Hillenbrand et al. (1995) database were presented to physiological IC units with characterized CFs, MTFs, and RVFs. Based on spike train data in response to 12 vowels, /iy, ih, ei, eh, ae, ah, aw, oo, uw, er, oa, uh/, each neuron was assessed on ability to distinguish vowels based on average rate and spike timing. The spike timing metric used was the rate-independent spike (RIS) distance, a measure of the temporal similarity between two spike trains independent of local rate (Satuvuori et al., 2017). Vowel classification accuracies were assessed on a vowel-by-vowel basis as well as by overall accuracy; classification based on timing was found to be universally more accurate than rate. Comparing individual neurons with different RVFs, no distinguishable patterns emerged in classification accuracies.

Using PCA, PC2 (direction-independent velocity sensitivity) was found to correlate positively to overall accuracy of classification using timing. Furthermore, direction bias at high-velocities (> 2 kHz) was also found to be positively correlated to accuracy of classification using rate or timing. However, in both cases the amount of variance explained was relatively low—although, in light of a similar result in chapter 2 that showed MTFs explaining more variance in SCHR responses than RVFs, this is a consistent finding.

Finally, model neurons with similar characteristics to select units from physiology were presented with Hillenbrand vowels stimuli. The classification accuracy of model vowels was compared to their correlates from physiology. Altogether, model neurons failed to predict classification trends in corresponding physiology neurons, suggesting that the IC chirp model is missing an unknown element. Alternatively, it could suggest that the IC chirp model, which was foremost conceptualized as way to simulate chirp direction bias, may need additional work to add direction-independent velocity sensitivity, the RVF feature towards which the highest correlation to classification accuracy was observed.

1.4 Overview of the thesis

Chapter 1 introduces the importance of sound features and peripheral nonlinearities in developing the basis of complex sound coding in the auditory system. Also, the importance of the IC as a point of integration is highlighted, since it serves as the first stage in which such a code of complex sounds might be realized.

Chapter 2 has been published in Hearing Research, and describes methods and results surrounding the physiological recordings and experiments undertaken to characterize and define chirp direction and velocity as a separate sound-feature towards which the IC is sensitive. This chapter introduces crucial tools such as the aperiodic chirp stimulus, RVF, and ROC-direction bias.

Chapter 3 has been published in the Journal of Computational Neuroscience. It describes the architecture of the chirp-sensitive IC model, the implementation of sequence detection in the octopus-cell stage, and the combination of chirp-sensitivity with AM tuning at the IC stage. Also, parameters for example model neurons spanning

the range of chirp-sensitive neurons observed in physiology are proposed, with the impact of changing individual parameters additionally explored.

Chapter 4 is unpublished work that compares responses of neurons recorded in physiology to corresponding neurons simulated by the chirp-sensitive model. Neurons are assessed in their ability to distinguish vowels based on rate and timing classification. A trend is established between chirp velocity and direction sensitivity and classification accuracy. Also, a method of quantifying and visualizing chirps contained within vowels is proposed.

Finally, in Chapter 5, we present a discussion of the overall work, as well as future directions to take research on chirp-sensitivity of the IC.

Bibliography

Adams, J. C. (1997). Projections from octopus cells of the posteroventral cochlear nucleus to the ventral nucleus of the lateral lemniscus in cat and human. Aud Neurosci, 3(4), 335-350.

Aitkin, L. M., & Reynolds, A. (1975). Development of binaural responses in the kitten inferior colliculus. Neuroscience Letters, 1(6), 315-319. https://doi.org/https://doi.org/10.1016/0304-3940(75)90019-1

Ashmore, J. (2008). Cochlear Outer Hair Cell Motility. Physiological reviews, 88(1), 173-210. https://doi.org/10.1152/physrev.00044.2006

Awan, S. N., Novaleski, C. K., & Rousseau, B. (2014). Nonlinear Analyses of Elicited Modal, Raised, and Pressed Rabbit Phonation. Journal of Voice, 28(5), 538-547. https://doi.org/https://doi.org/10.1016/j.jvoice.2014.01.015

Bal, R., & Oertel, D. (2001). Potassium Currents in Octopus Cells of the Mammalian Cochlear Nucleus. Journal of Neurophysiology, 86(5), 2299-2311. https://doi.org/10.1152/jn.2001.86.5.2299

Bharadwaj, H. M., Verhulst, S., Shaheen, L., Liberman, M. C., & Shinn-Cunningham, B. G. (2014). Cochlear neuropathy and the coding of supra-threshold sound [Hypothesis and Theory]. Frontiers in Systems Neuroscience, 8. https://doi.org/10.3389/fnsys.2014.00026

Bishop, J., & Keating, P. (2012). Perception of pitch location within a speaker's range: Fundamental frequency, voice quality and speaker sex. The Journal of the Acoustical Society of America, 132(2), 1100-1112. https://doi.org/10.1121/1.4714351

Carney, L. H. (2018). Supra-Threshold Hearing and Fluctuation Profiles: Implications for Sensorineural and Hidden Hearing Loss. Journal of the Association for Research in Otolaryngology, 19(4), 331-352. https://doi.org/10.1007/s10162-018-0669-5

Carney, L. H. (2024). Neural Fluctuation Contrast as a Code for Complex Sounds: The Role and Control of Peripheral Nonlinearities. Hearing research, 443, 108966. https://doi.org/https://doi.org/10.1016/j.heares.2024.108966

Cedolin, L., & Delgutte, B. (2005). Pitch of Complex Tones: Rate-Place and Interspike Interval Representations in the Auditory Nerve. Journal of Neurophysiology, 94(1), 347-362. https://doi.org/10.1152/jn.01114.2004

Colburn, H. S. (1973). Theory of binaural interaction based on auditory-nerve data. I. General strategy and preliminary results on interaural discrimination. The Journal of the Acoustical Society of America, 54(6), 1458-1470. https://doi.org/10.1121/1.1914445

Delgutte, B., & Kiang, N. Y. S. (1984). Speech coding in the auditory nerve: I. Vowel-like sounds. The Journal of the Acoustical Society of America, 75(3), 866-878. https://doi.org/10.1121/1.390596

Egan, J. P. (1975). Signal detection theory and ROC analysis. Academic Press. https://cir.nii.ac.jp/crid/1130282268734261760

Fant, G. (1971). Acoustic theory of speech production: with calculations based on X-ray studies of Russian articulations. Walter de Gruyter.

Fuzessery, Z. M. (1994). Response selectivity for multiple dimensions of frequency sweeps in the pallid bat inferior colliculus. Journal of Neurophysiology, 72(3), 1061-1079. https://doi.org/10.1152/jn.1994.72.3.1061

Fuzessery, Z. M., Razak, K. A., & Williams, A. J. (2011). Multiple mechanisms shape selectivity for FM sweep rate and direction in the pallid bat inferior colliculus and auditory cortex. Journal of Comparative Physiology A, 197(5), 615-623. https://doi.org/10.1007/s00359-010-0554-0

Fuzessery, Z. M., Richardson, M. D., & Coburn, M. S. (2006). Neural Mechanisms Underlying Selectivity for the Rate and Direction of Frequency-Modulated Sweeps in the Inferior Colliculus of the Pallid Bat. Journal of Neurophysiology, 96(3), 1320-1336. https://doi.org/10.1152/jn.00021.2006

Gittelman, J. X., Li, N., & Pollak, G. D. (2009). Mechanisms Underlying Directional Selectivity for Frequency-Modulated Sweeps in the Inferior Colliculus Revealed by In Vivo Whole-Cell Recordings. The Journal of Neuroscience, 29(41), 13030-13041. https://doi.org/10.1523/jneurosci.2477-09.2009

Gordon, M., & O'Neill, W. E. (1998). Temporal processing across frequency channels by FM selective auditory neurons can account for FM rate selectivity. Hearing research, 122(1), 97-108. https://doi.org/https://doi.org/10.1016/S0378-5955(98)00087-2

Heffner, H., & Masterton, B. (1980). Hearing in Glires: Domestic rabbit, cotton rat, feral house mouse, and kangaroo rat. The Journal of the Acoustical Society of America, 68(6), 1584-1599. https://doi.org/10.1121/1.385213

Heinz, M. G., Colburn, H. S., & Carney, L. H. (2001). Evaluating Auditory Performance Limits: I. One-Parameter Discrimination Using a Computational Model for the Auditory Nerve. Neural computation, 13(10), 2273-2316. https://doi.org/10.1162/089976601750541804

Heinz, M. G., Colburn, H. S., & Carney, L. H. (2001). Evaluating Auditory Performance Limits: II. One-Parameter Discrimination with Random-Level Variation. Neural computation, 13(10), 2317-2338. https://doi.org/10.1162/089976601750541813

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. The Journal of the Acoustical Society of America, 97(5), 3099-3111. https://doi.org/10.1121/1.411872

Hudspeth, A. J. (2014). Integrating the active process of hair cells with cochlear function. Nature Reviews Neuroscience, 15(9), 600-614. https://doi.org/10.1038/nrn3786

Huet, A., Batrel, C., Tang, Y., Desmadryl, G., Wang, J., Puel, J.-L., & Bourien, J. (2016). Sound coding in the auditory nerve of gerbils. Hearing research, 338, 32-39. https://doi.org/https://doi.org/10.1016/j.heares.2016.05.006

Johnson, D. H. (1980). The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. The Journal of the Acoustical Society of America, 68(4), 1115-1122. https://doi.org/10.1121/1.384982

Kewley-Port, D., Burkle, T. Z., & Lee, J. H. (2007). Contribution of consonant versus vowel information to sentence intelligibility for young normal-hearing and elderly hearing-impaired listenersa). The Journal of the Acoustical Society of America, 122(4), 2365-2375. https://doi.org/10.1121/1.2773986

Kim, D. O. (1986). Active and nonlinear cochlear biomechanics and the role of outer-hair-cell subsystem in the mammalian auditory system. Hearing research, 22(1), 105-114. https://doi.org/https://doi.org/10.1016/0378-5955(86)90088-2

Kim, D. O., Carney, L., & Kuwada, S. (2020). Amplitude modulation transfer functions reveal opposing populations within both the inferior colliculus and medial geniculate body. Journal of Neurophysiology, 124(4), 1198-1215. https://doi.org/10.1152/jn.00279.2020

Krips, R., & Furst, M. (2009). Stochastic Properties of Coincidence-Detector Neural Cells. Neural computation, 21(9), 2524-2553. https://doi.org/10.1162/neco.2009.07-07-563

Ladefoged, P., & Broadbent, D. E. (1957). Information Conveyed by Vowels. The Journal of the Acoustical Society of America, 29(1), 98-104. https://doi.org/10.1121/1.1908694

Langner, G., & Schreiner, C. E. (1988). Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. Journal of Neurophysiology, 60(6), 1799-1822. https://doi.org/10.1152/jn.1988.60.6.1799

Levitt, H. (1971). Transformed Up-Down Methods in Psychoacoustics. The Journal of the Acoustical Society of America, 49(2B), 467-477. https://doi.org/10.1121/1.1912375

Lu, H.-W., Smith, P. H., & Joris, P. X. (2022). Mammalian octopus cells are direction selective to frequency sweeps by excitatory synaptic sequence detection. Proceedings of the National Academy of Sciences, 119(44), e2203748119. https://doi.org/doi:10.1073/pnas.2203748119

Mitchell, P. W., Henry, K. S., & Carney, L. H. (2023). Sensitivity to direction and velocity of fast frequency chirps in the inferior colliculus of awake rabbit. Hearing research, 440, 108915. https://doi.org/https://doi.org/10.1016/j.heares.2023.108915

Mitchell, P. W. & Carney, L. H. (2024, in press). A Computational Model of Auditory Chirp-Velocity Sensitivity and Amplitude-Modulation Tuning in Inferior Colliculus Neurons. The Journal of Computational Neuroscience.

Oertel, D., Bal, R., Gardner, S. M., Smith, P. H., & Joris, P. X. (2000). Detection of synchrony in the activity of auditory nerve fibers by octopus cells of the mammalian cochlear nucleus. Proceedings of the National Academy of Sciences, 97(22), 11773-11779. https://doi.org/doi:10.1073/pnas.97.22.11773

Osen, K. K. (1969). Cytoarchitecture of the cochlear nuclei in the cat. Journal of Comparative Neurology, 136(4), 453-483. https://doi.org/https://doi.org/10.1002/cne.901360407

Oxenham, A. J., & Bacon, S. P. (2003). Cochlear Compression: Perceptual Measures and Implications for Normal and Impaired Hearing. Ear and Hearing, 24(5). https://journals.lww.com/ear-hearing/fulltext/2003/10000/cochlear_compression__perceptual_measures_and.2.aspx

Pollak, G. D., Gittelman, J. X., Li, N., & Xie, R. (2011). Inhibitory projections from the ventral nucleus of the lateral lemniscus and superior paraolivary nucleus create directional selectivity of frequency modulations in the inferior colliculus: A comparison of bats with other mammals. Hearing research, 273(1), 134-144. https://doi.org/https://doi.org/10.1016/j.heares.2010.03.083

Portfors, C. V. (2018). Chapter 7 - Processing of Ultrasonic Vocalizations in the Auditory Midbrain of Mice. In S. M. Brudzynski (Ed.), Handbook of Behavioral Neuroscience (Vol. 25, pp. 73-82). Elsevier. https://doi.org/https://doi.org/10.1016/B978-0-12-809600-0.00007-X

Rees, A., & Langner, G. (2005). Temporal Coding in the Auditory Midbrain. In J. A. Winer & C. E. Schreiner (Eds.), The Inferior Colliculus (pp. 346-376). Springer New York. https://doi.org/10.1007/0-387-27083-3_12

Sachs, M. B., Blackburn, C. C., & Young, E. D. (1988). Rate-place and temporal-place representations of vowels in the auditory nerve and anteroventral cochlear nucleus. Journal of Phonetics, 16(1), 37-53. https://doi.org/https://doi.org/10.1016/S0095-4470(19)30465-6

Sachs, M. B., & Young, E. D. (1979). Encoding of steady-state vowels in the auditory nerve: Representation in terms of discharge rate. The Journal of the Acoustical Society of America, 66(2), 470-479. https://doi.org/10.1121/1.383098

Schroeder, M. (1970). Synthesis of low-peak-factor signals and binary sequences with low autocorrelation (Corresp.). IEEE Transactions on Information Theory, 16(1), 85-89. https://doi.org/10.1109/TIT.1970.1054411

Siebert, W. M. (1965). Some implications of the stochastic behavior of primary auditory neurons. Kybernetik, 2(5), 206-215. https://doi.org/10.1007/BF00306416

Siebert, W. M. (1970). Frequency discrimination in the auditory system: Place or periodicity mechanisms? Proceedings of the IEEE, 58(5), 723-730. https://doi.org/10.1109/PROC.1970.7727

Steenken, F., Oetjen, H., Beutelmann, R., Carney, L. H., Koeppl, C., & Klump, G. M. (2022). Neural processing and perception of Schroeder-phase harmonic tone complexes in the gerbil: Relating single-unit neurophysiology to behavior. European Journal of Neuroscience, 56(3), 4060-4085. https://doi.org/https://doi.org/10.1111/ejn.15744

Stiebler, I., & Ehret, G. (1985). Inferior colliculus of the house mouse. I. A quantitative study of tonotopic organization, frequency representation, and tone-threshold distribution. Journal of Comparative Neurology, 238(1), 65-76. https://doi.org/https://doi.org/10.1002/cne.902380106

Vater, M., Covey, E., & Casseday, J. H. (1997). The columnar region of the ventral nucleus of the lateral lemniscus in the big brown bat (Eptesicus fuscus): synaptic arrangements and structural correlates of feedforward inhibitory function. Cell and tissue research, 289(2), 223-233. https://doi.org/10.1007/s004410050869

Yates, G. K. (1990). Basilar membrane nonlinearity and its influence on auditory nerve rate-intensity functions. Hearing research, 50(1), 145-162. https://doi.org/https://doi.org/10.1016/0378-5955(90)90041-M

Yost, W. A. (2013). Fundamentals of hearing : an introduction (Fifth edition. ed.). Brill.

Young, E. D., & Sachs, M. B. (1979). Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers. The Journal of the Acoustical Society of America, 66(5), 1381-1403. https://doi.org/10.1121/1.383532

Chapter 2: Sensitivity to Direction and Velocity of Fast Frequency Chirps in the Inferior Colliculus of Awake Rabbit

This chapter is published in Hearing Research.

2.1 Abstract

Neurons in the mammalian inferior colliculus (IC) are sensitive to the velocity (speed and direction) of fast frequency chirps contained in Schroeder-phase harmonic complexes (SCHR). However, IC neurons are also sensitive to stimulus periodicity, a prominent feature of SCHR stimuli. Here, to disentangle velocity sensitivity from periodicity tuning, we introduced a novel stimulus consisting of aperiodic random chirps. Extracellular, single-unit recordings were made in the IC of Dutch-belted rabbits in response to both SCHR and aperiodic chirps. Rate-velocity functions were constructed from aperiodic-chirp responses and compared to SCHR rate profiles, revealing interactions between stimulus periodicity and neural velocity sensitivity. A generalized linear model analysis demonstrated that periodicity tuning influences SCHR response rates more strongly than velocity sensitivity. Principal component analysis of rate-velocity functions revealed that neurons were more often sensitive to the direction of lower-velocity chirps and were less often sensitive to the direction of higher-velocity chirps. Overall, these results demonstrate that sensitivity to chirp velocity is common in the IC. Harmonic sounds with complex phase spectra, such as speech and music, contain chirps, and velocity sensitivity would shape IC responses to these sounds.

2.2 Introduction

The inferior colliculus (IC) receives convergent excitatory and inhibitory inputs from several auditory brainstem nuclei, and consequently displays a variety of complex response properties including frequency tuning (Davis, 2005), sensitivity to interaural differences (Yin et al., 2019), duration tuning (Casseday et al., 1994), and tuning to amplitude modulation (AM) (Langner and Schreiner, 1988; Krishna and Semple, 2000). Sensitivity of IC neurons to the velocity (speed and direction) of frequency sweeps has also been studied. Studies of bat IC responses report neurons that prefer, or even respond exclusively to, a particular sweep direction, a possible adaptation for echolocation (Suga, 1965; Fuzessery, 1994; Fuzessery and Hall, 1996; Gordon and O'Neill, 1998; Pollak et al., 2011). Additionally, sounds with statistics resembling vocal spectral-peak (formant) transitions have been used to produce IC spectrotemporal receptive fields (STRFs) in cats and bats (Escabí and Schreiner, 2002; Escabí et al., 2003; Andoni et al., 2007).

Recently, IC sensitivity to the velocity of sweeps has been explored in anesthetized gerbils (Steenken et al., 2022) and awake budgerigars (Henry et al., 2023). These studies used Schroeder-phase (SCHR) harmonic complexes, an advantageous stimulus for studying frequency-sweep sensitivity. The harmonic components of SCHR stimuli have flat magnitude spectra and curved phase spectra that create periodic acoustic waveforms with maximally flat temporal envelopes (Schroeder, 1970). The phase differences across harmonic components result in linear frequency sweeps within each fundamental period, extending from the lowest component, typically the fundamental frequency (F0) or its first harmonic, to the highest harmonic component.

As a tool for interrogating neural velocity sensitivity, the SCHR stimulus is unique in several ways. First, SCHR stimuli are notable for the speed of frequency sweeps, which travel from F0 to the maximum harmonic (here, 16 kHz) over the fundamental period (here, between 2.5 and 20 ms). The resulting velocities are faster than most commonly studied frequency-modulated signals, such as formant transitions (Liberman and Mattingly, 1989). We refer to these fast sweeps as SCHR chirps. Second, as a harmonic complex, the SCHR stimulus resembles voiced speech sounds. SCHR chirps emerge from phase differences between harmonic components. Similar phase differences are present in sounds like vowels, introduced into the glottal pulse by vocal-tract filtering. Therefore, neural sensitivity to chirp velocity would likely influence IC responses to speech sounds.

However, as harmonic stimuli, SCHR stimuli also have strong periodicity. IC neurons are notable for their tuning to the modulation frequency of amplitude-modulated (AM) sounds, often having enhanced or suppressed response rates to AM stimuli relative to an unmodulated stimulus (Krishna and Semple, 2000; Joris et al., 2004; Kim et al., 2020). It is unclear whether sensitivity to chirps requires a periodic context and whether the sensitivity is modulated by stimulus periodicity. Here, we use a new stimulus that isolated the velocity sensitivity from periodicity tuning. We recorded IC responses to both periodic SCHR stimuli and to aperiodic, random-chirp stimuli to assess velocity sensitivity. We explored the differences in prevalence of direction selectivity between the two stimuli and quantified the relative contributions of periodicity and velocity sensitivity in predicting SCHR response rates.

2.3 Methods

Extracellular, single-unit recordings were made in the central nucleus of the IC (ICC) of six female Dutch-belted rabbits (*Oryctolagus cuniculus*). Rabbits were obtained from Envigo (Denver, PA). The age of the animals during experiments ranged from 1-5 years, during which distortion-product otoacoustic emissions (DPOAEs) and IC neural thresholds were used to confirm normal hearing (Whitehead et al., 1992). Animals were housed in separate cages in the same room, with a 12-hour light cycle from 6 AM to 6 PM. Animals were removed from the study when their DPOAE magnitudes decreased or neural thresholds increased. All methods were approved by the University of Rochester Committee on Animal Resources.

2.3.1 Recordings

Recordings were made in daily two-hour sessions between 9 AM and 4 PM in sound-attenuated booths (Acoustic Systems, Austin, TX, USA). To prevent head movement of the awake rabbits, a headbar (metal or 3D-printed plastic) was surgically affixed to the skull using stainless-steel screws and dental acrylic. Rabbits were anesthetized using ketamine (66 mg/kg) and xylazine (2 mg/kg), administered intramuscularly, for headbar placement, craniotomy, and electrode placements (below). Rabbits were wrapped in a towel and secured in a custom-made chair to limit movement during recording sessions.

Recordings were made using four tetrodes, each consisting of four 18-μm platinum-iridium epoxy-coated wires (California Fine Wire Co., Grover Beach, CA, USA) and plated with platinum black to obtain impedances of approximately 0.1 – 0.5 MOhms. Tetrodes were surgically implanted in the ICC via a small (~2-mm diameter) craniotomy

in the top of the skull. Stainless-steel guide tubes protected the tetrodes, which were positioned just dorsal to the ICC and then advanced through the ICC using a manual microdrive (Five-drive, Neuralynx, Inc., Bozeman, MT, USA). Tetrodes were advanced before or after recording sessions. The microdrive and tetrodes were surgically replaced every few months, as needed, to vary the recording location. Neural recordings were made using an RHD recording system (Intan Technologies, LLC., Los Angeles, CA, USA), with Intan software running on a Windows PC.

Stimuli were generated using custom MATLAB code (Mathworks, Natick, MA, USA), a MOTU audio interface (16A, Mark of the Unicorn, Cambridge, MA, USA), and a Benchmark digital-to-analog converter and amplifier (DAC3 HGC, Benchmark Media Systems, Inc., Syracuse, NY, USA). Etymotic ER2 earphones (Etymotic Research, Inc., Elk Grove Village, IL, USA) were used to present sound stimuli via custom ear molds (Hal-Hen Company, Inc., Garden City Park, New York, USA). Prior to every session, stimuli were calibrated using ER-7C probe-tube or ER-10B+ microphones (Etymotic Research, Inc., USA). The calibration curve (magnitude and phase) was included in a pre-emphasis filter that compensated all stimuli for the frequency response of the acoustic system.

Action potential times and waveforms were extracted from neural recordings using custom MATLAB applications. To identify single-unit events, the voltage recording was filtered using a 4^{th}-order Butterworth bandpass filter (300 – 3000 Hz). An action potential was identified when the filtered voltage signal exceeded a threshold, defined as four standard deviations of the voltage signal. Action potentials were sorted into clusters, presumed to represent different neurons, using features of the waveforms, primarily the slope of the waveform repolarization (Schwarz et al., 2012). Single-units were identified

when less than 2% of the interspike intervals were shorter than 1 ms. Neurons studied in consecutive sessions were considered to be unique only when both tetrode location and response properties changed.

2.3.2 Stimuli

A set of stimuli was presented to define neural response properties, including frequency response maps, modulation transfer functions, SCHR sensitivity, and rate-velocity functions for aperiodic-chirp stimuli. To generate frequency response maps, a randomly-ordered sequence of 0.2-s-duration tones at frequencies ranging from 250 Hz – 16 kHz was presented at 13, 33, 53, and 73 dB SPL. Tones were presented either contralaterally or diotically with 10-ms raised-cosine on/off ramps, and were separated by 0.4 s of silence. Each stimulus condition was presented three times. Response maps, plots of average discharge rate vs. frequency at each stimulus level, were used to estimate characteristic frequency (CF), the frequency for which threshold was lowest.

Modulation transfer functions (MTFs) were estimated by computing the average rate in response to 100% sinusoidally amplitude-modulated noise as a function of modulation frequency, over a range of 2 – 350 Hz, with 3 steps per octave. The wideband noise (100 Hz – 10 kHz) was 1 s in duration, including 50-ms raised-cosine on/off ramps, presented diotically. Modulation frequencies were presented in random order, with 5 repetitions of each. Noise presentations were separated by 0.5 s of silence. Stimuli were presented at a spectrum level of 33 dB (overall level of 73 dB SPL). MTF categorization was based on criteria that were slightly modified from Kim et al. (2020). MTF types were categorized as band-enhanced (BE) if (1) the MTF had two consecutive points significantly greater than 120% of the unmodulated rate (r_{um}), and (2) the MTF did

not have more than one point significantly below 80% of r_{um}. Similarly, MTFs were categorized as band-suppressed (BS) if (1) the MTF had two consecutive points significantly less than 80% of r_{um}, and (2) the MTF did not have more than one point significantly above 120% of r_{um}. MTFs that had both enhanced and suppressed bands that met the above criteria were categorized as hybrid MTFs. In the population of MTFs studied (N = 335), 177 (52.8%) were BS, 73 (21.8%) were BE, 51 (15.2%) were hybrid, and 34 (10.2%) were flat.

SCHR stimuli were produced using the phase spectrum introduced in Schroeder (1970) (Fig. 2.1A-D). For a harmonic complex with n components, the phase of the nth component was described by $\theta_n = C\pi n(n+1)/N$, where N is the total number of harmonic components, and C is a scalar that describes the phase-direction and duty cycle of the stimulus. SCHR stimuli with a positive C-value will be referred to as downward SCHR, reflecting the direction of their frequency chirps. SCHR stimuli with a negative C-value will be referred to as upward SCHR. Note that the direction of SCHR chirp (upward or downward) is the opposite of its phase (negative or positive, respectively). A random sequence of SCHR stimuli was presented, with F0 equal to 50, 100, 200, or 400 Hz, and C equal to -1, -0.5, 0.5, or 1. For all cases, the frequency of the highest harmonic component was 16 kHz. The number of harmonic components, N, depended on F0, and ranged between 40 and 320. Thirty repetitions of each F0 and C combination were presented. Stimuli were presented diotically at 63 dB SPL, with a duration of 0.4 seconds, 0.6 seconds of silence between stimuli, and 25-ms raised-cosine on/off ramps. Average discharge rates over the stimulus duration were calculated for each F0 and C combination, excluding a 25-ms window after stimulus onset. Receiver-operating-characteristic (ROC) analysis (Egan, 1975) was used to evaluate

neural direction selectivity for SCHR conditions with equivalent F0 and C-values of opposite signs (e.g., F0 = 50 Hz, C = ±1). For this analysis, the response rates for each stimulus repetition were gathered, for chirps of equal speed but opposite direction. Then, ROC was used to measure the discriminability of chirp direction based on response rates of single stimulus repetitions. Neurons were deemed significantly biased toward one sweep direction if the ROC area under the curve (AUC) was above 70.7% or below 29.3%. The two separate ROC criteria were used to differentiate between selectivity for upward and downward chirps (akin to a two-tailed t-test). Neurons with AUCs above 70.7% were designated as selective for upward chirps (indicating greater mean response rate in response to upward chirps) and neurons with AUCs below 29.3% were designated as selective for downward chirps. The ROC criteria of 70.7% and 29.3%

were chosen because they are equivalent to the targeted threshold of a two-down, one-up psychophysical procedure (Levitt, 1971).

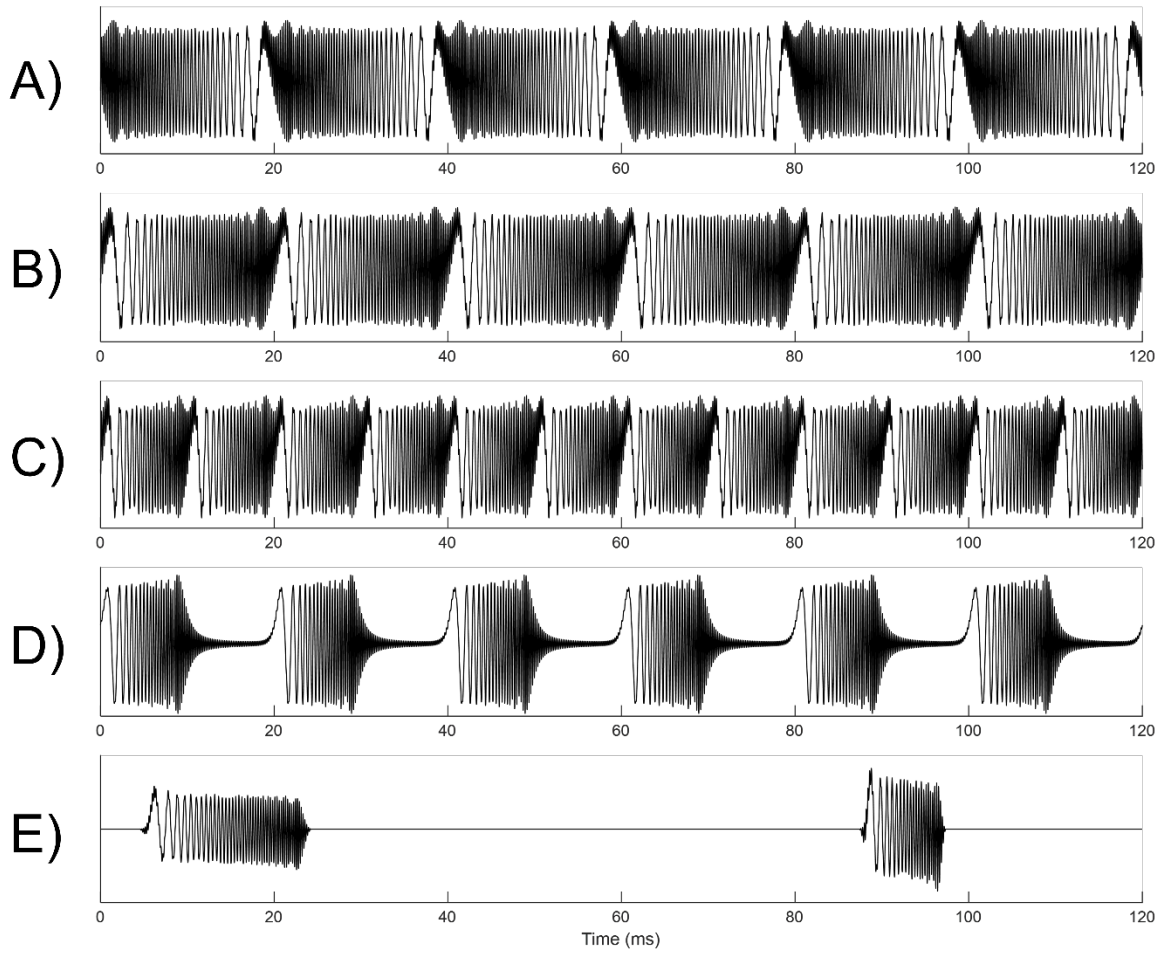

Figure 2.1 – Stimulus waveforms of SCHR and aperiodic-chirp stimuli. Note that the maximum harmonic in this figure is 5 kHz for clarity; in the actual stimuli, maximum harmonic was 16 kHz. (A) Downward SCHR with F0 = 50 Hz, C = +1. (B) Upward SCHR with F0 = 50 Hz, C = -1. (C) Upward SCHR with F0 = 100 Hz, C = -1. (D) Upward SCHR with F0 = 50 Hz, C = -0.5. (E) Two consecutive aperiodic chirps separated by silence.

Rate-velocity functions (RVFs) were based on responses to aperiodic-chirp stimuli consisting of isolated SCHR chirps separated by silent periods of variable duration. These chirps were single fundamental periods extracted from a SCHR stimulus, equivalent to one full SCHR chirp spanning F0 to the maximum harmonic, or

vice versa (Fig. 2.1E). To produce the stimulus waveforms, first, we generated a random sequence of chirp velocities that were equivalent to SCHR F0s of 25, 50, 100, 200, 400, and 600 Hz. Chirp direction (upward or downward, i.e., C = -1 and C = 1, respectively) was also randomly varied across stimuli. These parameters together resulted in aperiodic chirps with velocities ±0.40, ±0.80, ±1.59, ±3.16, ±6.24, and ±9.24 kHz/ms. Each combination of velocity and direction was presented a total of 42 times, in random order. To avoid periodicity, chirps were randomly spaced, using 40 – 60 ms uniformly distributed inter-chirp intervals. The full stimulus set was presented 20 times. Each individual chirp had raised-cosine on/off ramps, each with duration equivalent to 10% of chirp duration. Chirps were presented diotically at a sound level equal to $68\ \mathrm{dB\ SPL} - 10 \times \log_{10}(T/T_{ref})$, where T is the duration of the chirp, and T_{ref} = 2.5 ms (the duration of a SCHR chirp with F0 = 400 Hz). This scaling ensured that energy was equivalent across chirps of different durations.

Response rate was calculated by summing spikes over a 15-ms time window starting at an estimate of the neural latency (Fig. 2.2) based on the response to a 73-dB tone at CF (from the response map). In cases where a neuron responded to 73-dB SPL tones over a wide frequency range, the neural latency to a chirp was sometimes substantially shorter than would be predicted by the latency at CF, as expected if instantaneous frequencies well below or above CF excited the neuron. In these cases (approximately a quarter of the neurons), the latency estimates were adjusted based on the frequency range over which the neuron responded. ROC analysis of aperiodic-chirp rates was performed in an identical manner as for SCHR rates. Aperiodic-chirp rates were used to generate an RVF, which described how a neuron's rate changed as a function of chirp velocity (Figs. 2.3-4).

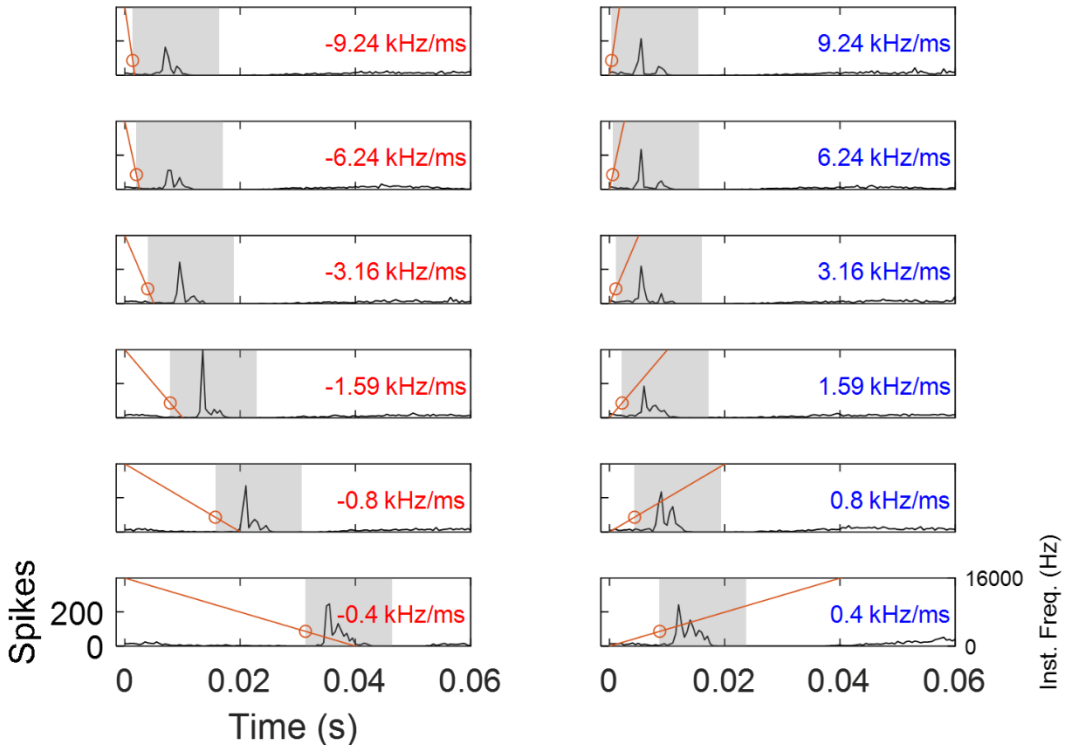

Figure 2.2 – A visual depiction of aperiodic-chirp stimulus rate calculation. Black lines illustrate PSTHs (bin size = 0.5 ms) of responses to chirps of different velocities. Orange lines indicate the instantaneous frequency of the aperiodic chirp over time. The left column (red text) shows downward chirps, and the right column (blue text) shows upward chirps. Orange circles indicate when the instantaneous frequency of the chirp passes CF (here, 3500 Hz). The shaded rectangles indicate the placement of the 15-ms window over which response rate was calculated.

2.3.3 Analysis

Principal Component Analysis (PCA) was performed to identify the salient features in RVFs, using the MATLAB function PCA (2021b, MathWorks). PCA analysis allowed grouping of RVFs by their shape to gain insight into underlying patterns or trends. RVFs were normalized by their individual peak rates before PCA analysis. SCHR responses were predicted based on MTFs and RVFs using generalized linear models (GLMs) (Nelder and Wedderburn, 1972). First, SCHR rates, MTFs, and RVFs were

normalized by each of their maximum rates. For the responses to periodic stimuli, (MTF and SCHR), cycle rate was used, defined as the average rate per fundamental period. Cycle rate (spikes/cycle) was necessary to compare responses to periodic stimuli and aperiodic chirps. For the aperiodic-chirp stimulus, rate was calculated using the method described above (Fig. 2.2). The periodicity and velocity of each SCHR stimulus was matched with the closest available modulation frequency and velocity from the MTF and RVF, respectively (data were not interpolated), creating a feature vector of the responses to periodicity and velocity changes. Then, a relationship between the periodicity and velocity feature vectors and SCHR rates was defined for the GLM, as follows:

$$R_S = b_P R_P + b_V R_V + b_I (R_P * R_V) + b_0 \quad , \qquad \text{(Eqn. 1)}$$

where R_P and R_V refer to rates of the periodicity and velocity feature vectors, respectively. R_S is the predicted response to SCHR stimuli in cycle rate, and b_n are the GLM coefficients. There were three terms in the GLM—a periodicity term, a velocity term, and an interaction term. GLM coefficients were estimated using the MATLAB function glmfit (2021b, MathWorks), with the distribution argument set to *normal* and constant argument set to *on*. The GLM coefficients were b_P, the periodicity coefficient, b_V, the velocity coefficient, b_I, the interaction-term coefficient, and b_0, a constant. A two-sided t-test was performed on the coefficients against the null hypothesis that the value of the coefficient was zero. After the SCHR rate prediction was calculated using estimated coefficients, the coefficient of determination (R^2) of the predicted and recorded SCHR cycle rates was determined to assess prediction quality. This process was repeated for two additional, simpler GLMs:

$R_S = b_1 R_P + b_0$, and (Eqn. 2)

$R_S = b_1 R_V + b_0$. (Eqn. 3)

These simpler GLMs depended only on one feature vector and thus could be used to assess the importance of periodicity or velocity alone. The coefficients of determination (R^2) of these two model predictions were obtained and used to measure the impact of the two features on neurons' SCHR responses.

The following equation was used to calculate the contribution of velocity to the total explainable variance:

$$Percent\ contribution\ of\ velocity\ to\ R^2 = \frac{\left({R^2}_{combined} - {R^2}_{periodicity}\right)}{{R^2}_{combined}} \times 100 \ , \quad \text{(Eqn. 4)}$$

where ${R^2}_{combined}$ and ${R^2}_{periodicity}$ refer to the coefficient of determination of the combined model (Eqn. 1) and periodicity-alone model (Eqn. 2), respectively. Equation 4 allowed the amount of additional variance explained by the inclusion of velocity (either in the velocity term or interaction term in Eqn. 1) to be directly quantified.

2.4 Results

2.4.1 Examples of responses to SCHR and aperiodic-chirp stimuli

Data were recorded over a span of 330 sessions and included recordings from 335 unique single units. Response properties of single-unit neurons are illustrated using frequency response maps, MTFs, SCHR rate functions, and RVFs (Figs. 2.3 and 2.4). Frequency response maps were used to determine CF; for neurons with spontaneous activity, these plots also revealed off-CF inhibitory regions. MTFs were categorized by shape (i.e., BE, BS, hybrid, or flat). SCHR rate functions were generated by summing

the total response over the duration of the periodic SCHR stimulus. Lastly, RVFs, which reflect velocity sensitivity in the absence of a periodic context, were constructed from aperiodic-chirp responses. Altogether, these stimulus responses were used to assess the interactions of different feature sensitivities, particularly to periodicity and velocity, in determining the SCHR responses.

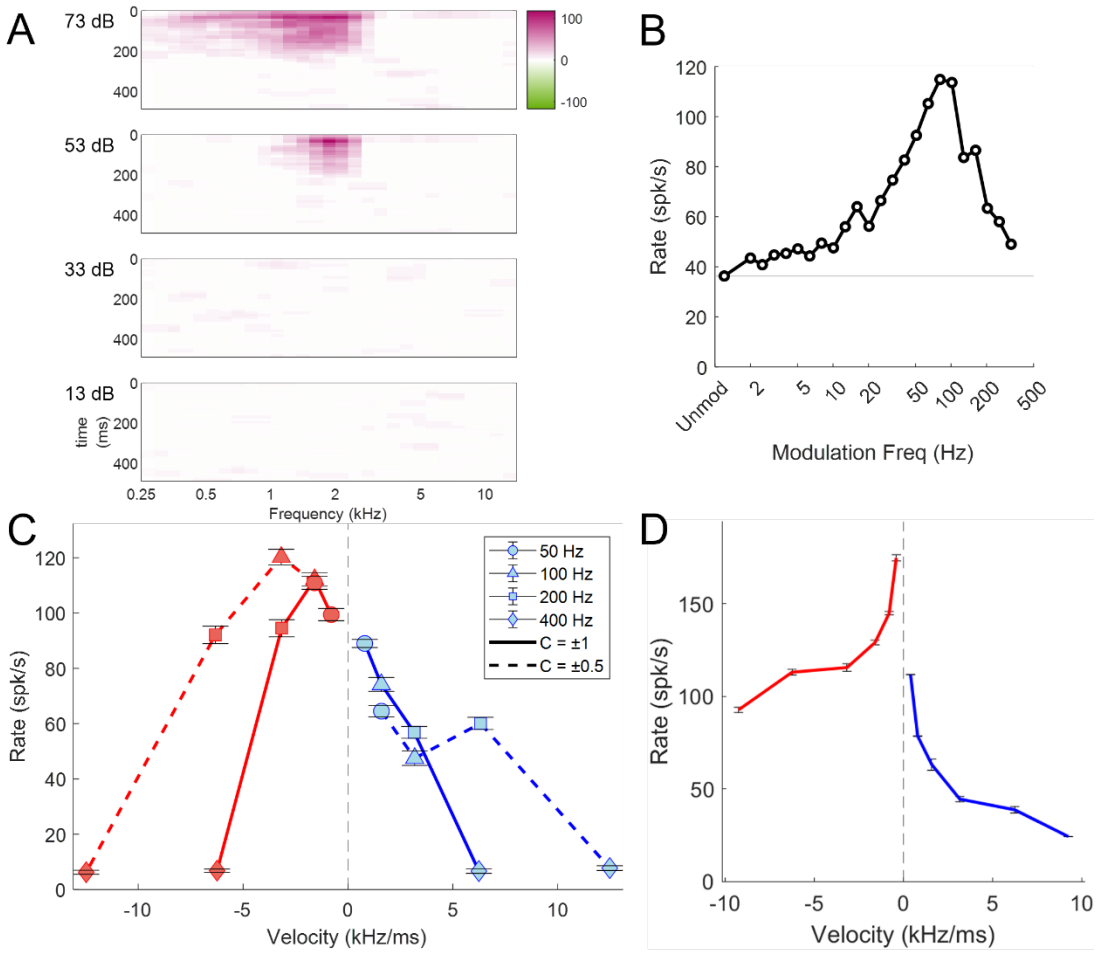

Figure 2.3 - Responses of unit R24TT2P8N4 to characterizing stimuli. (A) Response Map – Panels are arranged vertically by sound level of pure tones. Color plots reflect response rate versus frequency (x-axis) and time following tone onset (y-axis); magenta regions have increased activity relative to spontaneous rate, and green regions have decreased activity relative to spontaneous rate. (B) Modulation Transfer Function – rate versus modulation frequency. The grey horizontal line reflects unmodulated rate. (C) SCHR response rates plotted as a function of chirp velocity. Solid lines are used for C =

±1, and dashed lines are used for C = ±0.5. Line color denotes chirps of equivalent direction (red—downward chirp, blue—upward chirp). Symbols are used to reflect SCHR F0, with circles, triangles, squares, and diamonds corresponding to 50, 100, 200, and 400 Hz respectively. (D) Rate-velocity function – constructed using responses to aperiodic-chirp stimuli. Error bars are the standard error of rates across repetitions (20 each).

Example responses for two neurons are illustrated in Figs. 2.3 and 2.4. The neuron in Fig. 2.3 had a BE MTF (Fig. 2.3B) and a predominantly excitatory response map with a CF of 2 kHz (Fig. 2.3A). This neuron had chirp sensitivity in response to SCHR stimuli that was influenced by stimulus periodicity. In the SCHR rate function, for C = ±1 (Fig. 2.3C, solid lines) a greater response rate to downward chirps was observed for all but the highest velocity. This velocity corresponded to a 400-Hz F0, a modulation rate for which the neuron did not have an enhanced rate in its MTF. In comparison, the RVF (Fig. 2.3D) had a stronger rate selectivity for downward chirps at all velocities, including the high-velocity range at which the SCHR rates were not selective. Thus, the neuron's sensitivity to the velocity of SCHR chirps was impacted by stimulus periodicity, albeit both the SCHR responses and RVF were generally selective for the same chirp direction. It is notable that the C = ±0.5 SCHR responses (Fig. 2.3C, dashed lines) also differed from the responses to the aperiodic RVF, with rates that tapered off as modulation rate increased.

Other neurons' SCHR responses were more strongly impacted by their periodicity tuning, even showing opposite chirp direction selectivity to that of their RVF. An example of such a neuron is shown in Fig. 2.4 which had a BS MTF (Fig. 2.4B) with the suppressed band approaching zero rate. The response map (Fig. 2.4A) was tuned to 1.2 kHz with a low threshold (13 dB SPL), and a high-frequency inhibitory region at higher sound levels. This unit's SCHR response (Fig. 2.4C) and RVF (Fig. 2.4D) were

quite different from one another. The SCHR rate function had higher responses to downward chirps at most velocities. In contrast, the RVF displayed a higher rate for upward chirps for most velocity pairs. Notably, for C = ±1 SCHR stimuli (Fig. 2.4C, solid lines), the two lowest velocity-pairs corresponded to modulation rates in the trough region of the BS MTF. While the mechanism for this apparent direction reversal cannot be ascertained from the MTF alone, it is apparent that the periodic nature of the SCHR signal can change the chirp-direction sensitivity for some neurons.

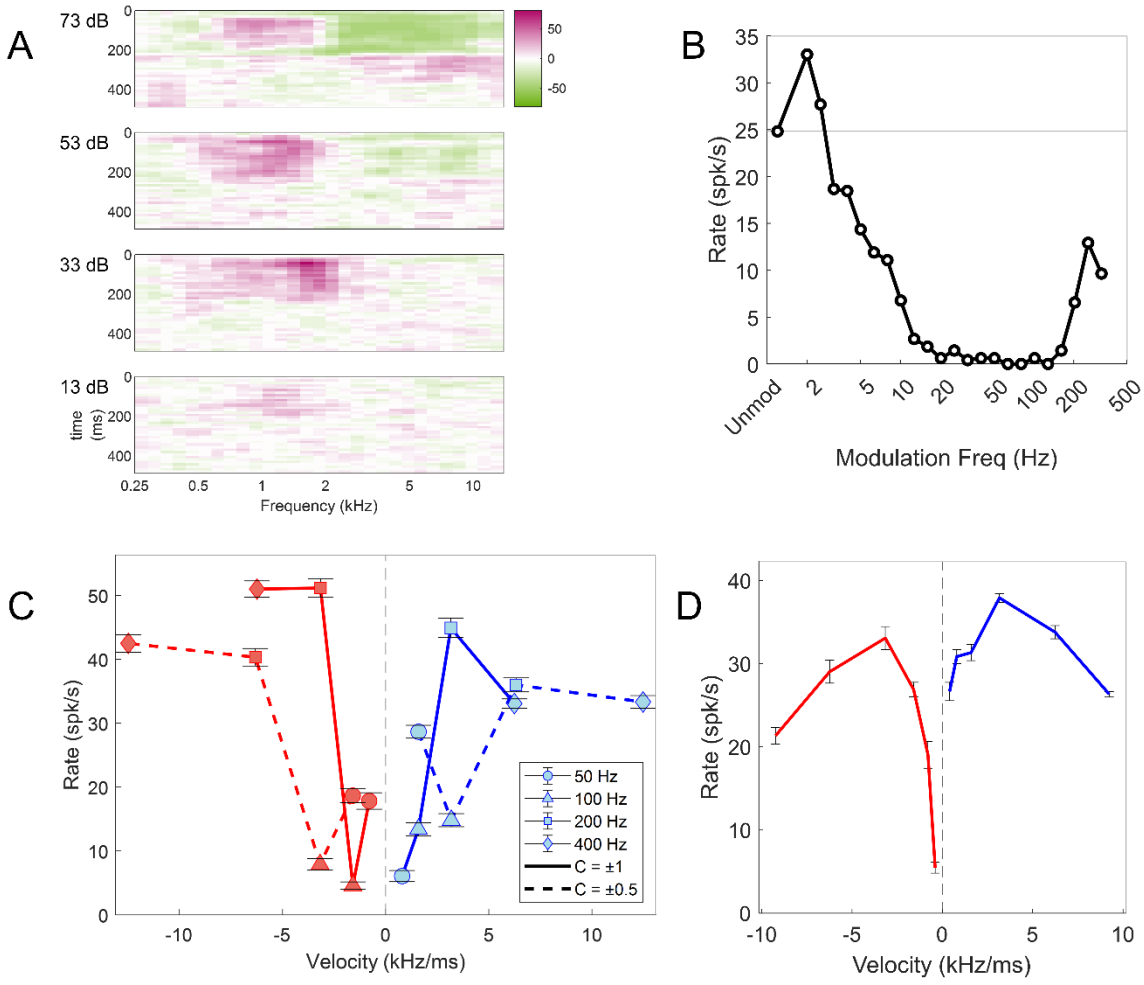

Figure 2.4 - Responses of unit R24TT4P7N2 to characterizing stimuli. (A) Response Map, (B) Modulation Transfer Function, (C) SCHR dot-raster plot, (D) Rate-Velocity Function. Figure description identical to Fig. 2.3.

2.4.2 The prevalence of direction selectivity is influenced by stimulus periodicity

The prevalence of direction selectivity in response to both SCHR and aperiodic-chirp stimuli was quantified for the population of recorded neurons using ROC analysis. SCHR response rates were organized by F0 and C magnitude and grouped by chirp direction (sign of C). Aperiodic-chirp responses were compared between chirps of equal speed and opposite direction. Neurons for which response rate allowed direction to be

discriminated 70.7% of the time via the ROC test were labeled direction-sensitive for the discriminated pair.

The ROC results across MTF groups suggest that prevalence of chirp-direction selectivity was impacted by the periodicity of the stimulus. When separated into groups based on MTF type, BS neurons were significantly less likely to be direction-selective to SCHR chirps than to aperiodic chirps ($\chi 2$ test of independence, Fig. 2.5). These significant differences were seen for F0s of 50 and 100 Hz, which correspond to typical modulation-frequency ranges for troughs in BS MTFs (Kim et al., 2020), and were likely a result of the SCHR-stimulus periodicity—the response to these SCHR F0s was suppressed, reducing sensitivity to chirp direction (as in Fig. 2.4). Furthermore, the prevalence of direction selectivity in BE and hybrid neurons was not significantly different between SCHR and aperiodic chirps for any chirp speed or SCHR F0 (Fig. 2.5). Responses to SCHR F0s that corresponded to peak modulation frequency values in BE and hybrid neurons were relatively unaffected, compared to responses to other SCHR F0s—for example, in Fig. 2.3C the direction bias at SCHR F0 50, 100, and 200 Hz was close to that of the corresponding velocities in the RVF (Fig 2.3D). However, the absence of direction bias at SCHR F0 400 Hz (Fig. 2.3C) demonstrates that chirp-direction selectivity may still be affected for an F0 outside the enhanced band. Overall, neurons were generally more sensitive to aperiodic chirps than to SCHR chirps, suggesting that stimulus periodicity may partially suppress chirp-velocity sensitivity in SCHR responses, especially in BS neurons.

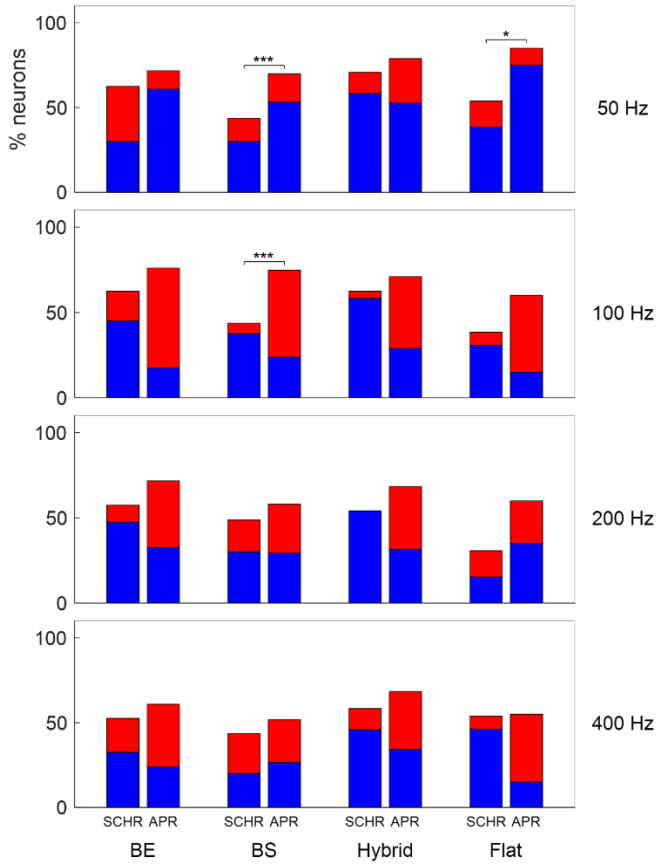

Figure 2.5 - Percent of direction-selective neurons determined by ROC analysis for different MTF shapes. For each MTF shape, the left bar shows the percent of neurons selective for SCHR chirps, and the right bar shows the percent selective for aperiodic chirps (labeled APR). Chart row reflects the stimulus F0 for SCHR selectivity, or the equivalent SCHR F0 for aperiodic chirp-direction selectivity (the corresponding absolute velocities are 0.80, 1.59, 3.16, and 6.24 kHz/ms). Blue bar-segments correspond to upward-selective neurons, and red bar-segments correspond to downward-selective neurons. Brackets indicate significant differences of direction selectivity prevalence between groups, as determined by the χ^2 test of independence, with asterisks indicating the level of significance (* $p < 0.05$, *** $p < 0.001$). p-values for each significant difference were 50 Hz—BS, $p = 0.0001$; 50 Hz—Flat, $p = 0.0496$; 100 Hz—BS, $p < 0.0001$.

Separating the ROC results by CF-range shows that stimulus periodicity impacts the prevalence of chirp-direction selectivity especially for higher-CF neurons (Fig. 2.6). Neurons in the high-CF (>6 kHz) range were direction-selective significantly more often for aperiodic chirps than for SCHR chirps, compared to low-CF (<3 kHz) and mid-CF (3-

6 kHz) neurons (χ2 test of independence, Fig. 2.6 caption). This trend in high-CF neurons was observed for 50-, 100-, and 200-Hz-equivalent chirps, but not for 400-Hz-equivalent chirps.

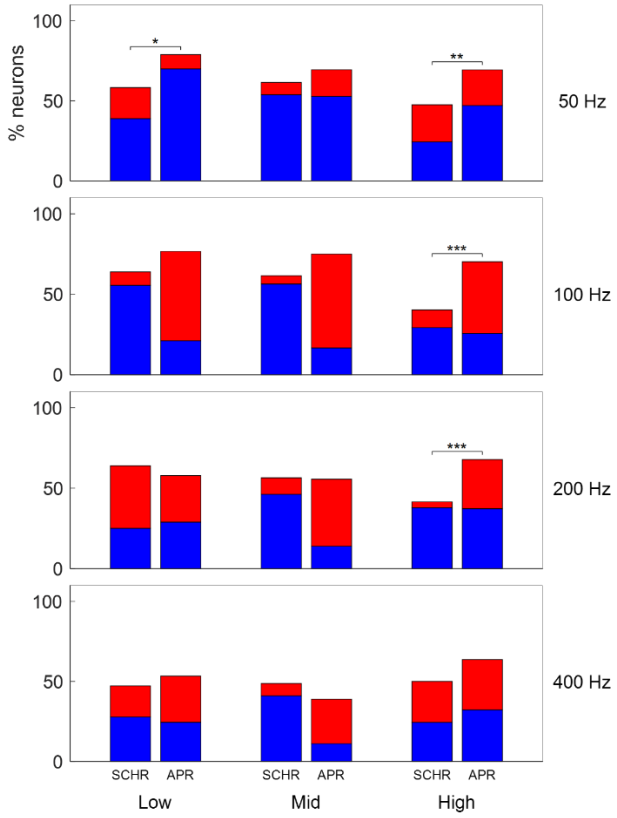

Figure 2.6—Percent of direction-selective neurons determined by ROC analysis for different CF groups—low (<3 kHz), mid (3-6 kHz), and high (>6 kHz). For each CF group, the left bar shows the percent of neurons selective for SCHR chirps, and the right bar shows the percent selective for aperiodic chirps (labeled APR). Chart row reflects the stimulus F0 for SCHR selectivity, or the equivalent SCHR F0 for aperiodic chirp-direction selectivity (the corresponding absolute velocities are 0.80, 1.59, 3.16, and 6.24 kHz/ms). Blue bar-segments correspond to upward-selective neurons, and red bar-segments correspond to downward-selective neurons. Brackets indicate significant differences of direction selectivity prevalence between groups, as determined by the χ^2 test of independence, with asterisks indicating the level of significance (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). p-values for each significant difference were 50 Hz—Low, $p = 0.0189$; 50 Hz—High, $p = 0.0018$; 100 Hz—High, $p < 0.0001$; 200 Hz—High, $p = 0.0002$.

To quantify the prevalence of chirp-selectivity, the number of units that were selective at a minimum of one condition was determined. For SCHR responses, 90.5% of neurons had at least one direction-selective F0/C pair. For responses to aperiodic chirps, 99.6% of neurons had at least one direction-selective velocity pair. To ensure that a reasonable threshold was used for determining direction selectivity, responses to each chirp-direction pair were shuffled for each neuron, and an ROC test was performed on the shuffled responses. After shuffling the responses 1000 times, 3.2% of the neurons were selective at a minimum of one SCHR F0/C condition, and 13.3% were selective at a minimum of one aperiodic-chirp velocity pair. It is clear that the high prevalence of chirp-direction selectivity could not be due to variance in the data, suggesting that the criterion used in the ROC test (70.07%) was sufficiently high.

2.4.3 Periodicity sensitivity has stronger influence on SCHR rate responses than does chirp-velocity sensitivity

We quantified the dependence of responses to SCHR stimuli on both AM (periodicity) tuning and velocity sensitivity. Using a GLM, MTFs and RVFs were used to predict SCHR responses. Figure 2.7 shows a visualization of this strategy for an example neuron (the same neuron depicted in Fig. 2.3). The SCHR response is shown as a surface plot on periodicity-velocity axes (Fig. 2.7A); darker colors correspond to higher response rates. This neuron responded more strongly to lower F0s and to downward SCHR. The shaded bars above and to the left of the SCHR surface are the MTF and RVF, respectively, and describe how the response rate depended on periodicity or velocity alone. Using a weighted combination of these two responses, the GLM predicted the SCHR response (Fig. 2.7B). By fitting an equation that included both periodicity and velocity terms (Eqn. 1), an accurate prediction of the recorded SCHR

data was made ($R^2 = 0.98$). Removing either the periodicity or velocity terms reduced the quality of the prediction (Fig. 2.7C, D). Predicting the SCHR response using the periodicity-only model (Eqn. 2) resulted in the disappearance of the distinct direction selectivity at lower F0s (Fig. 2.7C). As expected, a model using only periodicity tuning could not explain the observed direction selectivity. Conversely, a prediction of the SCHR responses using the velocity-only model (Eqn. 3) explained a relatively small amount of variance in the data ($R^2 = 0.30$). For most neurons (95%), more variance in the data was explained by the periodicity-only model than by the velocity-only model.

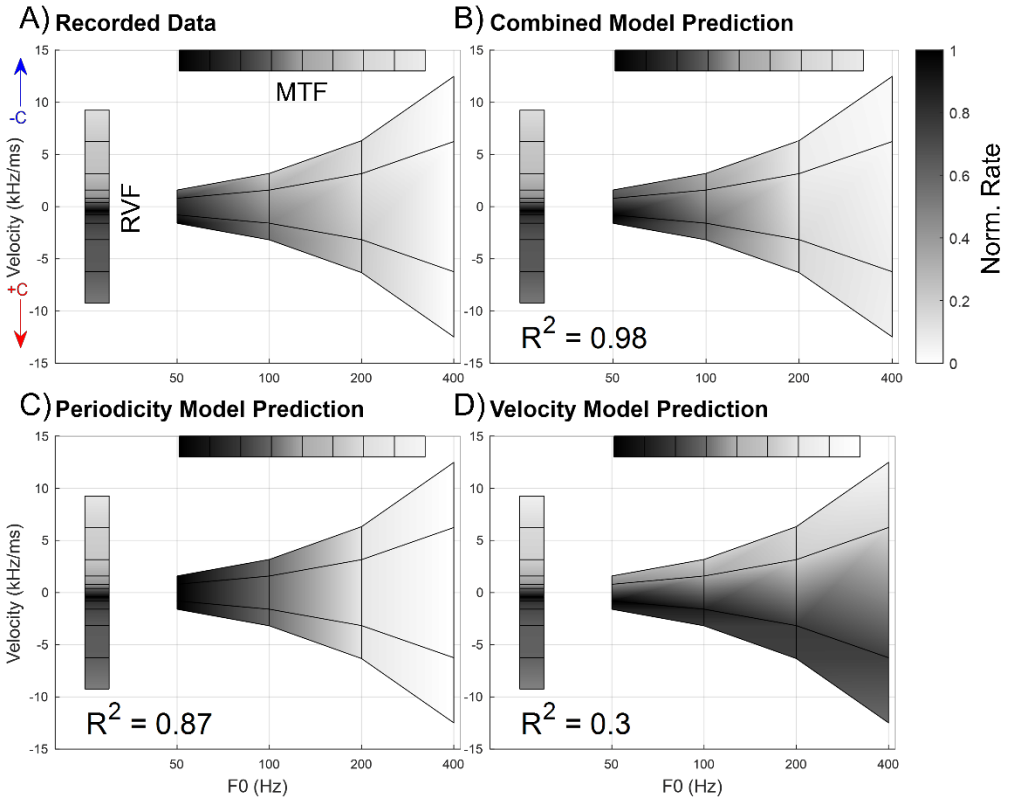

Figure 2.7 – GLM prediction of SCHR rates for an example neuron (R24TT2P8N4; the same neuron as in Fig. 2.3). (A) SCHR rates plotted on periodicity-velocity axes (center surface). Each of the 16 points on the surface plot correspond to an individual SCHR F0/C-value combination – for instance, four points have a SCHR F0 of 50 Hz. Their velocities vary with C-value; the points' C-values, from top to bottom, are -0.5, -1, 1, 0.5.

This surface plot demonstrates the effect of C-value on chirp velocity. The periodicity and velocity responses are plotted alongside the SCHR surface plot—these are the MTF and RVF, respectively. The modulation frequency range of the MTF was limited to the section relevant to the range of SCHR F0s (50 – 400 Hz). Rate is marked using color, with darker shades indicating larger rates. The rate scale of each surface plot is separately normalized by its own maximum rate—thus, shading does not reflect absolute rates. Surface plot colors are interpolated. (B) SCHR rates predicted using the combined GLM equation (Eqn. 1), which uses both periodicity and velocity cues. Coefficient of determination (R^2) between predicted and recorded SCHR rates is also shown. (C) SCHR rates predicted using the periodicity-only GLM equation (Eqn. 2), and corresponding R^2-value. (D) SCHR rates predicted using the velocity-only GLM equation (Eqn. 3), and corresponding R^2-value.

Figure 2.8 illustrates the interaction of stimulus periodicity and velocity in the combined GLM model. For aperiodic-chirp velocities that closely matched the velocities of SCHR chirps with F0 equal to 50 Hz, a large aperiodic-chirp response rate predicts a large SCHR response rate (blue line). For higher F0s, aperiodic-chirp response rate had a smaller impact on the predicted SCHR response rate, with slopes approaching zero (200- and 400-Hz lines, yellow and purple). Therefore, despite the RVF suggesting a strong selectivity for downward sweeps at all velocities, Fig. 2.8 shows that the interaction between velocity and periodicity sensitivities resulted in lower SCHR response rates for high F0s and suppressed direction selectivity. Similar velocity-periodicity interactions as illustrated in Fig. 2.8 were typical of many neurons in the study; the interaction term, b_I, for 24% of neurons was statistically significant in predicting SCHR response rates (two-sided t-test, $p < 0.05$). In comparison, the periodicity coefficient, b_P, was statistically significant for 40% of neurons, and the velocity coefficient, b_V, was significant for 7%.

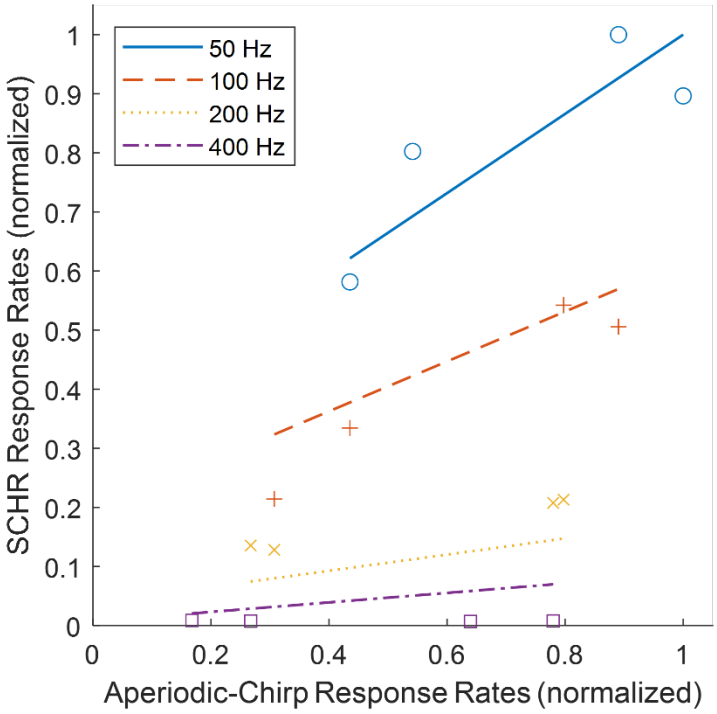

Figure 2.8 – SCHR response rates vs aperiodic-chirp response rates of equivalent velocity for R24TT2P8N4. GLM predictions are shown as lines. Blue solid line and circles correspond to 50 Hz, orange dashed line and plus signs correspond to 100 Hz, yellow dotted line and crosses correspond to 200 Hz, and purple dash-dotted line and squares correspond to 400 Hz. Both SCHR and aperiodic-chirp rates are normalized by their own maximum rate. This example neuron is also shown in Figs. 2.3 and 2.7.

Using the method outlined above (Fig. 2.7), an R^2 value was obtained for each of the three GLM models (Eqns. 2-4) by comparing model predictions to neural responses. The contribution of the velocity term to the total explainable variance was calculated as a percentage, using Eqn. 4. Figure 2.9 shows this percentage versus the variance explained by the combined model. For 95% of the neurons studied, periodicity was a more important feature than velocity in predicting SCHR response rates. In neurons with high explainable variance (80-100%), the contribution of velocity rarely exceeded 20%. Velocity was usually more important in neurons with a low overall explainable variance.

These results demonstrate that, while velocity sensitivity is critical to explain chirp-direction sensitivity, the periodicity tuning tends to dominate the SCHR response rates.

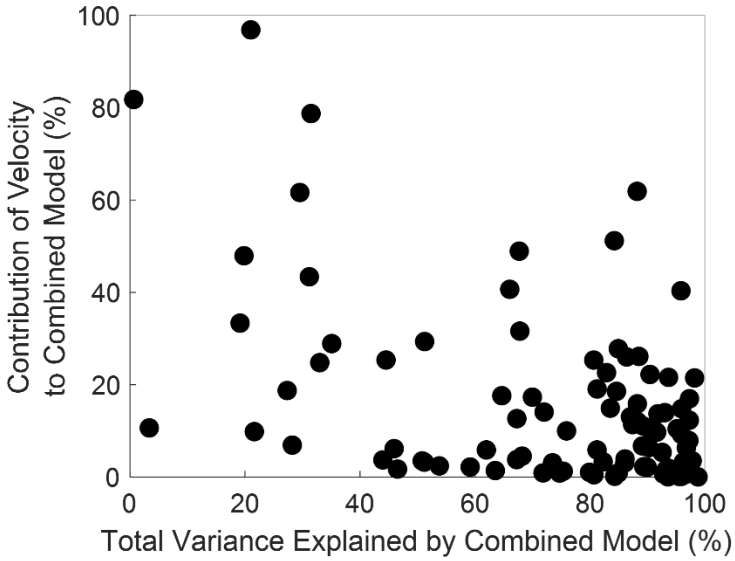

Figure 2.9 - Percent contribution of the velocity terms to the total explainable variance by the combined model vs. variance explained by the combined model.

2.4.4 RVFs were often direction-selective at slower velocities and direction-insensitive at higher velocities

Responses to aperiodic chirps isolate neural sensitivity to chirp velocity, so it is interesting to consider trends in these responses across cell groups. Figure 2.10 shows the results of ROC analysis for aperiodic-chirp responses, for neurons categorized by MTF type (Fig. 2.10A) and CF range (Fig. 2.10B). For both MTF type or CF range, there were no systematic differences between groups in prevalence of chirp-direction selectivity. Additionally, direction bias between both MTF groups and CF groups for a given chirp speed were comparable. However, one notable trend was that prevalence of selectivity decreased with increasing chirp speeds. Additionally, upward-selectivity was

more common for lower-speed chirps (0.80 kHz/ms absolute velocity and below), and downward-selectivity was more common for higher-speed chirps (1.59 kHz/ms absolute velocity and above). Overall, for aperiodic-chirp responses, there was not a trend across MTF types or CF, in either prevalence of direction selectivity or direction of bias. The direction selectivity observed was diverse, with both upward- and downward-selective units seen for every category.

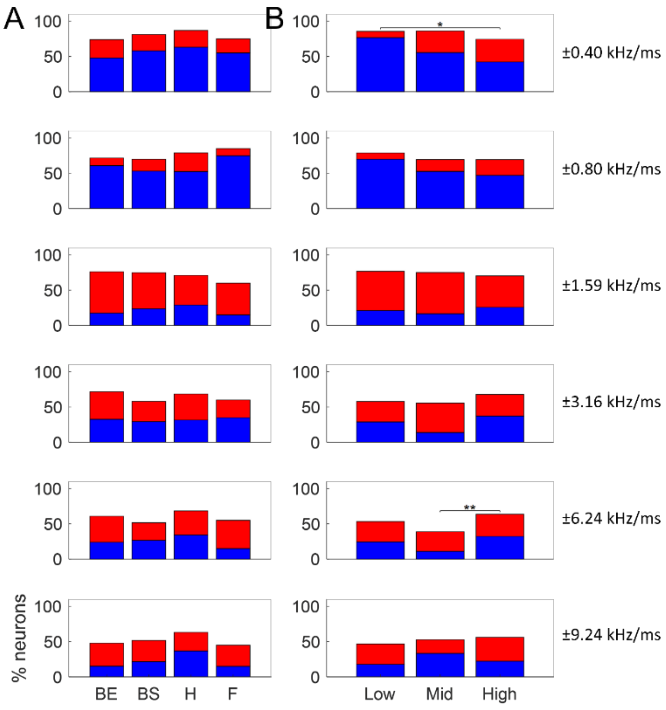

Figure 2.10 – Percent of direction-selective neurons determined by ROC analysis, by aperiodic-chirp velocity. (A) Neurons are grouped by MTF shape (BE – Band-Enhanced, BS – Band-Suppressed, H – Hybrid, F – Flat). (B) Neurons are grouped by CF range—low (<3 kHz), mid (3-6 kHz), and high (>6 kHz). Rows correspond to velocities of chirp duration, with velocities ±0.40, ±0.80, ±1.59, ±3.16, ±6.24, and ±9.24 kHz/ms corresponding to equivalent SCHR F0s of 25, 50, 100, 200, 400, and 600 Hz, respectively. Blue bar-segments correspond to upward-selective neurons, and red bar-segments correspond to downward-selective neurons. Brackets indicate significant differences of direction selectivity prevalence between groups, as determined by the χ^2 test of independence, with asterisks indicating the level of significance (* $p < 0.05$, ** $p < 0.01$). p-values for each significant difference were ±0.40 kHz/ms—Low-High, $p = 0.0481$; ±6.24 kHz/ms—Mid-High, $p = 0.0083$.

As demonstrated by the ROC analysis performed on aperiodic-chirp responses, significant direction selectivity to aperiodic chirps was commonplace. However, this analysis only addressed selectivity between direction of chirp-pairs and not sensitivity to chirp velocity. Although RVFs revealed velocity sensitivity to aperiodic chirps for individual neurons (Figs. 2.3-4), further analysis was needed to characterize trends in

velocity sensitivity for the population of neurons. In order to identify the most common RVF characteristics, principal component analysis (PCA) was performed on the set of RVFs obtained from the population of single-units (247 in total). This approach reduced the dimensionality of the RVF dataset and extracted the principal components, or the features that explained the most variability in the population.

Figure 2.11 shows the first three principal components, which together explained 74.0% of the variance in the RVFs. Principal component 1 (PC1) had relatively high weights across high-magnitude velocities (3.16 kHz/ms and faster), showing that rate responses to high-velocity chirps tended to be positively correlated, regardless of direction. In other words, neurons responded similarly to either direction of high-velocity chirp. PC1 explained 48.5% of variance in the RVF data. Principal component 2 (PC2) had high weights at low-magnitude velocities (1.59 kHz/ms and slower), showing that rate responses to low-velocity chirps also tended to be positively correlated, regardless of direction. PC2 explained 13.8% of variance in the data. Finally, PC3 had large positive weights at slow upward velocities and negative weights at slow downward velocities (1.59 kHz/ms and slower), demonstrating that responses to upward and downward velocity chirps tended to be negatively correlated at slow velocities. PC3 was the main PCA dimension reflecting direction selectivity. PC3 explained 11.7% of variance in the data.

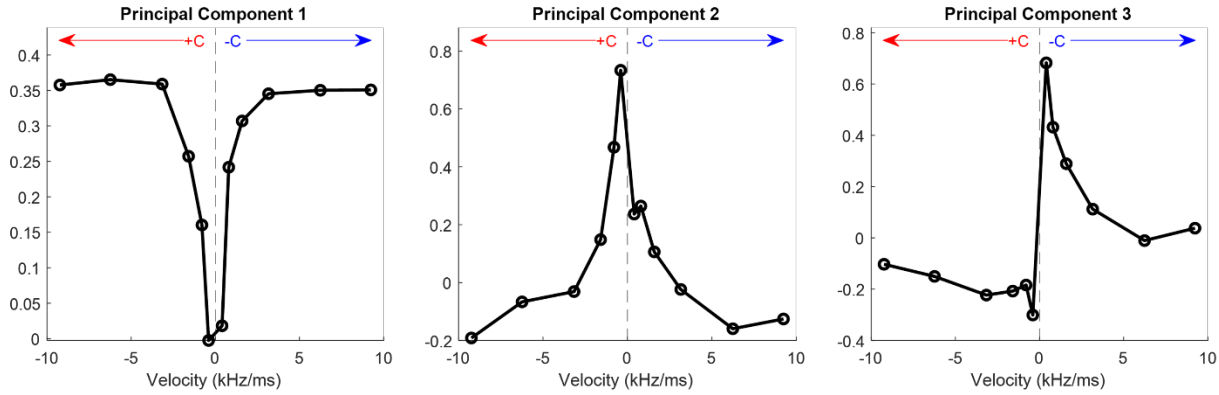

Figure 2.11 – Principal components derived from the set of RVFs (left, PC1; center, PC2; right, PC3). Y-axes reflect principal-component weights and are unitless.

The shapes of the first two principal components of the RVF data suggest that the most salient features of the RVF were the responses to high- and low-velocity chirps, rather than the direction of those chirps, with the boundary between these groups located at about 2 kHz/ms. PC3 showed that the direction selectivity was an important feature of these datasets; specifically, PC3 primarily represented selectivity to the direction of low-velocity chirps. In contrast, neural responses to chirps of high velocities depended more on chirp speed (magnitude of velocity) rather than direction, with the divide between low- and high-velocity chirps at about 2 kHz/ms. These results are consistent with the observation that direction selectivity to high-velocity chirps was not as common (Fig. 2.10).

Principal-component scores provide a measure of how well each RVF aligns with the principal components in Fig. 2.11, and thus are a useful way to divide the population based on type of velocity sensitivity. This approach also allowed a comparison of velocity sensitivity against other response properties, such as MTF type and CF. As shown in Fig. 2.12, there was not a clear relationship between the scores of principal

components 1-3 and either MTF type (symbol shapes) or CF (symbol colors). Furthermore, based on the scatter plots in Fig. 2.12, there were no clusters of neurons based on principal-component analysis – that is, RVF subgroups were not revealed by this analysis. This result suggests that a neuron's RVF was not directly related to its MTF or CF.

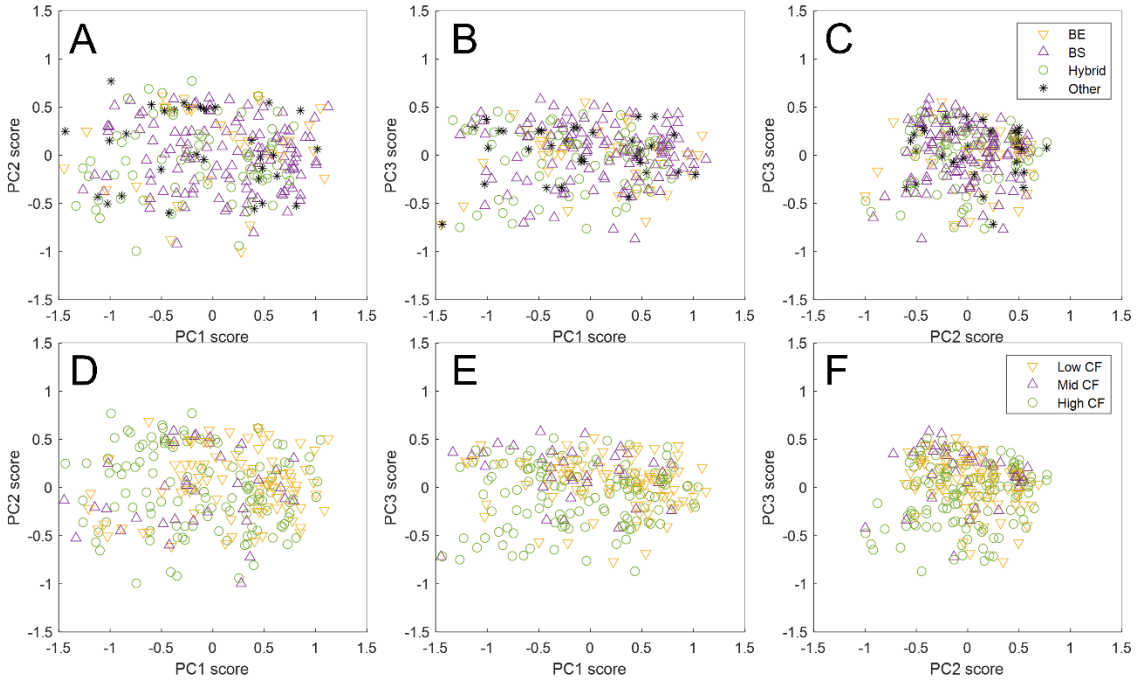

Figure 2.12 - Scatter plots depicting the Principal Component scores of each neuron, organized by MTF shape (A-C) and CF range (D-F). (A and D, PC2 vs PC1; B and E, PC3 vs PC1; C and F, PC3 vs PC2). In the top row (A-C), BE neurons are represented by yellow up-pointing triangles, BS neurons by purple down-pointing triangles, hybrid neurons by green circles, and other MTF shapes by black asterisks. In the bottom row (D-F), Low-CF neurons are represented by yellow up-pointing triangles, mid-CF by purple down-pointing triangles, and high-CF by green circles.

2.5 Discussion

This study demonstrated the diversity of sensitivity to the velocity of fast frequency chirps in the IC, for the first time in an awake rabbit model. Additionally, this

study introduced aperiodic-chirp stimuli as a way to isolate neural velocity sensitivity from periodicity tuning. ROC analysis was used to calculate the prevalence of chirp-direction selectivity for responses to SCHR and aperiodic chirps. BS neurons were direction-selective significantly less often for SCHR chirps than for aperiodic chirps, suggesting that chirp sensitivity was suppressed for SCHR stimuli with F0s in the trough of BS MTFs. For all MTF-types, periodicity tuning influenced the overall rate responses to SCHR stimuli. Additionally, GLM analysis showed that more variance in SCHR data could be explained by periodicity than by velocity. This study also showed that individual neurons had a wide variety of rate-response profiles to aperiodic chirps, described using RVFs. RVFs revealed that neurons were more commonly direction-selective to chirps with velocities below 2 kHz/ms than above, and responded to chirps of velocities greater than 2 kHz/ms in a more direction-insensitive way.

A common use of Schroeder-stimuli has been as a masker in psychophysical experiments evaluating detection of an increment in tone level. Notably, there is an impact of Schroeder-phase on the effectiveness of SCHR stimuli as maskers, with up to 20-dB lower thresholds (less masking) for downward-SCHR maskers than for upward-SCHR maskers (Smith et al., 1986). An explanation from the literature attributes this effect to asymmetry of filtering in the cochlea: phase-dispersive properties of the basilar-membrane (BM) result in a peakier temporal response to downward than for upward SCHR stimuli. The peakier response is then compressed at the output of the cochlea, resulting in higher average energy in upward-SCHR BM responses compared to downward-SCHR responses, and thus more effective masking by upward SCHR stimuli (Recio and Rhode, 2000; Summers et al., 2003). This logic can also be applied to explain rate-differences between SCHR-chirp directions. However, the results here

suggest that other factors underlie the widespread chirp-direction selectivity of IC single-units observed in this study. First, the phase-dispersion of the BM depends on CF (Ruggero et al., 1997; Carney et al., 1999; Shera, 2000), with neural masking differences in chinchilla observed primarily for CFs > 3-4 kHz (Recio, 2001). In this study, no consistent differences were observed in either the prevalence or direction of chirp-direction selectivity between low-CF (< 3 kHz) and mid- or high-CF (> 3 kHz) neurons in aperiodic-chirp response rates (Fig. 2.10b). Secondly, the cochlear dispersion hypothesis can only explain selectivity towards upward chirps; despite this, neurons with CF > 3 kHz could be selective for either direction of chirp (Fig. 2.10b), suggesting that cochlear phase-dispersion was not the basis of IC chirp sensitivity. Similarly, Henry et al. (2023) found diverse SCHR selectivity in the IC of budgerigars (parakeet), though a statistically significant gradient was observed from upward SCHR selectivity at CFs below 2-3 kHz to downward SCHR selectivity for higher CFs. These findings are also consistent with a recent study in gerbils (Steenken et al., 2022) that reported that average rates of AN fibers could not explain behavioral discrimination of SCHR stimuli, whereas temporal analyses of AN responses over short (1-ms) time windows could explain behavioral thresholds. Low correlations of AN rate to behavior provide further evidence against cochlear phase-dispersion driving IC selectivity. On the other hand, the presence of information for SCHR-direction discrimination in the relatively fine timing of AN fibers is consistent with neural mechanisms that combine temporal information across frequency channels. Finally, a recent study has shown that phase curvature of the extreme apical region of guinea pig and gerbil cochleae is different from other regions (Recio-Spinoso et al., 2023), complicating the statement that downward chirps always have peakier BM responses. This unique phase curvature may

influence processing of chirps containing low-frequency components, such as the SCHR and aperiodic chirps used in this study.

A potential mechanism to explain the diverse chirp-response profiles seen in this study may be inhibitory inputs to the IC. IC neurons receive excitatory and inhibitory inputs with varying frequencies and latencies that may interact to produce chirp-direction selectivity (Pollak et al., 2011). Interactions between excitatory and inhibitory inputs varying in latency have long been implicated in mechanisms for frequency-sweep selectivity in bats (Suga, 1965; Fuzessery and Hall, 1996; Gordon and O'Neill, 1998). These mechanisms are broadly dependent on sideband inhibitory regions in excitatory tuning curves and the resulting asymmetry of responses to upward and downward frequency sweeps—this inhibition may arise from a variety of IC inputs. While not all response maps showed inhibitory sidebands, it is interesting to examine those that did within the framework of this theory. For instance, the neuron in Fig. 2.4 had an above-CF inhibitory sideband, with a longer duration and shorter latency than the excitatory band at higher sound levels. The location of this inhibitory sideband would predict a stronger response to upward chirps. Indeed, the RVF of this neuron is mostly biased towards upward chirps. The contribution of off-CF inhibition observed in response maps to interpreting chirp-direction sensitivity in the IC population studied here was precluded by the low-spontaneous rates of most IC neurons. To more comprehensively assess the effect of inhibition on chirp sensitivity, response maps using pure tones in background noise could be measured—this would be an interesting direction for a future study.

Another mechanism for chirp-direction selectivity was proposed by Rall (1969), based on dendritic filtering of several excitatory inputs organized spatially by frequency. In Rall's framework, a sweep that activates distal dendrites first, and proximal dendrites

later, would temporally sum and generate a large synaptic input. This mechanism has been hypothesized to underlie the frequency-sweep direction selectivity of octopus cells in the posteroventral cochlear nucleus (Godfrey et al., 1975; Rhode and Smith, 1986; Lu et al., 2022). Notably, octopus cells project to the ventral nucleus of the lateral lemniscus (VNLL), which provides inhibitory innervation to the IC (Adams, 1997; Vater et al., 1997; Winer et al., 1995). Thus, direction selectivity of octopus cells, by way of the VNLL, may explain chirp-direction selectivity in the IC; in this scenario, responses to chirps in one direction would be inhibited by input from the VNLL and disinhibited by chirps in the opposite direction. The possibility that IC chirp sensitivity originates in octopus cells is particularly intriguing because, like IC neurons, octopus cells display a diversity of chirp response profiles and can be selective to either upward or downward chirps (Lu et al., 2022). Note that Lu et al. (2022) suggested that the diverse chirp-direction selectivity in octopus cells could be explained by sequence detection, with response magnitude determined by a sequence of inputs with both frequency-dependent delays and varied amplitudes, enabled by the relatively long hyperpolarization of low-threshold potassium channels. Another possibility is that chirp-direction selectivity may originate in stellate neurons of the cochlear nucleus, which have been noted to display strong-upward SCHR selectivity, at least for higher CFs (Recio and Rhode, 2000; Recio, 2001). However, because IC neurons can be selective for either upward or downward chirps (Fig. 2.10), cochlear nucleus stellates cannot be the only source of chirp sensitivity. Finally, another possibility is that chirp-direction selectivity may arise at the level of the IC, from an inhibitory interneuron with inputs extending across iso-frequency laminae, such as IC stellate neurons (Oliver, 1984). The literature suggests that such neurons may also have low-threshold potassium channels (Sivaramakrishnan and Oliver, 2001)

and could therefore be selective to a sequence of frequency inputs in a manner similar to that proposed for octopus cells.

The aperiodic-chirp stimulus was composed of discrete, single-chirps, whereas the SCHR stimulus was a longer, sustained harmonic tone. It is possible that the presence of silent gaps might partially explain differences in responses to the two stimuli. In the GLM analysis, the combined GLM with both periodicity and velocity terms sometimes could not explain a high percentage of the variance (Fig. 2.9), suggesting that additional factors shaped SCHR response rates. Long-term rate modulation in response to a sustained stimulus may be one of these factors. This long-term effect may be attributable to the auditory efferent system, and cannot be captured in brief aperiodic-chirp responses. Similar long-term changes in rate have been reported in response to other complex sounds, representing a possible direction for future research (Farhadi et al., 2021). Furthermore, one could imagine a future study designing an aperiodic-chirp stimulus containing no silent gaps.

SCHR-chirp sensitivity in the IC suggests a mechanism that may influence midbrain-level encoding of speech. When speech sounds, or any animal vocalizations, are produced, resonances within the vocal tract act as second-order filters, introducing phase shifts between components of harmonic sounds (Klatt, 1980). In the same way that SCHR chirps are produced by the relative phases of harmonic components, vocalizations contain chirp-like features to which IC neurons may be similarly sensitive. IC neurons in this study were sensitive to SCHR chirps and directions over a frequency range relevant to human speech. Chirps extended over the frequency range of formants (200 – 3000 Hz for F1 and F2), and SCHR F0s were within the range of F0s in human speech (about 100 to 400 Hz) (Ladefoged and Johnson, 2014). Therefore, IC sensitivity

to chirp velocity may influence midbrain coding of speech, music, and other complex harmonic sounds. Also, responses to these sounds may be affected by the interaction of periodicity tuning and chirp-velocity sensitivity observed in this study.

Acknowledgements

We acknowledge the assistance of Kristina S. Abrams, Johanna B. Fritzinger, and Dr. Swapna Agarwalla for help with physiological experiments. Supported by NIH-DC001641 and NIH-F31DC019816.

Bibliography

Adams JC (1997) Projections from octopus cells of the posteroventral cochlear nucleus to the ventral nucleus of the lateral lemniscus in cat and human. Aud Neurosci 3:335-350.

Andoni S, Li N, Pollak GD (2007) Spectrotemporal Receptive Fields in the Inferior Colliculus Revealing Selectivity for Spectral Motion in Conspecific Vocalizations. The Journal of Neuroscience 27:4882-4893.

Carney LH, McDuffy MJ, Shekhter I (1999) Frequency glides in the impulse responses of auditory-nerve fibers. The Journal of the Acoustical Society of America 105:2384-2391.

Casseday JH, Ehrlich D, Covey E (1994) Neural Tuning for Sound Duration: Role of Inhibitory Mechanisms in the Inferior Colliculus. Science 264:847-850.

Davis KA (2005) Spectral Processing in the Inferior Colliculus. In: International Review of Neurobiology, pp 169-205: Academic Press.

Egan JP, Egan JP (1975) Signal detection theory and ROC-analysis: Academic press.

Escabi MA, Schreiner CE (2002) Nonlinear spectrotemporal sound analysis by neurons in the auditory midbrain. Journal of Neuroscience 22:4114-4131.

Escabi MA, Miller LM, Read HL, Schreiner CE (2003) Naturalistic auditory contrast improves spectrotemporal coding in the cat inferior colliculus. Journal of Neuroscience 23:11489-11504.

Farhadi A, Jennings SG, Strickland EA, Carney LH (2021) A Closed-Loop Gain-Control Feedback Model for The Medial Efferent System of The Descending Auditory Pathway. In: ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp 291-295.

Fuzessery ZM (1994) Response selectivity for multiple dimensions of frequency sweeps in the pallid bat inferior colliculus. Journal of Neurophysiology 72:1061-1079.

Fuzessery Z, Hall J (1996) Role of GABA in shaping frequency tuning and creating FM sweep selectivity in the inferior colliculus. Journal of neurophysiology 76:1059-1073.

Godfrey DA, Kiang NYS, Norris BE (1975) Single unit activity in the posteroventral cochlear nucleus of the cat. Journal of Comparative Neurology 162:247-268.

Gordon M, O'Neill WE (1998) Temporal processing across frequency channels by FM selective auditory neurons can account for FM rate selectivity. Hearing research 122:97-108.

Henry K, Wang Y, Abrams K, Carney LH (2022) Mechanisms of masking by Schroeder-phase harmonic tone complexes in the budgerigar. The Journal of the Acoustical Society of America 151:A122-A122.

Joris PX, Schreiner CE, Rees, A (2004) Neural processing of amplitude-modulated sounds. Physiological reviews *84*:541-577.
Kim DO, Carney L, Kuwada S (2020) Amplitude modulation transfer functions reveal opposing populations within both the inferior colliculus and medial geniculate body. Journal of Neurophysiology 124:1198-1215.

Klatt DH (1980) Software for a cascade/parallel formant synthesizer. The Journal of the Acoustical Society of America 67:971-995.

Krishna BS, Semple MN (2000) Auditory Temporal Processing: Responses to Sinusoidally Amplitude-Modulated Tones in the Inferior Colliculus. Journal of Neurophysiology 84:255-273.

Ladefoged P, Johnson K (2014) A course in phonetics: Cengage learning.

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

Liberman AM, Mattingly IG (1989) A Specialization for Speech Perception. Science 243:489-494.

Lu H-W, Smith PH, Joris PX (2022) Mammalian octopus cells are direction selective to frequency sweeps by excitatory synaptic sequence detection. Proceedings of the National Academy of Sciences 119:e2203748119.

Nelder JA, Wedderburn RW (1972) Generalized linear models. Journal of the Royal Statistical Society: Series A (General) 135:370-384.

Oliver DL, Morest DK (1984) The central nucleus of the inferior colliculus in the cat. Journal of Comparative Neurology 222:237-264.

Pollak GD, Xie R, Gittelman JX, Andoni S, Li N (2011) The dominance of inhibition in the inferior colliculus. Hearing Research 274:27-39.

Rall W (1969) Time constants and electrotonic length of membrane cylinders and neurons. Biophysical Journal 9:1483-1508.

Recio A (2001) Representation of harmonic complex stimuli in the ventral cochlear nucleus of the chinchilla. The Journal of the Acoustical Society of America 110:2024-2033.

Recio A, Rhode WS (2000) Basilar membrane responses to broadband stimuli. The Journal of the Acoustical Society of America 108:2281-2298.

Recio-Spinoso A, Dong W, Oghalai JS (2023) On the Tonotopy of the Low-Frequency Region of the Cochlea. The Journal of Neuroscience 43:5172-5179.

Rhode WS, Smith PH (1986) Encoding timing and intensity in the ventral cochlear nucleus of the cat. Journal of Neurophysiology 56:261-286.

Ruggero MA, Rich NC, Recio A, Narayan SS, Robles L (1997) Basilar-membrane responses to tones at the base of the chinchilla cochlea. The Journal of the Acoustical Society of America 101:2151-2163.

Schroeder M (1970) Synthesis of low-peak-factor signals and binary sequences with low autocorrelation (Corresp.). IEEE Transactions on Information Theory 16:85-89.

Schwarz DM, Zilany MSA, Skevington M, Huang NJ, Flynn BC, Carney LH (2012) Semi-supervised spike sorting using pattern matching and a scaled Mahalanobis distance metric. Journal of Neuroscience Methods 206:120-131.

Shera CA (2001) Frequency glides in click responses of the basilar membrane and auditory nerve: Their scaling behavior and origin in traveling-wave dispersion. The Journal of the Acoustical Society of America 109:2023-2034.

Sivaramakrishnan S, Oliver DL (2001) Distinct K Currents Result in Physiologically Distinct Cell Types in the Inferior Colliculus of the Rat. The Journal of Neuroscience 21:2861-2877.

Smith BK, Sieben UK, Kohlrausch A, Schroeder MR (1986) Phase effects in masking related to dispersion in the inner ear. The Journal of the Acoustical Society of America 80:1631-1637.

Steenken F, Oetjen H, Beutelmann R, Carney LH, Koeppl C, Klump GM (2022) Neural processing and perception of Schroeder-phase harmonic tone complexes in the gerbil: Relating single-unit neurophysiology to behavior. European Journal of Neuroscience 56:4060-4085.

Suga N (1965) Analysis of frequency-modulated sounds by auditory neurones of echo-locating bats. J Physiol 179:26-53.

Summers V, Boer Ed, Nuttall AL (2003) Basilar-membrane responses to multicomponent (Schroeder-phase) signals: Understanding intensity effects. The Journal of the Acoustical Society of America 114:294-306.

Vater M, Covey E, Casseday J (1997) The columnar region of the ventral nucleus of the lateral lemniscus in the big brown bat (Eptesicus fuscus): synaptic arrangements and structural correlates of feedforward inhibitory function. Cell and tissue research 289:223-233.

Whitehead ML, Lonsbury-Martin BL, Martin GK (1992) Evidence for two discrete sources of 2f1−f2 distortion-product otoacoustic emission in rabbit: I. Differential dependence on stimulus parameters. The Journal of the Acoustical Society of America 91:1587-1607.

Winer JA, Larue DT, Pollak GD (1995) GABA and glycine in the central auditory system of the mustache bat: structural substrates for inhibitory neuronal organization. Journal of Comparative Neurology 355:317-353.

Yin TCT, Smith PH, Joris PX (2019) Neural Mechanisms of Binaural Processing in the Auditory Brainstem. Compr Physiol 9:1503-1575.

Chapter 3: A Computational Model of Auditory Chirp-Velocity Sensitivity and Amplitude-Modulation Tuning in Inferior Colliculus Neurons

This chapter is published in the Journal of Computational Neuroscience.

3.1 Abstract

We demonstrate a model of chirp-velocity sensitivity in the inferior colliculus (IC) that retains the tuning to amplitude modulation (AM) that was established in earlier models. The mechanism of velocity sensitivity is sequence detection by octopus cells of the posteroventral cochlear nucleus, which have been proposed in physiological studies to respond preferentially to the order of arrival of cross-frequency inputs of different amplitudes. Model architecture is based on coincidence detection of a combination of excitatory and inhibitory inputs. Chirp-sensitivity of the IC output is largely controlled by the strength and timing of the chirp-sensitive octopus-cell inhibitory input. AM tuning is controlled by inhibition and excitation that are tuned to the same frequency. We present several example neurons that demonstrate the feasibility of the model in simulating realistic chirp-sensitivity and AM tuning for a wide range of characteristic frequencies. Additionally, we explore the systematic impact of varying parameters on model responses. The proposed model can be used to assess the contribution of IC chirp-velocity sensitivity to responses to complex sounds, such as speech.

3.2 Introduction

Natural sound stimuli, such as speech and music, are rich with spectral and temporal features to which auditory neurons are sensitive. The inferior colliculus (IC) has strong

rate tuning for complex sound features, such as amplitude modulations (AM). Recently, physiological studies have revealed that IC neurons have diverse sensitivity to the velocity of fast frequency sweeps, known as chirps, in both periodic (Steenken et al., 2022; Henry et al., 2023) and aperiodic (Mitchell et al., 2023) stimuli. The velocity of these chirps is much greater than that of more commonly considered sounds, such as formant transitions (Liberman and Mattingly, 1989). The majority of IC neurons are sensitive for chirp velocity, regardless of characteristic frequency (CF) or type of periodicity tuning (Mitchell et al., 2023). Computational models of the IC currently do not include chirp-velocity sensitivity. Velocity sensitivity arising in octopus cells of the cochlear nucleus (CN) (Lu et al., 2022), which inhibit the IC via the ventral nucleus of the lateral lemniscus (VNLL) (Adams, 1997; Vater et al., 1997), could potentially give rise to velocity sensitivity in the IC. Here, a computational model was used to test the hypothesis that a midbrain model with an inhibitory input from a velocity-sensitive octopus-cell can model the velocity sensitivity observed in IC neurons in addition to AM tuning.

Octopus cells are uniquely found in the posteroventral cochlear nucleus (PVCN) (Golding et al., 1999). These cells are excellent coincidence detectors with fine temporal resolution (Golding et al., 1995), entraining to individual periodic stimulus cycles up to 800 Hz while responding only to the onset of pure tones (Godfrey et al., 1975; Rhode and Smith, 1986). Additionally, octopus cells are distinguished by their wide dendritic fields, which extend across a range of auditory-nerve (AN) input frequencies (Osen, 1969). Frequency-dependent dendritic delays have been suggested to counteract latencies arising from the cochlear traveling wave and to thus improve coincidence detection in response to transient stimuli (Spencer et al., 2012). Sensitivity to the timing

of cross-frequency inputs with different amplitudes may also give rise to diverse sensitivity to chirp velocity (Lu et al., 2022), which is similar to that observed in IC neurons. Octopus cells have a broad range of CFs, as low as 0.2 kHz (Liberman, 1993). Finally, octopus cells are known to project to the contralateral VNLL, which in turn provides an inhibitory input to the IC (Adams, 1997; Vater et al., 1997). This fact, together with the broad range of responses of octopus cells to chirp velocities, makes them a potential source of chirp sensitivity in the IC.

Lu et al. (2022) posited that octopus cells function as sequence detectors, responding preferentially to dendritic inputs with different amplitudes that arrive in a certain temporal sequence. This mechanism depends on low-voltage-activated potassium (KL) channels, which are abundant in octopus cells (Bal and Oertel, 2001). Due to the slow recovery dynamics of the KL channels, dendritic inputs evoking both subthreshold and suprathreshold excitatory post-synaptic potentials (EPSPs) are followed by relatively long periods of hyperpolarization, preventing subsequent inputs from triggering action potentials. A suprathreshold input that normally evokes an action potential on its own will not do so when preceded by a subthreshold input. Therefore, a frequency sweep that triggers the suprathreshold EPSP before the subthreshold one will result in an action potential, whereas one that triggers the subthreshold EPSP before the suprathreshold one will not. This dependence upon the temporal sequence of inputs with different amplitudes, which are presumed to be tuned to different frequencies, was proposed to explain selectivity of octopus cells for chirp direction (Lu et al. 2022).

The modeling strategy used in this paper is rooted in work by Siebert (1965), who introduced a quantitative strategy to predict psychophysical performance as a function of stimulus parameters, using an analytical model for auditory-nerve responses. This

approach used statistical decision theory to define the limits of auditory discrimination based on a statistical description of neural responses. An important assumption for this approach is to treat neural responses as nonhomogeneous Poisson processes (NHPPs) (Rieke et al., 1999). Siebert's method has been employed for discrimination of tone frequency, level, and binaural cues, such as interaural time and level differences (Siebert, 1970; Colburn, 1973; Heinz et al., 2001a,b). This strategy was further developed by Krips and Furst (2009a,b), who demonstrated that Siebert's method can be extended into the central nervous system. Krips and Furst's (2009b) model cells are coincidence detectors (CDs) that receive multiple inputs and respond depending upon the relative timing of excitatory and/or inhibitory inputs. This general CD-based strategy is promising for modeling both the octopus cell's sequence-detection mechanism and, subsequently, the chirp sensitivity of IC neurons.

Here, we propose a model of IC chirp-sensitivity based on sequence detection of inhibitory octopus cells. First, we outline the model architecture within the Krips and Furst framework, describing the octopus-cell stage and then the IC stage. Then, we demonstrate the feasibility of the model in simulating IC neurons with physiologically plausible chirp sensitivity as well as AM tuning. Finally, we describe the parameters of the model and explore how parameter choice affected the sensitivity of the model cell. This work is a step towards addressing a gap in current computational models, which do not simulate sensitivity to chirps, despite its prevalence among IC neurons—over 90% are sensitive to chirp direction (Mitchell et al., 2023). A goal of this work is to simulate a diverse set of chirp-sensitive exemplar neurons using a single model architecture, rather than fitting the model to individual neuron responses.

In complex, harmonic sounds such as speech and music, phase differences between components give rise to chirps. Thus, inclusion of chirp sensitivity in computational models will improve predictions of responses to these realistic, perceptually important sounds.

3.3 Methods

Many modeling approaches exist for both octopus cells and IC neurons, including Hodgkin-Huxley models (octopus cells: Spencer et al., 2012; Manis and Campagnola, 2018; Lu et al., 2022; IC neurons: Cai et al., 1998), conductance-based models (octopus cells: Kalluri and Delgutte, 2003; Spencer et al., 2018; IC neurons: Hewitt and Meddis, 1994), and simpler phenomenological models (octopus cells: Rebhan and Liebold, 2021; IC neurons: Nelson and Carney, 2004). The model proposed here is based on work that extends statistical decision theory (Siebert, 1965, 1970; Colburn, 1973; Heinz et al., 2001a,b, 2002) to the central nervous system by generalizing auditory neurons as coincidence detectors that receive excitatory and/or inhibitory inputs (Krips and Furst, 2009a,b). This approach was selected for its flexibility, relatively low number of parameters, and ability to assign parameters to physiological correlates.

3.3.1. Model Architecture

3.3.1.1. Model Inputs

Krips and Furst (2009a,b) show that the output of a coincidence detector (CD) is a non-homogeneous Poisson process (NHPP) if it receives independent inputs that are NHPPs. This property of Krips and Furst's model CDs has the primary advantage of satisfying requirements for the use of statistical decision theory to estimate

psychophysical thresholds from model responses, namely that the statistics of discharge patterns are well-described and change as a function of the stimulus parameter of interest (Siebert, 1965; Heinz, 2001a,b). Additionally, Krips and Furst's method allows for the design of multi-stage model architectures that extend into the central nervous system, with NHPP statistics preserved at each stage.

The CD unit described by Krips and Furst (2009a,b) receives any number of independent inputs, either excitatory or inhibitory, each described by an instantaneous rate function, $\lambda(t)$. CD units can be defined by two basic interactions of inputs: excitatory-inhibitory (EI) and excitatory-excitatory (EE). A fundamental parameter of both EI and EE interactions is a temporal integration window, Δ. In an EI neuron, Δ describes the time window over which inhibition can suppress the response of the model neuron. In an EE neuron, Δ describes the time window within which excitation from multiple inputs facilitates the model response. To retain NHPP statistics in the output, Δ must be less than the refractory periods of the inputs.

The full model introduced here consisted of two distinct Krips-and-Furst CD models, an octopus cell and an IC cell (Fig. 3.1). Inputs to the octopus-cell stage were provided by a version of the Zilany et al. (2014) AN model that was modified to include gain control via the medial olivocochlear (MOC) efferent (Farhadi et al., 2023) and an improved approximation to the power-law synapse model (Guest and Carney, 2023). The efferent feedback in the AN model affected responses to sounds with modulated envelopes, including the aperiodic random chirp stimulus used here to characterize model neurons' chirp-velocity sensitivity. Additionally, inclusion of the MOC efferent pathways imparted more physiologically accurate responses to amplitude-modulated

noise over a wide dynamic range. The AN model always simulated high-spontaneous-rate fibers, which are the majority of AN fibers (Liberman, 1978). For simplicity, the excitatory input to the IC stage was also provided by a delayed AN response, representing direct inputs from the CN or those relayed through other brainstem nuclei.

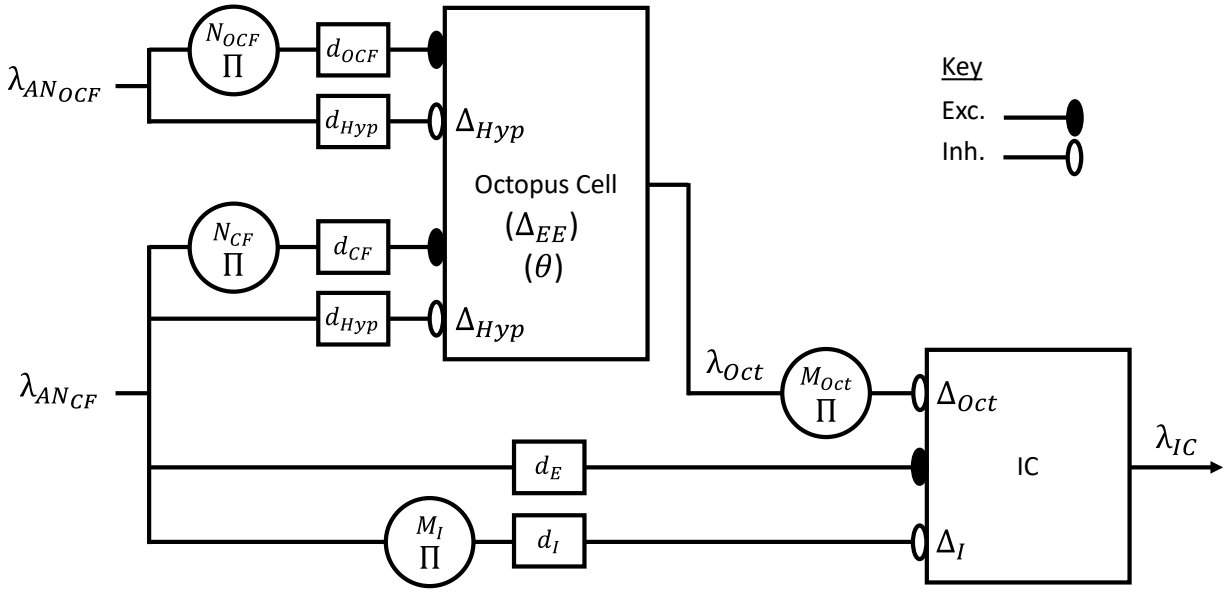

Figure 3.1 – Block Diagram of the model, showing excitatory and inhibitory inputs to both stages. AN input labels indicate whether they are CF or off-CF (OCF) with respect to the CF of the IC neuron. Rate functions are indicated by λ. AN fibers provide CF and off-CF (OCF) excitatory inputs to the octopus cell ($\lambda_{AN_{CF}}$ and $\lambda_{AN_{OCF}}$), delayed by d_{CF} and d_{OCF}, respectively. The numbers of excitatory inputs are N_{CF} and N_{OCF}. The excitatory inputs have an integration window Δ_{EE}. The AN fibers also provide an inhibitory input representing hyperpolarization of the cell due to opening of potassium channels, with delay d_{Hyp} and integration window Δ_{Hyp}. Finally, output of the octopus cell below a threshold θ was set equal to zero. The output of the octopus cell, λ_{Oct}, provides M_{Oct} inhibitory inputs to the IC stage, with integration window Δ_{Oct}. The IC stage also receives M_I on-CF inhibitory inputs ($\lambda_{AN_{CF}}$) with delay (d_I) and integration window Δ_I. Finally, the IC stage receives one CF excitatory input from $\lambda_{AN_{CF}}$, with delay (d_E).

The frequencies of the AN fiber inputs were defined as CF and off-CF (OCF), where CF was the desired characteristic frequency of the model IC cell. Note that whether the OCF frequency was above or below CF determined the direction of chirp selectivity. The following sections describe how the two model stages were portrayed

using the Krips and Furst (2009a,b) framework, with details provided for how parameter selection related to physiology.

3.3.1.2. Octopus-Cell Stage

The first stage of the model represented one aspect of octopus-cell responses, velocity sensitivity, based on the sequence detection theory posited by Lu et al. (2022). Sequence detection relies on the KL channels of octopus cells to provide hyperpolarization following either subthreshold or suprathreshold EPSPs. Excitatory inputs that arrive during KL hyperpolarizations did not produce action potentials. To mimic the time-course of hyperpolarization caused by KL channels, we used delayed inhibitory inputs to the octopus cell to represent KL hyperpolarizations. Note that these inputs do not represent the actual inhibitory inputs (from unknown sources) that have been described on octopus cell dendrites (Kreeger et al., 2024).

To implement sequence detection in its simplest configuration, two excitatory AN inputs were used, one subthreshold and one suprathreshold. The CF of the suprathreshold input matched the CF of the model IC cell. The frequency tuning of the off-CF subthreshold input (i.e., whether it was below or above CF) determined the direction of the chirp-velocity selectivity, as described below. The inputs had rate functions $\lambda_{AN_{CF}}$ and $\lambda_{AN_{OCF}}$, respectively. Additionally, the two inhibitory inputs representing KL hyperpolarization were delayed copies of the excitatory AN inputs. Note that since hyperpolarization always occurs after excitation, these inputs are not independent from one another (this issue will be further discussed below).

Direction-selectivity of the octopus cell was determined as follows: a chirp eliciting the suprathreshold CF input *before* the subthreshold OCF input resulted in an action

potential, because the suprathreshold excitation arrived before KL hyperpolarization could suppress the response. In contrast, a chirp of the opposite direction, eliciting the subthreshold OCF input *before* the suprathreshold CF input, resulted in suppression of the suprathreshold input by the KL hyperpolarization that followed the earlier subthreshold input, resulting in no response. In general, the cell was most responsive to stimuli that excited the suprathreshold (CF) input first. Thus, if the OCF input was tuned higher than IC CF, the octopus cell was selective for upward chirps. Alternatively, if the OCF input was tuned lower than IC CF, the octopus cell was selective for downward chirps (Lu et al., 2022). Figure 3.2 illustrates the sequence detection mechanism for two pairs of inputs with different CF ranges.

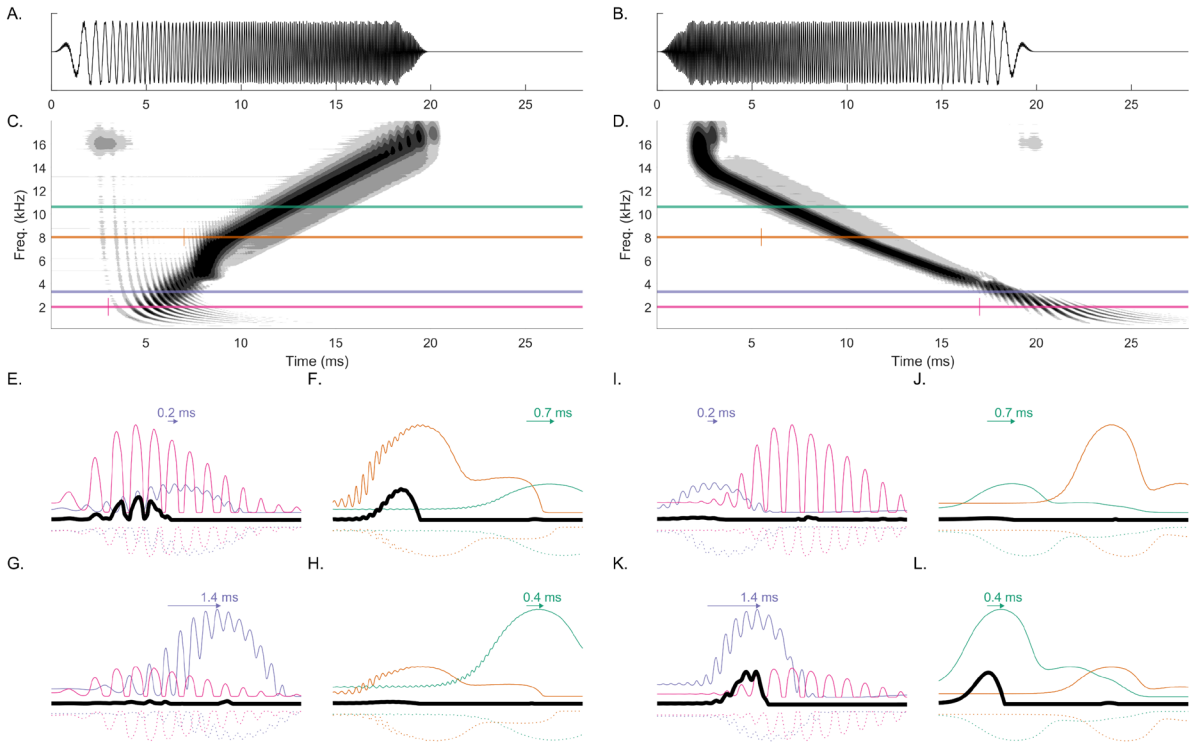

Figure 3.2 – Illustration of sequence-detection mechanism using example octopus-cell chirp responses. **A)** Upward chirp waveform (1.59 kHz/ms, 50 Hz to 16 kHz) **B)** Downward chirp waveform (-1.59 kHz/ms, 16 kHz to 50 Hz) **C)** Neurogram of AN model responses to example upward chirp. Gray shading (color bar) indicates rate function magnitude (lambda). Solid horizontal lines cut through the responses of individual fibers of different CFs (pink = 2 kHz, purple = 3.3 kHz, orange = 8 kHz, green = 10.6 kHz), corresponding to AN inputs to the example octopus-cell responses in E-L. Pink marker at 3 ms and orange marker at 7 ms mark the beginning of the plotted example responses in E-H. Note that the sound level used in this figure was 35 dB SPL (65 dB SPL used elsewhere) to simplify the shape of the neurogram by minimizing spread-of-excitation effects. **D)** Neurogram of AN model responses to example downward chirp. Pink marker at 17 ms and orange marker at 5.5 ms mark the beginning of the plotted example responses in I-L. **E-H)** Responses to an upward chirp of several example octopus cells (black), receiving suprathreshold (thick trace) and subthreshold (thin trace) inputs. Hyperpolarization traces are plotted as dotted lines. Labeled arrows depict delays applied to inputs with matching colors. **I-L)** Responses to a downward chirp of several example octopus cells, in the same format as E-H. Examples E and I receive a 2-kHz suprathreshold input (pink) and 3.3-kHz subthreshold input (purple), resulting in an octopus-cell model with upward chirp selectivity. Examples F and J receive an 8-kHz suprathreshold input (orange) and 10.6-kHz subthreshold input (green), resulting in upward chirp selectivity. Examples G and K receive a 3.3-kHz suprathreshold input (purple) and 2-kHz subthreshold input (pink), resulting in downward chirp selectivity. Examples H and L receive a 10.6-kHz suprathreshold input (green) and 8-kHz subthreshold input (orange), resulting in downward chirp selectivity.

To better temporally align AN responses with different CFs, which differ in latency, a delay parameter was imposed on both excitatory inputs, denoted as d_{CF} and d_{OCF}. Depending on the combination of input CFs, the delay was either applied only to the CF input ($d_{CF} > 0$ and $d_{OCF} = 0$) or only to the OCF input ($d_{OCF} > 0$ and $d_{CF} = 0$). A delay value that ensured the desired chirp-direction sensitivity across all chirp velocities was determined through parameter optimization (described in detail below). The function of the delays was to ensure two things: one, in response to chirps of the selected-for direction, suprathreshold CF input arrived sufficiently *before* the subthreshold OCF input to avoid suppression by OCF hyperpolarization, and two, in response to chirps of the opposite direction, suprathreshold CF input arrived sufficiently *after* the subthreshold OCF input for maximal suppressed by OCF hyperpolarization. In Figure 3.2E-L, arrows above the traces indicate delays that maximized desired direction-selectivity.

The sequence-detection mechanism was implemented using the framework in Krips & Furst (2009a,b), as follows: Consider only the excitatory inputs of the octopus cell ($\lambda_{AN_{CF}}$ and $\lambda_{AN_{OCF}}$). Let $N = N_{CF} + N_{OCF}$, be the number of excitatory AN inputs to the octopus cell, where N_{CF} represents the number of identical copies of $\lambda_{AN_{CF}}$, and N_{OCF} represents the number of identical copies of $\lambda_{AN_{OCF}}$. The cell responded when at least L inputs are active during an interval Δ. For the purposes of this sequence-detector model, $N_{CF} = L$ and $N_{OCF} = 1$, where $L > 1$ and Δ is Δ_{EE} (Fig. 3.1); therefore, the model cell responded only if activity occurred on at least $N - 1$ inputs during an interval Δ_{EE}.

To ensure the sequence-detection mechanism, the suprathreshold input, $\lambda_{AN_{CF}}$, must trigger an action potential in isolation, while $\lambda_{Oct_{OCF}}$, the subthreshold input, must not. Thus, the octopus cell can be thought of as a multiple-input EE cell that responds

either when receiving exactly N active inputs (all inputs, CF and OCF) or when receiving exactly $N-1$ active inputs (all CF inputs, but not the OCF input). Let us consider the first case, when the cell receives $L=N$ active inputs. The set of all inputs is $\{\lambda_1,\dots,\lambda_N\}$. The instantaneous rate is described by (Eqn. 4.21 in Krips and Furst, 2009b):

$$\lambda_{EE_L}(t)=\sum_{l=1}^{L}\lambda_l(t)\prod_{j=1,j\neq l}^{L}\int_{t-\Delta}^{t}\lambda_j(t')dt' \quad .$$

The cell responds at time t only when activity is observed on all $L=N$ inputs within the time interval $(t-\Delta,t)$. Note that here, $\Delta=\Delta_{EE}$.

Now consider the second case, when the cell receives exactly $N-1$ active inputs. Let l be the exact number of active inputs. If $L\leq l\leq N$, there are $\binom{N}{l}$ (N-choose-l) sets of active inputs possible. Given $1\leq i\leq\binom{N}{l}$, the ith set of active inputs is denoted as $\Psi_{l_i}=\{\lambda_1{}^{(i)},\dots,\lambda_l{}^{(i)}\}$. The complementary set of $N-l$ inactive inputs is denoted as $\Omega_{l_i}=\{\lambda_{l+1}{}^{(i)},\dots,\lambda_N{}^{(i)}\}$. Therefore, the instantaneous rate of a cell that responds when exactly l inputs are active, and $N-l$ inputs are not active, is described by (Eqn. 4.23 in Krips and Furst, 2009b):

$$\lambda_{EE_{=l}^{N}}(\Psi)=\sum_{i=1}^{\binom{N}{l}}\lambda_{EE_l}(\Psi_{l_i})\lambda_I(\Omega_{l_i}) \ ,$$

where $\lambda_{EE_l}(\Psi_{l_i})$ is the instantaneous rate of the EE cell receiving the set of active inputs Ψ_{l_i}, given by Eqn. 4.21 in Krips and Furst (2009b). Meanwhile, $\lambda_I(\Omega_{l_i})$ is the instantaneous rate of the set of inactive inputs Ω_{l_i}, described by (Eqn. 4.24 in Krips and Furst, 2009b):

$$\lambda_I(\Omega_{l_i}) = \prod_{j=l+1}^{N}\left(1 - \int_{t-\Delta}^{t} \lambda_{j^{(i)}}(t')dt'\right).$$

.

The octopus cell responds if $l = N$ or $l = N - 1$. The instantaneous rate of each of these two cases is found by substituting l into Eqn. 4.23 in Krips and Furst (2009b). Summing these two functions gives the final instantaneous rate of the multiple-input EE cell, $\lambda_{EE_L^N}(\Psi)$.

Next, the effect of the KL hyperpolarization inputs can be considered. These inhibitory inputs are copies of the excitatory inputs, $\lambda_{AN_{CF}}$ and $\lambda_{AN_{OCF}}$, delayed by d_{Hyp} seconds. Letting Ψ_{AN} be the set of all AN inputs (where again, $N = N_{CF} + N_{OCF}$), the full equation for the instantaneous rate of the octopus cell stage is described by (Eqn. 1):

$$\lambda_{Oct}(t) = \lambda_{EE_L^N(\Psi_{AN})} \cdot \left(1 - \int_{t-\Delta_{Hyp}}^{t} \lambda_{AN_{CF}}(t')dt'\right) \cdot \left(1 - \int_{t-\Delta_{Hyp}}^{t} \lambda_{AN_{OCF}}(t')dt'\right). \qquad (Eqn\ 1)$$

Finally, to ensure that the octopus cell has an “ideal onset” quality (Godfrey et al., 1975; Rhode and Smith, 1986; Oertel et al., 2000), a threshold θ was applied to λ_{Oct}, that is, samples of $\lambda_{Oct}(t)$ below θ were set to zero. The value of θ was chosen by observing rate functions of the octopus cell stage in response to click trains. The value θ = 50 spk/s eliminated activity between click cycles.

3.3.1.3. IC Stage

The second stage of the model represented a neuron in the IC that received excitatory input from the brainstem and inhibition from the octopus-cell via the VNLL, which gave it chirp-direction sensitivity. The IC model neuron also received a delayed

inhibitory input with the same CF as the excitatory input, as in the same-frequency inhibition and excitation (SFIE) model for AM tuning (Nelson and Carney, 2004).

The AM tuning of a neuron is characterized by a modulation transfer function (MTF), the average response rate versus modulation frequency. The SFIE model produces neurons with band-enhanced (BE) MTFs, which are characterized by increased excitation, with respect to unmodulated responses, over a band of modulation frequencies (Kim et al., 2020). For the IC stage here, the brainstem was not explicitly modeled, for simplicity, and the excitatory brainstem input was represented by a version of $\lambda_{AN_{CF}}$ that was delayed by d_E. The corresponding inhibition was also represented by a copy of $\lambda_{AN_{CF}}$ that was delayed by d_I. This inhibition had the associated parameters Δ_I, the integration window, and M_I, describing the number of times the inhibitory input was duplicated. To ensure the octopus-cell inhibition arrived before the excitatory input, the value of d_E was greater than 0. Additionally, to ensure the same-frequency inhibition arrived after the excitatory input, the inhibitory delay, d_I, was greater than the excitatory delay, d_E.

The inhibition from the octopus-cell stage, λ_{Oct}, had its own set of parameters: d_{Oct}, Δ_{Oct}, and M_{Oct}, for the delay, integration window, and number of inhibitory inputs, respectively. The instantaneous rate at the IC stage output was defined as

$$\lambda_{IC}(t) = \lambda_{AN_{CF}}(t) \cdot \left(1 - \int_{t-\Delta_I}^{t} \lambda_{AN_{CF}}(t')dt'\right)^{M_I} \cdot \left(1 - \int_{t-\Delta_{Oct}}^{t} \lambda_{Oct}(t')dt'\right)^{M_{Oct}} . \quad (Eqn\ 2)$$

Additionally, the final output λ_{IC} was half-wave rectified to prevent negative rates. As illustrated below, the model IC cell was sensitive to chirp direction and velocity and had BE AM tuning.

3.3.2. Stimuli

To validate the IC-model response characteristics, the following set of stimuli, similar to those used in the experimental study of Mitchell et al. (2023), were presented to model cells. Responses to pure tones at different levels and frequencies were used to generate response maps (RMs) and assess frequency tuning. Sinusoidally amplitude-modulated (SAM) noise was used to generate modulation transfer functions (MTFs), used to evaluate tuning to modulation frequency. Aperiodic chirp stimuli were used to generate rate-velocity functions (RVFs), which characterize sensitivity to direction and velocity of chirps. Additionally, click-train stimuli were used as an alternate method of generating MTFs, for the purpose of comparing octopus-cell model responses to physiological responses (Godfrey et al., 1975). Unless stated otherwise, all model response rates were determined by calculating the integral of the model rate function over the stimulus duration. Rate functions were constructed using the mean of responses to five stimulus repetitions. The input signal for each repetition was the mean of 10 statistically independent high-spontaneous-rate (HSR) AN-model responses for each frequency channel. The number of AN fibers per channel was chosen to align with the approximate number of HSR AN fibers that innervate each inner hair cell in the cochlea (Keithley and Schreiber, 1987).

RMs were produced using a series of 200-ms-duration tones with frequencies ranging from 250 Hz – 10 kHz, at 10, 30, 50, and 70 dB SPL. Tones had 10-ms raised-cosine ramps.

Responses to SAM noise were used to generate rate MTFs. Noise was 100% modulated over a range of modulation frequencies from 2 – 500 Hz. The noiseband

spanned 100 Hz – 10 kHz, had a spectrum level of 30 dB SPL (overall level of 70 dB SPL), and duration of 1000 s (including 50-ms raised-cosine ramps). MTFs were classified based on the rates in response to modulated relative to unmodulated stimuli. Here, the model was designed to produce BE MTF shapes.

An aperiodic chirp stimulus, introduced in Mitchell et al. (2023), was designed to characterize neural sensitivity to direction and velocity of fast frequency chirps using RVFs, defined as average rate versus the velocity of a linear frequency sweep. This stimulus is derived from the Schroeder-phase harmonic complex (Schroeder, 1970). To construct the aperiodic chirp stimulus, fundamental periods were extracted from a set of Schroeder-phase stimuli, with each period being equivalent to a linear frequency chirp. The set of chirp velocities used was identical to those in Mitchell et al. (2023): ±0.40, ±0.80, ±1.59, ±3.16, ±6.24, and ±9.24 kHz/ms. Note that these velocities are matched to those of Schroeder-phase harmonic complexes with fundamental frequencies of 25, 50, 100, 200, 400, and 600 Hz, respectively, and with the highest harmonic frequency at 16 kHz. A random sequence of chirps was generated, with each combination of direction and velocity presented a total of 42 times. To avoid periodicity, random spacing (40 – 60 ms) was introduced between chirp offsets and onsets. Raised-cosine ramps with durations equal to 10% of chirp duration were applied to each chirp. The sound level of each chirp was set to $65\ \mathrm{dB\ SPL} - 10 \times \log_{10}(T/T_{ref})$, where T is the duration of the chirp, and T_{ref} = 2.5 ms (the duration of the ±6.24 kHz/ms chirp). This scaling ensured that energy was normalized among chirps of different durations. To construct the RVF, response rate was calculated by summing spikes over a 15-ms window centered at the peak of the neural response.

For click MTFs, click trains were generated with methods adapted from Godfrey et al. (1975). Rarefaction clicks, 0.1 ms in duration, were generated with rates from 2 – 500 Hz. Click level was approximately 130 dB peSPL, to match the stimuli described in Godfrey et al. (1975). Click MTFs were generated in the same manner as noise MTFs, with rate expressed as a function of click rate. To illustrate entrainment, a special rate calculation was performed to generate click MTFs. Instead of integrating the rate function over the response duration, the number of threshold crossings in the rate function was counted. To ensure each response was only counted once, a refractory period of 1 ms was included. For this calculation only, a threshold of 110 spikes/s was manually selected based on examination of the click-evoked rate functions (this threshold is distinct from the threshold θ applied to octopus cell output, described below).

3.3.3. Model Parameter Selection

The chirp-velocity sensitivity of the model IC cell was dependent upon the sensitivity of the octopus-cell inhibitory input. Therefore, the selection of octopus-cell parameters was important for generating model IC cells with physiologically appropriate chirp responses. Sensitivity towards chirp velocity and direction can be characterized by RVFs. While octopus-cells have heterogeneous chirp sensitivity (Lu et al., 2022), and thus would have a variety of RVF shapes, for the purpose of this study it was useful to consider two basic types, one selective for upward chirps and one for downward chirps.

Octopus-cell parameters were determined using the MATLAB parameter-optimization tool *fmincon* (2022a, MathWorks). This tool is designed to determine the parameters that minimize the output of a loss function. Here, the loss function was $1 -$

$corr(RVF_{mod}, RVF_{tem})$, where RVF_{mod} was the RVF of the model octopus cell, RVF_{tem} was the template RVF, and $corr$ was the linear correlation operation (note that this loss function was identical to maximizing the correlation between model and template RVFs). Two template RVFs were used: the upward-selective template had rates of one for positive velocities and rates of zero for negative velocities; the downward-selective template had rates of one for negative velocities and zero for positive velocities. These two simple RVF shapes were chosen to impart the most basic direction selectivity upon the octopus-cell RVFs.

Octopus-cell parameters yielding upward-selective and downward-selective RVFs were found for CFs of 1, 4, and 8 kHz (representing low, medium, and high IC CFs). To ease optimization, the octopus-cell parameter space was simplified to two free parameters: OCF, the frequency of the off-CF input, and a single delay that was applied to the higher-CF of the two AN inputs, either d_{CF} or d_{OCF}. Initial parameter values were randomly selected within each parameter's lower and upper bounds (Table 3.1). Note that the bounds for OCF depended on the desired direction-selectivity of the octopus cell. From here, *fmincon* optimized the free parameters that minimized the objective function and resulted in an RVF that most resembled the template RVF. Note that the optimal delay parameter yields the highest direction bias for the most velocity pairs, but this is not necessarily optimal for all velocities individually. Thus, the magnitude of direction bias in the RVF varies for different velocities.

Table 3.1 – Octopus-cell stage parameters and their respective values or ranges.

Parameter Name	Value (or range)
N_{CF}	3 or 4
N_{OCF}	1
OCF	CF/3–3CF Hz
d_{CF} OR d_{OCF}	0–2 ms
Δ_{EE}	1 ms
Δ_{Hyp}	2 ms
d_{Hyp}	0.4 ms
θ	50 spikes/s

For the remaining octopus-cell parameters, N_{CF}, Δ_{EE}, Δ_{Hyp}, and d_{Hyp}, a range of values was explored to optimize the chirp-sensitivity of octopus-cell RVFs. This exploration is summarized in Results (Figs. 3.7-10), and the default values for each parameter are given in Table 3.1.

Finally, IC-stage parameters d_{Oct}, M_{Oct}, d_I, and M_I, were manually selected to match the desired response properties, i.e., an IC cell receiving upward-selective octopus-cell inhibition had parameters to maximize downward-selectivity in its RVF. Parameters of all IC cells were selected to yield BE MTFs. IC velocity-sensitivity was primarily affected by octopus-cell inhibition parameters (d_{Oct}, M_{Oct}), and periodicity tuning was primarily affected by SFIE inhibition parameters (d_I, M_I). The impact of varying these IC-stage parameters is summarized in Figs. 3.11-14. For simplicity, Δ_{Oct} and Δ_I were both set equal to 1 ms.

3.4 Results

3.4.1. Octopus-Cell Stage Responses

Tones and click stimuli were used to confirm that response properties of the octopus-cell stage were consistent with physiological recordings (Godfrey et al., 1975;

Rhode and Smith, 1986). The responses of an upward-sensitive octopus cell (λ_{Oct}) with CF = 4 kHz illustrate a rate function with a strong onset response to a pure tone at CF, followed by rates near zero (Fig. 3.3A).

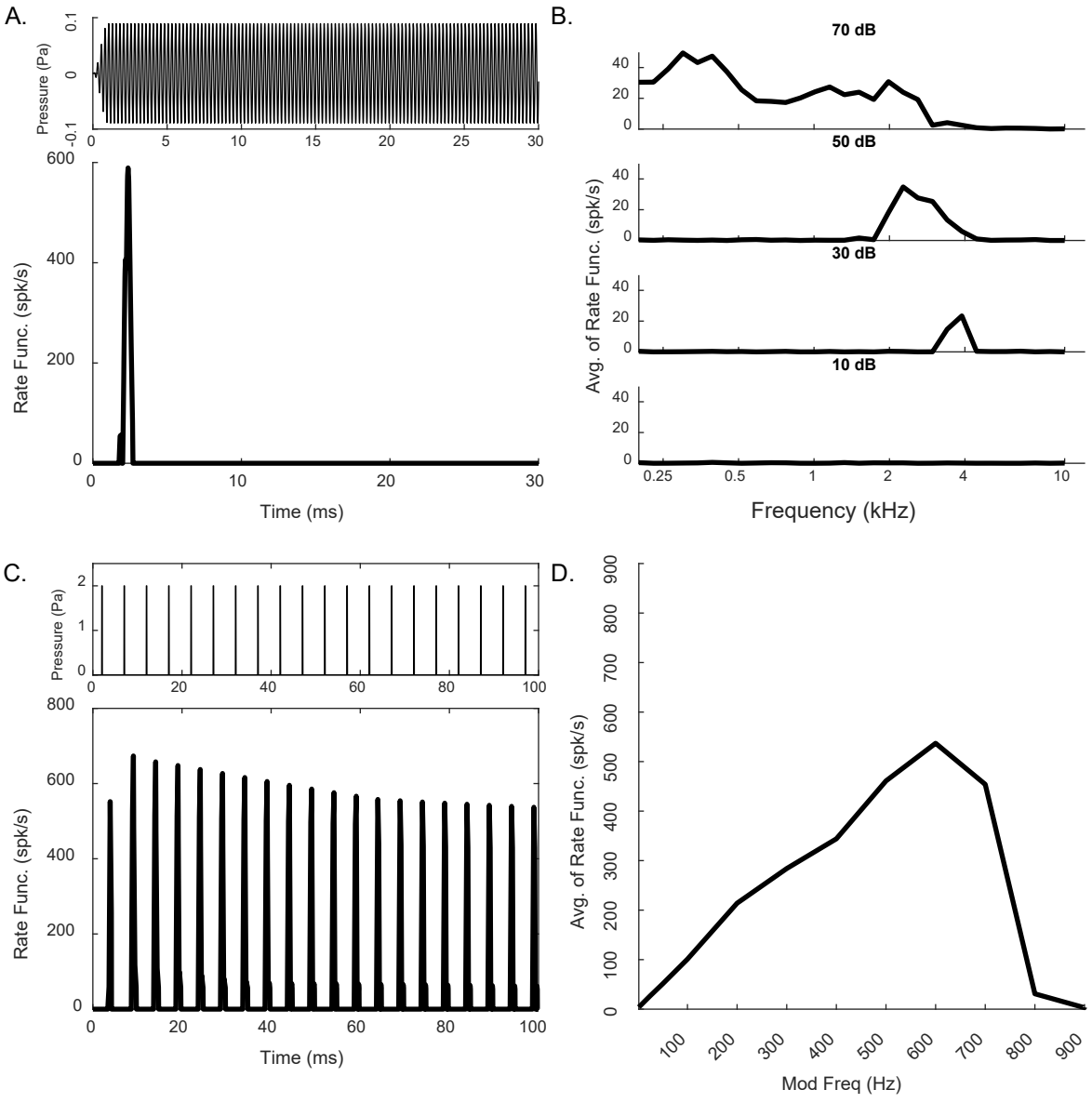

Figure 3.3 – Responses of an example model octopus cell to tone and click stimuli. A) Rate function (λ_{Oct}) in response to a pure tone at CF, B) Pure-tone response map, illustrating CF at 4 kHz, C) Rate function in response to a click train (200 Hz, peak level 130.6 dB peSPL), and D) Click MTF for 2–900 Hz. For this panel only, to illustrate

entrainment, threshold-crossings of the rate-function (C) were counted to approximate action potentials (threshold = 110 spk/s).

The pure-tone RM (Fig. 3.3B) reflects the CF of 4 kHz, with a broader frequency-response at 50 and 70 dB SPL. In response to a click train, a peak in the rate function was observed for every click; the amplitude of the rate function increased at the beginning of the response but leveled off with time (Fig. 3.3C). With increasing click rate, the response rate entrained until 600 Hz, and stopped responding at 900 Hz (Fig. 3.3D).

3.4.2. Example Neurons

Example neurons with both upward and downward chirp-direction sensitivity were produced with low, medium, and high CFs. Parameter values for example neurons are provided in Table 3.2. Note that these parameters were manually selected to result in example model neurons with substantial chirp sensitivity and AM tuning that was representative of IC recordings (Mitchell et al., 2023). Optimal fitting of this large set of parameters to actual neural responses may be possible but is beyond the scope of this study.

Table 3.2 – Parameter values for example neurons

CF (kHz)	IC direction	OCF (kHz)	d_{CF} (ms)	d_{OCF} (ms)	N_{CF}	M_{Oct}	M_I	d_E (ms)	d_I (ms)	Δ_{Oct} (ms)	Δ_I (ms)
1	Down	2.21	0	0.45	4	12	8	1	3.5	1	1
1	Up	0.90	1.00	0	4	12	16	1.4	3.4	1	1
4	Down	5.33	0.30	0	3	6	16	0.5	3.0	1	1
4	Up	2.24	1.20	0	3	6	8	1.5	4.0	1	1
8	Down	10.69	0.40	0	3	3	8	1.2	3.7	1	1
8	Up	5.51	0.90	0	3	12	8	1.5	4.5	1	1

For model IC neurons with low CF (CF = 1 kHz) (Fig. 3.4), chirp-direction sensitivity was less prominent than for higher-CF neurons. Chirp-direction sensitivity is observed in the RVF plots by comparing the rates in response to positive and negative

chirp directions at each chirp speed. In the plots below, vertical dashed lines at ±1.59 and ±6.24 kHz/ms have been included for ease of comparison. The downward-sensitive IC neuron received upward-sensitive octopus-cell inhibition (Fig. 3.4A). The model octopus cell was upward-sensitive across all velocities in the RVF, whereas the model IC cell was downward-sensitive for chirps below ±3.16 kHz/ms, and was not direction sensitive at higher speeds (Fig. 3.4B). The MTF of the IC stage was BE (Fig. 3.4C).

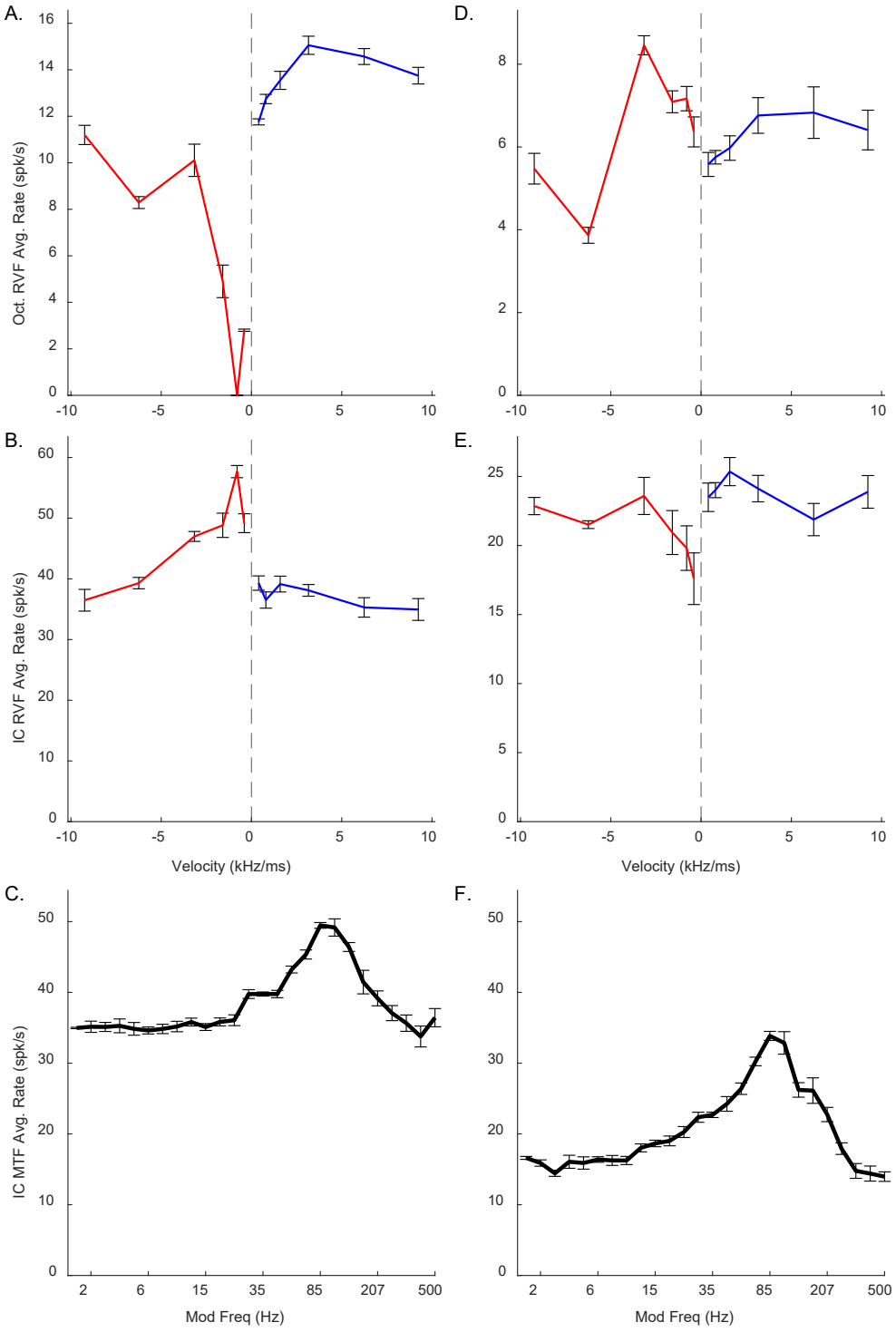

Figure 3.1 – Example low-CF (1 kHz) octopus-cell (A, D) and IC (B,C,E,F) model responses. A-C correspond to upward-sensitive octopus and downward-sensitive IC

models, D-E correspond to downward-sensitive octopus and upward-sensitive IC models. In RVFs, blue indicates response to upward velocities, red indicates response to downward velocities. Error bars indicate standard deviation of 5 model repetitions.

The upward-sensitive model IC neuron received downward-sensitive octopus-cell inhibition (Fig. 3.4D). The RVF of the octopus-cell inhibition was downward-sensitive, for velocities below ±6.24 kHz/ms; the RVF of the IC stage was similarly upward-sensitive for velocities below ±6.24 kHz/ms (Fig. 3.4E). The MTF of the IC stage was BE (Fig. 3.4F). Notably, for both low-CF examples, N_{CF} was set to 4, resulting in stronger chirp-sensitivity (N_{CF} was set to 3 for model neurons with higher CFs, below). The error bars indicate the standard deviation of 5 model trials.

Chirp-sensitive medium-CF (CF = 4 kHz) model neurons (Fig. 3.5) had MTFs with more prominent peaks and RVFs with larger rate-differences between directions than the low-CF neurons (Fig. 3.4). The RVF of the upward-sensitive medium-CF octopus cell had large rate-differences for all velocity pairs (Fig. 3.5A). The rate-differences in the downward-sensitive IC RVF are also large (Fig. 3.5B). The IC MTF (Fig. 3.5C) was BE, with a well-defined peak at about 100 Hz. In contrast to the upward-sensitive octopus cell (Fig. 3.5A), the downward-sensitive medium-CF octopus-cell (Fig. 3.5D) was upward-sensitive at low chirp speeds (<3.16 kHz/ms), but not at high speeds (>6.24 kHz/ms). Similarly, the IC RVF (Fig. 3.5E) is downward-sensitive for low chirp speeds, and upward-sensitive at high speeds. The IC MTF is BE, with a BMF of about 100 Hz (Fig. 3.5 F). For both medium-CF model cells, direction-sensitivity was strongest for velocities below ±3.16 kHz/ms.

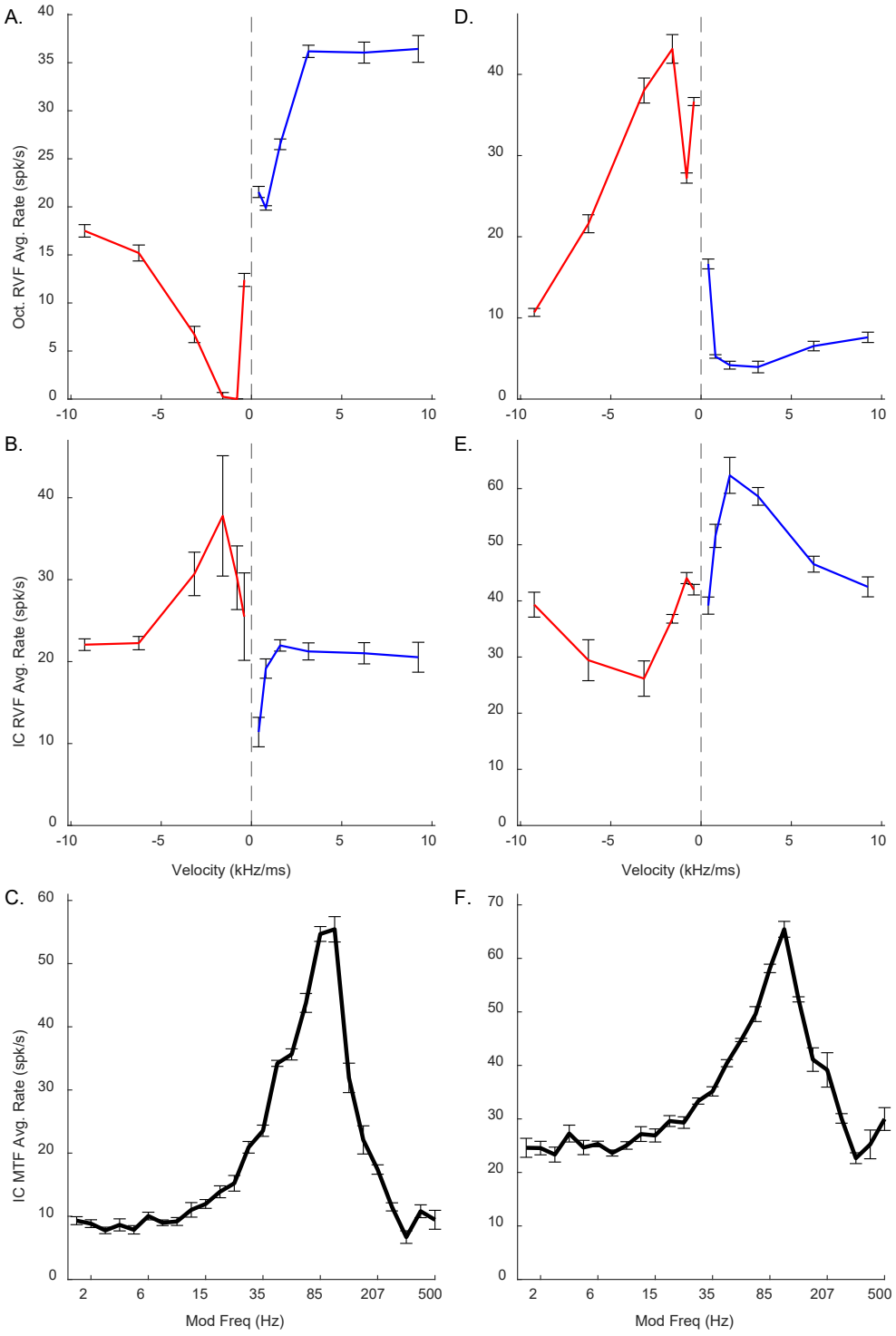

Figure 3.5 – Example medium-CF (4 kHz) octopus-cell (A, D) and IC (B,C,E,F) model responses. A-C correspond to upward-sensitive octopus and downward-sensitive IC

models, D-E correspond to downward-sensitive octopus and upward-sensitive IC models. In RVFs, blue indicates response to upward velocities, red indicates response to downward velocities. Error bars indicate standard deviation of 5 model repetitions.

Chirp-sensitive high-CF (CF = 8 kHz) IC neurons (Fig. 3.6) had slightly smaller rate-differences between directions than CF = 4 kHz (Fig. 3.5). The RVF of the octopus-cell (Fig. 3.6A) that inhibited the downward-sensitive IC neuron was upward-sensitive at all velocities, with response rate peaking at +0.80 kHz/ms. The corresponding IC RVF (Fig. 3.6B) is downward-sensitive at all velocities, but with smaller rate differences than the medium-CF example (Fig. 3.5B). The IC MTF is BE (Fig. 3.6D), with BMF around 40 Hz. Finally, the downward-sensitive octopus cell has strong chirp-direction sensitivity for all velocities (Fig. 3.6D). The IC RVF (Fig. 3.6E) is upward-sensitive, except for ±0.80 kHz/ms, for which it is slightly downward-sensitive. Also notable is the large rate-difference for ±0.40 kHz/ms, despite the comparatively smaller rate difference in the octopus cell RVF (Fig. 3.6E). The IC MTF is BE (Fig. 3.6F), with a BMF of approximately 40 Hz.

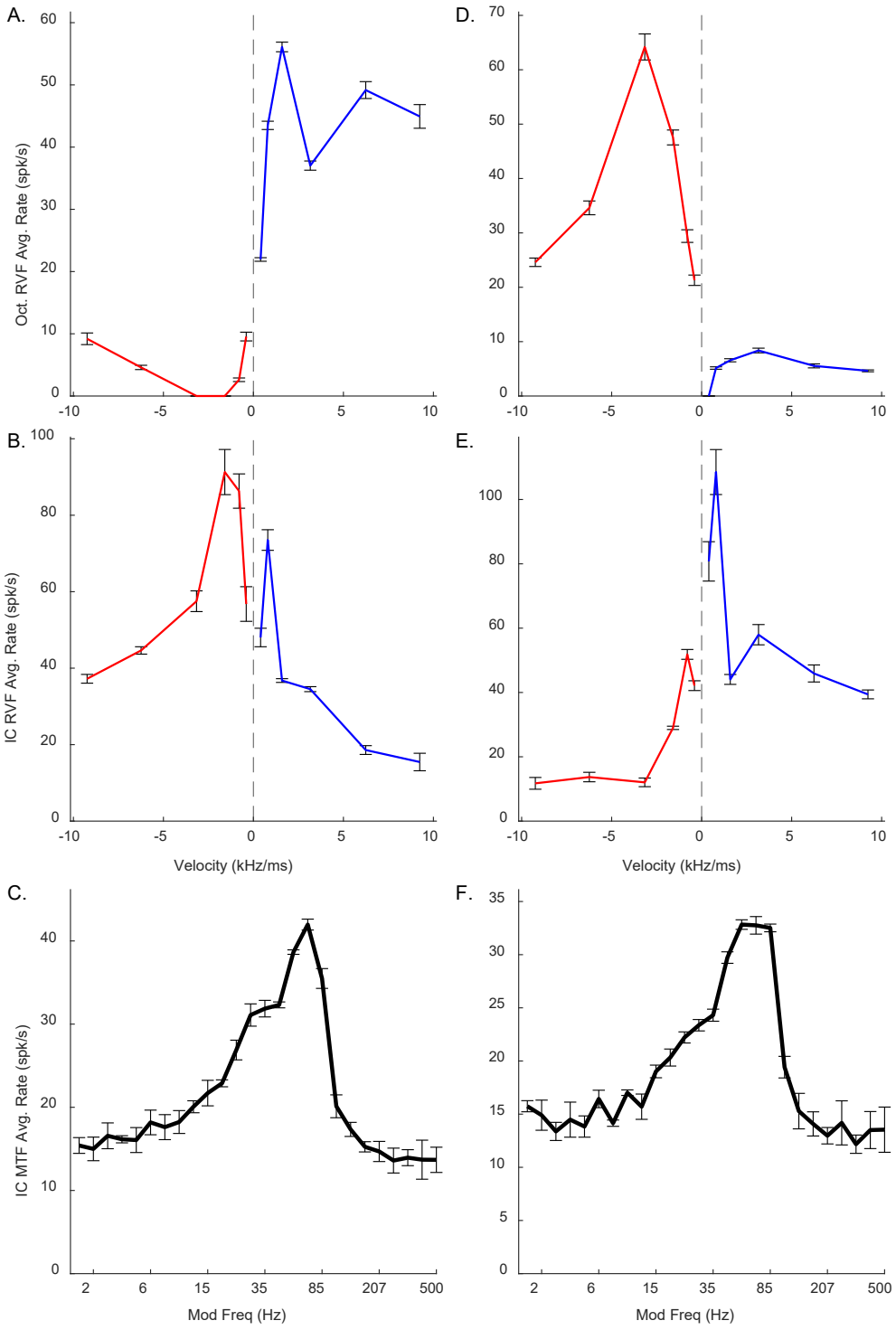

Figure 3.6 – Example high-CF (8 kHz) octopus-cell (A, D) and IC (B,C,E,F) model responses. A-C correspond to upward-sensitive octopus and downward-sensitive IC

models, D-E correspond to downward-sensitive octopus and upward-sensitive IC models. In RVFs, blue indicates response to upward velocities, red indicates response to downward velocities. Error bars indicate standard deviation of 5 model repetitions.

3.4.3. Effect of Varying Parameters

Parameters for example cells were selected with the goal of maximizing chirp-direction sensitivity and BE MTF tuning. The following section illustrates the contribution of each parameter to the model responses to chirps and AM noise. In each case, all parameters except one were held constant to the values listed in Table 3.2.

N_{CF} represented the number of identical on-CF excitatory inputs arriving to the octopus cell, and $N_{CF} = L$, the number of active inputs required for the octopus cell to respond. For the low-CF, upward-sensitive, octopus-cell example, decreasing N_{CF} from 4 to 2 resulted in increased rates in response to all stimuli (Fig. 3.7A), ultimately making the model IC neuron less downward-sensitive (Fig. 3.7B), and reducing the amplitude of the peak in the BE MTF (Fig. 3.7C). Conversely, increasing N_{CF} from 4 to 6 reduced the octopus-cell response rates (Fig. 3.7A), resulting in less octopus-cell inhibition to the IC cell, and again a less downward-chirp sensitive RVF (Fig. 3.7B). It was apparent that there was an optimal value for N_{CF} that maximized the desired chirp-direction sensitivity. For mid-and-high-CF neurons, this value was 3, whereas for low-CF neurons, N_{CF} = 4 resulted in the strongest chirp-direction sensitivity.

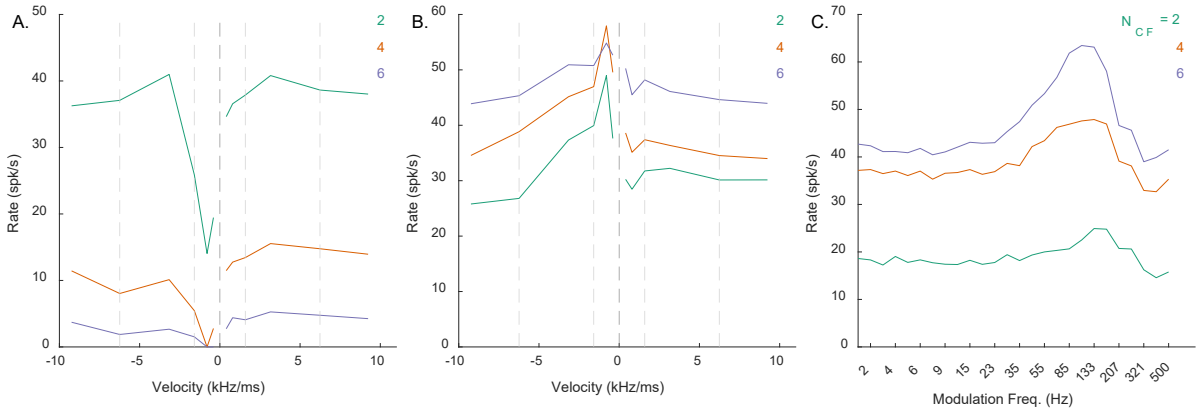

Figure 3.7 – Impact of varying N_{CF} on responses of an example low-CF (1 kHz) model cell with downward-sensitive IC output (Green – N_{CF} = 2, orange – N_{CF} = 4, purple – N_{CF} = 6). Responses for N_{CF} = 4 are also shown in Figs. 3.4A-C. A) octopus cell RVF; B) IC RVF; C) IC MTF.

In general, Δ described the integration windows of CDs. Per Krips and Furst (2009b), this value should be smaller than the refractory period of the neuron. However, the integration window for the EE inputs to the octopus cell (Δ_{EE}) had to be 1 ms or greater to capture the desired chirp-direction sensitivity in either the octopus RVF (Fig. 3.8A, purple line) or the IC RVF (Fig. 3.8B, purple line); smaller integration windows resulted in less sensitive neurons (Fig. 3.8B, orange and green lines). Additionally, the integration window of the hyperpolarization inhibition (Δ_{Hyp}) had to be relatively long to adequately suppress the excitatory signals (Fig. 3.9). Values of Δ_{Hyp} less than approximately 2 ms resulted in octopus-cell RVFs that were not direction sensitive (Fig. 3.9A, green and orange), and ultimately non-sensitive IC RVFs (Fig. 3.9B, green and orange). Implications of these integration window values will be discussed below. Finally, the value of the hyperpolarization delay, d_{Hyp}, affected octopus-cell chirp sensitivity (Fig. 3.10A) and sensitivity of the IC RVF (Fig. 3.10B), with short delay associated with reduced chirp-direction sensitivity.

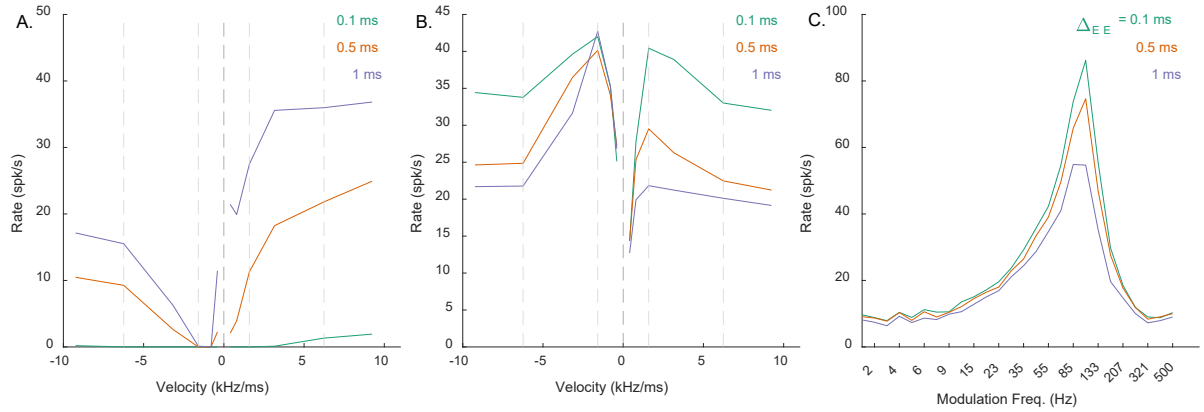

Figure 3.8 – Impact of varying Δ_{EE} on responses of an example mid-CF (4 kHz) model cell with downward-sensitive IC output (Green – Δ_{EE} = 0.1 ms, orange – Δ_{EE} = 0.5 ms, purple – Δ_{EE} = 1 ms). Responses for Δ_{EE} = 1 ms are also shown in Figs. 3.5A-C. A) octopus cell RVF; B) IC RVF; C) IC MTF.

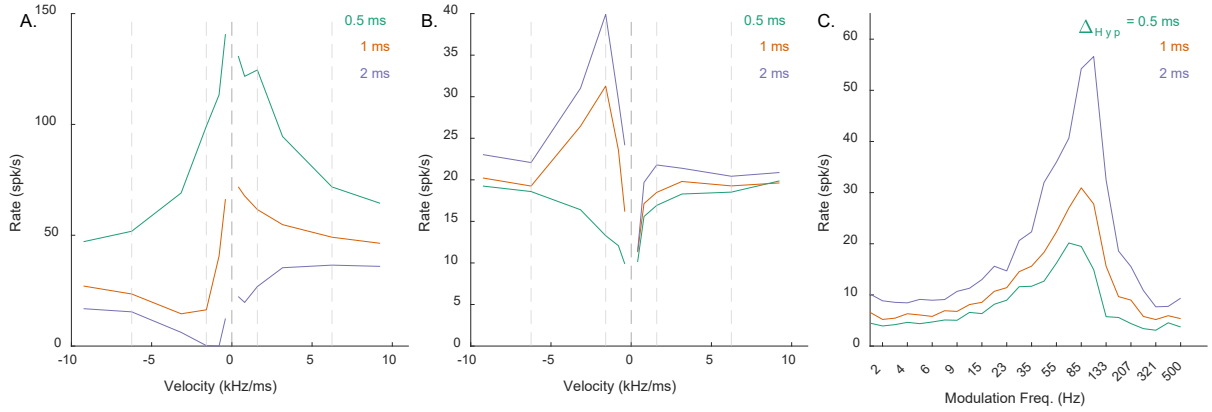

Figure 3.9 – Impact of varying Δ_{Hyp} on responses of an example mid-CF (4 kHz) model cell with downward-sensitive IC output (Green – Δ_{Hyp} = 0.5 ms, orange – Δ_{Hyp} = 1 ms, purple – Δ_{Hyp} = 2 ms). Responses for Δ_{Hyp} = 2 ms are also shown in Figs. 3.5A-C. A) octopus cell RVF; B) IC RVF; C) IC MTF.

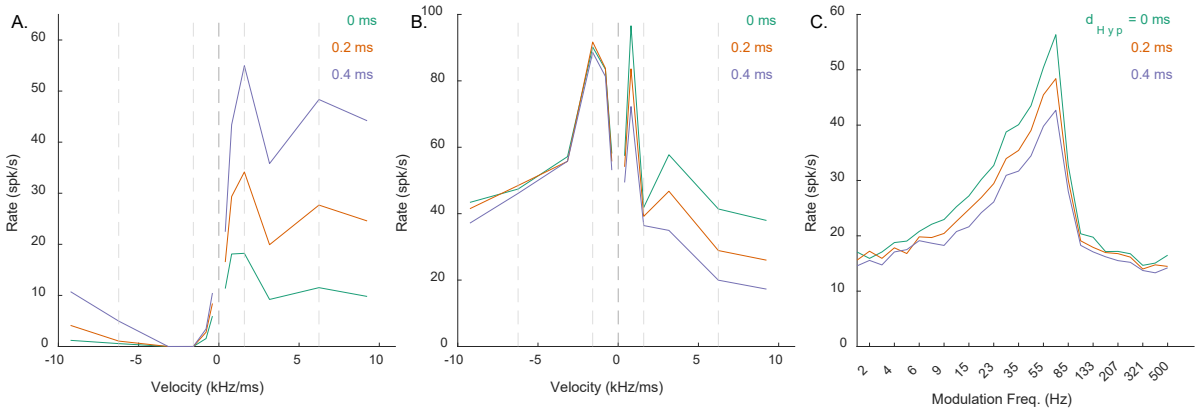

Figure 3.10 – Impact of varying d_{Hyp} on responses of an example high-CF (8 kHz) model cell with downward-sensitive IC output (Green – d_{Hyp} = 0 ms, orange – d_{Hyp} = 0.2 ms, purple – d_{Hyp} = 0.4 ms). Responses for d_{Hyp} = 0.4 ms are also shown in Figs. 3.6A-C. A) octopus cell RVF; B) IC RVF; C) IC MTF.

At the IC stage, chirp-sensitivity was primarily controlled by the parameters M_{Oct} and d_E, the number of inhibitory octopus-cell inputs and the delay of the excitatory input relative to the octopus-cell inputs. For the mid-CF, upward-sensitive, example IC neuron, when M_{Oct} was set at zero, there was no impact of the octopus cell on the RVF (Fig. 3.11A), but the MTF had a large rate at the peak (Fig. 3.11B). As M_{Oct} increased, the RVF became upward-sensitive (Fig. 3.11A), but the MTF rate decreased (Fig. 3.11B), demonstrating that M_{Oct} selection must balance the desired chirp-direction sensitivity and prominence of the BE MTF.

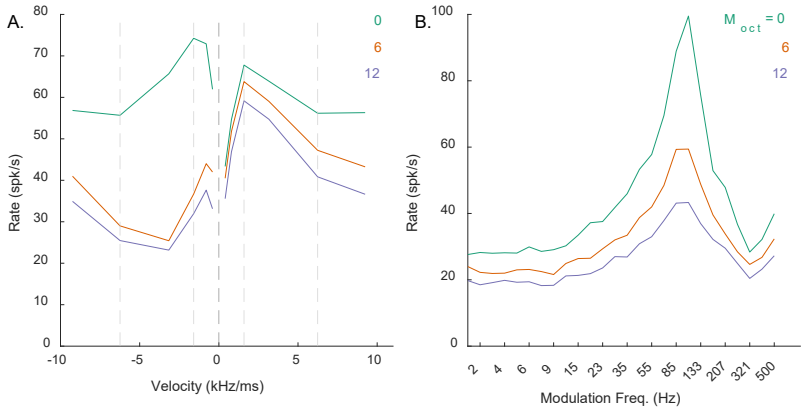

Figure 3.11 – Impact of varying M_{Oct} on responses of an example mid-CF (4 kHz) model cell with upward-sensitive IC output (Green – M_{Oct} = 0, orange – M_{Oct} = 6, purple – M_{Oct} = 12). Responses for M_{Oct} = 6 are also shown in Figs. 3.5D-F. A) IC RVF; B) IC MTF.

To maximize IC chirp-direction sensitivity, d_E must allow octopus-cell inhibition to arrive sufficiently before excitation. For the example mid-CF, downward-sensitive neuron, d_E = 0.5 ms resulted in the largest downward-chirp sensitivity (Fig. 3.12A). Using a higher value of d_E, such as 1 ms, reduced downward-chirp sensitivity (Fig. 3.12A).

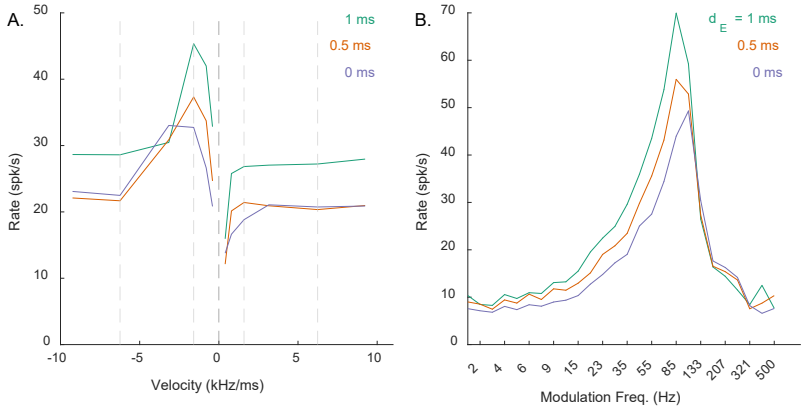

Figure 3.12 – Impact of varying d_E on responses of an example mid-CF (4 kHz) model cell with downward-sensitive IC output (Green – d_E = 1 ms, orange – d_E = 0.5 ms, purple – d_E = 0 ms). Responses for d_E = 0.5 ms are also shown in Figs. 3.5A-C. A) IC RVF; B) IC MTF.

The AM-tuning of the model IC neuron was controlled by M_I and d_I: M_I defined the number of inhibitory CF inputs the neuron received, and d_I defined the delay of the CF inhibition relative to the octopus-cell inputs, where $d_I - d_E$ was the delay between CF excitation and CF inhibition. If M_I were zero, the tuning of the MTF was not affected by same-frequency inhibition (Fig. 3.13B), instead having a flat or possibly band-suppressed MTF shape. Increasing M_I gave the MTF a BE shape and sharpened the peak (Fig. 3.13B), and, as expected, reduced the response rates across all stimuli (Fig. 3.13A).

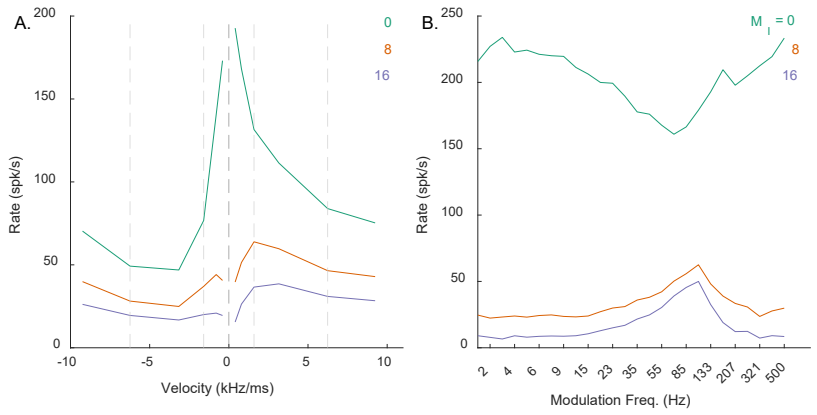

Figure 3.13 – Impact of varying M_I on responses of an example mid-CF (4 kHz) model cell with upward-sensitive IC output (Green – M_I = 0, orange – M_I = 8, purple – M_I = 16). Responses for M_I = 8 are also shown in Figs. 3.5D-F. A) IC RVF; B) IC MTF.

The delay between CF excitation and inhibition, $d_I - d_E$, determined the location of the MTF peak. For the example mid-CF neuron, as this value increased, the BE peak shifted to lower modulation frequencies (Fig. 3.14B). The RVF was also affected when the delay was small (Fig. 3.14A), illustrating an interaction between the peak modulation frequency of the MTF and the shape of the RVF.

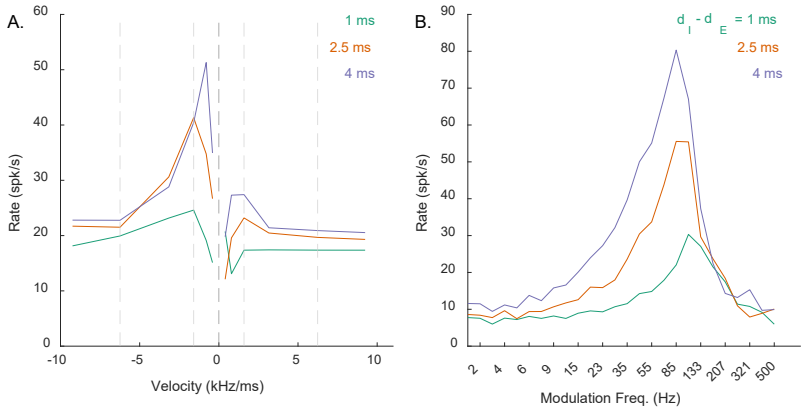

Figure 3.14 – Impact of varying $d_I - d_E$ on responses of an example mid-CF (4 kHz) model cell with downward-sensitive IC output (Green – $d_I - d_E$ = 1 ms, orange – $d_I - d_E$ = 2.5 ms, purple – $d_I - d_E$ = 4 ms). Responses for $d_I - d_E$ = 2.5 ms are also shown in Figs. 3.5A-C. A) IC RVF; B) IC MTF.

3.5 Discussion

These results describe a computational model for both chirp-sensitivity and periodicity tuning in the IC. Chirp-sensitivity in the model originated in model PVCN octopus cells, which also had characteristics such as O_I tone responses and click entrainment. Depending on parameter values, IC cells with sensitivity towards either chirp direction and with BE-type MTFs could be simulated for low, medium, and high CFs. Chirp-sensitivity and periodicity tuning were largely controlled by separate inhibitory parameters in the IC. Model parameters had systematic effects on IC RVFs and MTFs, allowing responses to be tuned.

Responses to tone and click stimuli confirmed that the octopus-cell stage was consistent with this cell type's physiological responses (Fig. 3.3), although the primary purpose of this stage was to provide a chirp-sensitive input to the IC. Octopus cells have an ideal onset (O_I) response to high-frequency tones, with one well-timed response at tone onset (Godfrey et al., 1975; Rhode and Smith, 1986). The model octopus cell responded to a tone at CF with a peak in the rate function shortly after tone onset,

followed by no activity. Frequency-response areas of octopus cells are broad, consistent with their wide-dendritic fields (Osen, 1969; Rhode et al., 1983). O_I cells tend to have thresholds greater than 30 dB, with much higher rates at frequencies lower than CF and at high sound levels (Rhode and Smith, 1986; Rhode, 1994), as does the model RM (Fig. 3.3 B). Octopus cells are also characterized by entrainment to modulated stimuli such as click trains, responding with one precisely timed action potential for every cycle for frequencies up to 500-800 Hz (Godfrey et al., 1975; Rhode, 1994; Oertel et al., 2000), similar to model responses (Fig. 3.3C). Finally, the model click MTF (Fig. 3.3D) increased monotonically up to 600 Hz, a slightly lower frequency than observed in O_I neurons, which entrained up to 700 Hz to clicks of a comparable level (Godfrey et al., 1975).

The results of this modeling study agree with the physiological results of our previous study of the responses of rabbit IC neurons to chirp stimuli (Mitchell et al., 2023). Diverse chirp sensitivity was observed across all CF ranges (Mitchell et al., 2023). Here, we show that it was possible to choose model parameters to produce chirp sensitivity for a similar range of CFs. In physiology, chirp-direction sensitivity is more common towards low-speed chirps (< 2 kHz/ms) than towards high-speed chirps (> 2 kHz/ms), a result also observed in model responses (Fig. 3.4-6). For many example IC model responses, rate-differences between the fastest chirp-pairs were smaller than for slower chirp-pairs (Fig. 3.5B), and sometimes displayed bias towards the opposite direction (Fig. 3.5E). This property of RVFs at high chirp-velocities, also observed in some physiological responses (e.g., Fig. 2.4D, Mitchell et al., 2023), may occur when chirps of opposing directions begin to resemble each other as representations in the AN become more click-like.

Also, Mitchell et al. (2023) suggested that chirp-direction sensitivity and periodicity tuning in the IC originate from different neural mechanisms that create two distinct feature sensitivities. This conclusion was echoed in the modeling results—chirp-sensitivity in the model was strongest when octopus-cell inhibition arrived about 1 ms before the IC excitatory input (Fig. 3.12), and BE MTF peaks were highest when the same-frequency inhibition arrived more than 1 ms after IC excitation (Fig. 3.14). That these inhibitions functioned best when their timings did not coincide suggests that they come from separate inputs.

In the six example model neurons shown (Fig. 3.4-6), the low-CF octopus cell RVFs (Fig. 3.4) display less uniform direction bias than high-CF octopus cell RVFs (Fig. 3.6). However, direction sensitivity is not observed less frequently in low-CF neurons—rather, this difference is likely a result of the parameter-fitting process. The difference between CF and optimal OCF is much closer for a low-CF octopus cell than for a high-CF one. For the 1-kHz, upward-selective neuron, the spacing between CF and OCF is only 100 Hz, while for the 8-kHz, upward-selective neuron, it is 2.5 kHz. The chirp stimulus passes through these frequencies linearly, requiring much more precise fitting of the delay parameter d_{CF} for the low-CF octopus cell to achieve the same direction bias as the high-CF one.

Krips and Furst (2009a,b) showed that CD responses are NHPPs for an integration window Δ much smaller than the refractory period of their inputs. They suggested a Δ of 200 μs for EI CDs and 20 μs for EE CDs (Krips and Furst, 2009a). The rationale for the limitation of Δ is to prevent multiple spikes from the same input from occurring within the integration window, thereby triggering a response without a coincidence from multiple inputs. However, here, a small Δ_{EE} (≤ 100 μs) resulted in the

model octopus cell, and subsequently the model IC neuron, having greatly reduced chirp-direction sensitivity (Fig. 3.8). A short Δ_{Hyp}, the hyperpolarization integration window, also reduced chirp sensitivity (Fig. 3.9). Due to the long integration windows used in the model, which appear to be necessary to ensure chirp-direction sensitivity, the model output would not be a NHPP. This is a limitation of the current model, potentially preventing its use in estimating psychophysical detection thresholds for various stimulus parameters. However, the prerequisite for long integration windows would be expected for a model of a chirp-direction sensitive neuron with widely spaced CF inputs. Conceptually, there is a natural trade-off between difference in input CF and length of integration window—when inputs are far apart in CF, they necessarily require a larger Δ. Also, KL hyperpolarizations have relatively long timescales compared to the suggested integration windows (Golding et al., 1995). Long Δ_{Hyp} may be unavoidable if treating KL hyperpolarization as an inhibitory input, as was done here. In the future, it might be possible to identify sets of model parameters that allow shorter integration windows but still retain chirp-sensitivity. Furthermore, hyperpolarization inhibition cannot be treated as independent from its corresponding excitatory input as it is entirely conditional upon the excitatory activity. In future work, this limitation could be addressed by combining the “locked” excitatory AN input and inhibitory hyperpolarization into a complex “excitatory-inhibitory” signal more closely representing the full post-synaptic response.

Parameters for the example neurons were selected, using a combination of parameter optimization and manual selection, to maximize direction-sensitivity. The values of these parameters align well with their physiological correlates. For instance, octopus cells provide early onset inhibition to the IC via the VNLL (Covey and Casseday,

1991). Intracellular recordings in VNLL and IC cells showed that IC cells that received VNLL inhibition often displayed an early inhibition before action potentials in response to sounds (Nayagam et al., 2005). In the present study, inhibition arriving from the octopus-cell stage 0.5-1 ms before the excitatory input maximized the chirp sensitivity (Fig. 3.12). Additionally, if this inhibition arrived too early or too late relative to excitation, chirp sensitivity was diminished. Similarly, the CFs of AN inputs to the octopus cell stage were aligned with experimental and modeling studies, which show that these CFs can come from a wide range of frequencies: Spencer et al. (2012) estimated CFs of AN inputs to octopus cells based on physiological recordings in cats, and determined they can range from 1.5 – 40 kHz (Godfrey et al., 1975; Rhode and Smith, 1986).

In this paper, octopus cells were proposed as a source of chirp-velocity sensitivity for IC cells; however, alternative mechanisms have been proposed for frequency-modulation (FM) sensitivity. For example, Pollak et al. (2011) summarized two mechanisms other than VNLL-inhibition that could explain sensitivity of IC neurons to FM chirps. One of these is the classical explanation for FM sensitivity, based on asymmetry in the timing and frequency of excitation and inhibition (Fuzessery and Hall, 1996; Gordon and O'Neill, 1998; Andoni et al., 2007). This asymmetry is revealed by spectrotemporal receptive fields (STRFs), a technique using spike-triggered averaging to generate a kernel used to identify excitatory and inhibitory regions. STRFs have been shown to predict the sensitivity of chirp-direction sensitive neurons in bat IC (Andoni et al., 2007). Another hypothesis for chirp sensitivity in the IC proposes that cells with high input resistances and long time constants could be sensitive to asymmetry in input magnitudes, rather than input timing (Gittelman et al., 2009). The modular nature of

Krips and Furst's modeling strategy may facilitate exploration of these additional chirp-sensitivity mechanisms.

In physiological recordings, chirp-sensitive neurons with band-suppressed (BS) MTFs, characterized by lower rates in response to modulated stimuli compared to unmodulated stimuli, are at least as common as BE ones (Mitchell et al., 2023). It may be possible to model chirp-sensitive neurons with BS MTFs by using a similar strategy to Carney et al. (2015), which used an inhibitory input from a BE model cell. Implementing an inhibitory interneuron would require careful calibration of additional timing parameters but may be a useful advancement towards understanding the responses of all chirp-sensitive IC neurons.

The results shown here involved manual selection of IC parameters, with the goal of maximizing direction-sensitivity in the RVF; other response characteristics, such as the salience of MTF tuning, could be increased at the expense of direction-sensitivity. In general, the model parameter space is open-ended, with a potential to simulate neurons with differing response features. A strategy of parameter fitting could eventually be employed to simulate actual IC neuron recordings. Additionally, keeping in mind the sensitivity of these neurons to interaural differences, one possible future direction could be to add binaural inputs to the model, with the VNLL inhibition to the IC driven by contralateral octopus cells (Vater et al., 1997).

The model presented here for IC chirp-velocity sensitivity and AM tuning provides a tool for investigating the contribution of velocity sensitivity to complex sounds, such as speech responses. In speech stimuli, phase shifts due to vocal-tract filtering (Klatt, 1980) would result in frequency chirps within pitch periods. An IC model that is sensitive to

chirp velocity may improve the accuracy of predictions of physiological responses to speech sounds. Given the ubiquity of such neurons in the IC (Mitchell et al., 2023), such a study would be important in elucidating the processing of speech in the midbrain.

Acknowledgements

This work was funded by NIH-F31DC019816, NIDCD-R01-001641 and NIDCD-R01-010813.

Bibliography

Adams, J. C. (1997). Projections from octopus cells of the posteroventral cochlear nucleus to the ventral nucleus of the lateral lemniscus in cat and human. Aud Neurosci, 3(4), 335-350.

Andoni, S., Li, N., & Pollak, G. D. (2007). Spectrotemporal Receptive Fields in the Inferior Colliculus Revealing Selectivity for Spectral Motion in Conspecific Vocalizations. The Journal of Neuroscience, 27(18), 4882-4893. https://doi.org/10.1523/jneurosci.4342-06.2007

Bal, R., & Oertel, D. (2001). Potassium Currents in Octopus Cells of the Mammalian Cochlear Nucleus. Journal of Neurophysiology, 86(5), 2299-2311. https://doi.org/10.1152/jn.2001.86.5.2299

Cai, H., Carney, L. H., & Colburn, H. S. (1998). A model for binaural response properties of inferior colliculus neurons. I. A model with interaural time difference-sensitive excitatory and inhibitory inputs. The Journal of the Acoustical Society of America, 103(1), 475-493. https://doi.org/10.1121/1.421100

Carney, L. H., Li, T., & McDonough, J. M. (2015). Speech Coding in the Brain: Representation of Vowel Formants by Midbrain Neurons Tuned to Sound Fluctuations. Eneuro, 2(4), ENEURO.0004-0015.2015. https://doi.org/10.1523/eneuro.0004-15.2015

Colburn, H. S. (1973). Theory of binaural interaction based on auditory-nerve data. I. General strategy and preliminary results on interaural discrimination. The Journal of the Acoustical Society of America, 54(6), 1458-1470. https://doi.org/10.1121/1.1914445

Covey, E., & Casseday, J. (1991). The monaural nuclei of the lateral lemniscus in an echolocating bat: parallel pathways for analyzing temporal features of sound. The Journal of Neuroscience, 11(11), 3456-3470. https://doi.org/10.1523/jneurosci.11-11-03456.1991

Farhadi, A., Jennings, S. G., Strickland, E. A., & Carney, L. H. (2023). Subcortical auditory model including efferent dynamic gain control with inputs from cochlear nucleus and inferior colliculus. The Journal of the Acoustical Society of America, 154(6), 3644-3659. https://doi.org/10.1121/10.0022578

Fuzessery, Z. M., & Hall, J. C. (1996). Role of GABA in shaping frequency tuning and creating FM sweep selectivity in the inferior colliculus. Journal of Neurophysiology, 76(2), 1059-1073. https://doi.org/10.1152/jn.1996.76.2.1059

Gittelman, J. X., Li, N., & Pollak, G. D. (2009). Mechanisms Underlying Directional Selectivity for Frequency-Modulated Sweeps in the Inferior Colliculus Revealed by In Vivo Whole-Cell Recordings. The Journal of Neuroscience, 29(41), 13030-13041. https://doi.org/10.1523/jneurosci.2477-09.2009

Godfrey, D. A., Kiang, N. Y. S., & Norris, B. E. (1975). Single unit activity in the posteroventral cochlear nucleus of the cat. Journal of Comparative Neurology, 162(2), 247-268. https://doi.org/https://doi.org/10.1002/cne.901620206

Golding, N., Robertson, D., & Oertel, D. (1995). Recordings from slices indicate that octopus cells of the cochlear nucleus detect coincident firing of auditory nerve fibers with temporal precision. The Journal of Neuroscience, 15(4), 3138-3153. https://doi.org/10.1523/jneurosci.15-04-03138.1995

Golding, N. L., Ferragamo, M. J., & Oertel, D. (1999). Role of Intrinsic Conductances Underlying Responses to Transients in Octopus Cells of the Cochlear Nucleus. The Journal of Neuroscience, 19(8), 2897-2905. https://doi.org/10.1523/jneurosci.19-08-02897.1999

Gordon, M., & O'Neill, W. E. (1998). Temporal processing across frequency channels by FM selective auditory neurons can account for FM rate selectivity. Hearing research, 122(1), 97-108. https://doi.org/https://doi.org/10.1016/S0378-5955(98)00087-2

Guest, D. R., & Carney, L. H. (2023). A fast and flexible approximation of power-law adaptation for auditory computational models. bioRxiv, 2023.2011.2030.569467. https://doi.org/10.1101/2023.11.30.569467

Heinz, M. G., Colburn, H. S., & Carney, L. H. (2001). Evaluating Auditory Performance Limits: I. One-Parameter Discrimination Using a Computational Model for the Auditory Nerve. Neural computation, 13(10), 2273-2316. https://doi.org/10.1162/089976601750541804

Heinz, M. G., Colburn, H. S., & Carney, L. H. (2001). Evaluating Auditory Performance Limits: II. One-Parameter Discrimination with Random-Level Variation. Neural computation, 13(10), 2317-2338. https://doi.org/10.1162/089976601750541813

Heinz, M. G., Colburn, H. S., & Carney, L. H. (2002). Quantifying the implications of nonlinear cochlear tuning for auditory-filter estimates. The Journal of the Acoustical Society of America, 111(2), 996-1011. https://doi.org/10.1121/1.1436071

Henry, K. S., Wang, Y., Abrams, K. S., & Carney, L. H. (2023). Mechanisms of masking by Schroeder-phase harmonic tone complexes in the budgerigar (Melopsittacus undulatus). Hearing research, 435, 108812. https://doi.org/https://doi.org/10.1016/j.heares.2023.108812

Hewitt, M. J., & Meddis, R. (1994). A computer model of amplitude-modulation sensitivity of single units in the inferior colliculus. The Journal of the Acoustical Society of America, 95(4), 2145-2159. https://doi.org/10.1121/1.408676

Kalluri, S., & Delgutte, B. (2003). Mathematical Models of Cochlear Nucleus Onset Neurons: I. Point Neuron with Many Weak Synaptic Inputs. Journal of Computational Neuroscience, 14(1), 71-90. https://doi.org/10.1023/A:1021128418615

Keithley, E. M., & Schreiber, R. C. (1987). Frequency map of the spiral ganglion in the cat. The Journal of the Acoustical Society of America, 81(4), 1036-1042. https://doi.org/10.1121/1.394675

Kim, D. O., Carney, L., & Kuwada, S. (2020). Amplitude modulation transfer functions reveal opposing populations within both the inferior colliculus and medial geniculate body. Journal of Neurophysiology, 124(4), 1198-1215. https://doi.org/10.1152/jn.00279.2020

Klatt, D. H. (1980). Software for a cascade/parallel formant synthesizer. The Journal of the Acoustical Society of America, 67(3), 971-995. https://doi.org/10.1121/1.383940

Kreeger, L. J., Honnuraiah, S., Maeker, S., Shea, S., Fishell, G., & Goodrich, L. V. (2024). An Anatomical and Physiological Basis for Coincidence Detection Across Time Scales in the Auditory System. bioRxiv, 2024.2002.2029.582808. https://doi.org/10.1101/2024.02.29.582808

Krips, R., & Furst, M. (2009). Stochastic properties of auditory brainstem coincidence detectors in binaural perception. The Journal of the Acoustical Society of America, 125(3), 1567-1583. https://doi.org/10.1121/1.3068446

Krips, R., & Furst, M. (2009). Stochastic Properties of Coincidence-Detector Neural Cells. Neural computation, 21(9), 2524-2553. https://doi.org/10.1162/neco.2009.07-07-563

Liberman, A. M., & Mattingly, I. G. (1989). A Specialization for Speech Perception. Science, 243(4890), 489-494. https://doi.org/doi:10.1126/science.2643163

Liberman, M. C. (1978). Auditory-nerve response from cats raised in a low-noise chamber. The Journal of the Acoustical Society of America, 63(2), 442-455. https://doi.org/10.1121/1.381736

Liberman, M. C. (1993). Central projections of auditory nerve fibers of differing spontaneous rate, II: Posteroventral and dorsal cochlear nuclei. Journal of Comparative Neurology, 327(1), 17-36. https://doi.org/https://doi.org/10.1002/cne.903270103

Lu, H.-W., Smith, P. H., & Joris, P. X. (2022). Mammalian octopus cells are direction selective to frequency sweeps by excitatory synaptic sequence detection. Proceedings of the National Academy of Sciences, 119(44), e2203748119. https://doi.org/doi:10.1073/pnas.2203748119

Manis, P. B., & Campagnola, L. (2018). A biophysical modelling platform of the cochlear nucleus and other auditory circuits: From channels to networks. Hearing research, 360, 76-91. https://doi.org/https://doi.org/10.1016/j.heares.2017.12.017

Mitchell, P. W., Henry, K. S., & Carney, L. H. (2023). Sensitivity to direction and velocity of fast frequency chirps in the inferior colliculus of awake rabbit. Hearing research, 440, 108915. https://doi.org/https://doi.org/10.1016/j.heares.2023.108915

Nayagam, D. A. X., Clarey, J. C., & Paolini, A. G. (2005). Powerful, Onset Inhibition in the Ventral Nucleus of the Lateral Lemniscus. Journal of Neurophysiology, 94(2), 1651-1654. https://doi.org/10.1152/jn.00167.2005

Nelson, P. C., & Carney, L. H. (2004). A phenomenological model of peripheral and central neural responses to amplitude-modulated tones. The Journal of the Acoustical Society of America, 116(4), 2173-2186. https://doi.org/10.1121/1.1784442

Oertel, D., Bal, R., Gardner, S. M., Smith, P. H., & Joris, P. X. (2000). Detection of synchrony in the activity of auditory nerve fibers by octopus cells of the mammalian cochlear nucleus. Proceedings of the National Academy of Sciences, 97(22), 11773-11779. https://doi.org/doi:10.1073/pnas.97.22.11773

Osen, K. K. (1969). Cytoarchitecture of the cochlear nuclei in the cat. Journal of Comparative Neurology, 136(4), 453-483. https://doi.org/https://doi.org/10.1002/cne.901360407

Pollak, G. D., Gittelman, J. X., Li, N., & Xie, R. (2011). Inhibitory projections from the ventral nucleus of the lateral lemniscus and superior paraolivary nucleus create directional selectivity of frequency modulations in the inferior colliculus: A comparison of bats with other mammals. Hearing research, 273(1), 134-144. https://doi.org/https://doi.org/10.1016/j.heares.2010.03.083

Rebhan, M., & Leibold, C. (2021). A phenomenological spiking model for octopus cells in the posterior–ventral cochlear nucleus. Biological Cybernetics, 115(4), 331-341. https://doi.org/10.1007/s00422-021-00881-x

Rhode, W. S. (1994). Temporal coding of 200% amplitude modulated signals in the ventral cochlear nucleus of cat. Hearing research, 77(1), 43-68. https://doi.org/https://doi.org/10.1016/0378-5955(94)90252-6

Rhode, W. S., Oertel, D., & Smith, P. H. (1983). Physiological response properties of cells labeled intracellularly with horseradish peroxidase in cat ventral cochlear nucleus. Journal of Comparative Neurology, 213(4), 448-463. https://doi.org/https://doi.org/10.1002/cne.902130408

Rhode, W. S., & Smith, P. H. (1986). Encoding timing and intensity in the ventral cochlear nucleus of the cat. Journal of Neurophysiology, 56(2), 261-286. https://doi.org/10.1152/jn.1986.56.2.261

Rieke, F., Warland, D., Van Steveninck, R. d. R., & Bialek, W. (1999). Spikes: exploring the neural code. MIT press.

Schroeder, M. (1970). Synthesis of low-peak-factor signals and binary sequences with low autocorrelation (Corresp.). IEEE Transactions on Information Theory, 16(1), 85-89. https://doi.org/10.1109/TIT.1970.1054411

Siebert, W. M. (1965). Some implications of the stochastic behavior of primary auditory neurons. Kybernetik, 2(5), 206-215. https://doi.org/10.1007/BF00306416

Siebert, W. M. (1970). Frequency discrimination in the auditory system: Place or periodicity mechanisms? Proceedings of the IEEE, 58(5), 723-730. https://doi.org/10.1109/PROC.1970.7727

Spencer, M., Grayden, D., Bruce, I., Meffin, H., & Burkitt, A. (2012). An investigation of dendritic delay in octopus cells of the mammalian cochlear nucleus [Original Research]. Frontiers in Computational Neuroscience, 6. https://doi.org/10.3389/fncom.2012.00083

Spencer, M. J., Meffin, H., Burkitt, A. N., & Grayden, D. B. (2018). Compensation for Traveling Wave Delay Through Selection of Dendritic Delays Using Spike-Timing-Dependent Plasticity in a Model of the Auditory Brainstem [Original Research]. Frontiers in Computational Neuroscience, 12. https://doi.org/10.3389/fncom.2018.00036

Steenken, F., Oetjen, H., Beutelmann, R., Carney, L. H., Koeppl, C., & Klump, G. M. (2022). Neural processing and perception of Schroeder-phase harmonic tone complexes in the gerbil: Relating single-unit neurophysiology to behavior. European Journal of Neuroscience, 56(3), 4060-4085. https://doi.org/https://doi.org/10.1111/ejn.15744

Vater, M., Covey, E., & Casseday, J. H. (1997). The columnar region of the ventral nucleus of the lateral lemniscus in the big brown bat (Eptesicus fuscus): synaptic arrangements and structural correlates of feedforward inhibitory function. Cell and tissue research, 289(2), 223-233. https://doi.org/10.1007/s004410050869

Zilany, M. S. A., Bruce, I. C., & Carney, L. H. (2014). Updated parameters and expanded simulation options for a model of the auditory periphery. The Journal of the Acoustical Society of America, 135(1), 283-286. https://doi.org/10.1121/1.4837815

Chapter 4: Examining the Impact of Chirp-Sensitivity on Vowel Coding in the Inferior Colliculus

4.1 Abstract

The inferior colliculus (IC) is an important brain region to understand neural encoding of complex sounds due to its diverse sound-feature sensitivities and capacity to utilize peripheral nonlinearities. Fast spectrotemporal changes, known as chirps, are contained within pitch-periods of natural vowels. Recent physiological studies in rabbit IC demonstrate that IC neurons are sensitive to chirp direction and velocity. Here, we use a combination of physiology and modeling strategies to assess the impact of chirp-sensitivity on vowel coding. Neural responses to vowel stimuli were recorded and vowel token discrimination was evaluated based on average rate and spike timing. Response timing was found to result in higher classification accuracy than rate. Additionally, rate bias towards low-velocity chirps, independent of chirp direction, was shown to correlate with higher vowel classification accuracy based on timing. Also, direction bias in response to chirps of high velocity was shown to correlate with vowel classification accuracy based on rate and timing. Responses to natural vowel tokens of individual neurons were simulated using an IC model with controllable chirp sensitivity—responses of upward-biased, downward-biased, and non-direction-biased were generated. Manipulating chirp sensitivity clearly influenced response profiles across natural vowel tokens and model neuron discrimination, though more work is needed relating model responses to similar neurons in physiological recordings.

4.2 Introduction

Vowels carry vital linguistic information and are an important signal for speech perception (Kewley-Port et al., 2007). Acoustically, vowels are defined by the shape of their harmonic spectra, and by the fundamental frequency (F0) which is determined by the pitch of the speaker's voice. Vowels can be distinguished from one another by the frequencies of the spectral peaks, or formants, resulting from vocal-tract filtering (Fant, 1960). The first two formant frequencies, referred to as F1 and F2, are sufficient for vowel identification (Hillenbrand et al, 1995). It is important to understand how the auditory system converts these acoustic spectral peaks into a neural code by which vowel discrimination is accomplished. However, non-linearity in the auditory system shapes neural representation of these signals such that the underlying mechanisms supporting vowel coding at various levels of the auditory system are not well understood. Traditionally, studies of speech coding in the auditory nerve (AN) have established average rate and temporal fine structure to be primary candidates for vowel identification (Sachs and Young, 1979; Young and Sachs, 1979), but both of these fail at moderately high sound levels (Delgutte and Kiang, 1984a) or in background noise (Delgutte and Kiang, 1984b), suggesting that a more comprehensive vowel code may take shape elsewhere in the auditory system.

The inferior colliculus (IC) is a nearly obligatory synapse in the ascending auditory pathway (Aitkin and Phillips, 1983). In addition to spectral tuning to a characteristic frequency (CF) (Schreiner and Langner, 1997), IC neurons commonly display sensitivity to amplitude modulation (AM) (Joris et al., 2004). Modulation transfer functions (MTFs), which depict neural rate versus modulation frequency (Kim et al., 2020), reveal that the best modulation frequencies (BMFs) of IC neurons span the F0-

range of speech (Langner, 1992; Krishna and Semple, 2000). Sensitivity to F0 periodicity has been suggested to underlie a representation of vowels in the IC by converting neural fluctuations in the periphery to rate coding of formants (Carney et al., 2015; Carney, 2024), and is an example of how speech coding can emerge from sensitivity to sound features.

Beyond AM tuning, most IC neurons are also sensitive to the velocity of fast frequency chirps resulting from phase-differences between components in harmonic sounds (Steenken et al., 2023; Henry et al., 2023; Mitchell et al., 2023). This sensitivity is reflected in large rate differences in response to chirps having opposite directions but identical speeds. Sensitivity to chirp direction has been observed in the majority of IC neurons tested (Mitchell et al., 2023); yet, the implications of chirp-velocity sensitivity for vowel processing are unknown. Phase-differences between harmonics, which are associated with frequency chirps, are a feature of natural speech due to the resonant characteristics of the vocal-tract filter. A related feature of vowels, the group-delay functions that track time delay versus frequency, indicate peaks in time lag at and around formant frequencies (Bozkurt et al., 2006; Rajan et al., 2013).

In this paper, we test the hypothesis that IC chirp sensitivity affects neural coding of vowels using physiological and modeling methods. Examining the responses of IC neurons to natural vowels, we evaluate the impact of chirp-direction bias on average rate and temporal coding of vowels. Finally, using a model of chirp-sensitive IC neurons, we simulate physiologically plausible responses to vowel stimuli and compare them for model cells with and without chirp sensitivity.

4.3 Methods

In order to test the hypothesis that chirp-sensitivity of IC neurons impacts responses to vowels, we employed a combination of physiology and modeling methods.

4.3.1 Neural Recordings

Physiological methods are described in detail in Mitchell et al. (2023). Briefly, extracellular neural recordings were made in the central nucleus of the inferior colliculus (ICC) of awake Dutch-belted rabbits (*Oryctolagus cuniculus*) using tetrodes. Data were collected in a total of 5 animals with normal hearing, assessed using distortion product otoacoustic emissions (DPOAEs). All methods were approved by the University of Rochester Committee on Animal Resources.

An initial craniotomy procedure was conducted to attach the headbar and implant the microdrive and initial tetrodes in the ICC. Multiple surgeries were conducted to remove and replace the tetrodes. All surgeries were performed with anesthesia, using ketamine (66 mg/kg) and xylazine (2 mg/kg), administered intramuscularly. Time between tetrode replacement surgeries was approximately 1-3 months; in between these, tetrodes were advanced and retracted using the implanted microdrive, as needed, to locate new neurons. Data collection was conducted daily in 2-hour sessions in a sound-attenuated booth. Sound was delivered to the rabbits via custom-made earmolds. At the beginning of every session, the system was calibrated to compensate the stimuli for the frequency response of the acoustic system.

Voltage recordings were made using an RHD recording system (Intan Technologies, LLC., Los Angeles, CA, USA) and Intan software. To identify single-unit action potentials (spikes), the voltage recording was filtered using a 4^{th}-order Butterworth

bandpass filter (300 – 3000 Hz). Spikes were identified when the voltage recording exceed a threshold defined as four standard deviations of the signal. Features of the spike waveforms were used to sort spikes into clusters, primarily the slope of repolarization (Schwarz et al., 2012). Clusters were identified as single-unit neural responses when less than 2% of the inter-spike intervals were shorter than 1 ms. Neurons identified in consecutive sessions were considered unique only if both tetrode location and response properties changed.

4.3.2 Stimuli

Frequency response maps (RMs) were used to assess characteristic frequency (CF), defined as the frequency that elicited the highest response rate at threshold level. To generate RMs, a series of 0.2-s-duration tones were presented at different levels (13-, 33-, 53-, and 73-dB SPL) and frequencies (250 Hz—16 kHz) in random order. Each tone was presented 3 times, either contralaterally or diotically, and included 10-ms raised-cosine on/off ramps. Tones were separated by 0.4 s of silence.

Modulation transfer functions (MTFs) were used to assess neural sensitivity to amplitude modulation (AM). To generate the stimulus, a 1-s-duration wideband noise (100 Hz—10 kHz) was generated for each repetition at 33 dB spectrum level (overall level of 73 dB SPL). Then, the noise was 100% sinusoidally amplitude modulated, with modulation frequency in a range 2-350 Hz, with 3 steps per octave. Each modulation frequency was presented a total of 5 times in random order. Stimuli were presented diotically, with 50-ms raised-cosine on/off ramps. MTFs were categorized based on their shape into one of four categories: band-enhanced (BE) neurons have significantly increased rate for a band of modulation frequencies relative to

unmodulated rate (Kim et al, 2020; Mitchell et al., 2023). Band suppressed (BS) neurons had significantly decreased rate for a band of modulation frequencies relative to unmodulated rate. Hybrid (H) neurons had both a band of increased and a band of decreased rates relative to unmodulated rate. Finally, flat (F) neurons had no significantly different bands compared to unmodulated.

Rate-velocity functions (RVFs) were used to evaluate neural chirp-sensitivity (Mitchell et al., 2023). To generate the stimulus, a set of chirps were generated with velocities ±0.40, ±0.80, ±1.59, ±3.16, ±6.24, and ±9.24 kHz/ms (equivalent to Schroeder-harmonic complex of fundamental frequency (F0) 25, 50, 100, 200, 400, and 600 Hz, respectively, with instantaneous frequencies spanning from F0 to 16 kHz). To normalize energy in these variable-duration chirps, stimuli were assigned a sound level equal to $68 \text{ dB SPL} - 10 \times \log_{10}(T/T_{ref})$, where T is the duration of the chirp, and T_{ref} = 2.5 ms (duration of the chirp of ±6.24 kHz/ms). Additionally, raised-cosine on/off ramps were applied with duration equivalent to 10% of chirp duration. Then, chirps were presented in a random order, separated by 40—60 ms of silence to ensure aperiodicity. Each chirp was presented a total of 840 times. Response rate was calculated by summing spikes over a 15-ms time window starting at an estimate of the neural latency based on the response to a 73-dB tone at CF (from the response map).

Vowel stimuli were from Hillenbrand et al. (1995), a database that contains English vowels from a variety of speakers. All twelve vowels included in this database were presented, including /iy, ih, ei, eh, ae, ah, aw, oo, uw, er, oa, uh/. The stimuli consisted of 200 ms of the steady-state center portion of the vowels. A 25-ms raised-cosine on/off ramp was applied. Stimuli were presented diotically, at 68 dB SPL, and

with 30 repetitions per vowel. Responses presented here were for the three speakers of ID “M03”, “M40”, and “W39” (designated in the Hillenbrand dataset), who had average fundamental frequencies of 95, 148, and 202 Hz, respectively.

4.3.3 Vowel Component Decomposition

In order to evaluate the chirp cues contained in spoken vowels, the magnitude and phase spectra of the stimuli were estimated, as follows (Yasi, 2004; Ramamurthy and Raghavan, 2013). First, each 200-ms-duration vowel stimulus was divided into four 50-ms segments, and the Welch power spectrum was evaluated for each. The peaks of the power spectrum were identified up to 3.5 kHz—these peak frequencies were the initial estimates of harmonic components. Note that to reject spurious, non-harmonic peaks, power spectrum peaks that differed by less than 65% of average F0 were removed.

Next, a set of bandpass filters was designed to isolate each harmonic. Filters were designed with a 20-Hz passband centered around each component frequency, and a stopband cutoff above and below the component frequency by F0/2. The Matlab function “kaiserord” was used to generate Kaiser-window filter parameters for a finite impulse response (FIR) filter. The Kaiser window was chosen for its linear phase response. The Matlab function “fir1” was used to generate the filter.

The response of the filter to the original signal was approximately the waveform of the isolated harmonic component. Using the Matlab function “fmincon”, the frequency, magnitude, and phase of the harmonic were estimated, using the sum of squared errors as the objective function and a starting value of the frequency based on the power

spectrum peak. To prevent phase optimization from becoming stuck at the bounds, phase was constrained to -$2\pi - 2\pi$, and later assigned equivalent phase between -$\pi - \pi$.

The analysis was repeated for each peak in the power spectrum up to 3.5 kHz to estimate the vowel spectrum. To illustrate the fast frequency chirps within the vowel stimuli, due to phase transitions in the spectra near formants, a synthetic version of the vowel was generated that had a uniform magnitude spectrum and the estimated phase spectrum. To illustrate the within-pitch-period frequency chirps contained within each resulting synthetic vowel, spectrograms were generated using the Matlab function "spectrogram". These were generated using Hamming windowing over 600-sample segments, and allowing for 590 samples of overlap between segments, with an overall sampling rate of 48828 samples/sec. Finally, formant frequencies of the Hillenbrand vowels were identified using Praat analysis.

4.3.4 Response Analysis and Classification

To assess how IC neural responses could distinguish vowels, two classification matrices were constructed based on responses to each speaker (F0), one using average rate and one using temporal information. The average rate in response to a given repetition of one vowel was calculated, and the overall average rates across repetitions in response to each of the 12 vowels was calculated, excluding the current repetition. Each repetition was classified as the vowel for which the absolute difference between the single-repetition rate and overall average rate was minimal.

Vowel classification based on spike timing used a strategy developed by Satuvuori et al. (2017) called rate-independent spike (RIS) distance, a measure of the temporal similarity of two spike trains that is unaffected by average rate. Briefly, the

method involves matching each spike to its closest neighbor in the other spike train. A profile of each spike train is constructed describing temporal similarity to the other train on a sample-by-sample basis. Averaging the two profiles and dividing by the local average rate of the two spike trains yields a final distance estimate (Satuvuori et al., 2017). The RIS distance is a value between 0 and 1 that evaluates similarity solely based on spike timing (Satuvuori and Kreuz, 2018). For classification, the RIS distance for every possible pair of spike trains was calculated (Fig. 4.1). Each repetition was classified as the vowel with the smallest average RIS distance, excluding self-comparisons.

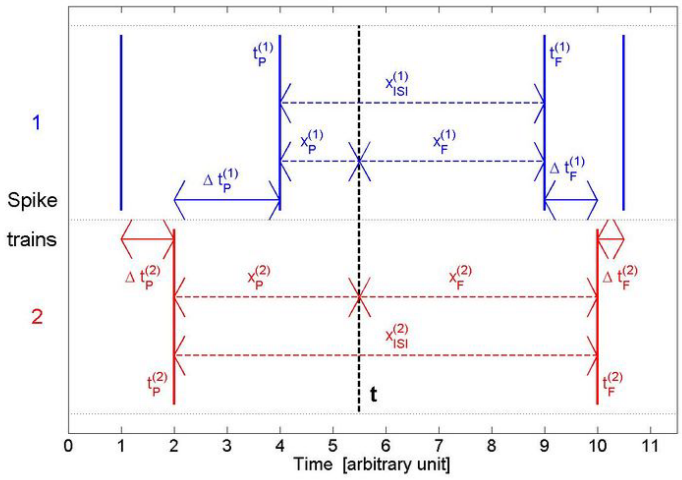

Figure 4.1 – An illustration of the calculations performed in the RI-SPIKE distance metric (Graphic obtained from Kreuz et al., 2012). In the calculation, for each time-sample (indicated by vertical dashed line), several functions describe spike timing on either spike trains. This includes the minimum distance between spikes between the two trains for the previous spike (t_P) and the following spike (t_F), the interval between the current sample and the previous spike (x_P) or the following spike (x_F), and the inter-spike interval at the current time sample (x_{ISI}). Together, these functions form the profile from which the RI-SPIKE distance is calculated.

For both methods, a classification matrix was generated by tallying the classifier results on true vowel vs. predicted vowel axes. Vowel-by-vowel accuracy was assessed using the formula

$$Acc_{Vow} = \frac{TP + TN}{TP + TN + FP + FN}$$

Here, true positives (TP) is the number of correct predictions of the vowel (the square on the diagonal). True negatives (TN) is the number of correct rejections (all other squares on the diagonal). False positives (FP) is the number of incorrect predictions of the vowel (all squares in the vowel's column except the diagonal). False negatives (FN) is the number of incorrect rejections of the vowel (all squares in the vowel's row except the diagonal).

Additionally, the overall accuracy of the classifier can be calculated by dividing the total number of correct predictions (sum of squares on the diagonal) by the total number of incorrect predictions (sum of squares not on the diagonal).

To extract prominent features of the population of RVFs, as in Mitchell et al. (2023), Principal Component Analysis (PCA) was conducted using the Matlab function "PCA". RVFs were normalized by individual peak rates before analysis. Note that principal component analysis performed in this study used a different set of neurons than those used in Mitchell et al. (2023), although some overlapped. Thus, the exact shape of the first three principal components is not the same between the two studies. However, the resulting principal components are similar between studies and the same interpretation can be applied to PCA using both sets of neurons.

Additionally, a metric based on receiver-operating characteristic (ROC) analysis (Egan, 1975) was used to measure the extent of direction selectivity in high-velocity (< 2 kHz/ms) chirp responses, independent of direction. ROC was used to measure the discriminability of chirp direction based on single-chirp repetition response rates for chirps of equivalent speeds (absolute value of velocity) but opposite directions. Calculating the area-under-the-curve of the resulting ROC function gave the direction bias (DB) per velocity pair as a number from 0 to 1, where 0 indicated downward bias, and 1 indicated upward bias. To obtain high-velocity direction bias (HVDB), the DBs of the velocities greater than 2 kHz/ms (3.16, 6.24, and 9.24 kHz/ms) were averaged. Then, HVDB was calculated using $HVDB = 2 * abs(\overline{DB} - 0.5)$, where $\overline{DB}$ is the mean of DBs. HVDB described direction bias, independent of direction, on a scale from 0 to 1, with 1 indicating strong direction bias.

4.3.5 Modeling

The model for IC-cell chirp-sensitivity is detailed in Mitchell and Carney (2024). Briefly, the mechanism of the model depends on a chirp-direction-sensitive inhibition of the IC by the ventral nucleus of the lateral lemniscus (VNLL) originating from octopus cells of the posteroventral cochlear nucleus (PVCN). The model uses a strategy described by Krips and Furst (2009), in which non-homogeneous Poisson processes (NHPPs) representing neural inputs are combined by coincidence detectors (CDs)—the responses of these CDs are also NHPPs (Krips and Furst 2009). Thus, the responses of IC neurons can be modeled starting from auditory-nerve (AN) model inputs by combining them using inhibitory and excitatory CD stages.

In the Mitchell and Carney (2024) model, there are two model stages: an octopus-cell stage and an IC stage. The octopus-cell stage receives two AN inputs of different frequencies, on-CF (the CF of the final IC neuron) and off-CF (OCF). Identical CF inputs together numbering N_{CF} effectively increase input magnitude, such that the CF input alone is suprathreshold. Additionally, both AN inputs elicit a delayed hyperpolarization input that acts like an effective inhibition—hyperpolarization suppresses excitatory output equivalently for both CF and OCF inputs. Thus, a mechanism that detects input sequence emerges: a chirp eliciting the CF input first is likely to result in an excitatory response. Meanwhile, a chirp eliciting the OCF input before the CF input will have the excitatory response suppressed by the coinciding hyperpolarization.

The IC stage receives one on-CF excitatory input, and two inhibitory inputs: one from the octopus-cell model and one relayed from an on-CF AN model. The on-CF excitation and inhibition are delayed by delay parameters d_E and d_I, where the difference between the two tunes the final neuron's MTF in a manner described by the same-frequency inhibition-excitation (SFIE) model (Nelson and Carney, 2004). The output of the IC-model stage has chirp-sensitivity opposite to that of the inhibitory octopus stage, and it has a BE MTF (Mitchell and Carney, 2024).

In order to test the hypothesis that chirp sensitivity impacts vowel classification, six model neurons were created. Three of these model neurons had 2-kHz CFs, and three had 1-kHz CFs, to directly compare the model to similar units for which physiological recordings were available. For both CF groups, one neuron had upward bias, one neuron had downward bias, and one neuron was non-selective (NS). These direction biases resulted from different octopus-cell-stage parameters, or in the case for

NS neurons, no octopus-cell-stage inhibition. The parameters of these model neurons are described in Table 4.1. Note that the upward and downward-biased 1-kHz-CF model neurons are identical to those described in Mitchell and Carney, 2024 (their Fig. 2.4). Additionally, to ensure comparable average rates across vowels between the model neurons and units from physiology, model neuron rate functions were normalized such that their average rates across vowels match (2-kHz-CF model neurons were matched to unit 1, with average vowel rate of 127.03 spk/s; 1-kHz-CF model neurons were matched to unit 4, with average vowel rate of 88.50 spk/s). Rate multipliers are reported in Table 4.1.

Table 4.1 – Parameter values for model neurons. For NS neurons, non-applicable parameters are listed as “N/A”.

CF (kHz)	Direction bias	OCF (kHz)	d_{CF} (ms)	d_{OCF} (ms)	N_{CF}	M_{Oct}	M_I	d_E (ms)	d_I (ms)	Δ_{Oct} (ms)	Δ_I (ms)	Rate Mult.
2	Up	1.00	0	0.45	3	10	10	1.2	3.7	1	1	4.77
2	Down	3.00	1.00	0	3	10	10	1.2	3.7	1	1	4.16
2	NS	N/A	0.30	0	N/A	0	10	1.2	3.7	1	1	2.61
1	Up	0.90	0	0.45	4	12	16	1.4	3.4	1	1	4.80
1	Down	2.21	1.00	0	4	12	8	1.4	3.5	1	1	2.31
1	NS	N/A	N/A	N/A	N/A	0	8	1.4	3.9	1	1	2.37

In order to apply RI-SPIKE distance metric to model neurons, spike times were generated based on rate functions output by the model. For each time sample, a random number was generated (range 0 – 1, uniformly distributed, Matlab’s “rand” function). A spike was generated at that time if the random number was less than the value of the rate function divided by the sampling rate (100,000 samples/sec). Additionally, an absolute refractory period of 1 ms was applied, so that no spikes were generated regardless of the random input.

4.4 Results

To test for the presence of chirps in vowels, center-tokens of Hillenbrand-vowel stimuli were analyzed using component-phase decomposition (Yasi, 2004; Ramamurthy and Raghavan, 2013). Figure 4.1 depicts magnitude spectra (left) and several periods of synthesized vowels (right) produced by this analysis (/aw/, /ih/, and /uw/, all 95-Hz-F0 speaker). For each of the formant frequencies, F1-F3, there is an associated inflection in the vertical white bands of the synthesized vowel, indicating a local chirp (or frequency sweep). This inflection is characterized by positive (upward) chirps just below the formant frequencies and/or negative (downward) chirps just above the formant frequencies—this pattern is consistent with peaks in group delays that have been reported to occur near formant frequencies (Bozkurt et al., 2006; Rajan et al., 2013). Note that a synthetic vowel with zero-phase harmonics would appear as a series of vertical lines in this visualization, reflecting no chirp features.

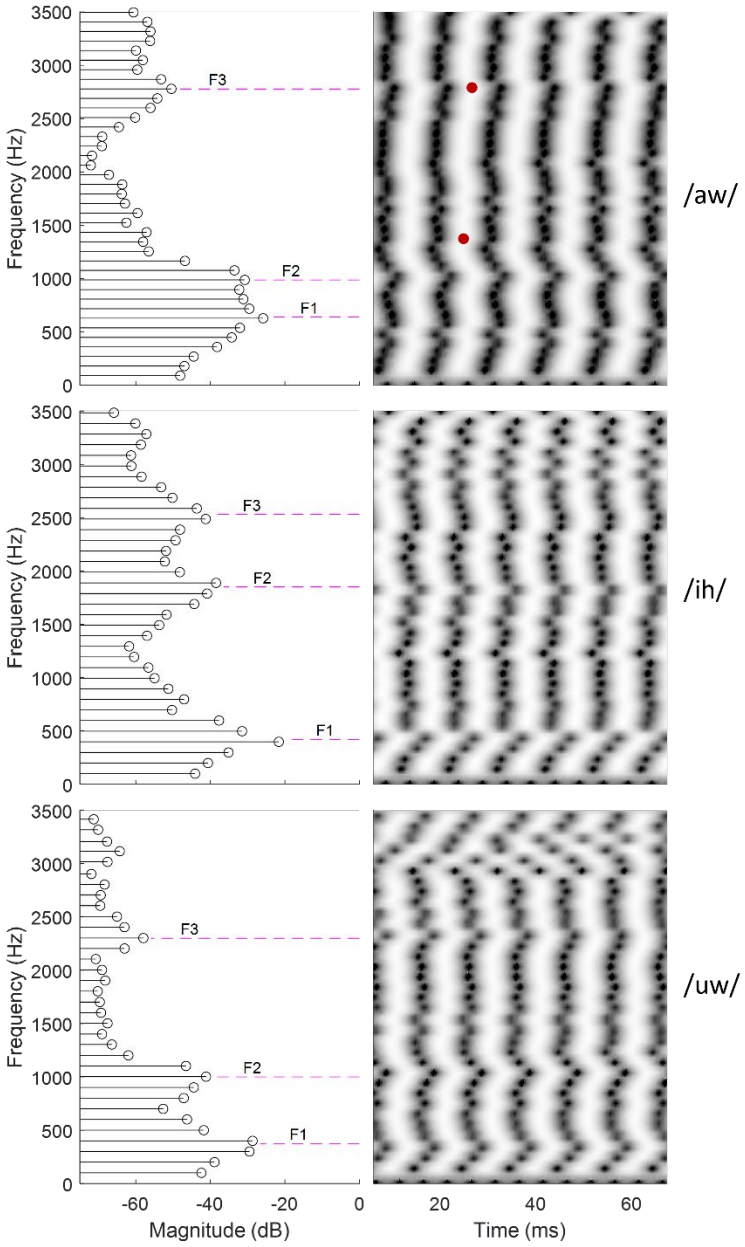

Figure 4.2 – Component-phase decomposition analysis for three representative vowels, /aw/, /ih/ and /uw/, from a speaker with 95 Hz F0. Left—magnitude spectra of natural vowel tokens resulting from vowel component decomposition. Magenta dashed lines indicate formant frequencies F1, F2, and F3 (matched by frequency to the corresponding synthetic vowel). Right—synthesized vowel resulting from component decomposition. The magnitude of these synthesized vowels were flat. Red dots indicate example temporal amplitude peaks at frequencies 1.38 kHz and 2.81 kHz (see text).

Chirps associated with these inflections are low-velocity and span a limited frequency range when compared to chirping stimuli such as Schroder complexes and aperiodic chirp stimuli. However, the frequency regions between formant inflections include chirps more comparable to these stimuli. The cyan dots on the top spectrogram in Fig. 4.2 mark temporal amplitude peaks at two frequencies, 1.38 kHz and 2.81 kHz, between F2 and F3 in the vowel /aw/. There is a time difference of 1.80 ms between these points, and a frequency difference of 1.43 kHz, making the local velocity of their spectrotemporal difference 0.79 kHz/ms: a comparable, albeit slightly slower, velocity to chirps presented for RVFs (which had a range of 0.80—9.24 kHz/ms). However, note that local velocities vary between F2 and F3, sometimes changing directions and thus approaching near-vertical slopes. Therefore, 0.79 kHz/ms reflects the average chirp velocity between these two frequencies, with local chirp velocities exceeding this average and approaching infinity.

Much like the representative vowels in Fig. 4.2, all vowel stimuli have chirp inflections at formant frequencies, and faster chirps between formants with equivalent velocities to those used in RVFs. These spectrotemporal features represent a rich set of cues to which chirp-sensitive neurons may be sensitive. To guide later comparisons of these features to neural CFs, Table 4.2 indicates the frequencies of F1 and F2 for the set of 95-Hz-F0 vowels (identified via Praat analysis).

Table 4.2 – Frequencies of F1 and F2 for 95-Hz-F0 vowel stimuli.

Vowel	IY	IH	EI	EH	AE	AH	AW	OO	UW	ER	OA	UH
F1 (Hz)	372	425	410	563	638	771	666	487	402	449	460	610
F2 (Hz)	2305	1880	2100	1658	1788	1216	1033	1190	1025	1560	812	903

Response properties of individual IC neurons toward sound features such as frequency, AM, and spectrotemoral chirps are diverse; their responses to vowel stimuli are expected to depend on a combination of these feature sensitivities. Figure 4.3 shows the response profiles of five example neurons—CF, MTF, RVF, and vowel rate profiles, alongside classification matrices based on average rate and timing. These neurons are useful case studies to illustrate how response properties may contribute to vowel classification.

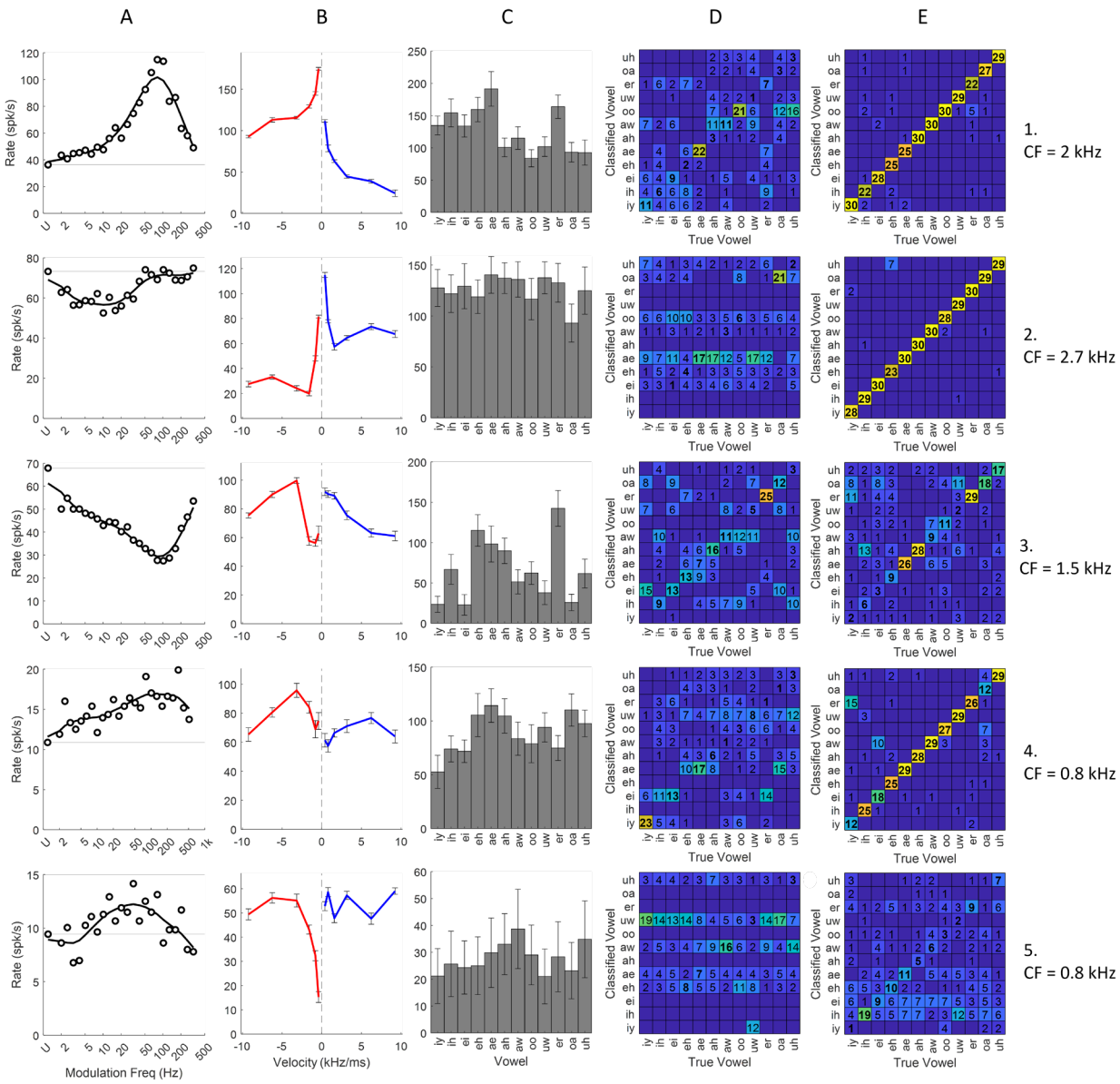

Figure 4.3 – Response profiles of five representative neurons with CFs in the typical F1-F2 range (< 3 kHz). Column A: MTFs based on AM noise. U on the x-axis indicates rate in response to unmodulated noise. Solid black curve depicts smoothed MTF based on raw data points. Column B: RVFs, blue indicates upward chirp response, red indicates downward chirp response. Error bars depict standard errors. Vertical dashed gray line indicates 0 velocity. Column C: Average rate histograms in response to 30 repetitions of twelve vowels, 95 Hz F0, taken from Hillenbrand et al. (1995) database. Error bars indicate standard deviation. Column D and E: Classification matrices using average rate (Column D) and timing (Column E) metrics. Number of classifications is indicated by boxed numbers; boxes with no number had no classifications. True positive

classifications along unity line are indicated with bold text. Classification matrices are based on 30 repetitions of 95-Hz-F0 vowel stimuli.

CFs of the selected neurons are below 3 kHz, around the frequency range of F1 and F2. Column A of Fig. 4.3 contains MTFs—Neuron 1, 4 and 5 are BE, and Neuron 2 and 3 are BS. Column B of Figure 4.3 shows RVFs; Neuron 1's RVF has a consistent and strong bias toward downward chirps, while Neuron 2 has consistently upward-selective RVF. Neuron 3's RVF has mixed chirp-direction sensitivity, flipping direction bias at 2 kHz/ms. Finally, Neurons 4 and 5 are examples of identical CFs but opposite RVF direction-bias, downward (Neuron 4) and upward (Neuron 5). These example neurons are representative of the population in that chirp-direction bias is not more common among neurons of a specific MTF shape or CF. In all cases, these response properties contribute to neural coding of vowels. Vowel response rates are influenced by proximity of CF to formant frequencies. BE neurons encode F0-related amplitude fluctuations differently than BS neurons, as well as neural fluctuations related to formant frequencies. Lastly, given the presence of chirps similar to those in Fig. 4.2, an assortment of chirp cues may also influence coding.

To assess the performance of vowel discrimination in IC neurons, classification matrices were generated for each speaker-F0 and for two classification metrics, average rate and timing. The classification matrices in Fig. 4.3 are based on a 95-Hz-F0 speaker. In Fig. 4.3, column D, classification matrices based on average rate are shown for the example neurons. These grids show the number of single-repetition responses that were classified as one of twelve vowels, versus the true vowel used to elicit that response. Classifier performance can be assessed by looking at true positives, which are accurate, positive predictions of vowels (denoted using bold text). In Fig. 4.3, column C, the

average rate profiles are shown, which were used to produce the rate classification matrices. Generally, the rate classifier does well those vowels that are rate extrema—for instance, Neuron 3 has particularly high accuracy classifying the vowel /er/ due to its high average response rate. Neuron 4 accurately classifies the vowel /iy/ due to its low average response rate. Average rate classification is comparatively inaccurate for those vowels that are not extrema.

In Fig. 4.3, column E, classification matrices based on timing are shown, specifically the rate-independent SPIKE distance (Satuvuori et al., 2017). Timing classification generally outperforms average rate, with individual accuracy reaching 100% (all 30 repetitions correctly classified) for many vowels. These classification matrices indicate that for many neurons (Neuron 1, 2, 4) a vowel coding scheme based on timing would be able to correctly classify all twelve vowels presented here.

Classification matrices can be summarized using overall accuracy, which sums the total number of correct classifications (bold text) and divides by the total number of classifications. Using overall accuracy, the performance of the timing-based classifier of Neurons 1, 2 and 4 (overall accuracy of 90.8%, 95.8%, and 80.3%, respectively) is higher than the rate-based classifier (overall accuracy of 27.5%, 15.0%, and 21.1%, respectively). Even for neurons where timing classification overall accuracy is low (Neuron 2 and 5, 44.4% and 22.8%, respectively), it outperforms rate classification (31.7% and 10.3%). Overall classification is a useful metric to distill the performance of a classifier to one metric; however, note that overall accuracies can be expected to be quite low relative to vowel-specific accuracies. Given twelve outcomes, a classifier operating at chance would have an overall accuracy of 8.3%. In comparison to chance, a

classifier able to reliable identify certain vowels like Neuron 3, Column E, still has overall accuracy much greater than chance (44.4%).

By evaluating classification performance per neuron with a single metric, overall accuracy, systematic trends of classification performance versus RVF features can begin to be identified. Principal component analysis was performed on the RVFs of the subset of neurons with vowel responses to identify their dimensions of highest variability. Principal components 1 through 3 (PC1, PC2, and PC3) are shown in Figure 4.4 (leftmost column). Together, PC1, PC2 and PC3 explained 95.0% of the RVF variance.

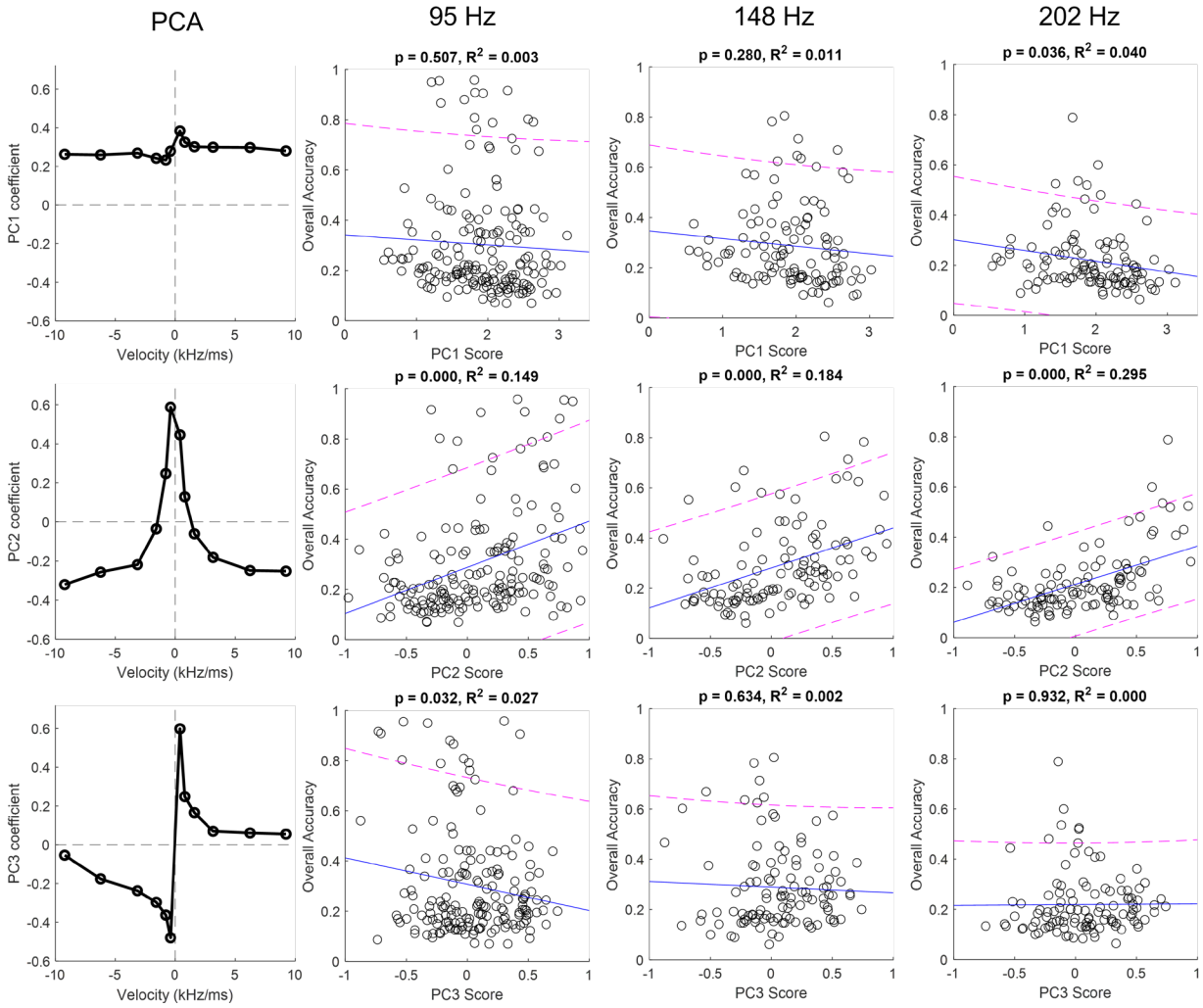

Figure 4.4 – Results of principal component analysis (PCA) on neural RVFs, and subsequent scatter plots and regression analysis comparing overall accuracy based on timing to principal components scores. Leftmost column: Principal components 1, 2 and 3, plotted on loadings-versus-velocity axes. Dashed lines mark zero on both axes to assist PCA interpretation. Scatter plots show overall accuracy versus principal component scores for individual neurons, organized by speaker-F0 (columns) and principal component (rows). Blue regression lines reflect data trend—magenta dashed lines indicate 95% confidence intervals. For each fit, a p-value and R^2 is reported; p-values are the result of ANOVA, testing the hypothesis that the slope coefficient of the linear fit is non-zero. R^2 reflects variance explained by the linear fit.

PC1 (Fig. 4.4, left column, top row) reflected an overall average rate feature, explaining 88.2% of the RVF variance, suggesting that the most prominent difference between RVFs was absolute rate. All loadings for PC1 having positive values, indicating

that all rates vary together in PC1. PC2 (Fig. 4.4, left column, middle row), showed a tendency for low-velocity rates to covary, and inversely vary with high-velocity rates. In other words, neurons with high PC2 tend to exhibit high rates in response to low-velocity chirps and low rates in response to high-velocity chirps, regardless of direction. PC2 explained 4.4% of RVF variance. Finally, PC3 (Fig. 4.4, left column, bottom row) reflects RVF direction-sensitivity, especially at low velocities. Neurons with high PC3 are positive-chirp biased. PC3 explains 2.5% of RVF variance. It can be seen from these results that RVF rates vary in absolute velocity—that is, velocity of chirp irrespective of direction—more than chirp direction.

Quantifying principal component scores for each neuron, trends between RVF features and overall accuracy can be established. In Fig. 4.4, scatter plots showing overall accuracy versus principal component scores are shown, separated by vowel speaker (F0, indicated by column). For each, a regression line was fit to the data, the quality of fit described by a p-value (ANOVA, testing the null hypothesis that linear slope coefficient ≠ 0) and R^2 (coefficient of determination, describing variance explained). Most notably, PC2 is correlated with overall accuracy for all speakers. P-values were consistently < 0.001—furthermore, R^2 indicate 14.9%, 18.4%, and 29.5% variance explained for vowel accuracies with speakers 95, 148, and 202 Hz respectively. In comparison to PC2, PC1 and PC3 show little to no correlation with overall accuracy. However, note that PC1 vs. 202-Hz vowel accuracy and PC3 vs. 95-Hz vowel accuracy have significant p-values, using a significance level of 0.05.

The overall accuracy values in Fig. 4.4 were taken from classification based on timing. In comparison, there were no notable trends in overall accuracy based on average rate versus principal-component scores.

Based on the data in Fig. 4.4, high chirp-direction bias at low chirp-velocities (PC3) has no relationship to overall accuracy—on the contrary, neurons for which low chirp-velocities elicit high response rates irrespective of direction (i.e. have high PC2) tend to have higher overall accuracy. However, these principal components do not directly assess the impact of high-velocity chirp direction bias on accuracy; thus, an alternative metric is necessary. High-velocity direction bias metric, described in detail in Methods section 4.2.4, quantifies the average direction bias in the response rates to the three fastest velocity pairs, ±3.16, ±6.24, and ±9.24 kHz/ms. Figure 4.5 summarizes trends of overall accuracy versus high-velocity direction bias, based on both rate and timing.

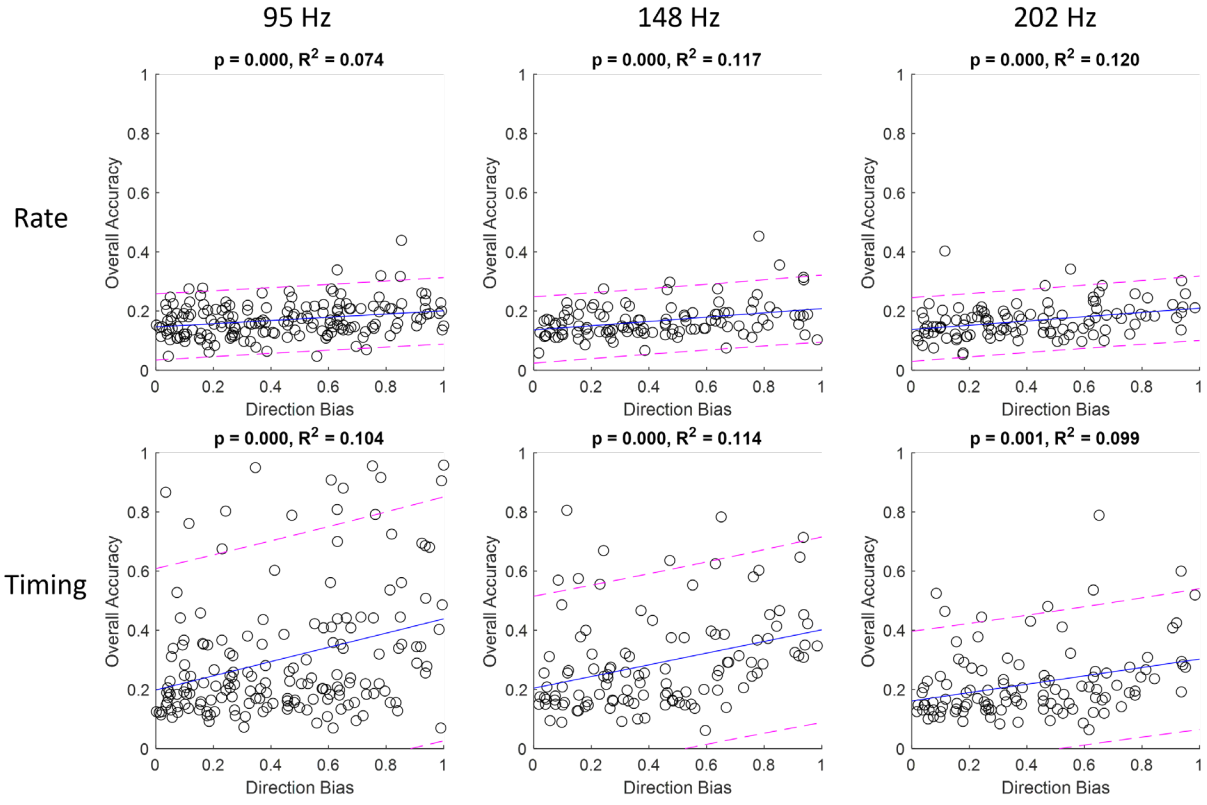

Figure 4.5 – Scatter plots and regression analysis of overall accuracy (based on rate and timing classification) versus high-velocity direction bias. Plots are organized by speaker-F0 (column) and classification metric (row). Blue regression lines reflect data trend—

magenta dashed lines indicate 95% confidence intervals. For each fit, a p-value and R^2 is reported; p-values are the result of ANOVA, testing the hypothesis that the slope coefficient of the linear fit is non-zero. R^2 reflects variance explained by the linear fit.

The trend of high-velocity direction bias versus overall accuracy based on rate is positive, validated by p-value < 0.001 for all speakers (Fig. 4.5, top row). However, the amount of variance explained by this metric is low, at 7.4%, 11.7%, and 12.0% for 95 Hz, 148 Hz, and 202 Hz speakers, respectively. Similarly, high-velocity direction bias is positively correlated with overall accuracy based on timing (Fig. 4.5, bottom row); p-values < 0.01 for all speakers. Similar to rate-based accuracy, the percentage of variance explained by timing-based accuracy is low, at 10.4%, 11.4%, and 9.9% for 95 Hz, 148 Hz, and 202 Hz speakers, respectively. However, note the difference in range between rate- and timing-based overall accuracy; rate-based accuracy is consistently low, whereas some individual neurons have near-100% overall accuracy based on timing.

Vowel classification was assessed for chirp-sensitivity model IC cells (Mitchell and Carney, 2024) to assess if they aligned with trends observed in the individual (Fig. 4.3) and population (Fig. 4.4-5) physiological data. In Figure 4.6, the response profile and classification matrices are shown for three model IC neurons with 2 kHz CF, one biased towards upward chirps (top row), one biased towards downward chirps (middle row), and one without chirp direction bias (non-selective, NS) (bottom row). Model parameters are given in Methods section 4.2.5.

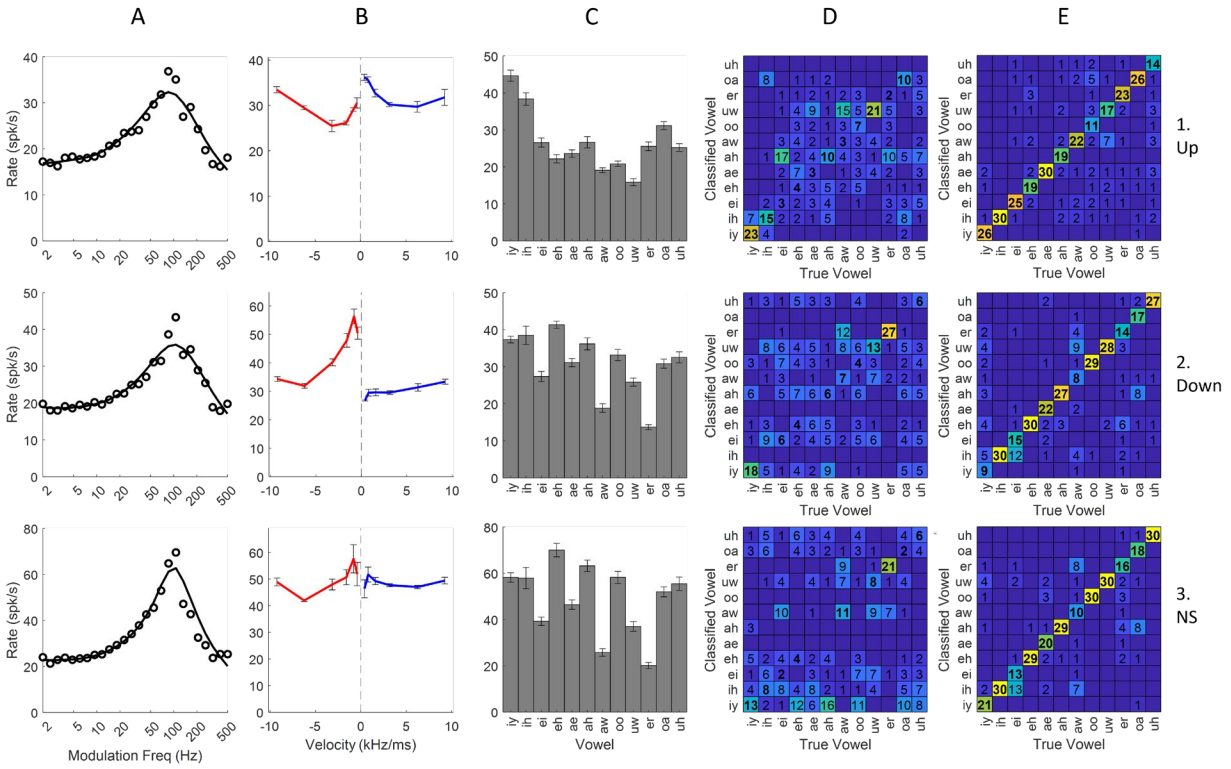

Figure 4.6 – Response profiles of three model neurons, all with 2-kHz CF. Neurons differ by octopus-cell inhibition; in Row 1, responses of an upward-biased model neuron receiving downward-biased octopus-cell inhibition are shown. In Row 2, responses of a downward-biased model neuron receiving upward-biased octopus-cell inhibition are shown. Row 3 shows responses of a non-selective (NS) model neuron that received no octopus-cell inhibition. Stimulus responses are otherwise organized identically to Figure 4.3.

Due to a limitation of the model, the simulated neurons' MTFs are all BE, with peaks around 100 Hz modulation frequency (Fig. 4.6, Column A). This matches the F0 of vowel stimuli used to generate the pictured classification matrices (95 Hz). Simulated neurons' RVFs (Fig. 4.6, Column B) differed solely in the presence or nature of octopus-cell-stage inhibition—the upward and downward biased neurons received inhibition from oppositely-biased octopus-cell model cells. The NS neuron (bottom row) instead received no inhibition from the octopus-cell stage.

Vowel rate histograms (Fig. 4.6, Column C) depict the average rate profile of the neurons' response to vowels. The downward-biased neuron's vowel rate profile (Row 2) is similar to the NS neuron's (Row 3); relatively, the upward-biased neuron's profile (Row 1) has larger differences to the NS neuron's, with rate minima and maxima changing completely, among other dissimilarities. Notably, average rate across vowels differs by neuron. In order to compare vowel classification between model neurons, and also to compare these model neurons to a similar unit from physiology (Fig. 4.3, Row 1), model rate functions were multiplied by a normalizing factor such that the average vowel rate for all three model neurons was equivalent to this physiology unit (values reported in Methods).

Average rate-based classification matrices are shown in Fig. 4.6, Column D, generated using responses to 95-Hz-F0 vowels. Vowels with the most true positives tend to be extrema in the rate profiles. Timing-based classification matrices are also shown (Fig. 4.6, Column E). Between model neurons, specific vowel classification differed in accuracy—in rate-based classification, /er/ is identified reliably for the downward-biased and NS neurons, whereas the upward-biased neuron identifies /iy/ and /uw/ reliably instead. Using timing, accuracy of classification for some vowels such as /aw/ improves in the upward-biased neuron compared to the downward-biased or NS neurons.

Overall accuracies for rate-based classification (28.1%, 25.3%, and 20.9% for upward, downward, and NS neurons, respectively) remain low compared to timing-based classification (72.8%, 71.1%, and 76.7% for upward, downward, and NS neurons, respectively) for these model neurons. In comparison to NS, the addition of octopus-cell inhibition benefited rate-based classification, but decreased overall accuracy for timing-

based classification. However, note that vowel-by-vowel accuracy did not always conform to these trends in overall accuracy.

In Figure 4.7, response profiles and classification matrices are shown for three model IC neurons with 1 kHz CF: upward-biased (top row), downward-biased (middle row), and NS (bottom row). For this group, average vowel rates were normalized by that of Neuron 4 in Fig. 4.3.

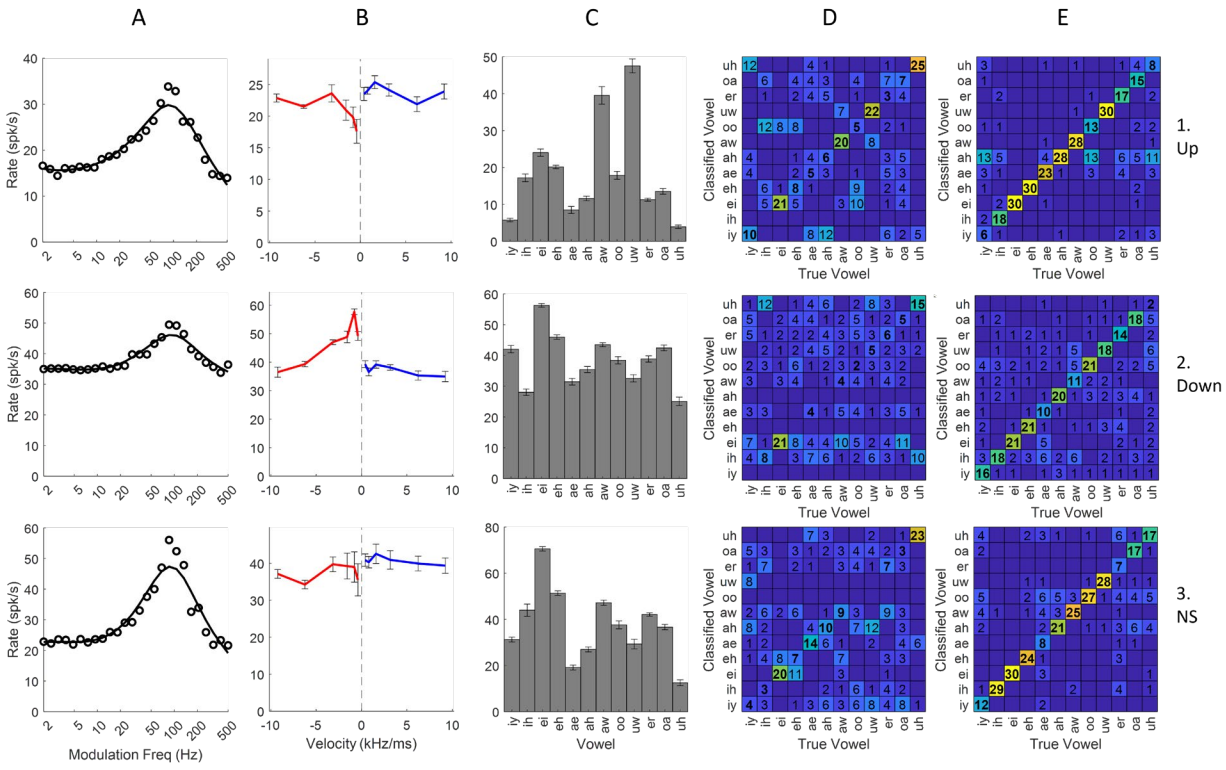

Figure 4.7 - Response profiles of three model neurons, all with 1-kHz CF. Neurons differ by octopus-cell inhibition and RVF bias (upward, downward, and NS are shown on the top, middle, and bottom rows, respectively). Stimulus responses are otherwise organized identically to Figure 4.3.

All model neuron MTFs are BE (Fig. 4.7, Column A). Owing to the lower CF (1 kHz), the direction bias in the model RVFs (Fig. 4.7, Column B) is not as pronounced as for the 2-kHz-CF neurons depicted in Fig. 4.6. Nonetheless, the three model neurons

display markedly different vowel rate profiles (Column C), despite their similar MTFs and identical CFs. In the rate-based classification matrices (Column D), the vowel rate profiles again predict classification accuracy. Note that all model neurons display high classification accuracy for /ei/ and /uh/, but the upward-biased neuron also shows high accuracy for /aw/ and /uw/. Timing-based classification matrices (Column E) show vowel-by-vowel accuracy differences as well, /ae/, /aw/ and /uh/ are particularly prominent examples.

4.5 Discussion

In this paper, we tested the hypothesis that chirp sensitivity of IC neurons influence neural discriminability of natural vowel tokens. Using a strategy of component decomposition, we identified chirps within vowel stimuli of a velocity range relevant to IC chirp velocity sensitive neurons. Examining confusion matrices generated based on rate and timing metrics, vowel classification based on recordings made of IC units in physiology was not readily predicted by RVF alone. Also, vowel classification based on timing was consistently better than classification based on rate. A population-level analysis revealed a relationship of overall accuracy based on timing to Principal Component 2 (PC2), a chirp-directionless feature of RVFs related to a rate-bias toward low velocities, opposed to high velocities. No relationship between low-velocity chirp direction bias (PC3) was found with overall accuracy; however, high-velocity chirp bias was found to have a statistically significant relationship to overall accuracy based on rate and timing. Finally, a chirp-velocity sensitive model was used to examine the vowel classification performance of model cells with and without octopus-cell inhibition. Vowel rate profiles differed between otherwise identical model neurons receiving different

octopus-cell inhibition. Velocity sensitivity impacts neural response profiles across natural vowel tokens, though further work is needed to relate model responses to examples in physiology.

Another method of interrogating the impact of chirp-sensitivity on vowel representation is to directly compare model neuron results. In Fig. 4.5, featuring 2-kHz-CF model neurons, the vowel rate histogram (Column C) appears to differ more greatly between upward and NS neurons than downward and NS neurons. This is corroborated by the rate-based classification matrices (Fig. 4.5, Column D), which feature different salient vowels, as well as by the timing-based classification matrices (Fig. 4.6, Column E), in which the upward neuron displays markedly different classification accuracies than downward or NS for the vowels /aw/, /oo/, and /ei/, among others. It is interesting to note that in Fig. 4.1, it was demonstrated that the vowel /aw/ (95 Hz F0) contained an upward chirp between 1.38 kHz and 2.81 kHz. It is possible that the presence of this cue led to an elevated classification based on timing for the upward-biased, 2-kHz-CF neuron. It is nonetheless clear that different chirp-sensitivities have an effect on vowel-by-vowel classification accuracy.

A question this work attempts to answer is whether fast spectrotemporal chirps contained in vowels, resulting from the phase differences between harmonics due to the vocal-tract filter or group-delay functions, contributes to a representation of formant frequencies of natural vowels in the midbrain. Results of analysis on vowel responses of individual neurons in physiology alone are inconclusive, due to a lack of controlled comparisons to similar neurons that differ by direction bias. Ideally, by introducing responses from model neurons, comparisons to such physiological units should be possible. For example, the CF of physiological neuron 1 (Fig. 4.2) is 2 kHz, MTF is BE,

and its RVF is downward biased—we would expect its vowel rate histogram to be similar to a neuron with similar response properties. However, comparing this unit's vowel rate histogram to that of model neuron 2 in Fig. 4.5, it is apparent that the profiles do not match. Similar lack of correspondence can be seen between physiology neuron 4 (Fig. 4.2) and 1-kHz model neuron 2 (Fig. 4.6), and physiology neuron 5 (Fig. 4.2) and 1-kHz model neuron 1 (Fig. 4.6). One would expect basic response properties of CF and MTF shape to drive vowel rate profiles, for instance, proximity of CF to formant frequencies would drive higher response rates. The fact that physiology and model neurons that share these properties differ greatly between vowel rate profiles, regardless of chirp-velocity sensitivity, suggest that there are additional factors unaccounted for in the model.

Of the two RVF features displaying trends with overall accuracy, PC2 reflected a chirp-directionless property, instead describing an RVF shape that has high rates in response to low chirp velocities, and comparatively low rates in response to high chirp velocities. Note that while individual neurons with high PC2 scores may still be direction-biased for some velocities, this suggests that absolute velocity (i.e. speed) of chirp is more indicative of vowel accuracy than direction bias. It is interesting to consider that the shape outlined by PC2 would match the RVF of a hypothetical neuron responding only to energy at CF—chirps that take a long time to pass through CF (i.e. low velocity chirps) would elicit relatively higher rates than faster (high velocity) chirps. In the aperiodic chirp stimulus upon which RVFs are constructed, energy between chirps of different velocities is normalized by adjusting chirp amplitude; however, this normalization may not completely counteract this effect. However, the effect of high-velocity direction bias on overall accuracy (Fig. 4.4) should not be understated, suggesting that direction bias

remains an important aspect of RVFs to consider. It is clear that a future version of the model should add ways to more precisely control directionless chirp-velocity and high-velocity direction bias of model neurons, so as to better reflect these trends in physiology.

Bibliography

Aitkin, L. M., & Phillips, S. C. (1984). Is the inferior colliculus and obligatory relay in the cat auditory system? Neuroscience Letters, 44(3), 259-264. https://doi.org/https://doi.org/10.1016/0304-3940(84)90032-6

Bozkurt, B., Couvreur, L., & Dutoit, T. (2007). Chirp group delay analysis of speech signals. Speech Communication, 49(3), 159-176. https://doi.org/https://doi.org/10.1016/j.specom.2006.12.004

Carney, L. H. (2024). Neural Fluctuation Contrast as a Code for Complex Sounds: The Role and Control of Peripheral Nonlinearities. Hearing research, 443, 108966. https://doi.org/https://doi.org/10.1016/j.heares.2024.108966

Carney, L. H., Li, T., & McDonough, J. M. (2015). Speech Coding in the Brain: Representation of Vowel Formants by Midbrain Neurons Tuned to Sound Fluctuations. Eneuro, 2(4), ENEURO.0004-0015.2015. https://doi.org/10.1523/ENEURO.0004-15.2015

Delgutte, B., & Kiang, N. Y. S. (1984). Speech coding in the auditory nerve: I. Vowel-like sounds. The Journal of the Acoustical Society of America, 75(3), 866-878. https://doi.org/10.1121/1.390596

Delgutte, B., & Kiang, N. Y. S. (1984). Speech coding in the auditory nerve: V. Vowels in background noise. The Journal of the Acoustical Society of America, 75(3), 908-918. https://doi.org/10.1121/1.390537

Egan, J. P., & Egan, J. P. (1975). Signal detection theory and ROC-analysis. Academic press.

Fant, G. (1971). Acoustic theory of speech production: with calculations based on X-ray studies of Russian articulations. Walter de Gruyter.

Henry, K., Wang, Y., Abrams, K., & Carney, L. H. (2022). Mechanisms of masking by Schroeder-phase harmonic tone complexes in the budgerigar. The Journal of the Acoustical Society of America, 151(4), A122-A122. https://doi.org/10.1121/10.0010849

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. The Journal of the Acoustical Society of America, 97(5), 3099-3111. https://doi.org/10.1121/1.411872

Joris, P. X., Schreiner, C. E., & Rees, A. (2004). Neural Processing of Amplitude-Modulated Sounds. Physiological reviews, 84(2), 541-577. https://doi.org/10.1152/physrev.00029.2003

Kewley-Port, D., Burkle, T. Z., & Lee, J. H. (2007). Contribution of consonant versus vowel information to sentence intelligibility for young normal-hearing and elderly hearing-impaired listenersa). The Journal of the Acoustical Society of America, 122(4), 2365-2375. https://doi.org/10.1121/1.2773986

Kim, D. O., Carney, L., & Kuwada, S. (2020). Amplitude modulation transfer functions reveal opposing populations within both the inferior colliculus and medial geniculate body. Journal of Neurophysiology, 124(4), 1198-1215. https://doi.org/10.1152/jn.00279.2020

Kreuz, T. (2012). SPIKE-distance. Scholarpedia, 7(12), 30652.

Krips, R., & Furst, M. (2009). Stochastic Properties of Coincidence-Detector Neural Cells. Neural computation, 21(9), 2524-2553. https://doi.org/10.1162/neco.2009.07-07-563

Krishna, B. S., & Semple, M. N. (2000). Auditory Temporal Processing: Responses to Sinusoidally Amplitude-Modulated Tones in the Inferior Colliculus. Journal of Neurophysiology, 84(1), 255-273. https://doi.org/10.1152/jn.2000.84.1.255

Langner, G. (1992). Periodicity coding in the auditory system. Hearing research, 60(2), 115-142. https://doi.org/https://doi.org/10.1016/0378-5955(92)90015-F

Mitchell, P. W., Henry, K. S., & Carney, L. H. (2023). Sensitivity to direction and velocity of fast frequency chirps in the inferior colliculus of awake rabbit. Hearing research, 440, 108915. https://doi.org/https://doi.org/10.1016/j.heares.2023.108915

Mitchell, P. W. & Carney, L. H. (2024, in press). A Computational Model of Auditory Chirp-Velocity Sensitivity and Amplitude-Modulation Tuning in Inferior Colliculus Neurons. The Journal of Computational Neuroscience.

Nelson, P. C., & Carney, L. H. (2004). A phenomenological model of peripheral and central neural responses to amplitude-modulated tones. The Journal of the Acoustical Society of America, 116(4), 2173-2186. https://doi.org/10.1121/1.1784442

Rajan, P., Kinnunen, T. H., Pohjalainen, J., Alku, P., Bimbot, F., Cerisara, C., Fougeron, C., Gravier, G., Lamel, L., & Pellegrino, F. (2013). Using group delay functions from all-pole models for speaker recognition.

Ramamurthy, S., & Raghavan, M. (2013, 8-10 Oct. 2013). Filter design for synthesis of musical notes: A multidimensional feature-based approach. 2013 IEEE International Conference on Signal and Image Processing Applications,

Sachs, M. B., & Young, E. D. (1979). Encoding of steady-state vowels in the auditory nerve: representation in terms of discharge rate. The Journal of the Acoustical Society of America, 66(2), 470-479.

Satuvuori, E., & Kreuz, T. (2018). Which spike train distance is most suitable for distinguishing rate and temporal coding? Journal of Neuroscience Methods, 299, 22-33. https://doi.org/https://doi.org/10.1016/j.jneumeth.2018.02.009

Satuvuori, E., Mulansky, M., Bozanic, N., Malvestio, I., Zeldenrust, F., Lenk, K., & Kreuz, T. (2017). Measures of spike train synchrony for data with multiple time scales. Journal of Neuroscience Methods, 287, 25-38. https://doi.org/https://doi.org/10.1016/j.jneumeth.2017.05.028

Schreiner, C. E., & Langner, G. (1997). Laminar fine structure of frequency organization in auditory midbrain. Nature, 388(6640), 383-386. https://doi.org/10.1038/41106

Schwarz, D. M., Zilany, M. S. A., Skevington, M., Huang, N. J., Flynn, B. C., & Carney, L. H. (2012). Semi-supervised spike sorting using pattern matching and a scaled Mahalanobis distance metric. Journal of Neuroscience Methods, 206(2), 120-131. https://doi.org/https://doi.org/10.1016/j.jneumeth.2012.02.013

Steenken, F., Oetjen, H., Beutelmann, R., Carney, L. H., Koeppl, C., & Klump, G. M. (2022). Neural processing and perception of Schroeder-phase harmonic tone complexes in the gerbil: Relating single-unit neurophysiology to behavior. European Journal of Neuroscience, 56(3), 4060-4085. https://doi.org/https://doi.org/10.1111/ejn.15744

Yasi, J. (2004). An Algorithm for Extracting the Relative Phase of Harmonics from a Periodic Digital Signal. unpublished)< http://online. physics. uiuc. edu/courses/phys 199pom/NSF_REU_Reports/2004_reu/Joe_ Yasi_Final_Paper. pdf.

Young, E. D., & Sachs, M. B. (1979). Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers. The Journal of the Acoustical Society of America, 66(5), 1381-1403.

Chapter 5: Summary and Discussion

This thesis quantified the prevalence and nature of chirp direction and velocity sensitivity in the IC, proposed and demonstrated the feasibility of a novel computation model with chirp sensitivity, and examined the implications of chirp sensitivity to IC vowel coding. Throughout this research, we tested the hypothesis that chirp cues contained within complex sounds like speech represent cues that impact IC responses to these sounds. Evidence that chirp sensitivity is a widespread, independent sound-feature sensitivity was identified in physiological recordings, which made use of a novel stimulus and direction-bias analysis technique. A computational model was designed, including a physiologically-plausible mechanism for IC chirp sensitivity, while building upon and preserving features of existing models. Finally, the impact of chirp sensitivity on midbrain representation of vowels was explored, demonstrating a significant relationship between directionless velocity sensitivity and vowel classification accuracy.

5.1 Summary and Novel Results

Chapter 2 details the development of original experimental aims from preliminary extracellular IC recordings in response to SCHR stimuli. The chapter outlines the nature of these recordings and the indications that they represented a sound-feature sensitivity to chirp direction and velocity. The aims of the physiological experiments are discussed, primarily the necessity to separate sensitivity to chirps from the periodicity cue provided by SCHR stimuli. Then, a novel stimulus to interrogate this aim is presented, the aperiodic chirp stimulus, alongside a novel characterizing technique in the RVF, which illustrates chirp response rate as a function of velocity. Through the use of GLM, MTF was found to explain a larger percentage of explainable variance in SCHR responses

than RVFs, although RVFs alone explained chirp direction selectivity in SCHR responses. Additionally, ROC analysis was used to quantify chirp selectivity towards a particular direction—this analysis showed that the majority of neurons surveyed showed direction selectivity in response to chirps of at least one speed. Finally, PCA was used to identify a divide in the population of RVFs at 2 kHz—RVFs tend to have similar response rates to low-velocity (< 2 kHz) chirps regardless of direction, and also similar response rates to high-velocity (> 2 kHz) chirps regardless of direction.

Chapter 3 presents a computational model of IC neuron chirp-sensitivity alongside AM tuning. The chapter begins with a review of potential mechanisms that might serve as the origin of IC chirp sensitivity, and then identifies inhibition of the IC by octopus cells of the PVCN (via the VNLL) as a promising answer. The chapter continues by detailing sequence detection, the mechanism by which octopus cells are sensitive to chirp velocity. The most notable result of this chapter is the model itself. First, an explanation is given of the modeling strategy, first proposed by Krips and Furst (2009). The chosen strategy allows for simulation of and hypothesis testing of many model architectures due to its modular nature—the strategy allows for the simulation of neurons in the central nervous system by preserving the statistics of inputs and outputs, which are NHPPs. Each neural stage is a coincidence detector, receiving some combination of excitatory and inhibitory inputs. Importantly, this modeling strategy allows for a model organization that generally controlled chirp-sensitivity and AM tuning through separate inhibitory inputs to the IC stage. Sequence detection in octopus cells is implemented using this framework and serves as one inhibitory input to the IC. The IC also received a delayed inhibition of the same frequency as excitation, a mechanism shown to produce BE MTFs (Nelson and Carney, 2004). The final outcome of this work is a model capable

of accurately simultaneously simulating diverse chirp-sensitivity observed in physiological work through the use of different model parameters. Example model neurons are given mirroring the CF range of neurons recorded from in physiology. Finally, the impact of parameter modulation on model responses is explored.

Chapter 4 combines physiology and modeling techniques in the previous two studies to analyze the impact of chirp sensitivity on IC representation of vowels. The chapter begins with a brief overview of the need for alternative vowel coding mechanisms beyond rate or TFS in the AN. The chapter also highlights that the IC is a promising location for speech coding due to its diverse sound-feature sensitivities, as well as its unique capability to take advantage of nonlinearities of the periphery. Then, the possibility of chirp cues for formant frequencies existing within vowels is presented, with the presence of these related to the phase inflections resulting from vocal tract resonances, also described as group-delay peaks. The first outcome of this study demonstrated that vowels, originally from the Hillenbrand (1995) database, do contain chirp cues that plausibly could encode location of formant frequencies. Secondly, results of vowel classification analysis in physiological units were shown—this classification was either based on average rate or timing. Timing was universally a more accurate classifier than rate. Furthermore, classification of physiological units were compared to similar model neurons with varied chirp-sensitivities. Although a clear relationship between vowel accuracy and RVF was not observed for these model neurons, vowel-by-vowel classifications as well as average rate profiles changed depending on RVF direction bias. Finally, statistically-significant positive correlations were identified between directionless chirp-velocity sensitivity in the RVF and overall timing classification

accuracy, as well as high-velocity direction bias and overall classification accuracy (both rate and timing).

5.2 Future Work

Data recorded in multiple species (Morrison et al., 2018; Henry et al., 2022; Steenken et al., 2022; Mitchell et al., 2023) show that IC neurons possess remarkable selectivity to chirp direction and velocity in responses to laboratory stimuli. The research included in this thesis demonstrates chirp sensitivity has the potential to greatly impact IC responses, and at minimum represents a necessary aspect to include in the pursuit of a comprehensive understanding of IC physiology. In this section, we present future avenues of research that might extend the findings of this work towards broader applications in auditory neuroscience.

5.2.1 Alternative chirp-mechanisms for modeling

In Chapter 4, directionless velocity-sensitivity in the RVFs in physiological recordings was found to be significantly predictive of overall vowel classification accuracy based on timing. Specifically, this RVF feature was based on PCA, and indicated that neurons responded with similar rates to low-velocity chirps, regardless of direction, as well as similar rates to high-velocity chirps, regardless of direction. Additionally, RVFs that showed high response rates to low-velocity chirps and low response rates to high-velocity chirps were associated with high overall vowel-classification accuracy.

These results were difficult to corroborate using the chirp model. While differences did exist between model IC neurons with upward, downward, and non-

selective RVFs, a systematic relationship between chirp direction and overall accuracy was not apparent. Furthermore, attempts to directly interrogate the effect of directionless velocity-sensitivity on classification accuracy of model neurons were ineffectual. Also, while the chirp model could demonstrably produce RVFs of either upward or downward direction bias for a large range of CFs, all types of velocity sensitivity were not necessarily possible at every CF—for instance, chirp direction bias at high velocities became harder to achieve with low CFs (evident in example model neuron responses given in Chapter 3 and Chapter 4). This limitation made direct analysis of high-velocity direction bias difficult, another observed trend in the physiological data of vowel classification accuracy.

Altogether, these limitations suggest a blind spot in the mechanism of the model, perhaps rooted in its conception. From the outset, the goal of the model was to accurately portray direction and velocity sensitivity to chirps—in practice, the results of modeling work were framed in a way that emphasized feasibility of direction bias over velocity sensitivity per se. Based on the evidence presented in Chapter 4, the chirp model must be updated to better control directionless velocity sensitivity and high-velocity direction bias in order to more meaningfully examine the impact of chirp sensitivity to vowel coding.

While octopus-cell inhibition alone is probably not sufficient to explain directionless velocity sensitivity, chirp sensitivity in IC neurons is recognized to likely originate from multiple different mechanisms combining in variable ways between specific neurons, among which octopus-cell inhibition remains a well-suited candidate (Pollak et al., 2011; Pollak et al., 2013). As such, one possible improvement upon the existing chirp model might be to introduce an additional off-CF input in the octopus-cell

stage, as suggested in Chapter 4. Two off-CF input of equivalent frequency-difference might impart a symmetry in the RVF that potentially would result in a direction-insensitive, velocity-sensitive IC neuron.

Of remaining theories on formation of chirp-sensitivity in the IC, they can broadly be divided into two theories (Pollak et al., 2013): that the exact timing of excitation and inhibition shapes chirp-sensitivity (Covey and Casseday, 1999; Fuzessery and Hall, 1996; Gordon and O'Neill, 2000), or that the overall magnitudes of excitation versus inhibition are more important (Gittelman et al., 2009; Gittelman and Pollak, 2011). Fortunately, given the flexible nature of the Krips and Furst modeling framework, future models seeking to incorporate these mechanisms likely would not have to start anew. For instance, theories concerning timing generally argue that the presence of an off-CF inhibitory field result in asymmetrical responses to different chirp directions (Brimijoin et al., 2005; Andoni et al., 2007; Kuo and Wu, 2012). The Krips and Furst framework employed in the chirp model described in this work is well-suited to the task of proposing additional off-CF inhibitory inputs. Additionally, such an input may help control high-velocity chirp sensitivity, due to its ability to impart direction bias on IC RVFs more directly than can be achieved using octopus-cell inhibition (which is comparatively more indirect). Another timing-based mechanism, early on-CF inhibition, may help a future model explain directionless chirp-velocity sensitivity (Fuzessery et al., 2006; Fuzessery et al., 2011). This mechanism proposes a "fast-pass" chirp filter, in which chirps passing through the neuron's CF too slowly are inhibited while faster chirps are not (Fuzessery et al., 2006). A future model including this mechanism would be capable of directly assessing the impact of RVF PC2 on vowel classification accuracy.

Theories concerning differences in magnitude between excitation and inhibition take a different perspective on the role of timing, pointing out that timing differences that affect the interaction between excitatory and inhibitory post-synaptic potentials (EPSPs and IPSPs) can ultimately change the response amplitude, which may be more important to the target cell than timing would normally be expected to (Gittelman and Pollak, 2011). Once again, the model framework presented is appropriate for such a mechanism; after all, the sequence detection mechanism in the octopus-cell stage is also based on how timing of EPSPs can result in either suprathreshold or subthreshold response. In future work, an IC chirp model that includes some or all of these alternative mechanisms will represent a more precise tool to examine the precise impact of certain RVF characteristics on vowel responses than the current model.

5.2.2 Alternative vowel stimuli for increased phase-control

In Chapter 4, we examined the responses of both physiological and model IC units to real speech stimuli, originally from the Hillenbrand (1995) dataset. While this approach had advantages, including ensuring the most realistic possible stimulus was presented, there were a number of experimental controls lacking that became apparent afterwards. For instance, following the visualization and quantification of the chirp-cues surrounding formant frequencies in vowel stimuli, it became desirable to directly alter those cues, potentially by adjusting relative phase of harmonic components. One path future work might take is the use of synthetic vowels as a primary experimental stimulus.

Use of synthetic vowels to specifically examine the impact of phase has previously been explored in psychophysical studies to evaluate the impact of hearing loss on vowel identification (Leek et al., 1987; Traunmüller, 1987). Other studies have

manipulated vowel phase to assess formant phase contrast on vowel identification (Molis et al., 2013) or to design a stimulus suitable for investigating speech encoding using the auditory brainstem response (ABR) (“peaky speech”; Polonenko and Maddox, 2021). Such phase manipulations are particularly of interest in the context of responses by chirp-sensitive IC neurons. A future study adopting similar stimuli may be able to directly test the hypothesis that chirp-sensitive neurons respond to vowel phase.

5.2.3 Chirp sensitivity in IC neuron types based on molecular markers

In the studies presented in this thesis, we classified neurons solely based on the nature of their responses to characterizing stimuli—chiefly CF, MTF type, and RVF. The long-term nature of our electrode implantation methods made spatial analysis of neural populations possible, but difficult, with inexactitude of repeated implantations and purported tetrode depths contributing to this difficulty (although, at least one study (Morrison et al., 2018) reports a relationship between IC CF and FM responses features).

Recently work has focused on identifying neuron types of the IC based on molecular markers and circuit characteristics (Goyer et al., 2019; Silveira et al., 2020; Drotos and Roberts, 2024). In particular, three molecular markers (vasoactive intestinal peptide, VIP; Goyer et al., 2019; neuropeptide Y, NPY; Silveira et al., 2020; excitatory cholecystokinin, CCKE; Kreeger et al., 2021) have emerged as relating to different IC neuron types. These types are hypothesized to illuminate circuit motifs in the IC that have heretofore been inaccessible. Future work is currently underway to identify how these neuron types may respond to chirps (Drotos and Roberts, 2024). The completion

of this work would enable the potential identification of chirp-sensitivity subtypes and focus modeling efforts toward neural chirp mechanisms supported by circuit motifs.

5.2.4 Chirp-sensitivity in Hearing Loss

The most common public health concern related to hearing is hearing loss—typically, noise-related damage in the cochlea or its synapses results in increased audiometric thresholds and increased difficulty understanding speech, especially in noise (Baer et al., 1993; Makary et al., 2011). Notably, age-related hearing loss is a prevalent pathology only in part related to traumatic noise exposure (Makary et al., 2011). Development of novel therapies and technologies that might assist people with hearing loss retain functional speech comprehension is a priority for the betterment of public health.

Evidence suggests a link between hearing loss and detection thresholds of tones in SCHR maskers—listeners with normal hearing may have better tone-detection thresholds in -SCHR maskers than +SCHR, whereas listeners with hearing loss had identical thresholds (Summers, 2000). Additionally, in a study that utilized synthetic vowels with varied phase contrast at formants, normal-hearing and hearing-impaired listeners demonstrated ability to identify vowels solely based on phase-shifts, although normal-hearing listeners were more sensitive to minor phase differences (Molis et al., 2013). If sensitivity to chirps and phase can be conceptualized as aspects of the same cue, it is possible that chirp-sensitivity is degraded by hearing loss. By this logic, future hearing aids may be able alter the phase-spectra of incoming vowels or other complex sounds (such as music), exaggerating the presence of chirp cues, which may then be more discriminable by chirp-sensitive neurons of the IC. Such a strategy might

eventually be employed by advanced machine-learning signal-processing algorithms that aim to compensate for hearing loss in a practical and computationally-efficient way realistic towards the processing-power available to hearing-assistive devices (Leer et al., 2024).

Bibliography

Andoni, S., Li, N., & Pollak, G. D. (2007). Spectrotemporal Receptive Fields in the Inferior Colliculus Revealing Selectivity for Spectral Motion in Conspecific Vocalizations. The Journal of Neuroscience, 27(18), 4882. https://doi.org/10.1523/JNEUROSCI.4342-06.2007

Baer, T., Moore, B. C., & Gatehouse, S. (1993). Spectral contrast enhancement of speech in noise for listeners with sensorineural hearing impairment: effects on intelligibility, quality, and response times. J Rehabil Res Dev, 30(1), 49-72.

Brimijoin, W. O., & O'Neill, W. E. (2005). On the prediction of sweep rate and directional selectivity for FM sounds from two-tone interactions in the inferior colliculus. Hearing research, 210(1), 63-79. https://doi.org/https://doi.org/10.1016/j.heares.2005.07.005

Casseday, J. H., Ehrlich, D., & Covey, E. (1994). Neural Tuning for Sound Duration: Role of Inhibitory Mechanisms in the Inferior Colliculus. Science, 264(5160), 847-850. https://doi.org/10.1126/science.8171341

Drotos, A. C., & Roberts, M. T. (2024). Identifying neuron types and circuit mechanisms in the auditory midbrain. Hearing research, 442, 108938. https://doi.org/https://doi.org/10.1016/j.heares.2023.108938

Fuzessery, Z. M., & Hall, J. C. (1996). Role of GABA in shaping frequency tuning and creating FM sweep selectivity in the inferior colliculus. Journal of Neurophysiology, 76(2), 1059-1073. https://doi.org/10.1152/jn.1996.76.2.1059

Fuzessery, Z. M., Razak, K. A., & Williams, A. J. (2011). Multiple mechanisms shape selectivity for FM sweep rate and direction in the pallid bat inferior colliculus and auditory cortex. Journal of Comparative Physiology A, 197(5), 615-623. https://doi.org/10.1007/s00359-010-0554-0

Fuzessery, Z. M., Richardson, M. D., & Coburn, M. S. (2006). Neural Mechanisms Underlying Selectivity for the Rate and Direction of Frequency-Modulated Sweeps in the Inferior Colliculus of the Pallid Bat. Journal of Neurophysiology, 96(3), 1320-1336. https://doi.org/10.1152/jn.00021.2006

Gittelman, J. X., Li, N., & Pollak, G. D. (2009). Mechanisms Underlying Directional Selectivity for Frequency-Modulated Sweeps in the Inferior Colliculus Revealed by

In Vivo Whole-Cell Recordings. The Journal of Neuroscience, 29(41), 13030-13041. https://doi.org/10.1523/jneurosci.2477-09.2009

Gittelman, J. X., & Pollak, G. D. (2011). It's About Time: How Input Timing Is Used and Not Used To Create Emergent Properties in the Auditory System. The Journal of Neuroscience, 31(7), 2576. https://doi.org/10.1523/JNEUROSCI.5112-10.2011

Gordon, M., & O'Neill, W. E. (1998). Temporal processing across frequency channels by FM selective auditory neurons can account for FM rate selectivity. Hearing research, 122(1), 97-108. https://doi.org/https://doi.org/10.1016/S0378-5955(98)00087-2

Goyer, D., Silveira, M. A., George, A. P., Beebe, N. L., Edelbrock, R. M., Malinski, P. T., Schofield, B. R., & Roberts, M. T. (2019). A novel class of inferior colliculus principal neurons labeled in vasoactive intestinal peptide-Cre mice. eLife, 8, e43770. https://doi.org/10.7554/eLife.43770

Henry, K., Wang, Y., Abrams, K., & Carney, L. H. (2022). Mechanisms of masking by Schroeder-phase harmonic tone complexes in the budgerigar. The Journal of the Acoustical Society of America, 151(4), A122-A122. https://doi.org/10.1121/10.0010849

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. The Journal of the Acoustical Society of America, 97(5), 3099-3111. https://doi.org/10.1121/1.411872

Kreeger, L. J., Connelly, C. J., Mehta, P., Zemelman, B. V., & Golding, N. L. (2021). Excitatory cholecystokinin neurons of the midbrain integrate diverse temporal responses and drive auditory thalamic subdomains [Article]. Proceedings of the National Academy of Sciences of the United States of America, 118(10), Article e2007724118. https://doi.org/10.1073/pnas.2007724118

Kuo, Richard I., & Wu, Guangying K. (2012). The Generation of Direction Selectivity in the Auditory System. Neuron, 73(5), 1016-1027. https://doi.org/10.1016/j.neuron.2011.11.035

Leek, M. R., Dorman, M. F., & Summerfield, Q. (1987). Minimum spectral contrast for vowel identification by normal-hearing and hearing-impaired listeners. The Journal of the Acoustical Society of America, 81(1), 148-154. https://doi.org/10.1121/1.395024

Leer, P., Jensen, J., Tan, Z. H., Østergaard, J., & Bramsløw, L. (2024). How to Train Your Ears: Auditory-Model Emulation for Large-Dynamic-Range Inputs and Mild-to-Severe Hearing Losses. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 32, 2006-2020. https://doi.org/10.1109/TASLP.2024.3378099

Makary, C. A., Shin, J., Kujawa, S. G., Liberman, M. C., & Merchant, S. N. (2011). Age-Related Primary Cochlear Neuronal Degeneration in Human Temporal Bones. Journal of the Association for Research in Otolaryngology, 12(6), 711-717. https://doi.org/10.1007/s10162-011-0283-2

Mitchell, P. W., Henry, K. S., & Carney, L. H. (2023). Sensitivity to direction and velocity of fast frequency chirps in the inferior colliculus of awake rabbit. Hearing research, 440, 108915. https://doi.org/https://doi.org/10.1016/j.heares.2023.108915

Mitchell, P. W. & Carney, L. H. (2024, in press). A Computational Model of Auditory Chirp-Velocity Sensitivity and Amplitude-Modulation Tuning in Inferior Colliculus Neurons. The Journal of Computational Neuroscience.

Molis, M. R., Diedesch, A., Gallun, F., & Leek, M. R. (2013). Vowel Identification by Amplitude and Phase Contrast. Journal of the Association for Research in Otolaryngology, 14(1), 125-137. https://doi.org/10.1007/s10162-012-0352-1

Morrison, J. A., Valdizón-Rodríguez, R., Goldreich, D., & Faure, P. A. (2018). Tuning for rate and duration of frequency-modulated sweeps in the mammalian inferior colliculus. Journal of Neurophysiology, 120(3), 985-997. https://doi.org/10.1152/jn.00065.2018

Pollak, G. D. (2013). The dominant role of inhibition in creating response selectivities for communication calls in the brainstem auditory system. Hearing research, 305, 86-101. https://doi.org/https://doi.org/10.1016/j.heares.2013.03.001

Pollak, G. D., Gittelman, J. X., Li, N., & Xie, R. (2011). Inhibitory projections from the ventral nucleus of the lateral lemniscus and superior paraolivary nucleus create directional selectivity of frequency modulations in the inferior colliculus: A comparison of bats with other mammals. Hearing research, 273(1), 134-144. https://doi.org/https://doi.org/10.1016/j.heares.2010.03.083

Polonenko, M. J., & Maddox, R. K. (2021). Exposing distinct subcortical components of the auditory brainstem response evoked by continuous naturalistic speech. eLife, 10, e62329. https://doi.org/10.7554/eLife.62329

Silveira, M. A., Anair, J. D., Beebe, N. L., Mirjalili, P., Schofield, B. R., & Roberts, M. T. (2020). Neuropeptide Y Expression Defines a Novel Class of GABAergic Projection Neuron in the Inferior Colliculus. The Journal of Neuroscience, 40(24), 4685. https://doi.org/10.1523/JNEUROSCI.0420-20.2020

Steenken, F., Oetjen, H., Beutelmann, R., Carney, L. H., Koeppl, C., & Klump, G. M. (2022). Neural processing and perception of Schroeder-phase harmonic tone complexes in the gerbil: Relating single-unit neurophysiology to behavior. European Journal of Neuroscience, 56(3), 4060-4085. https://doi.org/https://doi.org/10.1111/ejn.15744

Traunmüller, H. (1987). Phase Vowels. In M. E. H. Schouten (Ed.), The Psychophysics of Speech Perception (pp. 377-384). Springer Netherlands. https://doi.org/10.1007/978-94-009-3629-4_30