Mechanisms that Underlie Poorer Binaural Outcomes in Patients with Asymmetrical Hearing and Bilateral Cochlear Implants

by

Sean R. Anderson

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Communication Sciences and Disorders)

at the UNIVERSITY OF WISCONSIN-MADISON

2022

Date of final oral examination: 12/15/2021

The dissertation is approved by the following members of the Final Oral Committee:
Ruth Y. Litovsky, Professor, Communication Sciences and Disorders
Phillip Smith, Professor, Department of Neuroscience
Matthew I. Banks, Professor, Department of Neuroscience
Benjamin Parrell, Assistant Professor, Communication Sciences and Disorders
Mathias Dietz, Professor, Medical Physics, Universität Oldenburg

Acknowledgements

There are a great many people to whom I am indebted for their wisdom, guidance, compassion, and love, and I will not be able to do them justice here. I would like to thank Drs. Ruth Litovsky and Matthew Goupell most of all for their mentorship in hearing science and for taking a chance on me as a student. I would also like to thank my many other mentors in science over the years, including Drs. Lina Reiss, Donata Oertel, Alan Kan, Mathias Dietz, Sara Duran, Yonghee Oh, Kim Nickerson, Yi Ting Huang, Peter Hoekje, Jackie Morris, Brian Thomas, Andrew Mickley, and Chris Turner. I would like to thank my labmates and peers who helped make these projects possible with their thoughtful feedback, patience as I grew as a student, and help, including Shelly Godar, Tanvi Thakkar, Jasenia Hartman, Stephen Dennison, Emily Burg, Ellen Peng, Ela Warnecke, Xin Zhou, Lukas Suveg, Jonas Klug, Helen Heinermann, and Coral Dirks. I would like to thank my mentees for their wisdom and patience as I learned to teach, including Caroline Swords, Emily Bott, Gammon Gresham, Jiwon Kim, Chris Naber, and Skyler DeMorett. I would like to thank the great many people who showed me kindness with their interest and guidance during my training, including Drs. Monita Chatterjee and David Landsberger. I would like to thank my friends and family for their support through the challenges that presented themselves during graduate school, especially my grandmother, mother, Elizabeth, Emily, Luna, and the friends I made at Summit Avenue Cooperative and Atlas Improv Company.

Finally, I would like to thank the Department of Communication Sciences and Disorders as well as the University of Wisconsin-Madison for the training and funding provided to me during my studies. This work was supported by NIH-NIDCD F31 DC018483-01A1 awarded to myself, NIH-NIDCD R01-DC003083 awarded to Ruth Litovsky, and NIH-NICHD U54HD090256 awarded to the Waisman Center.

Table of Contents

Abstract

Bilateral cochlear implants (BiCIs) result in several benefits, including improvements in speech understanding in noise and sound source localization. However, the amount of benefit received by each patient varies considerably. One reason for this variability is difference between the two ears' hearing function, i.e., interaural asymmetry. Thus far, investigations of interaural asymmetry have been highly specialized within various topic areas. The goal of this dissertation is to provide a new framework under which these studies can be integrated synergistically. This framework begins with a binaural signal, ends with a behavioral response, and consists of two interim stages: encoding and decoding of binaural information. Encoding of binaural cues is represented using excitation-inhibition (EI) of signals from the left ear and right ear according to where the signal originated in space. Decoding of binaural cues is represented using predictive coding, which assumes that expectations based on prior sensory experience and transmitted via top-down connections are compared against sensory input (i.e., bottom-up connections). This new framework can be applied to physiological, behavioral, and modeling studies, resulting in a clearer understanding of the implications of interaural asymmetry and improved strategies for optimizing patient interventions.

Chapter 1: Introduction and Statement of Problem

I. INTRODUCTION

In the United States, approximately 37.6% of listeners with hearing loss have asymmetrical thresholds favoring one ear [1]. However, the vast majority of research (e.g., human psychophysics, electrophysiology, and computational modeling) on binaural hearing assumes interaural symmetry. This is concerning, particularly as the number of individuals with hearing loss continues to increase [2], as do the number of individuals who receive cochlear implants (CIs) [3].

It is current clinical practice to preserve bilateral sound input, increasing the number of patients who receive bilateral CIs (BiCIs) [4,5]. These patients are advised to receive BiCIs because they improve sound source localization and speech understanding in noise relative to unilateral implantation (e.g., [6–8]). Due to factors associated with hearing loss (e.g., deterioration of the auditory periphery) and implantation (e.g., electrode placement), interaural asymmetry is pervasive and its sources are heterogeneous. Recent research has identified that interaural asymmetry limits the benefits of BiCIs (e.g., [9–12]) and contributes to difficulty navigating complex sound environments (e.g., [10,13–15]). Most studies addressing interaural asymmetry focused on identifying its sources and proposing how to counteract them, but the language and point of view in different topic areas can vary widely, making it difficult to come to a consensus about how to optimize patient outcomes.

Finally, it is important to recognize that auditory processing derives from two interdependent channels: the left and right ear. Historically, the binaural hearing literature has assumed that inputs from the two ears to binaural processors act independently, and that binaural hearing is therefore optimal because of their independence [16]. In other words, that information one ear can be ignored at will. However, information from both ears is inextricably linked by the auditory system, such that processing of information from each ear is not truly independent, and recent evidence has shown that listening with two rather than one ear in specific conditions (e.g., when both ears are spectro-temporally degraded) can impair task performance in listeners with normal-hearing (NH) [17–22] and hearing loss [10,14,15,23–27]. Understanding that processing between the ears is interdependent and recognizing the pervasiveness of interaural asymmetry are both essential for devising strategies to improve patient outcomes.

The goal of the present chapter is to review the literature concerning interaural asymmetry for listeners with BiCIs in order to present a conceptual framework that integrates the findings from a wide variety of studies and methodologies, including human psychophysics and animal electrophysiology. Because the literature concerning interaural asymmetry has focused on specific causes (e.g., unilateral stimulation during development) or manifestations (i.e., interaural mismatch in electrode placement) of asymmetries, the framework and previous literature are discussed in parallel. This allows appropriate context to be applied by the reader. Section I.B presents the overarching framework. Sections II and III present its components.

A. Definitions

Before proceeding further, it is important to provide several operational definitions. The phrase “interaural asymmetry” is meant to refer to *any* stimulus-independent difference in the representation or perception of sounds between the left and right ear. Examples of interaural asymmetry include differences in hearing thresholds, differences in loudness growth, differences in speech understanding scores, differences in place-of-stimulation of the auditory nerve in patients with CIs, and differences in the spectro-temporal pattern conveyed by the auditory nerve. While this review focuses primarily upon listeners with BiCIs, individuals who experience hearing loss and receive bilateral hearing aids, a unilateral CI and hearing aids, or hybrid CIs are also likely to be impacted by these kinds of asymmetries. Thus, this review is focused on characterizing the impacts of many different kinds of interaural asymmetry, creating a framework that is flexible and applicable to all of them, and that can be generalized to other clinical populations. Interaural asymmetry does not represent a singular state, or a point on a continuum between "symmetric" and "asymmetric." Instead, it represents a collection of states or continua that contribute to a measurable outcome: differences between the ears.

Binaural benefits are perceptual advantages provided to listeners when they have access to two ears instead of one. These are usually encapsulated under improved sound source localization in the horizontal plane and increased speech understanding in noise when there is a spatial separation between target and masking sounds compared to when they are co-located (i.e., spatial release from masking). There are some monaural components involved with sound source localization (loudness cues and spectral cues) and improvements associated with speech understanding in noise (head shadow). Binaural benefits include binaural redundancy or summation, which results from having two copies of a signal. Some experiments focus on binaural squelch, which is the benefit attained by having a second copy of the masker when the target and masker are spatially separated in addition to head shadow and binaural summation. Benefits of bilateral hearing are expected to be primarily driven by binaural cues (e.g., [28,29]). There are two primary binaural cues: interaural level and time differences (ILDs and ITDs, respectively). These cues result as a function of changes in sound as they reach the ears according to their horizontal locations in space. Moreover, these cues are conveyed in the temporal fine-structure and envelope of a sound. Temporal fine-structure refers to fast, instantaneous changes in pressure over time, and the temporal envelope refers to more gradual pressure changes over time.

Spatial cues (i.e., ITDs and ILDs) are a subset of cues available to distinguish sound sources from one another in complex auditory environments. Throughout this dissertation “cues” refer to information conveyed in a sound that may or may not be well-represented by the auditory system, but are suspected to be useful for distinguishing between sound sources. In contrast, we will also discuss features, which refer to the internal representation or perception of sounds that can be used to distinguish between sound sources. Examples of sound features related to spatial cues are perceived location and perceived width of a sound image. Thus, cues refer to aspects of a signal (i.e., external to the listener), and features will refer to perceptual qualities (i.e., internal perceptions), both of which are intimately related and involved with segregation of sound sources in complex environments. These ideas will be discussed in greater context in the sections that follow.

B. Conceptual framework for binaural processing

Figure 1 shows the conceptual framework applied throughout this chapter. It is based upon the model used by Jens Blauert [30]. This framework applies to behavioral studies in humans or other animals, as the input is a binaural signal, and the output is a behavioral response. The first motivation for this particular framework is its simplicity. All perceptual processes require several basic components: an encoding stage where cues in the physical stimuli are represented by the nervous system, and a decoding stage where the representations of those stimuli are integrated into a broader perceptual context. Degraded or sub-optimal processing at either of these stages could therefore limit outcomes for listeners. We focus primarily upon encoding of binaural cues and decoding of more general auditory cues (e.g., monaural spectro-temporal fluctuations). Decoding could also be applied to multisensory integration, and though that is not addressed in the present review, we desired to use a model of decoding that was compatible with multi-sensory integration. Inspiration was drawn from the distinction between energetic and informational masking (for review, see [31]). Energetic masking can be attributed to impairments in task performance due to physical obstruction of cues in the stimulus and is therefore related to encoding. Informational masking can be attributed to confusability of target/masker stimuli and is therefore related to decoding. The specific component models of encoding and decoding appropriate for listeners with BiCIs are discussed in the Section II and III.

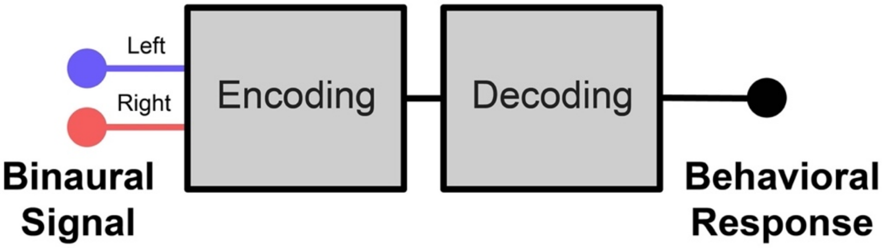


Fig. 1. A conceptual framework for addressing auditory processing. The x-axis corresponds to time. The endpoints represent signal onset and behavioral response offset, respectively.

Inspiration for this framework was also drawn from historical perspectives related to a typically developing ear advantage (i.e., hemispheric specialization) and hemispheric neglect that occurs after a stroke. Listeners with BiCIs who experience extreme interaural asymmetries show some evidence of hemispheric neglect similar to that observed in patients who experience strokes (e.g., auditory "extinction," or the ignoring of input from one ear when both ears are stimulated simultaneously [13,15]). Unlike the binaural hearing literature, the right ear advantage (i.e., auditory hemispheric specialization) literature has assumed and assessed asymmetries in function between the sides of the body. Two competing hypotheses were proposed related to hemispheric specialization in the literature, referred to throughout as structural and attentional hypotheses. The structural hypothesis supposes that hemispheric preferences and ear advantage occur as a result of bottom-up sensory processing due to stronger connections from one side of the body to the cerebral structures that process this information [32]. The attentional hypothesis supposes that hemispheric preferences occur as a result of an imbalance of the attentional network, preferring one side of the cortex [33]. These two competing hypotheses nicely parallel the encoding and decoding stages within the conceptual framework proposed here and are discussed in greater depth in Section II and III, respectively. Critically, due to the different sources of interaural asymmetry, some patients may be more strongly affected by decrements in encoding of auditory information while others are more strongly affected by attention. In other words, problems with both encoding and decoding are suspected to play a role in patient outcomes.

II. MODEL FOR ENCODING: THE LATERAL SUPERIOR OLIVE

A. Motivation and justification

Several principles were established to distinguish encoding from decoding in Fig. 1:

1. Binaural encoding should occur in early stages of auditory processing.
2. Output of the encoder should correspond directly with the stimulus parameters being encoded.
3. The binaural encoder should be limited by the input to the auditory system (i.e., monaural input from the auditory periphery).

4. Performance on behavioral tasks must be separable from problems with the attentional system. Therefore, limitations of the encoder should be indicated using simple behavioral tasks that are trivially taxing to the attentional system.
5. Reliable or consistent bias in the encoder could be corrected by plasticity during decoding, so long as there is sufficient experience and feedback.

The circuits thought to encode binaural cues in the NH system are well-characterized (for review, see [34]) and meet the criteria described above. However, there are differences between listeners with BiCIs and those with NH that warrant a distinct model of encoding. Firstly, most commercially available CI processing strategies discard temporal fine-structure [35], making it impossible to use binaural cues conveyed there. Recent measurements evaluating the output of CI processors confirm that only ILDs and ITDs in the temporal envelope are preserved [36]. Secondly, even when temporal fine-structure is available using research devices, listeners with BiCIs do not seem to gain benefits from fine-structure ITDs [37–39], and localization seems to be dominated by ILDs [40–42]. Thirdly, most CI arrays only reach the basal portion of the cochlea [43]. Fourthly, stimulation provided by CIs is made up of transient electrical pulses, which are modulated in amplitude corresponding to the envelope of the sound being represented [35]. Listeners with BiCIs show reliable evidence of “rate limitations,” meaning that there is a decline in envelope ITD sensitivity with pulse rates (i.e., fine-structure) or amplitude modulation (AM) rates above 300 Hz [44–46].

One circuit in the auditory brainstem seems particularly appropriate for representing binaural cues for listeners with BiCIs: The lateral superior olive (LSO; Fig. 2). The LSO is the second and third synapse from the ipsilateral and contralateral sides of the brainstem, respectively, indicating that it occurs early in auditory processing (Principle 1). The LSO is thought to have evolved to encode for processing of ITDs in transient sounds and perform pre-processing for ILD coding in the inferior colliculus [47,48]. Accordingly, it shows sensitivity to ITDs conveyed by broadband transients [49,50] and temporal envelope [48,51] (Principle 2). The LSO receives excitatory and inhibitory input from the ipsilateral and contralateral ear, respectively [52] and acts as an anti-coincidence detector, responding with the lowest spike rates at 0-µs ITDs, and showing ITD-dependent rates of firing as the ITD is increased or decreased [49] (Principle 2). Time-constants in the LSO circuit are some of the fastest in the central nervous system [48,53], like the medial superior olive that is suspected to code for fine-structure ITDs [54]. Most cells in the LSO have high characteristic frequencies [55–57] (Principle 2). High-frequency cells in the LSO show a low-pass characteristic [58–60] for envelope ITD sensitivity, consistent with the rate limitations observed in listeners with BiCIs, as well as those observed in listeners with NH listening to high-frequency transients [61] (Principle 2 and 3). These perceptual rate limitations are also present for monaural temporal discrimination for listeners with CIs [9,62,63]. Monaural temporal sensitivity is predictive of binaural sensitivity at the same rates [9] suggesting a common mechanism that may limit binaural sensitivity (Principle 4) that parallels processing in the LSO (Principle 2). The LSO is also sensitive to changes in ILDs, with ILDs favoring the ipsilateral and contralateral ear resulting in greater and lesser spike rates, respectively [52,64] (Principle 2). Thus, issues associated with deafness or CI stimulation that affect LSO processing likely result in changes to the encoding of ITDs conveyed by transients and the temporal envelope, as well as ILDs. Finally, LSO responses adapt to the ITDs and ILDs applied over time, indicating a mechanism that could account for reliable bias in its input [65] (Principle 5).

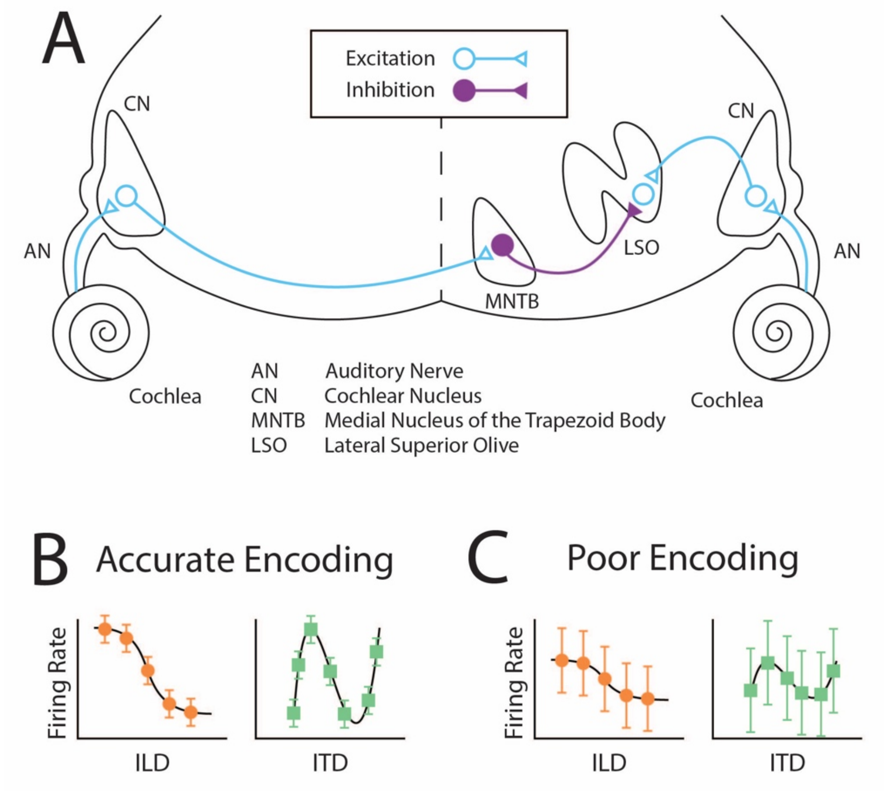


Fig. 2. (A) Model encoding circuit. Stimulation arrives at the auditory nerve via electrical pulses and travels through action potentials to the LSO. (B) LSO output for accurately encoded binaural cues. The x-axis corresponds to ILD in dB or ITD in µs. The y-axis corresponds to the spike rate measured by single unit recording. Spike rates are modulated by binaural cue, showing distinct responses for different ILD or ITD magnitudes. (C) LSO output for poorly encoded binaural cues. The x-axis corresponds to ILD in dB or ITD in µs. The y-axis corresponds to the spike rate measured by single unit recording. Spike rates are weakly modulated by binaural cue, showing indistinct responses for different ILD or ITD magnitudes.

In order to explore the limitations of the LSO with hearing loss and BiCIs, we first need to address the LSO circuit. Figure 2 shows an illustration of the LSO circuit. The cells in the cochlear nucleus that provide excitatory ipsilateral input to the LSO are still debated [66,67].

One candidate is spherical bushy cells. The contralateral input arrives through the medial nucleus of the trapezoid body (MNTB), and is conveyed first by globular bushy cells [68,69]. Both spherical and globular bushy cells show improved phase locking relative to the auditory nerve for the temporal fine-structure [60,70,71] and envelope [72] of stimuli. These cells act as monaural coincidence detectors, firing a spike when inputs are temporally coincident within some window of sensitivity. Monaural and binaural coincidence detectors have characteristically low input resistances that result from low-voltage gated potassium channels, resulting in very short time constants [54]. For spherical bushy cells, it is thought that precisely timed pre-excitatory inhibition facilitates spiking as inputs capitalize on the repolarization channels that are already open [73]. This same sort of early, facilitating inhibition is present in LSO cells [49]. For globular bushy cells, it is thought that temporal coincidence occurs as a result of oversampling the auditory nerve input, i.e., by taking inputs from many auditory nerve fibers [74,75]. The role of inhibition in globular bushy cells is less clear, but blocking inhibition increases overall spike rate in globular bushy cells (e.g., [76]). Excitatory globular bushy cell input to the MNTB is converted to inhibition, which shares extremely similar temporal characteristics to globular bushy cells [77]. Thus, shifts in the number of auditory nerve fibers or balance between excitation and inhibition in the circuitry would likely result in poorer phase locking of bushy cells, and therefore less temporal precision of inputs to the LSO. The balance of excitation and inhibition is also clearly important for the LSO.

Figure 2(B) and 2(C) show illustrations of the types of changes in output that one would expect with deterioration of the LSO circuit, or more generally the encoding system. Of particular concern is the distinctness between varying magnitudes of ITD or ILD and the output of the encoder. In recordings that measure the firing rate of neurons, the distinctness between neural responses associated with different spatial cues is usually measured via neural d’ [78] or mutual information [79,80]. Neural d’ is computed as a ratio of the difference in means between distributions of firing rates relative to their variance. This is based on the ubiquitous statistic d’ from the psychophysics literature [81]. Mutual information is an information theoretic approach that computes the information (in bits) about the binaural cue from the output of the encoder. In other words, mutual information tells the researcher how much is known about the binaural cue being presented based upon the output of the encoder.

No previous reports have directly investigated the effects of different sources of interaural asymmetry in the auditory periphery on LSO responses. This probably has to do with the challenges associated with making recordings in the superior olivary complex. However, much is understood about the circuitry that can be used to make reasonable predictions about how LSO output would change based upon deterioration of earlier structures. The LSO and any neurons sensitive to microsecond ITDs have such specialized mechanisms that any decrement earlier in the auditory pathway is suspected to lead to poorer binaural processing. Various studies have been conducted concerning sources of interaural asymmetry that have implications for LSO processing. The rest of Section II is dedicated to a discussion of these sources.

B. Auditory deprivation

Auditory deprivation for patients with BiCIs has received considerable attention in the literature. We chose to divide this section into two parts because early and late auditory deprivation are suspected to have unique impacts on the auditory periphery and brainstem. A confounding factor with respect to disentangling the effects of earlier and later periods of auditory deprivation is that CI surgery for children was not approved by most insurance companies in the United States until recently, so many listeners with early onset of deafness and

BiCIs are young. In contrast, listeners who receive CIs in adulthood may have experienced relatively short or long periods of auditory deprivation. These likely indicate two clinical sub-populations. Critically, auditory deprivation can occur in one or both ears. Many listeners with BiCIs undergo sequential implantation [3,82], meaning that they have a different period of auditory deprivation for each ear. Distinctions between bilateral and unilateral auditory deprivation will be made throughout the text.

1. Early in development: Central auditory changes

The auditory brainstem is thought to undergo activity-dependent reorganization during a critical window of early development [83,84]. This reorganization coincides with increasing head size of children during development, which changes the relationship between binaural cues and physical locations in space, requiring the auditory system to remain plastic and re-map cues to spatial locations [85]. Accordingly, physical development corresponds with elevated ITD discrimination thresholds in listeners with NH [85], which later decrease once the head stops growing.

Early exposure to coherent ITDs is important for auditory brainstem development [86]. For listeners with BiCIs, an early onset of deafness is associated with greater errors in sound source localization [87–93] and less sensitivity to ITDs [93,94]. However, these listeners demonstrate sensitivity to ILDs, which do not require the same temporal precision as ITDs. One challenge in these types of correlational studies is that many factors predictive of binaural performance in children are interrelated (e.g., experience with BiCIs, age at onset of deafness, duration without CI stimulation).

Alternative approaches include animal models of early onset of deafness. In rabbits stimulated with BiCIs, early bilateral deafness is associated with poorer tuning to ITDs in the inferior colliculus [95,96]. The inferior colliculus receives direct projections from the LSO [97] and shows similar ILD and ITD tuning characteristics [48], suggesting that it is reflective of earlier stages of binaural processing. Providing coherent ITDs via BiCI stimulation early in development restores sensitivity to ITDs in congenitally deaf rabbits and rats [80,98,99].

Similar effects are observed in humans and other animals who experienced intermittent or prolonged periods of asymmetric hearing loss early in life. Human listeners with congenital, unilateral hearing loss who did not receive coherent ITDs during childhood, even after unilateral hearing loss is corrected, show reduced sensitivity to binaural masking level differences up to two years after surgery [100]. Guinea pigs and ferrets with intermittent, asymmetric hearing loss induced by earplugs showed poorer localization after ear plugs were removed [101,102]. Using the same paradigm, chinchillas showed poorer correspondence between ILD and spike rate in the inferior colliculus when the earplug was removed [79]. Results from these studies concerning early deafness on responses in the inferior colliculus suggest a shift in the balance of excitation and inhibition. Inhibition in the inferior colliculus appears to increase (i.e., ratio between excitation and “suppression” [95]). Moreover, the proportion of cells showing sensitivity to binaural stimulation decreases drastically, and is accompanied by a proportional increase in sensitivity to monaural stimulation [79]. Children who grow up with bilateral deafness and receive a unilateral CI show a similar bias in their auditory brainstem responses [103,104]. This will be discussed in more detail in Section 3.4 concerning hemispheric lateralization.

2. Throughout the lifespan: Peripheral deterioration

Recent evidence in adults with BiCIs demonstrated that ITD sensitivity is predicted by the interaction between duration of bilateral hearing impairment (i.e., period of time between hearing loss onset and age at second implant) and period of time with at least one CI [105]. In other words, the decrement in performance associated with prolonged bilateral hearing impairment was counteracted by a longer period of time with at least one CI, and vice versa. This suggests auditory input, and especially bilateral auditory input, is important to preserve binaural circuitry even after development.

In contrast to early onset of deafness, deafness that occurs later in life where central structures are intact followed by auditory deprivation is thought to limit binaural outcomes primarily by deterioration to the existing circuitry. Long periods of auditory deprivation in humans and other animals are associated with deterioration of dendritic processes and cell death of auditory nerve fibers [106–110]. This cell death can occur somewhat uniformly across the auditory nerve or in large sub-populations called “dead regions” [111]. Prolonged auditory deprivation also leads to demyelination of auditory nerve axons [110,112]. The etiology of loss is associated with differing amounts of auditory nerve fiber survival assessed from the temporal bones of cadavers, with listeners who experience early deafness demonstrating the least surviving auditory nerve fibers [108]. However, an important caveat is that this may be confounded with age at onset of hearing loss or prolonged auditory deprivation. Noise-induced hearing loss has been associated with demyelination of auditory nerve fibers [113,114]. Thus, demyelination of auditory nerve fibers, deterioration of dendrites, and death of neurons may represent the pathophysiological changes that occur in response to hearing loss and auditory deprivation.

These changes are associated with temporal response changes in the auditory nerve. Loss of peripheral dendrites is correlated with shifts in the latency of action potential initiation in models of auditory nerve fibers with CI stimulation [115]. Loss and demyelination of peripheral dendrites is also associated with poorer refractory properties of the auditory nerve of mice and rats [116,117]. Since bushy cells refine the temporal precision of auditory nerve fibers, it seems likely that the temporal fidelity of their output would decrease with decreasing temporal fidelity of auditory nerve fibers. In particular, loss of auditory nerve fibers has been related to poorer temporal response properties of models of globular bushy cells [118]. Shifts in excitation and inhibition may be more likely to impact the responses of spherical bushy cells, and loss of auditory nerve fibers may be more likely to affect globular bushy cells. These changes in the temporal response properties before reaching the LSO likely lead to poorer responses of LSO cells with ITDs, which rely on temporal coincidence.

Most of the changes discussed so far have involved the peripheral auditory system, but some central changes are likely occur in listeners who experience a later onset of deafness. The amount of myelin on MNTB axons is regulated by activity during development and into adulthood [119]. Axon diameter, in contrast, increases in an activity-dependent fashion during development [119]. Axon diameter and myelination both affect the conduction velocity of action potentials, where more heavily myelinated and wider axons convey action potentials more quickly, e.g., [120]. Both of these factors have also been associated with decreased maximum firing rate of MNTB cells in a computational model [119].

C. Differences in dynamic range and loudness growth

Dynamic range for BiCI listeners represents the amount of electrical current corresponding to a comfortably loud level minus the amount of current required to detect a sound in quiet. Smaller dynamic range has been associated with poorer sensitivity to ITDs within the same listener [121]. That is, the ITD sensitivity of a particular electrode pair can be predicted by the dynamic range of both electrodes. Similarly, as the AM depth of a stimulus decreases so does ITD sensitivity for listeners with BiCIs [7,122]. Studies in listeners with NH have shown that interaurally symmetric or asymmetric reductions in AM depth are associated with less sensitivity to interaural phase differences and ITDs in the envelope [Chapter 2, Chapter 4]. It is well-documented that the dynamic range of CI electrodes varies across the electrode array (e.g., [121,123]). This can be due to many factors, including distance from the electrode to the modiolus and the health of the auditory nerve [124]. Thresholds measured using "focused stimulation" strategies, which employ negative currents on neighboring electrodes to restrict the neural populations being excited, have been proposed as an index of the quality of the interface between CI electrodes and auditory nerve fibers [125]. This may be related to dynamic range of the electrodes.

Loudness growth differs depending upon the electrode being stimulated within the same ear [126]. Presumably, differences in loudness growth would also be observed between the electrodes in either ear, though this has not been tested systematically. Balancing of bilateral loudness at a comfortable level does not alter lateralization of ITDs, but does improve lateralization of ILDs under conditions of large interaural place-of-stimulation mismatch [127]. To our knowledge, no attempts have ever been made to balance loudness growth within or between ears for listeners with BiCIs.

Differences in loudness growth and dynamic range between ears may be analogous to stimulus-independent, time-varying differences in sound level in each ear for stimuli with complex temporal envelopes like speech. Random, time-varying amplitude fluctuations increase the amount of interaural decorrelation, obscuring ITDs and ILDs in the signal at any moment in time. The effect of interaural decorrelation on binaural processing has been studied in listeners with NH. These studies found that as the amount of interaural decorrelation increases, ITD sensitivity declines [128] and the ability to understand speech in noise decreases [29]. Relatedly, as interaural decorrelation increases, the perceived width of a sound image increases or becomes more diffuse [129], which could contribute to the limited binaural benefits observed under such conditions.

D. Position of CI array in the cochlea

The distance between the CI electrode array and auditory nerve fibers varies depending upon the electrode type, cochlear anatomy, and surgical outcomes [43,130,131]. Greater distance from the auditory nerve fibers results in higher detection thresholds [124] and is also thought to lead to increased spatial spread as the current travels from CI electrodes to auditory nerve fibers. Therefore, multi-electrode stimulation that is used in clinical processing strategies can result in spectro-temporal smearing of the signal, referred to as channel interaction. Reduced spectro-temporal fidelity is associated with poorer speech understanding in listeners with CIs [132,133]. In simulations with listeners who have NH, reduced spectro-temporal fidelity leads to decreased spatial unmasking whether it is interaurally symmetric [20] or asymmetric [19].

Electrode arrays are surgically placed in the scala tympani of the cochlea, where they are in close proximity to the auditory nerve and there are minimal physical obstructions before current reaches the fibers. Particularly troublesome are translocations of CI electrode arrays into other scala of the cochlea. Scalar translocations are one of the strongest predictors of speech intelligibility in that ear [130,131]. Scalar translocations result in substantially further distance between the electrode array and auditory nerve fibers and require current to travel through additional tissue, increasing resistance [134]. One study evaluating the prevalence of interaural scalar mismatches found them for at least one electrode in one ear in 50% of listeners [43].

E. Place-of-stimulation mismatch

The LSO relies on frequency-matched input from the two ears [52]. A recent study investigated the binaural interaction component in BiCI listeners, thought to reflect activity in the LSO [135,136], and found that it was modulated by ITD and interaural place-of-stimulation mismatch [137]. A smaller effect of ITD on the binaural interaction component was found with increasing interaural place-of-stimulation mismatch. This finding corresponds strongly with decreased ITD sensitivity and smaller intracranial lateralization ranges (i.e., shifts in the perceived locations of sounds within the head) of listeners with BiCIs [43,138,139]. Increased interaural place-of-stimulation mismatch is also associated with less spatial fusion of sounds, or more frequent reporting of multiple sound images [138,139], and poorer sensitivity to interaural decorrelation [140]. Place-of-stimulation mismatch also results in reduced spatial release from masking in simulations with listeners who have NH [141].

Unlike other sources of interaural asymmetry, place-of-stimulation mismatch likely reflects the binaural system having fewer "looks" (i.e., neurons processing binaural cues) at the binaural cues conveyed in the stimulus. Thus, it seems highly likely that spectro-temporal degradations would add insult to injury, making it extremely challenging for the LSO and other frequency-matched binaural structures to accurately complete binaural computations. Consistent with this hypothesis, a recent study with simulations in listeners with NH showed that reducing AM depth and increasing interaural place-of-stimulation mismatch have additive effects on ITD lateralization [Chapter 4]. Predicted lateralization of ITDs using a model of the LSO showed reasonable correspondence with trends observed in the behavioral results. That is, as interaural place-of-stimulation mismatch increased and temporal fidelity decreased simultaneously, lateralization of ITDs becomes worse at a faster rate than either decrement in isolation. This is particularly concerning since interaural place-of-stimulation mismatch is a common problem for listeners with BiCIs [43]. Other studies showed reduced interaural decorrelation detection with increasing interaural place-of-stimulation mismatch in listeners with NH and BiCIs [140]. That is, image width, thought to be the primary perceptual feature used in decorrelation detection [129], may be much more difficult to distinguish when the sound image is already quite diffuse. For listeners who receive interaurally decorrelated inputs due to differences in loudness growth between the ears, the sound image with 100% interaural correlation in the signal would have a correlation closer to zero in their brainstem. Listeners with NH show poorer sensitivity in decorrelation detection experiments relative to a reference condition with an interaural correlation of 0% [140,142,143]. Listeners with BiCIs show reduced sensitivity to interaural decorrelation compared to listeners with NH, making it difficult to assess whether there are different patterns between reference conditions in either group [140,144].

F. Summary

Binaural processing in the LSO shows excellent correspondence with trends observed in listeners with BiCIs. This section overviewed sources of interaural asymmetry, their suspected impacts on the inputs to the LSO and its processing of binaural cues, and their effects on perception. Importantly, each of these sources of interaural asymmetry are suspected to interact with one another and affect binaural processing in the LSO or similar binaural structures. It is unlikely that these sources of interaural asymmetry occur in isolation and only recently has evidence begun to accumulate suggesting that they may interact with one another to produce especially deleterious effects on patient outcomes [Chapter 2, Chapter 4].

III. MODEL FOR DECODING: PREDICTIVE CODING

A. Motivation and justification

Several principles were established to disambiguate decoding from encoding in Fig. 1:

1. Decoding must occur after encoding begins.
2. Changes in decoding must be distinct from encoding. In other words, encoding must be demonstrated as intact or stimuli must be presented in such a way that they cannot be attributed purely to changes in encoding.
3. Increasing task difficulty is assumed to only affect decoding (not encoding).
4. Decoding may be demonstrated as different encoding “rules” (i.e., relationships between stimulus parameters and behavioral responses) under different stimulus conditions that should not affect the encoder.

The goal of this section is to use a broader approach to encapsulate changes in decoding associated with interaural asymmetry. Thus, compared to Section II where we proposed one circuit that could be assessed to measure or predict changes in encoding, in this section, we propose a framework called predictive coding consisting of a canonical, cascading series of circuits. Predictive coding supposes that the brain generates predictions about the outside world that are updated in light of new, aberrant sensory input (first described in the visual system by [145]). For our purpose, we assume that information is conveyed in the form of action potentials, which are temporally distinct and carried via frequency-specific monaural or binaural channels. Figure 3(A) is meant to reflect a single unit in a broader network of interconnected units, each responsible for representing different perceptual features. Here, “features” are meant to represent qualities of a sound source that make it distinct from other sources and to which listeners are sensitive. This network of circuits that represent features would consist of cells in at least the auditory cortex, inferior temporal cortex, intraparietal lobule, posterior dorsal field, and hippocampus [146]. An extremely detailed model of this network is not necessary to generate predictions about the relationship between interaural asymmetry and decoding.

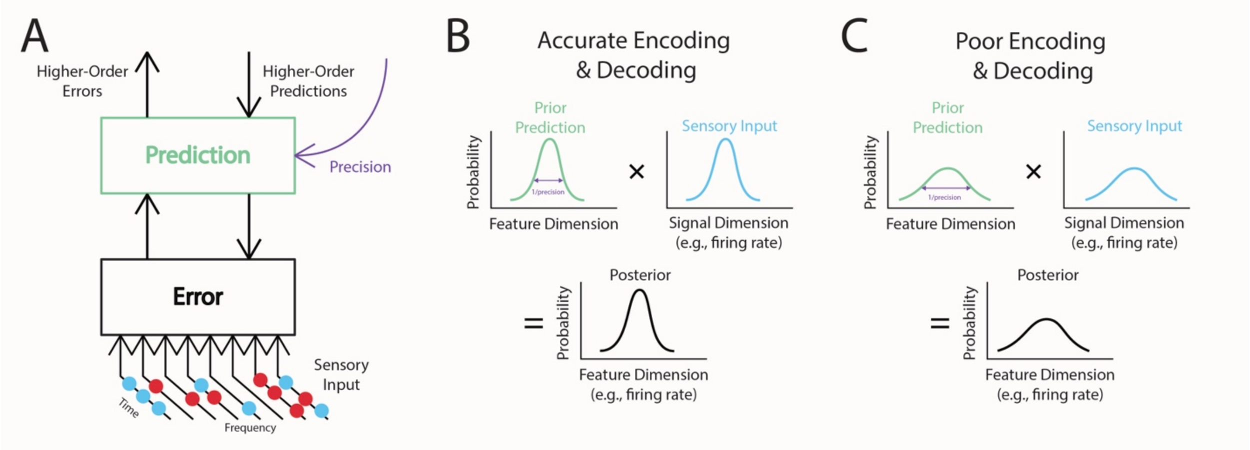


Fig. 3. (A) Predictive coding circuit. Inputs (i.e., circles) represent action potentials corresponding to stimulation of the left (blue) and right (red) ear over time. Input channels are frequency specific and responsive to either monaural or binaural stimulation. Feedbackward predictions, modulated by the degree of certainty or attention (i.e., precision) are compared against feedforward sensory input. If predictions and sensory input conflict, an error signal is sent and predictions are updated. (B) Accurate encoding and decoding. Panels represent prior, sensory input, and posterior decoding distributions, respectively. A less variable distribution corresponds to higher certainty or distinctiveness between sound features and cues. (C) Same as B, except for poor encoding and decoding.

The cortical basis for Fig. 3A comes from the “canonical microcircuit,” which is thought to be a general pattern of connectivity and activation prevalent throughout sensory cortex, and takes the form of a cortical column [147]. Thus, the single unit in Fig. 3(A) is meant to represent a cortical column that encodes one feature of an auditory stimulus, whose features become increasingly complex for later stages of auditory processing. The network represents a series of interconnected cortical columns that convey various hierarchical features used to distinguish one sound source from another. Models of predictive coding have already been used to describe changes associated with hearing loss (e.g., [148]).

Predictive coding lends itself to other conceptual models of auditory scene analysis. For example, Shinn-Cunningham described auditory scene analysis in terms of auditory objects that compete for attention [149]. In that model, a listener can attend to a source of interest using cues encoded during earlier stages of processing (e.g., pitch and location cues). The single unit of this framework is an auditory object composed of different features. In the predictive coding framework, the auditory object is represented by a chain of interconnected cortical columns, whose activity can be modulated with attention according to the source’s features (Fig. 3(A)). However, as the representations of those features become more or less distinct (narrower and wider distributions in Fig. 3(B) vs. Fig. 3(C)), so would the boundaries between different sound sources. Mathematically, this can be represented using Bayesian inference, where a prior prediction (i.e., the internal perceptual model) is updated according to the data (i.e., aberrant sensory input) [145]. Columns further upstream in auditory processing, then, represent hyperparameters of other distributions or probabilistic processes (i.e., distributions of distributions). Predictive coding frameworks have also been applied to speech and language processing (e.g., [150]) and language development (e.g., [151]), suggesting that it may represent a general scheme for information processing in the brain. Studies of spoken language suggest that words with a high number of competitors due to similar features are less accurately identified and take longer to identify [152], similar to our predictions about features used to distinguish sound sources. Thus, one benefit of shifting from a purely conceptual model to predictive coding is that predictive coding can be used to form hypotheses about the neural representations of sounds and sound features, which can be tested experimentally using a wide variety of approaches and stimuli.

Interaural asymmetry represented in this decoding stage will likely be associated with both binaural and monaural encoding issues in listeners with BiCIs. Rather than focusing on specific insults that lead to interaural asymmetry, the rest of this section focuses on systems or processes thought to be involved directly or indirectly in auditory predictive coding. Some specific contributors to interaural asymmetry (e.g., auditory deprivation) will be discussed within each sub-section.

B. Auditory object formation and fusion

Auditory object formation is the process by which various features of a sound are combined into a singular, fundamental unit of auditory perception (for review, see [149,153,154]). Auditory objects carry information that can be derived or parsed. Auditory object formation is thought to occur on two different time scales: short windows of time when spectro-temporal components bind or “fuse” together based on similar properties, and through the formation of auditory “streams” that can be tracked over time. Common onset time has been identified as particularly important for identifying which spectro-temporal segments belong to the same sound source [155]. Auditory streams are updated over time according to the internal predictive model. Both spectro-temporal fusion and auditory stream formation are important for perception, though relatively little attention has been paid to auditory stream formation in the binaural literature. Thus, this section will focus mainly upon different kinds of binaural fusion.

As mentioned previously, the ears represent interdependent channels, meaning that the signals in one are mixed with the other both in physical space and throughout binaural neurons in the central auditory system. Studies using a wide variety of paradigms suggest that information from each ear is mandatorily integrated, even when this is disadvantageous to listening in listeners with NH [17,18,20–22] and hearing loss [13–15,23–27]. In headphone experiments, worsening performance in the presence of contralateral stimulation is most widely referred to as contralateral interference, meaning that information in one ear interferes with the other. In listeners with BiCIs as well as those with hearing loss who use hearing aids, it is difficult to distinguish between failures of attention to one ear and failures of auditory object formation.

Listeners with BiCIs fuse the pitch of bilaterally presented stimuli over very disparate places-of-stimulation, corresponding to frequencies ranges over an octave [156]. This same pattern is observed in listeners who use hearing aids [157,158]. Recent evidence shows a negative linear relationship between the pitch fusion range across interaural frequency differences and the benefit attained from differences in fundamental frequency of target and masker for listeners who use hearing aids [25]. In other words, as the range of frequencies fused between ears increases, the benefit of frequency differences between speakers decreases. Additionally, there is a positive linear relationship between the fusion of pitch and vowels with interaural frequency disparities in listeners who use hearing aids [26], implying that listeners might fuse vowels from different words in complex auditory environments. Fusing stimuli that represent different words provides a perceptual mechanism that may explain poor speech in noise performance for listeners with hearing loss. These experiments could not explicitly account for broader ranges of interaural pitch and vowel fusion based upon unilateral pitch perception, suggesting a central auditory mechanism. Listeners with hearing loss who use hearing aids and/or CIs also show different psychometric functions associated with vowel continua (e.g., nine step continuum of mixtures of the vowels /IH/ and /EH/) between the ears [24]). When presented with stimuli bilaterally, the psychometric function does not always correspond to the psychometric function of either ear. This implies that the spectro-temporal profiles assigned to specific vowels differ between ears and are decoded differently when listening unilaterally compared to listening bilaterally. This is consistent with simulations of interaurally asymmetric temporal fidelity in listeners with NH, which show that listeners are more likely to perceive a single word when stimuli are temporally degraded and more likely to misunderstand the word being presented [Chapter 6].

The literature concerning spatial fusion with listeners who use BiCIs is not consistent. Adults with BiCIs demonstrate slightly greater amounts of fusion over larger amounts of interaural place-of-stimulation mismatch in ITD lateralization experiments compared to NH [138,139,159–161]. They also show fusion for very large ITDs [159,160] compared to listeners with NH. Similarly, for listeners with BiCIs, very large ITDs (≥2000 μs) can be used to lateralize stimuli further to the left and right than smaller, more physiologically plausible ITDs (≤800 μs) [46,93,162]. These results suggest that over-fusing sounds according to interaural place-of-stimulation mismatch might be associated with poorer lateralization of ITDs [138,139,163], sound source localization [164], and spatial release from masking [141]. This is also supported by wider central masking functions, where electrodes from disparate places-of-stimulation in the ear opposite the target result in higher detection thresholds [27,159]. On the other hand, adults with BiCIs show less fusion compared to listeners with NH for time-delayed pairs of stimuli with opposite ITDs (i.e., echo thresholds of lead-lag pairs) in precedence effect experiments [165]. Moreover, children with BiCIs, particularly those with the earliest onset of deafness, tend not to fuse interaurally place-matched stimuli even with 0-μs ITDs [166,167].

Assuming that place-of-stimulation is not a very salient cue to distinguish sounds, and that ITDs are not very useful as they are conveyed via clinical processors, both sets of findings could be indicative of a similar underlying problem. That is, because the cues conveyed cannot be used effectively to segregate between sound sources, the auditory system relies more heavily on other cues (e.g., ILDs) to segregate between sounds and treats inputs from each ear as always of the same or different source origin. In other words, the distinction between "one" or "two" sounds when the cues being presented are not salient may be somewhat arbitrary.

C. Attention

Attentional shifting and modulation begins after auditory object formation, and eventually works in parallel with auditory object formation for ongoing stimuli. Auditory streams compete for attention, which is allocated using features of auditory objects. Thorough reviews of object-based auditory attention were published previously [154,168], and some of their findings are summarized here. Allocation of attention depends upon accurate representation of cues in the encoding stage (Fig. 1). Thus, when features of different sounds are less distinct or attention is drawn elsewhere, performance in difficult listening conditions becomes poorer. Auditory attention can make use of auditory objects' features to sort them into the foreground or background. This attention may manifest as suppression of neural responses related to objects in the background, as well as amplifying and refining tuning of neural responses related to objects in the foreground.

Interaural asymmetry of attention in listeners with BiCIs has only recently begun to be explored. Results suggest attentional differences between the ears when there are sufficiently challenging task conditions. Many of these studies have used binaural unmasking tasks, where a mixture of target and masking speech is presented to one ear and compared against conditions where the other ear is also presented with a copy of the masker. Listeners with NH usually show an improvement in the latter condition, presumably because the masking speech is fused across ears (e.g., [19,20]) resulting in a centered image (ITD = 0 μs, ILD = 0 dB) and the target speech is perceived toward the ear in which it is presented (ITD = ±∞ μs, ILD = ±∞ dB). Listeners with BiCIs who experienced a prolonged period of deafness, on the other hand, frequently show a *decrement* in performance when a copy of the masker is added to the other ear (i.e., contralateral interference [13]). In this and related studies, the result of adding a copy of the masker to the opposite ear is dependent upon the ear to which the target is presented. When target speech is presented to the better ear (defined by speech recognition scores), some listeners show improvements in speech perception, although these are typically smaller than those observed in listeners with NH. When target speech is presented to the poorer ear and there is significant interaural asymmetry, contralateral interference occurs for listeners with BiCIs [13–15], simulations in NH [19], and listeners with a CI in one ear and NH in the other ear [23]. Simulations of BiCI processing in NH also show evidence of increased listening effort via pupillometry [169]. This is consistent with the idea of a mandatory shift in attention toward the more salient source.

Interestingly, patients who experience strokes demonstrate a similar difficulty allocating attention to the ear contralateral to the lesion when there is stimulation in the ear ipsilateral to the lesion. This is referred to as auditory extinction [170,171]. Auditory extinction is modulated by the relative onset time of the stimulation of either ear [172], where the ear presented first was more likely to be correct, especially if it was ipsilateral to the lesion. Contralateral interference in listeners with BiCIs may therefore be indicative of a major failure of the attentional network that presents as auditory extinction. That is, in order to detect the failure of attention for these listeners, both ears must be stimulated and the worse ear must contain the target. It may also be that certain information in the signal(s) is prioritized. Unconscious shifting of attention to particular information in sounds is supported by a study showing that identification of whether or not a sound is moving is poorer with stimuli that have speech envelopes for CI simulations in listeners with NH [173]. This suggests that the speech processing and localization pathways may compete for resources during behavioral tasks.

An alternative explanation of contralateral interference presented in the previous section is that when stimuli are sufficiently spectro-temporally degraded, changes in auditory object formation occur. This could be due to an over-fusion of unrelated speech between the ears, resulting in an obscured sound image making it more difficult to parse for speech information. This could also occur due to an under-fusion of the masking speech in each ear, effectively introducing an additional masker. One important middle step has been to explore the fusion of speech stimuli for monosyllabic words [Chapter 6] and vowels [24,26]. Results from the former imply that over-fusion occurs when stimuli are spectro-temporally degraded, that the worse ear is ignored, and that the worse ear interferes with access to speech information in the better ear.

In conclusion, it seems likely that both auditory object formation and attention are affected by interaural asymmetry. Particularly indicative of an attentional problem are studies with listeners who have a CI in one ear and NH in the other [23] and simulations in NH [19] showing contralateral interference. In this case, the signals in each ear are so different that over-fusion seems very unlikely. Together these findings suggest that changes in object formation and problems allocating attention likely interact with one another to produce poorer bilateral speech outcomes when listeners demonstrate interaural asymmetry. They also suggest that it may be necessary to tax the attentional system by introducing masking stimuli in spatially distinct channels (e.g., the opposite ear) before interaural asymmetry is apparent [13,15].

D. Cortical lateralization and specialization

In listeners with NH, some specialization of speech and language function is present. A review on this topic has been published previously [174] and results are summarized here. Auditory hemispheric lateralization in listeners with NH is typically associated with a right ear advantage in listening experiments [175,176] and greater activation identified in cortical imaging [177]. Classical studies have also shown right ear advantage associated with simulated ablation via anesthesia [32]. Right ear advantage is slightly less prevalent in individuals who are left-handed, but both groups tend to show right ear advantage over left ear advantage [178]. Interestingly, this advantage can still be seen via ILDs favoring the left ear [176] and when stimuli have ITDs up to ~60-90 ms [175]. The amount of advantage depends highly upon the task used and is sometimes difficult to reproduce in the same groups of listeners [178–180].

Many listeners with BiCIs have a perceptually indicated “better ear” (e.g., [6,9,11,13]). This is one type of interaural asymmetry. Most often, the better ear corresponds to the first-implanted ear. Different factors predict speech understanding in either ear, largely stemming from differences associated with the auditory periphery. Thus, while we discussed asymmetries in attention in Section III.C, this section will be dedicated to the asymmetric representation of sounds in either side of the cortical and subcortical auditory structures.

Two excellent reviews have addressed the effects of unilateral auditory deprivation during development on binaural stimulation [181,182] and are partially summarized here. Children with BiCIs who have an extensive history of unilateral stimulation show an extreme shift toward bilateral cortical representations of stimuli from the first-implanted ear [183–185]. The same finding has been replicated in congenitally deaf cats [182,186]. These studies show auditory deprivation during early stages of development is associated with a repurposing of cells in the contralateral cortical hemisphere to represent ipsilateral input, and can only be mediated by early CI implantation [186,187]. At lower levels of processing, single units of the inferior colliculus in chinchillas who experienced mild, intermittent hearing loss via earplugging [79] and asymmetries in brainstem function observed in children who are congenitally deaf and receive a unilateral CI [104] show the same types of reorganization. Reorganization favoring the non-deprived side of the body is demonstrated in classical studies of visual deprivation [188]. Thus, hemispheric reorganization during sensitive periods of development may represent one pathophysiological mechanism that underlies sensory asymmetry. Accordingly, research in children indicates that the best outcomes are observed in listeners with the least history of asymmetric auditory input [189].

An important finding demonstrates that no consistent ear advantage is found in children [190] or adults [6,11] who are simultaneously, bilaterally implanted. That is, ear advantage in listeners with BiCIs may have more to do with factors related to neural health as well as the interface between their CI electrode array and auditory nerve fibers. Some significant differences in speech understanding take time to emerge as listeners learn to use their CIs [11], suggesting that listeners may adapt to one ear over time. Thus, rather than attention as in listeners with NH, other factors associated with the auditory periphery likely contribute to differences in hemispheric representations of sound. One example that is likely to have a very strong effect on interaural asymmetries in speech outcomes is scalar translocation [130,131]. It should be noted that differences in psychophysical sensitivity and predictors of auditory nerve health vary depending upon the place-of-stimulation even within the same ear as well as across the ears (e.g., [9,62,123,124,191–194], Chapter 3, Chapter 5). Thus, interaural asymmetry and unilateral deprivation are not synonymous. Unilateral auditory deprivation, especially during development, is a condition that has an extraordinarily high chance of resulting in interaural asymmetry and whose the pathophysiology has been well-documented. Since it is likely that many of the same peripheral issues may occur for patients who are sequentially implanted and patients who are simultaneously implanted, it is important not to overlook how peripheral factors contribute to interaural asymmetry for listeners with prolonged unilateral deprivation. One particularly important question is whether psychophysically based predictors of neural health are sensitive to the pathophysiological changes associated with auditory deprivation. It may be possible to save time and resources if these predictors provide similar information about spectro-temporal processing as more complicated and intensive assessments of cortical lateralization. It remains unclear whether differences in sensitivity to spectro-temporal cues are reflected in cortical activity. In other words, interaural asymmetry in the auditory periphery leading to poorer binaural outcomes may not necessarily be reflected in measures of cortical activity.

E. Auditory experience and training

Listeners with BiCIs who are implanted in adulthood show a period of plasticity during which they adapt to the new mode of stimulation. With respect to pitch, this is characterized by adaptation of the place-of-stimulation to pitch map, which is different for patients with CI than individuals with acoustic hearing [195,196]. With respect to speech understanding, accuracy increases and seems to plateau around three to five years of CI use [197]. This plasticity does not necessarily imply that both ears will improve at a similar rate. For example, adult listeners who were simultaneously implanted demonstrated significant differences in performance between ears that emerged only after one year of experience [11]. Children with BiCIs tend to have the best outcomes over time when they are implanted shortly after onset of deafness, improving over time for sound source localization or lateralization [88,89,91,198] and speech understanding [199], but not spatial release from masking [200]. As mentioned earlier, these factors are often interrelated in children with BiCIs, so it is difficult to discern exactly how deafness and cochlear implantation interact with the normal developmental trajectory. Together with Section II.C and III.D, these studies imply that auditory deprivation, especially during early development, may interact with experience to produce interaural asymmetries.

Recent evidence suggests that providing coherent ITDs to the developing auditory system might restore access to ITD sensitivity later in life [80,98,99]. In these experiments, animals who retained ITD sensitivity despite early deafness regularly used ITDs to localize sounds with operant conditioning throughout their development. Thus, it may be that the utility of ITDs and sensitivity of the midbrain was based as much on training as exposure to coherent cues. One important future direction, especially when making attempts to restore access to spatial cues for listeners with BiCIs, may be to involve training. A lack of training or experience may help explain why benefits are not observed in listeners with BiCIs when ITDs in the temporal fine-structure are preserved [37–39], particularly because they may have learned to rely on other cues for sound source localization. When considered in the context of a patient with significant interaural asymmetry, it is not surprising that many listeners report removing one or both hearing devices during the day [201–204] because they may feel that their worse ear interferes with or frustrates listening. Thus, rather than asking patients to passively listen in these difficult configurations with the hope that device acceptance and performance may improve, auditory training may provide a promising and motivating alternative. Results from listeners with unilateral CIs and simulated asymmetric loss via earplugging suggest listeners are able to re-weight cues based on auditory training [205,206]. Pilot testing of another training procedure suggests that it can also lead to improvements in spatial hearing outcomes for listeners with BiCIs [207].

F. Summary

Decoding was proposed as a process whereby ongoing perceptual predictions are refined by new, unexpected sensory input. This process is thought to occur by using features associated with different auditory objects to modulate attention to a desired source of interest. In listeners with interaural asymmetry, auditory object formation may be compromised and attention may be drawn toward the clearer signal. Both of these processes may be facilitated by an overrepresentation of the better ear throughout the ipsilateral hemisphere, and a lack of training or relevant experience may compound or facilitate continued interaural asymmetry. While unilateral deprivation has received considerable attention, it is unlikely to be the only factor contributing to overrepresentation of the better ear and is not the sole predictor of ear advantage.

Chapter 2: Asymmetric Temporal Envelope Encoding: Within- and Across-Ear Envelope Comparisons

Reproduced from Anderson, S. R., Kan, A., & Litovsky, R. Y. (2019). Asymmetric Temporal Envelope Encoding: Within- and Across-Ear Envelope Comparisons, *J Acoust Soc Am, 146(2),* 1189-1206, with the permission of AIP Publishing.

I. INTRODUCTION

Individuals with hearing impairment who are fitted with hearing aids or cochlear implants (CIs) generally perform more poorly than normal-hearing (NH) listeners on speech in noise tasks [8,208]. However, performance varies widely across individuals with hearing impairment, where some listeners demonstrate performance in the NH range and others perform much more poorly. This variability may be related to how well listeners can segregate target speech from background noise, which involves the use of source segregation cues [149,209]. It is likely that sensitivity to and efficacy of source segregation cues is intimately related to how well these cues are encoded at the auditory periphery. That is, if the signal is poorly encoded in early stages of the auditory system, then access to source segregation cues will be limited through the rest of the central auditory system.

To date, the ability to compare simultaneously presented temporal information across cochlear places-of-stimulation with asymmetries in the state of temporal encoding across frequency and/or ears has not been directly addressed. Each stage of the auditory system has developed specialized mechanisms to extract information necessary for the spectro-temporal comparisons involved with source segregation. Thus, it may be that the ability to process auditory stimuli and extract information necessary for source segregation is limited by the stages of auditory processing with greatest decrement. We therefore hypothesized that the ability to compare temporal information across frequency and/or ears may be limited by the place-of-stimulation with the poorest state of encoding. In this experiment, we attempted to quantify the effect of a reduction of temporal information by reducing the amplitude modulation depth, for amplitude-modulated (AM) tones. Reducing AM depth can be viewed as a reduction in the stimulus dynamic range, which is suspected to lead to less phase locking of the auditory nerve [210,211]. We predicted that reducing AM depth at one cochlear place will reduce the ability of the auditory system to compare temporal envelope fluctuations.

A. Hearing impairment leads to peripheral degradations

The fidelity of temporal input to the auditory system of individuals with hearing impairment can vary widely across the frequency spectrum and between right and left ears according to etiology of disease, duration of deafness, and the type of hearing device being used. Thus, access to temporal information used to segregate sound sources may be limited for some or many frequencies in one or both ears within an individual.

Long durations of deafness and etiology of disease have been linked to deterioration of the auditory nerve (for review, see [112]). One consequence of long durations of deafness is demyelination of axons, reduced arborization of dendrites, and shrinkage of cell bodies (e.g., [106]). Several studies suggest that long durations of deafness, which result in deterioration of dendrites and demyelination of axons, lead to poorer phase locking of the auditory nerve (e.g., [116,117]).

Additional studies in animals indicate that long durations of deafness (> 6 months) resulted in a reduction in the number of spiral ganglion cells in the auditory nerve [116]. Human temporal bone studies confirmed that spiral ganglion cells tended to decrease as durations of deafness increased [108] and depended on etiology [110]. Spectral holes, or non-uniform losses in spiral ganglion cells in the auditory nerve, have also been demonstrated in populations with sensorineural hearing impairment [109]. Together, results from studies in animals and humans suggest that the ability to convey temporal information can vary highly across or within individuals due to etiology of disease or long durations of deafness.

Losses of auditory nerve fibers lead to an increase in threshold for the compound action potential and deterioration of peripheral processes that is associated with increased threshold for individual auditory nerve fibers [115–117]. Sensorineural hearing loss due to damage of outer hair cells results in a change in loudness growth, increases in threshold, and reduced dynamic range (for review, see [212]). Similarly, poor interface between CI electrodes and auditory nerve fibers is associated with reduced dynamic range in CI users [125,126]. Thus, dynamic range of auditory nerve fibers is reduced in listeners with hearing impairment and may also vary with place-of-stimulation, depending upon the status of the auditory nerve. This suggests that the same temporal information (rate and depth of temporal fluctuations) must be represented in a smaller stimulus space. Accordingly, the central auditory system may lose access to information in the stimulus (i.e., via changes in spiking with the rate and depth of envelope fluctuations) because of the limitations in the peripheral auditory system.

Compared with acoustical stimulation, electrical stimulation leads to different neural representations for the same temporal information in stimuli. Modeling studies and extracellular recordings from the auditory nerve show that electrical stimulation of auditory nerve fibers gives rise to very little temporal jitter compared to normal, acoustic stimulation [213,214]. This difference in temporal envelope encoding due to acoustical vs. electrical stimulation suggests that listeners that receive both methods of stimulation must reconcile differences in the representation of temporal information during sound source segregation. For example, the introduction of hybrid CIs has led to electric and acoustic hearing within the same ear. Thus, for patients with a hybrid CI, the same temporal envelope in the acoustic stimulus may have very different representations in the auditory nerve within the same ear. Similarly, for listeners with bimodal hearing or single-sided deafness that are fitted with a CI in the deaf ear (SSD-CI), acoustic and electric inputs must be integrated across the ears. Differences in temporal envelope encoding due to hearing device may be compounded by interactions with the previous factors, like duration of deafness and etiology of disease. It has even been suggested that introducing stochasticity to electric hearing via high rates of stimulation may improve hearing outcomes in listeners with CIs (e.g., [214]), thereby making electrical stimulation more similar to acoustical stimulus representations. While it is well known that listeners with bimodal hearing, hybrid CIs (e.g., [215]), SSD-CI, or bilateral CIs (e.g., [14]) perform worse than NH listeners in the same speech understanding tasks, the functional implications of differences in temporal encoding have never been studied systematically.

While hearing impairment can lead to differences in the physiology in each ear, and by extension the representation of incoming signals, how these differences impact real world listening where multiple sound sources are often present is unclear. However, it is known that temporal information in the envelope contributes to sound source segregation in several different ways and differences in fidelity of temporal encoding within and/or across ears may have a substantial impact on one’s ability to segregate sound sources. In the next section, the role of temporal envelope cues in sound source grouping and segregation will be discussed briefly.

B. Temporal envelope information facilitates grouping and segregation of sounds

Temporal information in a signal can be decomposed into two categories: Envelope and fine structure. Broadly, the envelope refers to slow, gradual changes in the stimulus over time, whereas the fine structure refers to faster changes that fall within the envelope. From a signal processing perspective, signals can be decomposed as the product of the temporal envelope (akin to amplitude modulation) and fine structure (akin to carrier). The temporal envelope is preserved in CI processing [35] and the role of the temporal envelope in speech perception has been studied in depth (e.g., [216–218]). Temporal envelope fluctuations make up two aspects of information in speech, corresponding to slower and faster fluctuations (for review, see [219]). For low AM rates, particularly those below 30-50 Hz, listeners perceive changes in loudness over time during a single stimulus presentation. However, at higher rates (above 50-90 Hz), listeners perceive a change in pitch.

Temporal envelope fluctuations are suspected to contribute to sound grouping and segregation in several ways (for review, see [220]). Repetitive fluctuations in the envelope over time result in the formation of an auditory stream for listeners with NH [221–223] and CIs [224,225]. Also, co-modulation of temporal envelopes across frequency leads to better speech reception of sine wave speech [226] and could be used to group sounds [227]. Conversely, signal disparities can be a cue for segregation. For example, listeners with NH are sensitive to phase disparities of sinusoidally amplitude-modulated tones [228]. Further, binaural envelope decorrelation, which is perceived as a diffuseness in the head or “roughness” [129,229] may be related to detection of sounds with differing phase in binaural masking level difference experiments. It has been demonstrated that NH listeners are sensitive to both monaural [230] and binaural [231] envelope decorrelation. Listeners with CIs are also sensitive to binaural envelope decorrelation [140,229], but their sensitivity to monaural envelope decorrelation has not been studied.

C. Purpose of current study

Representation of temporal information important for sound grouping and segregation likely varies depending upon place-of-stimulation for listeners with hearing impairment. Hence, it is likely that these spectro-temporal comparisons in listeners with hearing-impairment must be completed with asymmetric representations of temporal information. Most research concerning sound source segregation in NH listeners has been conducted with individuals that have relatively symmetric representations of temporal information within and across ears and has focused on primarily symmetric representations of temporal information in the stimulus. Sound source segregation research in individuals with hearing impairment tends to focus on the type of prosthesis being used in either ear: bilateral hearing aids (e.g., [158,208]), bimodal hearing (e.g., [157]), single-sided deafness and electrical-acoustic integration [215,232], or bilateral CIs [8,14]. Beyond this, most studies do not address asymmetries in temporal encoding for all places-of-stimulation despite evidence that asymmetries exist. Thus, the contributions of asymmetric temporal encoding on the ability to extract and compare information from co-occurring envelope fluctuations have not been explored systematically. It is possible that, given the heterogeneity in fidelity of temporal encoding across places-of-stimulation in the hearing-impaired auditory system, the limit to the central auditory system is imposed by places-of-stimulation providing poorest temporal encoding.

One way to understand the impact of these differences is to provide poorer temporal encoding of one sound when compared against a second sound in NH listeners. We attempt to model varying states of temporal envelope encoding (better and poorer) by reducing the depth of

AM when presenting the stimulus to NH listeners. Data from the auditory nerve of cats indicated that as modulation depth decreased, the synchronization of firing decreased for all AM rates between 50 – 250 Hz [210]. Further, smaller AM fluctuations are homologous to the reduced dynamic range observed with poor status of the interface between CI electrodes and the auditory nerve [125,126].

In this work, we use a new experimental paradigm similar to decorrelation detection to probe the auditory system. The goal of this paradigm was to explore listeners' ability to compare the rate of AM across ears and place-of-stimulation when temporal encoding at one cochlear place-of-stimulation was degraded. We hypothesized that the auditory system is limited in its ability to compare temporal envelopes by the fidelity of temporal encoding in the worse place-of-stimulation. Therefore, if amplitude modulation depth is lower in one cochlear place-of-stimulation in one ear, thereby reducing the fidelity of temporal encoding, then the sensitivity to differences in AM rate across ears and/or cochlear places-of-stimulation should decrease.

II. EXPERIMENT 1: A NORMAL-HEARING MODEL OF POOR TEMPORAL ENCODING

A. Motivation

The purpose of this experiment was to determine if reduced AM depth resulted in poorer ability to discriminate between AM rates presented to a single place-of-stimulation in NH listeners. Reducing AM depth increased rate discrimination thresholds for square wave amplitude-modulated noise in NH listeners and listeners with hearing impairment [233]. Additionally, using noise carriers, it is possible that high frequency hearing loss could account for better performance with some listeners compared to others (for review, see [234]). The only experiment that has tested the effects of AM depth on rate discrimination using tone carriers varied AM depth relative to detection threshold [235]. Thus, it was necessary to determine whether the same effects hold for the stimuli employed in Experiment 2 of this paper.

B. Methods

1. Listeners

Twelve NH listeners (age 19-25 years; mean 22.4 years) participated in this experiment. All listeners had absolute detection thresholds less than or equal to 20 decibels hearing level (dB HL) for octave-spaced frequencies between 250 and 8000 Hz in the ear being tested. Before completing the experiment, all listeners provided informed consent. All procedures were approved by the Human Listeners Institutional Review Board of the University of Wisconsin-Madison.

2. Stimuli & procedures

Stimuli were sinusoidally amplitude-modulated (SAM) pure tones with carrier frequencies of 4000 or 7260 Hz. These carrier frequencies were chosen because they are disparate, with carrier frequency and sidebands falling outside equivalent rectangular bandwidths for each center frequency based on the equations provided by Moore and Glasberg [236]. Therefore, they were expected to result in a very small degree of overlap in stimulation on the basilar membrane when the two SAM tones were played simultaneously in the same ear for Experiment 2. Additionally, using carrier frequencies above 1500-3000 Hz should result in phase locking only to the envelope and not carrier tone. For each listener, one cochlear place-of-stimulation was chosen by using a carrier frequency of 4000 or 7260 Hz in the left or right ear. Six listeners were tested with carrier frequencies of 4000 Hz and the other six were tested with carrier frequencies of 7260 Hz. Seven listeners were tested in the left and five in the right ear. Each SAM tone was presented at 65 dB sound pressure level A-weighted (dB(A)) for 600 ms. Stimuli were generated in MATLAB and presented using a Tucker-Davis Technologies System3 with RP2.1, HB7, and PA5 units (digital processor, amplifier, and attenuator, respectively) through Sennheiser HD600 open-back headphones. All testing took place in a double-walled, sound-attenuating booth (Industrial Acoustics Company, Inc.).

Two AM depths were tested: 50% and 20%, corresponding to -6 and -14 dB with respect to 20 $\log_{10}(m)$, where m is the modulation depth in proportion. The AM depth of 20% was chosen because pilot experiments (data not shown) suggested this was the AM depth for which listeners began to show difficulty with discriminating AM rate but remained able to detect AM. The AM depth of 50% was chosen as a comparison to provide a control where listeners easily discriminated AM rates without making AM depths extremely different from one another (e.g., 20% vs. 100% AM depth). Trials were blocked by AM depth condition and the order of AM depths was counterbalanced across listeners. One practice block was given with 50% AM depth prior to testing. Minimal training was desired so that results could be more easily generalized to listeners without extensive training in other experiments. Testing took approximately 0.5 to 1 hours for each listener.

The experiment employed an “oddball” three-interval, two-alternative forced-choice task where listeners were instructed to choose the highest AM rate from either the second or third interval. The first interval always consisted of a standard AM rate of 10, 30, or 90 Hz. The second and third intervals presented the same standard rate as the first interval and a higher (oddball) AM rate, respectively, or vice versa with equal probability. The inter-stimulus interval was 300 ms within each trial. Visual feedback of the correct answer was given after each trial.

Three interleaved two-down, one-up adaptive tracking staircases corresponding to standard AM rates of 10, 30, and 90 Hz were used. Three blocks of three staircases were collected per listener: one practice block with 50% AM depth, and two other blocks with 50% and 20% AM depth. Other than the first (practice) block, no data were excluded from analysis for Experiment 1. On each trial, one of the staircases that was not previously completed was selected to be presented to the subject with equal probability. For each staircase, 12 turnarounds were collected to estimate AM rate discrimination threshold. To ease the task of listeners and prevent confusion about the task throughout testing, step sizes were determined following the parameter estimation by sequential testing (PEST) rules outlined in Litovsky [237]. The maximum step size for each standard was 16/3, 16, and 48 Hz, and the minimum was 1/3, 1, and 3 Hz (i.e., the same relative change with the standard AM rates of 10, 30, and 90). The maximum oddball AM rate possible was 5 times the standard and the minimum was 31/30 of the standard. Discrimination thresholds for AM rate were computed using a weighted maximum likelihood procedure by fitting a two-parameter logistic function to all data within each staircase and weighting each point by the number of observations. Curve fitting was completed using version 2.5.6 of the psignifit toolbox in MATLAB via the method described by Wichmann and Hill [238]. AM thresholds are presented as the base 10 logarithm of change in AM rate divided by standard AM rate ($\Delta f/f$) at threshold. This was done to normalize thresholds, allowing for fairer comparisons across standard AM rates.

C. Results

The goal of Experiment 1 was to determine the effects of depth of AM on AM rate discrimination thresholds. For reference, thresholds for each subject in each condition are shown in Supplementary Fig. 1 and demonstrate the considerable variability across listeners.

Summarized results are plotted in Fig. 1. The x-axis of Fig. 1 corresponds to the standard AM rate tested. The y-axis corresponds to the difference in threshold between the 50% and 20% AM depth conditions; positive values indicate poorer performance with 20% AM depth. Error bars shown in the plot correspond to 95% confidence intervals for the difference in threshold between the 50% and 20% AM depth conditions. Change in threshold was expressed in units of log(Δf/f) as in Eq. 1.

$$x = log_{10}(\frac{\Delta f_m}{f_m}) \qquad (1)$$

to normalize between the different standard AM rates. A value of 0 would indicate that threshold (Δf) was equal to the standard AM rate, and a value of 1 would indicate that threshold was 10 times the standard AM rate. A mixed-effects analysis of variance (ANOVA) was completed with fixed effects of AM depth and standard AM rate within-listeners, as well as presentation ear and carrier frequency across-listeners, and log(Δf/f) threshold as the dependent variable. A random effect of listener was included to account for variability across listeners. Thresholds for AM rate increased as AM depth decreased from 50% to 20% [F(1,55) = 4.912, $p < .05$]. Anecdotal reports from most listeners suggested that the 20% AM depth condition was noticeably more difficult and required closer attention to complete the task.

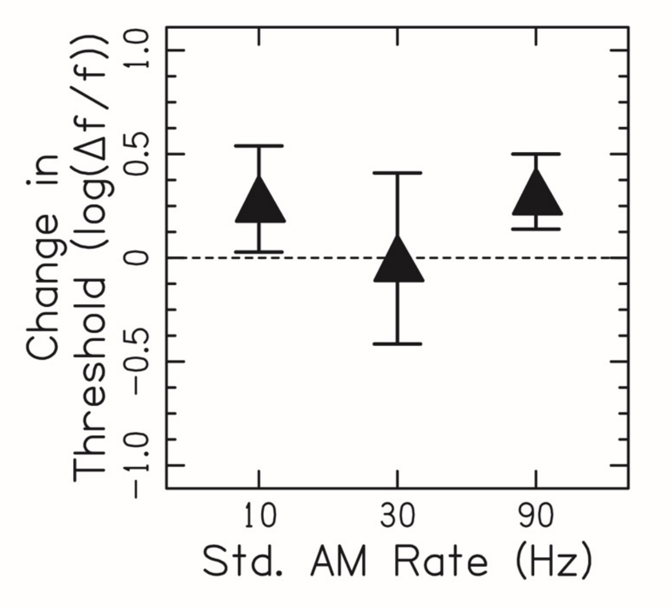


FIG. 1. 95% confidence interval for change in threshold between the 20% and 50% AM depth conditions. The x-axis corresponds to the three possible standard AM rates (i.e., the AM rate repeated twice in the three-interval, two-alternative forced-choice task). The y-axis corresponds to the difference in threshold between the 20% and 50% AM depth conditions. Confidence intervals were computed assuming a t distribution. Confidence intervals for standard AM rates of 10 and 90 Hz fall outside of zero, suggesting that there was a significant increase in threshold as AM depth was reduced from 50% to 20%.

Standard AM rate also significantly affected threshold [F(2,55) = 4.908, $p < .05$], with the 10 Hz standard AM rate having a significantly higher threshold than 90 Hz [t(55) = 3.123, $p < .01$], but no significant differences between 10 and 30 Hz [t(55) = 1.776, p = .188] or 30 and 90 Hz [t(55) = 1.348, p = .375]. All post-hoc tests were completed by comparing estimated marginal means with Tukey adjustment for multiple comparisons. There was no statistically significant interaction between the effects of standard AM rate and AM depth [F(2,55) = 1.278, p = .287].

The large decrease in change in threshold plotted in Fig. 1, suggestive of an interaction between AM rate and AM depth, is due primarily to subject TQB. Subject TQB had a substantially lower threshold for 20% AM depth compared to 50% (see Supplementary Fig. 1), and may have been due to an ill fit of the maximum likelihood function to listener TQB's adaptive tracking data.

To rule out effects of differences between carrier frequencies, the choice of carrier frequency was counterbalanced across listeners. The effects of ear [F(1,9) = 0.555, p = .478] and carrier frequency [F(1,9) = 0.090, p = .772] were not statistically significant across individuals.

D. Discussion

Experiment 1 investigated whether the ability to discriminate changes in AM rate presented over sequential intervals becomes worse as the depth of AM decreases. Thresholds for AM rate discrimination relative to three standard rates (10, 30, and 90 Hz) were estimated for two different AM depths: 20% and 50%. Results indicated that when AM depth decreased, there was an increase in threshold for discriminating changes in AM rate, consistent with the prediction that poorer encoding of the temporal envelope results in poorer ability to discriminate changes in rate of fluctuations. This result is consistent with previous reports in NH and listeners with hearing impairment using square wave amplitude-modulated noise [233].

There are several reasons why a reduction in AM depth might result in more difficulty in discriminating AM rate for SAM tones (for review, see [239]). As mentioned previously, smaller AM depth resulted in poorer phase locking to AM in extracellular recordings of the auditory nerve from the cat [210] and computational models [240]. Thus, one interpretation of our result suggests that poorer phase locking of the auditory nerve would worsen psychophysical performance. This conclusion is supported by studies with listeners that have CIs, showing a significantly higher percentage of correct discriminations for square wave and sawtooth waveforms, which have sharp onsets [241]. Similar improvements in performance due to sharper envelopes have been shown in tasks involving interaural timing difference discrimination in listeners with NH and CIs [242–245]. These improvements in performance were presumably driven by an improvement in phase locking by the auditory nerve.

Another reason reduced AM depth might worsen the ability to discriminate AM rates presented sequentially relates to sidebands in the magnitude spectrum. Lower AM depths result in lower magnitude of sidebands in the magnitude spectrum, and higher AM rates increase the distance between the center frequency band and sidebands. However, data from [234] suggest that when modulation depth is presented near detection threshold, sidebands would have been useful in discrimination only for the highest AM rates presented in this study (roughly >300 Hz for a carrier of 4000 Hz). In their experiments, sidebands were audible at substantially lower AM depths than those used here. From Supplementary Fig. 1 it appears that only participants TJI at 20% and 50% AM depth and TSF at 20% AM depth had thresholds with the standard AM rate of 90 Hz in the frequency region where sidebands would be a useful cue.

Additionally, it is thought that non-linear frequency processing in the cochlea introduces combination tones that vary depending upon the relationship between carrier frequency and higher AM rates [246–249]. Lee [246] attempted to mask combination tones by introducing low-pass noise in one of the conditions in her experiment. Adding masking noise increased threshold for AM rate discrimination at higher AM rates (160 Hz and 320 Hz when the carrier frequency was chosen randomly, 160 Hz at 500 Hz carrier frequency). However, the magnitude of combination tones is not computable, so the addition of masking noise may have masked difference tones at lower frequencies or interfered with the listener’s ability to assess pitch. Low-frequency noise was also added in experiments concerning AM detection with the goal of masking combination tones [248,249]. Results from the Appendix of Eddins [250] suggest that low-frequency distortion products would not contribute to AM detection for the rates used in our study (though the stimuli in the former study were noise and quasifrequency-modulated).

Results from Experiment 1 also indicate that thresholds were significantly higher for the 10 Hz compared to the 90 Hz standard AM rate. Increasing thresholds as the AM decreases below 100 Hz are relatively consistent with previous literature on rate discrimination in general (see data reviewed by [239,251]). Lee [246] reported that thresholds were relatively similar across AM rate when expressed in terms of Δf/f using SAM tones. However, the data in experiment one of their manuscript show that each of the three subjects have a small but consistent increase in threshold between the 10 or 20 Hz AM rate compared against the 80 or 160 Hz AM rate for the 4000 Hz carrier. These standard AM rates from Lee [246] were the most similar to the present experiment and indicate a consistent pattern with our data. Similarly, Lemańaska and colleagues [252] describe that their data demonstrate a near-miss to Weber's law (or a change in threshold depending upon standard AM rate).

While the results were qualitatively consistent with previous reports [233,235,246,252], AM rate discrimination thresholds were much higher on average. In the present study, thresholds at 50% AM depth were 2-7 times greater than those of Lemańaska and colleagues [252] and Füllgrabe and Lorenzi [235], and roughly 10 times higher than those from Lee [246]. It is important to note that several studies using SAM tones at 100% AM depth [252], with varying AM depth [235], SAM noise with 100% AM depth [252], and square wave AM with varying AM depth [233] also showed higher thresholds than those reported by Lee [246]. One important difference between the present study and that of Lee [246]and Lemańaska and colleagues [252] is that prior studies used 100% modulation depth, whereas here the two modulation depths were 50% and 20%. Decreasing AM depth by 30% for SAM noise resulted in an approximate 2-3 fold increase in thresholds in previous studies [233,235], suggesting that the difference between our results and previous results could have been driven at least in part by AM depth differences. Results from our experiment led to a similar or slightly smaller magnitude effect of AM depth. One other possibility for this difference from previous reports is that listeners used different listening strategies for each standard rate, and interleaving standard AM rates in this study made this task more difficult than previous tasks. In agreement with this idea, the study by Lee [246] found that when carrier frequencies were randomized within the same block of trials, thresholds for discrimination increased, although thresholds were still lower than in the present study. It is also possible that loudness across AM rate played a role in the results of this experiment. The potential impact of loudness cues is explored in the Discussion of Experiment 2. Finally, it should be noted that the listeners in the present experiment did not receive extensive training on the task.

III. EXPERIMENT 2: COMPARING TEMPORAL ENVELOPE ACROSS PLACES-OF-STIMULATION WITH ASYMMETRIC TEMPORAL ENCODING

A. Motivation

Experiment 1 established that decreased modulation depth of SAM tones led to poorer performance on an AM rate discrimination task for AM rates presented sequentially (over three intervals) in NH listeners. In this experiment, we applied these findings to examine whether the ability to make simultaneous comparisons across two places-of-stimulation might be affected by temporal envelope representations in one cochlear place-of-stimulation.

It was predicted that the auditory system relies on similar quality of representations of temporal information between places-of-stimulation. Thus, if the temporal encoding is worse in one place-of-stimulation, then the ability to make these comparisons should be limited by the state of temporal encoding in the worse place-of-stimulation. Therefore, it was hypothesized that if AM depth was reduced in one of two places-of-stimulation, then sensitivity to differences in AM rate presented between places-of-stimulation would decrease. Further, it was predicted that if listeners compared AM rates at the same place-of-stimulation across the ears, listeners would have greater sensitivity to differences in AM rate overall due to the presence of an additional binaural beat cue in the envelope [253]. Binaural beats refer to the perception of a moving or diffuse sound source when the AM rate differs by a small number of cycles per second, introducing an interaural timing difference that varies over time. The addition of a contralateral cue in a previous rate discrimination experiment has also improved ability to discriminate AM rate when presented in sequence, as in Experiment 1 [239], supporting the notion that binaural AM rate discrimination is improved relative to tasks like that used in Experiment 1.

B. Methods

1. Listeners

Eleven listeners (age 18-26; mean 21.2 years) met the same hearing screening criteria as in Experiment 1, but with the addition of having 10 dB or less asymmetry in absolute pure-tone thresholds across the ears.

2. Stimuli & procedures

Stimuli were presented using the same equipment as Experiment 1. SAM tones of 600 ms duration with carrier frequencies of 4000 or 7260 Hz were played to the left and/or right ear (depending on the condition being tested). The starting phase of the envelope for each SAM tone was randomly selected from a uniform distribution between 0 and 2π radians for each presentation. The level was set to 65 dB(A) overall for any stimuli presented monaurally and attenuated by 6 dB in both ears when presented binaurally to result in similar loudness for monaural and binaural stimuli.

This experiment used a one-interval, two-alternative forced-choice task, where the listener was presented with SAM tones in two different cochlear places-of-stimulation and they responded by indicating whether the two AM rates were the same or different. Listeners were given visual feedback after each response. The AM rate in one place-of-stimulation was a fixed standard of either 10 or 90 Hz and the other (variable) place-of-stimulation received an AM rate equal to or greater than the standard rate. A psychometric function was measured over varying values of log(Δf/f). There was a 0.5 probability of both places-of-stimulation having an equal AM rate on each trial.

There were three possible place-of-stimulation pairing configurations: Same place-of-stimulation across-ears, different place-of-stimulation across-ears, and different place-of-stimulation within-ears. Both the task and AM pairing configurations are outlined in Fig. 2 for a "different" trial (where the AM rate differed between places-of-stimulation).

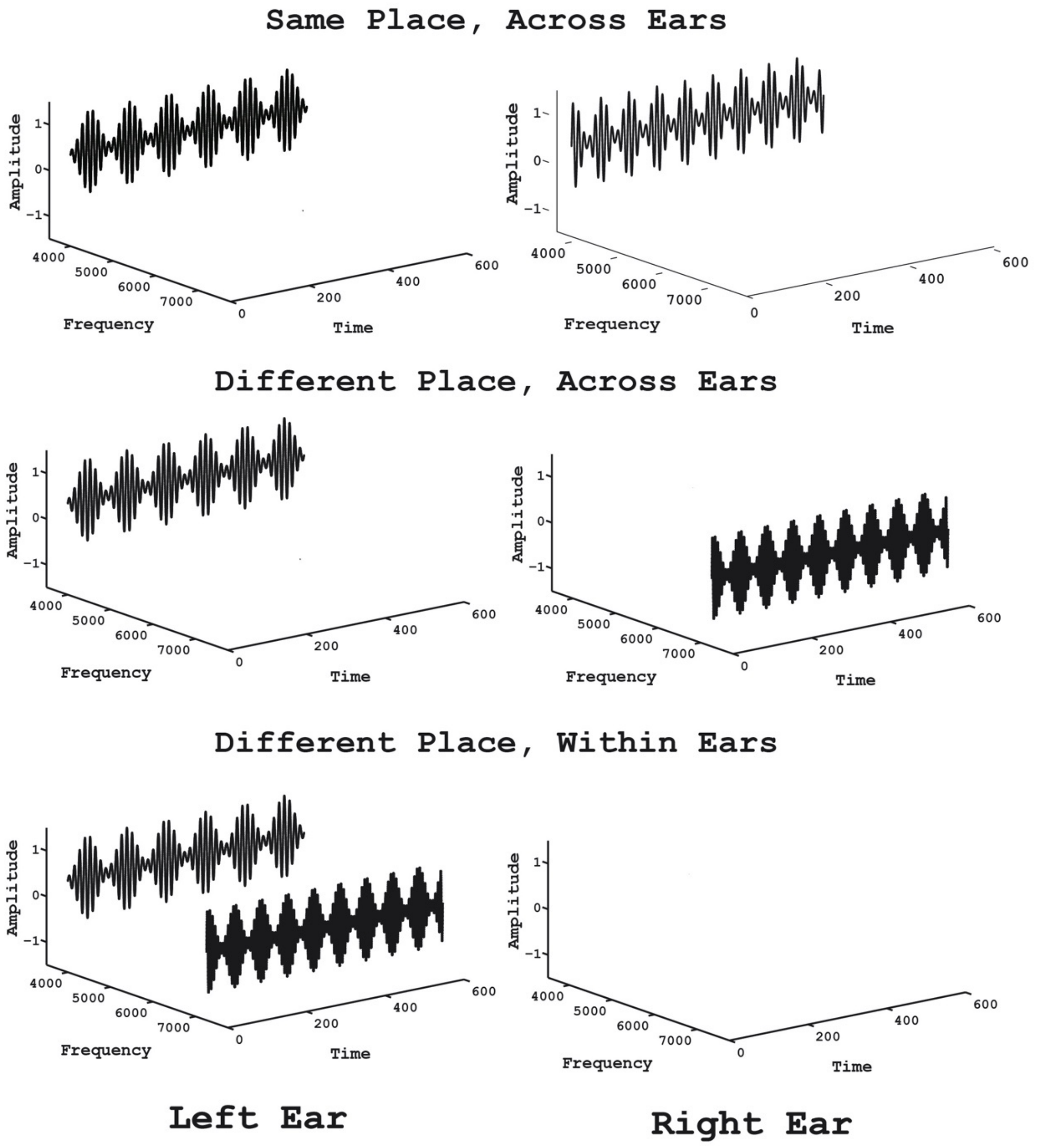


FIG. 2. Illustration of tone pair configurations. The x-axis represents time (each stimulus had a duration of 600 ms). The y-axis represents carrier frequency (of either 4000 or 7260 Hz). The z-axis represents relative amplitude of the stimulus. Each row represents a different AM rate pairing condition. The left and right column correspond to the left and right ear, respectively. Same Place, Across Ears: Carrier frequencies were equal in both ears (either 4000 or 7260 Hz). This configuration presumably resulted in the perception of binaural beats and represents

temporal envelope comparisons that occur for matched place-of-stimulation across the ears. Different Place, Across Ears: Different carrier frequencies were used in each ear and this configuration represents temporal envelope comparisons completed across spectrum and the ears. Different Place, Within Ears: Different carrier frequencies were used within the same ear and this configuration represents across-spectrum comparisons of the temporal envelope.

A split-plot experimental design was used with AM configuration as the whole-plot variable, and the AM rate pairing configuration and AM depth as the split-plot variables (Table I). Within each listener, the place-of-stimulation containing the variable AM rate was counterbalanced. For example, in the same place, across ears condition depicted in Fig. 2, half of the time the variable AM rate was presented on the left side, and the other half on the right side. By extension, in the first and second halves of the trials, the standard AM rate was presented to the right and left, respectively. This was done to account for any differences in performance depending on where the standard AM rate was delivered. Location of the standard AM rate was blocked to prevent confusion for listeners (for example, see Table I).

TABLE I. Example experimental conditions. Each row corresponds to a block of 420 trials. On each trial, the variable AM rate place-of-stimulation would present the standard AM rate with a probability of 0.5 or one of the variable rates shown in Table II. In this example, the left ear with a 4000 Hz carrier was chosen for the AM depth manipulation; then in half of all conditions the AM depth was 50% and the rest was 20% for that ear. In the same place, across ears configuration, the other place-of-stimulation was 4000 Hz in the right ear. For the different place, across ears configuration, the other place-of-stimulation was 7260 Hz in the right ear. For the different place, within ears configuration, the other place-of-stimulation was 7260 Hz in the right ear. In all of these cases, the "other place-of-stimulation" was presented with 50% AM depth.

		Variable AM Rate			**Std. AM Rate**		
Block	**Place-of-Stimulation Pairing**	**Ear**	**Frequency (Hz)**	**AM Depth (%)**	**Ear**	**Frequency (Hz)**	**AM Depth (%)**
1	Same Place, Across Ears	L	4000	50	R	4000	50
2	Same Place, Across Ears	L	4000	20	R	4000	50
3	Same Place, Across Ears	R	4000	50	L	4000	50
4	Same Place, Across Ears	R	4000	50	L	4000	20
5	Different Place, Across Ears	L	4000	50	R	7260	50
6	Different Place, Across Ears	L	4000	20	R	7260	50
7	Different Place, Across Ears	R	7260	50	L	4000	50
8	Different Place, Across Ears	R	7260	50	L	4000	20
9	Different Place, Within Ears	L	4000	50	L	7260	50
10	Different Place, Within Ears	L	4000	20	L	7260	50
11	Different Place, Within Ears	L	7260	50	L	4000	50
12	Different Place, Within Ears	L	7260	50	L	4000	20

All AM pairing configurations were repeated, once with 50% AM depth in both places-of-stimulation, and once with 20% AM depth in one place-of-stimulation and 50% AM depth in the other. The carrier frequency and ear of the place-of-stimulation with 20% AM depth, and order of AM pairing configurations were counterbalanced across listeners. Only one place-of-stimulation was chosen at the outset for each listener to receive the AM depth manipulation (either left or right ear with 4000 or 7260 Hz). The corresponding places-of-stimulation in all other configurations were then chosen accordingly. An example of the order and conditions tested is shown in Table I. For this example, the left ear with a 4000 Hz carrier was chosen for the AM depth manipulation; then in half of all conditions the AM depth was 50% and the rest were 20% for that ear. In the same place, across ears configuration, the other place-of-stimulation was 4000 Hz in the right ear. For the different place, across ears configuration, the other place-of-stimulation was 7260 Hz in the right ear. For the different place, within ears configuration, the other place-of-stimulation was 7260 Hz in the right ear. In all of these cases, the “other place-of-stimulation” was presented with 50% AM depth.

A seven-point psychometric function (see Table II for AM rates) was collected, with each point representing a change in AM rate relative to 10 or 90 Hz standard AM rates. Thirty repetitions were collected per variable AM rate in each condition, resulting in 840 trials per condition, with three possible AM pairing configurations and two possible AM depth conditions. This resulted in a total of 5040 stimulus presentations throughout the experiment. Testing took approximately 6 hours, with 1-2 hours of training. Testing was divided over 2-3 sessions on different days.

TABLE II. AM rates corresponding to each log(Δf/f) for Experiment 2. The columns show the extent of AM roving for 10 and 90 Hz standard AM rates. The N/A row refers to values taken on by the standard AM rate (technically log(Δf/f) = -∞).

	Std. Rate: 10 Hz			**Std. Rate: 90 Hz**		
log(Δf/f)	**Min. (Hz)**	**Mean (Hz)**	**Max. (Hz)**	**Min. (Hz)**	**Mean (Hz)**	**Max. (Hz)**
N/A (Std.)	7.50	10.00	13.33	67.50	90.00	120.00
-0.71 ± 0.28	8.95	11.94	15.92	80.59	107.45	143.27
-0.42 ± 0.28	10.37	13.82	18.43	93.31	124.41	165.88
-0.12 ± 0.28	13.16	17.54	23.39	118.40	157.86	210.48
0.17 ± 0.28	18.65	24.87	33.16	167.87	223.83	298.44
0.47 ± 0.28	29.49	39.33	52.43	265.44	353.93	471.90
0.76 ± 0.28	50.87	67.83	90.44	457.85	610.47	813.96
1.06 ± 0.28	93.03	124.05	165.39	837.31	1116.41	1488.55

The goal in every trial was to correctly identify whether the AM rate was the same in both places-of-stimulation. If AM rates were different, the variable place-of-stimulation had a higher AM rate. Trials for both of the standard AM rates were interleaved within one block. The AM rate on each trial was roved according to Table II. Interleaving standard AM rate and rate roving were both completed to discourage listeners from using a single-place-of-stimulation strategy to complete the task (e.g., responding “different” for high AM rates and “same” for low AM rates). Rate roving was not completed for one listener (THK) because they were tested before rate roving was implemented. We attempted to ensure that participants were completing the task by comparing the two SAM tones to one another. If participants exhibited non-monotonic functions during the beginning of a new trial, they were given a break, a practice block where only the two highest values of log(Δf/f) were included, and data for that block were collected de novo. Unfortunately, some raw data still exhibited non-monotonic trends.

C. Results

In this experiment, pairs of SAM tones were presented across ears at the same or different places-of-stimulation, or within the same ear at different places along the cochlea (see Fig. 2). Examples of raw data from two listeners in the same place, across ears condition are presented in Fig. 3 showing that performance could be quite variable. Some listeners performed extremely well (e.g., listener TIG), with a small amount of bias and large number of correct responses. Others performed very poorly (e.g., listener THK), with substantial bias toward choosing “different” and small number of correct responses. As a reminder, THK was tested before rate AM roving was implemented. Individuals varied considerably, and some psychometric functions were non-monotonic. Raw data from all listeners are provided in Supplementary Fig. 2. Notably, non-monotonic psychometric functions were more likely to occur in the 20%:50% AM depth condition.

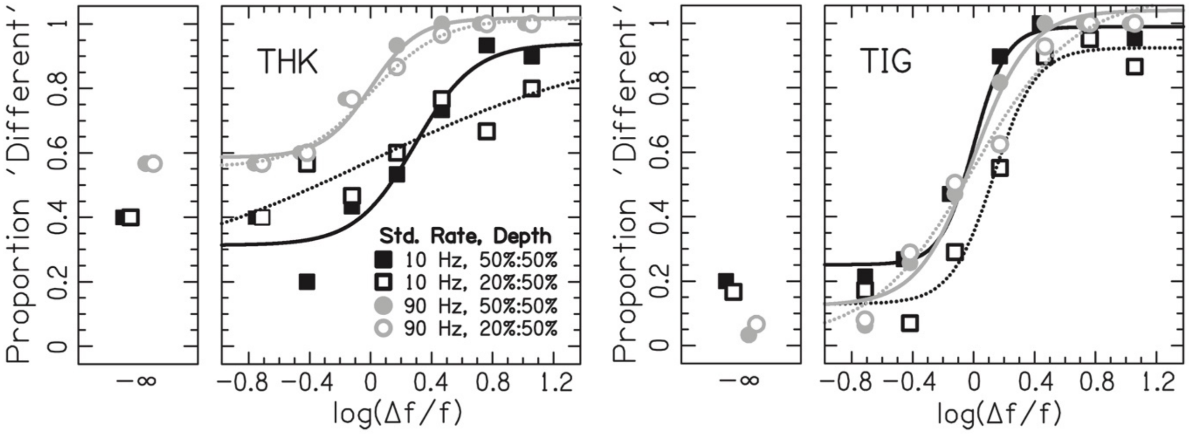


FIG. 3. Example raw data from two listeners in the same place, across ears condition. Listener codes for each individual are given in the top left corner. Open shapes and dotted lines correspond to the 20%:50% AM depth condition, and closed shapes and solid lines correspond to the 50%:50% AM depth condition. Performance for 10 Hz standard rate is shown in black and 90 Hz standard rate are shown in grey. The y-axis corresponds to the proportion of "different" responses across all trials. The x-axis corresponds to the difference in AM rate between the standard and variable AM rate (see Table 1 for values of AM rates in Hz). The small panel on the left (a log(Δf/f) of -∞) represents the proportion of "different" responses when AM rates were equal. Ideal performance occurs when the proportion of "different" responses is 0 for the small left panel, and 1 for the larger, right panel. Sensitivity (d') can be calculated directly from the raw data. For raw data for all listeners, please see Supplementary Fig. 2.

Figure 4 shows differences in sensitivity by AM pairing configuration, AM depth in each ear, and log(Δf/f). Proportion of correct responses were converted to d' [81]. The data in Fig. 4 suggest that sensitivity was greater for standard AM rates of 90 compared with 10 Hz, demonstrated in all three AM pairing configurations by a higher asymptotic sensitivity for the 90

Hz standard AM rate conditions. Further, average sensitivity was greater in the 50%:50% compared against the 20%:50% AM depth conditions for all AM pairing configurations, demonstrated by a rightward shift for the open symbols and dashed lines.

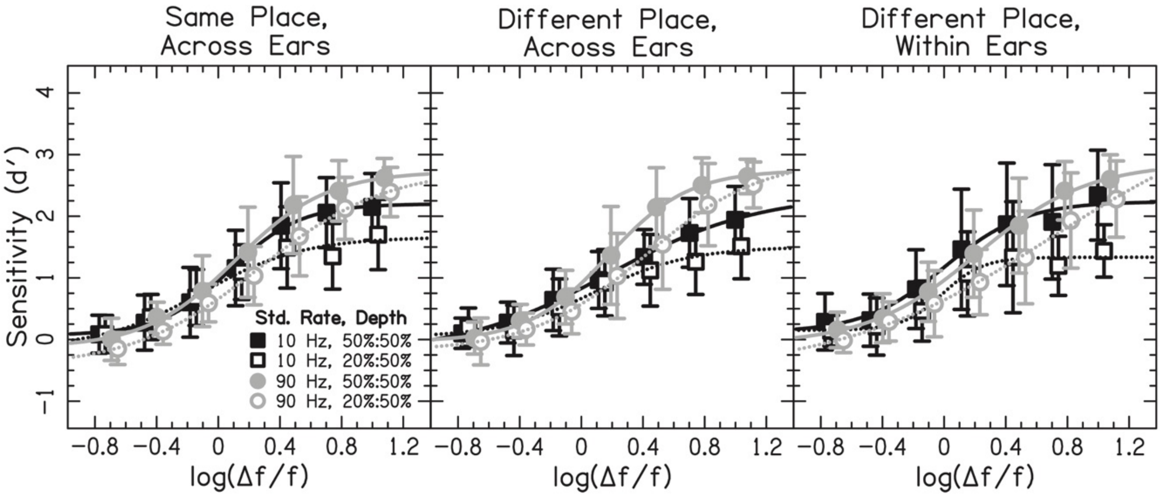


FIG. 4. Mean ± one standard deviation sensitivity across listeners for all three conditions. Open shapes and dotted lines correspond to the 20%:50% AM depth condition, and closed shapes and solid lines correspond to the 50%:50% AM depth condition. Performance for 10 Hz standard rate is shown in black and 90 Hz standard rate are shown in grey. Each panel corresponds to a different AM pairing condition (see Fig. 2). The y-axis corresponds to sensitivity in d'[81]. The x-axis corresponds to the difference in AM rate between the standard and variable AM rate (see Table I for values of AM rates in Hz).

To determine the effects of log(Δf/f) on d' and interaction with blocking variables (standard AM rate, AM depth, and AM pairing condition), log(Δf/f) vs. d' were fit with a four-parameter logistic mixed effects model with each combination of blocking variables, where the mid-point was a random effect by listener (i.e., data from Fig. 4). Unfortunately, the covariance structure did not converge, and it would not be appropriate to report the results of the model. As a second approach, mixed effects ANOVA was completed with d' as the dependent variable and log(Δf/f), AM pairing configuration, AM rate standard, and reduction in AM depth as fixed effects, with listener as a random-effect. The effect of carrier frequency and ear for the place-of-stimulation across listeners were also explored as fixed effects. Mixed effects ANOVAs fit fewer parameters compared to the four-parameter logistic model, a substantially less complex covariance structure, and converged without issue in this case. Results indicated that log(Δf/f) had a significant effect on d' [F(6,842) = 402.711, $p < .0001$] on average across all blocking variable combinations. This was unsurprising because sensitivity to differences in AM rate should increase as the degree of difference between standard and variable AM rates increases.

To confirm that using different carrier frequencies representing the place-of-stimulation with poorer phase locking did not affect performance, a fixed effect of carrier frequency was estimated in the ANOVA. Effects of carrier frequency were tested because previous experiments investigating the ability to discriminate interaural time differences in the envelope has been reported to change depending upon carrier frequency (e.g., [242]). Results indicated that carrier frequency with the AM depth manipulation did not have a significant effect across listeners [F(1,9) = 0.319, $p = .586$]. This result, along with the non-significant effect of carrier frequency in Experiment 1, suggests that listeners were able to compare AM rates regardless of carrier frequencies employed in this experiment. It is important to note that only 11 people participated in Experiment 2 and carrier frequency with the AM depth manipulation was an across-listener variable, so variability across listeners could have masked an effect.

1. Effects of asymmetric temporal encoding

The primary purpose of this experiment was to explore if sensitivity to differences in AM rate between places-of-stimulation decreases when AM depth is reduced for one place-of-stimulation. Reduced AM depth significantly decreased d' [F(1,842) = 81.920, p < .0001], suggesting that sensitivity to differences in AM rate between places-of-stimulation decreased on average when depth was reduced in only one of the two places-of-stimulation. There was also a significant interaction between log(Δf/f) and reduced AM depth [F(1,842) = 4.569, p < .001], suggesting that the slope of the psychometric function changed when AM depth was reduced. Due to the limitations of using an ANOVA compared with a four-parameter logistic model, a direct comparison of the slope of psychometric functions fitted with the same model between AM depth conditions was not possible.

Mean thresholds for both standard AM rates and AM pairing configurations are reported in Fig. 5. Thresholds (.707 proportion correct) were estimated by fitting a four-parameter logistic function to the proportion of "different" responses across variable AM rates within each subject (data and fitted curves shown in Supplementary Fig. 2). Five thresholds (THQ different place, across ears; THQ different place, within ears; THS different place, within ears; THU different place, across ears; TJF same place, across ears) could not be estimated because the logistic curve did not cross .707 for log(Δf/f) values tested in the experiment (Supplementary Fig. 2). All thresholds in the same AM pairing configuration (for 20%:50% or 50%:50% AM depth) where thresholds could not be estimated were excluded from analyses for that subject. Thresholds are generally consistent with the results of the ANOVA, although the considerable variability across subjects is clear from Fig. 5.

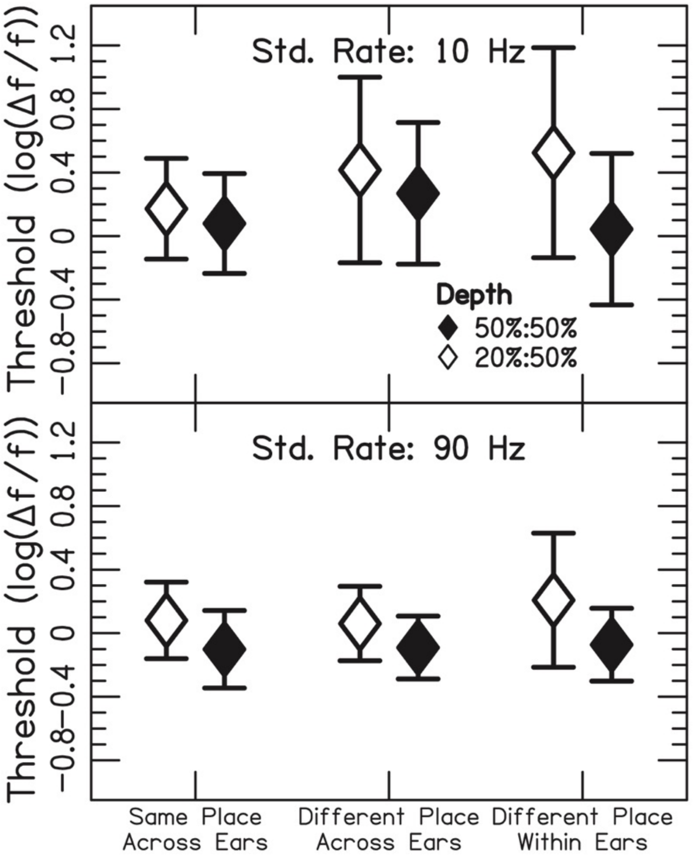


FIG. 5. Mean ± one standard deviation thresholds for each standard AM rate. Top and bottom rows correspond to standard AM rates of 10 and 90 Hz, respectively. Open and closed shapes correspond to the 20%:50% and 50%:50% AM depth conditions. The y-axis corresponds to threshold (defined as .707 proportion correct). The x-axis corresponds to the AM pairing condition (see Fig. 2). Thresholds were highly variable across listeners (for raw data for each listener, see Supplementary Fig. 2).

2. Effects of standard AM rate

In general, for AM rates above 50-90 Hz the listener experiences a pitch cue. Thus, the listening strategy in this experiment was expected to change depending upon the standard AM rate. Normalizing differences in AM rate using log(Δf/f) made it possible to directly compare performance between standard AM rates, and indicated a significant effect of standard AM rate [F(1,842) = 52.926, p < .0001] and interaction between log(Δf/f) and standard AM rate [F(1,842) = 15.966, p < .0001], where the 90 Hz standard AM rate had a higher average d'. The addition of a pitch cue may account for enhanced performance at 90 Hz. The interaction between log(Δf/f) and standard AM rate suggests that the slope of the psychometric function differed between 10 and 90 Hz. However, the interaction between standard AM rate and AM depth was not significant [F(1,842) = 0.006, p = .940], implying that the ability for listeners to compare AM rate between places-of-stimulation may be impaired when AM depth is reduced in one place-of-stimulation regardless of standard AM rate.

3. Effects of AM pairing configuration

Pairs of SAM tones were presented in three different configurations, representing the three primary ways that envelope fluctuations could be compared in real-world listening (outlined in Fig. 2). It was predicted that the same place, across ears condition might result in better performance overall because of the addition of a binaural beat cue [253]. There was not sufficient statistical evidence to conclude that AM pairing configuration affected d' [F(2,842) = 2.076, p = .127]. However, there was a significant interaction between standard AM rate and AM pairing configuration [F(2,842) = 3.901, p < .05]. This suggests that the ability to compare envelope fluctuations between places-of-stimulation may differ depending on which places-of-stimulation are compared for lower frequency envelope cues where pitch cues are not available. For 10 Hz, d' was significantly higher for: different place, within ears compared to the different place, across ears [t(842) = 2.863, p < .05], same place, across ears compared to the different place, across ears [t(842) = 2.871, p < .05], but not the different place, within ears compared to same place, across ears [t(842) = 0.007, p = 1.000] configurations. For 90 Hz, d’ no significant difference was observed between any AM pairing configurations: within ears, different place compared to the different place, across ears [t(842) = -.894, p = .948], same place, across ears compared to the different place, across ears [t(842) = -.064, p = 1.000], and the different place, within ears compared to same place, across ears [t(842) = -.830, p = .962] configurations. Post-hoc tests were completed using estimated marginal means with Tukey adjustment for multiple comparisons.

The 95% confidence intervals for change in threshold between the 20%:50% and 50%:50% AM depth conditions were created for each AM pair configuration and standard AM rate. Confidence intervals are displayed in Fig. 6 and suggest a change in the effect of AM depth with AM rate and AM pairing configuration. Results from confidence intervals are generally consistent with the results of the ANOVA. Though there was not a statistically significant interaction between reduced modulation depth and AM pairing configuration on average, it can be seen from Fig. 6 that a change in threshold between the 20%:50% and 50%:50% AM depth conditions was only observed for the different place, within ears condition for the standard AM rate of 10 Hz. With the standard AM rate of 90 Hz, there was a positive difference between the 20%:50% AM depth conditions in all AM pair configurations. Note that the ANOVA was completed using the entire psychometric function for each subject, while confidence intervals only show the change in threshold between 20%:50% and 50%:50% AM depth.

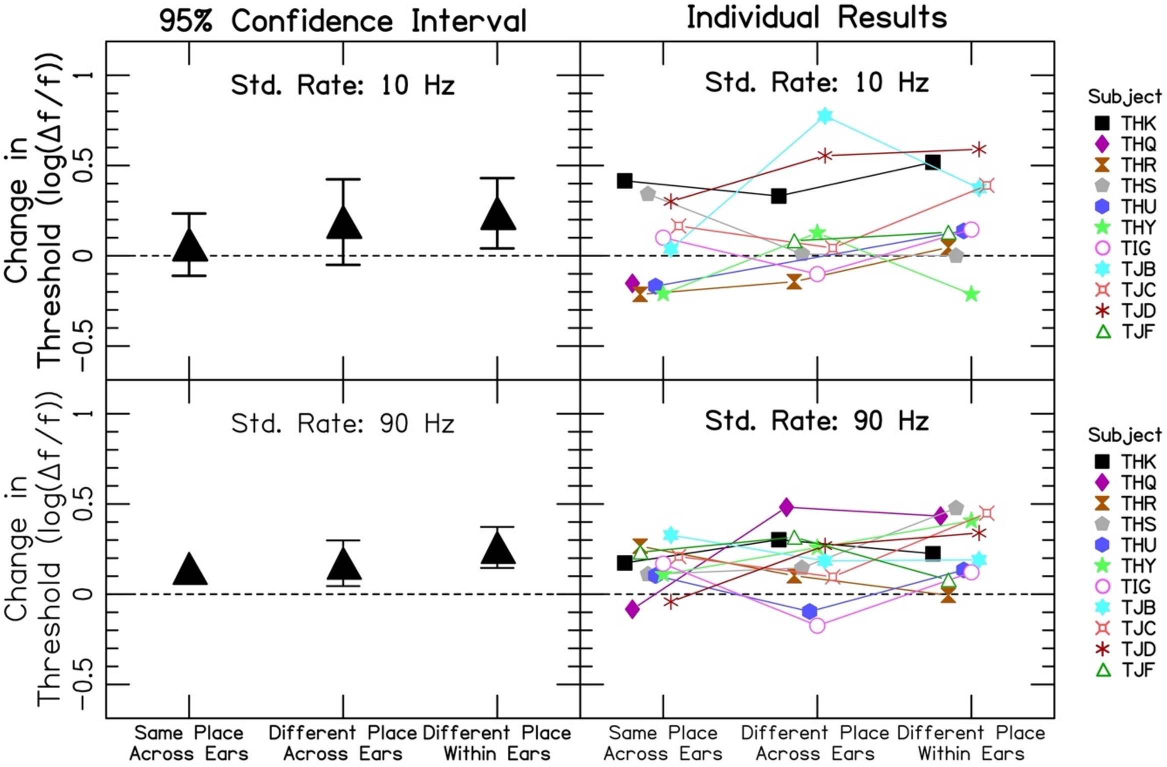


FIG. 6. 95% confidence intervals and individual results for change in threshold (.707 proportion correct) between the 20%:50% and 50%:50% AM depth conditions. The top and bottom rows correspond to standard AM rates of 10 and 90 Hz, respectively. The left and right columns correspond to 95% confidence intervals and individual results, respectively. The x-axis corresponds to the AM pairing configuration (see Fig. 2). The y-axis corresponds to the change in threshold between the 20%:50% and 50%:50% AM depth conditions. Values above zero indicated that listeners worsened on the task when AM depth was reduced from 50% to 20% in one place-of-stimulation. Confidence intervals were computed assuming a t distribution. Values above zero imply a significant increase in threshold from the 50%:50% to 20%:50% AM depth conditions. Thresholds for listeners were excluded if they fell above or below the values for log(Δf/f) tested in the experiment. Listener codes and corresponding symbols are given on the far right.

D. Discussion

Results from this experiment demonstrated that when comparing the temporal envelope between two places-of-stimulation, if the AM depth was reduced in one place-of-stimulation, then sensitivity to differences in rate of AM decreased. This finding held for both standard AM rates of 10 and 90 Hz. Moreover, there was an interaction between AM pairing configuration and standard AM rate. For 10 Hz standard AM rate, sensitivity is greatest when comparing temporal envelope fluctuations at a different place, across ears. For 90 Hz standard AM rate, sensitivity was similar across all AM pairing configurations.

Most listeners reported that Experiment 2 was exceptionally difficult. Thus, some preventative stimulus manipulations were foregone to prevent distraction or increased difficulty. It is possible that listeners could have relied on changes in loudness to discriminate between different AM rates. Zhang and Zeng [254] evaluated the effects of AM rate on loudness perception of SAM tones in NH listeners. Their results suggest that loudness may have changed for the AM rates employed in the present study. The best way to account for this would have been to rove the level in each cochlear place-of-stimulation independently. Instead, level was fixed within each block. However, recall that in Experiment 2 the AM rate was roved between trials and standard rates of 10 and 90 Hz were interleaved for all listeners except THK. If listeners experienced changes in loudness associated with AM rate between places-of-stimulation, the average loudness and difference in loudness should have changed on each trial. Further, the root-mean square level between different AM depths and AM rates was fixed. Finally, when asked to describe how they completed the task, most listeners indicated that they formed categories by AM rate, with the lowest rates corresponding to noticeable changes in loudness over time that could be compared between cochlear places-of-stimulation, and the highest AM rates corresponding to changes in pitch.

To address whether loudness could have confounded the results from Experiment 2, the loudness for individual SAM tones was estimated across AM rate using the loudness model from Moore and colleagues [255]. Results from the model are plotted in Fig. 7. Stimuli with a 4000 Hz carrier resulted in substantially greater estimated loudness than 7260 Hz. Estimated loudness was relatively consistent for most AM rates at 20% or 50% AM depth for both center frequencies. Estimated loudness decreased slightly as AM rates were increased from 10 to 100 Hz and increased substantially for the highest AM rates (mean of 610.47 and 1116.41 Hz) as in Zhang and Zeng [254], especially for 50% AM depths. Critically, in the results of Zhang and Zeng [254] AM rate had different effects on loudness at different modulation depths for SAM noise. Figure 7 indicates that estimated loudness was relatively consistent across rate for 20% and 50% AM depth, suggesting that loudness remained relatively consistent across AM rate for 50% and 20% AM depth for the five smallest values of log(Δf/f).

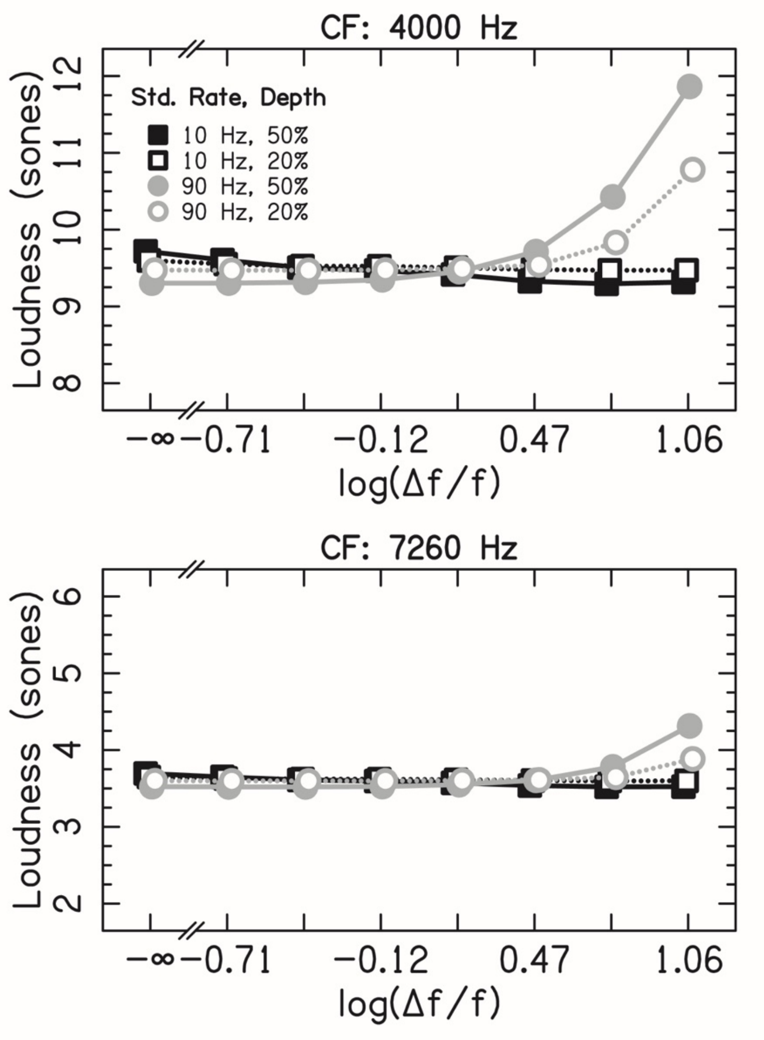


FIG. 7. Estimated loudness by AM rate. Loudness was estimated using the model from [255]. Results are plotted as in Fig. 4 for comparison, but each point represents a single AM rate (for specific AM rates, see Table II). The y-axis corresponds to loudness in sones. The top and bottom panels correspond to SAM tones with carrier frequencies of 4000 and 7260 Hz, respectively. Please note the difference in scale for each panel (7260 Hz resulted in much less loudness overall). Different scales were used to make the change in loudness across AM rate visually apparent.

To account for results that may have been driven by loudness differences, the results from three additional ANOVAs are included in Table III. For one ANOVA, the analysis was conducted with the two largest values of log(Δf/f) excluded to account for cases where loudness might have provided a useful cue. Since loudness differences did not occur for the largest values of log(Δf/f) for the standard AM rate of 10 Hz, two additional ANOVAs were computed for standard AM rates of 10 Hz and 90 Hz separately to evaluate the effects of log(Δf/f), AM depth, and AM pairing configuration. For the ANOVA corresponding to 90 Hz standard AM rate alone, the two highest values of log(Δf/f) were excluded from the analysis. Together, the new analyses attempting to account for estimated loudness changes due to AM rate suggest that loudness differences in the largest values of log(Δf/f) for standard AM rate of 90 Hz may have driven greater sensitivity compared to the standard AM rate of 10 Hz. Additionally, the second set of ANOVAs confirm that for the standard AM rate of 10 Hz, differences in sensitivity between AM pairing configurations were driven in part by the largest two values of log(Δf/f).

TABLE III. ANOVAs summarized by main effects and interactions. The second and fourth columns (labeled 10 and 90 Hz, and 90 Hz only) were computed excluding data for the two highest values of log(Δf/f).

	Original ANOVA		10 and 90 Hz		10 Hz only		90 Hz only	
Effect	**F-Statistic**	**p**	**F-Statistic**	**p**	**F-Statistic**	**p**	**F-Statistic**	**p**
log(Δf/f)	F(6,842) = 402.711	<.0001	F(4,596) = 241.261	<.0001	F(6,422) = 152.927	<.0001	F(4,298) = 199.959	<.0001
AM Depth	F(1,842) = 81.920	<.0001	F(1,598) = 32.900	<.0001	F(1,422) = 45.961	<.0001	F(1,298) = 37.688	<.0001
Std. Rate	F(1,842) = 52.926	<.0001	F(1,598) = 1.766	.184	--	--	--	--
Pairing Configuration	F(2,842) = 2.076	.127	F(2,598) = 2.224	.109	F(2,422) = 6.252	<.01	F(2,298) = 0.194	.824
Center Frequency	F(1,9)= 0.319	.586	F(1,9)= 0.403	.541	F(1,9)= 0.064	.806	F(1,9)= 0.088	.773
log(Δf/f) x AM Depth	F(1,842) = 4.569	<.001	F(4,598) = 3.653	<.01	F(6,422) = 5.429	<.0001	F(4,298) = 3.154	<.05
log(Δf/f) x Std. Rate	F(1,842) = 15.966	<.0001	F(4,598) = 3.653	<.01	--	--	--	--
AM Depth x Std. Rate	F(1,842) = 0.006	.940	F(1,598) = 3.223	.073	--	--	--	--
Std. Rate x Pairing Configuration	F(2,842) = 3.901	<.05	F(2,598) = 1.954	.143	--	--	--	--

It is also possible that loudness over time of stimulus presentation differed according to log(Δf/f). This reflects one possible cue that listeners could have used to perform the experiment. At very low rates, listeners perceived fluctuating loudness over time for a single SAM tone, so when multiple, low-rate SAM tones are presented simultaneously, it seems likely that the perceived loudness will vary with the relationship between the phase of each SAM tone. Additional experiments would be required to determine the role of loudness changes over stimulus presentation in this task. Specifically, loudness changes at low AM rates due to additive loudness from two components depending upon phase could also apply to a previous experiment where SAM tones were presented to different places-of-stimulation and AM phase was manipulated but AM rate was kept constant [228]. In the present experiment, AM rate roving and randomization of phase would have made it difficult to make decisions on each trial according to the additive loudness from both SAM tones.

Removing the two highest values of log(Δf/f) would have kept the maximum AM rate in Experiment 2 at 471.90 Hz (see Table II). Experiments by Kohlrausch and colleagues [234] imply that these sidebands are not useful in detecting AM rate in a task like that used in Experiment 1, so it is possible that sidebands were not very useful in this experiment.

Removing the two highest values of log(Δf/f) would also decrease the utility of combination tones. Since combination tones depend upon the relationship between the AM rate and carrier, the detectability of these tones may change with AM rate. Thus, one possible explanation for the difference in slope between the psychometric function for 10 Hz compared to 90 Hz standard AM rates is an increased access to combination tones at larger values of log(Δf/f) for 90 Hz. We do not feel that there is sufficient evidence concerning which combination tones could contribute most greatly affect discrimination of AM rate, but combination tones would be most audible at higher rates. Therefore, preservation of effects after excluding the highest rates tested in Experiment 2 provides some evidence to suggest that the effects of AM depth at 90 Hz were not due to difference tones. Experiments with patients that use CIs could address this in more detail since CIs directly stimulate the auditory nerve and do not rely on the cochlear filtering that produces combination tones.

The goal for the task in Experiment 2 was to require listeners to compare pairs of AM rates presented simultaneously between cochlear places-of-stimulation. For the different place, within ears AM pairing configuration in Experiment 2, it is possible that NH listeners could have used an overlapping region on the basilar membrane to make decisions on each trial [256]. The best way to account for this overlap would have been to use a low-level masking noise. During pilot testing, several listeners reported the inclusion of masking noise as being distracting, so no masking noise was used. Moreover, the carrier frequencies used in Experiment 2 were quite disparate (4000 or 7260 Hz). Performance in the different place, within ears condition was worse than other the different place, across ear conditions for 10 Hz (Fig. 5). Thus, the monaural overlap in excitation on the basilar membrane did not appear to provide any substantial advantage over the different place, across ears AM pairing configuration. In the same place, across ears condition, it is possible that listeners could use sidebands which change with AM rate to complete the task. A listener could detect a non-zero interaural level difference using the sidebands in either ear. Similarly, a masking noise would have been needed to prevent the use of sidebands but was not included.

Results were considerably variable across listeners (for example, see Fig. 3; for raw data, see Supplementary Fig. 2). Variability in naïve NH listeners has been documented on a variety of psychophysical tasks [257–259], though not with this specific psychophysical paradigm or that used in Experiment 1 [233,235,246,252]. It is possible that the task difficulty contributed to variability across listeners. Fortunately, the design and statistical analyses implemented in this study account for differences across listeners. This result suggests that a within-subject design will be important to consider if this experiment is implemented in listeners with hearing loss. Further, variability in performance of listeners with hearing loss would be expected and could be compared against NH listeners.

It was predicted that if one AM pairing configuration resulted in the best performance, it would be the same place, across ears condition because of the addition of a binaural beat cue at low values of log(Δf/f). Binaural beat cues from the envelope result when small differences in AM rate exist between each ear and the same center frequency is used [253]. Thus, binaural beat cues could provide an additional cue in the same place, across ears AM pairing configuration beyond those available for the other AM pairing configurations. Surprisingly, the different place, across ears AM pairing configuration led to better performance than the same place, across ears AM pairing configuration for the 10 Hz standard AM rate. Note that sensitivity to interaural timing differences is worse for 32 Hz (lowest rate tested) compared to 128 Hz with SAM tones [242]. It may be that a binaural beat cue is not useful for very low standard AM rates, or that it simply was not sufficient to result in better performance. Note that the magnitude of the effect of AM depth was smaller for the same place, across ears compared to the different place, within ears AM pairing configurations. The overall better performance of 90 Hz was most likely due to the addition of a pitch cue or sharper slope of the envelope for 90 Hz compared to 10 Hz [242–245]. This result also supports the notion that listeners did not use sidebands to complete the task, since sidebands would have been most useful in the same place, across ears AM pairing configuration.

The paradigm in experiment 2 was similar to decorrelation detection experiments [229–231], except that the differences in envelope across places-of-stimulation were deterministic, resulting in changes in the quality of sounds with increasing rate. It is possible that the task could have been completed using the amount of decorrelation between each place of stimulation, which could have changed with AM rate. To assess this possibility, the cross correlation was computed as in previous studies [229] and is reported in Supplementary Fig. 3. The only changes in correlation occurred at the standard AM rate and changed systematically with envelope phase. The cross correlation is relatively constant across values of log(Δf/f). Since the proportion of "different" AM rates in this experiment was similar when the standard AM rate was presented to each place-of-stimulation and when small values of log(Δf/f) were used (Fig. 3 and Supplementary Fig. 2), it seems that decorrelation did not provide a useful cue for listeners to complete the task. It should be noted, however, that the magnitude of change in cross-correlation with envelope phase decreased when the AM depth was 20%:50%. Thus, differences in phase across places-of-stimulation could be less useful if the AM depth is reduced in one place-of-stimulation (e.g., interaural timing differences in the envelope).

Previous research evaluating thresholds for discriminating between interaural time differences in the envelope changed with frequency [260]. Thus, it was important to determine whether the center frequency of the ear that received the AM depth manipulation affected results. With respect to Experiment 2, the same place, across ears AM pairing configuration might then be expected to see a change in sensitivity according to carrier frequency. In this case no effect of carrier frequency was found, suggesting that results could be interpreted similarly regardless of carrier frequency. The difference in our results could be related to the fact that listeners in

Bernstein and Trahiotis [260] were older than the undergraduate students that participated in this experiment.

IV. GENERAL DISCUSSION

In everyday auditory environments, listeners with hearing impairment struggle in separating speech from noise. It is not obvious which factors contribute most to poor performance listening in noise, and it may be that factors vary from one individual to another. One particular factor that is thought to vary within and across individuals is the fidelity of temporal envelope encoding in the auditory nerve in either ear. Even within the same listener, there may be places-of-stimulation that have poor to excellent temporal envelope encoding. Specifically, periodic fluctuations in the temporal envelope can be used to form auditory streams when presented in a sequence [222] and differences between the temporal envelopes are detectable within- [230] and across-ears [129,229] when presented simultaneously. While many factors can contribute to poorer temporal envelope encoding, at least some of these factors affect virtually all listeners with hearing-impairment.

This study investigated a NH simulation of poor temporal encoding at specific places-of-stimulation. The depth of AM was reduced for stimuli to represent poorer temporal envelope encoding. Experiment 1 demonstrated that performance discriminating between intervals with different AM rates worsened when AM depth was reduced from 50% to 20%. In Experiment 2, a new psychophysical task where listeners compared rates of AM across ears and places-of-stimulation was used (see Fig. 2). Results indicated that reducing AM depth from 50% to 20% worsened sensitivity (Fig. 6). Together, these results suggest that the ability to compare information in the envelope across places-of-stimulation might be impaired when the temporal encoding is poor for one of the places-of-stimulation.

The critical part of this conclusion is that when only one of the places-of-stimulation encodes temporal information poorly, the auditory system cannot make accurate comparisons of temporal envelope fluctuations. From the review by Grose and colleagues [220], it is clear that there are many examples where the temporal envelope can be used to discriminate between sound sources. This experiment focused specifically on the case of short duration, periodic sounds, with the unique advantage of comparing performance within and across the ears. This is the first study to systematically compare performance discriminating between AM rate within and across the ears and cochlear place. Importantly, while no effect of AM pairing configuration was shown here on average, sensitivity changed depending upon the AM pair configuration for low standard AM rates (10 Hz).

In the present study, listeners compared one place-of-stimulation against a 10 or 90 Hz standard AM rate, reflecting two speech-relevant perceptual processes: slower fluctuations that result in e.g., word segmentation cues and faster fluctuations that result in pitch cues [219]. Results from both experiments showed an effect of standard AM rate on the ability to judge differences in AM rate when presented sequentially (Experiment 1) and simultaneously presented across places-of-stimulation (Experiment 2), with listeners being more sensitive to changes at 90 Hz relative to 10 Hz.

A. Relation to listeners with hearing impairment

Listeners with hearing impairment can have highly-varying temporal representations in the auditory nerve. For example, for listeners that receive a CI and have either NH or use a hearing aid in the other ear, electric and acoustic information must be integrated across the ears to distinguish between different sound sources. Similarly, for listeners that receive a hybrid CI, highly phase-locked, electrically encoded information must be compared against acoustically encoded information within the same ear. For CI users, greater distance between the electrode array and auditory nerve has been related to increases in threshold, and likely results in poorer spectro-temporal representations [125]. Similarly, long durations of deafness are associated with poorer phase locking and loss of auditory nerve fibers, as well as deterioration of dendrites [106,109,116]. Thus, it is apparent that temporal representations can vary highly within the same individual; yet little research has systematically focused on the implications of differing temporal representations.

The present study suggests that the ability to make spectro-temporal comparisons across pairs of cochlear locations worsens when temporal encoding is poor at one place-of-stimulation. This result implies that comparisons of temporal information are limited by the worst temporal representation. Results from Experiment 2 suggest that the ability to compare temporal information within the same ear may be more heavily impacted by poor temporal encoding than comparisons across ears at low rates of modulation (see Fig. 6). This may be due to the addition of a binaural beat cue [253] in across-ear conditions, and requires further investigation in listeners with hearing impairment. It is important to note that simply because the change in threshold between AM depth conditions was smaller in the across-ear compared to the within-ear condition, this does not suggest that the functional implications are less severe. That is, the binaural system relies on precise timing information to detect the location of sounds.

Recent work in bilateral CI users suggests that sensitivity to binaural cues may be predicted by the ear with worse sensitivity to temporal information [9]. These results are similar in spirit to studies in listeners with NH, CIs, or single-sided deafness showing that as the envelope attack slope increases, which should result in highly-synchronous firing of the auditory nerve, listeners become increasingly sensitive to interaural timing differences in the envelope

[242–245]. None of the studies in NH have investigated what occurs when temporal envelope encoding is asymmetric across the ears, however. Thus, it may be important to more thoroughly investigate changes in binaural sensitivity under asymmetric temporal envelope encoding.

Some previous studies have demonstrated improved performance of individuals with hearing loss compared to NH in tasks involving AM detection using low AM rates (e.g., [261]). Additionally, some experiments have suggested that hearing loss might improve temporal resolution for stimuli with AM (e.g., [262]). It has been suggested that differences in performance are due to loudness recruitment and a lack of compression associated with hearing loss [261,263]. Some studies show an improvement in AM detection for listeners with hearing-impairment compared to NH when the same sound pressure levels (not sensation levels) are used [263], while others show no difference [261]. Regardless, these results suggest that the representation of AM rate may be asymmetric between different places-of-stimulation for individuals with hearing impairment. Further, results from the present study suggest that the ability to compare fluctuations in the temporal envelope will be limited by the place-of-stimulation with poorer representations of that stimulus.

Before this task or similar tasks are implemented in listeners with hearing loss in the future, investigators should consider several issues. Listeners with NH in this study exhibited considerable variability across individuals (see Fig. 3 for one example and Supplementary Fig. 1 and 2 for raw data). This suggests that an across-subjects experimental design using similar tasks to compare performance between groups would need a very large sample size to attain the necessary statistical power to detect effects. A more efficient approach may be to use a within-subjects experimental design and compare temporal envelope processing abilities across a variable within the same person as has been completed in several studies in individuals with hearing loss (e.g., [9,191,241,264]). In this study, investigators spent considerable time training listeners and discussing listeners' perception of changes in AM rate. This seemed to be an especially effective approach as it helped listeners understand the task and identify the perceptual changes they experienced as AM rate changed (e.g., rhythm, roughness, timbre, pitch). Experimenters were careful not to tell listeners what they "should" hear as AM rate changed. Finally, experimenters documented listeners' descriptions of changes in stimuli for specific ranges of AM rates. This was helpful if listeners became frustrated or began a new session on a different day.

B. Implications of asymmetric temporal representations

One example of how poorer signal encoding can interfere with speech in noise understanding is contralateral interference, where information from the poorer ear interferes with accessing information in the other ear [13,15,20,265]. This has been observed in NH listeners [20], but the implications for listeners with hearing impairment are not immediately clear. Within the bilateral CI population, it appears that longer durations of deafness might be related to poorer ability to ignore information in the worse ear [13,15]. Interference has also been demonstrated in subsets of patients with single-sided deafness [265].

The present experiments provide evidence to suggest that the ability to separate auditory objects will be negatively impacted by asymmetric temporal encoding. As this report demonstrates, the ability to make spectro-temporal comparisons across pairs of cochlear places-of-stimulation likely worsens when the amount of phase locking decreases for one place-of-stimulation. Poorer spectro-temporal comparisons have downstream implications for the model presented by [149], making auditory objects less salient and therefore less able to compete for attention. Listeners with hearing impairment exhibit extraordinary heterogeneity with respect to temporal encoding and performance on speech reception tasks in noise. More research is required to understand the implications of asymmetries in temporal encoding on patient outcomes to potentially improve patient care.

Previous studies have demonstrated that turning off electrodes where patients with CIs are insensitive to temporal cues can improve speech in noise understanding [191,264]. This paper provides one example of a sound source segregation mechanism that might be improved when electrodes for places-of-stimulation where the patient is not very sensitive to temporal fluctuations are turned off in a patient's programming.

C. Normal-hearing simulation of asymmetric temporal encoding

The experiments presented in this manuscript altered the representation of AM rate in the auditory nerve by reducing the depth of AM. The goal of this manipulation was to reduce the dynamic range of the stimulus and degree of phase locking for auditory nerve fibers. Changes in loudness as AM rate was varied could have been useful in performance. To address this potential confound, Fig. 7 shows estimated loudness due to changes in AM rate. The AM rates resulting in the largest differences in estimated loudness were not tested in Experiment 1 and additional ANOVAs are provided in the Discussion of Experiment 2 excluding cases with the largest differences. The results imply that listeners' sensitivity to differences in AM rate decreased when AM depth was reduced for both Experiment 1 and 2 even after accounting for changes in loudness.

It is difficult to determine the role of combination tones in the current experiments. While the existence of combination tones has been verified psychophysically and physiologically, there is not a widely-accepted model to account for the magnitude of each combination tone. In the experiments by Lee [246], the author suggests that that the difference tone (equal to the rate of AM) could contribute to AM rate discrimination in a paradigm similar to Experiment 1. If the difference tone is audible, then it would have presumably affected results in Experiment 2 as well. If participants in our experiments could use the difference tone to make judgments on differences in AM rate, then this would affect the higher AM rates used in this study (>90 Hz). However, it is also possible that the difference tone did not contribute much to perceptual results, or that additional combination tones could be used to make AM rate distinctions. Ultimately, the role of combination tones in AM rate discrimination requires further investigation.

D. Summary and conclusions

In real listening environments, sound sources span across frequency and are present in both ears. Segregating sound sources requires ongoing spectro-temporal comparisons within and across the ears [149,209,220]. Experiment 2 explored the simplest case; where listeners indicated whether pairs of places-of-stimulation were the same or different, representing the temporal envelope with good temporal fidelity in one ear and poorer temporal fidelity in the other ear. The results from this manuscript suggest that the accuracy of these spectro-temporal comparisons may be determined by the places-of-stimulation with poorest temporal encoding. Asymmetries in temporal encoding provide one mechanism to explain difficulties separating sound sources when listening in complex auditory environments for listeners with hearing impairment.

FOOTNOTES

1. See supplementary material at https://asa.scitation.org/doi/10.1121/1.5121423.

Chapter 3: Asymmetric Temporal Envelope Encoding: Within- and Across-Ear Envelope Comparisons in Listeners with Bilateral Cochlear Implants

I. INTRODUCTION

The perceptual separation of target sounds from background sounds (i.e., sound source segregation) underlies speech understanding in noise and is accomplished via multiple cues encoded by early stages of auditory processing [149,209]. Patients with bilateral cochlear implants (BiCIs) experience poorer speech understanding outcomes in noisy environments compared to listeners with normal-hearing (NH; e.g., [6,130]). This is likely related to the poorer sensitivity to many sound source segregation cues for listeners with BiCIs compared to listeners with NH (e.g., [46,62,105]). For listeners with CIs, sensitivity to sound source segregation cues is not only mediated by limitations inherent in CI hardware and software, but can also differ within the same individual depending upon the place along the auditory nerve being stimulated (e.g., [62,63,105,123,191,193,194,264]). Variability in sensitivity to segregation cues likely relates to the interface between CI electrodes and the auditory nerve and is further affected by auditory deprivation (resulting in deterioration of the auditory nerve), CI array type, and surgical factors (e.g., [125,130]). The combination of these factors affects the fidelity with which auditory information is preserved and accurately conveyed by the peripheral auditory system.

Many sound source segregation cues rely on a comparison of temporal information conveyed via different places-of-stimulation on the auditory nerve. This begs the question: does sensitivity to source segregation cues compared across different frequencies or ears correspond more closely to the place-of-stimulation on the auditory nerve with greater or lesser temporal fidelity? A recent report investigating sensitivity to binaural stimuli suggests that the ear with poorer temporal sensitivity predicts sensitivity to binaural cues [9], i.e., that binaural sensitivity is limited by the ability of the worse ear to process monaural temporal cues. The present study attempts to extend on these findings in listeners with BiCIs by exploring this relationship more generally, for different places-of-stimulation within the same ear or across ears. By using stimuli and cues that can be conveyed via present-day CI processing algorithms we aimed to characterize the relative sensitivity to temporal envelope differences across frequency (i.e., place-of-stimulation) in both ears.

A. Temporal envelope information in auditory scene analysis

Auditory signals contain two different types of temporal information. Faster fluctuations in intensity over time make up the temporal fine-structure. Slower changes in the average intensity over time make up the temporal envelope. Most CI processing algorithms discard temporal fine-structure and replace it with constant-rate, pulsatile stimulation, where amplitude modulation (AM) is applied according to the temporal envelope (e.g., [35]). Temporal envelope information is provided by a limited number of frequency bands that are conveyed by each electrode in the CI. Interference may occur between the temporal information presented via different electrodes, further limiting their spectro-temporal fidelity (for review, see [125]). Thus, from the beginning of auditory processing, the auditory system of listeners with CIs receives severely degraded auditory information. Temporal envelope information, compared to temporal fine-structure and frequency, is therefore well-preserved by CI processing and could be especially important to perception.

Auditory perception is characterized by serial processing of auditory information during which the temporal envelope may provide important segregation cues. Short spectro-temporal components form into auditory "streams," eventually banding together into auditory objects that compete for a listener's attention [149]. Listeners with CIs can use the temporal envelope to form auditory streams for sounds that occur at different points in time [224,225]. Temporal envelope cues are also used to distinguish between sounds that occur at the same time [209,220]. For example, when tones share similar temporal envelope fluctuations, they tend to be grouped together by the auditory system. If an off-frequency masker shares similar temporal envelope fluctuations with a target, then by extension, this masker is difficult to ignore. Listeners with CIs show evidence of this type of "interference" in temporal envelope detection and discrimination experiments [256,266–268]. If many off-frequency maskers share temporal envelope fluctuations that differ from the target, then the masker is easier to ignore. Listeners with CIs do not show evidence of this type of masking release [269,270]. Thus, previous experiments concerning the usefulness of temporal envelope cues in sound source segregation show mixed evidence in listeners with CIs and greater across-listener variability compared to those with normal hearing (NH).

The previous paradigms tested perceptual processes thought to be directly involved in perceptual grouping of sounds. An alternative approach is to explore sensitivity to differences in the temporal envelope between frequency or ears, then make predictions about the utility of these cues for perception. In one example, stimuli were presented simultaneously and listeners were asked to indicate whether these sounds have the same or different noise-based (i.e., stochastic) temporal envelope fluctuations. Listeners with NH show sensitivity to coherence of temporal envelopes in monaural [230] and binaural (e.g., [231]) hearing tasks. Listeners with BiCIs show poorer sensitivity to coherence of temporal envelopes binaurally compared to NH [140,229], but have not been tested monaurally.

The process of sound source segregation begins by grouping together stimuli with similar spectro-temporal structure. As demonstrated in the previous paragraphs, this process occurs within- and across-ears, across frequency and time. Interaural temporal envelope coherence is related to the amount of reverberation in an acoustic environment [271,272] and benefit in binaural masking level difference experiments (e.g., [143]), so it is sensible to use stimuli with stochastic envelopes in this context. However, stimuli with stochastic envelopes may have limited relevance to other sorts of listening conditions in the real world. In the present and a previously published experiment [273], we asked listeners to compare temporal patterns across spectral channels within- and across-ears. We employed deterministic (sinusoidal) envelopes to examine listeners' ability to judge differences between the shape of the envelope (i.e., rate of fluctuations) when predictably and parametrically manipulated.

B. Factors limiting access to the temporal envelope

Encoding of the temporal envelope with BiCIs may be limited by numerous listener-dependent factors before this information can be used to navigate auditory scenes. One important factor is deafness and prolonged auditory deprivation. Deprivation can lead to degraded auditory peripheral function, manifesting in the deterioration of dendritic processes on the auditory nerve and a loss of auditory nerve fibers evidenced from human temporal bones post-mortem (e.g., [108,109]) and animal models (e.g., [106,116,274]). A profound example that affects some listeners with CIs is extreme loss of nerve fibers in distinct regions of the auditory periphery, resulting in "spectral holes" [111]. The loss of auditory nerve fibers and degradation of their temporal fidelity compromises specializations in the auditory system to refine temporal precision. Over-representation of auditory nerve fibers plays a critical role in refining temporal precision at early stages of sound processing, such as the level of the brainstem (e.g., [54]), so the loss of fibers is deleterious to auditory processing. Together, studies of peripheral degradation after hearing loss suggest that fidelity of temporal encoding varies with the populations of neurons being excited.

Behavioral experiments show listener-dependent relationships between place-of-stimulation with CIs and monaural temporal sensitivity [9,62,63,193,194], as well as binaural temporal sensitivity [44,105]. These findings contrast to those from listeners with NH, who show better performance when stimulated closer to the cochlear apex at the same sound level for each place-of-stimulation (e.g., [61,275]). Neural health differs with place-of-stimulation [43,124,191,264] and experiments in listeners with CIs show that binaural sensitivity is related to monaural sensitivity to temporal information [9], suggesting that these factors are related.

There is also a similar “rate limitation” for monaural [9,62,63,193,194] and binaural stimulation [9,44–46,276], where pulse or AM rates near or above 300 Hz result in poorer sensitivity in listeners with CIs. To better understand the relationship between monaural rate sensitivity and binaural sensitivity, Ihlefeld and colleagues [9] examined monaural pulse rate discrimination (i.e., temporal pitch discrimination) and interaural time difference (ITD) discrimination in the same listeners with BiCIs. They tested three different places-of-stimulation in each ear and evaluated the relationship between monaural pulse rate and ITD discrimination at each place-of-stimulation. Crucially, their study showed that the ear with poorer (lower) pulse rate discrimination sensitivity was predictive of poorer (lower) ITD discrimination sensitivity. This suggests that when there is a place-of-stimulation yielding poor sensitivity to temporal information in one ear, there is a limitation placed on the ability to compare temporal information across the two ears.

A further important consideration is whether monaural temporal resolution in stimuli that are more akin to today's CI processors places a limitation on binaural sensitivity. Because that issue had not been investigated in listeners with NH, we recently tested the same relationship and manipulated the monaural resolution at one simulated cochlear place-of-stimulation. In that study, we reduced the fidelity of temporal cues by applying temporal fluctuations in a smaller dynamic range [273]. Instead of manipulating ITDs, the rate of temporal envelope fluctuations (AM rate) was varied in each ear. The results indicated that, indeed, the ear conveying poorer temporal information limited sensitivity to binaural cues.

The present study extended on the results of Ihlefeld and colleagues [9] by testing the same task and configurations as those used in [273]. We focused on cues conveyed via the temporal envelope, which are available in current CI processing algorithms and are therefore much more likely to be useful in everyday listening.

C. Study aims and hypotheses

In the present study we investigated sensitivity to differences in the rate of AM for pairs of stimuli played to two different electrodes in listeners with BiCIs. The goals of the present study were twofold: (1) determine the relative efficacy of temporal envelope comparisons for different configurations within- and across-ears, and (2) determine whether the place-of-stimulation yielding less temporal sensitivity was predictive of sensitivity to differences across places-of-stimulation (i.e., acted as a bottleneck). In accordance with the first goal, it was hypothesized that listeners with BiCIs would be most sensitive to differences in AM rate for pairs of stimuli presented to the same approximate place-of-stimulation in each ear because of the additional contribution of binaural cues. In accordance with the second goal, it was hypothesized that the place-of stimulation yielding least temporal sensitivity (Experiment 1) would be predictive of the sensitivity for pairs of electrodes (Experiment 2).

II. EXPERIMENT 1: SEQUENTIAL ENVELOPE DISCRIMINATION

A. Motivation

The goal of Experiment 1 was to assess the fidelity of AM rate coding by the auditory periphery for all four places-of-stimulation used in Experiment 2. We assumed that the best proxy of AM rate encoding would be AM rate discrimination thresholds. Thus, in Experiment 1, listeners discriminated between three stimuli played in a sequence, indicating the stimulus with highest AM rate. These results were used to predict the sensitivity to stimuli presented via pairs of electrodes simultaneously in Experiment 2. We hypothesized that there would be no consistent pattern in sensitivity at each place-of-stimulation (apex or base) across listeners in Experiment 1, and that listeners would be most sensitive to changes in AM rate near 100 Hz, consistent with previous AM rate discrimination experiments in listeners with CIs [193,194].

B. Methods

1. Listeners

Eleven listeners with bilateral CIs (age 51-80 years, with one listeners of 90+ years; mean 63 years) participated in this experiment. Demographics and CI processor information are listed in Table I. All listeners had at least 1 year of listening experience with bilateral CIs. All listeners provided informed consent and procedures were approved by the Health Sciences Institutional Review Board of the University of Wisconsin-Madison.

TABLE I. Listener demographics. All listeners used a 22 electrode array, where lower values indicate closer proximity to the base of the cochlea.

Listener ID	**Electrode (Apex/Base)**	**AM Depth (%)**	**Etiology**	**Age at Onset of Deafness (yrs)**	**Age at Implantation (L/R; yrs)**	**Age at Testing (yrs)**	**Internal Processor (L/R)**
IAJ	16 / 4	50	Progressive	5	51 / 58	71	CI24M / CI24R
IBF	16 / 4	50	Hereditary	38	56 / 54	65	CI24RE / CI24RE
IBK	16 / 4	50	Hereditary; noise exposure	53	63 / 69	80	CI24R / CI24RE
IBO	16 / 4	50	Otosclerosis ; sudden	20	45 / 42	52	CI24RE / CI24RE
IBZ	16 / 4	50	Unknown; sudden	38	40 / 38	51	CI24RE / CI24RE
ICD	16 / 4	50	Enlarged vestibular acqueduct	3	50 / 44	59	CI24R / CI24RE
ICI	16 / 4	50	Unknown	31	50 / 51	58	CI24RE / CI24RE
ICM	16 / 4	100	Unknown	23	57 / 58	64	CI512 / CI24RE
ICP	20 / 8	100	Unknown; progressive	3	46 / 49	54	CI24RE / CI24RE
ICS	16 / 4	100	Unknown	68	82 / 74	90+	CI24R / CI512
IDA	16 / 4	100	Progressive	8	47 / 46	51	CI24RE / CI24RE

2. Stimuli & procedures

Stimuli were sinusoidally amplitude-modulated (SAM) pulse trains. Pulse trains consisted of monopolar, biphasic pulses presented at a rate of 3000 pulses per second (pps) and a comfortable level. The pulse duration was 25 μs with an 8-μs interphase gap, except for listener ICP, who was presented with a pulse duration of 50 μs consistent with their clinical programming. High pulse rates were used in order to sample the envelope and increase stochastic firing of the auditory nerve. Studies on AM detection and pitch ranking in listeners with CIs have shown that four to five pulses per cycle are sufficient for psychophysical sensitivity and neural representation of AM [277–279]. Models of CI stimulation on auditory nerve fibers suggest that high rates of stimulation (≳3000 pps) result in more stochastic firing of model auditory nerve fibers [214]. Therefore, high pulse rates could delineate the electrode sites yielding poorest temporal resolution. Stimuli were presented to two different electrodes in each ear, one closer to the apex and another near the base of the cochlea. Stimuli were generated in MATLAB (Natick, MA) and presented via RF Generator XS research processor (Cochlea, Ltd., Sydney, Australia). Stimuli had 10-ms cosine onset- and offset-ramps. Testing was completed in a quiet, private room free from distractions.

Loudness balancing was completed to ensure that stimuli resulted in equal loudness on each electrode following procedures described in a previous methodological paper from our laboratory [280]. Briefly, the threshold (T), comfortable (C), and maximally comfortable (M) levels were determined using a 600-ms, constant-amplitude pulse train. Then, SAM pulse trains were presented at an AM rate between 100-160 Hz and an AM depth specific to that listener (in most cases 50%; see Table I for additional details) for all possible pairs of electrodes used in the present study. Stimulation level (in current units) was adjusted by the experimenter until the listener reported equal loudness for all possible pairs of electrodes. If listeners reported a change in loudness on later days of testing, these levels were adjusted. No loudness balancing was completed across the AM rates used in this study because previous experiments showed no change in loudness over AM rate [193,281].

Before testing began, listeners were familiarized with perceptual changes associated with increasing AM rate. Listeners were presented with AM rates of 10, 50, 90, 180, and 900 Hz and asked to describe their perception (e.g., compare sounds to something encountered in real life).

Listeners completed a three-interval, two-alternative forced-choice “oddball” discrimination task. On each trial, the first interval presented a standard AM rate (10, 30, or 90 Hz). The second or third interval (with equal probability) also presented the standard AM rate. The other interval consisted of an oddball (higher) AM rate, which was determined via adaptive tracking. An inter-stimulus interval of 300 ms separated each presentation and listeners were not allowed to repeat stimuli. Three two-down, one-up adaptive staircases corresponding to standard AM rates of 10, 30, and 90 Hz were interleaved such that on each trial the standard rate was randomly chosen from one of the remaining, unfinished staircases. A staircase was considered complete after 12 turnarounds. Step sizes for adaptive staircases were determined using a hybrid set of rules from the parameter estimation by sequential testing (PEST) and Levitt [282], as implemented in Litovsky [237] to minimize confusion [283]. The maximum step size for each standard AM rate was 16/3, 16, and 68 Hz, and the minimum was 1/3, 1, and 3 Hz such that the relative change was equal across standard AM rates (10, 30, and 90 Hz, respectively). Each adaptive staircase was initiated at five times the standard AM rate and this was the maximum oddball AM rate allowed during testing. The minimum AM rate tested was 31/30 times the standard AM rate. Testing was blocked by electrode (apical or basal in the left or right ear) and the order of testing was partially counterbalanced across listeners. Visual feedback was given after each trial. One practice block was given on the first electrode to be tested to ensure that listeners understood the task. Testing took approximately 2-3 hours for each listener.

3. Analysis

Thresholds for the oddball AM rate resulting in 71.1% correct discrimination were estimated by fitting a two-parameter sigmoidal curve bound between 50% and 100% using a weighted maximum likelihood procedure. Curve fitting was completed using version 2.5.6 of the psignfit toolbox in MATLAB [284]. Data were analyzed by converting discrimination threshold in Hz to $\Delta f_m/f_m$, where the numerator represents the rate presented at threshold (rate that could be discriminated) minus the standard AM rate and the denominator represents the standard AM rate, both in Hz. This was completed to express thresholds in terms of relative change from standard AM rate, allowing comparison across standard AM rates.

Data were analyzed using a mixed-effects analysis of variance (ANOVA) test. The dependent variable was threshold in $\log(\Delta f_m/f_m)$ units (i.e., proportional to the standard AM rate). A random intercept associated with listener was included to account for variability in means across individuals. Fixed-effects of standard AM rate treated categorically, electrode treated categorically, and their interaction were also estimated in the ANOVA in version 1.1-17 of the lme4 package of R [285]. Degrees of freedom were estimated using the Kenward-Roger approximation [286] to provide consistent estimates between ANOVAs and pairwise comparisons. All pairwise comparisons were completed using estimated marginal means with Tukey adjustment for multiple comparisons with version 1.3.0 of the emmeans package in R.

C. Results

The goal of Experiment 1 was to assess the sensitivity of listeners with BiCIs to stimuli delivered via all four electrodes employed in Experiment 2. Data from individuals are shown in Fig. 1 plotted as in Chatterjee and Oberzut [193] for ease of comparison. It was not possible to estimate a threshold from the data in some conditions because listeners achieved above or below 71.1% correct at all oddball AM rates ("CND" upper- and lower-lines, respectively, in panels of Fig. 1). These thresholds were therefore excluded from analysis. There was considerable variability across listeners, consistent with previous experiments examining AM rate discrimination in listeners with CIs [193,194]. Data are summarized across listeners in Fig. 2, where the mean and standard deviation of all measurable thresholds are reported.

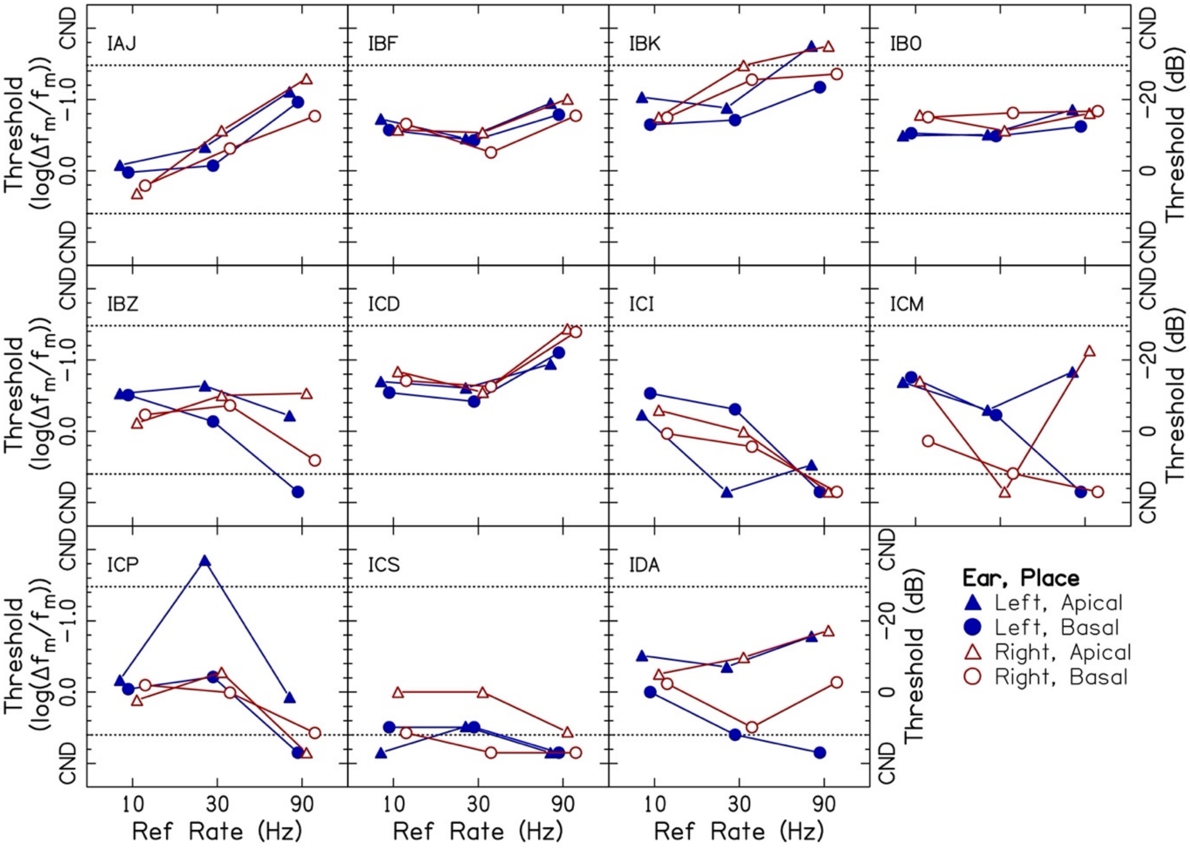


FIG. 1. Thresholds for sequential AM rate discrimination. Each panel corresponds to a different listeners whose subject code is given in the upper-left corner. The x-axis corresponds to the standard AM rate. The y-axis corresponds to the threshold, presented in reverse order such that

higher values indicate better performance as in [193]. Triangles and circles represent apical and basal places-of-stimulation, respectively. Open and closed shapes represent the left and right ear, respectively. “CND” indicates that a threshold could not be determined because performance fell above or below 71.1% correct.

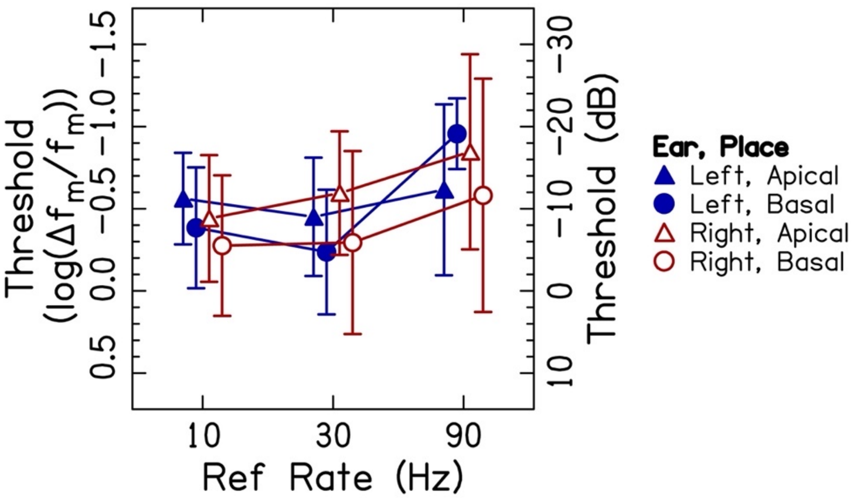


FIG. 2. Thresholds for sequential AM rate discrimination averaged across listeners in Fig. 1. The x-axis corresponds to the standard AM rate. The y-axis corresponds to the threshold, presented in reverse order such that higher values indicate better performance as in [193]. Triangles and circles represent apical and basal places-of-stimulation, respectively. Open and closed shapes represent the left and right ear, respectively.

Results of the mixed-effects ANOVA indicated that there were significant main effects of standard AM rate [$F(2,95) = 8.521$, $p < 0.001$] and place-of-stimulation [$F(1,95) = 8.004$, $p < 0.01$], but no significant interaction [$F(2,95) = 0.452$, $p = 0.637$]. Discrimination thresholds were best (lowest) for the 90-Hz standard AM rate compared to 30 Hz [t(95) = -3.875, p < 0.001] and 10 Hz [t(95) = -3.418, p < 0.01]. There was no significant difference between discrimination thresholds for the 30-Hz and 10-Hz standard AM rates [t(95) = 0.545, p = 0.849]. Stimulation of apical electrodes resulted in better (lower) AM rate discrimination thresholds compared to basal electrodes [t(95) = -2.829, p < 0.01].

D. Discussion

Experiment 1 investigated AM rate discrimination of SAM pulse trains presented sequentially via a single electrode. Previous experiments have examined sensitivity in AM rate discrimination tasks for listeners with CIs [193,194]. The present experiment was the first to assess AM rate discrimination in both ears of listeners with BiCIs to the best of the authors' knowledge. Thus, the present experiment had greater statistical power because two measurements were collected for a similar place-of-stimulation in each ear of the same participants.

The AM rate discrimination thresholds obtained in the present study were generally higher than those from previously published studies in listeners with CIs [193,194,241,281,287]. In a study of listeners with NH using SAM tones and the same behavioral task, we also showed that discrimination thresholds were higher than typically observed in that listener population [273]. Thus, it seems likely that something about the task made this experiment more challenging, resulting in worse performance. We suspect that the interleaving of standard AM rates within the same trial block may have played a role. This is consistent with findings of Lee [246] in listeners with NH showing that discrimination thresholds increased when the carrier frequency (and corresponding pitch) was chosen randomly from trial-to-trial. Though the AM depths employed in the present experiment (see Table I) were smaller than the full dynamic range for many listeners, other experiments have shown lower (better) discrimination thresholds when even smaller AM depths (around 20%) were used [193,194]. The present study used very high pulse rates (3000 pps), which is one major difference from previous studies. However, so long as the envelope is sufficiently sampled (four to five pulses per cycle; e.g., [279]), pulse rate has shown no consistent effects across listeners in previous studies [287], suggesting that very high pulse rates should not affect the comparability to other studies of AM rate discrimination in listeners with CIs. It should be noted that the experiment by Green and colleagues [287] had a small sample size and failed to show a difference in thresholds between 10- and 100-Hz standard AM rates, which has been demonstrated in most studies of AM rate discrimination in listeners with CIs using SAM [193,194,281].

Results from the present study showed that AM rate discrimination thresholds were best (lowest) at 90-Hz standard AM rate. This is consistent with previous results comparing lower standard AM rates to a 100-Hz standard AM rate using SAM [193,194]. There are several possible reasons for this result. The sharper slope of the temporal envelope near 100 Hz may lead to better encoding of temporal information (for detailed description, see Sec. IV.A). However, this conclusion is not supported by studies in listeners with CIs examining the differences in AM rate discrimination thresholds for stimuli with sharper or shallower temporal envelope slopes, showing no significant improvements of AM rate discrimination with sharper envelopes relative to 100 Hz or a higher AM rate [241,281]. Notably, binaural sensitivity for AM stimuli in listeners with BiCIs is also best around 100 Hz [46], suggesting that a similar mechanism may be involved.

The present experiment showed that apical stimulation resulted in better (lower) average discrimination thresholds to AM rates than basal stimulation. These results are somewhat consistent with Chatterjee and Peng [194], who found that a subset of listeners had better (lower) AM rate discrimination thresholds at apical electrodes. Their laboratory did not replicate this result when a similar paradigm was employed with a larger set of listeners [193]. Listeners with NH show no consistent effect of carrier frequency (and correspondingly, place-of-stimulation) on AM rate discrimination [246,273]. It should be noted that carrier frequency was only explored across listeners in Anderson and colleagues [273], not varied within the same individuals, so an effect could have easily been masked by across-listener variability. In contrast, modulation detection thresholds in listeners with NH are better (lower) for lower carrier frequencies [234], consistent with the findings in the present experiment. The role of differences between listeners with BiCIs and NH are discussed in more detail in Sec. IV.C.

III. EXPERIMENT 2: SIMULTANEOUS ENVELOPE DISCRIMINATION ACROSS PLACES-OF-STIMULATION

A. Methods

1. Listeners

All listeners that participated in Experiment 1 also participated in Experiment 2. However, not all listeners were able to complete the task. For more details, see Section III.B.

2. Stimuli & procedures

The stimuli and psychophysical procedures employed in the present study were similar to those used in Anderson and colleagues [273] except that AM depth was not varied systematically and carriers were electrical pulse trains. Testing procedures are explained succinctly here, but for greater details, please see [273].

Stimuli were presented using the same equipment as Experiment 1. The greatest difference between Experiment 1 and 2 was that in Experiment 2 two SAM pulse trains were presented together via different electrodes in the same interval, either within the same ear or in different ears. Each SAM pulse train was of 600 ms duration and random envelope phase, but identical fine-structure in both electrodes. All stimuli were presented at a comfortable level.

Experiment 2 consisted of a one-interval, two-alternative forced-choice task where listeners indicated whether the AM rates presented via both electrodes were the same or different. One electrode in an experimental block always presented the standard AM rate (10 or 90 Hz). The other electrode presented either the standard AM rate or a higher, variable AM rate with equal probability of 0.5. Variable AM rates of $\log(\Delta f_m/f_m)$, = -0.71, -0.42, -0.12, 0.17, 0.47, 0.76, and 1.06 were tested an equal number of times in each block to estimate a psychometric function. For most listeners, 30 repetitions per variable AM rate were tested. Listeners were not allowed to repeat stimuli and visual feedback was provided after each trial. Visual feedback was provided to give listeners maximal opportunities to learn the changes in perception associated with varying AM rate. The possible impact of visual feedback is discussed in more detail in Section III.C.1.a.

There were three possible ways to pair electrodes by varying the place-of-stimulation (apex or base) or ear (left or right). These pairing configurations were: (1) Same Place, Across Ears, (2) Different Place, Across Ears, and (3) Different Place, Within Ears. Broadly speaking, each configuration represents the different kinds of comparisons the auditory system can complete using the temporal envelope. For a graphical representation of these conditions, see Fig. 2 of [273]. No attempts were made to match place-of-stimulation between ears beyond using the electrode number. Thus, there may have been mismatches in the place-of-stimulation between ears (for review, see [45]), but it was assumed that the wide spread of current resulting from monopolar stimulation would lead to excitation of overlapping neural populations in each ear. Trials were blocked by AM pairing configurations according to a Latin square design across listeners to partially account for learning effects. For each listener, there were two possible electrode pairs in each AM pairing configuration. The order of the two pairs tested was chosen randomly. The electrode presenting the variable AM rate was randomly chosen for the first half of trials, and the other electrode in the pair presented the variable AM rate in the second half of trials. An example testing order is provided in Table II.

TABLE II. Example experimental conditions. Each row corresponds to a block of trials. On each trial, the variable AM rate place-of-stimulation would present the standard AM rate with a probability of 0.5 or a variable rate. In the third column referring to place-of-stimulation being stimulated, possible values include Apical (stimulation near the cochlear apex), Basal (stimulation near the cochlear base), Both (both apical and basal stimulation), or N/A (meaning not applicable because no stimulation was provided to that ear).

			Variable AM Rate		**Std. AM Rate**	
Block	**Place-of-Stimulation Pairing**	**Place-of-Stimulation (Left-Right)**	**Ear**	**Electrode**	**Ear**	**Electrode**
1	Same Place, Across Ears	Base-Base	L	4	R	4
2	Same Place, Across Ears	Apex-Apex	L	16	R	16
3	Same Place, Across Ears	Base-Base	R	4	L	4
4	Same Place, Across Ears	Apex-Apex	R	16	L	16
5	Different Place, Across Ears	Base-Apex	L	4	R	16
6	Different Place, Across Ears	Apex-Base	L	16	R	4
7	Different Place, Across Ears	Base-Apex	R	4	L	16
8	Different Place, Across Ears	Apex-Base	R	16	L	4
9	Different Place, Within Ears	Both-N/A	L	4	L	16
10	Different Place, Within Ears	Both-N/A	L	16	L	4
11	Different Place, Within Ears	N/A-Both	R	4	R	16
12	Different Place, Within Ears	N/A-Both	R	16	R	4

Listeners were familiarized with the task each time that a new AM pairing configuration was introduced. Familiarization took place with the pair of electrodes that was tested first. Listeners were presented with a 10-Hz AM rate in both electrodes. They were then asked to describe the sound to the best of their ability and compare it to some real-life examples. The experimenter documented these descriptions in case the listeners became confused during later testing. Next, the listeners were presented with 10 Hz in the reference (standard AM rate) electrode and 50 or 100 Hz in the variable electrode. Without exception, listeners described 10-50 and 10-100 Hz pairings as similar to one another (sometimes the latter, higher rate being higher in pitch). The process then was repeated, but with 90-Hz AM rate in the reference electrode and 450- or 900-Hz AM rate in the variable electrode. Finally, listeners completed a practice block of the experiment with the AM rates used in the familiarization procedure. The time spent during familiarization varied considerably across listeners and this sometimes took up to 4 hours to complete. During testing, if there was evidence that the listener became confused (e.g., a change in the shape of their psychometric function), they were given a practice block containing only the two highest variable AM rates and data were collected *de novo*. During analysis, psychometric functions from trial blocks that were not representative of other data from the same listeners were removed from analysis.

One of the challenges in designing this experiment was to limit the usefulness of attending to only the higher AM rate to make decisions on each trial. For example, the listener could have simply attended to the electrode with the variable AM rate and responded “different” if it increased above the standard AM rate, effectively shifting this task to something like that in Experiment 1. Several methodological choices were made to minimize the utility of such a listening strategy. Firstly, two psychometric functions corresponding to standard AM rates of 10 and 90 Hz were interleaved. Secondly, an overlapping range of variable AM rates was tested (the highest variable AM rates re: 10 Hz and lowest variable AM rates re: 90 Hz). Thirdly, a small amount of AM rate roving (±0.28 in $\log(\Delta f_m/f_m)$ units) was implemented on each trial.

3. Analysis

Psychometric curves were estimated using a four-parameter sigmoidal function in the Curve Fitting Toolbox of MATLAB and were only used for plotting purposes. Data were analyzed using a mixed-effects ANOVA including random-effect of listener and fixed-effects of variable AM rate ($\log(\Delta f_m/f_m)$) treated categorically, standard AM rate treated categorically, AM pairing configuration treated categorically, and their interactions. The dependent variable was sensitivity (in d' units transformed according to Eq. 1). Sensitivity in d' units was calculated using the number of hits, misses, correct rejections, and false alarms to estimate the difference between the mean distributions of the "signal" (variable AM rate in one electrode) and "no signal" (standard AM rate in both electrodes) trials [81]. There were an equal number of "signal" and "no signal" trials overall, meaning that a greater number of "no signal" trials (30 repetitions × 7 variable AM rates = 210 "no signal" presentations per standard AM rate) were used to compute d' at each value of $\log(\Delta f_m/f_m)$.

In a second set of analyses relating the results of Experiment 1 and 2, another mixed-effects ANOVA included fixed-effects of worse (higher) threshold in Experiment 1 (for one electrode in a pair, defined in more detail in Sec. III.B.3) treated continuously, variable AM rate ($\log(\Delta f_m/f_m)$) treated categorically, and standard AM rate treated categorically. These fixed-effects were used to predict sensitivity (in d' units transformed according to Eq. 1). To prevent using the same thresholds from Experiment 1 to predict different conditions in Experiment 2, separate ANOVAs were completed for each AM pairing configuration. For example, if the apical electrode in the left ear yielded the worst (highest) threshold of all electrodes in Experiment 1, then it would be used to predict the sensitivity of the apical pair of electrodes in the Same Place, Across Ears configuration and the left ear in the Different Place, Within Ears condition.

Degrees of freedom were estimated using the Kenward-Roger approximation [286]. All pairwise comparisons were completed using estimated marginal means with Tukey adjustment for multiple comparisons. The same statistical packages were used as described in Section II.A.3. For consistency with previous reports [9], an additional analysis was completed using the within-subjects correlation procedure described by Bland and Altman (fixed-effects ANOVA including intercepts for each listener) [288].

When models were fit with sensitivity in d’ units in the previously described methods, those models systematically underestimated the highest and lowest values of sensitivity, resulting in curvilinear, non-normally distributed model residuals. Thus, sensitivity in d’ units was transformed according to the top-half of a sigmoidal function as described in Eq. 1:

$$d'^{*}_{ij} = \frac{1}{1+e^{-(d'_{ij}-d'_{\min})}} \qquad (1)$$

where d'^{*}_{ij} is the transformed dependent variable for the *j*th observation from the *i*th listener, d'_{ij} is *j*th observation of sensitivity for the *i*th listener, and d'_{min} is the minimum value of sensitivity observed in the experiment across all listeners. Effectively, this equation compressed the highest values of sensitivity, which corrected some systematic overestimation of the dependent variable, but there was still some evidence of underestimation of residuals at the lowest sensitivities. This transform was chosen because it resulted in the best improvement of model diagnostics. Other transforms tested were: log, Box-Cox, several sigmoidal functions, and inverse d’. Only one of the inferences made as a result of these transformations changed from those that would have been made with the original data and is noted in Sec.III.B.2.

B. Results

The goal of the present experiment was to assess sensitivity to differences in AM rate across pairs of electrodes in various configurations for listeners with BiCIs. Individual data for listeners with BiCIs are shown in Fig. 3. The smaller panel on the left within each column shows the proportion of “different” responses when the same AM rate (standard AM rate ± rove) was presented to both electrodes. Therefore, the smaller panel on the left summarizes the bias toward responding “different.” Figure 3 indicates that there was substantial variability across listeners, both with respect to overall performance and effects of independent variables. For example, listener IBO showed relatively little bias toward responding “different” (on approximately 20% of trials), with well-behaved psychometric functions and steep slopes. In contrast, listener IDA showed strong bias toward responding “different” (on 40-50% of trials). Listener IBF shows consistent bias toward responding “different”, but only for the 10-Hz standard AM rate. Some non-monotonic trends or even reversed psychometric functions for individuals are shown in Fig. 3 (e.g., results from listeners ICD and IDA). Most listeners reported the task as very difficult, which may have contributed to variability in performance across individuals.

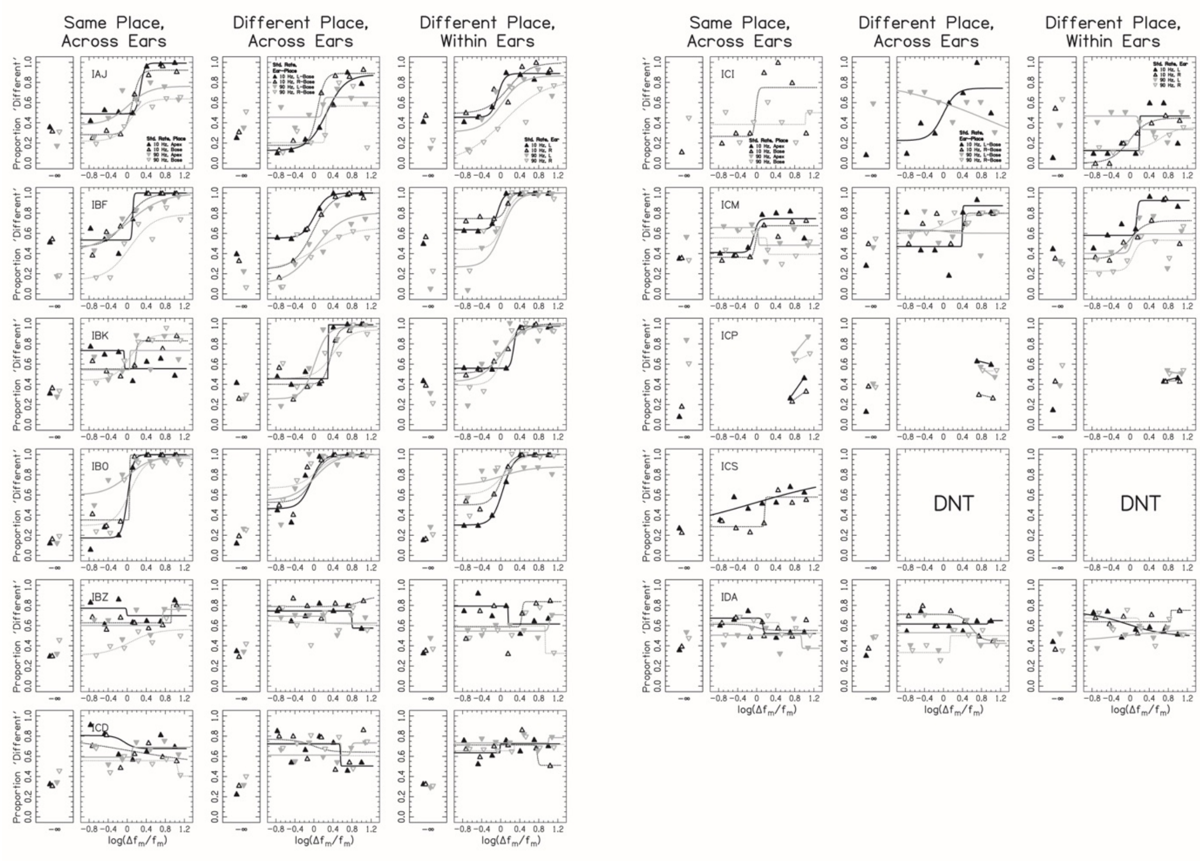


FIG. 3. Psychometric functions for AM rates presented simultaneously to pairs of electrodes. Each panel shows individual performance for the Same Place, Across Ears (first and fourth column), Different Place, Across Ears (second and fifth column), and Same Place, Across Ears (third and sixth columns) AM pairing configurations. Subject codes for each individual are given in the top left corner of the first and third column. Performance for 10 Hz standard rate is shown in black and 90 Hz standard rate are shown in grey. The y-axis corresponds to the proportion of “different” responses across all trials. The x-axis corresponds to the difference in AM rate between the standard and variable AM rate. The small panel on the left (a log($\Delta f_m/f_m$) of -∞) represents the proportion of “different” responses when AM rates were equal. Ideal performance occurs when the proportion of “different” responses is 0 for the small left panel, and 1 for the larger, right panel. Sensitivity (d’) can be calculated directly from the raw data. The electrode

pair indicated by the shape and line type is given in the figure legend in the top row and varies depending upon the AM pairing configuration.

Results are summarized across listeners in Fig. 4(A), which shows the average sensitivity (in d' units; i.e., z-transformed curves in Fig. 3 scaled by the amount of bias toward making a "different" response). Figure 4(B) shows a subset of the data from listeners with NH in a previously published experiment using the same paradigm with 50% AM depth for both places-of-stimulation [273]. Though there was also considerable variability in listeners with NH completing the same task with SAM tones, listeners with BiCIs showed worse performance (less sensitivity) and more poorly behaved psychometric functions. Despite these challenges for listeners with BiCIs, there was a significant effect of variable AM rate ($\log(\Delta f_m/f_m)$) on sensitivity [$F(6,623) = 38.626$, $p < 0.0001$], confirming that increasing differences in AM rate between each electrode in a pair resulted in greater sensitivity. The three largest variable AM rates ($\log(\Delta f_m/f_m)$) resulted in significantly greater sensitivity [$p < 0.001$ or 0.0001] compared to the smallest variable AM rate[2].

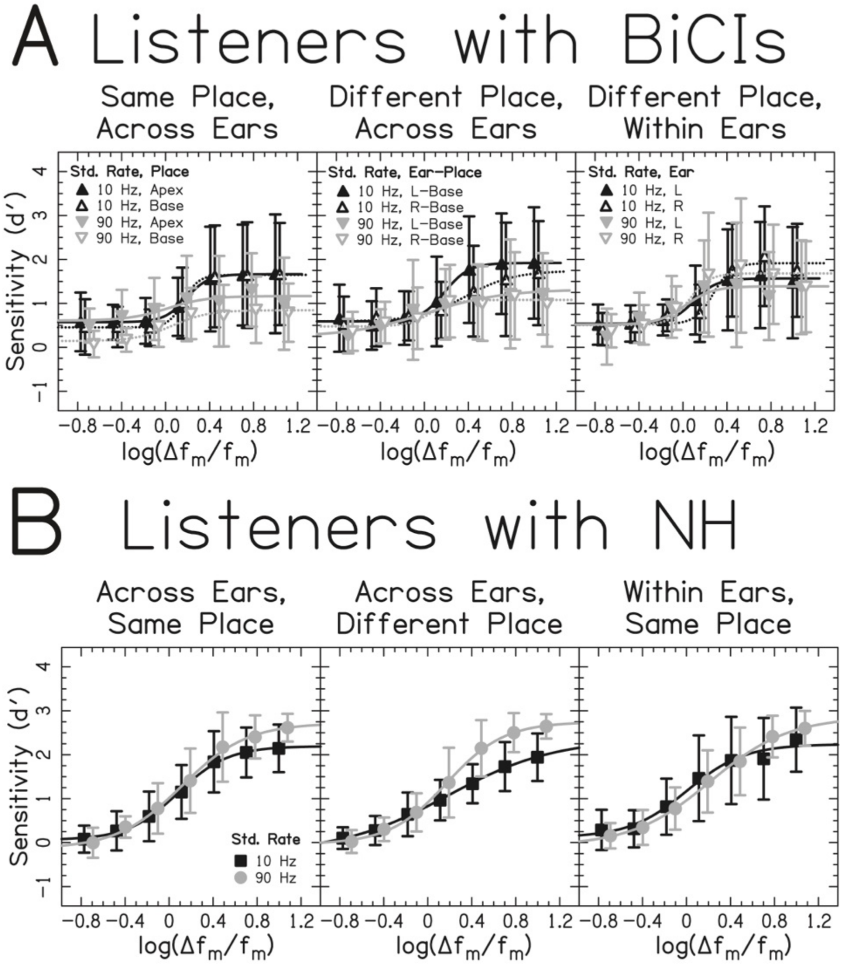


FIG. 4. Mean ± one standard deviation sensitivity across listeners for all three AM pairing configurations for listeners with (A) BiCIs and (B) NH. Performance for 10 Hz standard rate is shown in black and 90 Hz standard rate are shown in grey. Each panel corresponds to a different AM pairing configuration. The x-axis corresponds to the difference in AM rate between the standard and variable AM rate (see Table 1 for values of AM rates in Hz). The y-axis corresponds to sensitivity in d’ units. The electrode pair indicated by the shape and line type is given in the figure legend and varies depending upon the AM pairing configuration for listeners with BiCIs.

1. Effects of standard AM rate

Typically, standard AM rates near 100 Hz for SAM stimuli result in greater AM rate discrimination and binaural sensitivity for listeners with BiCIs [46,193,194]. Changes relative to the standard AM rate were normalized by measuring variable AM rates in $\log(\Delta f_m/f_m)$ units, allowing for fairer comparisons across standard AM rates. The higher standard AM rate tested in the present study (90 Hz) resulted in significantly *poorer* sensitivity [$F(1,623) = 14.797$, $p < 0.001$], in contrast to Experiment 1 and previous results using the same task with listeners with NH [273]. There was also standard AM rate × variable AM rate ($\log(\Delta f_m/f_m)$) interaction [$F(6,623) = 3.724$, $p < 0.01$], suggesting that the slope of the psychometric function changed with standard AM rate, consistent with results from listeners with NH [273].

2. Effects of AM pairing configuration

Three different AM pairing configurations were tested in the present study representing the possible ways in which the temporal envelope can be compared for listeners with bilateral hearing. The Same Place, Across Ears configuration was expected to result in greatest sensitivity because of the addition of a binaural beat cue (e.g., [253]). There was a significant effect of AM pairing configuration [$F(2,623) = 4.218$, $p < 0.05$]. In contrast to the hypothesis, pairwise comparisons indicated that sensitivity was significantly greater for the Different Place, Within Ears compared to the Same Place, Across Ears configuration [$t(623) = 2.88$, $p < 0.05$]. There was no significant difference between the Different Place, Within Ears and Different Place, Across Ears configuration [$t(623) = 1.713$, $p = 0.201$] or the Different Place, Across Ears and Same Place, Across Ears configurations [$t(623) = 1.175$, $p = 0.469$].

There was also no significant AM pairing configuration × standard AM rate interaction [F(2,263) = 2.335, p = 0.098]. However, it should be noted that this interaction was significant if the original (not transformed) values of sensitivity were included as the dependent variable.

There was no significant AM pairing configuration × variable AM rate ($\log(\Delta f_m/f_m)$) interaction [F(12,623) = 2.335, p = 0.988] or three-way AM pairing configuration × standard AM rate × variable AM rate ($\log(\Delta f_m/f_m)$) interaction [F(12,623) = 0.231, p = 0.997].

3. Role of the poorer place-of-stimulation

We hypothesized that in pairs of electrodes, the electrode yielding poorest temporal sensitivity (highest threshold) in Experiment 1 would predict the sensitivity of pairs of electrodes in Experiment 2. Two different sources of variability were suspected to play a role in the thresholds from Experiment 1: (1) variability due to poorer transmission of temporal information through the auditory system, and (2) variability over testing time unrelated to transmission of temporal information (e.g., attention, fatigue). Assuming that Experiment 2 would also be affected by poorer transmission of temporal information, we chose to include the higher thresholds from Experiment 1 of the pair tested at the same standard AM rate in Experiment 2 as a continuous fixed-effect in the model.

Results are summarized in Fig. 5. The x-axis shows the poorer (higher) AM rate discrimination threshold from Experiment 1 for the pair of electrodes tested in Experiment 2. The y-axis shows the average sensitivity across variable AM rates ($\log(\Delta f_m/f_m)$). Consistent with the hypothesis, there is a negative relationship these two variables, suggesting that as performance became worse (threshold increased) in Experiment 1, sensitivity decreased in Experiment 2. Figure 5 does not describe the substantial amount of variability in sensitivity observed within individuals due to changes in variable AM rate ($\log(\Delta f_m/f_m)$) that can be seen more clearly in Fig. 3 and 4.

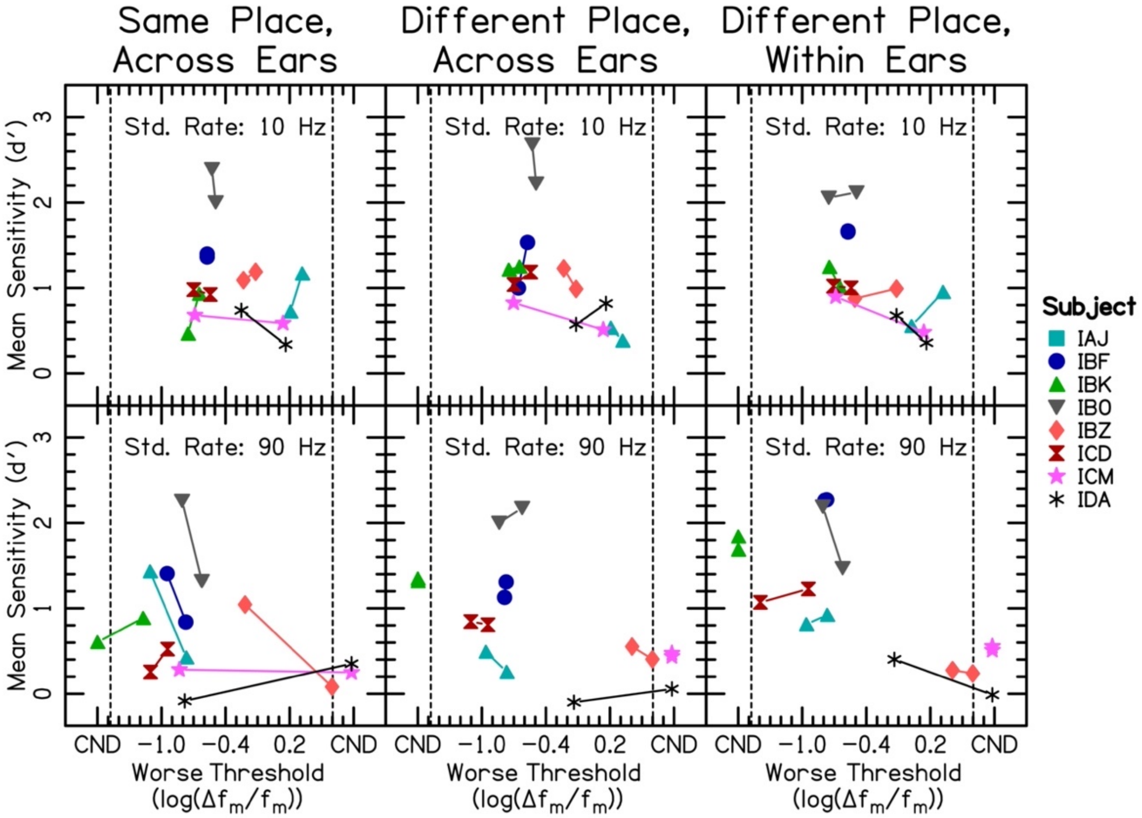


FIG. 5. Relationship between Experiment 1 and 2. Top and bottom rows correspond to standard AM rates of 10 and 90 Hz, respectively. Each column corresponds to the three AM pairing configurations. The x-axis corresponds to the threshold of the electrode yielding the poorer (higher) threshold in Experiment 1 or the pair tested in Experiment 2. The y-axis corresponds to the sensitivity of pairs of electrodes. The shape and color given in the figure legend on the right indicates the listener.

For convenience, results from the three ANOVAs for each AM pairing configuration are reported in Table III. Thresholds from Experiment 1 that could not be estimated ("CND" in Fig. 1) were excluded from analysis. Variable AM rate ($\log(\Delta f_m/f_m)$) significantly affected sensitivity in all three AM pairing configurations [$p < 0.0001$], consistent with the above sections. Threshold from Experiment 1 of the worse electrode in the pair significantly predicted sensitivity in Experiment 2 for the Same Place, Across Ears [$p < 0.01$], Different Place, Across Ears [$p < 0.01$], and Different Place, Within Ears [$p < 0.0001$] configurations. There were significant effects of standard AM rate for the Same Place, Across Ears [$p < 0.0001$] and Different Place, Across Ears [$p < 0.001$] configurations. In both cases, there was lower sensitivity associated with higher standard AM rates. This is mostly in agreement with the ANOVAs in Section III.C.2, which showed better sensitivity for the 10-Hz standard AM rate across all AM pairing configurations.

TABLE III. Results of mixed-effects ANOVAs relating the electrode yielding the worse threshold in Experiment 1 to sensitivity of the pair in Experiment 2. Dependent variable was sensitivity in d’ units transformed according to Eq. 1.

	Same Place, Across Ears		**Different Place, Across Ears**		**Different Place, Within Ears**	
Effect	**F-Statistic**	**p**	**F-Statistic**	**p**	**F-Statistic**	**p**
$\log(\Delta f_m/f_m)$	F(6,180) = 11.936	< 0.0001	F(6,166) = 9.702	< 0.0001	F(6,166) = 15.890	< 0.0001
Std. Rate	F(1,187) = 17.916	< 0.0001	F(1,173) = 13.890	< 0.001	F(1,173) = 2.705	0.102
Electrode Pair	F(1,185) = 2.547	0.112	F(1,166) = 1.216	0.272	F(1,167) = 21.520	0.630
Exp. 1 Threshold	F(1,170) = 7.556	< 0.01	F(1,157) = 8.002	< 0.01	F(1,143) = 21.520	< 0.0001
$\log(\Delta f_m/f_m)$ × Std. Rate	F(6,180) = 1.200	0.308	F(6,166) = 1.090	0.370	F(6,166) = 1.141	0.341

The ANOVA in Section III.C.2 yielded a significant standard AM rate × variable AM rate ($\log(\Delta f_m/f_m)$) interaction. This was not replicated in any of the three ANOVAs when thresholds from Experiment 1 were included. Including Experiment 1 thresholds in the model may have accounted for some of the variability associated with the standard AM rate × variable AM rate ($\log(\Delta f_m/f_m)$) interaction.

An alternative analysis approach provided by Ihlefeld and colleagues [9] is to fit a similar, fixed-effects model including listener as a fixed-effect and excluding interactions, finally computing within-subject correlations for Experiment 1 and 2. One benefit of this analysis method is that variability associated with independent variables (variable AM rate, standard AM rate, changes in mean associated with listener) is accounted for before the relationship with Experiment 1 thresholds is estimated. For consistency with their report, within-subject correlations are shown in Table IV based on the worst, mean, best, and difference between Experiment 1 thresholds in the pair. When using the worst threshold from Experiment 1 as the predictor, these correlations revealed significant within-subject associations for all AM pairing configurations, with Different Place, Within Ears yielding the strongest association. Since Ihlefeld and colleagues [9] evaluated sensitivity to binaural cues, the most similar AM pairing configuration is Same Place, Across Ears. Our within-subjects correlation was weaker than that of Ihlefeld et al.'s (-0.202 vs 0.33, respectively), and other configurations resulted in similar correlation coefficients (-0.194 and -0.294). Their study tested three pairs of electrodes rather than two, and tested four rates rather than two, which would have provided greater statistical power. Their study also did not need to account for factors varied in the binaural task (i.e., variable AM rate), which could add to residual variance.

TABLE IV. Within-listener correlations between Experiment 1 thresholds and results in Experiment 2. These correlations were calculated using the same procedure discussed in Bland and Altman [288], consistent with Ihlefeld and colleagues [9].

	Across Ears, Same Place		**Across Ears, Different Place**		**Within Ears, Different Place**	
Experiment 1 Threshold(s)	**Pearson r**	**p**	**Pearson r**	**p**	**Pearson r**	**p**
Max Threshold	-0.202	< 0.01	-0.194	< 0.05	-0.294	< 0.0001
Mean of Thresholds	-0.089	0.231	-0.018	0.817	0.028	0.721
Min Threshold	-0.092	0.220	0.030	0.698	0.025	0.743
Difference Between Thresholds	-0.187	< 0.05	-0.214	< 0.01	-0.288	< 0.001

C. Discussion

Results from the present experiment demonstrated that listeners with BiCIs were sensitive to differences in AM rate between pairs of electrodes for SAM pulse trains presented via: (1) Same Place, Across Ears, (2) Different Place, Across Ears, and (3) Different Place, Within Ears configurations. This is the first experiment to examine sensitivity to differences in the temporal envelope across places-of-stimulation in both ears for the same listeners with BiCIs. The results indicated surprisingly poor performance for listeners with BiCIs given that the temporal envelope is one of very few stimulus features that is well-preserved by CI processing, in agreement with several previous experiments. Experiments investigating comodulation masking release (i.e., benefit from similar temporal fluctuations in masker stimuli) showed no benefit for listeners with unilateral CIs, attributing the lack of benefit to poor overall performance, reduced spectral resolution, and confounding interference due to the presence of temporal modulations [269,270]. However, an additional factor may be reduced sensitivity to differences in the temporal envelope across spectral channels. Listeners with BiCIs showed less sensitivity to differences (i.e., coherence) in the temporal envelope presented to similar places-of-stimulation in each ear than listeners with NH in previous experiments [140,229]. While performance co-varies with acoustic hearing history, this was not sufficient to explain differences between groups [140]. Findings from the present experiment provide additional evidence to suggest that a lack of comodulation masking release could be due to a poorer ability to compare temporal envelope information between spectral channels (within- and across-ears) for listeners with CIs compared to NH.

Sensitivity to differences in AM rate for pairs of electrodes was further limited by the electrode yielding poorest sensitivity to temporal information. The evidence of this relationship was strongest in the Different Place, Within Ears configuration. The previous study by Ihelfeld and colleagues [9] showed the novel result that the sensitivity to temporal information in the worse ear predicted binaural sensitivity. Binaural comparisons in the auditory pathways are completed by an extremely specialized system that relies on precision of temporal inputs (e.g., [47,54]). For this reason, we expected that binaural sensitivity would be best predicted by the ear with poorer sensitivity to temporal information compared to other AM pairing configurations. Instead, these results imply that comparisons of temporal information across different frequency regions in the same ear are more highly impacted in listeners with BiCIs. If the binaural system is already considerably taxed, then it seems reasonable that listeners would show the smallest amount of benefit by improved temporal resolution relative to comparisons between pairs of electrodes within the same ear. Moreover, at the variable AM rates tested in the present experiment, the use "binaural beats" (cues that could have been useful in the Same Place, Across Ears condition) would only occur for the smallest values of variable AM rate, where listeners showed the poorest sensitivity. Thus, the stimuli in this experiment may be more relevant to the kinds of comparisons that are completed across frequency (for more information on this idea, see Sec. III.C.1.b).

1. Comparison with NH results

a. Overall performance

Compared to listeners with NH completing the same task with very similar stimuli [273], listeners with BiCIs showed considerably less sensitivity and more unreliable trends across variable AM rates (Fig. 3; Fig. 4). In several cases, listeners with BiCIs showed no sensitivity to differences in AM rate at all, and some showed reversed psychometric functions. More listeners with BiCIs (n=3 of 11; listeners ICI, ICP, and ICS) could not complete the task than listeners with NH (n=1 of 12; testing terminated after familiarization). Listeners with BiCIs tended to show bias (a greater number of "different" responses) for standard AM rates of 10 Hz compared to 90 Hz (Fig. 3 in this paper vs Supplementary Fig. 1 in Anderson et al. [46]). This bias was most pronounced for listener IBF and was not entirely consistent across listeners. For example, listener ICI showed greater bias toward the standard AM rate of 90 Hz (though their data were not included in the analyses). Many different factors may have played a role in these differences between groups. Sources of differences between listener groups are explored in greater detail in Sec. IV.C.

One factor that likely contributed to performance in both groups was the use of visual feedback. It was thought that providing visual feedback would help listeners anchor performance based upon trials where they could clearly hear two different AM rates being presented. However, if the listeners were completing the task correctly and were completely unbiased (i.e., reported "different" on zero trials where the same AM rate was presented to both electrodes), they would have responded "same" more often than "different" and would have been given feedback that, in fact, the correct answer was "different". These trials would have occurred when the variable AM rate was below threshold. Instead of anchoring performance in accordance with the goals of the experiment, visual feedback may have led listeners to question their responses and try new listening strategies. Ultimately, most listeners were still able to complete the task, but feedback may have played a role in situations where responses were counter-intuitive (e.g., large amounts of bias or reversed psychometric functions). Based on these results, it seems that the best course of action may have been to provide feedback during familiarization and remove it during the experiment. Listener-reported difficulty also played a role in the decision to include feedback during testing. Listeners with BiCIs and NH reported extreme difficulty with the task, so it is unclear whether excluding feedback would have even been possible.

b. Stimulus parameters contributing to sensitivity

The present study replicated all of the significant effects previously demonstrated in the same task using similar stimuli in listeners with NH (Fig. 4; see also Table III of Anderson et al., [273]) with the exception of a standard AM rate × AM pairing configuration interaction ($p = 0.098$ vs $p < 0.05$ in Anderson et al. [273]), though these effects were not always in the same direction. Note that this interaction would be significant in the present study if sensitivity in d' units was not transformed as in Eq. 1 in order to minimize violation of model assumptions. The present study also found a significant effect of AM pairing configuration, which was not observed in listeners with NH ($p < 0.05$ vs $p = 0.127$ in Anderson et al. [273]).

For listeners with NH, though it was not reported in their manuscript, greater sensitivity to the standard AM rate of 90 Hz was restricted to the Same Place, Across Ears and Different Place, Across Ears pairing configurations [273]. Listeners with BiCIs showed greater sensitivity to standard AM rate of 10 Hz on average for all AM pairing configurations. This may relate to how listeners performed the task. Most listeners with BiCIs reported the primary differences between SAM pulse trains at or above 90 Hz as a change in pitch (with higher AM rates resulting in higher pitch) until the highest AM rate, which took on a "smooth" quality. There can be large differences in sound quality between ears (e.g., pitch perception; [156]) and place-of-stimulation (e.g., pitch ranking; [280]) for listeners with BiCIs, so detecting a change in pitch associated with varying AM rate might be much more difficult, especially with interleaved standard AM rates and AM rate roving.

In listeners with NH, the Different Place, Within Ears and Same Place, Across Ears configurations resulted in significantly higher sensitivity than the Different Place, Across Ears, and was restricted to the 10-Hz standard AM rate. There was no difference between AM pairing configurations at the 90-Hz standard AM rate. In listeners with BiCIs, the Different Place, Within Ears configuration resulted in significantly higher sensitivity than the Different Place, Across Ears and Same Place, Across Ears configurations on average (i.e., across both standard AM rates). Combining the results from listeners with NH and BiCIs, these findings suggest that spatial or spectral proximity may affect the ability to compare temporal information between places-of-stimulation. The presence of binaural beats, or at least a more functional binaural system in NH, might explain part of this difference between groups.

In the previously published study of listeners with NH, additional ANOVAs were completed in Experiment 2 excluding the two greatest variable AM rates ($\log(\Delta f_m/f_m)$) relative to the 90-Hz standard AM rate because they likely resulted in greater loudness than other stimuli in the experiment. In the present study, the two greatest variable AM rates ($\log(\Delta f_m/f_m)$) would have resulted in a minimum of two pulses per sinusoidal cycle for listeners with BiCIs. The auditory system does not "know" to reconstruct stimuli using sine waves from discrete pulses like a computer does. For example, with only two pulses per cycle, when the envelope phase was 0, the listener would be presented with a SAM pulse train with pulses at max and min level. When the envelope phase was π, the listener would be presented with a constant-amplitude pulse train. Thus, the quality of a sinusoidally amplitude-modulated pulse train with only two pulses per cycle could change drastically with its sinusoidal phase. During familiarization some listeners with BiCIs anecdotally described stimuli at highest AM rates (means of 1116 Hz; range of 837-1489 Hz) as changing in quality across presentations, whereas lower AM rates remained consistent across presentations. This confounding variable may have also contributed to the overall better performance at the standard AM rate of 10 Hz.

Based on psychophysical and physiological experiments, four to five pulses per cycle are recommended [277–279]. Thus, in order to more fairly compare results in the present study with NH results and to account for undersampling of sinusoids, we replicated the analysis in [273] excluding the two highest variable AM rates ($\log(\Delta f_m/f_m)$) from the dataset, with the exception that sensitivity was transformed as in Eq. 1. The results are reported in Table V. Similar to the results in NH, excluding the highest variable AM rates ($\log(\Delta f_m/f_m)$) resulted in a loss of significance of standard AM rate ($p = 0.090$ vs $p = 0.184$ in Anderson et al.[273]), indicating that the effect was driven primarily by the two highest variable AM rates. Listeners with NH only showed an effect of AM pairing configuration at 10-Hz standard AM rate. In contrast, listeners with BiCIs only showed an effect AM pairing configuration at the 90-Hz standard AM rate.

TABLE V. ANOVAs from Experiment 2 summarized by main effects and interactions. The second and fourth columns (labeled 10 and 90 Hz, and 90 Hz only) were computed excluding data for the two highest values of $\log(\Delta f_m/f_m)$.

	Original ANOVA		**10 and 90 Hz**		**10 Hz only**		**90 Hz only**	
Effect	**F-Statistic**	**p**	**F-Statistic**	**p**	**F-Statistic**	**p**	**F-Statistic**	**p**
$\log(\Delta f_m/f_m)$	F(6,623) = 38.626	< 0.0001	F(4,443) = 33.289	< 0.0001	F(6,308) = 27.228	< 0.0001	F(4,218) = 21.089	< 0.0001
Std. Rate	F(1,623) = 14.797	< 0.001	F(1,443) = 2.924	0.090	-	-	-	-
Pairing Configuration	F(2,623) = 4.218	< 0.05	F(2,443) = 2.785	0.063	F(2,308) = 0.248	0.781	F(2,218) = 5.006	< 0.01
$\log(\Delta f_m/f_m)$ × Std. Rate	F(6,623) = 3.724	< 0.01	F(4,443) = 3.494	< 0.01	-	-	-	-
$\log(\Delta f_m/f_m)$ × Pairing Configuration	F(12,623) = 0.307	0.988	F(8,443) = 0.387	0.927	F(12,308) = 0.200	0.998	F(8,218) = 0.517	0.843
Std. Rate × Pairing Configuration	F(2,623) = 2.335	0.098	F(2,443) = 1.529	0.218	-	-	-	-
Three-way interaction	F(12,623) = 0.231	0.997	F(8,443) = 0.262	0.978	-	-	-	-

To summarize, these analyses showed that stimulus parameters affected listeners with NH and BiCIs in a mostly similar way. The standard AM rate of 90 Hz resulted in better sensitivity for listeners with NH, in agreement with binaural studies showing better sensitivity for SAM stimuli [46,242,243]. Note that listeners with BiCIs also show reduced binaural sensitivity for SAM stimuli at low rates [46]. Assuming that binaural comparisons were more robust in listeners with NH, greater sensitivity for the Same Place, Across Ears restricted to 90-Hz standard AM rate is consistent with the broader literature. In contrast, listeners with BiCIs showed greater sensitivity at the 10-Hz standard AM rate. This effect was driven by the two highest variable AM rates ($\log(\Delta f_m/f_m)$) for both groups of listeners, and for listeners with BiCIs, this may relate to an undersampling of the highest AM rates used in the study. Both listeners with NH and BiCIs showed greater sensitivity of the Different Place, Within Ears compared to other configurations. It may be more ecologically valid to compare temporal structure across different frequency bands when sounds are not spatially separated, which could have therefore resulted in better sensitivity.

IV. GENERAL DISCUSSION

Patients with BiCIs are faced with many different challenges that limit access to sound source segregation cues. One important factor that has received ample attention in the last decade is the interface between CI electrodes and the auditory nerve, where a poor interface creates a bottleneck to sound source segregation cues. In the present study, we were interested in whether differences in the temporal fidelity of sounds conveyed via different portions of the auditory nerve contribute to the ability to compare temporal features across frequency (i.e., place-of-stimulation) and ears for listeners with BiCIs. Using an intuitive approach, Experiment 1 investigated monaural temporal thresholds (AM rate discrimination) as a proxy for the fidelity of the temporal information presented to two different places-of-stimulation in each ear (Sec. II). Experiment 2 then used these thresholds to predict the sensitivity to differences in AM rate between pairs of electrodes for simultaneously presented SAM pulse trains (Sec. III). The results indicated that the worse (higher) threshold in Experiment 1 successfully predicted sensitivity of the pair in Experiment 2. This result held whether the electrodes were paired within- or across-ears, at the same or different places-of-stimulation, but was most predictive for different places-of-stimulation within the same ear.

Interestingly, the degree of asymmetry in thresholds from Experiment 1 (i.e., the difference in threshold between the electrodes) was similarly effective at predicting sensitivity for pairs of electrodes (Table IV). Asymmetry in AM rate discrimination thresholds was only moderately correlated with the worse AM rate discrimination threshold, with Pearson correlation coefficients of -0.426, 0.210, and 0.237 for the three AM pairing configurations. This suggests that a place-of-stimulation yielding poorer temporal information, the difference in the fidelity of temporal information, or both may play a role in how accurately listeners are able to compare temporal envelope fluctuations across place-of-stimulation in both ears.

Results from the present experiment extend on those showing that pulse rate discrimination is predictive of binaural sensitivity [9] in several important ways. Firstly, high rates of stimulation with CIs are suspected to result in more stochastic firing of the auditory nerve [214]. Stochasticity in auditory nerve fibers is combatted by oversampling of auditory nerve fibers by specialized cells in the brainstem, resulting in improved phase locking (e.g., [54]). Very high rates of stimulation may therefore be useful in identifying areas of the auditory nerve that are “unhealthy” because of an increased need for accurate output from more nerve fibers. Secondly, AM stimuli used in the present experiment align more closely with the stimuli experienced in the real world. Low standard AM rates were used, which span a wide range of perceptual phenomena. Slower temporal envelope fluctuations are thought to be useful for speech segmentation and crucial to speech understanding (~1-15 Hz), result in the perception of loudness changes or "flutter" (5-15 Hz), result in a "rough" character (10-300 Hz), and do not elicit a pitch cue (≲100 Hz; for review, see [219] and Fig. 9 of [289]). Faster temporal envelope fluctuations result in the perception of temporal pitch, which is thought to be useful for discriminating voices based upon fundamental frequency [216,290]. Thirdly, results from the present experiment show that the bottleneck to temporal processing is not simply a binaural phenomenon. The binaural system relies on extremely precise inputs, and even when ideal stimuli are used, listeners with BiCIs show extremely variable sensitivity to binaural cues [44,45,105]. Thus, one might expect that any additional tax to the binaural system could worsen performance. Our results show that this bottleneck exists for all spectro-temporal comparisons, which has critical implications for real-world listening. For example, some programming strategies have shown improvement when poorly performing electrodes (defined differently according to experiment) are turned off (e.g., [191,264]), up to some limit where too many electrodes are deactivated and performance worsens [291]. The findings from the present experiment provide crucial support for those studies, showing that the ability to compare temporal information can be limited by the electrode yielding poorest performance or the difference in performance between electrodes.

A. Rate limitations

Experiments routinely show that higher pulse rates result in poorer discrimination of temporal pitch and interaural timing differences for listeners with CIs (an "upper-rate" limitation; e.g., [9,44,45,63]). Much less attention has been provided to the lower-rate limitation with respect to AM rate discrimination and sensitivity to interaural timing differences [46,193,194]. These lower-rate limitations for AM stimuli have been also been shown in listeners with NH [46,242,243,251,273], suggesting that there is some overlap in the mechanisms involved. The study by Ihlefeld and colleagues [9] showed that the upper-rate limitation correlates well between monaural and binaural processing. The present study and that by Anderson and collegaues [273] extended on those results, showing that the lower-rate limitation might not apply for SAM stimuli with large rate differences. While the results indicated that performance was worse for the 10-Hz standard AM rate in Experiment 1 (Sec. II), performance was significantly *better* for the 10-Hz standard AM rate in Experiment 2 (Sec. III). This effect was no longer present when the highest AM rates were removed from analysis, suggesting it may have been under-sampled by electrical pulse trails. Thus, Experiment 2 (Sec. III) indicates that performance was either better or no different for the 10- vs 90-Hz AM rates.

The lower-rate limitation probably has to do with the slope of the envelope for listeners with NH. Previous experiments in listeners with NH showed that a sharper slope of the envelope (like that near 100 Hz) results in better discrimination of interaural timing differences (e.g., [242,243,245,292,293]). In contrast, listeners with CIs show no benefit of sharp slope on AM rate discrimination [241,281] or interaural timing difference discrimination [245].

B. Temporal envelope conundrums

Recordings from the auditory nerve in animals implanted with CIs have shown substantially better phase-locking to stimuli than animals with NH [294,295]. Based on this finding, one would expect that listeners with CIs would have better sensitivity to some temporal cues than listeners with NH. However, this is not true of the vast majority of listeners with CIs, who demonstrate considerable variability depending upon the listener. A small subset of listeners show exceptional performance, better than that of many listeners with NH, and are often referred to as “star” listeners (e.g., [63,140]). Thus, it seems likely that factors associated with deafness limit the representation of temporal information beginning at the auditory periphery. Recent finding show that interaural timing differences are represented with similar fidelity when presented via acoustic or CI stimulation to gerbils with NH [296], supporting this hypothesis. Many different attempts have been made to assess the status of the auditory nerve. Some of these include the use of objective measures to predict temporal encoding (e.g., electrically evoked compound action potential; for review, see [297]), examining the location of the electrode array within the cochlea (e.g., [130,131]), relating psychophysical measures as in the present experiment (e.g., [9,192,264,298]), amongst others. Collectively these studies show that no measure rules all, and when combined with the apparent contrast of animal and human results, suggest that there are heterogeneous outcomes in patients because there are heterogeneous underlying physiological deficits.

There is also a dichotomy in experiments investigating the relative importance of temporal envelope cues to listeners with CIs. On one hand, without access to the temporal envelope, speech becomes unintelligible for simulations of CI processing [217,218]. On the other hand, listeners show little-to-no benefit of or sensitivity to temporal envelope cues to unmask sounds (e.g., [269,270]). The question remains: Why are listeners with CIs not extremely sensitive to a sound source segregation cue to which they have considerably greater access than many other cues? While our results cannot directly speak to this question, it seems plausible that temporal envelope cues are not very useful in real-world listening, or that they are most useful when combined with good frequency resolution across spectral channels.

C. Limitations

The results of the present study were compared in-depth to a similar study in listeners with NH [273]. While trends were somewhat similar, poorer performance of listeners with BiCIs could be attributed to many different listener-dependent factors. The first obvious differences between listeners with NH and BiCIs is the nature of electrical stimulation and the effects of deafness. Listeners with BiCIs were also presented with stimuli that were loudness balanced. While it was not assessed, the listeners with NH likely experienced considerable differences in loudness according to carrier frequency (their Fig. 7; [273]). Increasing level results in better temporal representations of sounds (e.g., phase-locking of the auditory nerve; [210]). Thus, using loudness-balanced stimuli makes it difficult to compare the effects of place-of-stimulation (base vs apex) in listeners with BiCIs to listeners with NH, along with the fact that place-of-stimulation was a within-subjects factor for listeners with BiCIs but not NH. Listeners with BiCIs were much older (mean 63 years) on average from listeners with NH (mean 22 years). Aging independent of hearing loss leads to degradation of sensitivity to temporal cues (e.g., [46,299,300]).

Another factor that is a serious limitation to CI research is the demographic homogeneity of the participants. Highly controlled laboratory settings that require listeners to take extended time away from home or work heavily bias the populations of patients who participate in research. It is likely that this plays a role in the bias toward older age-at-testing for listeners with CIs shown in many studies. Our laboratory anonymized and pooled demographic data from adults with CIs and NH that have participated in all previous experiments and showed substantial bias toward White, non-Latinx participants (77%), and underrepresentation of Black (4%) and Indigenous (0.2%) participants. Recent research from other fields of healthcare demonstrated health disparities associated with race, with Black and Indigenous people of color in the United

States being at greatest risk (e.g., [301–303]). Further concerning is underrepresentation of people of color in audiology. As of 2019, the American Speech-Language-Hearing Association (ASHA) reported that 92% of participant audiologists were White [304]. By excluding people of color from CI experiments, researchers lack awareness of the issues that afflict these clinical populations and contribute to the lack of representation in hearing science. It is difficult to predict how this limitation exactly affects the interpretation of our results as it is a grave limitation in the CI literature more generally. Many of the measures in Sec. IV.B. have not yet been used to relate to sensitivity to psychophysical cues in the clinic where it is easier to sample from groups of listeners more racially representative of listeners with hearing loss. Thus, it is crucial that researchers and clinicians keep an open mind to the challenges faced by patients in different demographic groups, especially based upon race, and do not overgeneralize results from experiments.

D. Summary and conclusions

Together from Experiment 1 and 2, several conclusions can be drawn with respect to sensitivity to AM rate in listeners with BiCIs:

1. Listeners showed better (lower) thresholds to sequential presentations of AM rate at apical electrodes compared to basal electrodes, extending on previous results in listeners with unilateral CIs (Fig 2; [193,194,241,281,287]).
2. Several listeners were unable to identify differences in AM rate for SAM pulse trains presented to two electrodes simultaneously (Fig. 3), suggesting that the usefulness of the temporal envelope may be limited in listeners with CIs, in agreement with previous experiments [269,270].

3. Listeners were most sensitive to differences in AM rate for simultaneously presented stimuli when electrodes were paired at the different places-of-stimulation within the same ear compared to the same or different places-of-stimulation across the ears (Table V; Fig. 4).
4. Listeners were most sensitive to AM rate presented sequentially (Fig. 2) when the standard AM rate was 90 Hz compared to 10 Hz, in agreement with previous experiments in listeners with CIs using a different task [193,194] and NH using the same tasks [273].
5. The electrode with a poorer AM rate discrimination threshold was predictive of the sensitivity of pairs of electrodes to differences in AM rate (Table III and IV; Fig. 5), in agreement with Ihlefeld and colleagues [9]. The difference in AM rate discrimination thresholds was similarly effective at predicting sensitivity of the pair (Table IV). This finding held when stimuli were presented to pairs electrodes at the same approximate place-of-stimulation in both ears or different places-of-stimulation within the same ear, but was not statistically significant for pairs of electrodes at different places-of-stimulation in both ears.

FOOTNOTES

1. The transform used in Eq. 1 is also equivalent to a specific case of the logistic transform where the growth rate parameter is 1 and the midpoint is the minimum value of sensitivity.

2. All other analysis in this manuscript is based upon the Type I Sum of Squares from the ANOVA or pairwise-comparisons via estimated margin means, except for this point, which was assessed using the summary of the linear model in terms of regression rather than ANOVA.

Summarizing in terms of linear regression has the convenient property that categorical variables are expressed relative to the intercept (baseline condition).

Chapter 4: Asymmetric Temporal Envelope Encoding: Lateralization and Interactions with Envelope Shape and Interaural Place-of-Stimulation Mismatch

I. INTRODUCTION

Binaural cues provide essential information for listeners to determine the location of a sound source in the horizontal plane and distinguish target speech from background noise. These cues consist of interaural time and level differences (ITDs and ILDs, respectively). Listeners with normal-hearing (NH) rely more heavily upon ITDs at low frequencies to localize broadband sound sources [305] and derive greater advantage when target and masking sounds are spatially separated if ITDs are provided compared to ILDs [29,306,307]. The benefits of ITDs to listeners with NH have motivated researchers over the years to attempt to provide ITDs via cochlear implant (CI) processors and improve ITD sensitivity, but the latter demonstrate extraordinary variability with respect to their sensitivity to ITDs. This occurs even under ideal stimulus conditions when the timing of pulses to electrodes in the two ears is carefully controlled, and loudness balanced stimuli are used [44,45,93,105,159,160]. Many factors limit listeners' access and sensitivity to ITDs conveyed via bilateral CIs (BiCIs), related to both the hardware and software associated with BiCIs, as well as patient histories and outcomes.

A. Limitations in access to ITDs

Sound processing algorithms in cochlear implants (CIs) disregard a sound's temporal fine-structure, or fast changes in displacement over time [35], thereby only allowing access to ITDs provided in the temporal envelope, or slower, gradual changes in displacement over time [36]. Because BiCIs are not coordinated in commercially available sound processors, random and stimulus-independent ILDs may be introduced to sounds by automatic gain control algorithms [308,309]. Similarly, automatic gain control and uncoordinated CI clocks may lead to the introduction of spurious ITDs [36].

Listeners with BiCIs demonstrate sensitivity to ITDs for low rates (<300 Hz) in the temporal envelope [46,276,310,311]. However, these listeners do not rely on ITDs for sound source localization [40–42]. Sensitivity to ITDs has been shown to vary considerably across individuals, and relates to several factors that vary from patient-to-patient, e.g., [9,93,105]. Several of these factors make up the "electrode-neuron interface," or the relationship between CI electrodes that stimulate auditory nerve fibers to convey auditory information to the brain (for review, see [125]). While many studies investigating the role of the electrode-neuron interface in patient outcomes focus upon unilateral stimulation, differences between the ears likely play a role in binaural outcomes, generating different kinds of asymmetries between the ears. Figure 1 shows some examples of several factors that may affect representations of, and therefore sensitivity to, binaural cues and especially ITDs. These consist of factors that lead to spectral and temporal asymmetries.

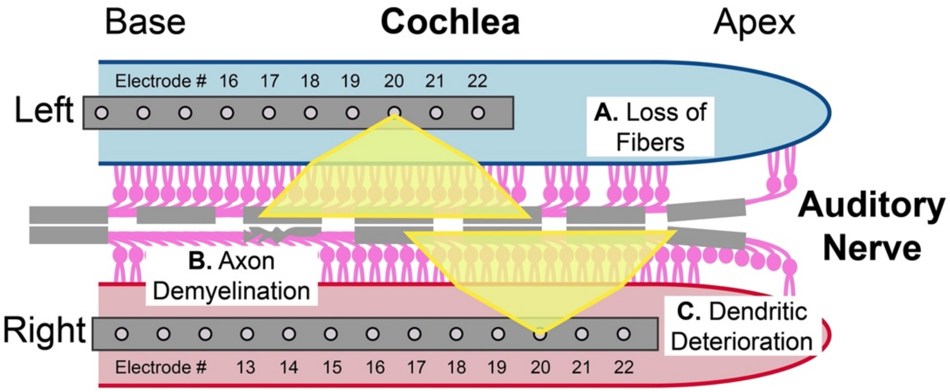


FIG. 1. Illustration of the electrode-neuron interface. This illustration shows examples of factors that can affect the spectro-temporal representations of sounds in each ear for listeners with BiCIs.

Sensitivity to ITDs also depends upon the parameters of the stimulus. A classical example is that sensitivity to ITDs in the temporal fine-structure is greater for low frequencies, which formed the basis of “duplex theory” of sound source localization (i.e., ITDs are only used to localize low-frequency sounds) [312]. Because listeners with BiCIs receive processed stimulation, they may not be limited by the same physiological limitations as listeners with NH, and therefore show no consistent differences in sensitivity to ITDs across regions corresponding to different frequencies [44,105]. In contrast, listeners with NH show greatest sensitivity to fine-structure ITDs below 1500 Hz e.g., [313] and envelope ITDs e.g., [61]. For both listeners with NH and BiCIs, when presented with pulsatile, band-limited stimulation, ITD sensitivity declines above rates of 300 Hz [44,45,61]. Sensitivity to ITDs is also greater for stimuli with a sharp attack time in listeners with NH [242,243,245,292,293,314], but may be related to other factors for listeners with BiCIs e.g., off-time [245]. This final point is especially important for interpreting the stimuli to which a listener may retain ITD sensitivity and how CI processing strategies may be able to most effectively convey ITDs.

B. Binaural encoding

An elementary understanding of the binaural circuits in the brainstem can be used to make predictions about how changes in the electrode-neuron interface affects the encoding of ITDs. A detailed review is provided by [34]. Briefly, the firing rate of cells in the medial superior olive (MSO) varies with fine-structure ITDs [315]. Further, MSO cells have a “best delay” or “best phase” to which they respond maximally, which is thought to correspond to a characteristic delay in the arrival of inputs from the ipsilateral and contralateral ear to the action potential initiation site (for review, see [316]). MSO cells act as coincidence detectors [54], responding when excitatory input from the ipsilateral and contralateral ear arrive within a given time window. They are accordingly referred to as excitatory-excitatory (EE) cells. The firing rate of the lateral superior olive (LSO) varies with envelope ITDs [51]. Cells in the LSO receive excitatory input from the ipsilateral ear and inhibitory input from the contralateral ear [52,53,64] and are referred to as excitatory-inhibitory (EI) cells. This means that LSO cells act as a sort of anti-coincidence detector, responding minimally when sounds arrive with 0-μs ITD and maximally at larger ipsilateral or contralateral ITDs.

Critically, both MSO and LSO cells have extremely short time constants, meaning that their spike rate is only modulated when the timing of inputs is highly coordinated [48,53,54]. This is accomplished by a high degree of phase locking, or an increased probability of firing a spike when inner hair cell stereocilia are maximally, positively deflected. Phase locking occurs for the fine-structure and envelope. For listeners with CIs, hair cells are bypassed and auditory nerve fibers instead phase lock to pulsatile, electric stimulation. Each problem associated with the electrode-neuron interface in Fig. 1 is expected to cause problems with phase locking and ITD encoding. For example, uniform [109] or site-specific losses (i.e., dead regions [111]) of auditory nerve fibers limit number of auditory nerve fiber inputs provided to the binaural pre-processors in the cochlear nucleus, which relies on oversampling of inputs by many auditory nerve fibers to retain its exquisite sensitivity to ITDs e.g., [74]. Similarly, axon demyelination leads to poorer phase locking in models of auditory nerve fibers [317]. Deterioration of dendrites lead to an increase in the latency of the action potential in model auditory nerve fibers [115], and could thus introduce stimulus-independent ITDs to neural responses. Collectively, these would result in temporal degradations in the stimulus, which may be either symmetric or asymmetric.

Thus, listeners with BiCIs likely experience a large amount of interaural temporal asymmetries that must be navigated by the binaural system, thereby limiting its performance. For example, loss of auditory nerve fibers, axon demyelination, and dendritic deterioration have been related to the duration of deafness in studies of human temporal bones [110] and animal models of hearing loss [106,117,274]. Many listeners with BiCIs are implanted during adulthood receive their CIs sequentially [3,82], suggesting that the duration of deafness differs for each ear. Similarly, loudness growth depends upon the electrode being stimulated in listeners with BiCIs [126,127,318,319]. Together, these studies suggest that for envelope ITDs with interaural temporal asymmetry, inputs to binaural nuclei may be decorrelated, resulting in a diffuse sound image [129], poorer sensitivity to ITDs [128,320], and poorer spatial unmasking of speech [29]. Moreover, the extent of this effect may depend upon the stimulus, as stimuli with a sharp attack time are thought to initiate "looks" of the binaural system to ITDs in the stimulus [321,322], improving sensitivity to ITDs [292,293,323].

Studies have assessed the impact of interaural mismatches in place-of-stimulation on ITD discrimination and lateralization. Lateralization represents the extent to which listeners perceived a sound to the left or right inside of their head as binaural cues, in this case ITDs, are varied. Lateralization is therefore thought to reflect the utility of spatial cues to impact perception. Mismatches in the place-of-stimulation that result after CI surgery lead to poorer sensitivity to ITDs [144,324] and poorer lateralization [138,139]. Interaural frequency disparities for listeners with NH, resulting in interaural place-of-stimulation differences, also lead to poorer sensitivity to ITDs conveyed in the temporal envelope [163,325–328] and worse lateralization [138,139,163]. Interaural frequency mismatch also leads to a smaller binaural interaction component in listeners with NH [137], thought to reflect processing in the LSO [135].

Both EE and EI type models have been used successfully to predict binaural performance using stimuli that are relevant for listeners with CIs, that is, pulses of high-frequency, acoustic transients, e.g., [61,137,329,330]. The role of the MSO and LSO remains contested in the literature for listeners with BiCIs (for review, see [331]). Most modeling studies have focused on interaurally symmetric representations of stimuli. The exception is the study by Brown and colleagues [137], which investigated the effect of interaural place-of-stimulation mismatch and successfully predicted its effects on the binaural interaction component using a model of the LSO. Thus, one important dimension that remains to be tested is interaural temporal asymmetry, as well as its interaction with interaural place-of-stimulation mismatch.

C. Summary and Hypotheses

1. Motivation and goals

Sensitivity to binaural cues might be limited by the ear with poorer temporal representations in listeners with BiCIs [9] and simulations in NH [273]. These differences are presumably mediated by differences in the extent of auditory deprivation and electrode-neuron interface in each ear. Interaural temporal asymmetries resulting from the electrode-neuron interface and auditory deprivation are compounded by spectral asymmetries introduced by place-of-stimulation mismatch. Leveraging our knowledge of binaural processing, the present study aimed to explore the effects of interaural temporal and spectral asymmetry in temporal representations on the lateralization of envelope ITDs and using models of binaural processing that were developed with an assumption of interaural symmetry. We used stimuli with different envelope shapes to determine whether interaural temporal and spectral mismatch have different effects. It is difficult to control or account for degrees of interaural asymmetry in listeners with hearing loss because the sources are so numerous and patient outcomes are highly heterogeneous. Therefore, the present study tested listeners with NH using a simulation of interaural asymmetry in temporal representations (see Sec. I.C.2).

The present study investigated whether the ear with worse temporal fidelity was predictive of poorer ITD lateralization. The proof of concept was tested in Experiment 1 (Sec. II). In Experiment 2 and 3 (Sec. III and IV, respectively), we investigated interactions between interaural asymmetry of temporal representations with two additional stimulus parameters known to affect sensitivity to envelope ITDs: (1) attack time of the temporal envelope and (2) interaural mismatch in place-of-stimulation. In each experiment, lateralization was compared against predictions of existing models of binaural processing.

Most of the research concerning ITD sensitivity have used low-rate, pulsatile stimulation, which tends to yield optimal sensitivity to ITDs in listeners with NH [46,242,243] and BiCIs [44–46]. The present study used high-rate, acoustic pulse trains simulating the spread of excitation that occurs with electrical stimulation and applied a second-order amplitude modulation (AM) as would occur during CI stimulation (see Sec II.B.2). The stimuli used in these experiments were meant to be more relevant for listeners who use BiCIs in the real world, whose auditory system is not stimulated at rates below 300 pulses per second (pps) and only receive envelope ITDs through their clinical processors [36]. The present study used whole-waveform ITDs, which would preserve ITDs in the temporal-fine structure. However, sensitivity to and lateralization of high-rate acoustic pulse trains is poor [61,275,314].

2. Approach and hypotheses

Temporal representations in each ear were manipulated by controlling the AM depth of the second-order modulator. When AM depth is small, auditory nerve fibers demonstrate poorer phase locking to the temporal envelope, e.g., [210]. Figure 2 shows the degree of phase locking of model auditory nerve fibers from Bruce and colleagues [332] to 500 pulse-per-second acoustic pulse trains with a 100-Hz sinusoidal second-order modulator (see Sec. II. B. for more details).

The results suggest, consistent with auditory nerve data using sinusoidally amplitude-modulated (SAM) tones, that phase locking to the second-order modulator decreases with decreasing AM depth. As the second-order AM depth decreases, stochasticity increases for the auditory nerve, providing noisier input to binaural pre-processing neurons in the cochlear nucleus.

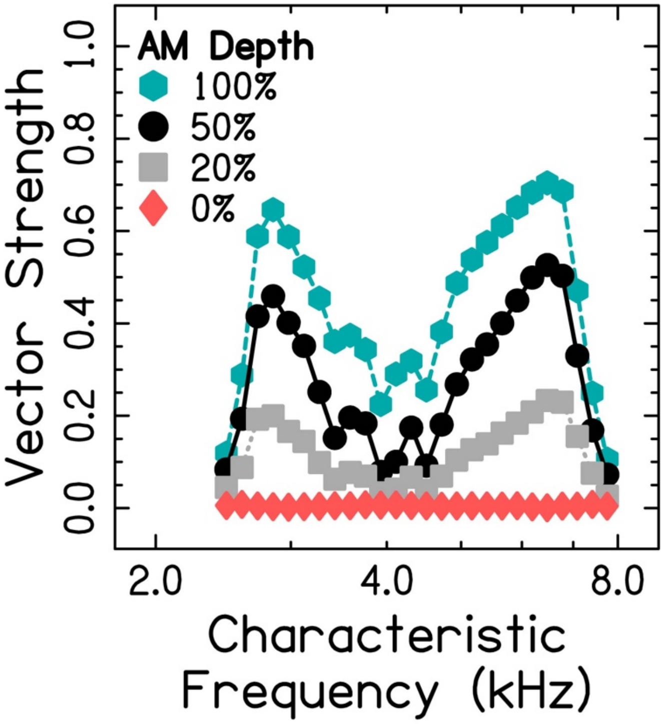


FIG. 2. Phase locking of model auditory nerve fibers. Model, medium spontaneous rate auditory nerve fibers [332] were used to predict changes in phase locking with varying AM depth of high-rate, amplitude-modulated acoustic pulse trains based on 10000 simulations. The x-axis corresponds to the characteristic frequency of each fiber. The y-axis corresponds to the vector strength [333]. Color and shape correspond to the AM depth of the second-order modulator as described in the figure legend.

Then, due to this temporally stochastic input to the cochlear nucleus, the input to the MSO and LSO should contain greater neural interaural decorrelation. Consistent with this hypothesis, several previous studies have shown that sensitivity to envelope ITDs or interaural envelope phase decreases as the AM depth in both ears decreases [314,327,334]. Similarly, a recent study by Anderson and colleagues [273] showed that sensitivity to interaural differences in AM rate decreased as AM depth decreased *in one ear*. It was therefore hypothesized that if the second-order AM depth was decreased in one or both ears, envelope ITDs would be lateralized to a lesser extent. It was further hypothesized that AM depth would have additive effects with shallow envelope attack (Experiment 2; Sec. III) and interaural place-of-stimulation mismatch (Experiment 3; Sec. IV) because these stimulus parameters also increase the amount of interaural decorrelation.

II. EXPERIMENT 1: SYMMETRIC VS. ASYMMETRIC AM DEPTH

A. Motivation

The goal of Experiment 1 was to determine whether lateralization of a sound is limited by the ear with poorer temporal fidelity. Thus, three experimental conditions were devised: Conditions with two "good" or "bad" ears, and an interaurally asymmetric condition with one "good" and "bad" ear. To avoid floor or ceiling effects, 50% and 20% AM depth were used. These were the same AM depths tested by Anderson et al. [273].

B. Methods

1. Listeners

Ten listeners with NH participated in this experiment (age 19-25; mean age: 22.4 years). All listeners had hearing thresholds ≤20 decibels hearing level (dB HL) for octave frequencies between 0.25 to 8 kHz, with no difference in threshold between the left and right ear greater than

10 dB at any frequency. Before participating in the experiment, all listeners provided informed consent. All procedures were approved by the Health Sciences Institutional Review Board of the University of Wisconsin-Madison.

2. Stimuli and equipment

Stimuli were Gaussian-enveloped tones with a center frequency of 4 kHz and second-order, sinusoidal AM. Briefly, Gaussian-enveloped tones are generated by creating a train of Gaussian pulses and multiplying the result by a sine wave of the desired duration and center frequency. More details regarding Gaussian-enveloped tones can be found in [163,335]. The stimuli used in the present experiment were generated at a rate of 500 pps with 3-mm bandwidth, i.e., 3 dB roll-off at 3 mm about the region stimulated along the cochlea [336], and had a duration of 500 ms. Then, second-order sinusoidal AM was applied by multiplying the constant-amplitude pulse train using Eq. 1:

$$y(t) = (1 + m\sin(2\pi f_m t))\, x(t) \qquad (1)$$

where $y(t)$ is the instantaneous amplitude of the SAM pulse train at time t, m is the modulation depth (in proportion, ranging from 0 to 1, and set either to 0.2 or 0.5), f_m is the modulation frequency fixed at 100 Hz, t is a vector of time values, and $x(t)$ is the instantaneous amplitude of the constant-amplitude pulse train at time t. Stimuli were presented at a level of 65 dB sound pressure level A-weighted [dB(A)]. SAM pulse trains were gated with 10-ms cosine ramps at stimulus onset and offset. Whole-waveform ITDs of 0, ±100, ±200, ±400, or ±800 μs were then applied. Finally, 61 dB(A) masking noise that began 100 ms prior to stimulus onset and persisted until 100 ms post stimulus offset was provided to prevent listeners from exploiting low-frequency non-linear distortions [293]. The noise had two cut-off frequencies: the first at 200 Hz where the roll-off was -3 dB/octave and the second at 1000 Hz where the roll-off was -18 dB/octave. The AM depth in each ear is described as 20%:20%, 20%:50%, or 50%:50% depth, representing the depth in each ear. In the 20%:50% condition, the ear with smaller depth could have been in the left or right ear.

Stimuli were presented through ER-2 insert earphones. Stimuli were generated in MATLAB and presented using a Tucker-Davis Technologies System3 with RP2.1, HB7, and PA5 units (digital processor, amplifier, and attenuator, respectively). All testing took place in a double-walled, sound-attenuating booth (Industrial Acoustics Company, Inc.).

3. Procedure

All listeners were familiarized with lateralization prior to beginning the experiment. First, listeners were shown an illustration of the graphical user interface (a cartoon face) used to judge the intracranial location of sounds. The experimenter pressed buttons on the screen further to the left or right, resulting in the presentation of sounds with a left- or right-leading ITD, respectively. Listeners were then allowed to use this interface until satisfied and encouraged to compare stimuli ITDs near 0 μs (i.e., lateralized closer to the center of the head and therefore more difficult to distinguish). Listeners were then given a pre-test where the experimenter presented stimuli with left- or right-leading ITDs and asked listeners to indicate the side on which the sound was perceived. Finally, listeners were tested with five repetitions of 0 and ±800 μs ITDs in random order. Stimuli had 50%:50% AM depth. They could repeat the stimuli as many times as desired. If listeners responded with approximately 50% lateralization range, then they proceeded to the rest of the experiment. If not, they were tested with five repetitions of 0, ±800, and ±1600 μs ITDs, and then again with 0 and ±800 μs ITDs. In all cases, by the time that the rest of testing began, their lateralization range was at least approximately 50% for 50%:50% AM depth.

Throughout the rest of testing, 20 repetitions per ITD were collected for 50%:50%, 20%:50%, and 20%:20% AM depth in each ear. Half of listeners were tested with 20% AM depth in the left ear for interaurally asymmetric conditions, and the other half with 20% AM depth in the right ear (both denoted 20%:50% throughout this manuscript). Testing was blocked by AM depth configuration and counterbalanced across listeners. This resulted in a total of 20 repetitions × 9 ITDs × 3 AM depth conditions = 540 trials, which were completed in approximately two hours. Listeners were allowed to repeat stimuli as many times as desired and submit final answers before beginning the next trial. Listeners were allowed to respond with multiple source locations. Only 0.5% of trials included multiple source responses. Thus, only the most prominent source as indicated by the listener was considered during analysis.

4. Analysis and modeling

a. Analysis. Listeners' response locations for each stimulus were converted into numerical values within the range of ±10, corresponding to the left (negative) and right (positive) sides. Lateralization responses were assessed using a mixed-effects analysis of variance model (ANOVA) with version 1.1-21 of the lme4 package in R [285]. Raw responses from each listener were included as the dependent variable, predicted by ITD, parameters tested in the experiment (for Experiment 1, AM depth), and all possible interactions. The null hypothesis for each factor in the model was that the mean lateralization response was equal across all values of that factor. Then a significant effect of ITD, for example, which was treated as a nominal variable in the model, would indicate that there was a statistically significant difference between at least one pair of ITDs. Post-hoc comparisons were then used to determine for which ITDs this is true. A significant main effect of AM depth, for example, would indicate a bias toward the left or right associated for different AM depth conditions. Significant interactions between ITD and AM depth, for example, would indicate a difference in the shape of the ITD-lateralization function for at least one ITD between AM depth conditions. Thus, the ANOVAs included here were highly efficient analyses that capture many different aspects of the lateralization function within a single model and used all of the data in the experiment. They also accounted for the variance associated with each possible comparison. We included random intercepts associated with listener to account for typical biases associated with listener observed in lateralization experiments [337]. All post-hoc paired comparisons were completed using estimated-marginal means with Tukey adjustments for multiple comparisons using the *emmeans* package in version 3.6.0 of R. Degrees of freedom were estimated with Kenward-Roger approximation [286] in the *lmerTest* package of R.

An alternative approach would have been to compute sensitivity (e.g., d' [81]) for pairs of ITDs and comparing the ITD-sensitivity function across different stimulus parameters. However, in so doing, the experimenter collapses 20 repetitions at nine different ITDs to four sensitivity values, substantially reducing statistical power. More importantly, sensitivity is calculated in a pairwise fashion, meaning that it is always expressed relative to some other condition. Therefore, a sensitivity analysis like the one described would not account for the bias associated with a particular condition. It would also only account for the relative variance between the pairs of ITDs, not the variance of the response distribution.

Lateralization responses were fit using the function described in Eq. 2:

$$Lat(ITD) = A\,\mathrm{erf}\left(\frac{ITD - \mu_{ITD}}{\sqrt{2}\sigma}\right) - \mu_{Lat} \qquad (2)$$

where *Lat(ITD)* corresponds to the predicted value of lateralization at a particular ITD, *A* corresponds to maximum extent of lateralization achieved by the fit, erf corresponds to the error function which takes as input any real number and whose output is bounded between ±1, *ITD* corresponds to the value of the ITD in μs, μ_{ITD} corresponds to the horizontal shift, σ is related to the lateralization slope, and μ_{Lat} corresponds to the vertical shift. All curve fitting was completed using non-linear least squares via the curve fitting toolbox in MATLAB.

The lateralization range, or extent to which listeners perceived a sound to the left and right in the head, was assessed to provide a proxy of the impact of spatial cues on perception. Lateralization range was calculated as the absolute difference of the perceived left and right positions for each listener or output of the models discussed in Section II.B.4.b for ITD values of ±800 μs from the fitted curve. Additional ANOVAs were used to predict the effects of stimulus parameters on the lateralization range. They included random intercepts associated with listener.

b. Modeling. Two computational models of binaural processing were used to predict lateralization performance. They were chosen because they have been highly effective at predicting lateralization of trains of high-frequency, acoustic transients, they represent EE- and EI-type binaural processing, they are both available via open source software, and they are both highly computationally efficient. The first implementation was provided by open-source code[1] described in [338]. The Akeroyd model implementation is a frequency- and delay-weighted version of the Jeffress model [339], where delayed copies of the signal in each ear are compared and a correlation is computed. In other words, it computes a cross-correlation, and the centroid of this cross-correlation is then regressed against the data observed in an experiment. The particular implementation used in the present study involved the cross-product rather than cross-correlation (i.e., a normalized version of the cross-product). Examples of similar models have an extensive history that is beyond the scope of the present chapter. The particular parameters chosen for the Akeroyd model implementation used in the present experiment were motivated by the models used in [242,340], who measured sensitivity to and lateralization of transient, high-frequency stimuli similar to those used in the present experiment, except that our stimuli included second-order AM. The parameters are summarized in Table I. Frequency-weighting functions [341] and delay-weighting functions from [342] were used.

TABLE I. Akeroyd model implementation parameters. These parameters are specified in the "mcorrelogram" function in MATLAB.

Parameter	**Value**
Lower Frequency	2000
Upper Frequency	8000
Filters per ERB	1
Lower ITD	-2000
Upper ITD	2000
Transduction	"v=3" (halfwave rectification + power-law (^3) expansion of waveform)
Binaural Switch	Crossproduct

The second model implementation was provided in the supplementary code accompanying Klug and colleagues [329]. The Klug model implementation is based upon processing in the LSO of the brainstem, where the input from the ipsilateral and contralateral side is excitatory and inhibitory, respectively. Processing in the lateral superior olive results in sensitivity to ITDs for high-frequency stimuli [51], which can be modeled by counting the number of coincident spikes in a given time window [343]. We simply used the parameter defaults recommended in the paper and provided in their code. Klug and colleagues [329] showed a small improvement in the modeling of the data from [340]. Thus, both the Akeroyd and Klug model implementations seemed like ideal candidates for modeling the behavioral results in the present study.

The output of the Akeroyd model implementation returns the centroid ITD from the cross-product and the Klug model implementation returns the spike rate difference for each side of the head. In both cases, the output of the model was related to the mean lateralization response across listeners via linear regression. Upon pilot testing, it was found that both model implementations predicted biases that were inconsistent with behavioral performance. Recent physiological evidence indicates that the auditory brainstem adjusts its rate-ITD and rate-ILD function according to the distributions of the cues presented [65]. Moreover, intuition indicates that if someone is always biased by a specific amount in one direction, they could simply and predictably adjust for this bias over time. Thus, we first centered the model output by subtracting its mean. Then, when performing regressions, we included effects of stimulus conditions (e.g., AM depth), allowing the mean of the model output (i.e., intercept of the regression) to shift but retaining the relationship between ITD and model output. This resulted in improved fit of the model output to the data compared to fitting uncentered and inappropriately biased model output to the lateralization responses.

C. Results

1. Behavioral results

The goal of this experiment was to determine whether the ear with poorer temporal fidelity (i.e., smaller AM depth) dominates lateralization of envelope ITDs. Results averaged across listeners are shown Fig. 3, compared against predictions from the Akeroyd (Fig. 3(A)) and Klug (Fig. 3(B)) model implementations. Individual lateralization functions were highly variable. ITD significantly impacted lateralization response [$F(8,5724) = 1098.561$, $p < .0001$], with significant differences between all possible pairwise combinations of ITDs ($p < .0001$). Results in Fig. 3 demonstrate no consistent bias toward the left or right side with interaurally symmetric or asymmetric AM depth, which would have been indicated by an upward- or downward-shift along the y-axis. However, there was a significant difference between AM depth conditions [$F(2,5277) = 7.251$, $p < .001$]. Post-hoc comparisons revealed that there were significant differences between 50%:50% and 20%:20% [mean difference: 0.363; $z = 3.330$, $p < .01$], as well as 20%:50% and 20%:20% AM depth [mean difference: 0.361; $z = 3.258$, $p < .01$], but not 50%:50% and 20%:50% [mean difference: 0.002; $z = 0.021$, $p = 1.00$]. Thus, stimuli with 50% AM depth tended to be lateralized slightly further to the right. The lateralization of the largest magnitude ITDs is slightly closer to the center of the head for stimuli with at least one ear containing 20% AM depth, which may have played a role in this finding. There was also a significant ITD × AM depth interaction [$F(16,5274) = 5.884$, $p < .0001$], suggesting that the shape of the lateralization function differed between AM depth conditions.

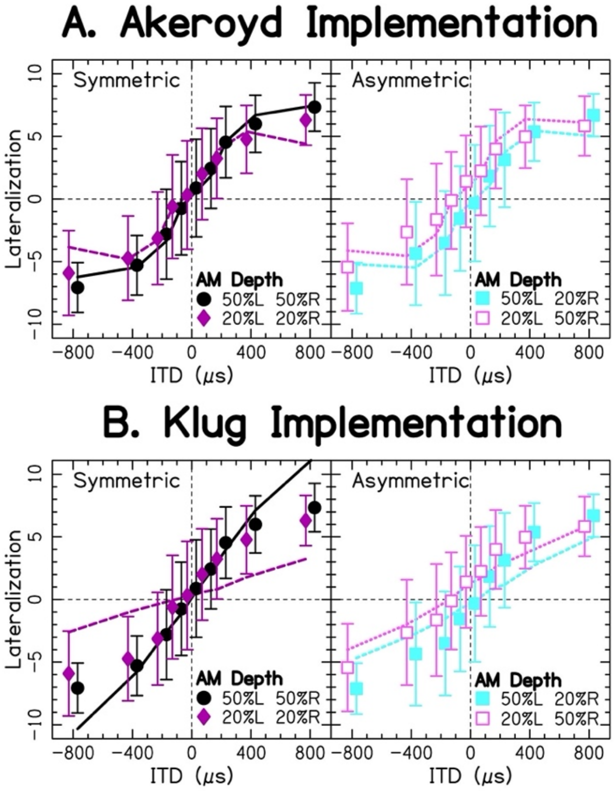


FIG. 3. Lateralization of ITDs by AM depth. The left and right panel correspond to symmetric and asymmetric AM depth conditions, respectively. The x-axis corresponds to the ITD presented on each trial. Negative values indicate a left-leading ITD. The y-axis corresponds to the extent of lateralization, with ±10 indicating lateralization to left or right the side of the head. By extension, a lateralization value of 0 represents the center of the head. Negative values indicate a leftward perception. Points show the mean were across listeners ± one standard deviation. (A) Curves represent predictions from the Akeroyd model implementation. (B) Curves represent predictions from the Klug model implementation.

It was of particular interest to assess the effect of AM depth on lateralization range, since lateralization range conveys the impact of ITDs on listeners’ perception. Results are shown on average and for individuals in Fig. 4(A). There was a significant effect of AM depth [F(2,18) =

7.693, p < .01]. Post-hoc comparisons revealed that lateralization range was significantly larger for 50%:50% compared to 20%:50% [t(18) = 2.320, p < 0.01], as well as 50%:50% compared to 20%:20% AM depth [t(18) = 3.139, p < .05], but not 20%:50% and 20%:20% AM depth [t(18) = -0.468, p = 0.887].

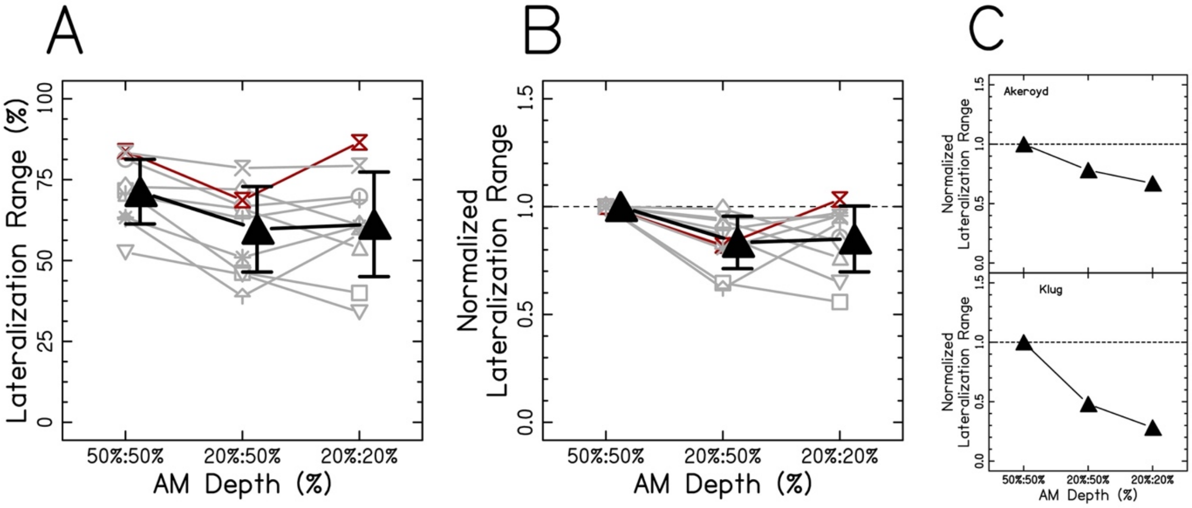


FIG. 4. Lateralization range by AM depth. The x-axis corresponds to the AM depth, and the y-axis corresponds to the lateralization range. (A) Mean data ± one standard deviation are shown in black. Individual performance is shown offset in grey or red. Listeners shown in red had a lateralization range that was greater than that observed with 50%:50% for 20%:50% or 20%:20% AM depth (contrary to hypotheses). Lateralization ranges were normalized to 50%:50% AM depth. (B) Predictions from models of binaural processing. Lateralization ranges were normalized to 50%:50% AM depth.

2. Modeling

The Akeroyd model implementation was highly accurate at predicting the lateralization responses averaged across listeners (adj. $R^2 = .949$; Fig. 3(A)). The Klug model implementation, on the other hand, showed exaggerated effects of AM depth (adj. $R^2 = .786$; Fig. 3(B)). Because individual performance was so variable, it was of interest to determine whether the trends in lateralization range with AM depth differed between models. Figure 4(B) show the lateralization ranges for the Akeroyd and Klug model implementations. Lateralization ranges were computed in the same fashion for the model predictions as for behavioral data (see Sec. II.B.4.a). Similar to Fig. 3, Fig. 4(B) shows exaggerated effects from the Klug compared to the Akeroyd model implementation, though for both models, directions of trends were in-line with behavioral results. Interestingly, both model implementations predicted a slightly larger lateralization range for the 20%:50% compared to the 20%:20% AM depth conditions, in contrast to the behavioral data.

D. Discussion

Results from Experiment 1 demonstrate an important proof-of-concept for lateralization of high-rate, amplitude-modulated pulse trains. That is, if *one* ear had smaller AM depth, which should reduce phase locking of model auditory nerve fibers (Fig. 2), lateralization range decreased. These results, combined with those from listeners with NH [273] and BiCIs [9] investigating sensitivity to interaural phase and ITDs, respectively, provide new evidence that the ear with poorer temporal representations limits binaural processing. The present study expands on the previous results by demonstrating that this limitation does not extend solely to discrimination, but to the utility of ITDs on affecting perception (i.e., lateralization).

III. EXPERIMENT 2: INTERACTIONS WITH ENVELOPE SHAPE

A. Motivation

Experiment 1 demonstrated the important proof of concept that one ear with reduced AM depth is enough to degrade ITD lateralization, and is not significantly different from both ears operating with poor temporal fidelity. In Experiment 2, we explored the relationship between different envelope shapes and interaurally asymmetric AM depth. In everyday listening, the sounds that we lateralize using envelope ITDs vary in their shape. That is, some sounds with characteristic envelope shapes may be easier to lateralize than others. Examples include transients and sounds with a sharp attack times [292,314]. Thus, to gain a clearer understanding of the impact of interaurally asymmetric temporal fidelity on a wider set of listeners' experiences, Experiment 2 evaluated two different envelope shapes with sharp or shallow attack times and identical magnitude spectra.

B. Methods

The same 10 listeners who participated in Experiment 1 participated in Experiment 2, using the same equipment and procedures. The only exception was that sawtooth waves were used as second-order modulators instead of sine waves. Two envelope shapes were tested. The first used a normal sawtooth wave with a sloping onset, referred to throughout the rest of the present manuscript as a "ramped" envelope. The second condition was a time-reversed version of the first, such that the onset of each cycle had a sharp attack with a sloping offset, and was referred to as "damped." These naming conventions were adopted from Dietz and colleagues [292]. Time-reversal of one signal ensured that the amplitude spectrum of both stimuli would be identical, and that only the phase spectrum would differ, ruling out the potential for spectral cues to explain the difference between envelope conditions. Testing consisted of 20 repetitions × 9

ITDs × 3 AM depths × 2 envelope shapes = 1080 trials, and took place over approximately four hours.

C. Results

1. Behavioral results

Results averaged across listeners are shown Fig. 5, compared against predictions from the Akeroyd (Fig. 5(A)) and Klug (Fig. 5(B)) model implementations. Individual lateralization functions were highly variable. ITD significantly impacted lateralization response [F(8,10917) = 43.331, p < .0001], with significant differences between all possible pairwise combinations of ITDs (p < .0001). Results in Fig. 5 suggest no consistent bias toward the left or right side with interaurally symmetric or asymmetric AM depth. However, there was a significant difference between AM depth [F(2,10918) = 30.715, p < .0001] and envelope shape [F(1,10918) = 100.598, p < .0001]. There was also a significant AM depth × envelope shape interaction [F(2,10918) = 43.868, p < .0001]. Post-hoc comparisons revealed that, for ramped stimuli, there were not significant differences between 50%:50% and 20%:50% [mean difference: -0.232; z = -2.042, p = .318], 50%:50% and 20%:20% [mean difference: -0.229; z = -1.996, p = .345], or 20%:50% and 20%:20% [mean difference: -0.003; z = -0.023, p = 1.000] AM depth. For damped stimuli, there were significant differences between 50%:50% and 20%:50% [mean difference: -1.221; z = -10.406, p < .0001] and 20%:50% and 20%:20% [mean difference: 1.154; z = 9.698, p < .0001], but not 50%:50% and 20%:20% [mean difference: -0.067; z = 0.562, p = .993] AM depth. This indicated that damped stimuli with 50%:50% and 20%:50% AM depth tended to be lateralized slightly to the left and right, respectively. There were also a significant ITD × AM depth [F(16,10917) = 6.335, p < .0001] and ITD × envelope shape [F(8,10917) = 8.782, p < .0001] interactions, as well as a significant ITD × AM depth × envelope shape interaction [F(16,10917)

= 10.534, p < .0001]. This result suggests that the lateralization function differed for AM depth conditions and envelope shape. Based on Fig. 5, this result was driven by a flattening of the lateralization function at lower AM depths for damped stimuli and a flat function in all AM depth conditions for ramped stimuli, which will be explored further in the analysis of lateralization range.

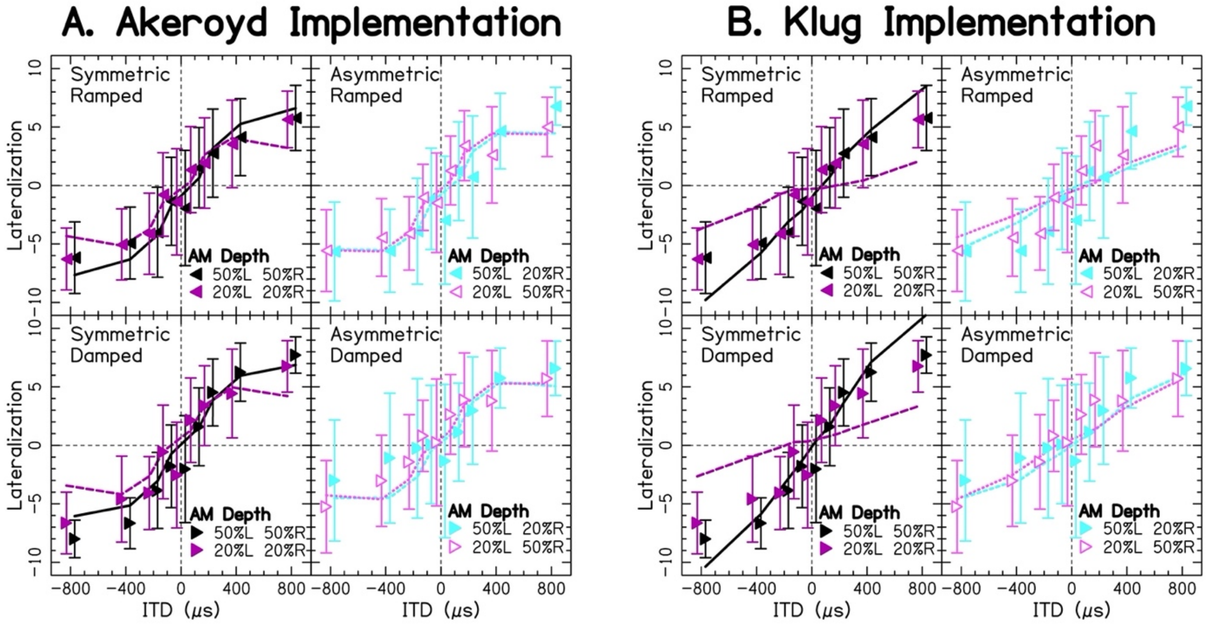


FIG. 5. Lateralization of ITDs by envelope shape and AM depth. Rows correspond to envelope shape. Columns correspond to symmetric and asymmetric AM depth conditions. The x-axis corresponds to the ITD presented on each trial. The y-axis corresponds to the extent of lateralization. Points show the mean across listeners ± one standard deviation. (A) Curves represent predictions from the Akeroyd model implementation. (B) Curves represent predictions from the Klug model implementation.

Lateralization range is shown in Fig. 6. Average results are shown in Fig. 6(A), and individuals are shown in Fig. 6(B). Model predictions are shown in Fig. 6(C). There was a significant effect of AM depth [F(2,45) = 8.161, p < .001], but not envelope type [F(1,45) = 3.139, p = .083]. There was a significant AM depth × envelope shape interaction [F(2,45) = 5.262, p < .01]. Post-hoc comparisons revealed that, for ramped stimuli, there were no significant differences between AM depth conditions. For damped stimuli, there were significant differences between 50%:50% and 20%:50% [t(45) = 5.051, p < .001], but not 50%:50% and 20%:20% AM depth [t(45) = 2.269, p = .228] or 20%:50% and 20%:20% AM depth [t(45) = 2.870, p = .065].

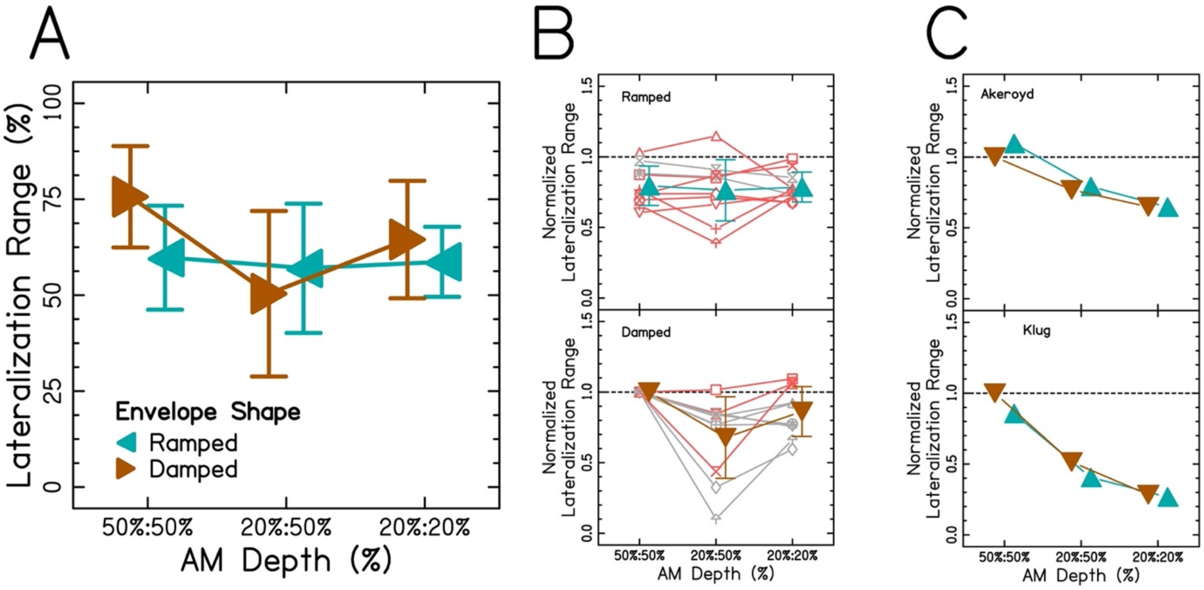


FIG. 6. Lateralization range by envelope shape and AM depth. The x-axis corresponds to the AM depth, and the y-axis corresponds to the lateralization range. (A) Average data ± one standard deviation. (B) Individual performance is shown offset in grey or red. Listeners shown in red had a lateralization range that was greater than that observed with 50%:50% for 20%:50% or 20%:20% AM depth (contrary to hypotheses). The top and bottom row show “ramped” and “damped” stimuli, respectively. Lateralization ranges were normalized to the damped envelope

at 50%:50% AM depth. (C) The top and bottom row show model predictions for the Akeroyd and Klug implementations, respectively. Lateralization ranges were normalized to the damped envelope at 50%:50% AM depth.

2. Modeling

The Akeroyd model implementation made highly accurate predictions of performance, especially for damped stimuli (adj. $R^2 = .885$; Fig. 5(A)), and slightly underestimated lateralization at the greatest values of ITDs for 20%:20% AM depth for both envelope shapes. The Klug model implementation shows slightly exaggerated effects of AM depth (adj. $R^2 = .743$; Fig. 5(B)), consistent with Experiment 1. Figure 6(C) shows that both model implementations predict effects of AM depth on the ramped stimuli, which is inconsistent with behavioral performance. Trends for AM depth with damped stimuli from both model implementations are consistent with behavioral performance, though the Klug model implementation shows exaggerated effects of AM depth.

D. Discussion

Experiment 2 investigated the interactions between interaurally symmetric or asymmetric temporal fidelity (i.e., AM depth) and sharp (damped) vs. shallow (ramped) attack times in stimulus envelopes for high-rate, amplitude-modulated acoustic pulse trains. We hypothesized that ramped stimuli would show greater reductions in lateralization range for the 50%:50% AM depth conditions relative to other AM depths. The results demonstrated an interaction between envelope shape and AM depth, in contrast to the hypothesis, where only stimuli with a sharp attack time (damped) demonstrated reduced lateralization range with decreasing AM depth in one ear. Ramped stimuli tended to show the smallest lateralization ranges, but there was no significant difference between envelope shapes.

Two studies investigated ITDs conveyed in amplitude-modulated, high-rate pulse trains for listeners with NH [321,323]. Stecker [321] used SAM wideband clicks at a rate of 500 pps, varied the ITD for each click in the stimulus, and asked listeners to report the perceived position of sounds. Then, he calculated a temporal weighting function (i.e., regressed the perceived location against the ITD for each click in the stimulus) to determine which clicks in the stimuli influenced the listeners' perceived location. His results showed that clicks at the onset were most highly weighted, regardless of the rate of sinusoidal AM (31.25, 62.5, and 125 Hz or constant-amplitude). Hu and colleagues measured ITD discrimination for 200 or 600 pps, band-limited acoustic pulses with a center frequency of 0.2, 0.6, or 4 kHz and sinusoidal AM [323]. The AM rates used in their experiment were 2 or 8 Hz. Their results demonstrated similar discrimination accuracy for onset and peak of the stimulus for the 200-pps pulse train, whether ITDs were presented via fine-structure or envelope. For the 600-pps pulse train, discrimination performance was more accurate for ITDs in the onset of the sound. Broadly speaking, both studies are in agreement with the psychophysical literature demonstrating that high-rate (~500 pps), constant-amplitude acoustic pulse trains result in a weighting of the onset of a sound, e.g., [344]. The results from Stecker [321] suggest that the slope of the onset of a click train should not affect onset weighting, and the results of Hu and colleagues [323] suggest that onset-weighting probably depends upon the pulse rate of the amplitude-modulated pulse train.

The present study is consistent with the notion that ITDs are sampled during transient increases in the temporal envelope for high-rate stimulation, when "enveloped-triggered looks" to ITDs are initiated ("restart" hypothesis; [321,322]). The lateralization range of 500-pps, constant-amplitude pulse train is similar to that of the shallow attack time (ramped) stimulus at all AM depths tested and the sharp attack time (damped) stimulus with smaller AM depth in one ear (see Appendix). Thus, the present study shows that lateralization of envelope ITDs using high-rate, amplitude-modulated pulse trains depends upon AM depth in either ear, and that the worse ear can limit lateralization performance, or the efficacy of envelope ITDs affecting perception. These results suggest that multiple “onsets” provided throughout a high-rate, amplitude-modulated pulse train can improve lateralization, but only if they are sufficiently large in AM depth and interaurally symmetric.

Both model implementations made characteristic errors associated with envelope type and AM depth: neither model accounted for their interaction. In other words, both models overestimated the contributions of ITDs in the stimulus offset. This is an important finding since, to our knowledge, no previous modeling efforts have been made to capture the effects of envelope shape on lateralization of high-rate, amplitude modulated pulse trains. Previous models have accurately predicted the relationship between ITD lateralization and envelope shape for AM applied to tone bursts, including the same Klug model implementation used here [293,329]. Interestingly, additional models have been used to explain the effects of sharp stimulus onsets or offsets in cells of the inferior colliculus, demonstrating differential effects of excitation and inhibition depending upon the envelope shape [244]. It may be that a similar approach is required for high-rate, amplitude-modulated pulse trains.

IV. EXPERIMENT 3: INTERACTIONS WITH PLACE-OF-STIMULATION MISMATCH

A. Motivation

The goal of Experiment 3 was to explore whether a different source of interaural asymmetry, place-of-stimulation mismatch, interacts with interaurally asymmetric temporal fidelity in lateralization of envelope ITDs. Interaural place-of-stimulation mismatch presumably decreases the number of input fibers from each ear that can be compared to compute an ITD. Decreasing AM depth in one or both ears is thought to introduce less consistent temporal responses of input fibers, simulating interaurally asymmetric temporal fidelity. Thus, the binaural system may have fewer “looks” at envelope ITDs when there is greater interaural place-of-stimulation mismatch (i.e., fewer neural populations overlapping in characteristic frequency) and increased stochasticity in the system due to poorer temporal fidelity (i.e., less temporal synchrony of inputs). Because interaural place-of-stimulation mismatch is prevalent in patients with BiCIs and due to surgical limitations rather than hearing history [43], it is highly likely that patients with BiCIs who experience interaurally asymmetric temporal fidelity would also be afflicted by place-of-stimulation mismatch.

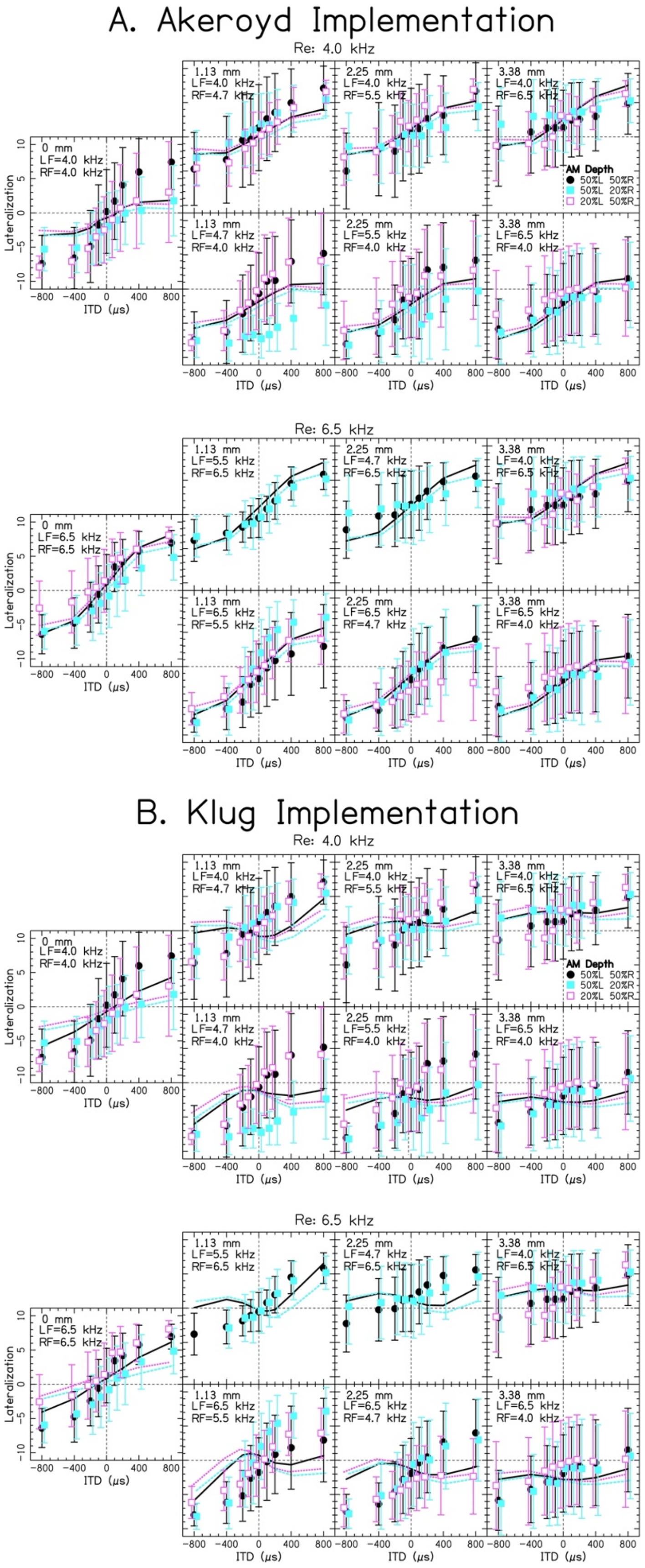
B. Methods

Sixteen new listeners participated in Experiment 3 (age 18-26; mean age: 21.3 years). All met the same audiological criteria as Experiment 1, using the same equipment and procedures. The only exceptions were that the carrier frequency of the Gaussian-enveloped tone was varied such that it resulted in differing degrees of interaural place-of-stimulation mismatch (0, 1.13, 2.25, or 3.38 mm) based on the frequency-to-place map from [336]. The resulting center frequencies were 4.00, 4.68, 5.51, and 6.46 kHz. The ear and direction of mismatch (by shifting up or down in frequency) was counterbalanced across listeners. Because of the additional number of combinations based on experimental variables, only 50%:50% and 20%:50% AM depth were tested. Testing consisted of 20 repetitions × 9 ITDs × 2 AM depths × 4 mismatches = 1440 trials, and took place over approximately six hours.

B. Results

1. Behavioral results

Results averaged across listeners are shown Fig. 7, compared against predictions from the Akeroyd (Fig. 7(A)) and Klug (Fig. 7(B)) model implementations. ITD significantly impacted lateralization response [F(8,23403) = 1121.769, p < .0001], with significant differences between all possible pairwise combinations of ITDs (p < .0001). Results in Fig. 7 demonstrate no consistent bias toward the left or right side with interaurally symmetric or asymmetric AM depth, but do show a bias toward the ear with the higher center frequency. There was no significant difference between AM depths [F(8,23403) = 3.083, p = .079], but there was a significant effect of place-of-stimulation mismatch [F(3,23403) = 11.879, p < .0001]. There was also a significant AM depth × place-of-stimulation mismatch interaction [F(3,23403) = 58.396, p < .0001]. Post-hoc comparisons revealed that there were significant differences between 50%:50% and 20%:50% for 0-mm mismatch [mean difference: 1.408; z = 11.222, p < .0001] and 3.38-mm mismatch [mean difference: -0.813; z = -6.519, p < .0001], but not 1.13-mm [mean difference: -0.346; z = -2.718, p = .117] or 2.25-mm mismatch [mean difference: 0.195; z = 1.529, p = .792]. This result indicates that stimuli with 50%:50% AM depth tended to be lateralized toward the right for 0-mm mismatch, and toward the left for 3.38-mm mismatch. There were also a significant ITD × AM depth [F(8,23403) = 20.558, p < .0001] and ITD × place-of-stimulation mismatch [F(24,23403) = 37.105, p < .0001] interactions, as well as a significant ITD × AM depth × place-of-stimulation interaction [$F(24,23403) = 3.850$, $p < .0001$]. This result suggests that the lateralization function differed for AM depth conditions and interaural place-of-stimulation mismatch. Based on Fig. 7, this result was driven by a flattening of the lateralization function for 20%:50% compared to 50%:50% AM depth, and a bias toward the ear with higher center frequency and flattening of the lateralization function for increasing amounts of interaural place-of-stimulation mismatch, which will be explored further in the analysis of lateralization range.



A. Akeroyd Implementation

B. Klug Implementation

FIG. 7. Lateralization of ITDs by interaural place-of-stimulation mismatch and AM depth. The top and bottom group of figures show lateralization relative to 4 and 6.5 kHz center frequency, respectively. Rows correspond to a higher frequency presented to the right or left ear, respectively. Columns correspond to interaural mismatch. The x-axis corresponds to the ITD presented on each trial. The y-axis corresponds to the extent of lateralization. Points represent the mean were averaged across listeners ± one standard deviation. (A) Curves represent predictions from the Akeroyd model implementation. (B) Curves represent predictions from the Klug model implementation.

Lateralization range is shown in Fig. 8. Average results are shown in Fig. 8(A), and individuals are shown in Fig. 8(B). Model predictions are shown in Fig. 8(C). There was a significant effect of AM depth [$F(1,105) = 21.626$, $p < .0001$] and place-of-stimulation mismatch [$F(3,105) = 32.818$, $p < .0001$]. There was not a significant AM depth × interaural place-of-stimulation mismatch interaction [$F(3,105) = 2.209$, $p = .091$]. Post-hoc comparisons revealed that there were significant differences between 0-mm and 3.38-mm [$t(105) = 9.207$, $p < .0001$], 0-mm and 2.25-mm [$t(105) = 4.239$, $p < .001$], 1.13-mm and 3.38-mm [$t(105) = 7.726$, $p < .0001$], 1.13-mm and 2.25-mm [$t(105) = 2.758$, $p < .05$], and 2.25-mm and 3.38-mm [$t(105) = 4.968$, $p < .0001$], but not 0-mm and 1.13-mm mismatch [$t(105) = 1.481$, $p = .452$]. For all but the smallest amounts of interaural place-of-stimulation mismatch, the lateralization range decreased with increasing mismatch.

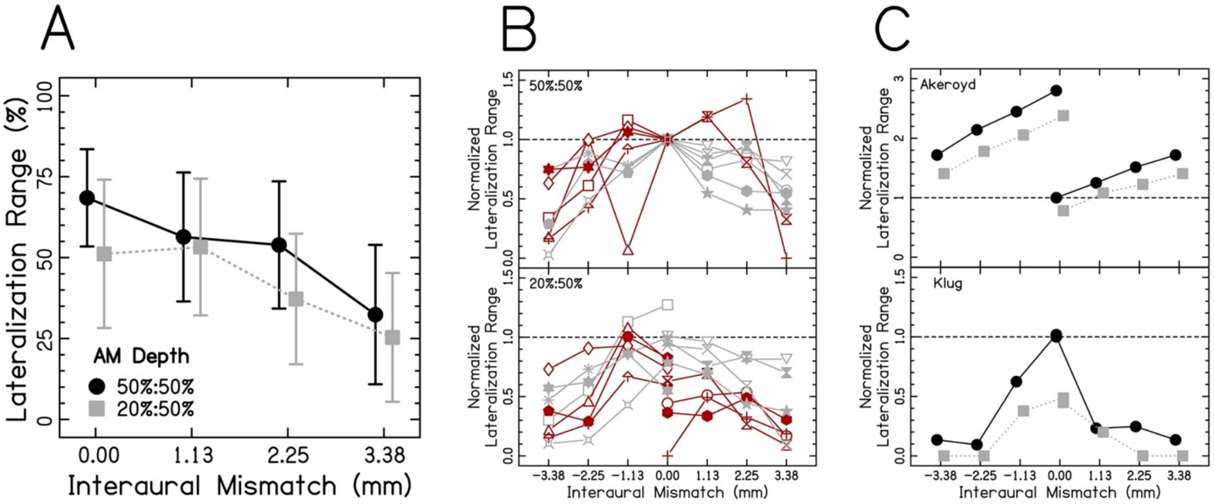


FIG. 8. Lateralization range by interaural place-of-stimulation mismatch and AM depth. The x-axis corresponds to interaural place-of-stimulation mismatch, and the y-axis corresponds to the lateralization range. (A) Average data ±1 standard deviation. (B) Individual performance is shown offset in grey or red. Listeners shown in red had a lateralization range that was greater than that observed with 0-mm for 1.13-, 2.25-, or 3.38-mm interaural place-of-stimulation mismatch (contrary to hypotheses). Lateralization ranges were normalized to 0-mm mismatch at 50%:50% AM depth. The top and bottom row show 50%:50% and 20%:50% AM depth, respectively. (C) The top and bottom row show model predictions for the Akeroyd and Klug implementations, respectively. Lateralization ranges were normalized to 0-mm mismatch at 50%:50% AM depth.

2. Modeling

The Akeroyd model implementation tended to underestimate lateralization (adj. R^2 = .743; Fig. 7). The Klug model implementation showed non-monotonicities of lateralization with varying ITD (adj. R^2 = .289). In contrast, the Klug model implementation accurately predicted the directions of the effects of AM depth and interaural place-of-stimulation mismatch on lateralization range (Fig. 8). The Akeroyd model implementation predicted increased lateralization range when one, and especially both, ears received stimuli with a center frequency of 6.5 kHz.

C. Discussion

Experiment 3 investigated the impact of interaural place-of-stimulation mismatch and temporal fidelity on envelope ITD lateralization. We hypothesized that there would be an interaction between these factors, with the greatest amounts of interaural place-of-stimulation mismatch yielding the greatest reductions in lateralization range associated with smaller AM depth in one ear. The results demonstrated no significant interaction, and instead showed that effects of interaural place-of-stimulation mismatch and AM depth on lateralization were additive. Lateralization was extremely poor when interaural place-of-stimulation was maximal and temporal fidelity was reduced in one ear. Thus, these findings are still consistent with the binaural system having fewer "looks" at noisier envelope ITDs with interaural place-of-stimulation mismatch and poor temporal fidelity in one or both ears. Critically, when interaural place-of-stimulation mismatch and reduced temporal fidelity co-occur, lateralization ranges decreased below the floor of performance relative to the same place-of-stimulation in both ears (see Appendix). Thus, place-of-stimulation seems to be the more greatly limiting factor, since envelope ITDs from extremely disparate places-of-stimulation may not be useful for lateralization [138,139,144,163].

Previous experiments investigating interaural place-of-stimulation mismatch have focused primarily on constant-amplitude pulse trains [138,139,144,163] and SAM tones [325–328]. The present study was the first to investigate high-rate, amplitude-modulated pulse trains.

Our results are consistent with those from previous studies, suggesting that the utility of ITDs decreases with increasing mismatch. Lateralization also tended to be biased toward the ear with higher frequency (Fig. 7), consistent with previous reports [138,163]. The present study shows that bias associated with place-of-stimulation mismatch is still present with in high-rate, amplitude modulated pulse trains, regardless of whether the AM depth is interaurally symmetric.

Both model implementations poorly predicted the effects of interaural mismatch for lateralization (Akeroyd adjusted $R^2 = .743$; Klug adjusted $R^2 = .289$). However, the trends in lateralization range with AM depth were consistent across both models. The Klug model implementation accurately predicted trends in lateralization range and the additive effects of AM depth, just to an exaggerated degree. Strangely, the Akeroyd model implementation predicted larger lateralization ranges for higher center frequencies, which this contrasts with predictions of ITD sensitivity with similar model implementations [61]. This effect did not appear to depend upon AM depth.

V. GENERAL DISCUSSION

1. Interaural asymmetries in temporal encoding

The experiments in the present manuscript investigated the relationship between interaurally asymmetric temporal fidelity in each ear with envelope shape and another source of interaural asymmetry: place-of-stimulation mismatch. Recent studies in listeners with BiCIs show that have one ear that was left unstimulated for long periods of time have poorer speech understanding [13], poorer localization performance [345], and poorer sensitivity to binaural cues [105]. Additionally, many listeners show large differences in each ear's ability to understand speech [10,13–15]. These factors have been associated with poorer speech understanding outcomes in listeners with BiCIs, those who use one CI and have NH in the other ear, those who use hearing aids, and simulations in NH [10,13–15,23–26]. These differences between the ears have been related to especially poor and/or abnormal speech understanding when both ears are stimulated simultaneously. Together, these and other studies [9,273] suggest that the poorer ear limits binaural outcomes.

It is difficult to define a strict set of criteria for interaural asymmetry. Anecdotally, many listeners with hearing loss report having a preferred ear, and much recent evidence shows large differences between ears. Approximately 37% of individuals with hearing loss have asymmetric hearing thresholds [1]. Electrical stimulation, compared to acoustic stimulation, leads to highly phase locked auditory nerve responses [213], suggesting that any comparisons of acoustically and electrically represented sounds would result in asymmetries in temporal coding. An increasing population of listeners is using a mixture of acoustic and electric stimulation. Recent studies have shown that receiving a CI in cases of unilateral deafness with NH in the other ear improves speech-in-noise outcomes [14,346], but only when target speech is in the better ear [23]. Listeners with a hearing aid in one ear and a CI in the other ear (i.e., bimodal hearing) also do better in with speech-in-noise tasks relative to the hearing aid alone [347,348]. While the focus of the present manuscript has been on interaural asymmetries, studies of the electrode-neuron interface assume that some places-of-stimulation yield the best spectro-temporal resolution, suggesting that the spectro-temporal fidelity changes within the same ear. Similarly, some listeners integrate acoustic and electric stimulation within-and across-ears. Efforts have been made to preserve low-frequency acoustic hearing for some listeners with combined electric-acoustic stimulation, demonstrating speech-in-noise benefits related to ITD sensitivity relative to one ear or using only hearing aids [215,349]. Thus, interaural asymmetry may be a pervasive problem for listeners with hearing loss, both with respect to the mode of stimulation and the outcomes in each ear. More concerning is that much recent evidence implies that interaural asymmetries limit binaural outcomes, and much of the conceptual and computational models associated with binaural processing have assumed interaural symmetry. At the very least, the present experiment demonstrates some ways that different kinds of interaural asymmetry can interact with other factors that affect envelope ITD lateralization and contribute to variability in performance for which may not be presently accounted.

B. Relevance to listeners with BiCIs

The present experiment simulated interaural asymmetry in temporal representations by manipulating the AM depth of a second-order modulator. This was meant to simulate poorer phase locking (Fig. 2) due to changes in the auditory periphery for listeners with BiCIs (Fig. 1), especially those who experience prolonged auditory deprivation. Small AM depth is analogous to reduced dynamic range at the level of the auditory nerve. Dynamic range as measured at the level of CI programming software is one psychophysical predictor for the electrode-neuron interface for listeners [123,125]. However, the "true" dynamic range of the auditory nerve (e.g., the relationship between input current to the neuron and spike rate) is not known for patients. Both increased distance of CI electrodes from the auditory nerve and peripheral deterioration are suspected to affect the dynamic range of auditory nerve fibers as measured at the level of CI programming software [115,125]. Binaural masking level difference and ITD discrimination thresholds both decrease with increasing dynamic range [121]. Similarly, increased AM depth for SAM pulse trains results in greater ITD sensitivity in listeners with BiCIs [122]. Thus, one implication of the results in the present study are that reduced dynamic range limits binaural processing.

The pulse rates used in the present study were lower than the clinical standard for listeners using most contemporary CI processors from Cochlear and Advanced Bionics, which tend to use pulse rates closer to 1000 pps. Cochlear implant processors do not preserve temporal-fine structure, so ITDs available through clinical processors are in the temporal envelope [36]. To our knowledge, very little research has used band-limited, high-rate, amplitude-modulated, acoustic pulse trains. This may have to do with the upper-limit of pulses that can be represented without temporal overlap [163]. That is, we chose to use 500 pps because that was near the maximum number of pulses with 3-mm bandwidth that could be presented without resulting in temporal overlap between each pulse.

Previous studies of envelope shape have demonstrated differences between listeners with NH and BiCIs. In particular, listeners with NH seem to be sensitive to the attack time of AM [242,243,245,292,293,314], where listeners with BiCIs show no difference between sharp and shallower attack times for monaural [241] or binaural tasks [245]. One important confound in those studies was that listeners with BiCIs were presented with amplitude-modulated pulse trains (i.e., their stimuli had pulsatile carriers), where listeners with NH were presented with stimuli that had tonal carriers. This sparse sampling of the envelope may have played a role in the lack of effect of envelope shape on BiCI performance. Figure 9 shows the lateralization ranges of listeners that participated in in the present experiment with a 500-pps acoustic pulse train carrier compared to previously published data with SAM tones [46]. Both studies used 100-Hz AM rates. The present study used 50% AM depth in both ears whereas Anderson and colleagues [46] used 100% AM depth. The data show extraordinarily similar distributions and no difference between groups according to a t-test [$t(36) = 0.338$, $p = .971$]. Data using interaurally symmetric AM suggest that there would be a considerable reduction in lateralization range for 100%

compared to 50% AM depth [314], which implies that pulsatile stimulation would improve lateralization of ITDs, or that AM depths above 50% provide no additional benefit. Two additional results from the present study shed important light on this confound between NH and BiCI studies. Firstly, by manipulating the depth of the second-order modulator, lateralization range decreased for listeners with NH in all three experiments. Reducing AM depth also decreases the slope, or smooths the stimulus attack. Secondly, listeners with NH had greater lateralization range for stimuli with a sharp attack in both ears. Together, these results imply that pulsatile stimulation is not sufficient to explain differences between listeners with BiCIs and NH. Thus, these differences may have to do with the high degree of phase locking characteristic of electrical stimulation [294,295], the older average age of participants in studies with BiCIs compared to NH [46,162], the nature of listeners' hearing loss, or a combination of all of these factors.

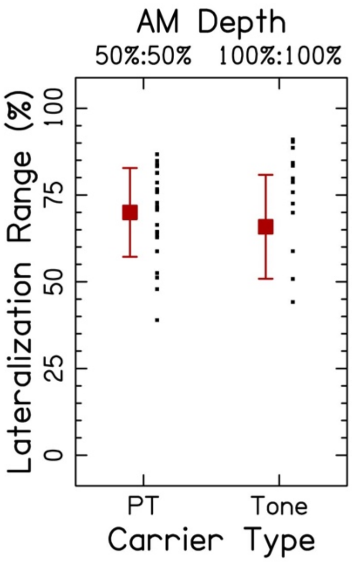


FIG. 9. Lateralization range by carrier type and AM depth. Data for SAM tones with 100%:100% AM depth reprinted from Anderson and colleagues [46] with permission. The x-axis

corresponds to the carrier, either a pulse train (PT) or 4-kHz sine tone, as well as the AM depth. The y-axis corresponds to the lateralization range.

For SAM pulse trains, on the other hand, listeners with NH and BiCIs show different utility of ITDs conveyed in different portions of the stimulus [323]. While listeners with NH weight the onset of high-rate, SAM pulse trains, listeners with BiCIs may weight the peak more heavily [323]. Interestingly, listeners with NH and BiCIs show very similar relationships between rate of sinusoidal AM rate and lateralization range [46]. The authors attributed this result to changes in the slope of AM, with 100-Hz AM rates yielding the largest lateralization range [46]. Because Hu and colleagues [323] used very low AM rates (2 or 8 Hz), it remains to be determined whether multiple "onsets" in high-rate, amplitude-modulated pulse trains are similarly useful for listeners with BiCIs.

The data modeled in the present study involved acoustic stimulation, and if a similar approach is to be used in future studies of listeners with BiCIs, it needs to be compatible with electrical stimulation. The Klug model implementation has two stages. The first is auditory nerve processing and the second is LSO processing. They used the Bruce model [332] for the auditory nerve input to the LSO, but this could simply be replaced by an auditory nerve model for CI stimulation. The Akeroyd model implementation includes peripheral processing stages, but it is not obvious how these could be replaced by a CI-like representation. Thus, the Klug model implementation may be more easily adapted to predict data from listeners with BiCIs.

One serious limitation of the Akeroyd model implementation is that it is largely level-invariant. However, listeners with NH show improved sensitivity to ITDs at higher levels [350]. Most studies of binaural processing in listeners with NH that measure ITD sensitivity to stimuli with AM rates near 100-Hz tend to show greater sensitivity with 4-kHz compared to higher carrier frequencies [61]. One explanation for this effect is differences in the modulation filter associated with different characteristic frequencies throughout the auditory system [58]. Listeners with BiCIs, on the other hand, do not show any consistent differences in ITD sensitivity with electrode location [44,45,105]. This could be due to any number of factors, including the nature of electrical stimulation or the etiology of deafness. A major difference between studies in listeners with BiCIs and those with NH is that loudness balancing is completed for most listeners with BiCIs prior to testing. Accordingly, a recent study showed no difference in ITD discrimination when the same sensation level was used across center frequencies for listeners with NH [351]. This effect could not be explained by a model that is level or loudness invariant. Interestingly, with the stimuli used in the present experiment, the Akeroyd model implementation predicted smaller lateralization ranges for higher carrier frequencies. This is inconsistent to similar models[2], which predict improved ITD sensitivity for frequencies closer to 4 kHz compared to those above 4 kHz [61].

C. Assumptions of symmetry in binaural models

In the present experiment, we compared the performance of two highly relevant models of binaural processing. The Akeroyd model implementation predicts lateralization based upon frequency- and delay-weighted cross-correlation, similar in spirit to the type of processing completed in the MSO. The model utilized by Klug and colleagues [329] has also been used to predict cellular output of the LSO [343]. Thus, both models are reflective of the types of processing that occur in the brainstem to code for envelope ITDs. Moreover, similar models have demonstrated highly accurate predictions of lateralization of constant-amplitude high-frequency acoustic pulse trains in previous studies [329,340].

The findings in this paper indicate that both the Akeroyd and Klug model implementations were well predictive of behavioral performance with these novel stimuli. The adjusted R^2 value for each experiment and model implementation is summarized in Table II. Both model implementations had different strengths, particularly in cases of interaural place-of-stimulation mismatch. That is, while the Akeroyd model implementation tended to do a better job of predicting mean lateralization performance across listeners, the Klug model implementation tended to demonstrate trends in lateralization ranges more closely corresponding to individual performance. Both models made erroneous predictions related to envelope shape; namely greater lateralization ranges than were observed for 20%:50% compared to 20%:20% AM depth. It is also worth mentioning that the Klug model implementation is based on stochastic sampling and requires considerably longer, more intensive computations to generate output. The Akeroyd model implementation is deterministic and highly computationally efficient.

TABLE II. Adjusted R^2 values for each model implementation by experiment.

	Adjusted R^2	
	Akeroyd Model Implementation	**Klug Model Implementation**
Experiment 1	0.949	0.786
Experiment 2	0.885	0.743
Experiment 3	0.743	0.289

For the sake of parsimony, most computational and statistical models attempt to minimize the number of parameters used. Thus, one strength of both of the model implementations used in the present experiment is their simplicity and reduced computational intensity relative to other more extensive biophysical models (for a review in the LSO, see [352]). The question remains of whether special considerations are necessary for interaural asymmetry. Certainly in the case of place-of-stimulation mismatch model performance could be improved, though it is not obvious what steps would be necessary to correct for this. Neither model implementation had been used to explicitly predict behavioral results involving sources of interaural asymmetry. The results from the present chapter demonstrate promise for both model implementations and underscore an important dimension on which binaural models should be validated: interaural asymmetry.

D. Limitations

The present study had several limitations that should be discussed. The sample sizes used here were relatively small, especially in light of the variability across listeners. Variability in performance is commonly observed in lateralization tasks using the same procedure as that used in the present experiment, even in the presence of normal and symmetric hearing thresholds [337]. Most results presented here are consistent with previously published experiments in listeners with NH and were discussed in more detail in the discussion section following each experiment. Even with only four listeners, as shown in the Appendix (Fig. 11(B)), across-group trends (e.g., effects of pulse rate on lateralization of constant-amplitude pulse trains) were consistent with previous studies [61,275].

It is difficult to directly generalize the results from the present experiment to patients with BiCIs. Participants in the present study were much younger on average than those tested in studies of listeners with BiCIs and aging is associated with smaller lateralization range [46,162]. While we focused primarily on simulating changes in the peripheral representations of sounds, these are likely accompanied by changes in the central auditory system of listeners with hearing loss. Finally, electrical stimulation via CI is suspected to result in highly synchronous auditory nerve firing [294,295]. It is unclear whether limitations to temporal encoding occur at the level of individual or population of nerve fibers in patients with BiCIs.

The present study used whole-waveform ITDs, which are not consistent with those preserved by BiCI output [36]. Other options would have been to jitter ITDs in the pulse train or fix the ITD of the constant-amplitude pulse train at 0 μs and only apply an ITD to the envelope. In order to attain a high pulse rate with 3-mm bandwidth acoustic pulse trains, there was very little time between pulses. It was therefore not possible to jitter ITDs with 500 pps acoustic pulse trains, so we prioritized using high pulse rates. Fixing ITDs of the constant-amplitude pulse train at 0 μs would artificially bias lateralization toward the center of the head and it would not be any more realistic than conveying coherent ITDs in the temporal fine-structure, since these temporal fine-structure ITDs show random bias in BiCI output [36]. From Fig. 11 in the Appendix, it is clear that listeners were still able to use ITDs from the pulsatile carrier to complete lateralization and attain a lateralization range of approximately 40%. Thus, our approach of applying whole-waveform ITDs represents a best case scenario for BiCI output.

We chose to simulate poorer temporal fidelity by reducing AM depth (akin to dynamic range for listeners with CIs). Even if our manipulation is not representative of changes in temporal representations due to phase locking of the auditory nerve, the findings are in agreement with studies in listeners with BiCIs showing reduced binaural sensitivity at lower AM depths [122] and smaller dynamic range in current levels at each electrode [121].

E. Summary and conclusions

Together from Experiment 1, 2, and 3 several conclusions can be drawn with respect to lateralization of envelope ITDs under conditions of interaural asymmetry:

1. Lateralization range associated with envelope ITDs was significantly smaller when one or both ears have reduced temporal fidelity (i.e., smaller AM depth; Fig. 4, 6, 8).
2. Lateralization of high-rate, amplitude-modulated pulse trains was facilitated by interaurally symmetric, sharp temporal onsets, and inhibited by sloping temporal onsets, in the temporal envelope (Fig. 6). This finding is consistent with weighting of the onset for high-rate, amplitude-modulated pulse trains in listeners with NH [321,323].
3. Limitations to lateralization from interaural place-of-stimulation mismatch and reduced AM depth were additive (Fig. 8).
4. Cross-correlation (i.e., EE-based) and EI-based models of binaural processing predicted lateralization of envelope ITDs for interaurally symmetric smaller AM depths, but tend to overestimate lateralization range for interaurally asymmetric AM depths (Fig. 4, 6, 8).
5. Cross-correlation (i.e., EE-based) and EI-based models of binaural processing did not accurately predict effects of envelope shape for high-rate, amplitude-modulated pulse trains (Fig. 6), predicting similar performance for stimuli with a temporally sharp onset or offset.
6. Cross-correlation (i.e., EE-based) and EI-based models of binaural processing poorly predicted effects of interaural place-of-stimulation mismatch, though the EI-based model accurately predicted trends in lateralization range.

VI. APPENDIX

It was important to characterize the floor of lateralization performance given that the pulse rate used in the present experiment was higher than that used in previous NH experiments and lower than that typical of CI stimulation. Much previous research has been dedicated to the "rate limitation" to binaural processing in listeners with NH (for review, see [61]) and BiCIs (for review, see [44,45]). Most of this research agrees with the consensus that the "rate limitation" is somewhere near 300 pps, meaning that sensitivity to ITDs decreases precipitously above this rate. However, some listeners with BiCIs maintain sensitivity to ITDs in constant-amplitude stimuli at very high pulse rates [61,275]. Thus, we tested a small group of listeners with NH using 50%:50%, 20%:20%, and 0%:0% AM depth to determine whether the lateralization range would decrease if stimuli were constant-amplitude (i.e., if AM depth was further reduced to 0%:0%).

We included ITDs of 0, ±100, ±200, ±400, ±800, ±1200, and ±1600 μs. The ±1200 and ±1600 μs ITDs would result in an interaural phase difference greater than π for the pulsatile carrier but less than π for the sinusoidal AM. Thus, including ±1200 and ±1600 μs ITDs would help determine whether the listeners attended to the pulsatile carrier or ITDs in other portions of the stimulus (i.e., at stimulus onset or ongoing in the envelope) to complete the task. If listeners attended to the pulsatile carrier, then the perceived intracranial location should move back toward the center of the head for ±1200 and ±1600 μs. If not, then the magnitude of lateralization should be equal or greater than the maximum magnitude of lateralization for ±1200 and ±1600 μs. Lateralization range was still assessed as the absolute difference in lateralization values from the fitted curves at ±800 μs.

Four listeners (age 20-24; mean age: 21.8 years) met the same audiological criteria as the experiments and the same equipment and procedures were used. Twenty repetitions per ITD were collected except for listener TZA, who only completed 10 repetitions per ITD for the 50%:50% and 0%:0% AM depth conditions because of time limitations. This resulted in 20 repetitions × 11 ITDs × 3 AM depths = 660 trials, which took approximately 1.5 hours to collect.

Individual lateralization curves are shown for each listener in Fig. 10. The results are similar and consistent with those in Experiment 1 for the 50%:50% and 20%:20% AM depth conditions. Fig 10 also demonstrates that ±1200 and ±1600 μs ITDs were lateralized to a similar position as ±800 μs ITDs. Lateralization ranges for differing AM depths are presented in Fig. 11(A). They show that lateralization range decreased slightly on average between 20%:20% and 0%:0% AM depth. This effect was driven by three of four listeners. Fig. 11(B) shows data from Anderson and colleagues [46] who tested lateralization of constant-amplitude acoustic pulse trains between 10 and 300 pps using the same interface and very similar stimuli as the present experiment. The lateralization ranges in Fig. 11(B) were all computed for ±1000 μs ITDs to be consistent with the previous manuscript, in contrast to the rest of the present manuscript where they were computed at ±800 μs. The data for 500 pps are from the listeners in the present experiment and demonstrate a consistent trend with the previously published data. This trend is in agreement with studies of ITD sensitivity using high-rate, high-frequency trains of transients [61,275].

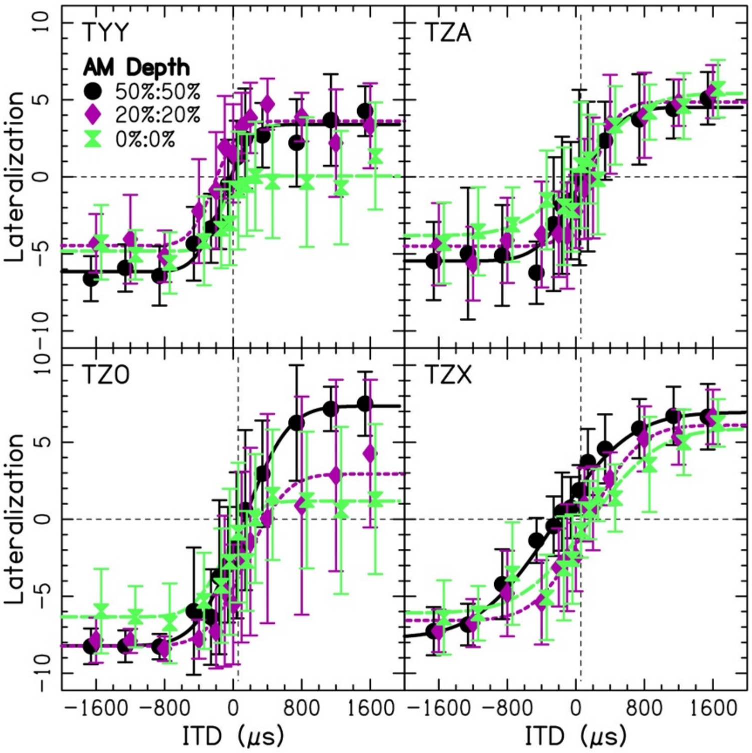


FIG. 10. Individual lateralization performance for varying, symmetric AM depth. Each panel corresponds to a listener whose subject code is given in the top-left corner. The x-axis corresponds to the ITD presented on each trial. The y-axis corresponds to the extent of lateralization, with mean ± 1 standard deviation shown by each point. Color indicates the AM depth condition.

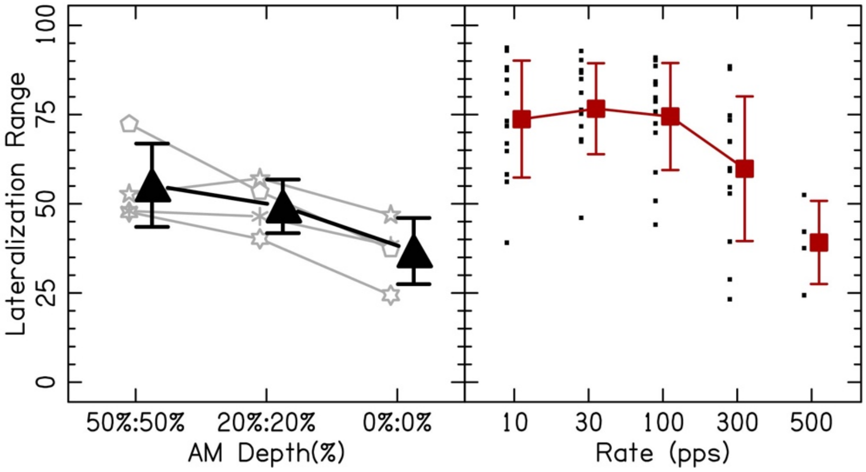


FIG. 11. Lateralization range. The y-axis corresponds to the lateralization range. Mean data ±1 standard deviation are by larger symbols offset to the right. Individual performance is shown offset in smaller symbols offset to the left. (A) The x-axis corresponds to the AM depth of amplitude-modulated pulse trains. (B) The x-axis corresponds to the pulse rate of constant-amplitude pulse trains.

The results presented in this appendix suggest two important pieces of information about how listeners completed the lateralization experiment. Firstly, because lateralization of ±1200 and ±1600 continued to be of large magnitude and was not toward the center of the head, it suggests that listeners relied on the onset, offset, or ongoing temporal envelope (i.e., second-order AM) of the sound for lateralization rather than ongoing ITDs in the pulse train. Results from the 0%:0% AM depth (i.e., constant-amplitude) in particular suggest that listeners may have been able to resolve interaural phase ambiguity using the ITDs at the onset and offset of stimuli. Secondly, these results suggest that lateralization range of 20%:20% AM depth was near but not at the floor of performance.

FOOTNOTES

1. The code for the Akeroyd model implementation is available via the following DOI: http://doi.org/10.17639/nott.320.

2. Most behavioral and modeling studies of center frequency have focused on ITD sensitivity. One study has explored the effects of high center frequencies on lateralization of envelope ITDs [353]. They used a 200-Hz sinusoidal AM with center frequencies of 2.2-, 3.2-, 4.2-, and 5.2-kHz carriers and 100% AM depth, showing a slight decrease in lateralization range computed at ±600 μs and averaged across listeners.

Chapter 5: Asymmetric temporal envelope encoding: Lateralization and modeling in listeners with bilateral cochlear implants

I. INTRODUCTION

Bilateral cochlear implants (BiCIs) are provided to patients with bilateral severe to profound hearing loss to improve sound source localization and speech understanding in noise (e.g., [6,11,345,354]). These benefits are suspected to result in part from interaural time and level differences (ITDs and ILDs, respectively). Even in the laboratory under conditions that elicit optimal ILD and ITD sensitivity, patients with BiCIs tend to show large amounts of variability between individuals [44,45,105]. There is also a high degree of variability within the same individual. For example, when listeners are stimulated at different places along the auditory nerve, sensitivity to binaural cues differs.

One factor that predicts ITD sensitivity within the same individual is their sensitivity to temporal cues in either ear [9]. That is, the ear with least temporal sensitivity is predictive of the ITD sensitivity of the pair of stimulating electrodes. The goal of the present study was to extend on these previous results by testing stimuli that were more realistic to everyday listening, are conveyed by the CI processor, and are most likely to elicit differences between ears using a paradigm that assess the impact of ITDs on perceived sound location.

A. Binaural cues conveyed by BiCIs

Binaural cues can be conveyed in two different portions of a signal: the temporal fine-structure (rapid/instantaneous changes in intensity over time) or the temporal envelope (characteristic, more gradual changes in the intensity over time). A recent study measured the ILDs and ITDs output by BiCI processors using clinical software [36]. They found that processors significantly reduced the average ILD and the ILDs across frequency. The ILDs are expected to be further obscured by independently operating automatic gain control (i.e., compression algorithms [308,309]). Smaller ILD magnitudes and non-monotonicities are related to the location of microphones on CI processors [355].

Gray and colleagues also showed that the relationship between location in the horizontal plane and ILD was often non-monotonic, sometimes favoring the opposite ear, particularly for lower frequencies [36]. Higher frequency electrodes exhibited larger differences in CI output to ILDs. While ITDs were more consistent across frequency, fine-structure ITDs were smaller in magnitude and artificially large (i.e., non-zero for central speaker locations). Envelope ITDs on the other hand showed good correspondence between the left and right ear. Most sound processing algorithms discard temporal fine-structure information [35]. However, even when temporal fine structure is preserved, listeners with BiCIs tend not to show any improvement on sound source localization [37–39]. At the rates provided by clinical processors (e.g., ~1000 pulses per second; pps), listeners are not sensitive to ITDs [44,45].

CI arrays generally span about 1 to 1.5 turns of the cochlea beginning at the cochlear base [43,130], meaning that they stimulate auditory nerve fibers corresponding to higher frequencies. Because of surgical limitations, mismatches in the interaural place-of-stimulation often occur for electrodes corresponding to electrodes representing the frequency information in each ear [43]. Most clinical processors use monopolar stimulation and are suspected to result in approximately 3-4.5 mm of spread of current before reaching the auditory nerve fibers [356]. This might restore overlap in frequency channels, making it possible to use ITDs (e.g., [138,139]). However, this would also lead to “smearing” of spectro-temporal fluctuations, making it more difficult to distinguish between sound sources [357].

B. Binaural encoding with BiCIs

Classically speaking, ITD coding in mammals is thought to occur in the medial superior olive of the brainstem [315,333]. However, there are several reasons to suspect that binaural cues are coded by the lateral superior olive (LSO) based upon the ways that binaural cues are represented with BiCIs. Cells in the LSO are most responsive to high frequencies [55–57], corresponding to the auditory nerve fibers stimulated by CI electrodes. Cells in the LSO show a low-pass characteristic for amplitude-modulated stimuli, meaning that they provide good phase locking until around 500 Hz, at which point phase locking decreases precipitously with increasing rate [58–60]. Spike rates of LSO cells are greatest for ipsilaterally favoring ILDs and decrease as the level on the contralateral side is increased [52,64]. The LSO is also sensitive to ITDs present in broadband clicks [49,50], similar to stimulation provided by CIs. Responses show the greatest spike rate near 0-μs ITDs and decreasing when the ITD is increased or decreased. The LSO is also responsive to ITDs in the temporal envelope [48,51], where ITDs are preserved by CI processors. Finally, the LSO has very short time constants [48,53], making it the ideal nucleus for representing ITDs of high-frequency transients and envelope ITDs.

The LSO relies on frequency-matched input [52]. Thus, when interaural mismatch in place-of-stimulation is present, the correspondence between LSO output and ITDs would be assumed to decrease. This result was confirmed using a model of the LSO and comparing it against the binaural interaction component in the brainstem [137]. This result is consistent with behavior, which shows a greater proportion of unfused spatial responses and smaller impacts of ITDs on perceived intracranial location with interaural place-of-stimulation mismatch in patients with BiCIs [138,139]. If the LSO encodes binaural cues for listeners with BiCIs, it is likely that challenges imposed by hearing loss and electrical stimulation would result poorer encoding. That is, if the input from either hemisphere were temporally degraded, then the output of the LSO would likely have poorer correspondence with spatial cues.

C. The monaural bottleneck

Limitations to processors and surgically implanted arrays provide some explanation as to patterns observed across patients. They do not, however, account for differences observed within the same individual. In listeners with normal-hearing (NH), sensitivity to ILDs is similar across frequency [358] and high-frequency envelope ITD sensitivity is greatest for apical places-of-stimulation (e.g., [61]). Listeners with BiCIs do not show consistent differences in sensitivity to ILDs or ITDs with place-of-stimulation across listeners [44,105]. Animal studies show similar sensitivity of inferior colliculus neurons to ITDs in animals reared with NH and those who receive BiCIs without deafness or auditory deprivation [296]. Thus, it seems likely that listener-dependent differences in sensitivity to ITDs of pairs of electrodes across different, interaurally matched place-of-stimulation are driven by effects of deafness and factors associated the position of each CI array within the cochlea.

Interestingly, sensitivity to monaural temporal cues with CIs also varies with place-of-stimulation depending upon the listener [9,62,193,194,281]. These tasks have used both detection of temporal fluctuations in the envelope and discrimination of the rate of temporal fluctuations in the pulse rate or temporal envelope. Rate discrimination shows a similar low-pass characteristic to that observed for ITD discrimination and LSO responses [9,62,63,193,194]. Accordingly, a previous experiment showed a relationship between monaural and binaural temporal sensitivity using low-rate, constant-amplitude pulse trains via direct stimulation [9]. Ihlefeld and colleagues' result suggests that the ear with worse temporal sensitivity is predictive of ITD sensitivity when accounting for a large range of stimulation rates.

D. Summary and hypotheses

The present study aimed to extend results from Ihlefeld and colleagues [9] in several important ways. Listeners with BiCIs tend not to use ITDs when other cues are available [37–39]. Thus, we studied lateralization to assess the impact of ITDs on perception, asking listeners to report their perceived intracranial location of sounds. Fine-structure ITDs are discarded by CI processing algorithms [35] and are not preserved with CI output [36]. Moreover, low pulse rates used in laboratory studies that yield ITD sensitivity result in poor speech understanding [359,360], suggesting only ITDs in the envelope would be useful for perception. Interestingly, at very high pulse rates (≳3000 pps), the auditory nerve is suspected to become more stochastic [214]. Thus, if one factor driving differences in sensitivity with place-of-stimulation is the loss of auditory nerve fibers associated with hearing loss and auditory deprivation (e.g., [106,108,110]) and the binaural pathway relies on many auditory nerve inputs to complete its comparisons between ears [74,75,343], then testing very high pulse rates might help delineate the electrode sites with the fewest fibers. It was hypothesized that the ear with the worse sensitivity to changes in amplitude-modulation (AM) rate would predict the range of lateralization responses using high-rate, amplitude-modulated stimuli.

II. EXPERIMENT 1: AM RATE DISCRIMINATION

A. Methods

1. Listeners

Eleven listeners with BiCIs participated in this experiment (age 18-80 years, with mean 57.09 years). Demographics and CI processor information for listeners are listed in Table I. All listeners had at least 1 year of listening experience with bilateral CIs. All listeners provided informed consent and procedures were approved by the Health Sciences Institutional Review Board of the University of Wisconsin-Madison.

TABLE I. Listener demographics. All listeners used a 22 electrode array, where lower values indicate closer proximity to the base of the cochlea.

Listener ID	**Apical Electrode (L/R)**	**Middle Electrode (L/R)**	**Basal Electrode (L/R)**	**Etiology**	**Age at Onset of Deafness (yrs)**	**Age at Implantation (L/R; yrs)**	**Age at Testing (yrs)**	**Internal Processor (L/R)**
IBF	21 / 20	12 / 14	4 / 7	Hereditary	38	56 / 54	65	CI24RE / CI24RE
IBK	20 / 22	12 / 13	4 / 6	Hereditary; noise exposure	53	63 / 69	80	CI24R / CI24RE
IBO	20 / 21	12 / 11	4 / 7	Otosclerosis; sudden	20	45 / 42	52	CI24RE / CI24RE
IBY	20 / 22	12 / 12	4 / 6	Unknown; progressive	41	43 / 47	48	CI24RE / CI512
ICC	20 / 20	12 / 13	4 / 6	Unknown; progressive	2	63 / 60	67	CI24RE / CI24RE
ICD	20 / 20	12 / 9	4 / 4	Enlarged vestibular acqueduct	3	50 / 44	59	CI24R / CI24RE
ICI	20 / 19	11 / 11	4 / 7	Unknown	31	50 / 51	58	CI24RE / CI24RE
ICJ	20 / 20	12 / 12	4 / 7	Unknown	25	60 / 60	63	CI512 / CI512
ICM	20 / 20	12 / 14	4 / 3	Unknown	23	57 / 58	64	CI512 / CI24RE
ICP	20 / 21	12 / 12	4 / 6	Unknown; progressive	3	46 / 49	54	CI24RE / CI24RE
IDD	20 / 20	12 / 11	4 / 4	Congenital	0.5	10 / 4	18	CI24RE / CI24R

2. Stimuli and equipment

Stimuli were monopolar, biphasic pulse trains presented at a rate of 4000-pps with sinusoidal AM at 100% AM depth. The pulse duration was 25 μs with an 8-μs interphase gap. Pulse trains were presented to one of six possible electrodes for each listener, closer to the apex, middle, or basal end of the cochlea in the left or right ear. The electrode being stimulated is provided in Table I. Sinusoidal AM was applied by multiplying the constant-amplitude pulse train using Eq. 1:

$$y(t) = (1 + \sin(2\pi f_m t))\, x(t) \qquad (1)$$

where $y(t)$ is the instantaneous amplitude of the sinusoidally amplitude-modulated (SAM) pulse train at time t, f_m is the modulation frequency fixed at 100 Hz, t is a vector of time values, and $x(t)$ is the instantaneous amplitude of the 4000-pps pulse train at time t. The stimuli used in the present experiment were presented at a comfortable level and had a duration of 300 ms. SAM pulse trains were gated with 10-ms cosine ramps at stimulus onset and offset.

Stimuli were generated in MATLAB (Natick, MA). They were presented through direct stimulation of CIs using the RF Generator XS research processor (Cochlear, Ltd., Sydney, Australia). All testing took place in a quiet, private room free from distractions.

3. Procedure

Loudness balancing was completed to ensure that stimuli resulted in equal loudness on each electrode following procedures described in a previous methodological paper from our laboratory [280]. Briefly, the comfortable (C) levels were determined using a 300-ms, constant-amplitude pulse train for even-numbered electrodes along the CI array in both ears. Current units for comfortable level were interpolated for odd-numbered electrodes. Stimulation level (in current units) was adjusted for each electrode by the experimenter until the listener reported equal loudness for all possible pairs of electrodes. The pitch-matched electrode was estimated for three pairs of electrodes in the left ear with the direct pitch comparison procedure from Litovsky and colleagues [280]. The pitch of six electrodes in the right ear per electrode in the left were compared against one another. The electrode in the right ear resulting in approximately the same pitch was chosen for the pair. Since the time of data collection, additional studies have demonstrated that pitch-matching is not a reliable way to ensure matched place-of-stimulation between ears [361–363]. However, ITD lateralization with monopolar stimulation is similar for several neighboring electrodes [138,139], probably due to large spreads of current, suggesting that performance should be similar to that for place-matched electrodes.

Next, threshold (T) and maximally comfortable (M) levels were determined for each electrode in each pair. SAM pulse trains were presented at an AM rate at 100 Hz for all possible pairs of electrodes used in the study. Stimulation level (in current units) was adjusted for each electrode by the experimenter until the listener reported equal loudness for all possible pairs of electrodes. No loudness balancing was completed across the AM rates used in this study because previous experiments showed no change in loudness over AM rate [193,281].

Before testing began, listeners were familiarized with perceptual changes associated with increasing AM rate. Listeners were presented with AM rates of 100, 250, 400, 500, 800, and 1000 Hz and asked to describe their perception (e.g., compare sounds to something encountered in real life). Responses were documented by the experimenter so that if the listener became confused during testing, additional familiarization could be completed efficiently.

Finally, listeners began the experiment. Testing was completed in blocks associated with each electrode, which were partially counterbalanced across listeners. Listeners completed a three-interval, two-alternative forced-choice “oddball” discrimination task. On each trial, the first interval presented a standard AM rate (100 or 400 Hz). The second or third interval (with equal probability) also presented the standard AM rate. The other interval consisted of an oddball (higher) AM rate, which was determined from one of six possibilities to estimate a psychometric function. The oddball interval had a rate of $\log(\Delta f_m/f_m)$ = -1.78, -1.34, -0.89, -0.45, -0.01, or 0.44 relative to the standard AM rate, and 30 repetitions per oddball AM rate were collected. This resulted in oddball rates of 101.7, 104.6, 112.8, 135.5, 198.5, and 373.0 Hz for the 100-Hz standard and 406.7, 418.5, 451.2, 542.0, 793.8, and 1492.0 Hz for the 400-Hz standard. An inter-stimulus interval of 300 ms separated each presentation and listeners were not allowed to repeat stimuli. Two interleaved psychometric functions were estimated simultaneously corresponding to standard AM rates of 100 and 400 Hz, interleaved across trials. The 100-Hz AM rate was selected to be consistent with the AM rate in the lateralization experiment. The 400-Hz AM rate was chosen because it is above the 300-Hz AM rate limit, and might be useful in disambiguating electrode sites yielding the greatest temporal sensitivity. Visual feedback was given after each trial. One practice block was given on the first electrode to be tested to ensure that listeners understood the task. Testing took approximately 6-8 hours for each listener.

4. Analysis

Percentage of correct responses were fitted with a two-parameter sigmoidal curve bound in version 3.5.1 of R using mixed-effects logistic regression, predicted by the oddball rate in $\log(\Delta f_m/f_m)$ units, standard AM rate, and place-of-stimulation using version 1.1-21 of the lme4 package in R [285]. Degrees of freedom were estimated using the Satterthwaite approximation [364]. Thresholds from psychometric functions were estimated using version 2.5.6 of the psignifit toolbox in MATLAB [238].

B. Results

The goal of this experiment was to evaluate monaural AM rate discrimination at three different places-of-stimulation in each ear. Results for each individual are shown in Fig. 1. Listener ICI could not complete the AM rate discrimination task at 400-Hz, so they were tested at 10- and 80-Hz standard AM rates instead. They were therefore excluded from the statistical analysis. Results of the mixed-effects logistic regression showed a significant effect of the oddball rate in $\log(\Delta f_m/f_m)$ units [$z = 14.733$, $p < .0001$]. There was also a standard AM rate [$z = -12.350$, $p < .0001$], where the logistic curve was shifted to the right for the 400-Hz relative to the 100-Hz standard AM rate, corresponding to lower percentage correct responses. There was no significantly significant difference in percent correct responses for middle [$z = 0.311$, $p = .749$] or basal [$z = 1.803$, $p = .071$] relative to apical places-of-stimulation. There was a significant interaction between the oddball rate and standard AM rate [$z = -6.001$, $p < .0001$], indicating a shallower slope for 400-Hz standard AM rate. There was also a significant interaction between the oddball rate and the intercept associated with basal places-of-stimulation [$z = 2.343$, $p < .05$], indicating a steeper slope with basal relative to apical places-of-stimulation. There were no other significant interactions.

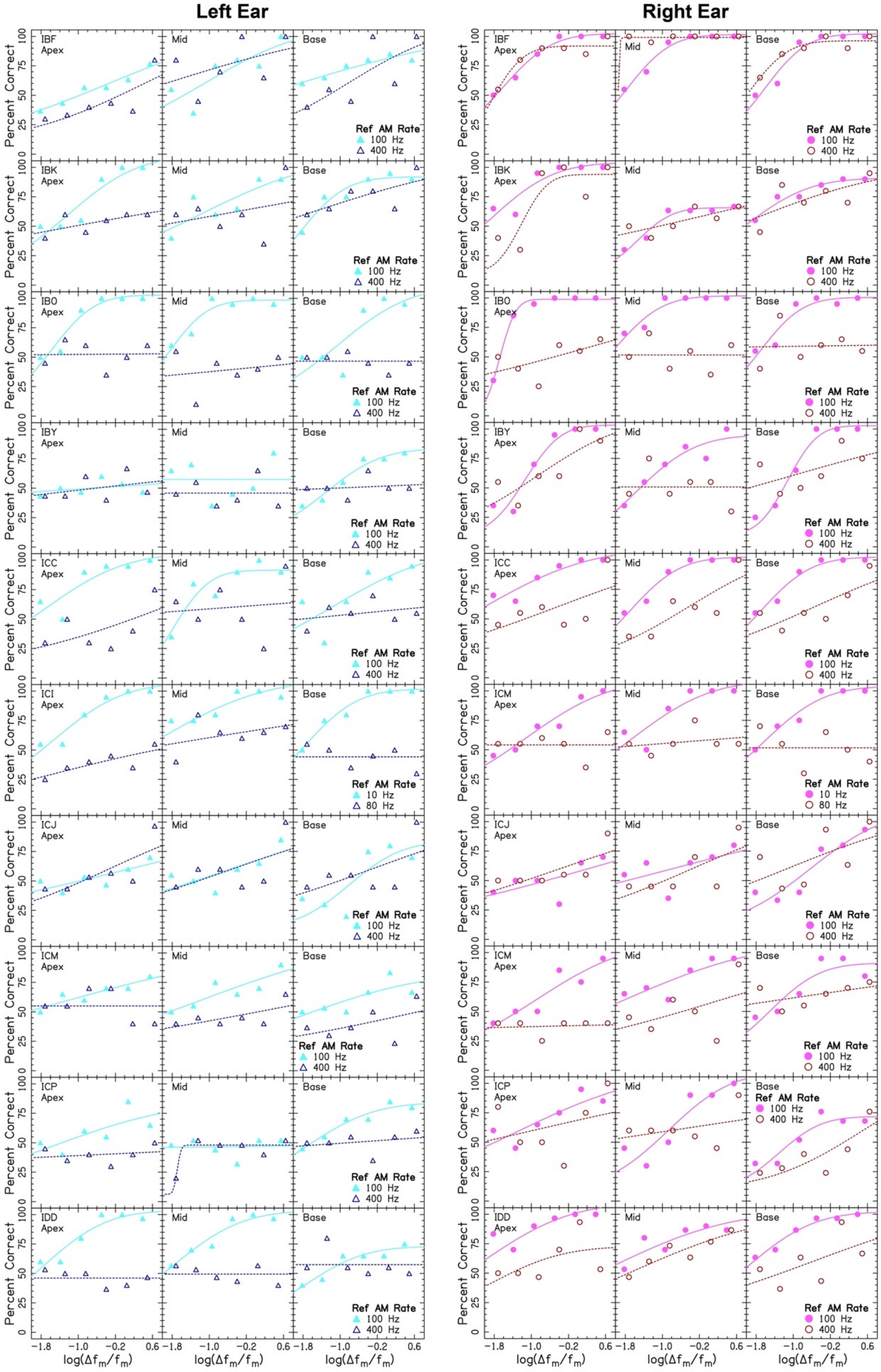

Left Ear
Right Ear
IBF
IBK
IBO
IBY
ICC
ICI
ICM
ICJ
ICP
IDD
Apex
Mid
Base
Ref AM Rate
100 Hz
400 Hz
10 Hz
80 Hz
Percent Correct
0
25
50
75
100
-1.8
-1.0
-0.2
0.6
log(Δf_m/f_m)

FIG. 1. Psychometric functions for AM rate discrimination. Each panel shows individual performance for the apical (left), middle (middle), and basal (right) electrodes. The left and right columns correspond to the left and right ears, respectively. Subject codes for each individual are given in the top left corner of the left-most panel. The y-axis corresponds to the percentage of correct responses across all trials. The x-axis corresponds to the oddball AM rate. Standard AM rates are indicated by the color and shape in the figure legend.

Thresholds were extracted from the psychometric function for each listener at 71.1% correct performance. Results for individuals are shown in Fig. 2, and average results are shown in Fig. 3. These results largely confirm those of the logistic regression, with the important exception that many thresholds could not be estimated for listeners, especially at the 400-Hz standard AM rate.

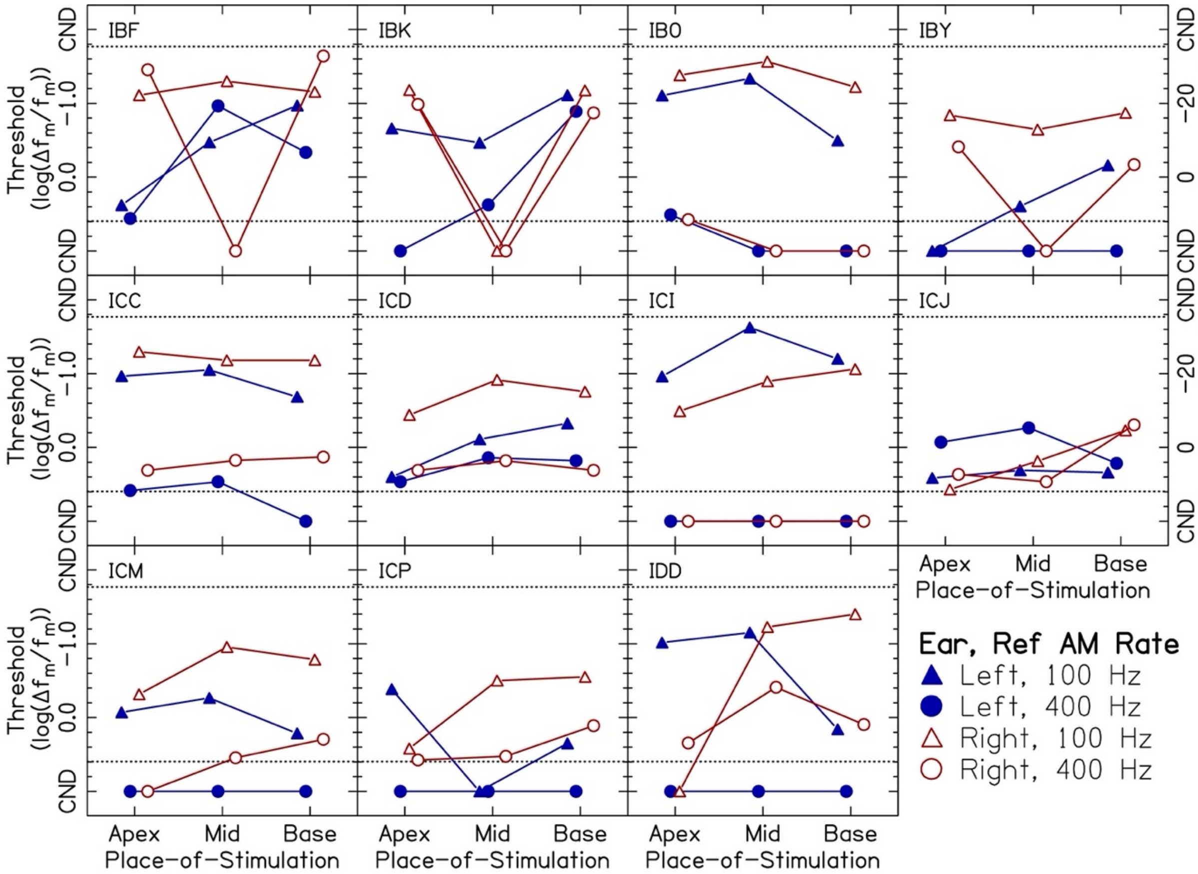


FIG. 2. Thresholds for AM rate discrimination. Each panel corresponds to a different listeners whose subject code is given in the upper-left corner. The x-axis corresponds to the standard AM rate. The y-axis corresponds to the threshold, presented in reverse order such that higher values indicate better performance as in [193]. Triangles and circles represent apical and basal places-of-stimulation, respectively. Closed and open shapes represent the left and right ear, respectively. “CND” indicates that a threshold could not be determined because performance fell above or below 71.1% correct.

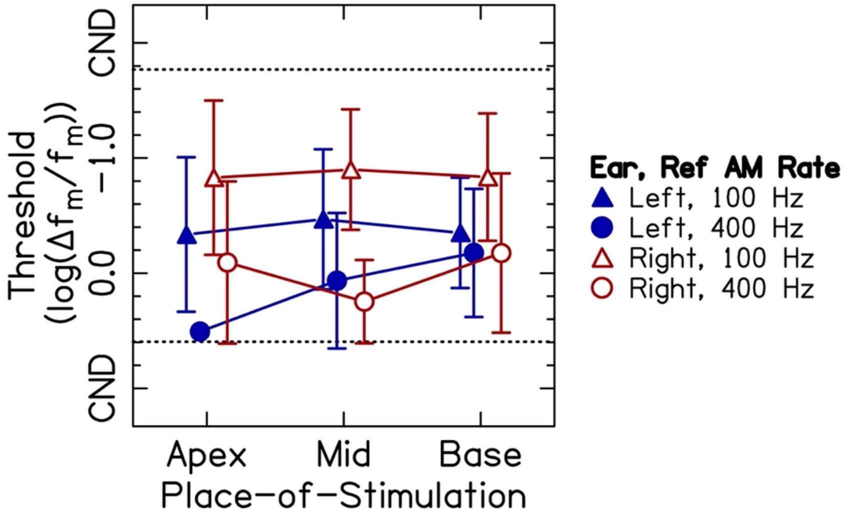


FIG. 3. Thresholds for AM rate discrimination averaged across listeners in Fig. 2. The x-axis corresponds to the standard AM rate. The y-axis corresponds to the threshold, presented in reverse order such that higher values indicate better performance as in [193]. Triangles and circles represent apical and basal places-of-stimulation, respectively. Closed and open shapes represent the left and right ear, respectively.

C. Discussion

Results from Experiment 1 demonstrated that monaural AM rate discrimination did not vary consistently with place of stimulation across listeners, consistent with previous results [193,194,281]. Additionally, AM rate discrimination was worse for the 400-Hz relative to the 100-Hz standard AM rate, consistent with studies of AM rate discrimination [193,194,281] and pulse rate discrimination [9,62,63]. One benefit of the study by Ihelfeld and colleagues and the method used here is that, in contrast to unilateral implantees, listeners with BiCIs can act as their own control, allowing two samples to be collected per person and increasing statistical power. It may be that discrimination relative to higher rates is indicative of temporal resolution associated with each electrode site. That is, 400 Hz is nearer to the suspected 300-Hz "rate limitation" in listeners with BiCIs. For example, some listeners exhibit "star" performance for high-rate monaural rate discrimination [62,63] and discrimination of binaural cues [44,140]. Results from these listeners might then suggest high degrees of temporal resolution relative to other listeners, which may facilitate accurate binaural processing. Accordingly, Ihlefeld and colleagues showed that monaural pulse rate and ITD discrimination decrease in a similar fashion for listeners with BiCIs [9].

Interestingly, there were consistent differences in the slope of the psychometric function with place-of-stimulation. Electrode locations closer to the base tended to have steeper psychometric functions. The effect size associated with this trend was small ($p<.05$), and there is no obvious reason to suspect that the base should have yielded steeper psychometric functions. Most previous studies concerning AM and pulse rate discrimination have only tested one rate (e.g., [9,62,63]) or estimated threshold (e.g., [193,194,281]). Thus, it may be that other aspects of the psychometric function are important indicators of performance between electrode locations. The trade-off is considerably longer testing time in tasks like those used in the present experiment.

III. EXPERIMENT 2: LATERALIZATION

A. Motivation and methods

Experiment 1 demonstrated no reliable differences in sensitivity to AM rate with place-of-stimulation. The goal of Experiment 2 was to explore the relationship between intracranial lateralization and monaural temporal sensitivity. Accordingly, we used the averaged percentage correct scores from the worse at the 400-Hz standard AM rate in Experiment 1 to predict lateralization responses in Experiment 2. The same listeners who participated in Experiment 1 participated in Experiment 2, using the same equipment and electrodes. Stimuli contained whole-waveform ITDs of 0, ±200, ±400, ±600, ±800, ±1000, and ±1200 μs.

Listeners' response locations for each stimulus were converted into numerical values within the range of ±10, corresponding to the left (negative) and right (positive) sides. These data were fit using the function described in Eq. 2:

$$Lat(ITD) = A\,\mathrm{erf}\left(\frac{ITD-\mu_{ITD}}{\sqrt{2}\sigma}\right) - \mu_{Lat} \qquad (2)$$

where *Lat(ITD)* corresponds to the predicted value of lateralization at a particular ITD, A corresponds to maximum extent of lateralization achieved by the fit, erf corresponds to the error function which takes as input any real number and whose output is a sigmoidal curve bounded between ±1, *ITD* corresponds to the value of the ITD in μs, μ_{ITD} corresponds to the horizontal shift, σ is related to the lateralization slope, and μ_{Lat} corresponds to the vertical shift. All curve fitting was completed using non-linear least squares via the curve fitting toolbox in MATLAB.

The lateralization range, or extent to which listeners perceived a sound to the left and right in the head, was assessed to provide a proxy of the impact of spatial cues on perception. Lateralization range was calculated as the absolute difference of the perceived left and right positions for each listener for ITD values of ±1200 μs from the fitted curve. Lateralization range was predicted by place-of-stimulation using mixed-effects ANOVAs with random-intercepts associated with listener. The lateralization range was predicted by the performance of the worse ear in Experiment 1 using mixed-effects linear regression, with fixed-effects of the percentage correct averaged across oddball rates with the 400-Hz standard from worse ear and random intercepts associated with listener. Degrees of freedom for both sets of analyses were based upon the Satterthwaite approximation [364]. Post-hoc comparisons were completed using estimated marginal means with Tukey corrections for multiple comparisons in version 1.3.0 of the *emmeans* package in R.

B. Results

Individual lateralization results are shown in Fig. 4 and average results are shown in Fig. 5. Similar to previous lateralization studies in listeners with BiCIs using singular electrode pairs [46,138,139,162], there was considerable variability across individuals. The results from Fig. 5 suggest that the range of lateralization responses might be smaller for apical places-of-stimulation. Consistent with this observation, there was a significant effect of place-of-stimulation with the mixed-effects ANOVA [$F(2,48) = 12.342$, $p < .0001$]. Post-hoc comparisons indicated that lateralization range was significantly greater for electrode pairs closer to the base than the apex [$t(48) = 4.025$, $p < .001$] and the middle compared to the apex [$t(48) = 4.535$, $p < .001$], but not the middle compared to the base [$t(48) = 0.511$, $p = .867$].

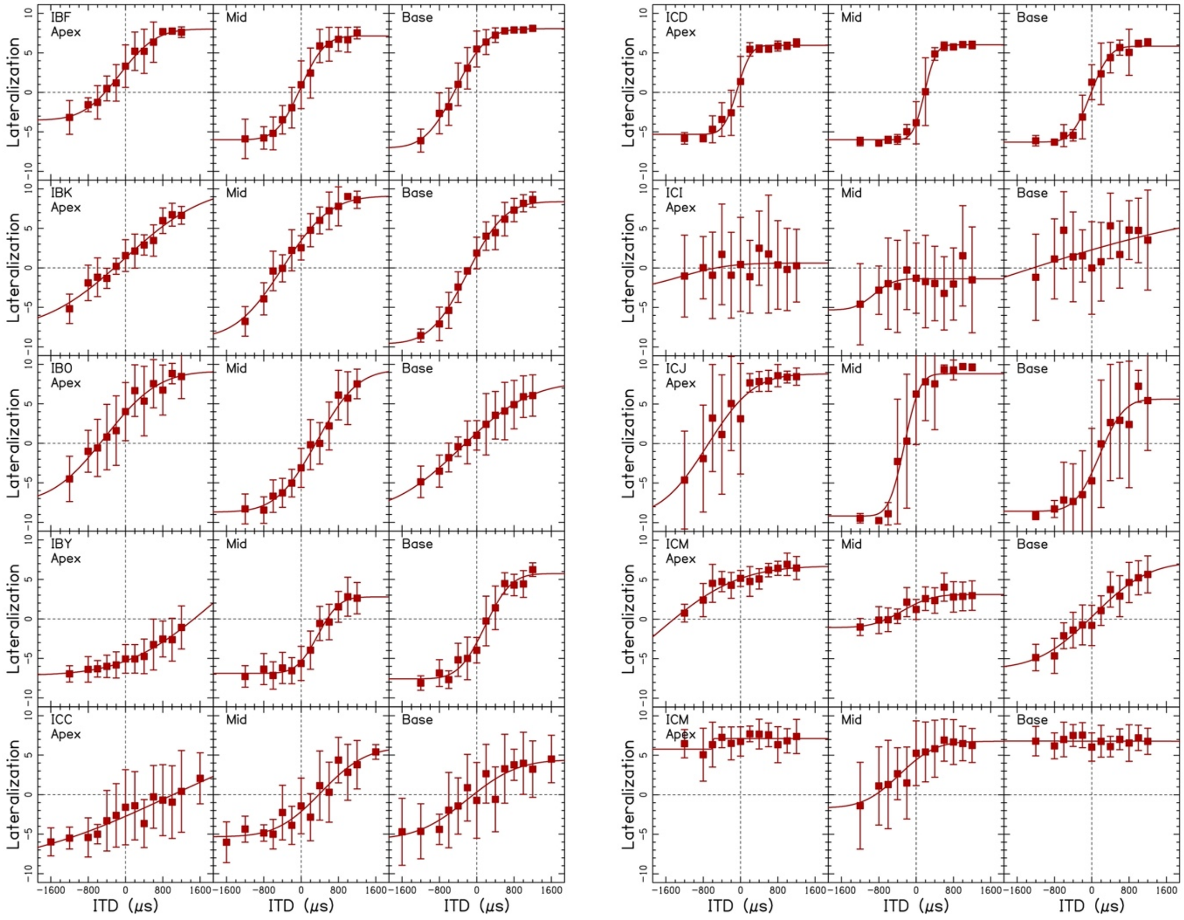


FIG. 4. Lateralization of ITDs by place-of-stimulation and listener. The left, middle, and right panel correspond to apical, middle, and basal electrode pairs, respectively. The subject code of each listener is given in the top-left corner or the left-most panel. The x-axis corresponds to the ITD presented on each trial. Negative values indicate a left-leading ITD. The y-axis corresponds to the extent of lateralization, with ±10 indicating lateralization to left or right the side of the head. A lateralization value of 0 represents the center of the head. Negative values indicate a leftward perception. Points shown were averaged across ITDs ± one standard deviation.

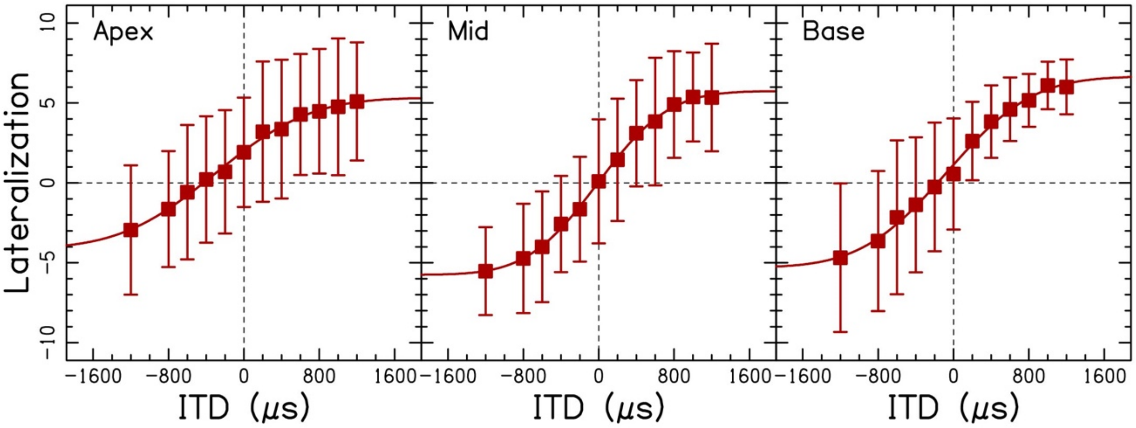


FIG. 5. Lateralization of ITDs by place-of-stimulation. The left, middle, and right panel correspond to apical, middle, and basal electrode pairs, respectively. The x-axis corresponds to the ITD presented on each trial. Negative values indicate a left-leading ITD. The y-axis corresponds to the extent of lateralization, with ±10 indicating lateralization to left or right the side of the head. A lateralization value of 0 represents the center of the head. Negative values indicate a leftward perception. Points shown were averaged across listeners ± one standard deviation.

It was of greatest interest to determine the relationship between monaural and binaural responses. Figure 6 shows the relationship between the mean percent correct discrimination of the worse ear in Experiment 1 for the 400-Hz standard AM rate and the lateralization range in Experiment 2. A mixed-effects regression determined that the percentage of correct responses for the 400-Hz standard AM rate in the worse ear significantly predicted the lateralization range [$t(25) = 2.494$, $p < .05$]. Interestingly, results from the 100-Hz standard AM rate were not predictive of lateralization range [$t(28) = 1.392$, $p = .175$]. Greater degrees of freedom for the ~100-Hz compared to the 400-Hz standard AM rate comes from the inclusion of listener ICI, who was tested with a standard AM rate of 80 Hz as they could not complete the task at higher rates. If listener ICI is excluded from the analysis, the conclusion with the ~100-Hz standard AM rate does not change [t(25) = 1.162, p = .256].

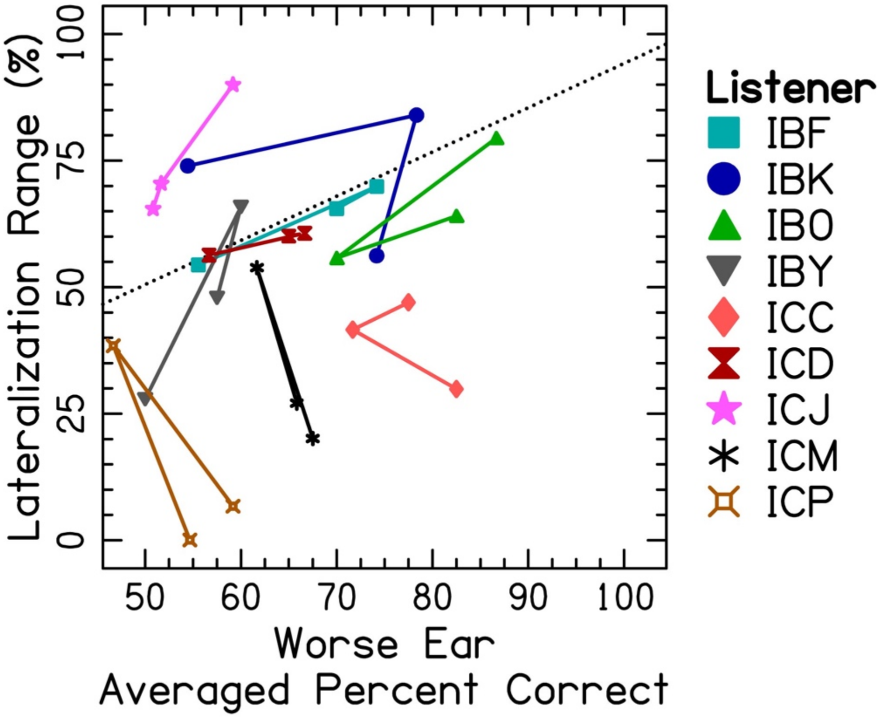


FIG. 6. Relationship between Experiment 1 and 2. The x-axis corresponds to the percent of correct responses from the ear yielding the poorer (lower) percentage of correct responses at the 400-Hz standard AM rate in Experiment 1. The y-axis corresponds to the lateralization range of the electrode pair. The shape and color given in the figure legend on the right indicates the listener. The dotted line shows the fixed-effect regression estimate from the mixed-effects model.

C. Discussion

The present study found a significant relationship between place-of-stimulation and ITD lateralization. This is inconsistent with the broader CI literature showing listener-dependent relationships between place-of-stimulation and ITD sensitivity [44,105]. Sensitivity was significantly worse at the apical compared middle and basal electrodes. This pattern is also inconsistent with listeners with NH, who show greatest ITD sensitivity to the lower frequencies (near 4 kHz) of high-frequency transients (e.g., [61]). Thus, given the small sample size included in the present study, it seems likely that our listeners happened to be less sensitive to ITDs at apical electrodes. Lateralization curves tended to show the greatest bias (i.e., upward or downward shifts; Fig. 4 and 5) for apical electrodes, which may have driven performance. This could be due to a lack of centering in the present experiment. However, ITD sensitivity does not seem to improve significantly even when stimuli are loudness balanced using ILD lateralization [144,319].

One other major difference between the present study and previous studies is that we used extremely high pulse rates (4000 pps). The literature is relatively scant concerning lateralization of envelope ITDs in listeners with BiCIs, but the ranges observed here are similar to previously published experiments using stimulation rates near 1000 pps with participants of a similar age [46,162]. It is possible that increased stochasticity of responses auditory nerve responses due to very high rate of stimulation [214] contributed to differences that are not usually observed using low-rate, constant-amplitude stimulation that yields the greatest ITD sensitivity.

IV. GENERAL DISCUSSION

The present study investigated the relationship between lateralization of envelope ITDs and sensitivity to AM rate in each ear. Consistent with previous results [9], we showed that the ear with poorer temporal sensitivity is predictive of binaural performance. Where previous studies have only explored the relationship between monaural and binaural temporal sensitivity [9,273], the present study expands those results to imply that sites with poorer temporal sensitivity may also decrease the utility of ITDs for perception. Interestingly, the present study did not replicate the result from Ihlefeld and colleagues [9] showing that rate discrimination with the same standard used for ITD discrimination was predictive of performance.

The present study only showed a relationship between monaural temporal sensitivity at the 400-Hz standard AM rate. There are several reasons this result could have occurred. The present study tested 400-Hz standard AM rates in Experiment 1 because it was thought that testing a rate above the 300-Hz "rate limitation" observed in patients with BiCIs would be a more robust predictor of temporal sensitivity. It is important to note that Ihlefeld and colleagues tested many different standard rates (100, 200, 300, and 500 pulses per second) within each listener. Thus, it may be that the present study was statistically underpowered to detect such an effect. Alternatively, it may be that lateralization range corresponds more closely with rate discrimination at higher standard AM rates.

A. Temporal sensitivity across the electrode array

The present study fits with the broader literature showing large amounts of variability in temporal sensitivity when providing stimulation via different electrodes in the electrode array. Psychophysical studies show varied sensitivity to AM detection and rate discrimination [191–194,241,264,281]. Similarly, the dynamic range varies across CI arrays [123,126,365]. Objective measures have been developed to reduce variability associated with perceptual results, which are also thought to reflect the temporal response properties of auditory nerve fibers (e.g., [124,365–367]). All of these measures have been proposed as means to predict the relative temporal sensitivity of auditory nerve fibers stimulated by different electrodes, the poorest of which could be deactivated to improve speech understanding.

Only recently have studies begun to expand these results to binaural stimulation. Rate discrimination predicts ITD sensitivity [9]. Dynamic range predicts sensitivity to binaural masking level differences [121]. In contrast to complicated tasks like speech, where many different features are likely useful for perception, the circuits involved in ITD coding are well-characterized and contain extremely specialized mechanisms to facilitate microsecond sensitivity. Therefore, one should expect that any decrement in temporal coding should lead to poor binaural outcomes, and that monaural temporal sensitivity should be a robust predictor of ITD sensitivity and utility. Results from the present and previous studies [9] suggest that sensitivity in the poorer ear is a robust but modest predictor of binaural performance. Listeners with BiCIs experience many different challenges to the binaural system (e.g., interaural place-of-stimulation mismatch, differences in loudness growth). Thus, it may be that combining predictors of temporal sensitivity in a larger testing battery provides more robust predictors of performance, and can be the most useful to researchers and clinicians.

B. Limitations

Data in the present study took approximately two days of testing to collect, and only three pairs of electrodes were assessed. Similar timelines were required in previous studies [9]. Thus, if tasks like these are to ever be clinically useful, a much more efficient version would need to be developed. Because of the extensive length of testing, we could only test a small group of listeners. To more extensively address the hypothesis (i.e., whether the worse ear limits binaural outcomes), given the small effect sizes observed across studies, a large cohort may be required. It may be more practical to use objective measures that are highly correlated with psychophysical sensitivity instead of extended psychophysical tasks. One of the major, related challenges is that lateralization performance is highly variable across individuals even with NH [337]. Then studies in listeners with BiCIs either need extremely large sample sizes to detect difference between groups or extensive measurements within the same individuals to acquire an appropriate baseline. Interestingly, data from the 400-Hz standard AM rate were predictive of ITD lateralization even when many listeners could not discriminate AM rates at a threshold of 71.1% correct. Listeners reported being quite frustrated by this task and it is unreasonable to ask listeners to the task at only the 400-Hz standard AM rate. It is likely that the 100-Hz AM rate helped to anchor performance and remind listeners of their task, as well as prevent them from giving up during the experiment.

C. Summary and conclusions

Together from Experiment 1 and 2, several conclusions can be drawn with respect to monaural and binaural temporal processing:

1. Listeners showed no consistent differences in AM rate discrimination between places-of-stimulation at 100- and 400-Hz standard AM rates, replicating results in listeners with unilateral CIs (Fig. 2, 3; [193,194,241,281]).
2. Listeners showed similar lateralization to envelope ITDs with extremely high pulse rates (4000 pps) compared to previous studies using lower pulse rates (Fig. 4, 5; [46,162]).

3. Listeners showed significantly smaller lateralization ranges for envelope ITDs at apical compared to middle and basal electrode pairs (Fig. 4, 5), in contrast to previous studies assessing ITD discrimination using constant-amplitude pulse trains [44,105].
4. AM rate discrimination relative to the 400-Hz standard in the poorer ear was predictive of lateralization range (Fig. 6), consistent with previous studies showing a small effect size relationship between monaural and binaural temporal processing [9]. This relationship did not hold for the 100-Hz standard AM rate.

Chapter 6: Asymmetric Temporal Envelope Encoding: Binaural Speech Fusion in Younger and Older Adults with Normal-Hearing

I. INTRODUCTION

Patients with bilateral severe to profound hearing loss can receive cochlear implants (CIs) to gain access to hearing. Bilateral CIs (BiCIs) improve sound source localization performance and speech understanding in noise relative to unilateral implantation (e.g., [6]). However, the extent of this benefit varies highly across patients [6,10–13,15,345].

The ability to distinguish between sources of interest in a complex auditory environment is facilitated by sound source segregation cues, which are used to modulate attention (e.g., [149,209]). Advantage listening with spatially separated relative to co-located target speech and masking stimuli (i.e., spatial unmasking) comes from several sources. Head-shadow provides an improved signal-to-noise ratio (SNR) in the ear opposite the masking stimulus. Binaural redundancy allows for multiple “looks” at the target stimulus. Finally, interaural differences that facilitate a perceived difference in location between target and masker stimulus are also suspected to improve performance. Some of the most convincing evidence of this final phenomenon comes from experiments that show improved speech understanding when a mixture of the target and masker are presented in one ear and a copy of the masking stimulus is added to the opposite ear, resulting in an improvement in performance (e.g., [8,15]). Models of spatial unmasking based on studies completed in listeners with normal-hearing (NH) often assume that the ears act as ideal and independent channels that can be used to cancel out the masking stimulus and attend to a target of interest (e.g., [16]). Similarly, studies presenting unrelated maskers to the ear opposite the target have been interpreted to suggest that listeners can ignore one ear without any decrement in performance [368,369]. These assumptions may not apply to patients with BiCIs, who often show marked asymmetry in speech understanding and spectro-temporal resolution between the ears.

A. Spectro-temporal resolution and speech

Measures of spectro-temporal resolution have been applied in patients with CIs to predict their speech understanding. One example includes different versions of a “spectral ripple” task where listeners are asked to discriminate between stimuli containing different spectro-temporal profiles (e.g., [298,357,370]). Results from spectral ripple tasks tend to predict speech understanding and they are interpreted as a proxy for spectro-temporal resolution. Consistent with these findings, spectral blurring of stimuli results in poorer speech understanding in listeners with CIs [371], and simulations of spectro-temporal smearing worsen speech understanding and spectral ripple discrimination in listeners with NH [357]. Other measures of psychophysical sensitivity to spectral or temporal cues have been used to identify electrodes that might limit spectro-temporal fidelity. Some examples include thresholds to focused stimulation strategies [123,125], temporal modulation detection or discrimination [191,192,264,372], and electrode discrimination [373]. The dynamic range of CI electrodes is predictive of temporal modulation detection thresholds [372]. Similarly, several studies have shown that smaller dynamic ranges are associated with poorer speech understanding in listeners with CIs [374,375] and simulations in listeners with NH [376].

One approach to combat poor spectro-temporal resolution is to deactivate electrodes yielding poor sensitivity. Studies have used computed topography (CT) scans [377] and electrically evoked compound action potentials [291,365,366,378] to this end. Because they do not rely on a listener’s perception, this group of approaches are referred to as objective measures. Objective measures are affected by the health of the auditory nerve and the distance of CI electrodes from the auditory nerve [124]. Collectively, studies addressing spectro-temporal resolution and speech understanding suggest that poor spectro-temporal resolution leads to poor speech outcomes for patients due to the interface between CI electrodes and the auditory nerve. The majority of this work focuses on unilaterally implanted patients.

Studies of patients with BiCIs and simulations in NH suggest that the poorer ear may limit performance in binaural tasks [9–13,15,273]. To address this more directly, some studies first indexed or manipulated temporal fidelity in both ears, then assessed sensitivity to binaural cues. The ear with poorer temporal fidelity was predictive of sensitivity to binaural cues [9,273]. Simulations in NH used asymmetric dynamic range (i.e., amplitude modulation depth), where smaller dynamic ranges in listeners with BiCIs have resulted in poorer sensitivity to binaural cues [121,122]. Other studies evaluating speech understanding in both ears suggest that differences between ears are associated with poorer benefits, like spatial separation of target speech and masking stimuli [11–13,15].

Spectro-temporal sensitivity need not be worse in one ear; differences between the ears that do not affect intelligibility are sufficient to degrade speech understanding and binaural benefit. For example, interaural place-of-stimulation mismatch for electrodes conveying the same frequency information results in decreased benefit when listening to spatially separated speech [141,379]. A spectral shift tends to worsen speech understanding for bilaterally presented target speech in patients with BiCIs [380] and simulations in NH [19,379]. Additional research shows that two different ears can limit the ability to understand vowels. This study evaluated vowel perception using continua between two vowels [24]. Reiss and colleagues first showed that listeners with hearing loss who use hearing aids and/or CIs can have considerably different vowel perception between the ears when stimuli were presented monaurally. They then showed that the majority of these listeners demonstrate interference in their ability to understand vowels presented to the left and right ear simultaneously. Because of the limited focus upon bilateral spectro-temporal resolution and how differences contribute to speech understanding, relatively little research has addressed the mechanisms that contribute to poorer speech understanding and limited bilateral benefit with interaural asymmetry.

B. Mechanisms underlying limited binaural benefit

Usually a copy of a masker in the ear opposite the target speech in spatial unmasking experiments results in improved performance for listeners with NH and BiCIs e.g., [8,14,15]. Experiments in patients with BiCIs who have marked asymmetry in speech understanding between the ears have demonstrated that they experience contralateral *interference* when target speech is presented to their poorer ear [10,13–15]. Listeners with a CI in one ear and NH in the other ear show the same pattern of performance [23]. In simulations of BiCIs, contralateral interference occurs when one or both ears have poor spectro-temporal resolution [19,20]. There have been two basic mechanisms proposed that are suspected to drive contralateral interference. The first is differences in how target and masking stimuli are perceptually segregated from one another, suggesting that they may instead be fused together (e.g., [20]). The second is a failure of attention, where it is more difficult to ignore the clearer stimulus (e.g., [19]).

Evidence of a relationship between the efficacy of sound source segregation cues and unmasking comes largely from studies assessing fundamental frequency differences between target and masking speech [24–26]. In these studies, the range of interaural frequency or electrode differences over which a single pitch was heard (i.e., fusion range) was first assessed. Then, it was used to predict performance in a speech understanding task. Interestingly, the fusion range is predictive of the benefit associated with fundamental frequency difference [25] and probability of fusing two different vowels into one [26]. Spatial unmasking is thought to be driven primarily by binaural spatial cues for listeners with NH [306]. While listeners are able to localize sound sources even if they are not fused in pitch [381], for speech stimuli, there may be mandatory integration of spectral content between the ears [382]. This integration of speech information is driven primarily by a common fundamental frequency. Interestingly, harmonically related spectral components are also fused in pitch when presented to different ears [383]. Because listeners with BiCIs receive limited harmonic content due to a smaller number of spectral channels and loss of temporal fine-structure information, it seems highly likely that they may be less sensitive to differences in fundamental frequency for formants at higher frequencies. Thus, it may be that speech is particularly susceptible to over-fusion of content from different speakers when stimuli are spectro-temporally degraded.

An attentional basis of better ear advantage and worse ear interference for spatial unmasking is supported by the trend that performance remains intact if the target is in the better ear. This has been demonstrated for listeners with BiCIs [10,13,15], listeners with one CI and one NH ear [14,23], and simulations of BiCIs in listeners with NH [19]. If contralateral interference resulted purely because of an inability to segregate target from masker, then it should not matter whether the target is in the better or poorer ear.

Right ear advantage has been well-documented in the auditory literature and is suspected to result from an attentional bias toward the right ear for typically developing listeners with NH [33,174]. Ear advantage varies depending upon several factors. The intensity in either ear can be used to modulate right ear advantage, with the right ear tending to show advantage even for moderately left-favoring interaural level differences [176,384]. The relative onset time of speech in each ear also contributes, with longer interaural delays resulting in smaller ear advantage since listeners are able to process speech in each ear [175]. Interestingly, increasing age is associated with an elevated right ear advantage [385]. Since most experiments concerning listeners with BiCIs tend to test older individuals, age is an important variable to account for in experiments concerning auditory spatial attention. Ear advantage is also modulated by the spectro-temporal fidelity of the stimulus in either ear [19]. Accordingly for listeners with hearing loss who are suspected to have differences in the spectro-temporal fidelity in each ear, many individuals report having a "better ear" that is not consistently on their right side.

C. Goals and hypotheses

It is becoming clearer in the literature that processing is not truly independent in each ear. Instead, it seems that information is integrated by the central auditory system and a highly efficient attentional network can be used to focus on a source of interest. Because of the inherent connection between sound source segregation and attention, it is difficult to disentangle both processes from one another and determine how they might affect patients. While there is a slight right ear advantage noted in the literature for listeners with NH, listeners with BiCIs can have considerably different speech outcomes between the ears. One of the major challenges of studies of interaural asymmetry in listeners with BiCIs is that extremely large sample sizes are required to account for differences between patients. Thus, it is sometimes more practical to simulate particular sources of asymmetry in listeners with NH to determine the impact on perception.

The present study aimed to critically evaluate the effects of interaurally asymmetric temporal fidelity on binaural speech perception using a task that could be used to explore both fusion and auditory attention. To meet this aim, we adapted a speech perception experiment from the 1970s exploring a phenomenon called "phonological fusion." In phonological fusion experiments, listeners were presented with two rhyming words to the left and right ear [386,387].

One word began with a stop consonant (e.g., /b/) and the other began with a liquid (e.g., /l/). Both words shared the same ending (e.g., -ed), and combining the stop and liquid into a cluster would generate a word in English (e.g., bled). In the original experiments, when words were generated using natural speech productions that were presented simultaneously, listeners would report hearing one word on approximately 30% of trials. The word reported could correspond to the fused word, the word in the left or right ear, or some other word unrelated to those presented. Thus, using this paradigm, it is possible to assess whether listeners fused the percept into one word, and whether they weighted the ears equally or unequally.

In the present experiment, we simulated CI processing using a vocoder and manipulated the temporal fidelity of the speech in each ear by varying its dynamic range. Further, we tested a group of listeners similar in age to a typical CI study cohort. We assessed phonological fusion as well as speech understanding of the same word in each ear or words with different vowels to evaluate a wide spectrum of performance. We hypothesized that listeners would report one word more often when stimuli had small dynamic range. We further hypothesized that the word reported would correspond to the right ear in symmetric dynamic range conditions or the better ear in asymmetric dynamic range conditions.

II. METHODS

A. Listeners and Equipment

Nine listeners (20-72 years; average 50.2 years) participated in the present study. Their estimated audiometric thresholds are presented in Fig. 1. Audiological thresholds were assessed for octave-spaced frequencies between 0.25 and 8.0 kHz using custom software in MATLAB. Most listeners had hearing thresholds of ≤25 decibels hearing level (dB HL), though some older participants had higher thresholds at 8 kHz. Several older listeners had somewhat large asymmetries in hearing thresholds for 8 kHz. All procedures were approved by the University of Wisconsin-Madison Health Sciences Institution Review Board. Listeners completed informed consent before participation began. One participant was left-handed.

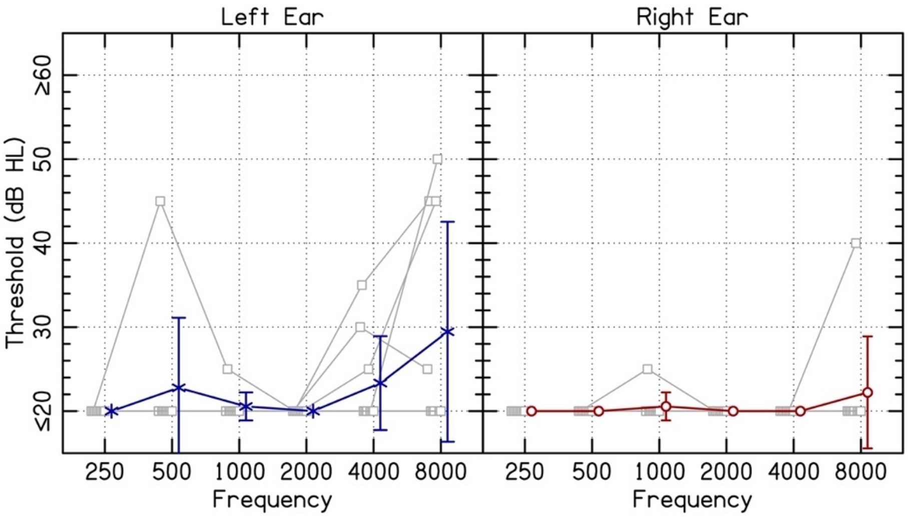


FIG. 1. Audiometric thresholds for the left and right ear. The panel on the left and right represent thresholds from the left and right ear, respectively. Average results are shown in blue or red, offset to the right, and error bars represent one standard deviation. Shapes offset to the left represent individual audiometric thresholds.

Testing was completed in the listeners' homes via home delivery by the experimenter. Equipment consisted of Sennheiser HD 280 Pro circumaural headphones, a Microsoft Surface tablet, a sound level data logger, and power supplies packaged into a small box. All testing was completed using automated, custom software written in MATLAB with the Microsoft Surface in kiosk mode. Kiosk mode with limited permissions was used to ensure that the listener could not see their data or use other software on the device. Stimuli were presented at a sampling rate of 44.1 kHz. Listeners were given written setup instructions and technical assistance was available from the experimenter via remote conference on video or telephone for the duration of the experiment. Before testing began, the sound level data logger was turned on to record the sound level in the room during testing in one minute increments.

B. Stimuli

Stimuli were a subset of monosyllabic words used in previous phonological fusion experiments [386,387]. They consisted of three sets of five words. Each set had a word with a stop consonant at the onset only (bed, pay, and go), two possible liquid consonants (/l/ and /r/), and both possible stop-liquid clusters (e.g., bled and bred).

The speech corpus was produced by one male speaker from the Midwest using standard American English. During the recording process, a metronome was used to assist in generating approximately 50 tokens of each word. Two of 50 tokens per word were selected such that the corpus had roughly similar duration and pitch. The duration and pitch were then manipulated in Praat until they were approximately equal. The resulting mean and standard deviation duration was 558 ± 37 ms. The resulting mean and standard deviation pitch was 101.9 ± 1.1 Hz and all stimuli fell within one semitone. Stimuli were recorded using an M-Audio Fast Track Pro interface and AKG C5900 microphone with pop filter. Stimuli were root-mean-square (RMS) level normalized.

Stimuli were vocoded in Praat with software available online[1]. Briefly, stimuli were bandpassed filtered into 16 frequency bands spaced between 250 and 8000 Hz. Frequency bands were evenly spaced and occupied equivalent cochlear space according to the Greenwood function [336]. The temporal envelope was extracted using half-wave rectification and a 300-Hz low-pass filter. The dynamic range of the temporal envelope was manipulated by compressing the extracted envelope to some percentage of its original value in dB. For example, if the dynamic range was equal to 60%, and the stimulus normally had a dynamic range of 30 dB (from a minimum level of 40 dB to maximum level of 70 dB), then the new dynamic range would be 18 dB (with a minimum level of 46 dB and a maximum of 64 dB). Therefore, the overall level remained equal between dynamic range conditions while the maximum and minimum levels within each band would be decreased and increased, respectively. Low-noise noise bands [388] with bandwidths equal to the filter bandwidth were then multiplied by resulting envelopes and summed across each frequency band. The resulting stimulus was RMS normalized to be equal in level to pink noise at 65 decibels sound pressure level, A-weighted [dB(A)].

C. Procedures

Listeners first confirmed that their computer set up looked like the instructions provided. Then, they confirmed that their headphones were on the correct sides of the head. This was completed by asking the listener the side on which a 250-Hz tone of 1000-ms duration was presented, with it being equally probable of being in the left or right earphone (one-interval, two-alternative forced-choice). The stimulus had 10-ms raised cosine onset- and offset-ramps. Two trials were presented at 70 dB(A). If the listener made two errors, they were instructed to reverse the earphones. If the listener made one error, they were instructed to try again. If the listener made no errors, they were allowed to continue. Listeners completed this check again before each block of experimental trials.

In the second equipment check, listeners confirmed that they were listening via headphones and not loudspeakers using a similar task to that used by Woods and colleagues

[389]. In this task, listeners were presented with three, 250-Hz tone bursts of 1000-ms duration with 500-ms inter-stimulus intervals. Their task was to choose the quietest burst (three-interval, two-alternative forced-choice). One tone burst was presented out of phase and at 70 dB(A). The levels of the other two tone bursts were 65 and 70 dB(A). Thus, if listeners were using loudspeakers and not headphones, destructive interference from the out-of-phase tone burst would reduce the sound level, making it the quietest. If instead listeners were using headphones, then the in-phase interval presented at 65 dB(A) would be the quietest. Six trials were completed. Listeners needed to achieve at least five out of six correct in order to progress to the next task. If they did not, they were asked to reconnect the headphones and the test was repeated.

Next, listeners an audiogram was collected. The task consisted of a presentation of one, two, or three pips of a pure tone with 10-ms cosine onset- and offset-ramps. Each pip had a duration of 300 ms separated by 200-ms inter-stimulus intervals. The task of the listener was to indicate the number of pips presented (three-alternative forced-choice). Testing followed standard step sizes of 10 dB-down and 5 dB-up, with a one-up, one-down adaptive rule. Levels were initiated at 70 dB HL for each frequency, beginning with 250 Hz in the left ear, increasing in frequency, and then progressing to the right ear. Threshold was determined by a listener achieving at least two out of three presentations at the same level correct.

Next, listeners completed familiarization and a series of pre-tests. Listeners were first allowed to listen to vocoded speech ad-hoc. A grid with the 15 stimuli in the corpus appeared on the screen. Listeners were allowed to play a word as many times as desired to the right ear. Stimuli were vocoded with 100% dynamic range. This was completed in order to gain some familiarity with vocoded speech. Then listeners were given a test where different tokens of each of the 15 words were presented one time to both ears and their task was to choose the word presented correctly (i.e., 15 alternative forced-choice). Different tokens were used so that listeners could not capitalize on arbitrary similarities between the ears due to using the same speech recording. Because they were also processed separately by the vocoder, the interaural correlation of the carriers in each frequency band was 0, resulting in a more diffuse sound image than if the interaural correlation were 1 [129]. The word was not allowed to be repeated, and listeners initiated the next trial when they were ready. A minimum criterion of 10 of 15 correct was enforced before listeners were allowed to progress to the next task. Then, the same test was given for stimuli vocoded with 40% dynamic range. No minimum criterion was established. Instead, the goal was simply for listeners to gain exposure to the easiest and most challenging stimuli presented during the experiment. The final pre-test consisted of pairs of stimuli with either the same word presented to each ear (10 trials) or words with different vowels presented to each ear (10 trials) using unprocessed (clean) speech. A minimum criterion of four one-word responses and four two-word responses was enforced before listeners were allowed to progress. If listeners failed to meet any criteria, they simply repeated the test until they successfully met the criteria.

Finally, listeners began experimental trials. Before each experimental block, listeners were informed that a longer block of testing was about to begin and that they could take a break if necessary. In experimental blocks, three types of trials were presented: (1) the same word using different tokens in each ear, (2) two words with different vowels in each ear, or (3) two rhyming words in each ear. Over the course of the experiment each word was tested 15 times in the “same word” trials (n=150) and each possible pairing was tested in the different vowel trials (n=150). The rhyming word trials consisted of two sub-types: phonological fusion and other trials. Phonological fusion trials consisted of a word beginning with a stop consonant and a word beginning with a liquid consonant, resulting in two pairs per set for the three sets, balanced so that each possible pairing was presented to the left and right ear, and each configuration was repeated five times (n = 2 × 3 × 2 × 5 = 60). Other trials consisted of non-fuseable pairs of rhyming words, with eight other pairings in each of the three sets, balanced so that each possible pairing was presented to the left and right ear, and each possible configuration was repeated two times (n = 8 × 3 × 2 × 2 = 96). Thus, conditions with the same vowel contained a similar number of trials (n=156). Listeners were tested with the following stimulus processing conditions: unprocessed, 100%:100%, 60%:60%, and 40%:40% interaurally symmetric dynamic ranges, and 60%:100%, 40%:100%, and 40%:60% interaurally asymmetric dynamic ranges. The ear with the smaller dynamic range was counterbalanced across participants. Each block had an equal number of trials from each vocoder and word-pair conditions, which consisted of 315 trials for the first nine blocks and 357 trials on the final block, resulting in a total of 3192 trials. Testing was scheduled over a four hour period and was able to be completed by most listeners during that time, including equipment assembly and disassembly. Chance performance in the task was 1/120 as there were 120 unique response combinations (105 combinations of two words and 15 single-word responses).

D. Analysis

All analyses were completed using mixed-effects analysis of variance (ANOVA) models in version 3.5.1 of R using version 1.1-17 of the *lme4* package to generate models and version 3.0-1 of the *lmerTest* package to estimate degrees of freedom using the Kenward-Roger approximation [286]. Each model included a random intercept associated with the listener and a fixed-effect of vocoder condition. Including the former allowed variation in mean performance due to difference between listeners to be accounted for in the model and not be attributed to residual error. Unless otherwise noted, the vocoder condition was agnostic of the ear receiving smaller dynamic range. Post-hoc comparisons were completed using estimated marginal means with Tukey adjustments for multiple comparisons using version 1.3.0 of the *emmeans* package in R.

III. RESULTS

The goal of the present experiment was to delineate the effects of binaural speech fusion and auditory attention in simulations of BiCIs with listeners who are representative in age to a CI cohort. To address this goal, we created interaurally symmetric and asymmetric conditions with varying temporal fidelity. The results are separated into three sections based upon the speech presented to the listener: (1) same word, (2) words with different vowels, and (3) rhyming/phonological fusion pairs.

A. Same word trials

Accuracy identifying the word presented and the number of words responded is shown in Fig. 2. Accurate responses were defined as those that included only one word and where that word corresponded to the word presented. The results of the ANOVA demonstrated significant effects of vocoder condition on percent correct [$F(6,48)=91.916$, $p<.0001$] and percentage of one-word responses [$F(6,48)=6.100$, $p<.0001$]. In the interaurally symmetric conditions in Fig. 2(A), the percentage of correct responses decreased as the temporal fidelity decreased. Post-hoc comparisons showed that there were significant differences between the unprocessed and 60%:60% [$t(48)=8.326$, $p<.0001$], unprocessed and 40%:40% [$t(48)=18.644$, $p<.0001$], 100%:100% and 60%:60% [$t(48)=7.136$, $p<.0001$], 100%:100% and 40%:40% [$t(48)=17.453$, $p<.0001$], and 60%:60% and 40%:40% [$t(48)=10.318$, $p<.0001$] conditions. We wanted to rule out the possibility that listeners began to perceive multiple words or guess more often as dynamic range was decreased. The percentage of one word responses demonstrates that listeners were more likely to report hearing two words at the smallest dynamic ranges. Post-hoc comparisons showed significant differences between the unprocessed and 40%:40% [t(48)=4.791, p<.001] and 100%:100% and 40%:40% [t(48)=4.275, p<.01].

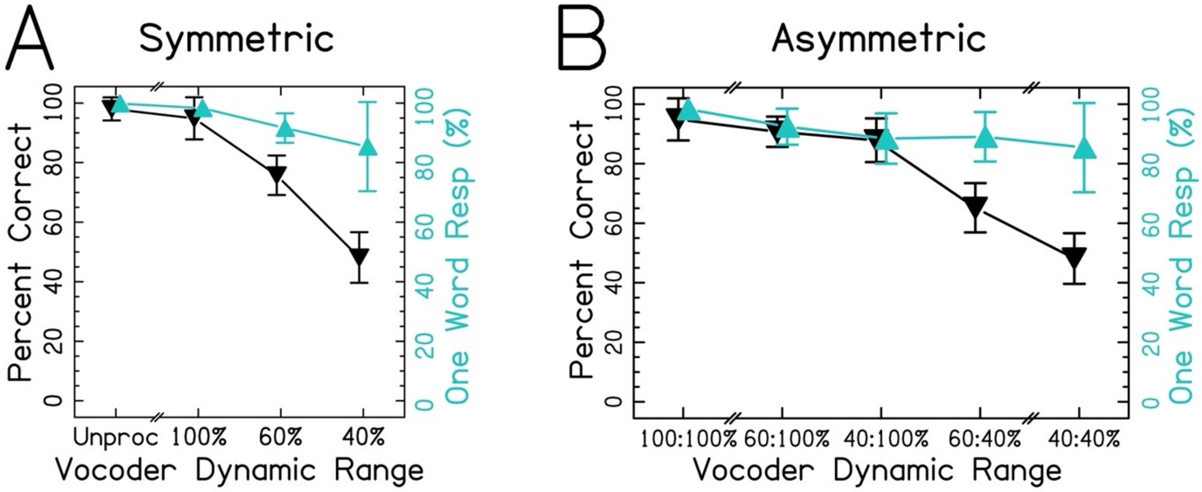
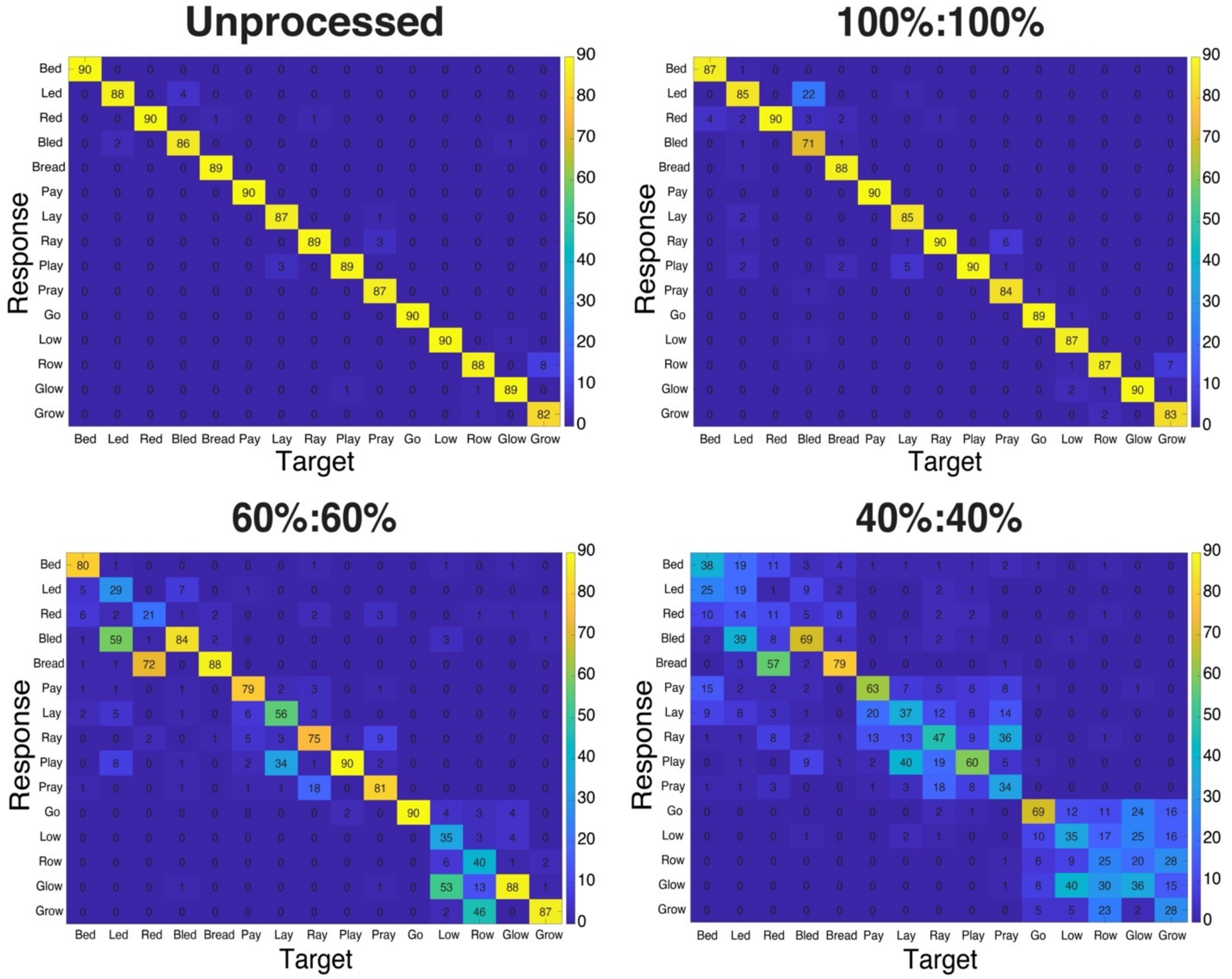


FIG. 2. Single word accuracy and number of words responded for (A) interaurally symmetric, and (B) interaurally asymmetric vocoding conditions. The x-axis corresponds to the vocoder condition. The y-axis represents the percentage of trials with at least one word accurately identified (shown in black) and the percentage of one-word responses (shown in teal).

The errors in interaurally symmetric conditions were further considered in Fig. 3, which shows confusions. Trials including one- and two-word responses were both included. The results suggest that, for 60%:60% dynamic range, listeners were more likely to confuse speech beginning with a liquid with stimuli containing a stop-liquid cluster with the same liquid. The same pattern of errors can be seen for 40%:40% dynamic range, though listeners were more likely to more likely to confuse all words sharing the same vowel, and more likely to make a vowel confusion. Additionally, listeners in the 40%:40% dynamic range conditions seemed to report words beginning with a liquid when a stop only was presented. Patterns of confusions will be explored further in the discussion.



Unprocessed

	Bed	Led	Red	Bled	Bread	Pay	Lay	Ray	Play	Pray	Go	Low	Row	Glow	Grow
Bed	90	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Led	0	88	0	4	0	0	0	0	0	0	0	0	0	0	0
Red	0	0	90	0	1	0	0	1	0	0	0	0	0	0	0
Bled	0	2	0	86	0	0	0	0	0	0	0	0	0	1	0
Bread	0	0	0	0	89	0	0	0	0	0	0	0	0	0	0
Pay	0	0	0	0	0	90	0	0	0	0	0	0	0	0	0
Lay	0	0	0	0	0	0	87	0	0	1	0	0	0	0	0
Ray	0	0	0	0	0	0	0	89	0	3	0	0	0	0	0
Play	0	0	0	0	0	0	3	0	89	0	0	0	0	0	0
Pray	0	0	0	0	0	0	0	0	0	87	0	0	0	0	0
Go	0	0	0	0	0	0	0	0	0	0	90	0	0	0	0
Low	0	0	0	0	0	0	0	0	0	0	0	90	0	1	0
Row	0	0	0	0	0	0	0	0	0	0	0	0	88	0	8
Glow	0	0	0	0	0	0	0	0	1	0	0	0	1	89	0
Grow	0	0	0	0	0	0	0	0	0	0	0	0	1	0	82

100%:100%

	Bed	Led	Red	Bled	Bread	Pay	Lay	Ray	Play	Pray	Go	Low	Row	Glow	Grow
Bed	87	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Led	0	85	0	22	0	0	1	0	0	0	0	0	0	0	0
Red	4	2	90	3	2	0	0	1	0	0	0	0	0	0	0
Bled	0	1	0	71	1	0	0	0	0	0	0	0	0	0	0
Bread	0	1	0	0	88	0	0	0	0	0	0	0	0	0	0
Pay	0	0	0	0	0	90	0	0	0	0	0	0	0	0	0
Lay	0	2	0	0	0	0	85	0	0	0	0	0	0	0	0
Ray	0	1	0	0	0	0	1	90	0	6	0	0	0	0	0
Play	0	2	0	0	2	0	5	0	90	1	0	0	0	0	0
Pray	0	0	0	1	0	0	0	0	0	84	1	0	0	0	0
Go	0	0	0	0	0	0	0	0	0	0	89	1	0	0	0
Low	0	0	0	1	0	0	0	0	0	0	0	87	0	0	0
Row	0	0	0	0	0	0	0	0	0	0	0	1	87	0	7
Glow	0	0	0	0	0	0	0	0	0	0	0	2	1	90	1
Grow	0	0	0	0	0	0	0	0	0	0	0	0	2	0	83

60%:60%

	Bed	Led	Red	Bled	Bread	Pay	Lay	Ray	Play	Pray	Go	Low	Row	Glow	Grow
Bed	80	1	0	0	0	0	0	1	0	0	0	1	0	1	0
Led	5	29	0	7	0	1	0	0	0	0	0	0	0	0	0
Red	6	2	21	1	2	0	0	2	0	3	0	0	1	1	1
Bled	1	59	1	84	2	0	0	0	0	0	0	3	0	0	1
Bread	1	1	72	0	88	0	0	0	0	0	0	0	0	0	0
Pay	1	1	0	1	0	79	2	3	0	1	0	0	0	0	0
Lay	2	5	0	1	0	6	56	3	0	0	0	0	0	0	0
Ray	0	0	2	0	1	5	3	75	1	9	0	0	0	0	0
Play	0	8	0	1	0	2	34	1	90	2	0	0	0	0	0
Pray	1	0	0	1	0	1	1	18	0	81	0	0	0	0	0
Go	0	0	0	0	0	0	0	0	2	0	90	4	3	4	0
Low	0	0	0	0	0	0	0	0	0	0	0	35	3	4	0
Row	0	0	0	0	0	0	0	0	0	0	0	6	40	1	2
Glow	0	0	0	1	0	0	0	0	0	1	0	53	13	88	1
Grow	0	0	0	0	0	0	0	0	0	0	0	2	46	0	87

40%:40%

	Bed	Led	Red	Bled	Bread	Pay	Lay	Ray	Play	Pray	Go	Low	Row	Glow	Grow
Bed	38	19	11	3	4	1	1	1	1	2	1	0	1	0	0
Led	25	19	1	9	2	0	0	2	1	0	0	0	0	0	0
Red	10	14	11	5	8	0	0	2	2	0	0	0	1	0	0
Bled	2	39	8	69	4	0	1	2	1	0	0	1	0	0	0
Bread	0	3	57	2	79	0	0	0	0	1	0	0	0	0	0
Pay	15	2	2	2	0	63	7	5	6	8	1	0	0	1	0
Lay	9	6	3	1	0	20	37	12	6	14	0	0	0	0	0
Ray	1	1	8	2	1	13	13	47	9	36	0	0	1	0	0
Play	0	1	0	9	1	2	40	19	60	5	1	0	0	0	0
Pray	1	1	3	0	0	1	3	18	8	34	0	0	0	0	0
Go	0	0	0	0	0	0	0	2	1	0	69	12	11	24	16
Low	0	0	0	1	0	0	2	1	0	0	10	35	17	25	16
Row	0	0	0	0	0	0	0	0	0	1	6	9	25	20	28
Glow	0	0	0	0	0	0	0	0	0	1	8	40	30	36	15
Grow	0	0	0	0	0	0	0	0	0	0	5	5	23	2	28

FIG. 3. Single word confusion matrices. Each panel corresponds to a different vocoder condition

given at the top. The x-axis represents the word presented. The y-axis represents the word(s) responded. Responses for all individuals are shown for all listeners. The color and number within each cell represents the number of responses given that correspond to that target and response.

Figure 2(B) shows accuracy for the interaurally asymmetric conditions, bounded by interaurally symmetric conditions with the largest and smallest dynamic ranges. The results indicate that listeners had high accuracy even when there were large asymmetries (i.e., 100%:40%) relative to conditions with small, symmetric dynamic ranges. Post-hoc comparisons showed that there were significant differences between the 100%:100% and 40%:60% [t(48)=11.089, p<.0001], 60%:100% and 40%:60% [t(48)=9.546, p<.0001], 60%:100% and 40%:40% [t(48)=15.910, p<.0001], 40%:100% and 40%:60% [t(48)=9.546, p<.0001], 40%:100% and 40%:40% [t(48)=14.831, p<.0001], and 40%:60% and 40%:40% [t(48)=6.364, p<.0001]. The percentage of one-word responses was similar to that observed for interaurally symmetric dynamic ranges, with percentages higher than the percent correct word identification. Post-hoc comparisons showed significant differences between only 100%:100% and 40%:100% [t(48)=3.268, p<.05]. Collectively, results from symmetric and asymmetric vocoding conditions suggest that speech understanding worsened with smaller dynamic ranges, remained high with at least one “good” ear, and that this worsening of performance was not due to incorrect segregation of speech in both ears.

B. Different vowel trials

Accuracy identifying the word presented and the number of words responded are shown in Fig. 4. In this case, the accuracy represents the probability of getting at least one word correct.

Presenting two words simultaneously was more cognitively demanding. Accordingly, the percentage of trials with both words reported correctly was substantially lower than those with at least one word correct. This increased difficulty was more likely to elicit ear advantage, where attention to one ear was prioritized. Thus, we wanted to determine whether information in one ear was preserved. The results of the ANOVA demonstrated significant effects of vocoder condition on percent correct [F(6,48)=38.084, p<.0001] and percentage of two-word responses [F(6,48)=19.601, p<.0001]. In the interaurally symmetric conditions in Fig. 4(A), the percentage of correct responses decreased as the temporal fidelity decreased in a similar fashion and to a greater extent than single word trials. Post-hoc comparisons showed significant differences between the unprocessed and 60%:60% [t(48)=5.391, p<.0001], unprocessed and 40%:40% [t(48)=11.998, p<.0001], 100%:100% and 60%:60% [t(48)=4.895, p<.001], 100%:100% and 40%:40% [t(48)=11.501, p<.0001], and 60%:60% and 40%:40% [t(48)=6.607, p<.0001] vocoder conditions. Interestingly, in contrast to single word trials, there was high correspondence between the number of words reported and the accuracy in reporting at least one word correct. Post-hoc comparisons showed significant differences between the unprocessed and 60%:60% [t(48)=3.387, p<.05], unprocessed and 40%:40% [t(48)=8.710, p<.0001], 100%:100% and 40%:40% [t(48)=8.392, p<.0001], and 60%:60% and 40%:40% [t(48)=5.323, p<.001] vocoder conditions.

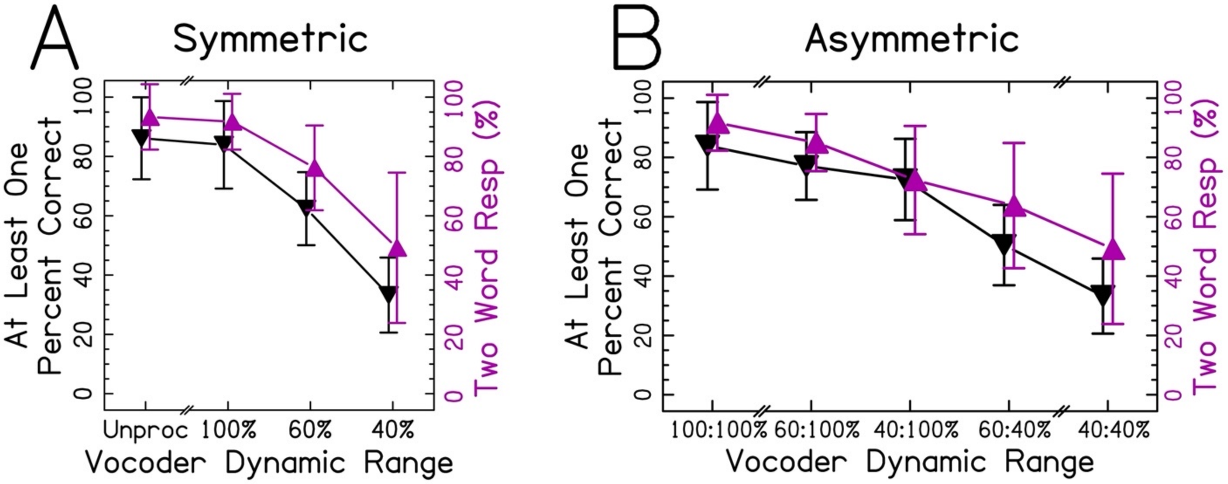


FIG. 4. Different vowel accuracy and number of words responded for (A) interaurally symmetric, and (B) interaurally asymmetric vocoding conditions. The x-axis corresponds to the vocoder condition. The y-axis represents the percentage of trials with at least one word accurately identified (shown in black) and the percentage of two-word responses (shown in purple).

We wanted to titrate the types of errors made in each symmetric vocoder condition. There were a total of 15 words × 10 possible pairings, or 150 combinations as well as 15 single-word response possibilities. Thus, a confusion matrix would be difficult to show with every possible combination. Instead, capitalizing on the small number of vowel errors made with single word trials, Fig. 5 shows vowel confusion matrices for interaurally symmetric vocoder trials. There were only three possible vowel combinations on each trial, but single vowel responses were also considered. As can be seen from Fig. 5, vowel errors were very rare. When listeners reported a single vowel, this usually corresponded to one of the vowels presented in one ear. The /ɛ/ and /eɪ/

pairs were the most likely to result in singular vowel responses. This may have to do with the fact that the /ɛ/ set had an additional /d/ cue at the end of each word.

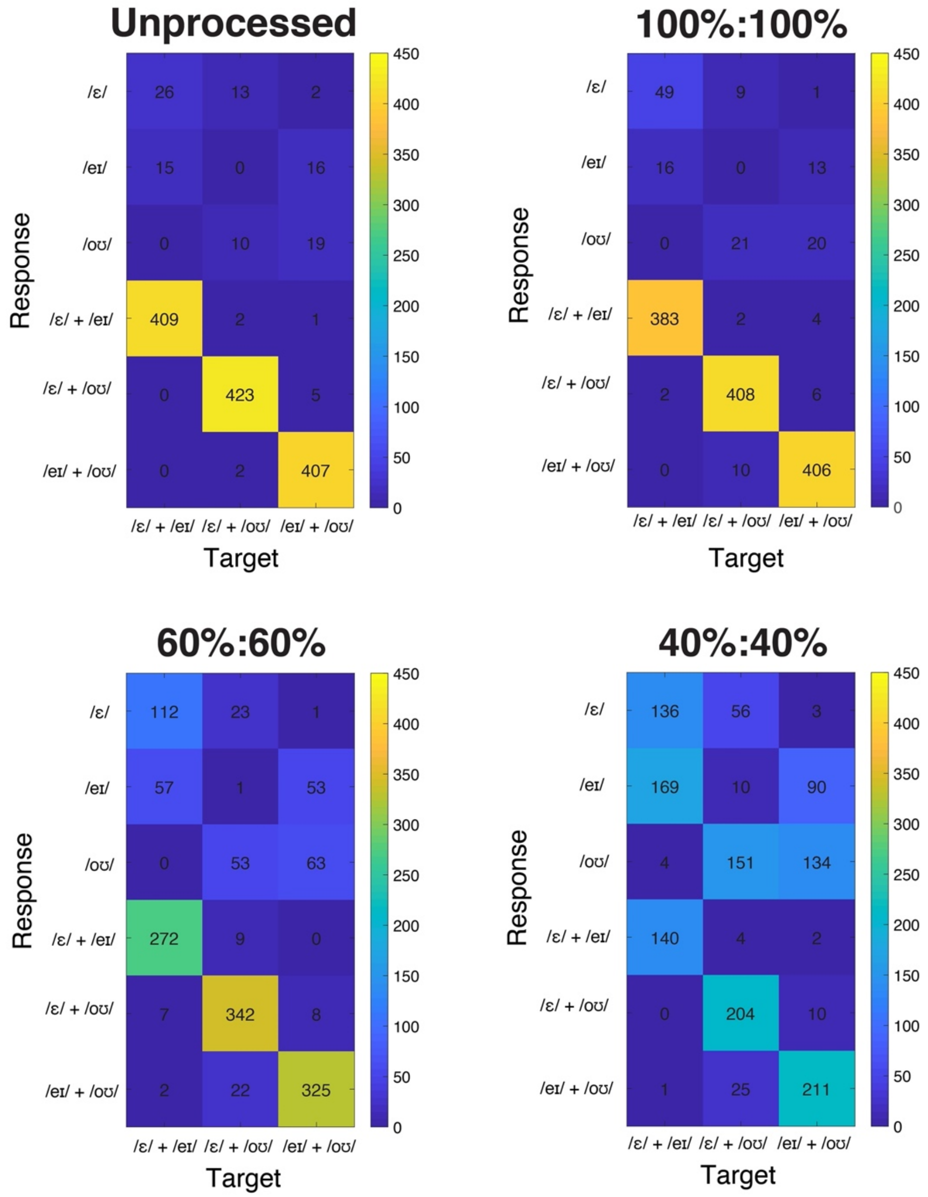


FIG. 5. Different vowel confusion matrices. Each panel corresponds to a different vocoder condition given at the top. The x-axis represents the vowels of the words presented. The y-axis

represents the vowel(s) of the word(s) responded. Responses for all individuals are shown for all listeners. The color and number within each cell represents the number of responses given that correspond to that target and response.

Figure 4(B) shows accuracy and two-word responses for the interaurally asymmetric conditions, bounded by interaurally symmetric conditions with the largest and smallest dynamic ranges. The results indicate that accuracy gradually decreased as the dynamic range in both ears decreased. Post-hoc comparisons showed significant differences between 100%:100% and 40%:60% [t(48)=7.600, p<.0001], 60%:100% and 40%:60% [t(48)=6.050, p<.0001], 60%:100% and 40%:40% [t(48)=9.951, p<.0001], 40%:100% and 40%:60% [t(48)=5.021, p<.001], 40%:100% and 40%:40% [t(48)=8.922, p<.0001], and 40%:60% and 40%:40% [t(48)=3.901, p<.01] dynamic range conditions. The percentage of two-word responses showed similar patterns. Post-hoc comparisons showed significant differences between 100%:100% and 40%:100% [t(48)=3.816, p<.01], 100%:100% and 40%:60% [t(48)=5.512, p<.0001], 60%:100% and 40%:60% [t(48)=4.194, p<.01], 60%:100% and 40%:40% [t(48)=7.074, p<.0001], and 40%:100% and 40%:40% [t(48)=4.575, p<.001] dynamic range conditions. Data from Fig. 4 show a slight amount of disagreement between accuracy and two-word responses in the condition with largest interaural asymmetry (40%:100%). This implies that listeners were more likely to report hearing one word, and more likely to have that word be accurate, when there were large degrees of interaural asymmetry. This is intuitive, since an effective strategy may be to ignore the poorer ear. Thus, it was of particular interest to explore the probability of responding with a correct response in the left or right ear.

Figure 6 shows the percentage of trials with the vowel corresponding to the left or right ear when only one word was responded. This decision was made based on Figs. 3 and 5, which showed that listeners were unlikely to make a vowel error. Thus, the vowel reported likely corresponds to the ear to which the listener was allocating attention. Fig. 6(A) shows a modest right ear vowel advantage across interaurally symmetric vocoder conditions, with the smallest dynamic ranges resulting in the greatest ear advantages. Fig. 6(B) and 6(C) show the same results for interaurally asymmetric vocoder conditions, with the better ear (i.e., ear with larger dynamic range) on the left and right, respectively. The results demonstrate that listeners were more likely to report a single vowel corresponding to the better ear, especially in the conditions with greatest interaural asymmetry and smallest overall dynamic range. Additionally, the results suggest that interaural asymmetric dynamic range interacts with right ear advantage, with smaller differences between left and right ear vowel responses when the better ear was on the left, and larger differences between left and right ear vowel responses when the better ear was on the right. Collectively, results from symmetric and asymmetric vocoding conditions suggest that speech understanding worsened with smaller dynamic range, became poor with at least one “poor” ear, and that worsening of performance was likely due to incorrect fusion of speech.

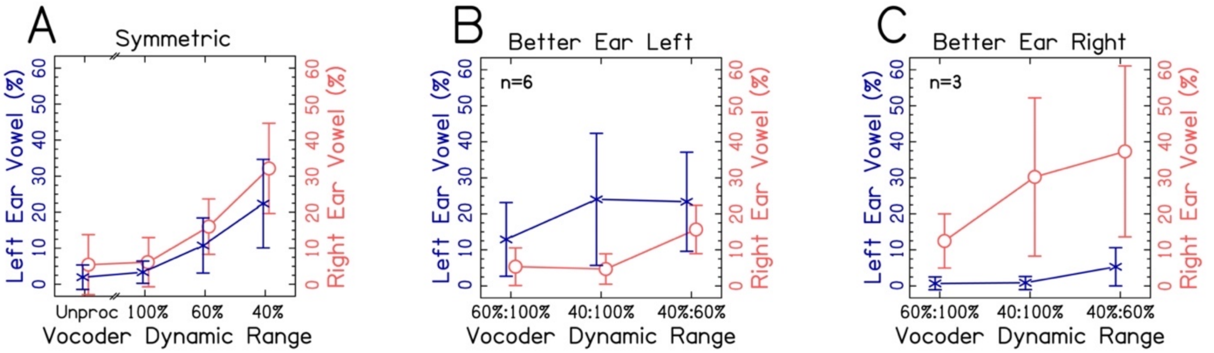


FIG. 6. Vowel responses by ear for different vowel trials for (A) interaurally symmetric or (B,C) interaurally asymmetric vocoder conditions. The x-axis corresponds to the vocoder condition. The y-axis corresponds to the percentage of trials where the vowel of the response was one word and came from the left (shown in blue) or right (shown in orange).

C. Rhyming word and phonological fusion trials

Figure 7 shows the percentage of trials with one or two word responses. The results of the ANOVA demonstrated significant effects of vocoder condition on the percentage of two-word responses [F(6,48)=5.987, p<.0001]. In the interaurally symmetric conditions in Fig. 7(A), the percentage of two-word responses decreased as the temporal fidelity decreased. Post-hoc comparisons showed significant differences between the unprocessed and 40%:40% [t(48)=4.072, p<.01] and 100%:100% and 40%:40% [t(48)=4.179, p<.01] vocoding conditions. The percentage of two-word responses tended to decrease as the dynamic range decreased in one or both ears. Post-hoc comparisons showed significant differences between the 100%:100% and

40%:100% [t(48)=3.327, p<.05] and 100%:100% and 40%:60% [t(48)=3.842, p<.01] dynamic range conditions.

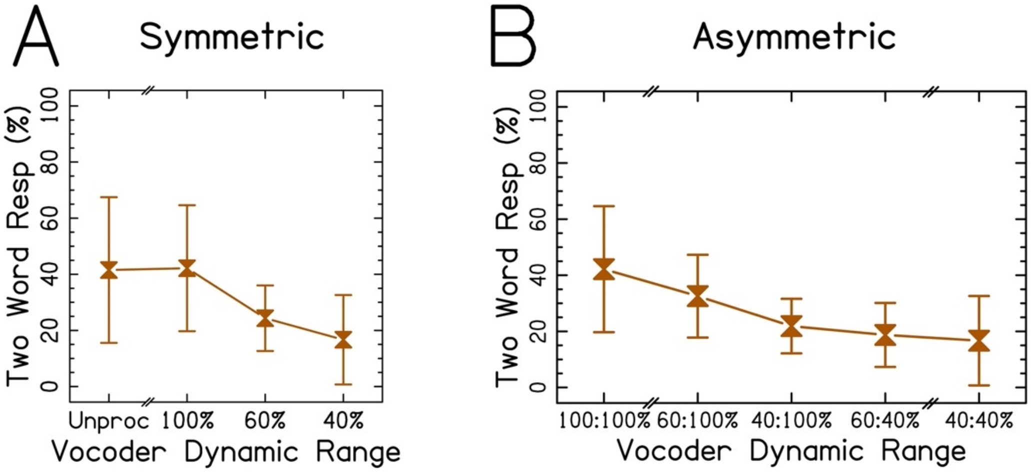


FIG. 7. Rhyming word number of words responded for (A) interaurally symmetric and (B) interaurally asymmetric vocoder conditions. The x-axis corresponds to the vocoder condition. The y-axis represents the percentage of trials with two-word responses (shown in purple).

Figure 8 shows an illustration of the phonological fusion trials and possible responses from listeners. On each trial, a word beginning with a stop consonant was presented to one ear and another starting with a liquid consonant was presented to the other ear. The words shared the same ending phonemes, and could be combined to generate a word in English. Thus, five categories of response could be indicated by listeners' responses (Fig. 8). "Ideal" responses represent the best case scenario, where a listener correctly indicated both words. The next best case scenario would be to respond with a "fused" response, where listeners still make use of the speech in both ears, but it is combined into a single word. The next best case scenario would be to ignore or prioritize speech in one ear, indicated by the biased left and right responses. Finally, the worst case scenario would occur when speech in both ears interferes with one another, preventing listeners from accessing speech in either.

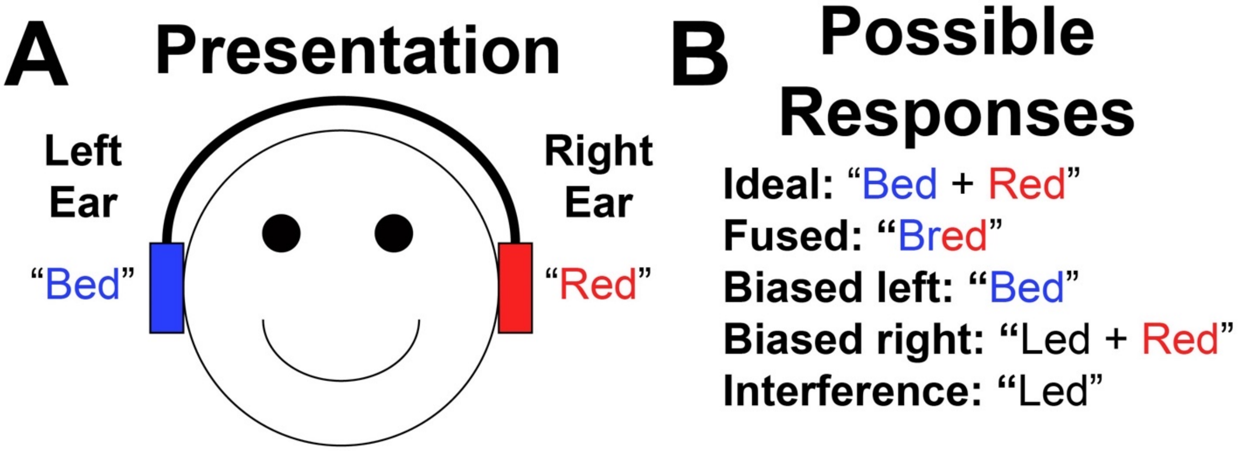


FIG. 8. Illustration of phonological fusion trials. (A) One example trial illustrating how stimuli were presented to listeners in a stop and liquid pair. (B) Possible response categories arranged from best (ideal) to worst (interference) performance.

Figure 9 shows the distribution of response categories depending for each vocoder condition. The results of the ANOVA demonstrated significant effects of vocoder condition on the percentage of ideal [F(3,24)=, p<.0001] and interference [F(3,24)=, p<.0001] responses. In the interaurally symmetric conditions in Fig. 9(A), the percentage of ideal responses decreased as the temporal fidelity decreased. Post-hoc comparisons showed significant differences in the percentage of ideal responses between the unprocessed and 60%:60% [t(24)=4.594, p<.001], unprocessed and 40%:40% [t(24)=5.886, p<.0001], 100%:100% and 60%:60% [t(24)=5.059, p<.001], and 100%:100% and 40%:40% [t(24)=6.351 , p<.0001] vocoder conditions. Similarly, post-hoc comparisons showed significant difference in the percentage of interference responses between the unprocessed and 40%:40% [t(24)=-7.610, p<.0001], 100%:100% and 40%:40% [t(24)=-6.847, p<.0001], 60%:60% and 40%:40% [t(24)=-5.186, p<.001] vocoder conditions.

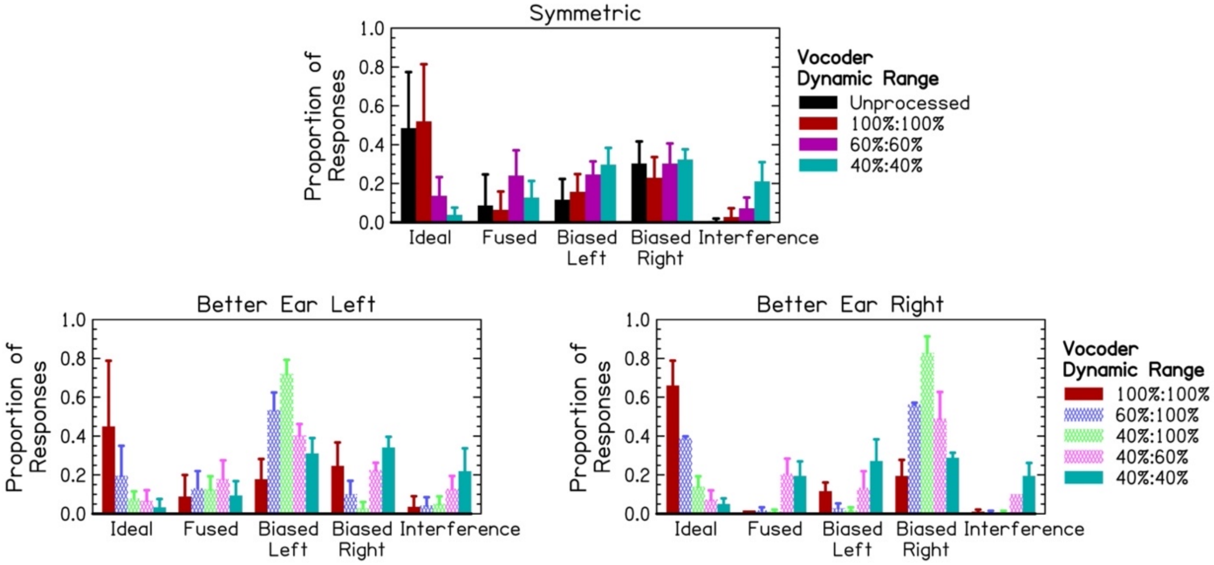


FIG. 9. Response categories by vocoder condition for interaurally symmetric (top row) and interaurally asymmetric (bottom row) vocoder conditions. The x-axis corresponds to the response category from Fig. 8. The y-axis corresponds to the proportion of responses. The color and pattern represents the vocoder condition given in the figure legend.

The bottom row of Fig. 9 shows the distributions of response categories for interaurally asymmetric vocoder conditions separated by better ear (i.e., ear with larger dynamic range), bounded by interaurally symmetric conditions with the largest and smallest dynamic ranges. Because of the small sample size in each asymmetric group (n=6 and n=3 for better ear on the left and right, respectively), no statistical comparisons were made. Some trends are apparent, however. When listeners were presented with a greater dynamic range on the left or right, this was associated with a greater bias toward responding with the stimulus presented to left or right ear, respectively. This effect was most pronounced when one ear was presented with 40% dynamic range and the other with 100% dynamic range. Intereference responses tended to increase with decreasing dynamic range in one or both ears, and seemed to be slightly more pronounced when the left ear contained the greater dynamic range.

IV. DISCUSSION

Patients with BiCIs often experience substantial differences in hearing outcomes associated with each ear. While bilateral implantation generally improves speech understanding in noise relative to unilateral implantation, there are some conditions under which listeners experience contralateral interference from the better ear [13–15,23]. Two putative mechanisms have been proposed in the literature related to the basis of contralateral interference. The first supposes that it results from a failure of sound source segregation cues to result in distinct auditory objects, e.g., [20,24,26]. The second supposes that contralateral interference results from a failure to allocate attention away from the source with greatest spectro-temporal fidelity and toward a more degraded sound source, e.g., [15,19].

Results from the present experiment suggest that poor segregation of sound sources and compromised attention both play a role and may interact with one another when temporal information is symmetrically or asymmetrically degraded. Listeners showed an increased number of one-word responses when rhyming words or words with dichotic vowels were presented, especially in cases where stimuli were symmetrically or asymmetrically temporally degraded. As stimuli became increasingly degraded in one or both ears, listeners also tended to report inaccurate word(s), even though results suggested that listeners shifted attention toward the right ear when stimuli were symmetrically degraded and toward the better ear when stimuli were asymmetrically degraded. These results are considered in greater context in the sections that follow.

A. Object-based auditory attention

Auditory attention is thought to be a process with serial and parallel processing stages [149,209]. First, similar spectral components are grouped into auditory streams. These streams compete for attention, which is modulated via source segregation cues. Finally, information can be extracted from sources of interest (e.g., language content). This is a relatively simple process for stimuli that occur in quiet with a definitive onset and offset time. However, real-world listening often occurs in complex auditory environments with competing background noise and ambiguous onset and offset times. Thus, it is suspected that listeners maintain an internal perceptual model that can be updated according to new sensory information e.g., [145]. Increasingly complex stimulus features are suspected to be extracted at later stages of sensory processing. These features are the product of interactions between internal predictions and incoming sensory input. Thus, for listeners who receive compromised or ambiguous sound source segregation cues, it is likely that the ability to maintain internal predictions, represent sound features, and allocate attention would be compromised.

Listeners with BiCIs show considerable evidence of a likelihood to fuse unrelated auditory information. Listeners perceive a singular pitch percept over a large disparity of electrodes between ears, corresponding to frequency differences up to one octave [156]. Abnormally large fusion ranges for interaural place-of-stimulation differences has been proposed by [196] as an adaptive process associated with the large degrees of mismatch associated with differences in degree of insertion for listeners with BiCIs [43]. Simulations in NH suggest that interaural mismatches in place-of-stimulation result in poorer spatial fusion [163] and speech fusion [390,391]. Accordingly, listeners with BiCIs perceive a singular spatial image over large interaural electrode disparities, even as the impact of binaural cues on perceived intracranial perception decreases [138,139,159–161]. Listeners with BiCIs will fuse stimuli presented with ITDs up to several milliseconds [159,160] and very large ITDs (~2 ms) are needed to achieve maximum intracranial lateralization [46,93,162].

Fusion experiments generally ask listeners to report the number of sounds perceived. One alternative approach is to assess a listener's ability to discriminate between sounds suspected to be perceptually segregated. In this case, segregation is not necessary to complete the task. Modifying the paradigm used by [273], Chapter 3 assessed BiCI listeners' ability to identify cases when the rate of temporal fluctuations differed between electrodes. The authors showed that listeners with BiCIs tend to be less accurate than those with NH, and highly biased in their responses toward responding with "same" or "different." This result implies a task-relevant bias of listeners. That is, if listeners are asked to respond with one or two sounds, they might be more biased toward responding one way based upon the task. This conclusion is supported by inconsistent bias toward responding "same" or "different" in listeners with NH when temporal fidelity was reduced. Then the high degrees of fusion in experiments might simply be indicative of poorer perceptual boundaries between features of the target and masking stimulus rather than a likelihood of perceiving one auditory object. For both the simulation in NH [Chapter 2] and experiment in listeners with BiCIs [Chapter 3], the poorer ear predicted performance. Spatial cues also have a lesser impact on perception in simulations of interaurally asymmetric temporal fidelity with NH [Chapter 4] and studies of listeners with BiCIs [9][Chapter 5]. Thus, it may be that temporal degradations in one ear are sufficient to interfere with segregation of sound sources.

The present study simulated compromised temporal fidelity by decreasing the dynamic range of stimuli. Reduced dynamic range is associated with poorer speech understanding in listeners with BiCIs [374,375] and simulations in NH [376]. It is also associated with poorer sensitivity to spatial cues in listeners with BiCIs [121,122] and listeners with NH [314][Chapter 4]. Reducing dynamic range is similar in spirit to the spectro-temporal smearing that is thought to occur in patients. Previous experiments addressing bilateral speech understanding and showing contralateral interference in simulations of CI processing have manipulated spectro-temporal fidelity by reducing the number of frequency bands [19,20]. The issue with this approach is that it is much less realistic. Increasing the number of maxima in peak-picking, N-of-M processing strategies beyond the eight maxima typically used in clinical practice improves speech understanding [133]. Reducing the number of spectral channels below eight is highly unlikely to be used in practice, and it is unlikely that listeners with BiCIs would be presented with such different numbers of frequency channels. It is possible that the number of “effective” channels is different between ears. The fewer than eight channels needed to demonstrate contralateral interference in NH stands in contrast to the benefits attained by patients who gain benefits with more than eight maxima in their programming strategy [133]. Recognizing that all vocoder experiments are highly artificial, reduced or asymmetric dynamic range may be a more realistic problem for patients.

B. Ear advantage

One interesting finding in the present study was that asymmetric temporal fidelity modulated the ear to which listeners attended. This is consistent with the results of [19]. Listeners with NH tend to show modest effects of right ear advantage, evidenced by greater accuracy or higher probability of reporting speech presented to the right ear compared to the left ear. This effect tends to become exaggerated as listeners get older [385].

Two classical hypotheses have been proposed associated with right ear advantage, based on structural biases in the left hemisphere [32] or biased attention [33]. For listeners with NH, the attentional hypothesis seems to provide a better explanation of patterns of performance [174]. It may be that shifts in attention also help to explain increased right ear advantage associated with aging. However, for listeners with neurological issues like those who have hearing loss, it may be that the structures conveying auditory information to either side of the brain are compromised. This conclusion is supported by one study showing that long periods of auditory deprivation in one ear are associated with speech understanding asymmetries and contralateral interference [13]. The present study and the study by Goupell and colleagues [19] simulated asymmetries in the spectro-temporal fidelity of sounds in listeners with NH, which resulted in a shift in attention toward the better ear. Thus, the conclusions of a structurally based ear advantage framework may be more appropriate for understanding interaural asymmetries in speech understanding for patients with BiCIs.

The results from single word trials (Fig. 3) suggest that ear advantage and interference effects observed in Fig. 9 could be due to consistent substitution errors (e.g., reporting a liquid when a stop and liquid were presented). In phonological fusion trials, this would be scored as a bias toward the left/right ear if the liquid matched the liquid presented, or an interference response if the liquid did not match that presented. Substituting a liquid for a stop-liquid cluster seems especially likely because of the manipulation used in the experiment. That is, decreasing the dynamic range would have smoothed the abrupt onset associated with a stop consonant. To address whether there was a consistent pattern of performance in both sets of trials, Fig. 10 shows the relationship between proportion of liquid responses when stops were presented for each listener in symmetric dynamic ranges in the same word and phonological fusion trials. The results show that smaller dynamic ranges resulted in a greater number of liquid-only responses when both a stop and liquid were presented. Together with Fig. 9, this suggests that listeners made a similar substitution error, reporting a liquid when both a stop and liquid were presented, especially at 100%:100% and 60%:60% dynamic range relative to the same word trials. Responding with a liquid for a stimulus containing a stop and liquid in same word trials was most common at 40%:40% dynamic range. While chance error was 0.83%, if listeners were able to understand the vowel and responded with one word, there was a 40% chance of guessing a liquid-only response. Figure 10 shows that several listeners responded with only liquids greater than 40% of the time at small dynamic ranges, suggesting that they demonstrated a consistent bias.

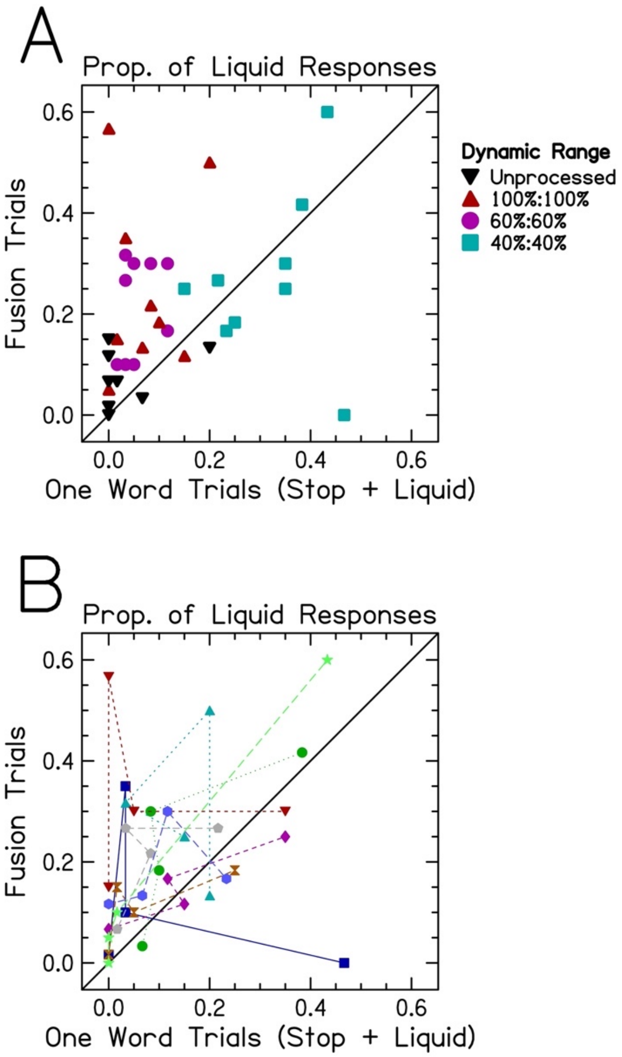


FIG. 10. Proportion of responses containing only liquid consonants for fusion and single word trials by (A) vocoder condition and (B) individual listener. The x- and y-axes correspond to the proportion of trials containing only liquid consonants for fusion and single word trials, respectively. (A) The color and shape represent the vocoder condition given in the figure legend. (B) The color and shape represent different listeners.

When symmetric dynamic range was changed from 100%:100% to 60%:60% or 40%:40%, it resulted in an approximately 35 or 45% decrease in the proportion of “ideal” responses, respectively (Fig. 9). Approximately 10% of this decrease can be explained by an increase in “fusion” responses, corresponding to a combination of the stop and liquid words. Another 10-15% can be explained by an increase in “biased left” responses. Another 10% can be explained by an increase in “biased right” responses. The remaining decrease in “ideal” responses is explained by an increase in “interference” responses, approximately twofold in 60%:60% and eightfold in 40%:40%. Combined with Fig. 10, these results suggest that during the fusion trials, listeners tended to report the liquid in the left or right ear slightly more often than the stop. Combined with Fig. 3, these results suggest that the most common error was for listeners to report the liquid consonant and not the stop from the word. However, most responses from listeners still contained a stop consonant. Thus, while reporting only the liquid was the most common error, this does not mean that listeners were unable to detect the presence of the stop consonant entirely.

C. Limitations

The greatest limitation to the generalizability of these results is probably the speech stimuli used, which were highly artificial. Speech tokens were presented with the same onset time, spoken by the same individual. It is important to note that in the original phonological fusion experiments, onset time played a large role in the perception of listeners. When the stop consonant preceded the liquid, listeners were even more likely to report a fused response [386]. In contrast, when the liquid preceded the stop, listeners were less likely to report a one-word response. Stimuli in the present experiment were presented entirely independent to each ear, rather than mixed as they would occur in the free-field. Other experiments exploring the effects of right ear advantage have varied relative level of the sound in each ear, showing changes in the percentage of correctly reported words [176,384]. Thus, varying the interaural level difference may be one way to help titrate the relative advantage listening with each ear. This is particularly relevant for listeners with BiCIs, who may show evidence of perceptual "extinction," a form of auditory neglect demonstrated by an inability to perceive sounds on one side when the opposite side is stimulated, e.g., [170]. This is evidenced by high levels of accuracy when speech is presented to either ear alone but large asymmetries when both ears are stimulated simultaneously [13,15]. Finally, because there were so few words in the present study, it is possible that some of these effects were driven by specific features. It may be that different phonemes are less or more likely to result in fusion and interference than those used here. It is highly likely that listeners used any features available to distinguish between words (e.g., /d/ at the end of the set with the /ɛ/ vowel).

Accordingly, the results of the present study should be treated as a proof of concept. With stimuli ideal to cause integration of auditory information, listeners were more likely report one word that did not correspond to the word presented to either ear. One strength of the phonological fusion procedure is that it provides a spectrum of performance to assess whether the information in either ear is used, integrated, ignored, or interferes with the other ear. Future experiments may be able to use more realistic stimuli to achieve similar goals.

The present experiment successfully collected high-quality data in a remote-testing context from older and younger listeners. However, there were several challenges that occurred during testing. Audiological data showed asymmetries for some of the older listeners, particularly at high frequencies. Fig. 11 shows the percent correct by block number over the course of the experiment. There was not a significant change in percentage of attaining at least one correct response [F(9,72)=1.7877, p=.085], but there was a significant effect of block number on percentage of responding with the number of words and words reported being correct [F(9,72)=3.602, p<.001]. The latter was driven by significantly worse performance for the last block compared to all but the first, seventh, and ninth blocks (p<.05). Because the experiment took so long, it is likely that listeners became tired or bored, which fit the anecdotal reports provided to the experimenter. Additionally, there were technical issues that had to be resolved during testing that would have been easier to address in the laboratory. Remote testing has unique challenges, but it may be an equitable path forward for working with populations who are unable to travel to the laboratory, and was effective for continuing to gather data during the COVID-19 pandemic.

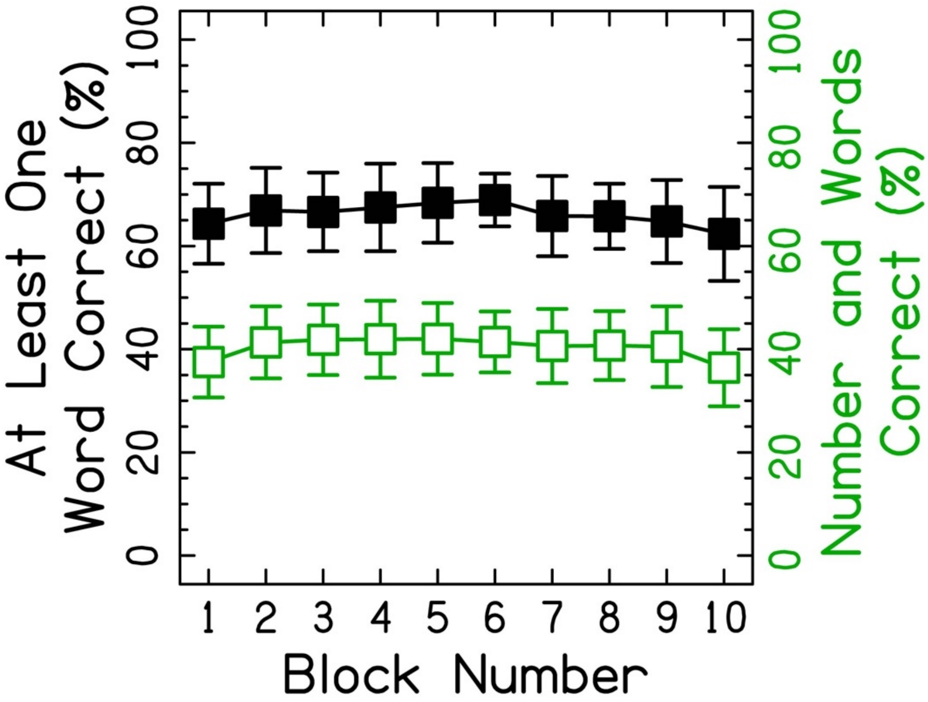


FIG. 10. Accuracy by trial block. The x-axis corresponds to the block number. The y-axis corresponds to the percentage of trials where at least one word was chosen correctly (shown in black) or the number of words and words chosen were correct (shown in green).

D. Summary and conclusions

Several conclusions can be drawn from the present experiment:

1. Listeners demonstrated poorer speech understanding with smaller, interaurally symmetric and asymmetric dynamic ranges (Fig. 2, 4), consistent with previous experiments in listeners with BiCIs [374,375] and simulations in NH [376].
2. Decreased accuracy was related to increased probability of one-word responses when speech with different vowels or rhyming words were presented to each ear (Fig. 4, 8).

3. In interaurally symmetric conditions when speech with different vowels were presented to each ear, listeners were more likely to report the stimulus in the right ear, especially when stimuli were temporally degraded (Fig. 6).
4. In interaurally asymmetric conditions when speech with different vowels were presented to each ear, listeners were more likely to report the stimulus in the better ear, which was especially pronounced if the better ear was on the right side (Fig. 6).
5. When phonologically fuseable pairs of words were presented to each ear, listeners were more likely to experience interference if one or both ears had small dynamic range (Fig. 9).

FOOTNOTES

1. Code is available at http://mattwinn.com/praat.html

Chapter 7: Conclusions and Discussion

I. IMPLICATIONS FOR RESEARCHERS AND CLINICIANS

Interaural asymmetry, revealed from differences in the spectro-temporal representations of sounds, psychophysical sensitivity, hemispheric representation, and speech understanding outcomes between the ears, is a highly prevalent problem for listeners with BiCIs. Differences between the auditory periphery of each ear lead to poorer encoding of the binaural cues used to distinguish between sound sources. Poor encoding leads to poorer auditory object formation, challenges allocating attention to sources of interest, and hemispheric reorganization toward the better ear. While the sources may overlap or vary, the manifestation is the same: one better performing ear and poorer binaural outcomes.

It is important to note that even listeners who demonstrate large interaural asymmetries in speech understanding generally do not experience a decrement from using both CIs compared to a unilateral CI in more realistic, free field speech tests [10]. Thus, bilateral implantation is unlikely to result in worse outcomes than unilateral implantation, but its benefits may be modulated by the factors discussed here. It is our hope that the framework and studies outlined in this review can be used to facilitate better patient outcomes and improvement of bilateral devices.

A. Interrelated sources of interaural asymmetry

While the literature has focused on specific conditions that induce interaural asymmetry, a central idea is that interaural asymmetry results from similarly manifested problems in the conceptual framework proposed here. Moreover, our goal is that this framework can be used to devise solutions for patients experiencing or likely to experience interaural asymmetry. While we aimed to address each topic separately, by no means do these sources of interaural asymmetry occur in isolation. For example, the number of surviving nerve fibers associated with etiology of hearing loss predicts the speech outcomes of listeners [109,197]. Patients with early onset of hearing loss and delayed implantation will experience deterioration of the auditory periphery, brainstem changes, hemispheric reorganization, and will also have less bilateral experience relative to age-matched peers. These listeners would be just as likely to experience scalar translocations and interaural place-of-stimulation mismatch.

The factor affecting interaural asymmetry that has received the most attention in the literature is auditory deprivation. There is consensus across clinicians and laboratory studies in animals and humans that providing consistent bilateral input, especially during development, is important for binaural outcomes [4,5,103,189]. Auditory deprivation is associated with poorer temporal response properties beginning at the level of the auditory periphery and a development of preferential bilateral processing of the better ear. Recent research shows promising results that processing and decrements in performance may be improved by experience with accurate binaural cues [80,98,99], but it would be ideal if this can be avoided in the first place. If not, then it would be ideal to incorporate auditory training to leverage any remaining plasticity.

Some issues associated with the auditory periphery may be improved with technological advances and improvements in surgical techniques. For example, robot assisted electrode insertion significantly reduces the number of translocations and amount of mechanical trauma to the cochlea during surgery [392]. The type of CI array also contributes to outcomes, with pre-curved arrays being more likely to result in translocation [43] and resulting in the smallest distance from the modiolus [130]. Deactivating electrodes suspected to degrade spectro-temporal representations by smearing their fluctuations improves speech understanding [191,264,291,365,377]. With new imaging approaches, it may also be possible to match interaural placement of CI arrays [361], or at least reallocate frequencies to electrodes based on similar interaural place-of-stimulation [43].

Differences in loudness growth or dynamic range have not been systematically addressed in the literature. It has been hypothesized that loudness growth differences result in increased interaural decorrelation [140,229], implying that resolving differences in loudness growth could result in considerable improvement of patient outcomes. This is one of the most promising avenues for future research. Similarly, linking the automatic gain control in each ear may improve binaural outcomes [308,309] and reduce the frequency of spurious spatial cues represented by current clinical processors [36]. Recent strategies have been devised to provide fine-structure ITDs, but demonstrate no consistent benefits on spatial hearing tasks [37–39]. Improvements to processors resulting in coherent and consistent cues may be most beneficial early in BiCI experience or may increase with training.

B. Assessing interaural asymmetry

Depending on the time and equipment available to researchers and clinicians, it may not be possible to intensively assess the sources of interaural asymmetry for listeners. It also may be difficult to determine at what level interaural asymmetry should be assessed. One particularly helpful trend in the literature is a systems-based approach, where the output of the "system" (i.e., behavioral responses) is used to make assumptions about the relevant underlying problems. One example includes measuring temporal sensitivity in each ear or electrode and predicting monaural or binaural outcomes (e.g., Chapter 3, Chapter 5, [9,191,192]). Another example is the various forms of the spectral ripple test, which is meant to provide a proxy of spectro-temporal resolution and has mainly been used to predict unilateral speech understanding (e.g., [133,298]). This could be a cost-effective and efficient avenue for researchers, clinicians, and patients.

This systems-based approach is much more straight forward for issues with encoding. For decoding, a few additional considerations should be made. Most importantly, some listeners with BiCIs do not demonstrate interaural asymmetries until they are given a sufficiently difficult task or are stimulated in both ears [13,15]. Measures need to be taken to demonstrate that the issues assumed to be associated with decoding are not in fact an encoding problem in disguise. Studies of contralateral masking and interference provide a useful example of one means to do this, by first measuring speech understanding in each ear alone, then comparing against results when both ears are stimulated. Finally, like the interrelated sources of interaural asymmetry, encoding and decoding problems are likely to manifest together since decoding's features depend upon encoding's cues. A disproportionate decrement in performance with no more than a slight challenge to the listener may be indicative of decoding problems compounding encoding problems.

C. Worse ear or degree of asymmetry?

One important topic that has not been addressed so far is whether poorer binaural outcomes associated with interaural asymmetry are a result of irreconcilable differences between two different ears or dominance of the worse ear. Results from patients are mixed, suggesting that the worse ear may be predictive of poorer sensitivity to binaural cues [9] [Chapter 5] or that the degree of difference and worse ear are similarly predictive [Chapter 3]. The best way to address this problem would be to test listeners who have good or poor, symmetric or asymmetric hearing outcomes. In practice, most participants in laboratory studies are high performers, so they tend to have at least one "good" ear. Thus, to investigate this question we conducted a series of studies simulating interaurally symmetric and asymmetric conditions with high and low temporal fidelity and measured lateralization, speech perception, and speech unmasking. In lateralization experiments, the worse ear predicted less utility of ITDs and results did not differ when both ears had low temporal fidelity [Chapter 4]. In speech experiments, interaural asymmetry and the average degree of temporal fidelity negatively affected outcomes [Chapter 6], suggesting that both play a role. In particular, it seems that interaural asymmetry may direct attention toward the better ear, ignoring the worse ear, and interfering with performance [19] [Chapter 6]. This may be facilitated by a fusion of sounds when one or both ears have low temporal fidelity [Chapter 2, Chapter 3, Chapter 6].

D. Challenges and suggested future directions

This dissertation attempts to provide a simple conceptual framework to bring together research on interaural asymmetry conducted using human or animal behavior, physiology, and computational modeling. This approach has strengths and weaknesses. The most significant weakness is the idea that encoding and decoding are independent of one another and occur in distinct physiological structures along the auditory pathway. As research on the efferent auditory system develops, it is becoming increasingly clear that encoding and decoding occur in parallel, and that encoding is shaped by ongoing decoding. One circuit or set of circuits will not be able to explain all of auditory or more general perceptual processing. Thus, the framework proposed here should be treated as a means to make predictions and devise solutions for patients, not as a model of all auditory perceptual phenomena.

One of the largest challenges associated with addressing interaural asymmetry is that the prevalence of different kinds of asymmetry is often unknown or ill-defined. A main argument in this dissertation is that interaural asymmetry is not a singularly dimensional spectrum from symmetric to asymmetric, or a dichotomous symmetric/asymmetric state. Instead, it represents a collection of states or continua (e.g., health of the auditory nerve in each ear, distance between

CI electrodes and auditory nerve fibers, interaural place-of-stimulation mismatch, degree of cortical lateralization) that generate similar outcomes (e.g., differences in speech understanding, difficulty using spatial cues, challenges segregating speech from noise). These different “types” of interaural asymmetry could be described using similar manifestations of the conceptual models presented here. A natural conclusion of such a framework is that no listener had purely interaurally symmetric hearing, supported by research showing modest right ear advantage in listeners with NH where interaural asymmetries are minimal and do not seem to impact NH listeners’ daily experience. The question that researchers should ask instead is whether it is reasonable to assume interaural symmetry, and if not, we hope that the framework proposed here can serve as a helpful guide in how to proceed further.

REFERENCES

1. Lin FR, Niparko JK, Ferrucci L. Hearing loss prevalence in the United States. Arch Intern Med. 2011;171: 1851–1852. doi:10.1002/gps.2627.6.
2. Haile LM, Kamenov K, Briant PS, Orji AU, Steinmetz JD, Abdoli A, et al. Hearing loss prevalence and years lived with disability, 1990-2019: Findings from the Global Burden of Disease Study 2019. Lancet. 2021;397: 996–1009. doi:10.1016/S0140-6736(21)00516-X
3. Peters BR, Wyss J, Manrique M. Worldwide trends in bilateral cochlear implantation. Laryngoscope. 2010;120: S17–S44. doi:10.1002/lary.20859
4. Gifford RH, Dorman MF. Bimodal Hearing or Bilateral Cochlear Implants? Ask the Patient. Ear Hear. 2018;40: 501–516. doi:10.1097/AUD.0000000000000657
5. Turton L, Souza P, Thibodeau L, Hickson L, Gifford R, Bird J, et al. Guidelines for best practice in the audiological management of adults with severe and profound hearing loss. Semin Hear. 2020;41: 141–245. doi:10.1055/s-0040-1714744
6. Litovsky R, Parkinson A, Arcaroli J, Sammeth C. Simultaneous bilateral cochlear implantation in adults: A multicenter clinical study. Ear Hear. 2006;27: 714–731. doi:10.1097/01.aud.0000246816.50820.42
7. van Hoesel RJM, Tyler RS. Speech perception, localization, and lateralization with bilateral cochlear implants. J Acoust Soc Am. 2003;113: 1617–1630. doi:10.1121/1.1539520
8. Loizou PC, Hu Y, Litovsky R, Yu G, Peters R, Lake J, et al. Speech recognition by bilateral cochlear implant users in a cocktail-party setting. J Acoust Soc Am. 2009;125: 372–383. doi:10.1121/1.3036175

9. Ihlefeld A, Carlyon RP, Kan A, Churchill TH, Litovsky RY. Limitations on monaural and binaural temporal processing in bilateral cochlear implant listeners. J Assoc Res Otolaryngol. 2015;16: 641–652. doi:10.1007/s10162-015-0527-7
10. Bakal TA, Milvae KDR, Chen C, Goupell MJ. Head Shadow, Summation, and Squelch in Bilateral Cochlear-Implant Users With Linked Automatic Gain Controls. Trend Hear. 2021;25: 1–17. doi:10.1177/23312165211018147
11. Mosnier I, Sterkers O, Bebear JP, Godey B, Robier A, Deguine O, et al. Speech performance and sound localization in a complex noisy environment in bilaterally implanted adult patients. Audiol Neurootol. 2009;14: 106–114. doi:10.1159/000159121
12. Yoon YS, Li Y, Kang HY, Fu QJ. The relationship between binaural benefit and difference in unilateral speech recognition performance for bilateral cochlear implant users. Int J Audiol. 2011;50: 554–565. doi:10.3109/14992027.2011.580785
13. Goupell MJ, Stakhovskaya OA, Bernstein JGW. Contralateral Interference Caused by Binaurally Presented Competing Speech in Adult Bilateral Cochlear-Implant Users. Ear Hear. 2018;39: 110–123. doi:10.1097/AUD.0000000000000470
14. Bernstein JGW, Goupell MJ, Schuchman GI, Rivera AL, Brungart DS. Having two ears facilitates the perceptual separation of concurrent talkers for bilateral and single-sided deaf cochlear implantees. Ear Hear. 2016;37: 289–302. doi:10.1097/AUD.0000000000000284
15. Goupell MJ, Kan A, Litovsky RY. Spatial attention in bilateral cochlear-implant users. J Acoust Soc Am. 2016;140: 1652–1662. doi:10.1121/1.4962378
16. Durlach NI. Equalization and cancellation theory of binaural masking-level differences. J Acoust Soc Am. 1963;35: 1206–1218. doi:10.1121/1.1918675

17. Schimmel O, van de Par S, Breebaart J, Kohlrausch A. Sound segregation based on temporal envelope structure and binaural cues. J Acoust Soc Am. 2008;124: 1130–1145. doi:10.1121/1.2945159

18. Piechowiak T, Ewert SD, Dau T. Modeling comodulation masking release using an equalization-cancellation mechanism. J Acoust Soc Am. 2007;121: 2111–2126. doi:10.1121/1.2534227

19. Goupell MJ, Eisenberg D, DeRoy Milvae K. Dichotic listening performance with cochlear-implant simulations of ear asymmetry is consistent with difficulty ignoring clearer speech. Atten Percept Psychophys. 2021;83: 2083–2101. doi:10.3758/s13414-021-02244-x

20. Gallun FJ, Mason CR, Kidd, Gerald J. The ability to listen with independent ears. J Acoust Soc Am. 2007;122: 2814–2825. doi:10.1121/1.2780143

21. Kidd G, Mason CR, Arbogast TL, Brungart DS, Simpson BD. Informational masking caused by contralateral stimulation. J Acoust Soc Am. 2003;113: 1594–1603. doi:10.1121/1.1547440

22. Zwislocki JJ. A Theory of Central Auditory Masking and Its Partial Validation. J Acoust Soc Am. 1971;52: 644–659. doi:10.1121/1.1913154

23. Bernstein JGW, Stakhovskaya OA, Jensen KK, Goupell MJ. Acoustic Hearing Can Interfere with Single-Sided Deafness Cochlear-Implant Speech Perception. Ear Hear. 2020;41: 747–761. doi:10.1097/AUD.0000000000000805

24. Reiss LAJ, Eggleston JL, Walker EP, Oh Y. Two ears are not always better than one: Mandatory vowel fusion across spectrally mismatched ears in hearing-impaired listeners. J Assoc Res Otolaryngol. 2016;17: 341–356. doi:10.1007/s10162-016-0570-z

25. Oh Y, Hartling CL, Srinivasan NK, Eddolls M, Diedesch AC, Gallun FJ, et al. Broad binaural fusion impairs segregation of speech based on voice pitch differences in a ‘ cocktail party ’ environment. BioRxiv. 2019; 1–39. doi:10.1101/805309

26. Reiss LAJ, Molis MR. Abnormal fusion of dichotic vowels across different fundamental frequencies in hearing-impaired listeners: An alternative explanation for difficulties with speech in background talkers. J Assoc Res Otolaryngol. 2021;22: 443–461. doi:10.1007/s10162-021-00790-7

27. Lin P, Lu T, Zeng F-G. Central masking with bilateral cochlear implants. J Acoust Soc Am. 2013;133: 962–969. doi:10.1121/1.4773262

28. Wightman FL, Kistler DJ. Monaural sound localization revisited. J Acoust Soc Am. 1997;101: 1050–1063. doi:10.1121/1.418029

29. Swaminathan J, Mason CR, Streeter TM, Best V, Roverud E, Kidd G. Role of binaural temporal fine structure and envelope cues in cocktail-party listening. J Neurosci. 2016;36: 8250–8257. doi:10.1523/JNEUROSCI.4421-15.2016

30. Blauert J. Introduction. 1st ed. Spatial Hearing: The Psychophysics of Human Sound Localization. 1st ed. Cambridge, MA: The MIT Press; 1983. pp. 1–35.

31. Kidd, Gerald J, Mason CR, Best VM, Gallun FJ, Durlach NI. Informational masking. In: Yost WA, Popper AN, Fay RR, editors. Springer Handbook of Auditory Research: Auditory Perception of Sound Sources. Boston, MA: Springer; 2006. pp. 143–189. doi:10.1055/s-0040-1715679

32. Kimura D. Functional asymmetry of the brain in dichotic listening. Cortex. 1967;3: 163–178. doi:10.1016/s0010-9452(67)80010-8

33. Kinsbourne M. The cerebral basis of lateral asymmetries in attention. Acta Psychol.

1970;33: 193–201. doi:10.1016/0001-6918(70)90132-0

34. Yin TCT, Smith PH, Joris PX. Neural mechanisms of binaural processing in the auditory brainstem. Compr Physiol. 2019;9: 1503–1575. doi:10.1002/cphy.c180036

35. Loizou PC. Speech processing in vocoder-centric cochlear implants. In: Møller AR, editor. Cochlear and Brainstem Implants Adv Otorhinolaryngol. Basel, Switzerland: Karger; 2006. pp. 109–143. doi:10.1159/000094648

36. Gray WO, Mayo PG, Goupell MJ, Brown AD. Transmission of Binaural Cues by Bilateral Cochlear Implants: Examining the Impacts of Bilaterally Independent Spectral Peak-Picking, Pulse Timing, and Compression. Trend Hear. 2021;25: 1–23. doi:10.1177/23312165211030411

37. Ausili SA, Agterberg MJH, Engel A, Voelter C, Thomas JP, Brill S, et al. Spatial Hearing by Bilateral Cochlear Implant Users With Temporal Fine-Structure Processing. Front Neurol. 2020;11: 1–13. doi:10.3389/fneur.2020.00915

38. Fischer T, Schmid C, Kompis M, Mantokoudis G, Caversaccio M, Wimmer W. Effects of temporal fine structure preservation on spatial hearing in bilateral cochlear implant users. J Acoust Soc Am. 2021;150: 673–686. doi:10.1121/10.0005732

39. Dennison SR, Jones HG, Kan A, Litovsky RY. The impact of synchronized cochlear implant sampling and stimulation on free-field spatial hearing outcomes: comparing the ciPDA research processor to clinical processors. Ear Hear. 2021; Forthcoming.

40. Aronoff JM, Yoon Y, Freed DJ, Vermiglio AJ, Pal I, Soli SD. The use of interaural time and level difference cues by bilateral cochlear implant users. J Acoust Soc Am. 2010;127: EL87-EL92. doi:10.1121/1.3298451

41. Grantham DW, Ashmead DH, Ricketts TA, Haynes DS, Labadie RF. Interaural time and

level difference thresholds for acoustically presented signals in post-lingually deafened adults fitted with bilateral cochlear implants using CIS+ processing. Ear Hear. 2008;29: 33–44. doi:10.1097/AUD.0b013e31815d636f

42. Seeber BU, Fastl H. Localization cues with bilateral cochlear implants. J Acoust Soc Am. 2008;123: 1030–1042. doi:10.1121/1.2821965

43. Goupell MJ, Noble JH, Phatak SA, Kolberg E, Cleary M, Stakhovskaya OA, et al. Computed-tomography estimates of interaural mismatch in insertion depth and scalar location in bilateral cochlear-implant users. medRxiv. 2021; 2021.02.26.21252533. doi:10.1101/2021.02.26.21252533

44. Laback B, Egger K, Majdak P. Perception and coding of interaural time differences with bilateral cochlear implants. Hear Res. 2015;322: 138–150. doi:10.1016/j.heares.2014.10.004

45. Kan A, Litovsky RY. Binaural hearing with electrical stimulation. Hear Res. 2015;322: 127–137. doi:10.1016/j.heares.2014.08.005

46. Anderson SR, Easter K, Goupell MJ. Effects of rate and age in processing interaural time and level differences in normal-hearing and bilateral cochlear-implant listeners. J Acoust Soc Am. 2019;146: 3232–3254. doi:10.1121/1.5130384

47. Joris PX, Trussell LO. The calyx of Held: A hypothesis on the need for reliable timing in an intensity-difference encoder. Neuron. 2018;100: 534–549. doi:10.1016/j.neuron.2018.10.026

48. Brown AD, Tollin DJ. Slow temporal integration enables robust neural coding and perception of a cue to sound source location. J Neurosci. 2016;36: 9908–9921. doi:10.1523/JNEUROSCI.1421-16.2016

49. Beiderbeck B, Myoga MH, Müller NIC, Callan AR, Friauf E, Grothe B, et al. Precisely timed inhibition facilitates action potential firing for spatial coding in the auditory brainstem. Nat Comm. 2018;9: 1–13. doi:10.1038/s41467-018-04210-y

50. Franken TP, Bondy BJ, Haimes DB, Goldwyn JH, Golding NL, Smith PH, et al. Glycinergic axonal inhibition subserves acute spatial sensitivity to sudden increases in sound intensity. Elife. 2021;10: 1–32. doi:10.7554/eLife.62183

51. Joris PX, Yin TCT. Envelope coding in the lateral superior olive. I. Sensitivity to interaural time differences. J Neurophysiol. 1995;73: 1043–1062. doi:10.1152/jn.1995.73.3.1043

52. Boudreau JC, Tsuchitani C. Binaural interaction in the cat superior olive S segment. J Neurophysiol. 1968;31: 442–454. doi:10.1152/jn.1968.31.3.442

53. Franken TP, Joris PX, Smith PH. Principal cells of the brainstem’s interaural sound level detector are temporal differentiators rather than integrators. Elife. 2018;7: 1–25. doi:10.7554/eLife.33854

54. Golding NL, Oertel D. Synaptic integration in dendrites: Exceptional need for speed. J Physiol. 2012;590: 5563–5569. doi:10.1113/jphysiol.2012.229328

55. Guinan JJ, Norris BE, Guinan SS. Single auditory units in the superior olivary complex: II: Locations of unit categories and tonotopic organization. Int J Neuro. 1972;4: 147–166. doi:10.3109/00207457209164756

56. Tsuchitani C. Input from the medial nucleus of trapezoid body to an interaural level detector. Hear Res. 1997;105: 165–176.

57. Tsuchitani C. Functional organization of lateral cell groups of cat superior olivary complex. J Neurophysiol. 1977;40: 296–318. doi:10.1152/jn.1977.40.2.296

58. Remme MWH, Donato R, Mikiel-Hunter J, Ballestero JA, Foster S, Rinzel J, et al. Subthreshold resonance properties contribute to the efficient coding of auditory spatial cues. Proc Natl Acad Sci U S A. 2014;111: 2339–2348. doi:10.1073/pnas.1316216111

59. Joris PX. Envelope coding in the lateral superior olive. II. Characteristic delays and comparison with responses in the medial superior olive. J Neurophysiol. 1996;76: 2137–2156. doi:10.1152/jn.1996.76.4.2137

60. Joris PX, Yin TCT. Envelope Coding in the Lateral Superior Olive. III. Comparison With Afferent Pathways. Am Physiol Soc. 1998;79: 253–269. doi:10.1152/jn.1998.79.1.253

61. Bernstein LR, Trahiotis C. Sensitivity to envelope-based interaural delays at high frequencies: Center frequency affects the envelope rate-limitation. J Acoust Soc Am. 2014;135: 808–816. doi:10.1121/1.4861251

62. Kong Y-Y, Deeks JM, Axon PR, Carlyon RP. Limits of temporal pitch in cochlear implants. J Acoust Soc Am. 2009;125: 1649–1657. doi:10.1121/1.3068457

63. Kong Y-Y, Carlyon RP. Temporal pitch perception at high rates in cochlear implants. J Acoust Soc Am. 2010;127: 3114–3123. doi:10.1121/1.3372713

64. Tollin DJ, Yin TCT. The coding of spatial location by single units in the lateral superior olive of the cat. II. The determinants of spatial receptive fields in azimuth. J Neurosci. 2002;22: 1468–1479. doi:10.1523/jneurosci.22-04-01468.2002

65. Lingner A, Pecka M, Leibold C, Grothe B. A novel concept for dynamic adjustment of auditory space. Sci Rep. 2018;8: 1–12. doi:10.1038/s41598-018-26690-0

66. Cant NB, Casseday JH. Projections from the anteroventral cochlear nucleus to the lateral and medial superior olivary nuclei. J Comp Neurol. 1986;247: 457–476. doi:10.1002/cne.902470406

67. Doucet JR, Ryugo DK. Axonal pathways to the lateral superior olive labeled with biotinylated dextran amine injections in the dorsal cochlear nucleus of rats. J Comp Neurol. 2003;461: 452–465. doi:10.1002/cne.10722

68. Smith PH, Joris PX, Yin TCT. Anatomy and physiology of principal cells of the medial nucleus of the trapezoid body (MNTB) of the cat. J Neurophysiol. 1998;79: 3127–3142. doi:10.1152/jn.1998.79.6.3127

69. Smith PH, Joris PX, Carney LH, Yin TCT. Projections of physiologically characterized globular bushy cell axons from the cochlear nucleus of the cat. J Comp Neurol. 1991;304: 387–407. doi:10.1002/cne.903040305

70. Joris PX, Smith PH. The volley theory and the spherical cell puzzle. Neuro. 2008;154: 65–76. doi:10.1016/j.neuroscience.2008.03.002

71. Joris PX, Carney H, Smith H, Yin TCT. Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency. J Neurophysiol. 1994;71: 1022–1036. doi:10.1152/jn.1994.71.3.1022

72. Rhode WS, Greenberg S. Encoding of amplitude modulation in the cochlear nucleus of the cat. J Neurophysiol. 1994;71: 1797–1825. doi:10.1152/jn.1994.71.5.1797

73. Keine C, Rübsamen R. Inhibition shapes acoustic responsiveness in spherical bushy cells. J Neurosci. 2015;35: 8579–8592. doi:10.1523/JNEUROSCI.0133-15.2015

74. Rothman JS, Young ED, Manis PB. Convergence of auditory nerve fibers onto bushy cells in the ventral cochlear nucleus: Implications of a computational model. J Neurophysiol. 1993;70: 2562–2583. doi:10.1152/jn.1993.70.6.2562

75. Spirou GA, Rager J, Manis PB. Convergence of auditory-nerve fiber projections onto globular bushy cells. Neuroscience. 2005;136: 843–863.

doi:10.1016/j.neuroscience.2005.08.068

76. Caspary DM, Backoff PM, Finlayson PG, Palombi PS. Inhibitory inputs modulate discharge rate within frequency receptive fields of anteroventral cochlear nucleus neurons. J Neurophysiol. 1994;72: 2124–2133. doi:10.1152/jn.1994.72.5.2124

77. Banks MI, Smith PH. Intracellular recordings from neurobiotin-labeled cells in brain slices of the rat medial nucleus of the trapezoid body. J Neurosci. 1992;12: 2819–2837. doi:10.1523/jneurosci.12-07-02819.1992

78. Smith ZM, Delgutte B. Sensitivity to Interaural Time Differences in the Inferior Colliculus with Bilateral Cochlear Implants. J Neurosci. 2007;27: 6740–6750. doi:10.1523/JNEUROSCI.0052-07.2007

79. Thornton JL, Anbuhl KL, Tollin DJ. Unilateral conductive hearing loss causes impaired auditory information processing in neurons in the central auditory system. Proc Int Symp Audit Audiol Res. 2013; 171–178.

80. Buck AN, Rosskothen-Kuhl N, Schnupp JW. Sensitivity to interaural time differences in the inferior colliculus of cochlear implanted rats with or without hearing experience. Hear Res. 2021;408: 108305. doi:10.1016/j.heares.2021.108305

81. Green DM, Swets JA. Signal Detection Theory and Psychophysics. 1st ed. Los Altos Hills, CA: Peninsula Publishing; 1966.

82. Holder JT, Reynolds SM, Sunderhaus LW, Gifford RH. Current Profile of Adults Presenting for Preoperative Cochlear Implant Evaluation. Trend Hear. 2018;22: 1–16. doi:10.1177/2331216518755288

83. Sanes DH, Bao S. Tuning up the developing auditory CNS. Curr Opin Neurobiol. 2009;19: 188–199. doi:10.1016/j.conb.2009.05.014

84. Takesian AE, Kotak VC, Sanes DH. Developmental hearing loss disrupts synaptic inhibition: Implications for auditory processing. Futur Neurol. 2009;4: 331–349. doi:10.2217/FNL.09.5

85. Litovsky RY, Ashmead DH. Development of binaural and spatial hearing. In: Gilkey RH, Anderson TR, editors. Binaural and Spatial Hearing in Real and Virtual Environments. Mahwah, NJ: Lawrence Erlbaum Associates; 1997. pp. 163–195. doi:10.1007/978-1-4614-1421-6_6

86. Seidl AH, Grothe B. Development of sound localization mechanisms in the Mongolian gerbil is shaped by early acoustic experience. J Neurophysiol. 2005;94: 1028–1036. doi:10.1152/jn.01143.2004

87. Van Deun L, van Wieringen A, Scherf F, Deggouj N, Desloovere C, Offeciers FE, et al. Earlier intervention leads to better sound localization in children with bilateral cochlear implants. Audiol Neurootol. 2010;15: 7–17.

88. Strøm-Roum H, Rødvik AK, Osnes TA, Fagerland MW, Wie OB. Sound localising ability in children with bilateral sequential cochlear implants. Int J Pediatr Otorhinolaryngol. 2012;76: 1245–1248. doi:10.1016/j.ijporl.2012.05.013

89. Steffens T, Lesinski-Schiedat A, Strutz J, Aschendorff A, Klenzner T, Rühl S, et al. The benefits of sequential bilateral cochlear implantation for hearing-impaired children. Acta Otolaryngol. 2008;128: 164–176. doi:10.1080/00016480701411528

90. Asp F, Mäki-Torkko E, Karltorp E, Harder H, Hergils L, Eskilsson G, et al. A longitudinal study of the bilateral benefit in children with bilateral cochlear implants. Int J Audiol. 2015;54: 77–88. doi:10.3109/14992027.2014.973536

91. Killan C, Scally A, Killan E, Totten C, Raine C. Factors affecting sound-source

localization in children with simultaneous or sequential bilateral cochlear implants. Ear Hear. 2019;40: 870–877. doi:10.1097/AUD.0000000000000666

92. Grieco-Calub TM, Litovsky RY. Sound localization skills in children who use bilateral cochlear implants and in children with normal acoustic hearing. Ear Hear. 2010;31: 645–656. doi:10.1097/AUD.0b013e3181e50a1d

93. Litovsky RY, Jones GL, Agrawal S, van Hoesel R. Effect of age at onset of deafness on binaural sensitivity in electric hearing in humans. J Acoust Soc Am. 2010;127: 400–414. doi:10.1121/1.3257546

94. Ehlers E, Goupell MJ, Zheng Y, Godar SP, Litovsky RY. Binaural sensitivity in children who use bilateral cochlear implants. J Acoust Soc Am. 2017;141: 4264–4277. doi:10.1121/1.4983824

95. Chung Y, Buechel BD, Sunwoo W, Wagner JD, Delgutte B. Neural ITD sensitivity and temporal coding with cochlear implants in an animal model of early-onset deafness. J Assoc Res Otolaryngol. 2019;20: 37–56. doi:10.1007/s10162-018-00708-w

96. Hancock KE, Chung Y, Delgutte B. Congenital and prolonged adult-onset deafness cause distinct degradations in neural ITD coding with bilateral cochlear implants. J Assoc Res Otolaryngol. 2013;14: 393–411. doi:10.1007/s10162-013-0380-5

97. Schofield BR. Superior olivary complex and lateral lemniscal connections of the auditory midbrain. In: Winer JA, Schreiner CE, editors. The Inferior Colliculus. New York, NY: Springer; 2005. pp. 132–154. doi:10.1007/0-387-27083-3_4

98. Sunwoo W, Delgutte B, Chung Y. Chronic bilateral cochlear implant stimulation partially restores neural binaural sensitivity in neonatally-deaf rabbits. J Neurosci. 2021;41: 3651–3664. doi:10.1523/JNEUROSCI.1076-20.2021

99. Rosskothen-Kuhl N, Buck AN, Li K, Schnupp JWH. Microsecond interaural time difference discrimination restored by cochlear implants after neonatal deafness. Elife. 2021;10: e5930005. doi:10.7554/eLife.59300

100. Wilmington D, Gray L, Jahrsdoerfer R. Binaural processing after corrected congenital unilateral conductive hearing loss. Hear Res. 1994;74: 99–114. doi:10.1016/0378-5955(94)90179-1

101. Clements M, Kelly JB. Auditory spatial responses of young guinea pigs (Cavia porcellus) during and after ear blocking. J Comp Physiol Psychol. 1978;92: 34–44. doi:10.1037/h0077424

102. Moore DR, Hine JE, Jiang ZD, Matsuda H, Parsons CH, King AJ. Conductive hearing loss produces a reversible binaural hearing impairment. J Neurosci. 1999;19: 8704–8711. doi:10.1523/JNEUROSCI.19-19-08704.1999

103. Gordon KA, Jiwani S, Papsin BC. What is the optimal timing for bilateral cochlear implantation in children? Cochlear Implant Int. 2014;12: S8–S14. doi:10.1179/146701011X13074645127199

104. Polonenko MJ, Papsin BC, Gordon KA. The effects of asymmetric hearing on bilateral brainstem function: Findings in children with bimodal (electric and acoustic) hearing. Audiol Neurootol. 2015;20: 13–20. doi:10.1159/000380743

105. Thakkar T, Anderson SR, Kan A, Litovsky RY. Evaluating the impact of age, acoustic exposure, and electrical stimulation on binaural sensitivity in adult bilateral cochlear implant patients. Brain Sci. 2020;10: 1–26. doi:10.3390/brainsci10060406

106. Leake PA, Hradek GT. Cochlear pathology of long term neomycin induced deafness in cats. Hear Res. 1988;33: 11–33. doi:10.1016/0378-5955(88)90018-4

107. Shepherd RK, Javel E. Electrical stimulation of the auditory nerve. I. Correlation of physiological responses with cochlear status. Hear Res. 1997;108: 112–144. doi:10.1016/S0378-5955(97)00046-4

108. Nadol, Joseph B J, Young Y-S, Glynn RJ. Survival of Spiral Ganglion Cells in Profound Sensorineural Hearing Loss: Implications for Cochlear Implantation. Ann Otol Rhinol Laryngol. 1989;98: 411–416. doi:10.1177/000348948909800602

109. Nadol JJ. Patterns of neural degeneration in the human cochlea and auditory nerve: Implications for cochlear implantation. Otolayngol Head Neck Surg. 1997;117: 220–228. doi:10.1016/S0194-5998(97)70178-5

110. Spoendlin H, Schrott A. Analysis of the human auditory nerve. Hear Res. 1989;43: 25–38. doi:10.1016/0378-5955(89)90056-7

111. Moore BCJ. Dead Regions in the Cochlea: Conceptual Foundations, Diagnosis, and Clinical Applications. Ear Hear. 2004;25: 98–116. doi:10.1097/01.AUD.0000120359.49711.D7

112. Shepherd RK, Hardie NA. Deafness-induced changes in the auditory pathway: Implications for cochlear implants. Audiol Neurootol. 2001;6: 305–318. doi:10.1159/000046843

113. Tagoe T, Barker M, Jones A, Allcock N, Hamann M. Auditory nerve perinodal dysmyelination in noise-induced hearing loss. J Neurosci. 2014;34: 2684–2688. doi:10.1523/JNEUROSCI.3977-13.2014

114. Wan G, Corfas G. Transient auditory nerve demyelination as a new mechanism for hidden hearing loss. Nat Comm. 2017;8: 1–13. doi:10.1038/ncomms14487

115. Goldwyn JH, Bierer SM, Bierer JA. Modeling the electrode-neuron interface of cochlear

implants: Effects of neural survival, electrode placement, and the partial tripolar configuration. Hear Res. 2010;268: 93–104. doi:10.1242/jcs.03292.Multiple

116. Shepherd RK, Roberts LA, Paolini AG. Long-term sensorineural hearing loss induces functional changes in the rat auditory nerve. Eur J Neurosci. 2004;20: 3131–3140. doi:10.1111/j.1460-9568.2004.03809.x

117. Zhou R, Abbas PJ, Assouline JG. Electrically evoked auditory brainstem response in peripherally myelin-deficient mice. Hear Res. 1995;88: 98–106. doi:10.1016/0378-5955(95)00105-D

118. Ashida G, Heinermann HT, Kretzberg J. Neuronal population model of globular bushy cells covering unit-to-unit variability. PLoS Comput. Biol. 2019. doi:10.1371/journal.pcbi.1007563

119. Sinclair JL, Fischl MJ, Alexandrova O, Heβ M, Grothe B, Leibold C, et al. Sound-evoked activity influences myelination of brainstem axons in the trapezoid body. J Neurosci. 2017;37: 8239–8255. doi:10.1523/JNEUROSCI.3728-16.2017

120. Gillespie MJ, Stein RB. The relationship between axon diameter, myelin thickness and conduction velocity during atrophy of mammalian peripheral nerves. Brain Res. 1983;259: 41–56. doi:10.1016/0006-8993(83)91065-X

121. Todd AE, Goupell MJ, Litovsky RY. The Relationship between Intensity Coding and Binaural Sensitivity in Adults with Cochlear Implants. Ear Hear. 2017;38: e128–e141. doi:10.1097/AUD.0000000000000382

122. Ihlefeld A, Kan A, Litovsky RY. Across-frequency combination of interaural time difference in bilateral cochlear implant listeners. Front Syst Neurosci. 2014;8: 1–12. doi:10.3389/fnsys.2014.00022

123. Long CJ, Holden TA, McClelland GH, Parkinson WS, Shelton C, Kelsall DC, et al. Examining the electro-neural interface of cochlear implant users using psychophysics, CT scans, and speech understanding. JARO - J Assoc Res Otolaryngol. 2014;15: 293–304. doi:10.1007/s10162-013-0437-5

124. Schvartz-Leyzac KC, Holden TA, Zwolan TA, Arts HA, Firszt JB, Buswinka CJ, et al. Effects of electrode location on estimates of neural health in humans with cochlear implants. J Assoc Res Otolaryngol. 2020;21: 259–275. doi:10.1007/s10162-020-00749-0

125. Bierer JA. Probing the electrode-neuron interface with focused cochlear implant stimulation. Trends Amplif. 2010;14: 84–95. doi:10.1177/1084713810375249

126. Bierer JA, Nye AD. Comparisons between detection threshold and loudness perception for individual cochlear implant channels. Ear Hear. 2014;35: 641–51. doi:10.1097/AUD.0000000000000058

127. Goupell MJ, Kan A, Litovsky RY. Mapping procedures can produce non-centered auditory images in bilateral cochlear implantees. J Acoust Soc Am. 2013;133: EL101-107. doi:10.1121/1.4776772

128. Buchholz JM, Le Goff N, Dau T. Localization of broadband sounds carrying interaural time differences: Effects of frequency, reference location, and interaural coherence. J Acoust Soc Am. 2018;144: 2225–2237. doi:10.1121/1.5058776

129. Whitmer WM, Seeber BU, Akeroyd MA. The perception of apparent auditory source width in hearing-impaired adults. J Acoust Soc Am. 2014;135: 3548–3559. doi:10.1121/1.4875575

130. Chakravorti S, Noble JH, Gifford RH, Dawant BM, O'Connell BP, Wang J, et al. Further evidence of the relationship between cochlear implant electrode positioning and hearing

outcomes. Otol Neurotol. 2019;40: 617–624. doi:10.1097/MAO.0000000000002204

131. Wanna GB, Noble JH, Carlson ML, Gifford RH, Dietrich MS, Haynes DS, et al. Impact of electrode design and surgical approach on scalar location and cochlear implant outcomes. Laryngoscope. 2014;124: S1–S7. doi:10.1002/lary.24728

132. Friesen LM, Shannon R V., Baskent D, Wang X. Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. J Acoust Soc Am. 2001;110: 1150–1163. doi:10.1121/1.1381538

133. Croghan NBH, Duran SI, Smith ZM. Re-examining the relationship between number of cochlear implant channels and maximal speech intelligibility. J Acoust Soc Am. 2017;142: EL537-EL543. doi:10.1121/1.5016044

134. Dong Y, Briaire JJ, Siebrecht M, Stronks HC, Frijns JHM. Detection of Translocation of Cochlear Implant Electrode Arrays by Intracochlear Impedance Measurements. Ear Hear. 2021;42: 1397–1404. doi:10.1097/AUD.0000000000001033

135. Benichoux V, Ferber A, Hunt S, Hughes E, Tollin D. Across species “natural ablation” reveals the brainstem source of a noninvasive biomarker of binaural hearing. J Neurosci. 2018;38: 8563–8573. doi:10.1523/JNEUROSCI.1211-18.2018

136. Laumen G, Ferber AT, Klump GM, Tollin DJ. The physiological basis and clinical use of the binaural interaction component of the auditory brainstem response. Ear Hear. 2016;37: e276–e290. doi:10.1097/AUD.0000000000000301

137. Brown AD, Anbuhl KL, Gilmer JI, Tollin DJ. Between-ear sound frequency disparity modulates a brain stem biomarker of binaural hearing. J Neurophysiol. 2019;122: 1110–1122. doi:10.1152/jn.00057.2019

138. Kan A, Stoelb C, Litovsky RY, Goupell MJ. Effect of mismatched place-of-stimulation on

binaural fusion and lateralization in bilateral cochlear-implant users. J Acoust Soc Am. 2013;134: 2923–2936. doi:10.1121/1.4820889

139. Kan A, Goupell MJ, Litovsky RY. Effect of channel separation and interaural mismatch on fusion and lateralization in normal-hearing and cochlear-implant listeners. J Acoust Soc Am. 2019;146: 1448–1463. doi:10.1121/1.5123464

140. Goupell MJ. Interaural envelope correlation change discrimination in bilateral cochlear implantees: Effects of mismatch, centering, and onset of deafness. J Acoust Soc Am. 2015;137: 1282–1297. doi:10.1121/1.4908221

141. Goupell MJ, Stoelb CA, Kan A, Litovsky RY. The effect of simulated interaural frequency mismatch on speech understanding and spatial release from masking. Ear Hear. 2018;39: 895–905. doi:10.1097/AUD.0000000000000541

142. Goupell MJ. The role of envelope statistics in detecting changes in interaural correlation. J Acoust Soc Am. 2012;132: 1561–1572. doi:10.1121/1.4740498

143. Goupell MJ, Litovsky RY. The effect of interaural fluctuation rate on correlation change discrimination. J Assoc Res Otolaryngol. 2014;15: 115–129. doi:10.1007/s10162-013-0426-8

144. Kan A, Litovsky RY, Goupell MJ. Effects of interaural pitch matching and auditory image centering on binaural sensitivity in cochlear implant users. Ear Hear. 2015;36: e62–e68. doi:10.1097/AUD.0000000000000135

145. Rao RPN, Ballard DH. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat Neurosci. 1999;2: 79–87. Available: http://neurosci.nature.com

146. Bizley JK, Cohen YE. The what, where and how of auditory-object perception. Nat Rev

Neurosci. 2013;14: 693–707. doi:10.1038/nrn3565

147. Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, Friston KJ. Canonical Microcircuits for Predictive Coding. Neuron. 2012;76: 695–711. doi:10.1016/j.neuron.2012.10.038

148. Kral A, Yusuf PA, Land R. Higher-order auditory areas in congenital deafness: Top-down interactions and corticocortical decoupling. Hear Res. 2017;343: 50–63. doi:10.1016/j.heares.2016.08.017

149. Shinn-Cunningham BG. Object-based auditory and visual attention. Trends Cogn Sci. 2008;12: 182–186. doi:10.1016/j.tics.2008.02.003

150. Lupyan G, Clark A. Words and the World: Predictive Coding and the Language-Perception-Cognition Interface. Curr Dir Psychol Sci. 2015;24: 279–284. doi:10.1177/0963721415570732

151. Ylinen S, Bosseler A, Junttila K, Huotilainen M. Predictive coding accelerates word recognition and learning in the early stages of language development. Dev Sci. 2017;20. doi:10.1111/desc.12472

152. Luce PA, Pisoni DB. Recognizing spoken words: The neighborhood activation model. Ear Hear. 1998;19: 1–36.

153. Griffiths TD, Warren JD. What is An Auditory Object? Nat Rev Neurosci. 2004;5: 887–892.

154. Shinn-Cunningham B, Best V, Lee AKC. Auditory object formation and selection. 60th ed. In: Middlebrooks JC, Simon J, Popper AN, Fay RR, editors. The Auditory System at the Cocktail Party Springer Handbook of Auditory Research. 60th ed. Cham, Switzerland: Springer; 2017. pp. 7–40. doi:10.1007/978-3-319-51662-2_2

155. Darwin CJ. Auditory Grouping. Trends Cogn Sci. 1997;1: 327–333. doi:10.1016/b978-012505626-7/50013-3

156. Reiss LAJ, Fowler JR, Hartling CL, Oh Y. Binaural pitch fusion in bilateral cochlear implant users. Ear Hear. 2018;39: 390–397. doi:10.1097/AUD.0000000000000497

157. Reiss LAJ, Ito RA, Eggleston JL, Wozny DR. Abnormal binaural spectral integration in cochlear implant users. J Assoc Res Otolaryngol. 2014;15: 235–248. doi:10.1007/s10162-013-0434-8

158. Reiss LAJ, Shayman CS, Walker EP, Bennett KO, Fowler JR, Hartling CL, et al. Binaural pitch fusion: Comparison of normal-hearing and hearing-impaired listeners. J Acoust Soc Am. 2017;141: 1909–1920. doi:10.1121/1.4978009

159. van Hoesel RJM, Clark GM. Psychophysical studies with two binaural cochlear implant subjects. J Acoust Soc Am. 1997;102: 495–507. doi:10.1121/1.419611

160. van Hoesel RJM, Clark GM. Fusion and lateralization study with two binaural cochlear implant patients. Ann Otol Rhinol Laryngol. 1995;104: 233–235.

161. Long CJ, Eddington DK, Colburn HS, Rabinowitz WM. Binaural sensitivity as a function of interaural electrode position with a bilateral cochlear implant user. J Acoust Soc Am. 2003;114: 1565–1574. doi:10.1121/1.1603765

162. Baumgärtel RM, Hu H, Kollmeier B, Dietz M. Extent of lateralization at large interaural time differences in simulated electric hearing and bilateral cochlear implant users. J Acoust Soc Am. 2017;141: 2338–2352. doi:10.1121/1.4979114

163. Goupell MJ, Stoelb C, Kan A, Litovsky RY. Effect of mismatched place-of-stimulation on the salience of binaural cues in conditions that simulate bilateral cochlear-implant listening. J Acoust Soc Am. 2013;133: 2272–2287. doi:10.1121/1.4792936

164. Suneel D, Staisloff H, Shayman CS, Stelmach J, Aronoff JM. Localization performance correlates with binaural fusion for interaurally mismatched vocoded speech. J Acoust Soc Am. 2017;142: EL276-EL280. doi:10.1121/1.5001903

165. Brown AD, Jones HG, Kan A, Thakkar T, Stecker GC, Goupell MJ, et al. Evidence for a neural source of the precedence effect in sound localization. J Neurophysiol. 2015;114: 2991–3001. doi:10.1152/jn.00243.2015

166. Salloum CAM, Valero J, Wong DDE, Papsin BC, Van Hoesel R, Gordon KA. Lateralization of interimplant timing and level differences in children who use bilateral cochlear implants. Ear Hear. 2010;31: 441–456. doi:10.1097/AUD.0b013e3181d4f228

167. Steel MM, Papsin BC, Gordon KA. Binaural fusion and listening effort in children who use bilateral cochlear implants: A psychoacoustic and pupillometric study. PLoS One. 2015;10: e0117611. doi:10.1371/journal.pone.0117611

168. Fritz JB, Elhilali M, David S V., Shamma SA. Auditory attention - focusing the searchlight on sound. Curr Opin Neurobiol. 2007;17: 437–455. doi:10.1016/j.conb.2007.07.011

169. DeRoy Milvae K, Kuchinsky SE, Stakhovskaya OA, Goupell MJ. Dichotic listening performance and effort as a function of spectral resolution and interaural symmetry. J Acoust Soc Am. 2021;150: 920–935. doi:10.1121/10.0005653

170. Deouell LY, Soroker N. What is extinguished in auditory extinction? Neuroreport. 2000;11: 3059–3062. doi:10.1097/00001756-200009110-00046

171. Tanabe H, Nishikawa T, Okuda J -i, Shiraishi J. Auditory extinction to nonverbal and verbal stimuli. Acta Neurol Scand. 1986;73: 173–179. doi:10.1111/j.1600-0404.1986.tb03260.x

172. Witte C, Grube M, Cramon DY v., Rübsamen R. Auditory extinction and spatio-temporal order judgment in patients with left- and right-hemisphere lesions. Neuropsychologia. 2012;50: 892–903. doi:10.1016/j.neuropsychologia.2012.01.029

173. Warnecke M, Litovsky RY. Signal envelope and speech intelligibility differentially impact auditory motion perception. Nat Sci Rep. 2021;11: 1–10. doi:10.1038/s41598-021-94662-y

174. Hiscock M, Kinsbourne M. Attention and the right-ear advantage: What is the connection? Brain Cogn. 2011;76: 263–275. doi:10.1016/j.bandc.2011.03.016

175. Wood S, Hiscock M, Widrig M. Selective attention fails to alter the dichotic listening lag effect: Evidence that the lag effect is preattentional. Brain Lang. 2000;71: 373–390. doi:10.1006/brln.1999.2271

176. Hugdahl K, Westerhausen R, Alho K, Medvedev S, Hämäläinen H. The effect of stimulus intensity on the right ear advantage in dichotic listening. Neurosci Lett. 2008;431: 90–94. doi:10.1016/j.neulet.2007.11.046

177. Tanaka K, Ross B, Kuriki S, Harashima T, Obuchi C, Okamoto H. Neurophysiological Evaluation of Right-Ear Advantage During Dichotic Listening. Front Psychol. 2021;12: 1–12. doi:10.3389/fpsyg.2021.696263

178. Hiscock M, Cole LC, Benthall JG, Carlson VL, Ricketts JM. Toward solving the inferential problem in laterality research: Effects of increased reliability on the validity of the dichotic listening right-ear advantage. J Int Neuropsychol Soc. 2000;6: 539–547.

179. Voyer D, Ingram JD. Attention, reliability, and validity of perceptual asymmetries in the fused dichotic words test. Laterality. 2005;10: 545–561. doi:10.1080/13576500442000292

180. Voyer D, Flight JI. Reliability and magnitude of auditory laterality effects: The influence

of attention. Brain Cogn. 2001;46: 397–413. doi:10.1006/brcg.2001.1298

181. Kumpik DP, King AJ. A review of the effects of unilateral hearing loss on spatial hearing. Hear Res. 2019;372: 17–28. doi:10.1016/j.heares.2018.08.003

182. Gordon K, Kral A. Animal and human studies on developmental monaural hearing loss. Hear Res. 2019;380: 60–74. doi:10.1016/j.heares.2019.05.011

183. Polonenko MJ, Papsin BC, Gordon KA. Delayed access to bilateral input alters cortical organization in children with asymmetric hearing. NeuroImage Clin. 2018;17: 415–425. doi:10.1016/j.nicl.2017.10.036

184. Yamazaki H, Easwar V, Polonenko MJ, Jiwani S, Wong DDE, Papsin BC, et al. Cortical hemispheric asymmetries are present at young ages and further develop into adolescence. Hum Brain Map. 2017;39: 941–954. doi:10.1002/hbm.23893

185. Gordon K, Henkin Y, Kral A. Asymmetric hearing during development: The aural preference syndrome and treatment options. Pediatrics. 2015;136: 141–153. doi:10.1542/peds.2014-3520

186. Kral A, Heid S, Hubka P, Tillein J. Unilateral hearing during development: hemispheric specificity in plastic reorganizations. Front Syst Neurosci. 2013;7: 1–13. doi:10.3389/fnsys.2013.00093

187. Kral A, Hubka P, Heid S, Tillein J. Single-sided deafness leads to unilateral aural preference within an early sensitive period. Brain. 2013;136: 180–193. doi:10.1093/brain/aws305

188. LeVay S, Wiesel TN, Hubel DH. The development of ocular dominance columns in normal and visually deprived monkeys. J Comp Neurol. 1980;191: 1–51. doi:10.1002/cne.901910102

189. Polonenko MJ, Papsin BC, Gordon KA. Limiting asymmetric hearing improves benefits of bilateral hearing in children using cochlear implants. Sci Rep. 2018;8: 1–17. doi:10.1038/s41598-018-31546-8

190. Koopmann M, Lesinski-Schiedat A, Illg A. Speech Perception, Dichotic Listening, and Ear Advantage in Simultaneous Bilateral Cochlear Implanted Children. Otol Neurotol. 2020;41: e208–e215. doi:10.1097/MAO.0000000000002456

191. Zhou N, Pfingst BE. Psychophysically based site selection coupled with dichotic stimulation improves speech recognition in noise with bilateral cochlear implants. J Acoust Soc Am. 2012;132: 994–1008. doi:10.1121/1.4730907

192. Garadat SN, Zwolan TA, Pfingst BE. Across-site patterns of modulation detection: Relation to speech recognition. J Acoust Soc Am. 2012;131: 4030–4041. doi:10.1121/1.3701879

193. Chatterjee M, Oberzut C. Detection and rate discrimination of amplitude modulation in electrical hearing. J Acoust Soc Am. 2011;130: 1567. doi:10.1121/1.3621445

194. Chatterjee M, Peng S-C. Processing F0 with cochlear implants: Modulation frequency discrimination and speech intonation recognition. Hear Res. 2008;235: 143–156. doi:10.1002/ana.22528.Toll-like

195. Reiss LAJ, Turner CW, Karsten SA, Gantz BJ. Plasticity in human pitch perception induced by tonotopically mismatched electro-acoustic stimulation. Neuro. 2014;256: 43–52. doi:10.1016/j.neuroscience.2013.10.024

196. Reiss LAJ, Ito RA, Eggleston JL, Liao S, Becker JJ, Lakin CE, et al. Pitch adaptation patterns in bimodal cochlear implant users: Over time and after experience. Ear Hear. 2015;36: e23–e34. doi:10.1097/AUD.0000000000000114

197. Blamey P, Artieres F, Başkent D, Bergeron F, Beynon A, Burke E, et al. Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants: An update with 2251 patients. Audiol Neurootol. 2012;18: 36–47. doi:10.1159/000343189

198. Zheng Y, Godar SP, Litovsky RY. Development of sound localization strategies in children with bilateral cochlear implants. PLoS One. 2015;10: e0135790. doi:10.1371/journal.pone.0135790

199. Dunn CC, Walker EA, Oleson J, Kenworthy M, Van Voorst T, Tomblin JB, et al. Longitudinal speech perception and language performance in pediatric cochlear implant users: The effect of age at implantation. Ear Hear. 2014;35: 148–160. doi:10.1097/AUD.0b013e3182a4a8f0

200. Litovsky RY, Misurelli SM. Does bilateral experience lead to improved spatial unmasking of speech in children who use bilateral cochlear implants? Otol Neurotol. 2016;37: e35–e42. doi:10.1097/MAO.0000000000000905

201. Fitzpatrick EM, Leblanc S. Exploring the Factors Influencing Discontinued Hearing Aid Use in Patients With Unilateral Cochlear Implants. Trends Amplif. 2010;14: 199–210. doi:10.1177/1084713810396511

202. Cox RM, Schwartz KS, Noe CM, Alexander GC. Preference for one or two hearing aids among adult patients. Ear Hear. 2011;32: 181–197. doi:10.1097/AUD.0b013e3181f8bf6c

203. Walden TC, Walden BE. Unilateral versus bilateral amplification for adults with impaired hearing. J Am Acad Audiol. 2005;16: 574–584.

204. McArdle RA, Killon M, Mennite MA, Chisolm TH. Are two ears not better than one? 2012;23: 171–181.

205. Firszt JB, Reeder RM, Dwyer NY, Burton H, Holden LK. Localization training results in

individuals with unilateral severe to profound hearing loss. Hear Res. 2015;319: 48–55. doi:10.1016/j.heares.2014.11.005

206. Keating P, Rosenior-Patten O, Dahmen JC, Bell O, King AJ. Behavioral training promotes multiple adaptive processes following acute hearing loss. Elife. 2016;5: e12264. doi:10.7554/eLife.12264

207. Tyler RS, Witt SA, Dunn CC, Wang W. Initial Development of a Spatially Separated Speech-in-Noise and Localization Training Program. J Am Acad Audiol. 2010;21: 390–403. doi:10.3766/jaaa.21.6.4.Initial

208. Festen JM, Plomp R. Effects of fluctuating noise and interfering speech reception threshold for impaired and normal hearing. J Acoust Soc Am. 1990;88: 1725–1736. doi:10.1121/1.400247

209. Bregman AS. Auditory Scene Analysis: The Perceptual Organization of Sound. Caimbridge, MA: Bradford Books, MIT Press; 1990. doi:10.1121/1.408434

210. Joris PX, Yin TCT. Responses to amplitude-modulated tones in the auditory nerve of the cat. J Acoust Soc Am. 1992;91: 215–232. doi:10.1121/1.402757

211. Zilany MSA, Bruce IC, Carney LH. Updated parameters and expanded simulation options for a model of the auditory periphery. J Acoust Soc Am. 2014;135. doi:10.1121/1.4837815

212. Oxenham AJ, Bacon SP. Cochlear compression: Perceptual measures and implications for normal and impaired hearing. Ear Hear. 2003;24: 352–366. doi:10.1097/01.AUD.0000090470.73934.78

213. Litvak LM, Delgutte B, Eddington DK. Improved temporal coding of sinusoids in electric stimulation of the auditory nerve using desynchronizing pulse trains. J Acoust Soc Am.

2003;114: 2079–2098. doi:10.1121/1.1612493

214. Rubinstein JTY, Wilson BS, Finley CC, Abbas PJ. Pseudospontaneous activity: Stochastic independence of auditory nerve fibers with electrical stimulation. Hear Res. 1999;127: 108–118. doi:10.1016/S0378-5955(98)00185-3

215. Gifford RH, Dorman MF, Skarzynski H, Lorens A, Polak M, Driscoll CLW, et al. Cochlear implantation with hearing preservation yields significant benefit for speech recognition in complex listening environments. Ear Hear. 2013;34: 413–425. doi:10.1097/AUD.0b013e31827e8163.Cochlear

216. Brungart DS. Informational and energetic masking effects in the perception of two simultaneous talkers. J Acoust Soc Am. 2001;109: 1101–1109. doi:10.1121/1.1345696

217. Drullman R, Festen JM, Plomp R. Effect of temporal envelope smearing on speech reception. J Acoust Soc Am. 1994;95: 1053–1064. doi:10.1121/1.408467

218. Shannon R V, Zeng F, Kamath V, Wygonski J, Ekelid M. Speech recognition with primarily temporal cues. Science (80-). 1995;270: 303–304. doi:10.1126/science.270.5234.303

219. Rosen S. Temporal information in speech: Acoustic, auditory and linguistic aspects. Phil Trans R Soc Lond B. 1992;336: 367–373. doi:10.1098/rstb.1992.0070

220. Grose JH, Hall JWI, Buss E. Across-channel spectral processing. Int Rev Neurobiol. 2005;70: 87–119. doi:10.1016/S0074-7742(05)70003-9

221. Dollezal L-V, Beutelmann R, Klump GM. Stream segregation in the perception of sinusoidally amplitude-modulated tones. PLoS One. 2012;7: e43615. doi:10.1371/journal.pone.0043615

222. Grimault N, Bacon SP, Micheyl C. Auditory stream segregation on the basis of amplitude-

modulation rate. J Acoust Soc Am. 2002;111: 1340–1348. doi:10.1121/1.1452740

223. Nie Y, Nelson PB. Auditory stream segregation using amplitude modulated bandpass noise. Front Psychol. 2015;6: 1–11. doi:10.3389/fpsyg.2015.01151

224. Chatterjee M, Sarampalis A, Oba SI. Auditory stream segregation with cochlear implants: A preliminary report. Hear Res. 2006;222: 100–107. doi:10.1016/j.heares.2006.09.001

225. Hong RS, Turner CW. Sequential stream segregation using temporal periodicity cues in cochlear implant recipients. J Acoust Soc Am. 2009;126: 291–299. doi:10.1121/1.3140592

226. Carrell TD, Opie JM. The effect of amplitude comodulation on auditory object formation in sentence perception. Percept Psyhophys. 1992;52: 437–445. doi:10.3758/BF03206703

227. Bregman AS, Abramson J, Doehring P, Darwin CJ. Spectral integration based on common amplitude modulation. Percept Psyhophys. 1985;37: 483–493. doi:10.3758/BF03202881

228. Strickland EA, Viemeister NF, Fantini DA, Garrison MA. Within- versus cross-channel mechanisms in detection of envelope phase disparity. J Acoust Soc Am. 1989;86: 2160–2166. doi:10.1121/1.398476

229. Goupell MJ, Litovsky RY. Sensitivity to interaural envelope correlation changes in bilateral cochlear-implant users. J Acoust Soc Am. 2015;137: 335–349. doi:10.1121/1.4904491

230. Richards VM. Monaural envelope correlation perception. J Acoust Soc Am. 1987;82: 1621–1630. doi:10.1121/1.395153

231. Bernstein LR, Trahiotis C. Discrimination of interaural envelope correlation and its relation to binaural unmasking at high frequencies. J Acoust Soc Am. 1992;91: 306–316. doi:10.1121/1.402773

232. Bernstein JGW, Iyer N, Brungart DS. Release from informational masking in a monaural competing-speech task with vocoded copies of the maskers presented contralaterally. J Acoust Soc Am. 2015;137: 702–713. doi:10.1121/1.4906167

233. Grant KW, Summers V, Leek MR. Modulation rate detection and discrimination by normal-hearing and hearing-impaired listeners. J Acoust Soc Am. 1998;104: 1051–1060. doi:10.1121/1.423323

234. Kohlrausch A, Fassel R, Dau T. The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers. J Acoust Soc Am. 2002;108: 723–734. doi:10.1121/1.429605

235. Füllgrabe C, Lorenzi C. The role of envelope beat cues in the detection and discrimination of second-order amplitude modulation (L). J Acoust Soc Am. 2003;113: 49–52. doi:10.1121/1.1523383

236. Moore BCJ, Glasberg BR. Derivation of auditory filter shapes from notched-noise data. Hear Res. 1990;47: 103–138.

237. Litovsky RY. Developmental changes in the precedence effect: Estimates of minimum audible angle. J Acoust Soc Am. 1997;102: 1739–1745. doi:10.1121/1.420106

238. Wichmann AF, Hill NJ. The psychometric function: II. Bootstrap-based confidence intervals and sampling. Percept Psychophys. 2001;63: 1314–29. doi:10.3758/BF03194545

239. Carlyon RP, Deeks JM. Limitations on rate discrimination. J Acoust Soc Am. 2002;112: 1009–1025. doi:10.1121/1.1496766

240. Zilany MSA, Bruce IC, Nelson PC, Carney LH. A phenomenological model of the synapse between the inner hair cell and auditory nerve: Long-term adaptation with power-law dynamics. J Acoust Soc Am. 2009;126: 2390–2412. doi:10.1121/1.3238250

241. Landsberger DM. Effects of modulation wave shape on modulation frequency discrimination with electrical hearing. J Acoust Soc Am. 2008;124: EL21-EL27. doi:10.1121/1.2947624

242. Bernstein LR, Trahiotis C. Enhancing sensitivity to interaural delays at high frequencies by using "transposed stimuli." J Acoust Soc Am. 2002;112: 1026–1036. doi:10.1121/1.1497620

243. Bernstein LR, Trahiotis C. How sensitivity to ongoing interaural temporal disparities is affected by manipulations of temporal features of the envelopes of high-frequency stimuli. J Acoust Soc Am. 2009;125: 3234–3242. doi:10.1121/1.3101454

244. Dietz M, Wang L, Greenberg D, McAlpine D. Sensitivity to interaural time differences conveyed in the stimulus envelope: Estimating inputs of binaural neurons through the temporal analysis of spike trains. J Assoc Res Otolaryngol. 2016;17: 313–330. doi:10.1007/s10162-016-0573-9

245. Laback B, Zimmermann I, Majdak P, Baumgartner W-D, Pok S-M. Effects of envelope shape on interaural envelope delay sensitivity in acoustic and electric hearing. J Acoust Soc Am. 2011;130: 1515–1529. doi:10.1121/1.3613704

246. Lee J. Amplitude modulation rate discrimination with sinusoidal carriers. J Acoust Soc Am. 1994;96: 2140–2147. doi:10.1121/1.410156

247. Ruggero MA, Robles L, Rich NC, Recio A. Basilar membrane responses to two-tone and broadband stimuli. Phil Trans R Soc Lond B. 1992;336: 307–315. doi:10.1098/rstb.1992.0063

248. Strickland EA, Viemeister NF. The effects of frequency region and bandwidth on the temporal modulation transfer function. J Acoust Soc Am. 1997;102: 1799–1810.

doi:10.1121/1.419617

249. Strickland EA, Dhar S. An analysis of quasi-frequency-modulated noise and random-sideband noise as comparisons for amplitude-modulated noise. J Acoust Soc Am. 2000;108: 735–742. doi:10.1121/1.429606

250. Eddins DA. Amplitude-modulation detection at low- and high-audio frequencies. J Acoust Soc Am. 1999;105: 829–837. doi:10.1121/1.426272

251. Krumbholz K, Patterson RD, Pressnitzer D. The lower limit of pitch as determined by rate discrimination. J Acoust Soc Am. 2000;108: 1170–1180. doi:10.1121/1.1287843

252. Lemańaska J, Sek̨ AP, Skrodzka EB. Discrimination of the amplitude modulation rate. Arch Acoust. 2002;27: 3–21.

253. Mcfadden D, Pasanen EG. Binaural beats at high frequencies. Science (80-). 1975;190: 394–396. doi:10.1126/science.1179219

254. Zhang C, Zeng F. Loudness of dynamic stimuli in acoustic and electric hearing. J Acoust Soc Am. 1997;102: 2925–2934. doi:10.1121/1.420347

255. Moore BCJ, Glasberg BR, Varathanathan A, Schlittenlacher J. A Loudness Model for Time-Varying Sounds Incorporating Binaural Inhibition. Trend Hear. 2016;20: 1–16. doi:10.1177/2331216516682698

256. Kreft HA, Nelson DA, Oxenham AJ. Modulation frequency discrimination with modulated and unmodulated interference in normal hearing and in cochlear-implant users. J Assoc Res Otolaryngol. 2013;14: 591–601. doi:10.1007/s10162-013-0391-2

257. Lutfi RA, Liu C. Individual differences in source identification from synthesized impact sounds. J Acoust Soc Am. 2007;122: 1017–1028. doi:10.1121/1.2751269

258. Kidd GR, Watson CS, Gygi B. Individual differences in auditory abilities. J Acoust Soc

Am. 2007;122: 418–435. doi:10.1121/1.2743154

259. Johnson DM, Watson CS, Jensen JK. Individual differences in auditory capabilities. I. J Acoust Soc Am. 1986;81: 427–438. doi:10.1121/1.394907

260. Bernstein LR, Trahiotis C. Detection of interaural delay in high-frequency sinusoidally amplitude-modulated tones, two-tone complexes, and bands of noise. J Acoust Soc Am. 1994;95: 3561–3567. doi:10.1121/1.409973

261. Schlittenlacher J, Moore BCJ. Discrimination of amplitude-modulation depth by subjects with normal and impaired hearing. J Acoust Soc Am. 2016;140: 3487–3495. doi:10.1121/1.4966117

262. Henry KS, Kale S, Heinz MG. Noise-induced hearing loss increases the temporal precision of complex envelope coding by auditory-nerve fibers. Front Neuro. 2014;8: 1–10. doi:10.3389/fnsys.2014.00020

263. Jennings SG, Chen J, Fultz SE, Ahlstrom JB, Dubno JR. Amplitude modulation detection with a short-duration carrier: Effects of a precursor and hearing loss. J Acoust Soc Am. 2018;143: 2232–2243. doi:10.1121/1.5031122

264. Garadat SN, Zwolan TA, Pfingst BE. Using temporal modulation sensitivity to select stimulation sites for processor MAPs in cochlear implant listeners. Audiol Neurootol. 2013;18. doi:10.1159/000351302

265. Bernstein JGW, Goupell MJ, Wess JM, Stakhovskaya OA, Brungart DS. Having two ears can facilitate or interfere with the perceptual separation of concurrent talkers for bilateral and single-sided deafness cochlear-implant listeners [Abstract]. J Acoust Soc Am. 2017;141: 4031. doi:10.1121/1.4989292

266. Chatterjee M. Modulation masking in cochlear implant listeners: envelope versus

tonotopic components. J Acoust Soc Am. 2003;113: 2042–2053. doi:10.1121/1.1555613

267. Chatterjee M, Oba SI. Noise improves modulation detection by cochlear implant listeners at moderate carrier levels. J Acoust Soc Am. 2005;118: 993–1002. doi:10.1121/1.1929258

268. Richardson LM, Busby PA, Clark GM, Richardson LM, Busby PA, Clark GM. Modulation detection interference in cochlear implant subjects. J Acoust Soc Am. 1998;442: 442–452. doi:10.1121/1.423248

269. Pierzycki RH, Seeber BU. Comodulation masking release in electric hearing. J Assoc Res Otolaryngol. 2014;15: 279–291. doi:10.1007/s10162-013-0433-9

270. Ihlefeld A, Shinn-Cunningham BG, Carlyon RP. Comodulation masking release in speech identification with real and simulated cochlear-implant hearing. J Acoust Soc Am. 2012;131: 1315–1324. doi:10.1121/1.3676701

271. Lavandier M, Culling JF. Speech segregation in rooms: Effects of reverberation on both target and interferer. J Acoust Soc Am. 2007;122: 1713–1723. doi:10.1121/1.2764469

272. Lavandier M, Culling JF. Speech segregation in rooms: Monaural, binaural, and interacting effects of reverberation on target and interferer. J Acoust Soc Am. 2008;123: 2237–2248. doi:10.1121/1.2871943

273. Anderson SR, Kan A, Litovsky RY. Asymmetric temporal envelope encoding: Implications for within- and across-ear envelope comparison. J Acoust Soc Am. 2019;146: 1189–1206. doi:10.1121/1.5121423

274. Zhou R, Assouline JG, Abbas PJ. Anatomical and physiological measures of auditory system in mice with peripheral myelin deficiency. Hear Res. 1995;88: 87–97. doi:10.1016/0378-5955(95)00104-c

275. Monaghan JJM, Bleeck S, McAlpine D. Sensitivity to envelope interaural time differences

at high modulation rates. Trend Hear. 2015;19: 1–14. doi:10.1177/2331216515619331

276. van Hoesel RJM, Jones GL, Litovsky RY. Interaural time delay sensitivity in Bilateral Cochear Implant Users: Effects of Pulse Rate, Modulation Rate, and Place of Stimulation. J Assoc Res Otolaryngol. 2009;10: 557–567. doi:10.1007/s10162-009-0175-x

277. Wilson BS. Thirty years of the British Journal of Audiology: Guest editorial. The future of cochlear implants. Br J Audiol. 1997;31: 205–225. doi:10.3109/03005369709076795

278. Busby PA, Tong YC, Clark GM. The perception of temporal modulations by cochlear implant patients. J Acoust Soc Am. 1993;94: 124–131. doi:10.1121/1.408212

279. McKay CM, McDermott HJ, Clark GM. Pitch percepts associated with amplitude-modulated current pulse trains in cochlear implantees. J Acoust Soc Am. 1994;96: 2664–2673. doi:10.1121/1.411377

280. Litovsky RY, Goupell MJ, Godar S, Grieco-Calub T, Jones GL, Garadat SN, et al. Studies on bilateral cochlear implants at the University of Wisconsin Binaural Hearing and Speech Lab. J Am Acad Audiol. 2012;23: 476–494. doi:10.3766/jaaa.23.6.9

281. Kreft HA, Oxenham AJ, Nelson DA. Modulation rate discrimination using half-wave rectified and sinusoidally amplitude modulated stimuli in cochlear-implant users. J Acoust Soc Am. 2010;127: 656–659. doi:10.1121/1.3282947

282. Levitt H. Transformed up-down methods in psychoacoustics. J Acoust Soc Am. 1971;49: 467–477. doi:10.1121/1.1912375

283. Macmillan NA, Creelman CD. Detection Theory: A User's Guide. 2nd Editio. Cambridge, England: Cambridge University Press; 2005.

284. Wichmann AF, Hill NJ. The psychometric function: I. Fitting, sampling, and goodness of fit. Percept Psyhophys. 2001;63: 1293–1313. doi:10.3758/BF03194544

285. Bates D, Machler M, Bolker B, Walker S. Fitting linear mixed-effects models using {lme4}. J Stat Soft. 2015;67: 1–48. doi:10.18637/jss.v067.i01

286. Kenward MG, Roger JH. Small sample inference for fixed effects from restricted maximum likelihood. Biometrics. 1997;53: 983–997. doi:2533558

287. Green T, Faulkner A, Rosen S. Variations in carrier pulse rate and the perception of amplitude modulation in cochlear implant users. Ear Hear. 2012;33: 221–230. doi:10.1097/AUD.0b013e318230fff8

288. Bland JM, Altman DG. Calculating correlation coefficients with repeated observations: Part 1-correlation within subjects. Br Med J. 1995;310: 446–446. doi:10.1785/gssrl.82.5.686

289. Joris PX, Schreiner CE, Rees A. Neural processing of amplitude-modulated sounds. Physiol Rev. 2004;84: 541–577. doi:10.1152/physrev.00029.2003

290. Oxenham AJ. Pitch Perception and Auditory Stream Segregation: Implications for Hearing Loss and Cochlear Implants. Trends Amplif. 2008;12: 316–331. doi:10.1177/1084713808325881

291. Schvartz-Leyzac KC, Zwolan TA, Pfingst BE. Effects of electrode deactivation on speech recognition in multichannel cochlear implant recipients. Cochlear Implant Int. 2017;18: 324–334. doi:10.1080/14670100.2017.1359457.Effects

292. Dietz M, Klein-Hennig M, Hohmann V. The influence of pause, attack, and decay duration of the ongoing envelope on sound lateralization. J Acoust Soc Am. 2015;137: EL137-EL143. doi:10.1121/1.4905891

293. Klein-Hennig M, Dietz M, Hohmann V, Ewert SD. The influence of different segments of the ongoing envelope on sensitivity to interaural time delays. J Acoust Soc Am. 2011;129:

3856–3872. doi:10.1121/1.3585847

294. Dynes SBC, Delgutte B. Phase-locking of auditory-nerve discharges to sinusoidal electric stimulation of the cochlea. Hear Res. 1992;58: 79–90. doi:10.1016/0378-5955(92)90011-B

295. Javel E, Shepherd RK. Electrical stimulation of the auditory nerve in man. Hear Res. 2000;140: 45–76. doi:10.1097/00043764-196705000-00018

296. Vollmer M. Neural processing of acoustic and electric interaural time differences in normal-hearing gerbils. J Neurosci. 2018;38: 6949–6966. doi:10.1523/JNEUROSCI.3328-17.2018

297. He S, Teagle HFB, Buchman CA. The Electrically Evoked Compound Action Potential : From Laboratory to Clinic. 2017;11: 1–20. doi:10.3389/fnins.2017.00339

298. Anderson ES, Oxenham AJ, Nelson PB, Nelson DA. Assessing the role of spectral and intensity cues in spectral ripple detection and discrimination in cochlear-implant users. J Acoust Soc Am. 2012;132: 3925–3934. doi:10.1121/1.4763999

299. Gallun FJ, McMillan GP, Molis MR, Kampel SD, Dann SM, Konrad-Martin DL. Relating age and hearing loss to monaural, bilateral, and binaural temporal sensitivity. Front Neuro. 2014;8: 1–14. doi:10.3389/fnins.2014.00172

300. Grose JH, Mamo SK. Processing of temporal fine structure as a function of age. Ear Hear. 2010;31: 755–760. doi:10.1097/AUD.0b013e3181e627e7

301. Abdelsattar ZM, Wong SL, Regenbogen SE, Jomaa DM, Hardiman KM, Hendren S. Colorectal cancer outcomes and treatment patterns in patients too young for average-risk screening. Cancer. 2016;122: 929–934. doi:10.1002/cncr.29716

302. Admon LK, Winkelman TNA, Zivin K, Terplan M, Mhyre JM, Dalton VK. Racial and

ethnic disparities in the incidence of severe maternal morbidity in the United States, 2012-2015. Obs Gynecol. 2018;132: 1158–1166. doi:10.1097/AOG.0000000000002937

303. Dess RT, Hartman HE, Mahal BA, Soni PD, Jackson WC, Cooperberg MR, et al. Association of Black Race with Prostate Cancer-Specific and Other-Cause Mortality. JAMA Oncol. 2019;5: 975–983. doi:10.1001/jamaoncol.2019.0826

304. ASHA. A demographic snapshot of audiologists. ASHA Lead. 2019;24: 32. doi:10.1044/leader.AAG.24082019.32

305. Macpherson EA, Middlebrooks JC. Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited. J Acoust Soc Am. 2002;111: 2219–2236. doi:10.1121/1.1471898

306. Bronkhorst AW, Plomp R. The effect of head-induced interaural time and level differences on speech intelligibility in noise. J Acoust Soc Am. 1988;83: 1508–1516. doi:10.1121/1.395906

307. Ihlefeld A, Litovsky RY. Interaural Level Differences Do Not Suffice for Restoring Spatial Release from Masking in Simulated Cochlear Implant Listening. PLoS One. 2012;7: 1–9. doi:10.1371/journal.pone.0045296

308. Archer-Boyd AW, Carlyon RP. Simulations of the effect of unlinked cochlear-implant automatic gain control and head movement on interaural level differences. J Acoust Soc Am. 2019;145: 1389–1400. doi:10.1121/1.5093623

309. Potts WB, Ramanna L, Perry T, Long CJ. Improving localization and speech reception in noise for bilateral cochlear implant recipients. Trend Hear. 2019;23: 1–18. doi:10.1177/2331216519831492

310. van Hoesel RJM. Sensitivity to binaural timing in bilateral cochlear implant users. J

Acoust Soc Am. 2007;121: 2192–2206. doi:10.1121/1.2537300

311. Noel VA, Eddington DK. Sensitivity of bilateral cochlear implant users to fine-structure and envelope interaural time differences. J Acoust Soc Am. 2013;133: 2314–2328. doi:10.1121/1.4794372

312. Lord Rayleigh OMPRS. XII. On our perception of sound direction. Lond Edinb Dubl Phil Mag. 1907;13: 214–232. doi:10.1080/14786440709463595

313. Zwislocki J, Feldman RS. Just Noticeable Differences in Dichotic Phase. J Acoust Soc Am. 1956;28: 860–864. doi:10.1121/1.1908495

314. Bernstein LR, Trahiotis C. Lateralization produced by envelope-based interaural temporal disparities of high-frequency, raised-sine stimuli: Empirical data and modeling. J Acoust Soc Am. 2011;129: 1501–1508. doi:10.1121/1.3552875

315. Yin TC, Chan JC. Interaural time sensitivity in medial superior olive of cat. J Neurophysiol. 1990;64: 465–488. doi:10.1152/jn.1990.64.2.465

316. Joris PX, van der Heijden M. Early Binaural Hearing: The Comparison of Temporal Differences at the Two Ears. Ann Rev Neuro. 2019;42: 433–457. doi:10.1146/annurev-neuro-080317-061925

317. Resnick JM, O’Brien GE, Rubinstein JT. Simulated auditory nerve axon demyelination alters sensitivity and response timing to extracellular stimulation. Hear Res. 2018;361: 121–137. doi:10.1016/j.heares.2018.01.014

318. Fu QJ. Loudness growth in cochlear implants: Effect of stimulation rate and electrode configuration. Hear Res. 2005;202: 55–62. doi:10.1016/j.heares.2004.10.004

319. Fitzgerald MB, Kan A, Goupell MJ. Bilateral Loudness Balancing and Distorted Spatial Perception in Recipients of Bilateral Cochlear Implants. Ear Hear. 2015;36: e225–e236.

doi:10.1097/AUD.0000000000000174

320. Saberi K, Takahashi Y, Konishi M, Albeck Y, Arthur BJ, Farahbod H. Effects of interaural decorrelation on neural and behavioral detection of spatial cues. Neuron. 1998;21: 789–798. doi:10.1016/S0896-6273(00)80595-4

321. Stecker GC. Temporal weighting functions for interaural time and level differences. V. Modulated noise carriers. J Acoust Soc Am. 2018;143: 686–695. doi:10.1121/1.5022785

322. Hafter ER, Buell TN. Restarting the adapted binaural system. J Acoust Soc Am. 1990;88: 806–812. doi:10.1121/1.399730

323. Hu H, Ewert SD, McAlpine D, Dietz M. Differences in the temporal course of interaural time difference sensitivity between acoustic and electric hearing in amplitude modulated stimuli. J Acoust Soc Am. 2017;141: 1862–1873. doi:10.1121/1.4977014

324. Poon BB, Eddington DK, Noel V, Colburn HS. Sensitivity to interaural time difference with bilateral cochlear implants: Development over time and effect of interaural electrode spacing. J Acoust Soc Am. 2009;126: 806–815. doi:10.1121/1.3158821

325. Blanks DA, Roberts JM, Buss E, Hall JW, Fitzpatrick DC. Neural and behavioral sensitivity to interaural time differences using amplitude modulated tones with mismatched carrier frequencies. J Assoc Res Otolaryngol. 2007;8: 393–408. doi:10.1007/s10162-007-0088-5

326. Blanks DA, Buss E, Grose JH, Fitzpatrick DC, Hall JW. Interaural time discrimination of envelopes carried on high-frequency tones as a function of level and interaural carrier mismatch. Ear Hear. 2008;29: 674–683. doi:10.1097/AUD.0b013e3181775e03

327. Nuetzel JM, Hafter ER. Discrimination of interaural delays in complex waveforms: Spectral effects. J Acoust Soc Am. 1981;69: 1112–1118. doi:10.1121/1.385690

328. Henning GB. Detectability of interaural delay in high-frequency complex waveforms. J Acoust Soc Am. 1974;55: 84–90. doi:10.1121/1.1928135

329. Klug J, Schmors L, Ashida G, Dietz M. Neural rate difference model can account for lateralization of high-frequency stimuli. J Acoust Soc Am. 2020;148: 678–691. doi:10.1121/10.0001602

330. Bernstein LR, Trahiotis C. Lateralization produced by interaural temporal and intensitive disparities of high-frequency, raised-sine stimuli: Data and modeling. J Acoust Soc Am. 2011;129: 1501–1508. doi:10.1121/1.3662056

331. Dietz M. Models of the electrically stimulated binaural system: A review. Netw Comput Neural Syst. 2016;27: 186–211. doi:10.1080/0954898X.2016.1219411

332. Bruce IC, Erfani Y, Zilany MSA. A phenomenological model of the synapse between the inner hair cell and auditory nerve: Implications of limited neurotransmitter release sites. Hear Res. 2018;360: 40–54. doi:10.1016/j.heares.2017.12.016

333. Goldberg JM, Brown PB. Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. J Neurophysiol. 1969;32: 613–636. doi:10.1007/978-1-4612-2700-7_3

334. Stellmack MA, Viemeister NF, Byrne AJ. Discrimination of interaural phase differences in the envelopes of sinusoidally amplitude-modulated 4-kHz tones as a function of modulation depth. J Acoust Soc Am. 2005;118: 346–352. doi:10.1121/1.1923370

335. Goupell MJ, Majdak P, Laback B. Median-plane sound localization as a function of the number of spectral channels using a channel vocoder. J Acoust Soc Am. 2010;127: 990–1001. doi:10.1121/1.3283014

336. Greenwood DD. A cochlear frequency-position function for several species—29 years

later. J Acoust Soc Am. 1990;87: 2592–2605. doi:10.1121/1.399052

337. Goupell MJ, Best V, Colburn HS. Intracranial lateralization bias observed in the presence of symmetrical hearing thresholds. JASA Express Lett. 2021;1: 104401. doi:10.1121/10.0006720

338. Akeroyd MA. Threshold differences for interaural time delays carried by double vowels. J Acoust Soc Am. 2003;114: 2167–2177. doi:10.1121/1.1611884

339. Jeffress LA. A place theory of sound localization. J Comp Physiol Psychol. 1948;41: 35–39. doi:10.1037/h0061495

340. Bernstein LR, Trahiotis C. Lateralization produced by interaural temporal and intensitive disparities of high-frequency, raised-sine stimuli: Data and modeling. J Acoust Soc Am. 2012;131: 409–415. doi:10.1121/1.3662056

341. Stern RM, Zeiberg AS, Trahiotis C. Lateralization of complex binaural stimuli: A weighted-image model. J Acoust Soc Am. 1988;84: 156–165. doi:10.1121/1.396982

342. Stern RM, Shear GD. Lateralization and detection of low-frequency binaural stimuli: Effects of distribution of internal delay. J Acoust Soc Am. 1996;100: 2278–2288. doi:10.1121/1.417937

343. Ashida G, Kretzberg J, Tollin DJ. Roles for Coincidence Detection in Coding Amplitude-Modulated Sounds. PLoS Comp Biol. 2016;12: e1004997. doi:10.1371/journal.pcbi.1004997

344. Hafter ER, Dye RH. Detection of interaural differences of time in trains of high-frequency clicks as a function of interclick interval and number. J Acoust Soc Am. 1983;73: 644–651. doi:10.1121/1.388956

345. Reeder RM, Firszt JB, Holden LK, Strube MJ. A longitudinal study in adults with

sequential bilateral cochlear implants: Time course for individual ear and bilateral performance. J Speech Lang Hear Res. 2014;57: 1108–1126. doi:10.1044/2014

346. Firszt JB, Holden LK, Reeder RM, Waltzman SB, Arndt S. Auditory abilities after cochlear implantation in adults with unilateral deafness: A pilot study. Otol Neurotol. 2012;33: 1339–1346. doi:10.1097/MAO.0b013e318268d52d

347. Gifford RH, Dorman MF, McKarns SA, Spahr AJ. Combined electric and contralateral acoustic hearing: Word and sentence recognition with bimodal hearing. J Speech Lang Hear Res. 2007;50: 835–843. doi:10.1044/1092-4388(2007/058)

348. Sheffield SW, Gifford RH. The Benefits of Bimodal Hearing: Effect of Frequency Region and Acoustic Bandwidth. Audiol Neurootol. 2014;19: 151–163. doi:10.1159/000357588.The

349. Gifford RH, Davis TJ, Sunderhaus LW, Menapace C, Buck B, Crosson J, et al. Combined Electric and Acoustic Stimulation With Hearing Preservation. Ear Hear. 2017;38: 539–553.

350. Dietz M, Bernstein LR, Trahiotis C, Ewert SD, Hohmann V. The effect of overall level on sensitivity to interaural differences of time and level at high frequencies. J Acoust Soc Am. 2013;134: 494–502. doi:10.1121/1.4807827

351. Goupell MJ, Fong S, Stakhovskaya O. The effect of envelope modulations on binaural processing. Hear Res. 2019;379: 117–127. doi:10.1016/j.heares.2019.05.003

352. Ashida G, Tollin DJ, Kretzberg J. Physiological models of the lateral superior olive. PLoS One Comp Biol. 2017;13: e1005903. doi:10.1371/journal.pcbi.1005903

353. Rule SJ. Lateralization of high frequency, amplitude modulated sounds as a function of carrier frequency. Can J Exp Psychol. 1994;48: 435–443. doi:10.1037/1196-

1961.48.3.435

354. Grantham DW, Ashmead DH, Ricketts TA, Labadie RF, Haynes DS. Horizontal-plane localization of noise and speech signals by postlingually deafened adults fitted with bilateral cochlear implants. Ear Hear. 2007;28: 524–541. doi:10.1097/AUD.0b013e31806dc21a

355. Jones HG, Kan A, Litovsky RY. The effect of microphone placement on interaural level differences and sound localization across the horizontal plane in bilateral cochlear implant users. Ear Hear. 2016;37: e341–e345. doi:10.1097/AUD.0000000000000297

356. Nelson DA, Donaldson GS, Kreft H. Forward-masked spatial tuning curves in cochlear implant users. J Acoust Soc Am. 2008;123: 1522–1543. doi:10.1121/1.2836786

357. Croghan NBH, Smith ZM. Speech Understanding With Various Maskers in Cochlear-Implant and Simulated Cochlear-Implant Hearing: Effects of Spectral Resolution and Implications for Masking Release. Trend Hear. 2018;22: 1–13. doi:10.1177/2331216518787276

358. Mills AW. Lateralization of High-Frequency Tones. J Acoust Soc Am. 1960;32: 132–134. doi:10.1121/1.1907864

359. Frijns JHM, Klop WMC, Bonnet RM, Briaire JJ. Optimizing the number of electrodes with high-rate stimulation of the clarion CII cochlear implant. Acta Otolaryngol. 2003;123: 138–142. doi:10.1080/0036554021000028126

360. Loizou PC, Poroy O, Dorman M. The effect of parametric variations of cochlear implant processors on speech understanding. J Acoust Soc Am. 2000;108: 790–802. doi:10.1121/1.429612

361. Bernstein JGW, Jensen KK, Stakhovskaya OA, Noble JH, Hoa M, Kim HJ, et al.

Interaural Place-of-Stimulation Mismatch Estimates Using CT Scans and Binaural Perception, But Not Pitch, Are Consistent in Cochlear-Implant Users. J Neurosci. 2021; In press.

362. Goupell MJ, Cosentino S, Stakhovskaya OA, Bernstein JGW. Interaural Pitch-Discrimination Range Effects for Bilateral and Single-Sided-Deafness Cochlear-Implant Users. J Assoc Res Otolaryngol. 2019. doi:10.1007/s10162-018-00707-x

363. Hu H, Dietz M. Comparison of Interaural Electrode Pairing Methods for Bilateral Cochlear Implants. Trend Hear. 2015;19: 1–22. doi:10.1177/2331216515617143

364. Satterthwaite FE. An approximate distribution of estimates of variance components. Int Biom Soc. 1946;2: 110–114. doi:10.1002/9780470057339.vai016

365. DeVries L, Scheperle R, Bierer JA. Assessing the Electrode-Neuron Interface with the Electrically Evoked Compound Action Potential, Electrode Position, and Behavioral Thresholds. JARO - J Assoc Res Otolaryngol. 2016;17: 237–252. doi:10.1007/s10162-016-0557-9

366. Schvartz-Leyzac KC, Zwolan TA, Pfingst BE. Using the electrically-evoked compound action potential (ECAP) interphase gap effect to select electrode stimulation sites in cochlear implant users. Hear Res. 2021;406: 108257. doi:10.1016/j.heares.2021.108257

367. Pfingst BE, Colesa DJ, Hembrador S, Kang SY, Middlebrooks JC, Raphael Y, et al. Detection of pulse trains in the electrically stimulated cochlea: Effects of cochlear health. J Acoust Soc Am. 2011;130: 3954–3968. doi:10.1121/1.3651820

368. Cherry EC. Some experiments on the recognition of speech, with one and with two ears. J Acoust Soc Am. 1953;25: 975–979. doi:10.1121/1.1907229

369. Brungart DS, Simpson BD. Within-ear and across-ear interference in a cocktail-party

listening task. J Acoust Soc Am. 2002;112: 2985–2995. doi:10.1121/1.1512703

370. Archer-Boyd AW, Southwell R V., Deeks JM, Turner RE, Carlyon RP. Development and validation of a spectro-temporal processing test for cochlear-implant listeners. J Acoust Soc Am. 2018;144: 2983–2997. doi:10.1121/1.5079636

371. Goehring T, Archer-Boyd AW, Arenberg JG, Carlyon RP. The effect of increased channel interaction on speech perception with cochlear implants. Nat Sci Rep. 2021;11: 1–9. doi:10.1038/s41598-021-89932-8

372. Fu QJ. Temporal processing and speech recognition in cochlear implant users. Neuroreport. 2002;13: 1635–1639. doi:10.1097/00001756-200209160-00013

373. Zwolan TA, Collins LM, Wakefield GH. Electrode discrimination and speech recognition in postlingually deafened adult cochlear implant subjects. J Acoust Soc Am. 1997;102: 3673–3685. doi:10.1121/1.420401

374. Spahr AJ, Dorman MF, Loiselle LH. Performance of patients using different cochlear implant systems: Effects of input dynamic range. Ear Hear. 2007;28: 260–275. doi:10.1097/AUD.0b013e3180312607

375. Firszt JB, Chambers RD, Kraus N. Neurophysiology of cochlear implant users II: Comparison among speech perception, dynamic range, and physiological measures. Ear Hear. 2002;23: 516–531. doi:10.1097/00003446-200212000-00003

376. Loizou PC, Dorman MF, Fitzke J. The effect of reduced dynamic range on speech understanding: Implications for patients with cochlear implants. Ear Hear. 2000;21: 25–31. doi:10.1097/00003446-200002000-00006

377. Noble JH, Gifford RH, Hedley-Williams AJ, Dawant BM, Labadie RF. Clinical evaluation of an image-guided cochlear implant programming strategy. Audiol Neurootol.

2014;19: 400–411. doi:10.1159/000365273

378. Garcia C, Goehring T, Cosentino S, Turner RE, Deeks JM, Brochier T, et al. The Panoramic ECAP Method: Estimating Patient-Specific Patterns of Current Spread and Neural Health in Cochlear Implant Users. J Assoc Res Otolaryngol. 2021;22: 567–589. doi:10.1007/s10162-021-00795-2

379. Xu K, Willis S, Gopen Q, Fu QJ. Effects of Spectral Resolution and Frequency Mismatch on Speech Understanding and Spatial Release from Masking in Simulated Bilateral Cochlear Implants. Ear Hear. 2020;41: 1362–1371. doi:10.1097/AUD.0000000000000865

380. Siciliano CM, Faulkner A, Rosen S, Mair K. Resistance to learning binaurally mismatched frequency-to-place maps: Implications for bilateral stimulation with cochlear implants. J Acoust Soc Am. 2010;127: 1645–1660. doi:10.1121/1.3293002

381. Scharf B. Localization of unlike tones from two loudspeakers. Sensation and Measurement. 1974. pp. 309–314. doi:10.1007/978-94-010-2245-3_30

382. Darwin CJ. Perceptual grouping of speech components differing in fundamental frequency and onset-time. Q J Exp Psychol. 1981;33: 185–207. doi:10.1080/14640748108400785

383. van den Brink G, Sintnicolaas K, van Stam WS. Dichotic pitch fusion. J Acoust Soc Am. 1976;59: 1471–1476. doi:10.1121/1.380989

384. Westerhausen R, Moosmann M, Alho K, Medvedev S, Hämäläinen H, Hugdahl K. Top-down and bottom-up interaction: manipulating the dichotic listening ear advantage. Brain Res. 2009;1250: 183–189. doi:10.1016/j.brainres.2008.10.070

385. Westerhausen R, Bless J, Kompus K. Behavioral laterality and aging: The free-recall dichotic-listening right-ear advantage increases with age. Dev Neuropsychol. 2015;40: 313–327. doi:10.1080/87565641.2015.1073291

386. Cutting JE. Aspects of phonological fusion. J Exp Psychol Hum Percept Perform. 1975;104: 105–120. doi:10.1037/0096-1523.1.2.105

387. Cutting JE. Auditory and linguistic processes in speech perception: Inferences from six fusions in dichotic listening. Psych Rev. 1976;83: 114–140. doi:10.1037/0033-295X.83.2.114

388. Pumplin J. Low-noise noise. J Acoust Soc Am. 1985;78: 100–104. doi:10.1121/1.392571

389. Woods KJP, Siegel MH, Traer J, McDermott JH. Headphone screening to facilitate web-based auditory experiments. Atten Percept Psychophys. 2017;79: 2064–2072. doi:10.3758/s13414-017-1361-2

390. Staisloff HE, Lee DH, Aronoff JM. Perceptually aligning apical frequency regions leads to more binaural fusion of speech in a cochlear implant simulation. Hear Res. 2016;337: 59–64. doi:10.1016/j.heares.2016.05.002

391. Aronoff JM, Shayman C, Prasad A, Suneel D, Stelmach J. Unilateral spectral and temporal compression reduces binaural fusion for normal hearing listeners with cochlear implant simulations. Hear Res. 2015;320: 24–29. doi:10.1016/j.heares.2014.12.005

392. Kaufmann CR, Henslee AM, Claussen A, Hansen MR. Evaluation of Insertion Forces and Cochlea Trauma following Robotics-Assisted Cochlear Implant Electrode Array Insertion. Otol Neurotol. 2020;41: 631–638. doi:10.1097/MAO.0000000000002608