ALTERATIONS IN THE SOUND LOCALIZATION PATHWAY RELATED TO IMPAIRED COCKTAIL-PARTY PERFORMANCE

EDITED BY: Achim Klug, Leonard Kaczmarek, Elizabeth Anne McCullagh and Dan Tollin

PUBLISHED IN: Frontiers in Neuroscience

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88976-159-3
DOI 10.3389/978-2-88976-159-3

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

ALTERATIONS IN THE SOUND LOCALIZATION PATHWAY RELATED TO IMPAIRED COCKTAIL-PARTY PERFORMANCE

Topic Editors:
Achim Klug, University of Colorado, United States
Leonard Kaczmarek, Yale University, United States
Elizabeth Anne McCullagh, Oklahoma State University, United States
Dan Tollin, University of Colorado, United States

Citation: Klug, A., Kaczmarek, L., McCullagh, E. A., Tollin, D., eds. (2022). Alterations in the Sound Localization Pathway Related to Impaired Cocktail-Party Performance. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88976-159-3

Table of Contents

frontiers | Frontiers in Neuroscience

EDITORIAL
published: 25 April 2022
doi: 10.3389/fnins.2022.902197

Editorial: Alterations in the Sound Localization Pathway Related to Impaired Cocktail-Party Performance

Elizabeth A. McCullagh[1]*, *Leonard K. Kaczmarek*[2,3], *Daniel J. Tollin*[4] *and Achim Klug*[4]

[1] *Department of Integrative Biology, Oklahoma State University, Stillwater, OK, United States,* [2] *Department of Pharmacology, Yale University, New Haven, CT, United States,* [3] *Department of Cellular and Molecular Physiology, Yale University, New Haven, CT, United States,* [4] *Department of Physiology and Biophysics, University of Colorado Anschutz Medical Campus, Aurora, CO, United States*

Keywords: sound localization, cocktail party effect, auditory, brainstem, interaural time and intensity differences

Editorial on the Research Topic

Alterations in the Sound Localization Pathway Related to Impaired Cocktail-Party Performance

Binaural and spatial hearing allows us to localize the source of a sound and to function in complex acoustic environments. In noisy environments we typically focus on one sound source, e.g., our conversation partner, and ignore background noises. The sound localization pathway in the auditory brainstem contributes to this ability by associating various competing sounds with their respective spatial channel. Normal hearing humans and many animals can localize and discriminate sound sources with a precision of just a few degrees. This is accomplished by comparing the interaural time difference (ITD) and interaural intensity difference (IID) that a sound creates between the two ears, which vary systematically with the location of the sound source in space. Even small alterations in this brainstem circuit can have major effects on one's ability to function in an acoustically busy environment. This special edition highlights some of these alterations in animal models and human subjects and discusses medical conditions associated with impaired hearing in noise.

OPEN ACCESS

Edited and reviewed by:
Robert J. Zatorre,
McGill University, Canada

***Correspondence:**
Elizabeth A. McCullagh
elizabeth.mccullagh@okstate.edu

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience

Received: *22 March 2022*
Accepted: *30 March 2022*
Published: *25 April 2022*

Citation:
McCullagh EA, Kaczmarek LK, Tollin DJ and Klug A (2022) Editorial: Alterations in the Sound Localization Pathway Related to Impaired Cocktail-Party Performance. Front. Neurosci. 16:902197. doi: 10.3389/fnins.2022.902197

DISORDERED SOUND LOCALIZATION PERFORMANCE AND MECHANISMS

Importantly, the alterations discussed here affect central neural circuits and are not dependent on decreased ability to detect sounds; in many cases, however, a patient might experience both peripheral and central hearing loss. A combination of alterations along the ascending auditory pathway makes diagnosis and treatment of this condition very challenging. A review by Gallun discusses the clear need for better diagnostic tools, including behavioral and neurophysiological tests to determine the specific alterations in any particular case.

Koerner et al. show the dependence and variance of physiological and behavioral measurements in a common form of age-related hearing loss alters CNS binaural circuits. Affected listeners have clinically normal hearing thresholds but struggle to understand speech in background noise. Alterations in the binaural system of these listeners cause impaired processing of temporally fast and precise binaural cues that can be detected with electrophysiological measurements that are directly related to the behavioral ability to decode binaural cues correctly. This suggests

that non-invasive physiological tests can potentially be used to quantify behavioral difficulties in affected listeners.

The ability to localize sounds is also severely disrupted in autism spectrum (ASD) disorder and inherited forms of intellectual disability such as Fragile X syndrome (FXS). FXS is caused by loss of Fragile X Mental Retardation Protein, FMRP, an mRNA-binding protein that controls translation and also regulates neural excitability by binding ion channels (Wu and Kaczmarek). An overwhelming majority of patients with FXS and ASD are hypersensitive to auditory stimuli and have difficulty in distinguishing speech sounds from background noise. As reviewed by Rotschafer, these abnormalities of auditory processing can often be detected by electric and magnetic signals recorded from the cerebral cortex of humans.

One of the biological factors altered in ASD may be the speed and precision at which auditory brainstem neurons propagate action potentials. Using a mouse model of FXS, Lucas et al. demonstrate that in the medial nucleus of the trapezoid body, a key brainstem relay for transmission of both IID and ITDs, loss of FMRP reduces both the diameter of axons and the thickness of the myelin sheath. A complementary computational investigation by Li et al., modeled such changes in myelin thickness and conduction velocity in a brainstem network. They analyzed firing patterns in response to sinusoidal tones and natural sounds and calculated tuning curves for ITDs in the medial superior olive, where the timing of inputs from the two ears is compared. The combined experimental/computational studies make a strong case that, by reducing the rate at which auditory information is propagated through the brainstem, impaired myelination disrupts accurate comparisons of ITDs in FXS.

In addition to genetic mutations, experiencing incorrect binaural cues during development may impair the high level of precision required for sound localization. The finding that abnormal early sound experiences can result in binaural neurons that incorrectly code for spatial location even in adulthood (Thornton et al.), underscores the importance of early interventions to hearing loss.

MODULATION OF BRAINSTEM CIRCUITRY IN COMPLEX ACOUSTIC ENVIRONMENTS

Wu and Kaczmarek review the modulation of potassium channels in auditory brainstem neurons in response to changes in the auditory environment. They describe how insights into the role of specific channels have come from human gene mutations that impair localization of sounds in space. Additionally, they review how short-term and long-term modulation of channels maximizes the extraction of auditory information. Among these channels is the Kv3.3 potassium channel, which is further discussed by Middlebrooks and Waters who describe a family with a Kv3.3 mutation. The affected family members exhibited severe loss of sensitivity for ITDs and ILDs, which almost certainly degrades their ability to segregate competing sounds in the real-world.

Middlebrooks and Waters further review the mechanisms by which human and animal listeners segregate competing sequences of sounds from sources separated by as little as 10°. Neurons in the auditory cortex tend to synchronize selectively to one of two such competing sequences. The ability to spatially resolve these stimuli depends on the binaural and monaural acoustical cues provided in the various experimental conditions. This contrasts with a different measure of sound-source localization, the minimum audible angle, which is constant across those conditions, suggesting that the central spatial mechanisms for stream segregation differ from those for sound localization.

Finally, Auerbach and Gritton review studies of the variety of different adaptive mechanisms by which information is extracted from complex acoustic and highly variable listening conditions. These mechanisms include both "bottom-up" gain alterations in response to changes in environmental sound statistics as well as "top-down" mechanisms that allow for selective extraction of specific sound features in a complex auditory scene. The review concludes with an examination of how hearing loss interacts with these gain control mechanisms.

In summary, this special edition highlights the importance of the auditory brainstem sound localization circuit in extracting sound source locations in space using a specialized area of the brain that is fine-tuned for temporal precision. Importantly, this circuit is functionally involved in disordered spatial hearing in complex conditions such as FXS, ASD, and aging among others.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

NIH 1R15HD105231-01 (EAM); NIH DC01919, NS102239, and NS111242 (LKK); NIDCD R01-011555 and NIDCD R01-017924 (DJT and AK); and NIDCD R01-018401 (AK).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Neuroscience

REVIEW
published: 16 September 2020
doi: 10.3389/fnins.2020.571095

Spatial Mechanisms for Segregation of Competing Sounds, and a Breakdown in Spatial Hearing

John C. Middlebrooks[1] and Michael F. Waters[2]*

[1] Departments of Otolaryngology, Neurobiology and Behavior, Cognitive Sciences, and Biomedical Engineering, University of California, Irvine, Irvine, CA, United States, [2] Department of Neurology, Barrow Neurological Institute, Phoenix, AZ, United States

We live in complex auditory environments, in which we are confronted with multiple competing sounds, including the cacophony of talkers in busy markets, classrooms, offices, etc. The purpose of this article is to synthesize observations from a series of experiments that focused on how spatial hearing might aid in disentangling interleaved sequences of sounds. The experiments were unified by a non-verbal task, "rhythmic masking release", which was applied to psychophysical studies in humans and cats and to cortical physiology in anesthetized cats. Human and feline listeners could segregate competing sequences of sounds from sources that were separated by as little as ∼10°. Similarly, single neurons in the cat primary auditory cortex tended to synchronize selectively to sound sequences from one of two competing sources, again with spatial resolution of ∼10°. The spatial resolution of spatial stream segregation varied widely depending on the binaural and monaural acoustical cues that were available in various experimental conditions. This is in contrast to a measure of basic sound-source localization, the minimum audible angle, which showed largely constant acuity across those conditions. The differential utilization of acoustical cues suggests that the central spatial mechanisms for stream segregation differ from those for sound localization. The highest-acuity spatial stream segregation was derived from interaural time and level differences. Brainstem processing of those cues is thought to rely heavily on normal function of a voltage-gated potassium channel, Kv3.3. A family was studied having a dominant negative mutation in the gene for that channel. Affected family members exhibited severe loss of sensitivity for interaural time and level differences, which almost certainly would degrade their ability to segregate competing sounds in real-world auditory scenes.

Keywords: spatial release from masking, interaural time difference (ITD), interaural level difference (ILD), cat, rhythmic masking release, cerebellar ataxia, Kv3.3

OPEN ACCESS

Edited by:
Dan Tollin, University of Colorado School of Medicine, United States

Reviewed by:
Frederick Jerome Gallun, Oregon Health & Science University, United States
Ian D. Forsythe, University of Leicester, United Kingdom

****Correspondence:***
John C. Middlebrooks
j.midd@uci.edu

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience

Received: *09 June 2020*
Accepted: *21 August 2020*
Published: *16 September 2020*

Citation:
Middlebrooks JC and Waters MF (2020) Spatial Mechanisms for Segregation of Competing Sounds, and a Breakdown in Spatial Hearing. Front. Neurosci. 14:571095. doi: 10.3389/fnins.2020.571095

INTRODUCTION

Everyday listening situations require us to isolate sounds of interest amid competing sounds. The classic example is the "cocktail party problem" (Cherry, 1953), but more quotidian examples include busy offices, classrooms, restaurants, etc. Spatial hearing has long been thought to aid in sorting out these complex auditory scenes. For instance, Cherry listed "the voices come from

different directions" as a likely factor in recognizing what one person is saying when others are speaking (Cherry, 1953). *Spatial release from masking* refers to the condition in which detection or recognition of a sound of interest, the *target*, is enhanced when the target source is separated in space from sources of competing sounds, the *maskers* (Hirsh, 1950; Zurek, 1993; Kidd et al., 1998).

Spatial hearing can be especially beneficial in the task of *stream segregation* (Shinn-Cunningham, 2005; Marrone et al., 2008). Stream segregation refers to the ability to sort temporally interleaved sequences of sounds into distinct perceptual streams. Cherry's early study can be regarded as an example of spatial stream segregation (SSS), in which two competing speech streams were more intelligible when presented through separate headphones than when the two streams were mixed and the combined sounds presented to one or both headphones (Cherry, 1953). More recent reports have argued that spatial cues are weaker segregation cues than are fundamental frequency or spectral envelope (reviewed by Moore and Gockel, 2002). Weak spatial effects, however, are most often found in studies of *obligatory* stream segregation in which performance of a psychophysical task requires a listener to fuse information across two or more streams that might be segregated by spatial or other cues (Boehnke and Phillips, 2005; Stainsby et al., 2011; Füllgrabe and Moore, 2012).

Robust spatial effects on stream segregation are observed in studies of *voluntary* stream segregation in which the listener must evaluate a single stream in the presence of competing sounds (e.g., Hartmann and Johnson, 1991; Ihlefeld and Shinn-Cunningham, 2008). In particular, SSS is important for tasks that require a listener to piece together the successive syllables from one talker while excluding sounds from other talkers (Shinn-Cunningham, 2005; Kidd et al., 2008; Marrone et al., 2008). That is the task of a listener in a real-world cocktail party, and it applies to many other everyday listening situations.

The purpose of this review is to synthesize observations from a series of reported experiments that isolated spatial attributes of stream segregation. We address the question: "What is going on in the brain under conditions in which competing sound sequences are heard as segregated"? We review psychophysical and physiological experiments in humans and cats, and we review a natural experiment in which sensitivity to fundamental cues for SSS was lost due to a gene mutation. The results of those studies suggest that the multiple auditory objects in a cocktail party or other complex auditory scene activate multiple distinct ensembles of neurons in a listener's auditory cortex, each ensemble synchronized to a particular auditory source.

SPATIAL STREAM SEGREGATION IN HUMANS AND IN AN ANIMAL MODEL

Psychophysical studies of spatial stream segregation have been conducted using human and feline listeners (Middlebrooks and Onsan, 2012; Javier et al., 2016). Experiments with normal-hearing human listeners are important because of the importance of SSS in solving everyday human hearing challenges. The use of an animal model has enabled parallel psychophysical and invasive physiological studies.

"Rhythmic masking release" was originally devised as a psychophysical test of stream segregation using headphone-presented dichotic cues (Sach and Bailey, 2004). Middlebrooks and Onsan (2012) adapted that task for the free field, isolating spatial contributions to stream segregation while eliminating pitch, spectral, and other putative streaming cues. Target and masker sequences were constructed of temporally interleaved sequences of noise bursts having identical long-term spectra but no temporal overlap. Humans and cats were required to discriminate between rhythms of target sequences in the presence of interleaved masker sequences presented from varying source locations. Success in discriminating the rhythms required perceptual segregation of target and masker streams.

The rhythmic patterns and the layout of stimulus sources for the human psychophysical task are shown schematically in **Figure 1**. In the depictions of the two rhythms, red and blue bars denote noise bursts from the target and masker sequences, respectively. In the illustration, the vertical offset of signal and masker sound bursts denotes a difference in horizontal source location. On each trial, the listener was required to report whether he or she heard Rhythm 1 or Rhythm 2. On trials in which signal and masker sources were separated sufficiently, the target rhythm tended to pop out from the masker, and the rhythm was clearly recognizable. In an animal version of the task, cats pressed a pedal to begin presentation of Rhythm 1. When they detected a change to Rhythm 2, they could release the pedal to receive a food reward (Javier et al., 2016).

Figure 2 shows examples of performance of an individual human for targets located at 0° and 40° (**Figures 2A,B**; Middlebrooks and Onsan, 2012) and for an individual cat for a target at 0° (**Figure 2C**; Javier et al., 2016). Performance for each masker location was given by the sensitivity index, *d'*, where values of *d'* around zero indicate random-chance performance, and values $\geq$1 were taken as above threshold (Green and Swets, 1966; MacMillan and Creelman, 2005). As expected, the human and feline listeners were unable to recognize the rhythms when the masker locations were close to the target locations of 0° (**Figures 2A,C**) or 40° (**Figure 2B**). For both species, however, performance improved markedly as the target-masker separation was increased to about 10° or greater.

Figure 3 shows the distributions of RMR thresholds of human and feline listeners from the two studies in various stimulus-passband and target-location conditions. Individual thresholds are denoted by symbols, and the boxes represent medians and 25th and 75th quartiles for each condition. The broad-band stimulus condition is represented by the left-most column of each panel. Notably, broadband SSS by human and cat listeners in the two studies was comparable in acuity. The median RMR thresholds in the broad-band condition with the target at 0° were 8.1° for the human listeners and 10.2° for the cats. The similarity in psychophysical results between humans and cats, at least in the broad-band condition, adds validity to the cat as a model for humans in invasive physiological studies. Differences between the

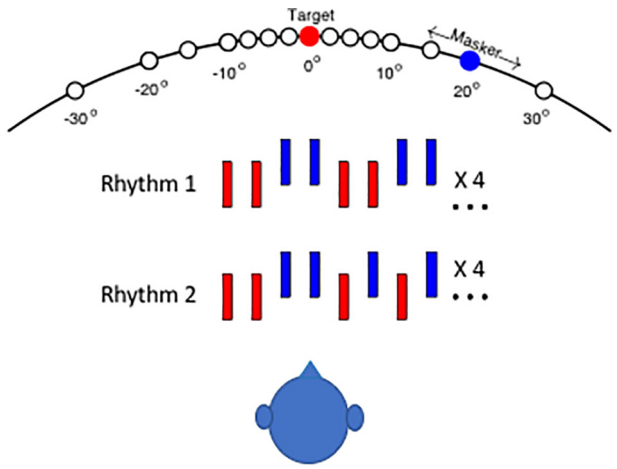

FIGURE 1 | Schematic of testing configuration and stimulus rhythms. The listener was seated in an anechoic chamber in the center of a horizontal array of small loudspeakers that were positioned at 5 or 10° intervals. The target source (denoted by red) was fixed at 0°, and the masker source location (blue) was varied parametrically in azimuth (i.e., the horizontal dimension). In the 40° target condition, the listener was turned to face 40° left, placing the target 40° to his/her right. Stimulus Rhythm 1 or 2 consisted of a sequence of noise bursts (denoted by red bars) that were interleaved with masker noise bursts (blue). The vertical offset of blue bars in the illustration indicates that target and masker sources could differ in azimuth.

species appeared when restricted stimulus passbands were tested, considered in the next section.

Performance for the human listeners was somewhat degraded when the target was displaced 40° to the side (**Figure 3B**; Middlebrooks and Onsan, 2012). The median RMR threshold in the broad-band condition increased from 8.1° for the straight-ahead, 0°, target to 11.2° for the target at 40°. That the threshold separations were wider for the lateral target is not surprising, given that the spatial rate of change of interaural difference cues tends to decline with increasing distance from the midline (Shaw, 1974; Kuhn, 1977). What is notable is that performance was not very much worse. A popular model of spatial representation in the auditory cortex has it that the location of a given stimulus is represented by the balance of activity between broadly tuned "opponent" neural populations tuned to the right or left half of space (Stecker et al., 2005; Phillips, 2008; Magezi and Krumbholz, 2010; Briley et al., 2013). In the measure of SSS in the 40°-target condition, however, the target and all of the masker location were restricted to the right hemifield of space, meaning that all the stimuli were primarily activating neurons in the left cortical hemisphere. That raises the possibility that listeners performed SSS primarily on the basis of computations within one cortical hemisphere, that is, with little or no inter-hemisphere comparison. That speculation is further supported by single-neuron recordings in cats (Middlebrooks and Bremen, 2013), presented in a later section.

This section has reviewed psychophysical experiments that demonstrated a robust spatial contribution to stream segregation, both in humans and cats. Relevant to the example of a cocktail party, the minimum spatial resolution of SSS reported for humans was somewhat narrower than the width of a human head at arm's length. Compared to a condition in which target and maskers were located around the frontal midline, human listeners showed only minor degradation of performance when all stimulus and masker source were restricted to one half of space. We now turn to the spatial acoustical cues that underlie SSS, which further inform notions of brain mechanism of SSS.

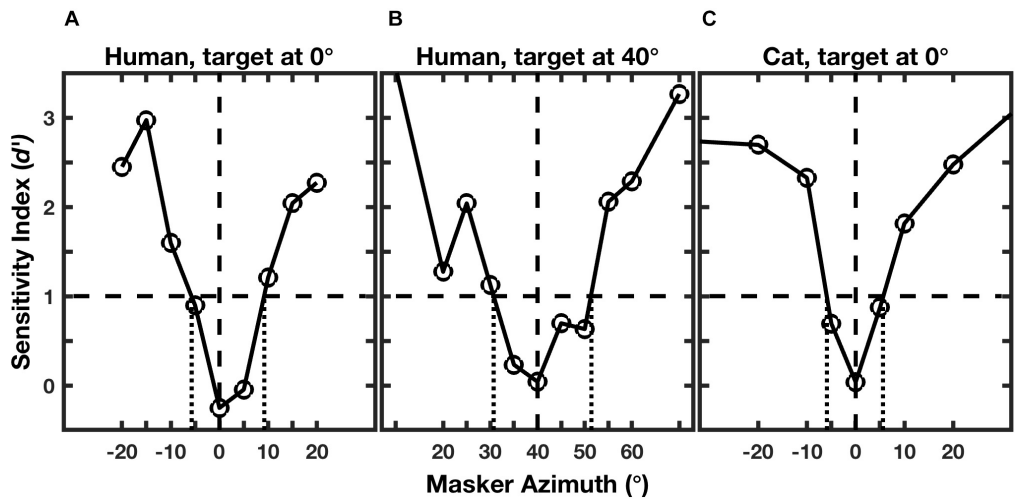

FIGURE 2 | Rhythmic masking release (RMR). The curves represent performance at rhythm identification as a function of masker location. Performance is given by the sensitivity index (*d'*), which is an unbiased measure that incorporates rates of correct and incorrect identifications (Green and Swets, 1966; MacMillan and Creelman, 2005). The horizontal dashed line indicates the threshold criterion of $d' = 1$, and the vertical dotted lines indicate the projections of threshold crossings to the masker-azimuth axis, giving the RMR thresholds. The three panels represent: **(A)** an individual human listener; **(B)** the same human listener turned to place the target 40° to the right; and **(C)** an individual feline listener with the target straight ahead. (From Middlebrooks and Onsan, 2012; and Javier et al., 2016).

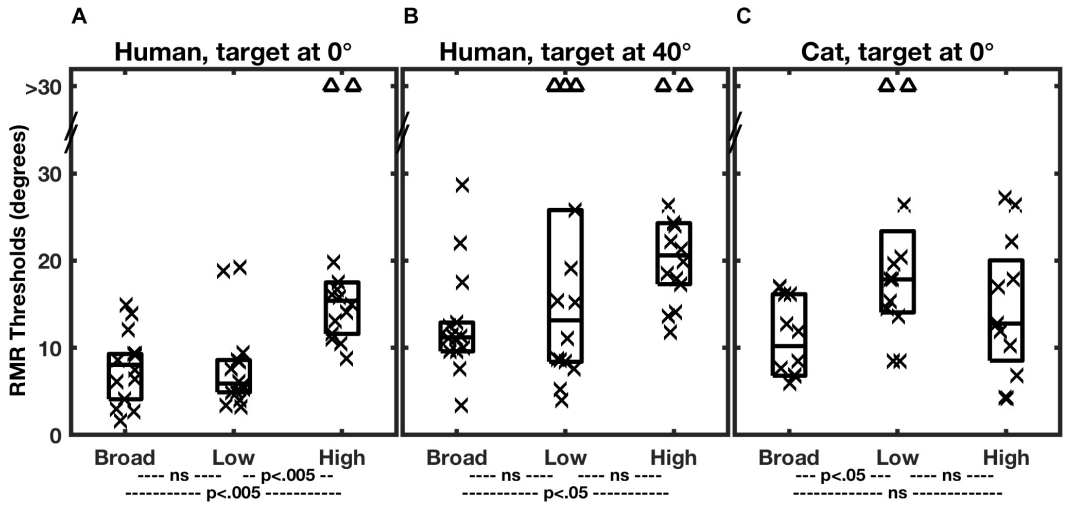

FIGURE 3 | Summary of RMR thresholds. In each panel, each column of a box and symbols represents the distribution of thresholds among 7 human listeners **(A,B)** or 6 feline listeners **(C)** in one passband condition. Thresholds for masker sources to the left and right of the target were combined, so that each listener is represented by two symbols. Symbols indicate individual thresholds, and boxes indicate 25th, 50th, and 75th percentiles. The triangles near the top of each panel denote instances in which threshold performance was not attained for a maximum target/masker separation of 30°. Broad, low, and high indicate the broadband, low-band, and high-band passband conditions described in the text. In each species/target location condition, thresholds varied significantly with passband. The statistical *p* values across the bottom of each panel indicate results of *post hoc* pair-wise comparisons of passband conditions with Bonferroni correction. (From Middlebrooks and Onsan, 2012; and Javier et al., 2016).

SPATIAL CUES FOR STREAM SEGREGATION

The locations of sound sources are not represented directly in the auditory periphery but must be inferred from spatial cues provided by the interaction of incident sounds with the acoustics of the head and external ears. The principal spatial cues in the horizontal dimension are interaural differences in the timing of cycle-by-cycle fine structure (ITD_{fs}) and interaural differences in sound pressure level (ILD), reviewed by Middlebrooks and Green (1991). Other possible spatial cues include interaural differences in the timing of sound envelopes (ITD_{env}), a monaural level sensitivity referred to as the "head-shadow effect", and spectral shape cues. The utility of various cues for spatial hearing varies with sound frequency, with ITD_{fs} cues being audible by humans only below ∼1.4 kHz (Brughera et al., 2013), and ILD cues generally increasing in magnitude with frequency increasing above 4 kHz. Identification of the spatial cues that support SSS has raised important insights into the brain mechanisms for SSS as well as providing some practical guidance for remediation of hearing impairment.

Middlebrooks and Onsan (2012) evaluated SSS performance by human listeners using stimuli that differentially favored ITD_{fs} or ILD cues; the control condition was SSS performance with a broadband stimulus, 400 to 16000 Hz in passband. Results from that study are shown in **Figures 3A,B**. The low-band stimulus, 400 to 1600 Hz, essentially eliminated ILD cues, leaving ITD_{fs} as the principal spatial cue in the horizontal dimension. In that condition, SSS performance was not significantly different from that in the control, broadband, condition. In contrast, the spatial acuity of SSS was markedly degraded in the high-band condition, which eliminated ITD_{fs} cues, leaving only ILD cues. Middlebrooks and Onsan (2012) interpreted those observations to mean that humans receive their highest-acuity spatial cues for SSS from ITD_{fs} cues.

A different result was obtained for cats by Javier and colleagues (2016; **Figure 3C**). The cats consistently showed degraded performance in the low-band condition (i.e., using ITD_{fs} cues) and control-level performance in the high-band condition, presumably using ILD cues. Those results were taken to indicate that cats receive their highest-acuity SSS cues from ILDs. Javier et al. (2016) suggested that the difference between humans and cats in use of ITD_{fs} and ILD cues could be accounted for in large part by differences in the sizes of the heads of the two species (Javier et al., 2016).

An additional interaural difference cue to consider is the interaural time difference in the envelopes of high-frequency sounds (ITD_{env}). In humans, Middlebrooks and Onsan (2012) evaluated stream segregation in high-frequency sounds (4000 to 16000 Hz) presented over headphones, manipulating ILD and ITD independently. The results of those experiments showed that high-frequency spatial stream segregation relies almost entirely on ILD cues, with only a slight synergy with ITD_{env} and only at the largest physiologically relevant ITDs, around 700 μs.

Studies of spatial release from masking have emphasized the importance of the monaural *head-shadow effect* (Bronkhorst and Plomp, 1988; Hawley et al., 2004). When a target and a masker are separated in space, shadowing by the head will result in a difference in the target-to-masker ratio at the two ears. In the RMR task, the head shadow could produce a systematic fluctuation between target and masker sound levels at each ear. Middlebrooks and Onsan (2012) tested conditions in which

the sound sequences varied randomly in sound level, thereby confounding any monaural level cues. The variable-level stimuli produced essentially no degradation in stream segregation, suggesting that the principal spatial cues are interaural difference cues (ITD_{fs} and/or ILD), which would not be confounded by the level variation.

The observation that normal-hearing listeners derive their highest-acuity SSS from ITD_{fs} cues (Middlebrooks and Onsan, 2012) is significant for the remediation of hearing impairment with hearing aids or cochlear implants. In general, hearing aids and cochlear implants do a poor job of transmitting ITD_{fs} information. Hearing aids introduce delays of as great as 10000 μs, and those delays can vary substantially across frequencies (e.g., Dillon et al., 2003). A device-imposed delay of, say, 7000 μs is about an order of magnitude larger than the maximum naturally occurring ITD_{fs}. Cochlear implant sound processors, on the other hand, transmit only the envelopes of sounds, eliminating temporal fine structure altogether. Moreover, when tested with laboratory processors, implant users show only limited sensitivity to temporal fine structure (e.g., Zeng, 2002; van Hoesel, 2007). The demonstration of the importance of ITD_{fs} cues for SSS should heighten the motivation for overcoming those failings in delivering temporal fine structure to hearing aid and cochlear implant users.

Given the results reviewed so far, one might question whether SSS should be regarded as a truly spatial phenomenon, or whether it merely reflects stream segregation on the basis of interaural differences. Middlebrooks and Onsan (2012) addressed that issue by presenting target and masker sources in the vertical midline. In that condition, interaural differences are negligible, and the principal spatial cues are spectral-shape cues provided by the elevation-specific filtering properties of the external ears (reviewed by Middlebrooks and Green, 1991). Those experiments demonstrated that spatial stream segregation is possible in elevation, i.e., in the absence of interaural difference cues. Nevertheless, they also revealed an unexpected dependence on the durations of the individual stimulus noise bursts that constituted the stimulus sequences (**Figure 4A**, right half of the panel). When the noise bursts were shortened to 10 ms in duration, the RMR task was impossible for most of the listeners, whereas that duration produced essentially no decline in horizontal resolution. When the burst duration was lengthened to 40 ms, however, stream segregation in elevation improved markedly, so that the median RMR threshold in elevation was not significantly different from that in azimuth. Those results indicate that SSS is not strictly an interaural-difference phenomenon. Nevertheless, they show that the mechanisms for deriving cues for SSS from spectral shapes appear to require greater temporal integration than do those for processing interaural cues.

The minimum audible angle (MAA) is a measure of the spatial acuity of sound-source localization. Middlebrooks and Onsan (2012) measured MAAs in the same human listeners that were tested for SSS; those MAA data are shown in **Figure 4B**. In the broadband, azimuth, condition (left-most box and symbols in **Figure 4**, panels A and B), nearly all the RMR thresholds were wider than the MAAs, although the distributions were contiguous. The most remarkable observation about the MAAs, however, is that the median values of MAAs in azimuth were largely constant across varying passbands and, in the vertical midline, were largely constant across burst durations. This contrasts with RMR thresholds (**Figure 4A**), which varied markedly across those stimulus conditions.

One might have entertained the hypothesis that static location discrimination (i.e., measured by MAA) and SSS draw spatial information from a common cortical spatial representation. That hypothesis, however, would predict that localization and SSS would show similar trends in spatial acuity across passband and burst-duration conditions. The results shown in **Figure 4** clearly refute that prediction. Based on those human psychophysical results, Middlebrooks and Onsan (2012) raised the possibility that SSS is derived from different cortical mechanisms than those that underlie sound-source localization. Location discrimination and SSS almost certainly rely on common mechanisms for low-level analysis of ITD_{fs}, ILD, and spectral shape. At more central levels, however, SSS appears to derive highest horizontal acuity from ITD_{fs} cues and to require greater temporal integration for use of spectral-shape cues for the vertical dimension.

Several additional lines of evidence support the view that the mechanisms that underlie SSS (or spatial release from masking) are distinct from those for source localization. First, neural recordings in anesthetized cats have demonstrated largely similar spatial sensitivity among several primary auditory cortical areas (Harrington et al., 2008). Nevertheless, reversible inactivation of a subset of those areas disrupts performance of a localization task (Malhotra et al., 2004), whereas inactivation of another area disrupts performance of a rhythm-discrimination task while preserving localization (Lomber and Malhotra, 2008). Second, a speech study demonstrated essentially equivalent spatial unmasking of speech by ITD and ILD cues across conditions that produced markedly different spatial percepts (Edmonds and Culling, 2005). Finally, a population of patients having a variety of cortical lesions displayed a dissociation between those who showed deficits in a lateralization task and others who showed impaired spatial release from masking (Duffour-Nikolov et al., 2012).

SPATIAL STREAM SEGREGATION IN THE ASCENDING AUDITORY PATHWAY

We now return to the question: "What is going on in the brain under stimulus conditions in which a listener could segregate interleaved sound sequences"? We consider two contrasting hypotheses. One is that the spatial relations of sound sources in the auditory scene are faithfully transmitted to early stages of the auditory cortex and that "higher" cortical mechanisms in some way segregate sounds based on that low-level cortical representation. The other view is that the job of spatial stream segregation is carried out by the auditory brainstem and that segregated streams are represented in the auditory cortex as distinct populations of activated neurons.

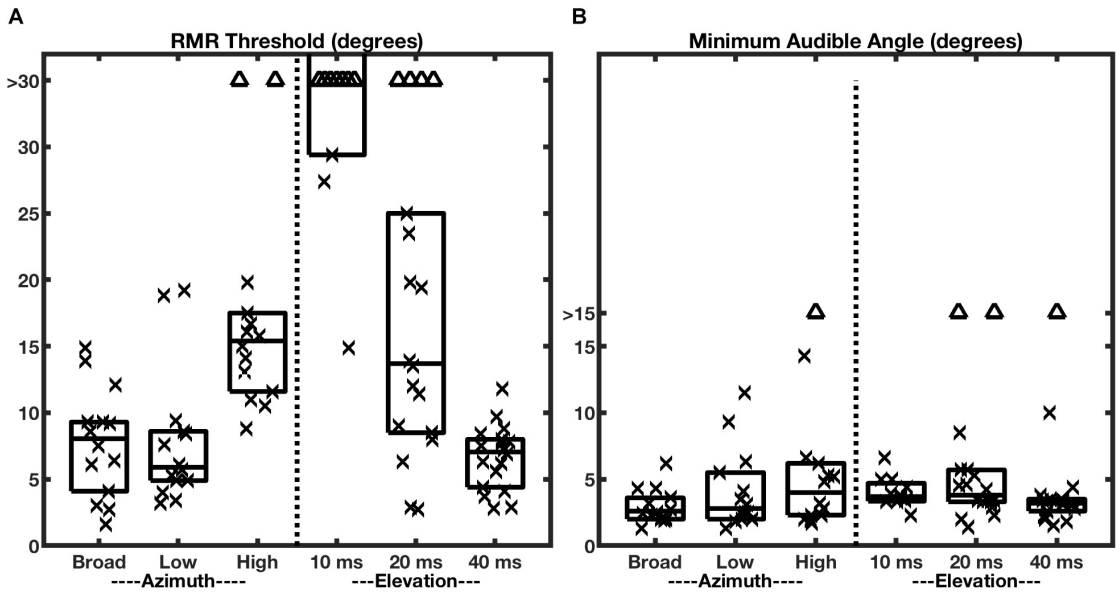

FIGURE 4 | RMR thresholds vary more across bandwidth and burst durations more than do MAAs. The two panels represent RMR thresholds **(A)** and MAAs **(B)** for human listeners. Each column of a box and symbols represents one passband condition in the horizontal (azimuth) dimension (burst duration 20 ms) or one burst-duration condition in the vertical midline (elevation); box and symbol conventions are as in **Figure 3**. The left halves of the panels show data for various passbands tested in the horizontal plane with the target at 0° azimuth. The right halves show data for various broadband burst durations tested in the vertical midline with the target at 0° elevation. (From Middlebrooks and Onsan, 2012).

Middlebrooks and Bremen (2013) tested those hypotheses by recording from single neurons in the primary auditory cortex (area A1) of anesthetized cats. The rationale was that higher-level cortical mechanisms are largely suppressed under anesthesia. For that reason, the first hypothesis, which demands higher-order cortical processing, would predict little or no spatial stream segregation in the anesthetized cortex. Conversely, the second hypothesis, which calls for stream segregation in the auditory brainstem, would predict that spatial stream segregation would be evident in the cortex under anesthesia.

Stimuli in the Middlebrooks and Bremen study consisted of trains of broad-band noise bursts presented from target and masker sources located in the horizontal plane, much as in the cat psychophysical experiments (Javier et al., 2016). Extracellular spikes recorded from cortical neurons tended to synchronize closely with the stimulus noise bursts. **Figure 5** shows post-stimulus-time histograms representing the responses of one well-isolated single neuron (Middlebrooks and Bremen, 2013). The left panels (**Figures 5A,C,E**) show the responses to sounds from a single source located straight ahead (0°) or at 40·contralateral or ipsilateral to the midline. Spike times were largely restricted to the 50-ms-wide time bins following the onsets of noise bursts. The spike rates of this neuron elicited by a train of noise burst from a single source showed essentially no sensitivity to the locations of sources across the 80° range shown in the illustration, as indicated by the similar heights of bars in panels 5A, C, and E.

The spatial sensitivity of neurons was substantially increased in the presence of a competing sound. The right panels in **Figure 5** show responses synchronized to a target fixed at 0° (denoted by red bars) and a masker (blue bars) presented from contralateral 40°, 0°, or ipsilateral 40°. In each condition, there was a robust response to the first noise burst in the sequence (at 0 ms), but the response to the second noise burst, at 200 ms, was weak or entirely suppressed. In the condition shown in **Figure 5D**, the target and masker were co-located at 0°; this is an identical condition to that shown in **Figure 5C** except that the rate of presentation of the noise bursts was doubled. At this higher presentation rate, the response to each burst was less than half of that at the slower rate, and the precision of synchrony was somewhat degraded. When the masker source was moved to ipsilateral 40°, however, there was a striking recovery of the response to the target and nearly complete suppression of the response to the masker (**Figure 5F**). Conversely, when the masker source was moved to contralateral 40°, the neural response was largely captured by the masker, with corresponding suppression of the response to the target (**Figure 5B**). Middlebrooks and Bremen (2013) took this pattern of responses as evidence for SSS in the responses of a single cortical neuron.

The responses of the neuron in **Figure 5** are shown in finer spatial detail in **Figure 6**; the three panels on the left of **Figure 6** show stimulus-synchronized spike counts measured for three target locations, with the target location for each panel denoted by the vertical dashed line. The blue lines represent counts of spikes that were synchronized to the masker as a function of masker location. The red lines represent counts synchronized to the fixed-location target indicated by the vertical red dashed line; those responses also varied as a function of masker location. The black lines, duplicated in

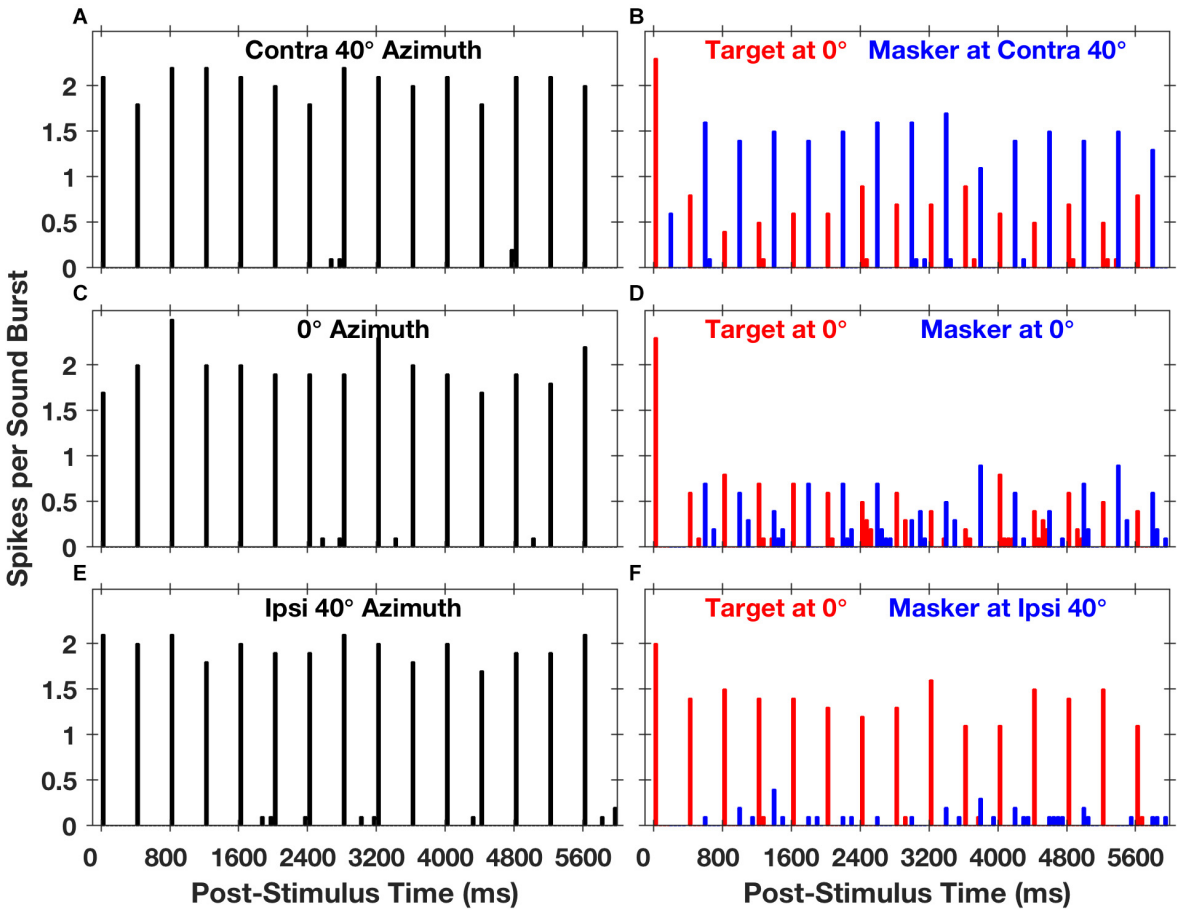

FIGURE 5 | Spike counts from an isolated single neuron synchronized to target and masker sound sequences. The plots are post-stimulus-time histograms of spikes synchronized to sequences of 5-ms sound bursts from a single source **(A,C,E)** at a rate of 2.5 bursts per second or to interleaved sequences from two sources **(B,D,F)** at an aggregate rate of 5 bursts per second. The histogram bars represent mean spike counts in 50-ms time bins, averaged over 20 repetitions. In the left column of panels, the sound source was located in the horizontal plane at contralateral 40° azimuth **(A)**, 0° azimuth **(C)**, or ipsilateral 40° azimuth **(E)**; contralateral and ipsilateral are with respect to the recording site in the right cortical hemisphere. In the right column of panels, red or blue bars represent spikes synchronized to the target or the masker, respectively. The target source was fixed in location at 0°, and the masker source was located at contralateral 40° **(B)**, 0° **(D)**, or ipsilateral 40° **(F)**. The condition in panel **(D)**, in which target and masker were co-located at 0°, is identical to the condition in panel **(C)** except that the sound-burst rate is doubled. Unit 1204.3.10. (From Middlebrooks and Bremen, 2013).

each of the panels, represent spike counts synchronized to a single source. When the target and masker sources were co-localized (i.e., when the blue line crossed the vertical dashed line), the target and masker spike counts were essentially identical, and both were strongly suppressed compared to the response to the single source; this is the condition shown in **Figure 5D**. Target and masker spike counts diverged markedly as the masker source was shifted away from the target source. In conditions of wide target/masker separation, the response synchronized to the target or masker could be equal in magnitude to the response to the single source. This unit was representative of the majority in the study in that the more contralateral sound source elicited a stronger response than did the more ipsilateral source; there was, however, a sizeable minority of units that favored the more ipsilateral source. Middlebrooks and Bremen (2013) showed that neurons exhibiting a similar preference for contralateral or ipsilateral sound sources tended to form preference-specific ensembles within the cortex.

The right column of panels in **Figure 6** shows the sensitivity with which the sounds synchronized to the target and masker could be segregated significantly on the basis of spike counts. In the illustrated example, supra-threshold sensitivity (i.e., $d' > 1$ or < -1) was observed in 5 of the 6 conditions of the masker at the minimum tested separation to the left or right of the target. This unit was representative of the finding that, in most cases, target/masker discrimination was more acute when the target source was located on the midline or in the ipsilateral half of space compared to when the target source was contralateral to the cortical recording site.

The unit in **Figures 5** and **6** was representative of essentially all those in the Middlebrooks and Bremen study in that its spatial sensitivity increased markedly when the target was presented with a competing sound source. In **Figure 6**, for example, the

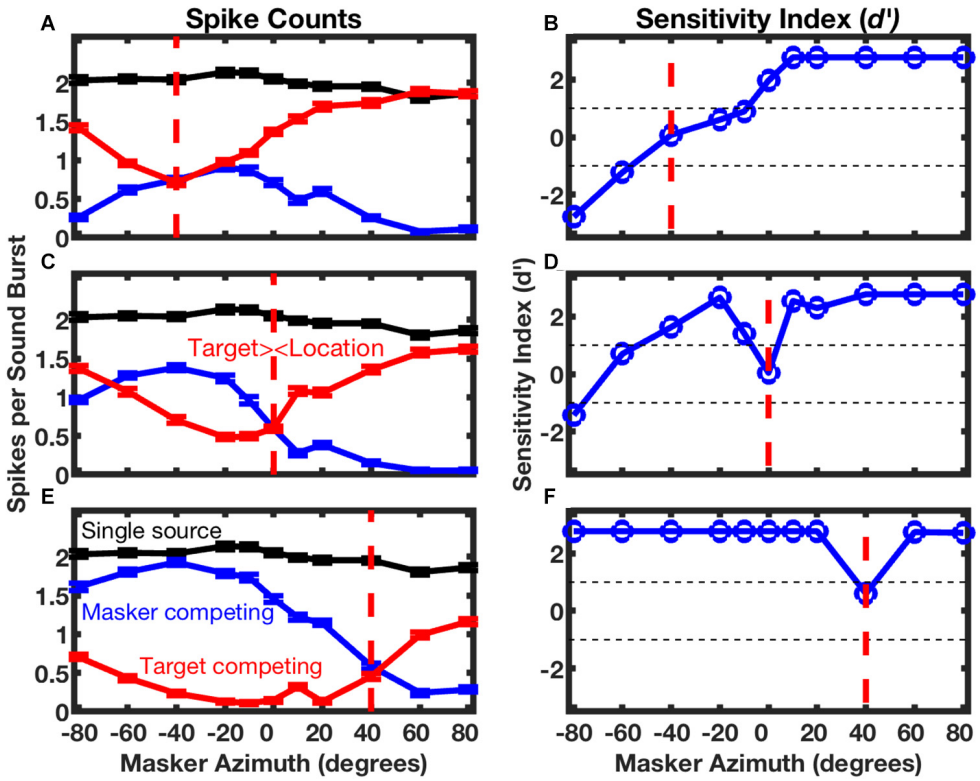

FIGURE 6 | Spatial stream segregation in the responses of a single neuron. Responses of the neuron represented in **Figure 5** are plotted here as a function of masker azimuth. The location of the target source was fixed at contralateral 40° **(A,B)**, 0° **(C,D)**, or ipsilateral 40° **(E,F)**, as indicated in each panel by a vertical dashed line. The left column of panels shows mean spike counts per sound burst synchronized to a single source (black, duplicated in each panel), or to competing target (red) or masker (blue) sources. Error bars are standard errors of the mean. The right column of panels plots the sensitivity index (*d'*) for discrimination of trial-by-trial mean spike rates synchronized to the target or masker. Positive values of *d'* denote cases in which there were more spikes synchronized to the more contralateral source. Unit 1204.3.10. (From Middlebrooks and Bremen, 2013).

blue and red lines, which represent conditions of competing target and masker sources, demonstrate substantially greater modulation by source location than does the black line, which represents the single-source condition. Across the sampled population, the breadth of tuning in azimuth narrowed by about 1/3 and the depth of modulation by changes in the masker location nearly doubled in the presence of a competing sound (Middlebrooks and Bremen, 2013).

Spatial stream segregation by neurons in the cat's primary auditory cortex tended to replicate the result that feline psychophysical performance is more acute with high- compared to low-frequency sounds (Javier et al., 2016). Middlebrooks and Bremen (2013) computed a metric of the strength of SSS. That metric varied significantly with the frequency tuning of neurons, indicating that SSS tended to be more robust among neurons tuned to frequencies in the upper half of the range sampled in the cat.

Middlebrooks and Bremen (2013) found that the spike counts synchronized to the target or masker in competing conditions could be modeled well by a linear expression that incorporated the spatial tuning to a single source and the magnitude of the forward suppression (or "attenuation") that could be measured in the co-localized condition. Forward suppression is a mechanism at one or more levels of the auditory pathway by which the response to one sound suppresses the response to a following sound. Middlebrooks and Bremen confirmed empirically that the forward suppression that they observed was not due to the simple habituation of responses of neurons in the auditory cortex. That observation suggested that forward suppression observed in the cortex is inherited from a sub-cortical level.

The conclusion of a sub-cortical origin of forward suppression is supported by measures of SSS and forward suppression at multiple levels of the rat ascending auditory pathway (Yao et al., 2015). In that study, SSS and forward suppression were essentially absent in the inferior colliculus at stimulus rates at which human and feline psychophysical listeners exhibit spatial stream segregation. Stream segregation and forward suppression first emerged at the level of the nucleus of the brachium of the inferior colliculus. Those phenomena also were robust in about 2/3 of neurons sampled in the ventral nucleus of the medial geniculate and were ubiquitous in the primary auditory cortex. The SSS strengthened at successive levels of the ascending auditory pathway, both due to increasing spatial sensitivity of neurons and increasing forward suppression. Tests of GABA inhibitors applied to the cortical surface

demonstrated that forward suppression is not due to synaptic inhibition at the level of the cortex. Instead, Yao and colleagues favored the view that forward suppression underlying stream segregation is most likely due to synaptic depression in the thalamocortical projection.

RHYTHMIC MASKING RELEASE IN THE AUDITORY CORTEX

The rhythmic masking release task that was employed in psychophysical experiments in humans (Middlebrooks and Onsan, 2012) and cats (Javier et al., 2016) demonstrated that human and feline listeners could discriminate rhythmic patterns when the target and masker sources were separated by around 10°. That is roughly the spatial acuity with which cortical neurons in the anesthetized cat auditory cortex could segregate streams of noise bursts from alternating source locations, according to the results from Middlebrooks and Bremen (2013). The latter authors extended that observation by testing the target-masker separation at which target rhythm could be identified on the basis of firing patterns of single cortical neurons.

In those empirical tests, stimuli consisted of sequences of broad-band noise bursts presented as Rhythm 1 or Rhythm 2, which were essentially equivalent to the broad-band condition in the human psychophysical experiments (Middlebrooks and Onsan, 2012). The target source was fixed at 0°, and the masker source was varied in azimuth. Neurons synchronized strongly to target or masker components of competing sounds. **Figure 7** shows post-stimulus-time histograms from a single well-isolated neuron in response to Rhythm 1 (top row of panels) or Rhythm 2 (bottom row) in three target/masker configurations (columns). The pattern of short bars across the top of each panel represents the stimulus rhythm, consisting in each case of four noise bursts from the target (red) and four from the masker (blue).

The response of that neuron was almost entirely suppressed when the target and masker were co-located (**Figures 7B,E**). Robust responses synchronized to the target or masker emerged when the masker was shifted to one or the other side of the target source. When the masker source was at contralateral 40° (**Figures 7A,D**), the neuron responded strongly only to temporally isolated masker bursts. That is, there were strong responses to a masker burst that followed a target burst, but no response to the second of two successive masker bursts. In contrast, when the masker source was at ipsilateral 40°, the response of the neuron was captured by the target sound bursts. In that condition, the response was restricted to target bursts that followed spatially distinct masker bursts, and there was no response to the second of two target bursts.

The identities of the two rhythms are evident by casual inspection of the histograms in **Figure 7**: there are strong responses at two post-stimulus times in response to Rhythm 1 and at three post-stimulus times in response to Rhythm 2. Middlebrooks and Bremen (2013) used multiple linear regression to evaluate the spike counts in each of 8 time bins (the regressor), solving for the appropriate rhythm, 1 or 2. **Figure 8A** shows the performance of a single unit in discriminating between stimulus Rhythms 1 and 2; the target was fixed at 0°, and the masker was varied in azimuth. When target and masker were co-located, performance was around chance level. When the masker was shifted to either side, however, performance rapidly improved. **Figure 8B** shows the distribution of *d'* values for the population of 57 well-isolated units that were tested in the Middlebrooks and Bremen (2013) study; the solid line plots the median, and the dashed lines show the 25th and 75th quartiles. Using a criterion of $d' = 1$, about 25% of neurons segregated streams from target and masker sources separated by as little as about 10°. That acuity of single cortical neurons is remarkably close to the psychophysical thresholds of feline (and human) listeners.

A BREAKDOWN IN SPATIAL HEARING

The auditory brainstem is well adapted for the fine temporal and intensive processing that is needed for use of interaural difference cues for spatial hearing. These adaptations include the end-bulbs of Held that terminate on the bushy cells of the anterior ventral cochlear nucleus (AVCN), the calyceal endings of Held in the medial nucleus of the trapezoid body (MNTB), and the specialized binaural nuclei of the superior olivary complex, the medial superior olive (MSO) and the lateral superior olive (LSO). All of those structures exhibit a high expression of high-threshold voltage-dependent potassium channels, specifically Kv3.1 and Kv3.3 (Grigg et al., 2000; Li et al., 2001; Chang et al., 2007); Kv3.3 subunits are highly expressed in the AVCN, MNTB, MSO, LSO, and central nucleus of the inferior colliculus (ICc), whereas Kv3.1 is largely restricted to the AVCN, MNTB, and ICc, with relatively little expression in the MSO and LSO (Li et al., 2001). The Kv3.1 and Kv3.3 channels permit rapid repolarization of action potentials, thereby supporting high spike rates and high temporal precision. In the mouse MNTB, either Kv3.1 or Kv3.3 subunits supported rapid repolarization, whereas Kv3.3 was essential for repolarization in the LSO (Choudhury et al., 2020).

Middlebrooks et al. (2013) took advantage of a natural experiment by testing psychophysical performance in human listeners who lack normal function of Kv3.3 channels. Autosomal dominant mutations in the gene encoding Kv3.3 have been identified in two kindreds, one in France (Herman-Bert et al., 2000) and one in the Philippines (Waters et al., 2005; Subramony et al., 2013). Both kindreds exhibit spinocerebellar ataxia 13 (SCA13), although the kindreds differ in channel properties. Study of the mutation in the Filipino kindred, *KCNC3*^{R420H}, in frog oocytes has demonstrated dominant negative suppression of potassium conductance (Waters et al., 2006). Middlebrooks et al. (2013) tested the hypothesis that disruption of normal Kv3.3 channel activity would also disrupt sensitivity to interaural difference cues.

Those authors tested 13 affected individuals in the Filipino family as well as control groups consisting of 6 unaffected

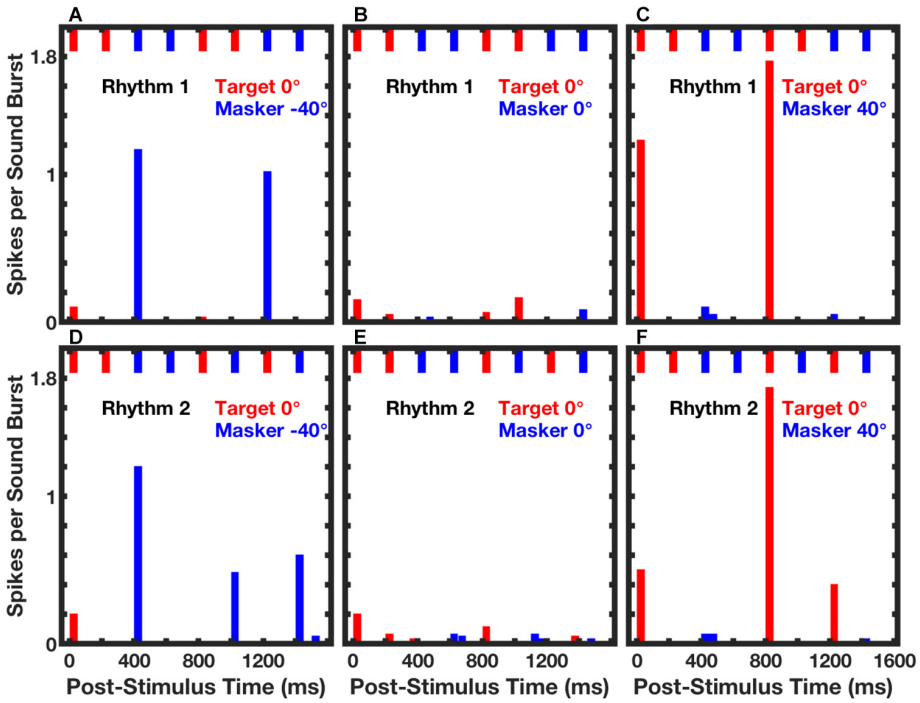

FIGURE 7 | Spike counts from an isolated single neuron synchronized to RMR stimuli. Post-stimulus-time histogram bars show mean spike counts synchronized to noise bursts from the target (red) or masker (blue) source. Data were averaged over 3 continuous repetitions of each rhythm in each of 10 trials. The upper and lower rows of panels represent responses to Rhythm 1 (top) and Rhythm 2 (bottom). The stimulus rhythm is represented by the row of short bars across the top of each panel. Across all panels, the target source was fixed at 0°. The masker source was located at contralateral 40° **(A,D)**, 0° **(B,E)**, or ipsilateral 40° **(C,F)**. Unit 1204.3.11. (From Middlebrooks and Bremen, 2013).

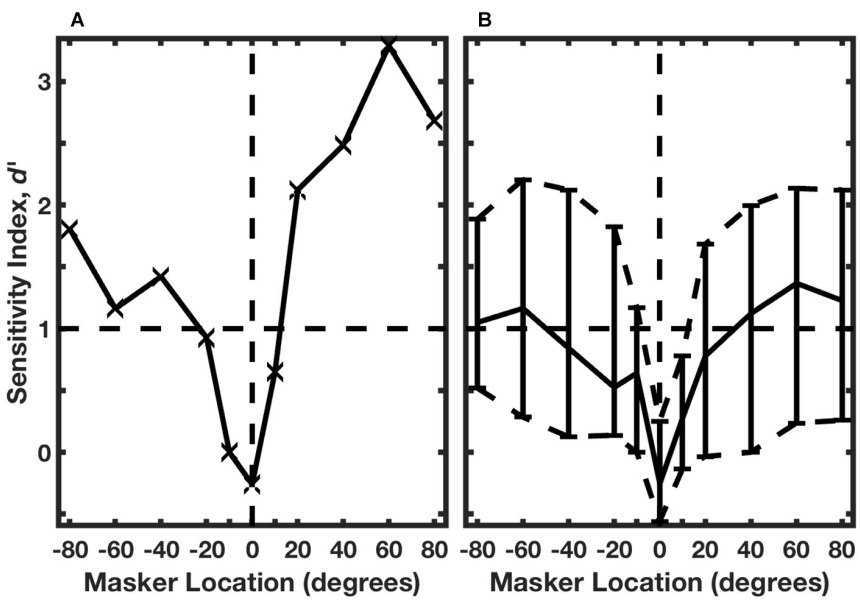

FIGURE 8 | Neural classification of rhythms. **(A)** An isolated single neuron. The target was fixed at 0°, and the masker source was varied parametrically. A regression procedure was used to identify Rhythm 1 or 2 based on the temporal patterns of neural spike counts. Performance is given by *d'*, based on trial-by-trial distributions of spike patterns across 10 trials. The horizontal dashed line indicates the RMR threshold criterion of $d' = 1$. Data are from the same unit represented in **Figure 7**. **(B)** Distribution of performance across 57 isolated single neurons. The solid curve shows the median, and dashed curves show the 25th and 75th percentiles. (From Middlebrooks and Bremen, 2013).

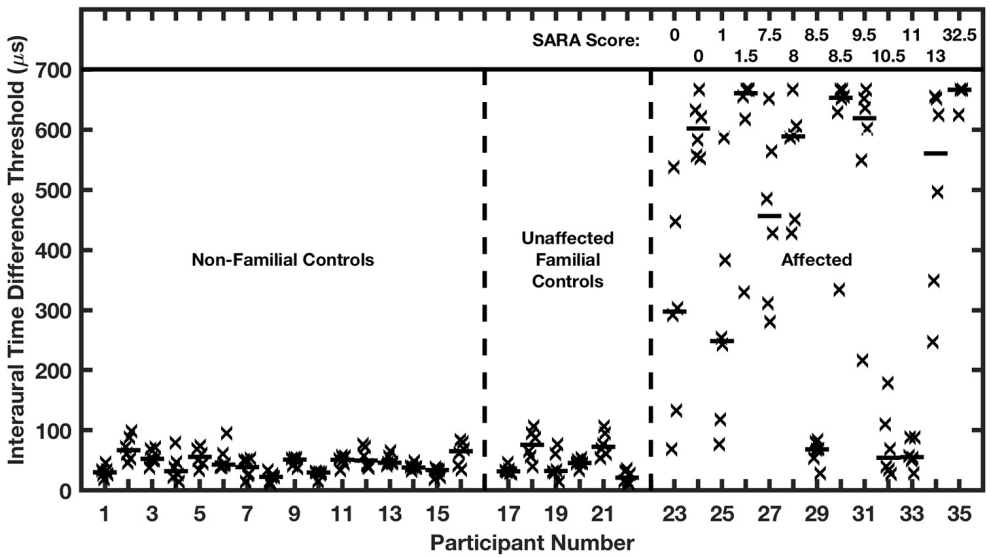

FIGURE 9 | Interaural time difference (ITD) thresholds are elevated in humans affected with a dominant negative mutation in potassium channel Kv3.3. Participants were assigned to 3 groups on the basis of the molecular testing of the *Kv3.3* gene: (1) age-matched non-familial controls; (2) unaffected familial controls; and (3) the affected group. The affected individuals are ranked left to right according to their SARA scores, shown at the top of the figure; SARA is an assessment of ataxia described in the text. Each vertical column of symbols represents 6 threshold measurements (X's) and a median (horizontal bar) for one participant. Median ITDs in the two control groups were not significantly different from each other and were comparable to published reports of ITDs of untrained listeners. The median ITDs in the affected group were significantly greater than those in the control groups. (From Middlebrooks et al., 2013).

family members and 16 non-related normal-hearing age-matched individuals. All of the affected participants were shown by molecular testing to be heterozygous for the mutated Kv3.3 gene. The family members were all evaluated for clinical signs of cerebellar ataxia. The clinical status was summarized by the Scale for the Assessment and Rating of Ataxia (SARA). Among the participants carrying the mutated gene, SARA scores ranged from 0 (asymptomatic) to 32.5 (severe disability).

The affected family members showed largely age-appropriate left- and right-ear pure-tone audiograms. None of the family members reported hearing disabilities or hearing-aid use. Dichotic (i.e., binaural) hearing tests utilized low- or high-passed stimuli that were designed to target, respectively, ITD_{fs} and ILD sensitivity and the corresponding pathways. On each trial, listeners heard two sounds and reported whether the second sound was to the left or the right of the first.

Nearly all the affected family members exhibited marked elevations of ITD and ILD thresholds. In the case of ITD (**Figure 9**), control groups showed median thresholds around 45 μs, which is comparable to published thresholds of untrained normal-hearing listeners (Wright and Fitzgerald, 2001). Conversely, ten of the 13 affected participants had ITD thresholds significantly higher than the thresholds of any of those in the control groups, mostly higher than 500 μs, which is near the maximum value produced by free-field sounds. The remaining 3 affected participants had median ITD thresholds of 68, 55, and 56 μs, which are within the distribution of control medians. Remarkably, there was no systematic correlation in the affected group between ITD thresholds and ataxia, as represented by SARA scores. Thresholds higher than 500 μs were exhibited by participants having the lowest (i.e., best: 0) or highest (32.5) SARA scores in the sample, and participants having SARA scores higher than 8 had ITD thresholds ranging from <100 to >650 μs. It is worth noting that SARA scores are based on fairly rudimentary motor exams, such that a score of 0 sometimes will be assigned in a case in which later, more precise, measures might reveal a gait disturbance or other signs of ataxia.

Thresholds for ILD detection were similarly elevated (**Figure 10**). All but 2 of the affected individuals had ILD thresholds that were 5 dB or greater, in contrast with the control groups having median values that all were 5 dB or less, averaging 2.5 dB. Again, there was no correlation in the affected group between ILD threshold and ataxia. Within the affected group, ITD thresholds of affected individuals correlated highly with their ILD thresholds. The high correlation between deficits in ITD and ILD sensitivity, and the absence of correlation with the severity of ataxia, suggests that expression of the mutant allele and selection of channel subtypes might differ between auditory and cerebellar pathways. Moreover, the presence of functional Kv3.3 subunits within voltage-gated potassium channels might be more or less essential for rapid repolarization in various structures, as has been demonstrated in the mouse LSO and MNTB (Choudhury et al., 2020).

It was not feasible for Middlebrooks et al. (2013) to test SSS in the affected listeners. Nevertheless, one could speculate that the deficits in ITD_{fs} and ILD sensitivity would severely

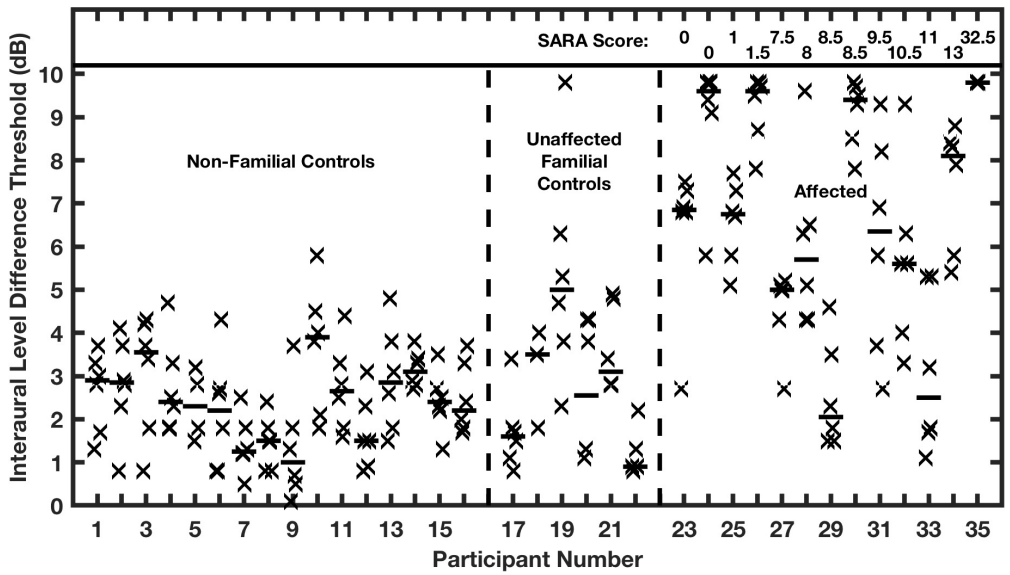

FIGURE 10 | Interaural level difference (ILD) thresholds are elevated in humans affected with a dominant negative mutation in potassium channel Kv3.3. ILD thresholds are shown for two control groups and the affected group. Other conventions are as in **Figure 9**. (From Middlebrooks et al., 2013).

impair SSS, leading to great difficulty in parsing conversations in the presence of competing sounds. In principle, spectral-shape cues could replace binaural cues to provide spatial information in the horizontal dimension. Tests of horizontal sound-source localization in the absence of binaural cues, however, have yielded rather mixed results (Belendiuk and Butler, 1977; Slattery and Middlebrooks, 1994; Wightman and Kistler, 1997; Macpherson and Middlebrooks, 2002) and we are aware of no tests of SSS in the horizontal dimension have evaluated the contribution of spectral-shape cues. The tests of SSS by Middlebrooks and Onsan (2012) in the horizontal plane and in the vertical midline suggest that the most robust, highest-acuity, SSS relies on binaural cues, specifically ITD_{fs} and ILD.

SUMMARY AND CONCLUSION

How does a listener piece string together the syllables from one talker amid the tangle of competing voices at a cocktail party or any other complex auditory scene? What are the brain mechanisms that enable such a task? In this review, we have focused on a series of experiments that were unified by use of a common psychophysical task, "rhythmic masking release", and its corresponding stimulus set. Here, we summarize some of the key conclusions from those previous studies.

In psychophysical tests, listeners successfully segregated interleaved sound sequences that differed only in the locations of target and masker. This suggests that spatial hearing would be highly beneficial in isolating a single talker amid other competing sounds. Among potential acoustical spatial cues, the best psychophysical SSS performance was provided by interaural difference (binaural) cues, particularly ITD_{fs} in humans. Nevertheless, SSS was possible for locations in the vertical midline, where interaural cues are negligible. This indicates that SSS is not strictly a binaural phenomenon.

Elementary acoustical cues for spatial hearing are analyzed in specialized nuclei of the auditory brainstem. The high-voltage voltage-gated potassium channel, Kv3.3, is particularly important for brainstem processing of ITD_{fs} and ILD. In a human kindred bearing a dominant negative mutation in the gene for the Kv3.3 channel, affected individuals showed a lack of sensitivity for ITD_{fs} and ILD, which almost certainly would severely impair their use of spatial hearing in everyday complex listening situations. Physiological studies in animal models demonstrate that SSS is derived from spatial and forward-suppression mechanism in the auditory brainstem, emerging in full force in the thalamo-cortical projection. Single neurons in the primary auditory cortex of the cat exhibit SSS with spatial acuity comparable to psychophysical listeners. The observation that SSS is observed in an early cortical level in the presence of anesthesia, i.e., in the absence of higher-level cortical processes, further supports the view that brainstem and thalamocortical mechanisms have already done the work of sorting interleaved sequences of sounds into activity in multiple distinct populations of cortical neurons.

In the cat auditory cortex, neurons that synchronize preferentially to the leftmost of a pair of sound sources tend to cluster apart from those that synchronize to the rightmost source. To the degree that the cat results can be generalized to humans, the single-neuron results provide a picture of what might be going on in our cocktail-party listener's brain when he or she attempts to focus on a speech stream from a particular talker. We speculate that the speech stream of interest would activate one or more

ensembles of mutually synchronized neurons that would be distinct from ensembles synchronized to other speech streams, background music, clinking glasses, etc. The listener, then, could use higher-level auditory or pre-frontal mechanisms to shine a light on the neural ensemble(s) representing the talker of interest. One hopes that this view of active cortical mechanisms can be tested, with or without a cocktail in hand, in future studies in behaving animals.

AUTHOR CONTRIBUTIONS

JM wrote the manuscript with comments and suggestions from MW. Both authors contributed to the article and approved the submitted version.

FUNDING

JM's work was supported by the NIDCD.

ACKNOWLEDGMENTS

We are pleased to thank our colleagues in this work, particularly Ewan Macpherson who introduced us to the RMR task, Zekiye Onsan and Beth McGuire for collecting human (Zekiye) and feline (Beth) psychophysical data, and collaborators Lauren Javier, Peter Bremen, and Justin Yao. We also thank Tetsuo Ashizawa, S. H. Subramony, Elizabeth Brooks, Joy Durana, and Praneetha Muthumani for assistance in collecting auditory data from the Filipino family.

REFERENCES

Belendiuk, K., and Butler, R. A. (1977). Spectral cues which influence monaural localization in the horizontal plane. *Percept. Psychophys.* 22, 353–358. doi: 10.3758/bf03199700

Boehnke, S. E., and Phillips, D. P. (2005). The relation between auditory temporal interval processing and sequential stream segregation examined with stimulus laterality differences. *Percept. Psychophys.* 67, 1088–1101. doi: 10.3758/bf03193634

Briley, P. M., Kitterick, P. T., and Summerfield, A. Q. (2013). Evidence for opponent process analysis of sound source location in humans. *J. Assoc. Res. Otolaryngol.* 14, 83–101. doi: 10.1007/s10162-012-0356-x

Bronkhorst, A. W., and Plomp, R. (1988). The effect of head-induced interaural time and level differences on speech intelligibility in noise. *J. Acoust. Soc. Am.* 93, 1508–1516. doi: 10.1121/1.395906

Brughera, A., Dunai, L., and Hartmann, W. M. (2013). Human interaural time difference thresholds for sine tones: the high-frequency limit. *J. Acoust. Soc. Am.* 133, 2839–2855. doi: 10.1121/1.4795778

Chang, S. Y., Zagha, E., Kwon, E. S., Ozaita, A., Bobik, M., Marone, M. E., et al. (2007). Distribution of Kv3.3 potassium channel subunits in distinct neuronal populations of mouse brain. *J. Comp. Neurol.* 502, 953–972. doi: 10.1002/cne.21353

Cherry, C. E. (1953). Some experiments on the recognition of speech, with one and two ears. *J. Acoust. Soc. Am.* 25, 975–979. doi: 10.1121/1.1907229

Choudhury, N., Linley, D., Richardson, A., Anderson, M., Robinson, S. W., Marra, V., et al. (2020). Kv3.1 and Kv3.3 subunits differentially contribute to Kv3 channels and action potential repolarization in principal neurons of the auditory brainstem. *J. Physiol.* 598, 2199–2222. doi: 10.1113/jp279668

Dillon, H., Keidser, G., O'Brien, A., and Silberstein, H. (2003). Sound quality comparisons of advanced hearing aids. *Hear. J.* 56, 30–40. doi: 10.1097/01.hj.0000293908.50552.34

Duffour-Nikolov, C., Tardif, E., Maeder, P., Bellmann Thiran, A., Block, J., Frischknecht, R., et al. (2012). Auditory spatial deficits following hemispheric lesions: dissociation of explicit and implicit processing. *Neuropsychol. Rehabil.* 22, 674–696. doi: 10.1080/09602011.2012.686818

Edmonds, B. A., and Culling, J. F. (2005). The spatial unmasking of speech: evidence for within-channel processing of interaural time delay. *J. Acoust. Soc. Am.* 117, 3069–3078. doi: 10.1121/1.1880752

Füllgrabe, C., and Moore, B. C. J. (2012). Objective and subjective measures of pure-tone stream segregation based on interaural time differences. *Hear. Res.* 291, 24–33. doi: 10.1016/j.heares.2012.06.006

Green, D. M., and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. New York, NY: Wiley.

Grigg, J. J., Brew, H. M., and Tempel, B. L. (2000). Differential expression of voltage-gated potassium channel genes in auditory nuclei of the mouse brainstem. *Hear. Res.* 140, 77–90. doi: 10.1016/s0378-5955(99)00187-2

Harrington, I. A., Stecker, G. C., Macpherson, E. A., and Middlebrooks, J. C. (2008). Spatial sensitivity of neurons in the anterior, posterior, and primary fields of cat auditory cortex. *Hear. Res.* 240, 22–41. doi: 10.1016/j.heares.2008.02.004

Hartmann, W. M., and Johnson, D. (1991). Stream segregation and peripheral channeling. *Music Percept.* 9, 155–184. doi: 10.2307/40285527

Hawley, M. L., Litovsky, R. Y., and Culling, J. F. (2004). The benefit of binaural hearing in a cocktail party: effect of location and type of interferer. *J. Acoust. Soc. Am.* 115, 833–843. doi: 10.1121/1.1639908

Herman-Bert, A., Stevanin, G., Netter, J. C., Rascol, O., Brassat, D., Calvas, P., et al. (2000). Mapping of spinocerebellar ataxia 13 to chromosome 19q13.3-*q*13.4 in a family with autosomal dominant cerebellar ataxia and mental retardation. *Am. J. Hum. Genet.* 67, 229–235. doi: 10.1086/302958

Hirsh, I. J. (1950). The relation between localization and intelligibility. *J. Acoust. Soc. Am.* 22, 196–200. doi: 10.1121/1.1906588

Ihlefeld, A., and Shinn-Cunningham, B. G. (2008). Spatial release from energetic and informational masking in a selective speech identification task. *J. Acoust. Soc. Am.* 123, 4369–4379. doi: 10.1121/1.2904826

Javier, L. K., McGuire, E. A., and Middlebrooks, J. C. (2016). Spatial stream segregation by cats. *J. Assoc. Res. Otolaryngol.* 17, 195–207. doi: 10.1007/s10162-016-0561-0

Kidd, G. Jr., Best, V. and Mason, C. R. (2008). Listening to every other word: examining the strength of linkage variables in forming streams of speech. *J. Acoust. Soc. Am.* 124, 3793–3802. doi: 10.1121/1.2998980

Kidd, G. Jr., Mason, C. R., Rohtla, T. L., and Deliwala, P. S. (1998). Release from masking due to spatial separation of sources in the identification of nonspeech auditory patterns. *J. Acoust. Soc. Am.* 104, 422–431. doi: 10.1121/1.423246

Kuhn, G. F. (1977). Model for the interaural time differences in the azimuthal plane. *J. Acoust. Soc. Am.* 62, 157–167. doi: 10.1121/1.381498

Li, W., Kaczmarek, L. K., and Perney, T. M. (2001). Localization of two high-threshold potassium channel subunits in the rat central auditory system. *J. Comp. Neurol.* 437, 196–218. doi: 10.1002/cne.1279

Lomber, S. G., and Malhotra, S. (2008). Double dissociation of 'what' and 'where' processing in auditory cortex. *Nat. Neurosci.* 11, 609–616. doi: 10.1038/nn.2108

MacMillan, N. A., and Creelman, C. D. (2005). *Detection Theory: A User's Guide*, 2nd Edn, Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Macpherson, E. A., and Middlebrooks, J. C. (2002). Listener weighting of cues for lateral angle: the duplex theory of sound localization revisited. *J. Acoust. Soc. Am.* 111, 2219–2236. doi: 10.1121/1.1471898

Magezi, D. A., and Krumbholz, K. (2010). Evidence for opponent-channel coding of interaural time differences in human auditory cortex. *J. Neurophysiol.* 104, 1997–2007. doi: 10.1152/jn.00424.2009

Malhotra, S., Hall, A. J., and Lomber, S. G. (2004). Cortical control of sound localization in the cat: unilateral cooling deactivation of 19 cerebral areas. *J. Neurophysiol.* 92, 1625–1643. doi: 10.1152/jn.01205.2003

Marrone, N., Mason, C. R., and Kidd, G. Jr. (2008). Tuning in the spatial dimension: evidence from a masked speech identification task. *J. Acoust. Soc. Am.* 124, 1146–1158. doi: 10.1121/1.2945710

Middlebrooks, J. C., and Bremen, P. (2013). Spatial stream segregation by auditory cortical neurons. *J. Neurosci.* 33, 10986–11001. doi: 10.1523/jneurosci.1065-13.2013

Middlebrooks, J. C., and Green, D. M. (1991). Sound localization by human listeners. *Ann. Rev. Psychol.* 42, 135–159. doi: 10.1146/annurev.ps.42.020191.001031

Middlebrooks, J. C., Nick, H. S., Subramony, S. H., Advincula, J., Rosales, R. L., Lee, L. V., et al. (2013). Mutation in the Kv3.3 voltage-gated potassium channel causing spinocerebellar ataxia 13 disrupts sound-localization mechanisms. *PLoS One* 8:e76749. doi: 10.1371/journal.pone.0076749

Middlebrooks, J. C., and Onsan, Z. A. (2012). Stream segregation with high spatial acuity. *J. Acoust. Soc. Am.* 132, 3896–3911. doi: 10.1121/1.4764879

Moore, B. C. J., and Gockel, H. (2002). Factors influencing sequential stream segregation. *Acta Acust.* 88, 320–332.

Phillips, D. P. (2008). A perceptual architecture for sound lateralization in man. *Hear. Res.* 238, 124–132. doi: 10.1016/j.heares.2007.09.007

Sach, A. J., and Bailey, P. J. (2004). Some characteristics of auditory spatial attention revealed using rhythmic masking release. *Percept. Psychophys.* 66, 1379–1387. doi: 10.3758/bf03195005

Shaw, E. A. G. (1974). Transformation of sound pressure level from the free field to the eardrum in the horizontal plane. *J. Acoust. Soc. Am.* 56, 1848–1861. doi: 10.1121/1.1903522

Shinn-Cunningham, B. G. (2005). "Influences of spatial cues on grouping and understanding sound," in *Proceedings of the Forum Acusticum, August 29, 2005*, Budapest.

Slattery, W. H. III, and Middlebrooks, J. C. (1994). Monaural sound localization: acute versus chronic unilateral impairment. *Hear. Res.* 75, 38–46. doi: 10.1016/0378-5955(94)90053-1

Stainsby, T. H., Fullgrabe, C., Flanagan, H. J., and Waldman, S. K. (2011). Sequential streaming due to manipulation of interaural time differences. *J. Acoust. Soc. Am.* 130, 904–914. doi: 10.1121/1.3605540

Stecker, G. C., Harrington, I. A., and Middlebrooks, J. C. (2005). Location coding by opponent neural populations in the auditory cortex. *PLoS Biol.* 3:e78. doi: 10.1371/journal.pone.000078

Subramony, S. H., Advincula, J., Perlman, S., Rosales, R. L., Lee, L. V., Ashizawa, T., et al. (2013). Comprehensive phenotype of the p.A*rg*420his allelic form of spinocerebellar ataxia type 13. *Cerebellum* 12, 932–936. doi: 10.1007/s12311-013-0507-6

van Hoesel, R. J. M. (2007). Sensitivity to binaural timing in bilateral cochlear implant users. *J. Acoust. Soc. Am.* 121, 2192–2206. doi: 10.1121/1.2537300

Waters, M. F., Fee, D., Figueroa, K. P., Nolte, D., Müller, U., Advincula, J., et al. (2005). An autosomal dominant ataxia maps to 19q13: allelic heterogeneity of SCA13 or novel locus? *Neurology* 65, 1111–1113. doi: 10.1212/01.wnl.0000177490.05162.41

Waters, M. F., Minassian, N. A., Stevanin, G., Figueroa, K. P., Bannister, J. P. A., Nolte, D., et al. (2006). Mutations in the voltage-gated potassium channel KCNC3 cause degenerative and developmental CNS phenotypes. *Nat. Genet.* 38, 447–451. doi: 10.1038/ng1758

Wightman, F. L., and Kistler, D. J. (1997). Monaural sound localization revisited. *J. Acoust. Soc. Am.* 101, 1050–1063. doi: 10.1121/1.418029

Wright, B. A., and Fitzgerald, M. B. (2001). Different patterns of human discrimination learning for two interaural cues to sound-source location. *Proc. Natl. Acad. Sci. U.S.A.* 98, 12307–12312. doi: 10.1073/pnas.211220498

Yao, J. D., Bremen, P., and Middlebrooks, J. C. (2015). Emergence of spatial stream segregation in the ascending auditory pathway. *J. Neurosci.* 35, 16199–16212. doi: 10.1523/jneurosci.3116-15.2015

Zeng, F. G. (2002). Temporal pitch in electric hearing. *Hear. Res.* 174, 101–106. doi: 10.1016/s0378-5955(02)00644-5

Zurek, P. (1993). "Binaural advantages and directional effects in speech intelligibility," in *Acoustical Factors Affecting Hearing Aid Performance*, eds G. A. Studebaker and I. Hochberg (Boston, FL: Allyn and Bacon), 255–276.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 27 October 2020
doi: 10.3389/fnins.2020.578566

Age-Related Deficits in Electrophysiological and Behavioral Measures of Binaural Temporal Processing

Tess K. Koerner[1], Ramesh Kumar Muralimanohar[1,2], Frederick J. Gallun[1,2] and Curtis J. Billings[1,2]*

[1] VA RR&D National Center for Rehabilitative Auditory Research, VA Portland Health Care System, Portland, OR, United States, [2] Department of Otolaryngology/Head and Neck Surgery, Oregon Health & Science University, Portland, OR, United States

Binaural processing, particularly the processing of interaural phase differences, is important for sound localization and speech understanding in background noise. Age has been shown to impact the neural encoding and perception of these binaural temporal cues even in individuals with clinically normal hearing sensitivity. This work used a new electrophysiological response, called the interaural phase modulation-following response (IPM-FR), to examine the effects of age on the neural encoding of interaural phase difference cues. Relationships between neural recordings and performance on several behavioral measures of binaural processing were used to determine whether the IPM-FR is predictive of interaural phase difference sensitivity and functional speech understanding deficits. Behavioral binaural frequency modulation detection thresholds were measured to assess sensitivity to interaural phase differences while spatial release-from-masking thresholds were used to assess speech understanding abilities in spatialized noise. Thirty adults between the ages of 35 to 74 years with normal low-frequency hearing thresholds were used in this study. Data showed that older participants had weaker neural responses to the interaural phase difference cue and were less able to take advantage of binaural cues for speech understanding compared to younger participants. Results also showed that the IPM-FR was predictive of performance on the binaural frequency modulation detection task, but not on the spatial release-from-masking task after accounting the effects of age. These results confirm previous work that showed that the IPM-FR reflects age-related declines in binaural temporal processing and provide further evidence that this response may represent a useful objective tool for assessing binaural function. However, further research is needed to understand how the IPM-FR is related to speech understanding abilities.

Keywords: aging, electrophysiology, interaural phase difference, binaural processing, IPM-FR, temporal processing, auditory steady-state response, auditory evoked potential

OPEN ACCESS

Edited by:
Achim Klug,
University of Colorado, United States

Reviewed by:
Pekcan Ungan,
Hacettepe University, Turkey
Fernando R. Nodal,
University of Oxford, United Kingdom

****Correspondence:***
Tess K. Koerner
Tess.Koerner@va.gov

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience

Received: *30 June 2020*
Accepted: *25 September 2020*
Published: *27 October 2020*

Citation:
Koerner TK, Muralimanohar RK, Gallun FJ and Billings CJ (2020) Age-Related Deficits in Electrophysiological and Behavioral Measures of Binaural Temporal Processing. Front. Neurosci. 14:578566. doi: 10.3389/fnins.2020.578566

INTRODUCTION

Accurate processing of binaural information is key to sound source localization and the detection of target signals in background noise. One cue used for binaural processing results from differences in the time of arrival of an auditory signal at the two ears. The interaural phase differences (IPDs) between the signal at each ear are detected at the level of the brainstem. The ability to

use these binaural IPD cues is dependent on accurate neural firing to the rapid fluctuations in signal amplitude over time, termed temporal fine structure, and the accurate comparison of these temporal cues between the ears. Recent research has shown that aging can impact the ability to process binaural temporal fine structure information independent of hearing loss, resulting in reduced IPD sensitivity (Ross et al., 2007a; Grose and Mamo, 2010, 2012b; Hopkins and Moore, 2011; Gallun et al., 2013, 2014; Papesh et al., 2017; Füllgrabe and Moore, 2018; Vercammen et al., 2018) and deficits understanding speech in background noise (Füllgrabe et al., 2015; Papesh et al., 2017). While the exact cause of this age-related decline in temporal processing is unknown, disruptions in neural synchrony, a slowing of neural activity, a loss of cochlear afferent synapses, deficits in the central integration of binaural information, and/or deficits in the central encoding of binaural information have been known to occur with aging (He et al., 2008; Grose and Mamo, 2010; Ruggles et al., 2012; King et al., 2014; Shaheen et al., 2015; Whiteford et al., 2017; Parthasarathy and Kujawa, 2018; Wu et al., 2019).

Recent reports have focused on the use of non-invasive electrophysiological measures to assess the effects of age on the neural encoding of binaural temporal cues (Ross et al., 2007a; Ross, 2008; Wambacq et al., 2009; Grose and Mamo, 2012a; Ozmeral et al., 2016; Papesh et al., 2017; Eddins and Eddins, 2018; Vercammen et al., 2018; Ungan et al., 2020). For example, Ross et al. (2007a) and Papesh et al. (2017) recorded transient event-related potentials in response to a 180° IPD using magnetoencephalography and electroencephalography, respectively. The IPD cue was embedded in the temporal fine structure of the stimulus by shifting an amplitude modulated (AM) tone from a diotic to a dichotic presentation. This work showed that aging impacted the neural encoding of the IPD cue (Ross et al., 2007a; Papesh et al., 2017) and that age-related changes in neural responses to the IPD cue were associated with individual behavioral limits of IPD discrimination (Ross et al., 2007a), which suggests that this type of measure may represent a robust tool for the neurophysiological assessment of binaural temporal processing abilities.

Recently, there has been growing interest in the use of a new electrophysiological measure called the interaural phase modulation-following response (IPM-FR), which has been developed as a more efficient method to assess the neural encoding of IPDs (Haywood et al., 2015; McAlpine et al., 2016; Undurraga et al., 2016). Similar to neural responses measured by Ross et al. (2007a; 2007b) and Papesh et al. (2017), this response is elicited by a shift in the phase of an AM carrier tone at the two ears. However, rather than a single transient response evoked by a single phase shift, the IPM-FR is a steady-state response to periodic shifts in phase embedded in the temporal fine structure of an ongoing AM tone, which results in a higher number of neural responses over a shorter period of time. This response can be objectively assessed through spectral analysis of the electroencephalographic response in the frequency bin corresponding to the rate at which the phase changes. This provides an additional advantage over the electrophysiological measure used by Papesh et al. (2017), which requires the detection of transient response peaks in the time domain.

To date, only one study has examined whether the IPM-FR is sensitive to the effects of age on binaural temporal processing (Vercammen et al., 2018). Similar to findings from studies that used different electrophysiological measurement techniques (Ross et al., 2007a; Ross, 2008; Ozmeral et al., 2016; Papesh et al., 2017; Ungan et al., 2020), Vercammen et al. (2018) showed that the neural encoding of IPD cues tends to be stronger in younger participants compared to older participants and, along with several other studies (Haywood et al., 2015; Undurraga et al., 2016; Vercammen et al., 2018; Parthasarathy et al., 2020), showed that IPM-FR responses tend to be weaker when behavioral detection of the IPD cue is poor. While these previous studies have established that the IPM-FR is likely reflective of behavioral IPD sensitivity, the stimuli previously used to assess behavioral IPD discrimination thresholds were dichotic AM stimuli analogous to those used to elicit the IPM-FR (Haywood et al., 2015; Undurraga et al., 2016; Vercammen et al., 2018; Parthasarathy et al., 2020). Assessing relationships between the IPM-FR and other behavioral measures of binaural temporal processing will determine whether associations between the IPM-FR and behavior can generalize to other stimuli and tasks that assess IPD sensitivity.

In addition, examining relationships between electrophysiological responses and measures of speech understanding in noise can provide information about neural processes that underlie speech perception deficits. For example, Papesh et al. (2017) showed that variability in neural responses to the IPD cue described above were more predictive of participants' spatial release-from-masking abilities than age and/or hearing loss. In other words, the electrophysiological measure used by Papesh et al. (2017) was better able to reflect the integrity of binaural temporal processing mechanisms that are important for speech understanding in spatialized noise than participant factors such as age and estimated hearing sensitivity. In addition to recently confirming that the IPM-FR reflects behavioral IPD sensitivity, Parthasarathy et al. (2020) was the first to examine relationships between IPM-FRs and speech perception. Speech understanding was assessed using a competing digits task where a string of digits spoken by target and competing speakers were presented diotically. As expected, the IPM-FR did not represent a link to speech understanding abilities due to the absence of binaural cues in the behavioral task (Parthasarathy et al., 2020). Therefore, to date, it is unknown whether the IPM-FR is also a neural correlate of functional speech understanding abilities in realistic listening environments that require the use of binaural cues. An exploration of additional relationships between neural and behavioral measures will allow for a better understanding of age-related declines in neural processes underlying behavioral measures of binaural function and will provide information about what types of tasks draw upon similar neural resources.

Therefore, while the current study was designed to confirm the effects of age on the IPM-FR, it primarily aimed to evaluate relationships between the IPM-FR and several new behavioral measures of binaural temporal processing to (1) confirm that the IPM-FR is reflective of IPD sensitivity and (2) determine whether variability in the IPM-FR is related to speech understanding

in noise abilities. Many common behavioral tests of binaural temporal processing tend to require extensive training periods on the part of the examiner and the listener as well as a high number of stimulus repetitions to obtain reliable estimates of binaural sensitivity (Stecker and Gallun, 2012). These issues have motivated recent research efforts that focus on new implementations of existing laboratory tests which would not require extensive resources, time, or training on the part of the experimenter or participant (Gallun et al., 2013; Füllgrabe and Moore, 2017, 2018; Jakien et al., 2017; Füllgrabe et al., 2018; Jakien and Gallun, 2018; Hoover et al., 2019; Lelo de Larrea-Mancera et al., 2020). Hoover et al. (2019) recently adapted a dichotic frequency modulation (FM) detection task that uses a frequency modulated signal that is inverted in phase at one ear relative to the other to create IPD cues (Grose and Mamo, 2012b; Whiteford et al., 2015). Performance on this task has been shown to be impacted by age, such that older listeners tend to have higher dichotic FM detection thresholds than younger- and middle-aged listeners, reflecting potential age-related declines in temporal fine structure processing (Grose and Mamo, 2012b) or central binaural integration processes (Whiteford et al., 2017). Hoover et al. (2019) showed that, compared to other behavioral tests designed to assess temporal fine structure processing, the FM detection tasks were among the most consistent and efficient measures of binaural processing in a group of young, normal-hearing participants. Several studies have also established the use of a measure of spatial release from masking using speech stimuli for the assessment of binaural function (Gallun et al., 2013; Jakien et al., 2017; Jakien and Gallun, 2018). Spatial release from masking refers to the increased ability to detect a target sound or speech stream of interest when it is spatially separated from one or more maskers (Hawley and Litovsky, 2004; Gallun et al., 2005). Studies have shown that older listeners tend to receive less benefit from the spatial separation of target and background speech streams compared to younger listeners (Gallun et al., 2013; Jakien et al., 2017; Jakien and Gallun, 2018). Spatial release-from-masking thresholds appear to be consistent across different modes of testing, including testing under headphones in a virtual space using traditional laboratory equipment (Jakien et al., 2017) as well as a tablet-based automated rapid-testing version of the task recently developed in the Portable Automated Rapid Testing application (PART; Gallun et al., 2018; Lelo de Larrea-Mancera et al., 2020). Taken together, evidence from these previous studies suggests that measures of binaural FM sensitivity and spatial release from masking may represent efficient and reliable testing tools that are sensitive to age-related changes in binaural function, which motivated their selection for use in the current study.

An additional practical aim of the current study was to explore potential differences in IPM-FRs recorded using different stimulus parameters. Since the stimulus used to elicit the IPM-FR is amplitude modulated, a steady-state response that follows the AM rate occurs in addition to the steady-state neural response to the phase reversal rate. This envelope following response, commonly known as the auditory steady-state response (ASSR), provides a measure of neural phase locking at the place of stimulation equal to the carrier frequency. The ability to concurrently elicit the ASSR along with the IPM-FR is an additional advantage of this electrophysiological measure, as the ASSR may serve as an estimate of hearing sensitivity (Dimitrijevic et al., 2002; Picton et al., 2003), temporal envelope processing, as well as an index of recording and response quality. Previous studies that have examined the IPM-FR have used stimuli that were amplitude modulated at lower (∼40 Hz; Haywood et al., 2015; Undurraga et al., 2016; Parthasarathy et al., 2020) and higher (∼80 Hz; Vercammen et al., 2018) rates. Although still a source of ongoing investigation (Coffey et al., 2019), it is generally thought that ASSRs elicited using a higher AM rate are primarily generated from more subcortical brainstem structures, including the superior olivary complex and inferior colliculus, while ASSRs elicited using a lower AM rate activate overlapping brainstem structures as well as additional neural generators located in the auditory cortex (Giraud et al., 2000; Herdman et al., 2002; Korczak et al., 2012). To date, no studies have directly compared IPM-FRs elicited using different modulation rates in the same individuals. While it is known that changing the AM rate of the stimulus impacts the strength of the ASSR (Levi et al., 1993), it is unknown if, or how, changing the AM rate may impact the IPM-FR. An analysis of this type may provide information regarding the optimization of specific IPM-FR stimulus parameters for a more efficient or reliable assessment of the neural encoding of IPD cues or for a stronger neurophysiological link to behavioral performance.

This work aimed to (1) confirm previous findings that the IPM-FR, measures of binaural FM detection, and measures of spatial release from masking are sensitive to the effects of age on binaural temporal processing; (2) evaluate associations between the IPM-FR and performance on behavioral measures of IPD sensitivity and speech perception; and (3) explore differences in IPM-FRs elicited using different AM rates. It was predicted that, consistent with previous work (e.g., Grose and Mamo, 2012b; Gallun et al., 2013, 2014; Papesh et al., 2017; Füllgrabe and Moore, 2018; Vercammen et al., 2018; Ungan et al., 2020), age would have a significant effect on each neural and behavioral measure of binaural temporal processing, such that older participants would have reduced IPM-FRs, reduced dichotic FM thresholds, and reduced speech understanding abilities compared to younger participants. In addition, it was predicted that older participants would show reduced benefit from the addition of binaural cues in both the FM detection tasks and the spatial release-from-masking tasks compared to younger participants. It was also expected that the strength of the IPM-FR would be predictive of performance on these behavioral measures, suggesting that each measure depends on overlapping neural mechanisms for binaural temporal processing. Finally, while it was unknown whether or not the use of different AM rates would impact the IPM-FR, this manipulation was anticipated to provide important practical guidance on the degree to which AM rate influences IPM-FR strength and its relationship with other measures. Taken together, the results of this work will have important implications for the development and use of measures designed to assess binaural temporal processing abilities.

MATERIALS AND METHODS

Participants

Participants included 30 adults (11 female, 19 male) who ranged in age from 35 to 74 years (mean age: 62.3 years). All participants were right-handed and were native speakers of American English. No participants reported taking medications that impacted sleep or had mood-altering affects. Pure-tone hearing thresholds for each participant as well as mean pure-tone thresholds averaged across participants are depicted for the right and left ears in **Figure 1**. All participants had hearing thresholds within normal limits (≤25 dB HL) at 500 Hz as measured by a standard pure-tone audiological assessment. This inclusion criterion was specifically chosen because 500 Hz was the carrier frequency used in our neural and behavioral measures of IPD sensitivity. Hearing thresholds at higher frequencies ranged from within normal limits to moderate-to-severe sensorineural hearing loss. No participants had asymmetrical hearing thresholds, as defined as a difference in four-frequency pure-tone averages (0.5, 1, 2, and 4 kHz) greater than 15 dB across ears. All participants provided informed consent prior to their participation in the study and all participants were paid for their participation. This work was approved by the joint Institutional Review Board of the Department of Veterans Affairs Portland Health Care System and Oregon Health & Science University.

Procedure

Neural Measures

Interaural phase modulation-following responses were elicited with a 500-Hz tone that was 100% sinusoidally amplitude modulated (AM) at either 40.8 or 81.6 Hz. The modulation envelope of the stimulus remained diotic, but the phase of the carrier frequency was square-wave modulated at a rate of 6.8 Hz such that when the instantaneous phase at one ear was +45°, the phase in the other ear was −45°, creating an IPM depth of ±90°. The 6.8-Hz IPM rate was chosen based on the results of McAlpine et al. (2016), who showed that when compared to slower rates, the 6.8-Hz IPM produced a steady-state neural response and that this neural response was stronger than those elicited by faster IPM rates. Similarly, the single IPM depth of ±90° was chosen because it was previously shown to produce the largest IPM-FR compared to those elicited by higher or lower IPM depths (Haywood et al., 2015; Undurraga et al., 2016). The IPD was introduced at the zero-amplitude minimum in the AM cycle (**Figure 2**), which helped ensure that neural responses to this stimulus were to shifts in the temporal fine structure, and not due to monaural off-frequency cues (brief broadening of the activation pattern at the level of the cochlea) or modulations in the stimulus envelope. These stimulus parameters are similar to those previously used by Haywood et al. (2015) and Undurraga et al. (2016). The dichotic test stimulus was presented for 5 s and was repeated 75 times per recording block. Seventy-five repetitions of a 5 s diotic control stimulus were also presented in an alternating manner within each recording block with an inter-stimulus interval of 20 ms. This stimulus was identical to the dichotic stimulus described above, except that it had a zero IPD. In other words, phase transitions of equal magnitude occurred at a rate of 6.8 Hz but were the same in both ears. This diotic control stimulus has been used previously to examine the degree to which neural responses at the IPM rate in the dichotic stimulus can be attributed to the introduction of the IPD

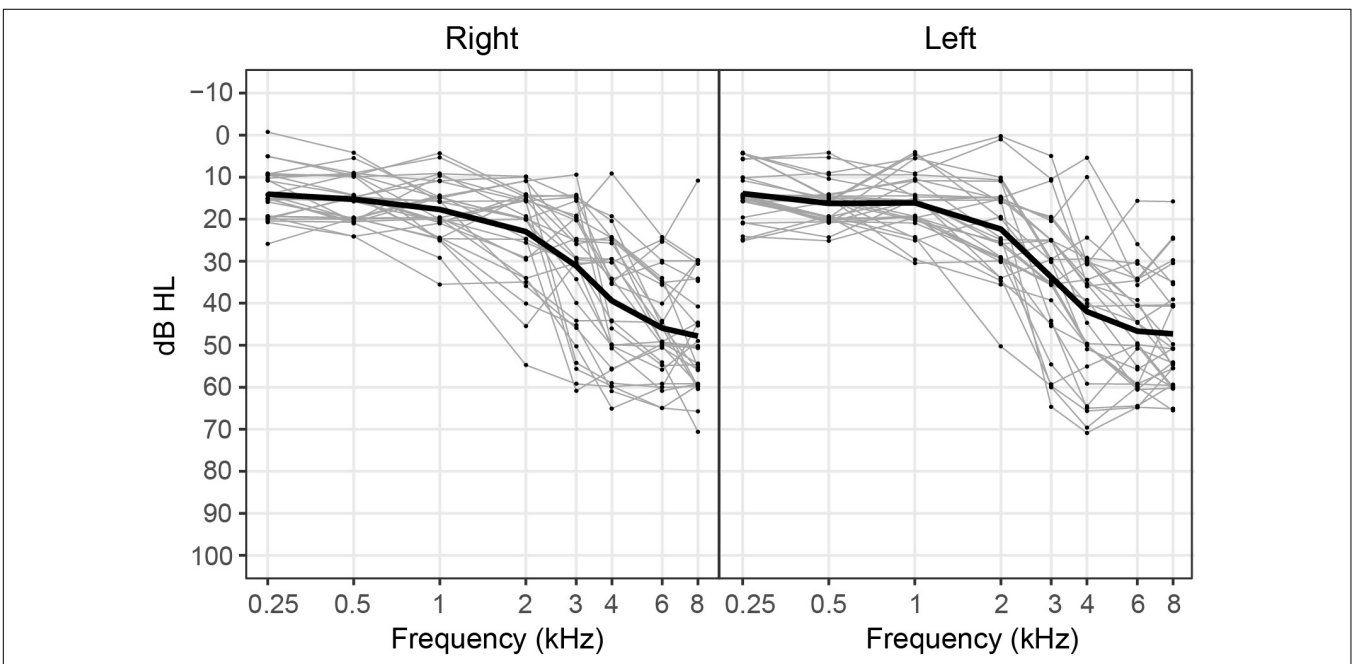

FIGURE 1 | Pure-tone hearing thresholds for each participant (gray lines) and mean hearing thresholds averaged across participants (black line) plotted for the right and left ears. Note that data points depicting individual hearing thresholds are jittered within 1 dB for ease of visualization.

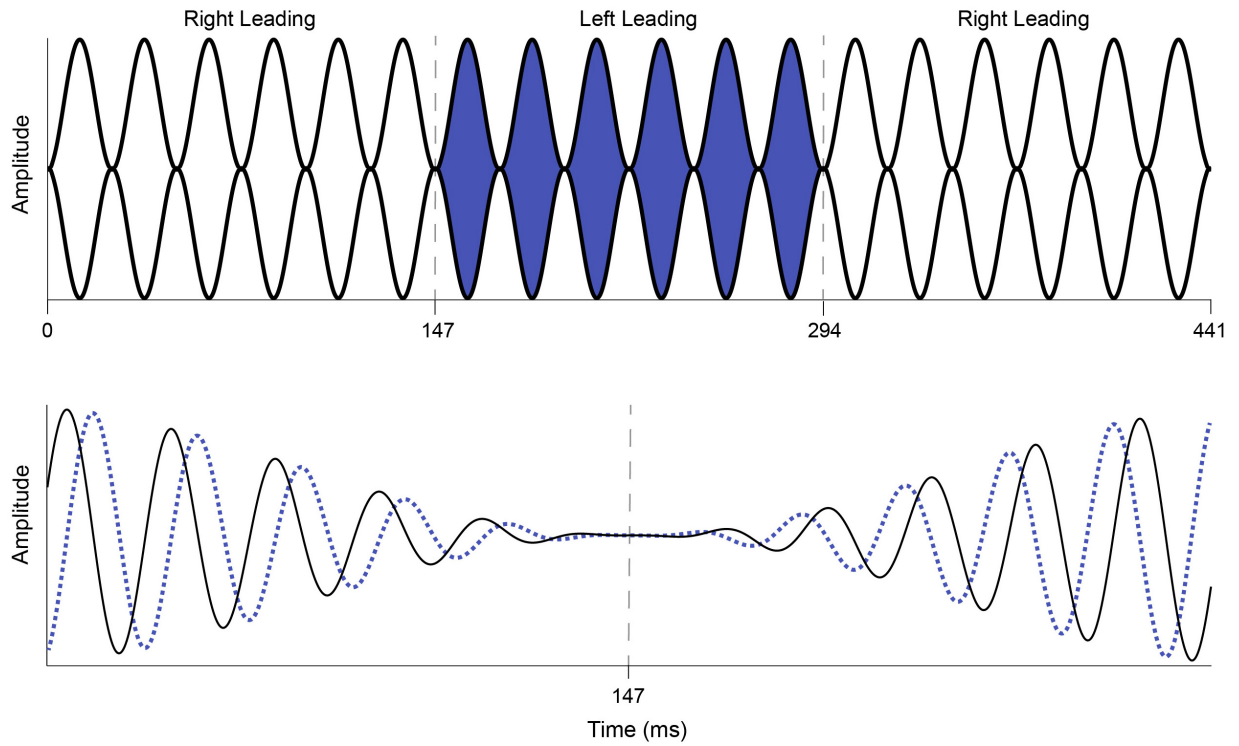

FIGURE 2 | Illustration of the dichotic AM stimulus used to elicit the IPM-FR. The top panel represents a 40.8-Hz AM tone that shifts from right leading (white) to left leading (blue) and back to right leading (white) at an IPM rate of 6.8 Hz. The interaural phase modulations depicted in this figure occur at 147 and 294 ms and are indicated by dashed gray lines. The bottom panel depicts an expanded view of the first phase change in the dichotic AM stimulus shown above. The right ear leads in phase (+90° IPD, solid black line) until the phase switch at 147 ms, at which point the left ear leads in phase (−90° IPD, dotted blue line). The phase change occurs at the zero crossing to avoid audible artifacts.

and not to monaural phase cues (Ross et al., 2007a,b; Haywood et al., 2015; Undurraga et al., 2016). While this diotic control stimulus can serve the same purpose in the current study, it was mainly used to calculate the signal-to-noise ratio (SNR) of the IPM-FR, as described below. Each recording block containing alternating diotic and dichotic stimuli was presented twice, for a total of 150 presentations of each dichotic and diotic stimulus. Separate blocks were recorded at each AM rate, for a total of four recording blocks. The presentation order of each recording block was randomized across participants.

Stimuli were presented bilaterally at 80 dB SPL through Etymotic ER2 insert earphones using Neuroscan software (Compumedics Neuroscan Stim2/Scan 4.5; Charlotte, NC, United States) in an acoustically treated and electrically shielded booth. IPM-FRs were recorded from a 64-channel tin-electrode cap (Electro-Cap International, Inc.; Eaton, OH, United States) with the ground electrode on the forehead and the reference electrode at Cz. Responses were analog low-pass filtered on-line at 200 Hz and were converted using an analog-to-digital sampling rate of 1,000 Hz. Application and preparation of the electrode cap, along with IPM-FR recording, lasted approximately 2 h. During recording, participants were seated in a comfortable chair, asked to relax, minimize eye and muscle movements, and watch a movie of their choice with subtitles.

The electrophysiological data was processed in Neuroscan Edit (Neuroscan, 2007). Eye-blink artifacts were corrected offline using eye movement information collected from the horizontal and vertical planes from electrodes located inferior to and at the outer canthi of both eyes. A spatial, singular value decomposition was used to calculate the amount of covariation between a vertical eye channel and each electrode. This vertical eye blink activity was then removed from each electrode on a point-by-point basis. The remaining ongoing response recorded in each condition was then high-pass filtered at 2 Hz and epoched from 0 to 5000 ms relative to stimulus onset. Epochs exceeding $\pm 100\ \mu V$ were rejected from analysis. The spectrum of each five second epoch was computed using a 10,000 point Fast Fourier Transform (FFT) which provided a resolution of 0.2 Hz. Mean response magnitudes were calculated by vector averaging these FFT values across all epochs in each condition. IPM-FRs were obtained as the spectral magnitude in the 6.8-Hz bin while ASSRs were obtained as the spectral magnitude in the 40.8- or 81.6-Hz bins for each participant from the electrode at the right mastoid (M2). This electrode site was chosen for analysis based on a previous report that IPM-FR and ASSR magnitudes tended to be largest across participants at the right mastoid (Haywood et al., 2015), which is a pattern consistent with neural responses collected in the current study. The magnitude of the neural activity in the 6.8-Hz bin

in the diotic condition at electrode M2 was used as a control condition. An estimate of signal-to-noise ratio of the IPM-FR was calculated for each participant using magnitudes in the 6.8-Hz bin in both the dichotic and diotic conditions, using the formula below:

$$\text{dB SNR} = 20log_{10}(\frac{V_{dichotic}}{V_{diotic}})$$

The measured IPM-FR is thought to consist of energy related to the steady-state response to the IPM rate as well as neural activity that is thought to randomly vary in phase and amplitude over time. Therefore, both the magnitude of the response at 6.8 Hz and this SNR metric were used as measures of the IPM-FR in an attempt to better characterize response strength.

Behavioral Measures

The behavioral binaural FM detection tasks and spatial release-from-masking tasks were completed using test batteries included in PART (Lelo de Larrea-Mancera et al., 2020) on an iPad with Sennheiser 280 Pro circumaural headphones (Wedemark, Germany). All stimuli were presented at a level of 80 dB SPL and all testing was completed in a sound-treated and electrically shielded booth. Each behavioral task was repeated twice and the presentation order was randomized across participants.

Diotic and dichotic FM detection thresholds were measured using a four-interval, two-cue, two-alternative forced choice procedure (Hoover et al., 2019; Lelo de Larrea-Mancera et al., 2020). Each standard stimulus was an unmodulated pure tone and the target stimulus was a pure tone carrier with a 6.8-Hz sinusoidal modulator that was presented in either the 2nd or 3rd stimulus interval on a touchscreen iPad display. This modulation rate was chosen to match the modulation rate of the stimulus used to elicit the IPM-FR, described above. All standard and target stimulus presentations were 400 ms long and had carrier frequencies that were randomly selected from a uniform distribution that ranged from 460 to 540 Hz. Randomizing the carrier frequency in this way lessens the possibility that participants use a place cue to detect the frequency modulated target stimulus from the unmodulated pure-tone standard stimuli. All stimuli had inter-stimulus intervals of 250 ms. Modulation depth was adaptively varied in logarithmic steps using the algorithm described in Lelo de Larrea-Mancera et al. (2020). This staircase procedure estimates the lowest rate, in Hz, at which a given individual can just detect the frequency modulation. In the diotic testing condition, the target FM signal was the same across both ears with a starting phase of 0 radians. Participants were instructed to detect the interval with the frequency modulated, "warbling" stimulus. In the dichotic testing condition, the target stimulus contained monaural FM that was out of phase at the two ears, with modulator starting phases of 0 and π radians. The resulting interaural phase modulation created dynamic interaural time difference (ITD) cues which created the percept of the signal moving in the head between the two ears at a rate of 6.8 Hz. The dichotic FM threshold, in Hz, corresponds to the smallest ITD that can be detected in this stimulus. An example of a dichotic FM stimulus is provided in **Figure 3**. Note that to improve visualization of the dichotic FM, the stimulus depicted in **Figure 3** has a lower carrier frequency and higher modulation rate compared to the dichotic FM stimulus used in the current study. Participants were not given any practice trials for either the diotic FM or dichotic FM conditions. Participants received feedback on each trial that indicated whether their selection was correct or incorrect. Detection thresholds were log transformed. The better of the two diotic FM detection thresholds and the better of the two dichotic FM detection thresholds were chosen for analysis. In addition, a difference score was calculated to estimate benefit from the addition of binaural cues for FM detection. For this calculation, the difference in individual dichotic FM and diotic FM detection thresholds was taken relative to individual performance in the diotic FM condition (Grose and Mamo, 2012b).

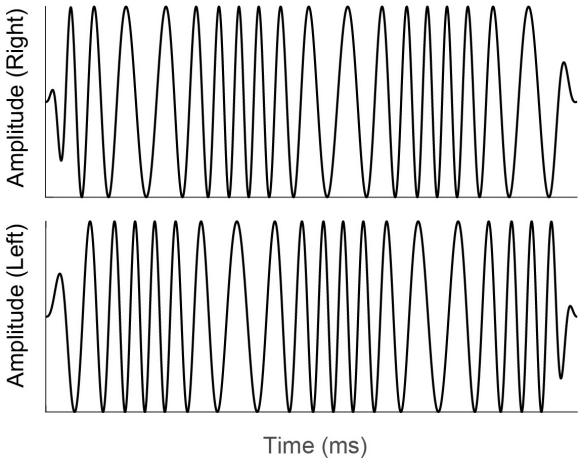

FIGURE 3 | Illustration of a dichotic FM stimulus similar to that used in the current study. Note that monaural FM that is out of phase at the right (top panel) and left (bottom panel) ears creates dynamic interaural time difference cues. Participants were instructed to select the stimulus interval with the interaural time difference cues that was perceived as a signal moving in the head between the two ears.

Spatial release from masking was measured using sentences from three male speakers from the Coordinate Response Measure (Bolia et al., 2000). The target and masker sentences were presented in colocated and spatially separated listening conditions, which were used to calculate spatial release from masking (Gallun et al., 2013; Jakien et al., 2017). Sentences were of the form: "Ready (CALL SIGN) go to (COLOR) (NUMBER) now." Participants were instructed to choose the appropriate color and number combination associated with the call sign "Charlie," which was always spoken by the talker located at 0°. Participants made choices on a color-number grid that was presented on the touchscreen iPad display. Distractor speakers each used one of seven different callsigns, such as "Eagle" or "Baron," and different color number combinations from those spoken by the target speaker. The distractors were located at +45° and −45° in the separated listening condition and at 0° in the more difficult colocated listening condition. The locations of the target talker and each distractor talker were simulated by convolving the anechoic sentences with the head-related impulse responses measured at those locations in the horizontal plane (see Gallun et al., 2013). To familiarize participants with the

response format prior to testing, a short practice session was provided in which the target "Charlie" sentences were spoken without any competing distractor speakers. Progressive tracking was used in each testing repetition as described in Gallun et al. (2013), which involves reducing the target-to-masker ratios from 10 to −8 dB in 2 dB steps. Participants were given two trials at each target-to-masker ratio. Feedback was provided on each trial to indicate whether the response was correct or incorrect. Target-to-masker thresholds (in dB), which approximate the point at which performance is 50%, were calculated by subtracting the number of correct responses from the starting target-to-masker ratio of 10 dB (see Gallun et al., 2013 for further details). Thresholds for the separated and colocated conditions were averaged across the two separate testing repetitions for each participant. Spatial release from masking, in dB, was calculated as the difference in threshold from the colocated to the spatially separated listening condition.

Analysis

Separate linear regression models were created in R (R Core Team) using the *nlme* package (Pinheiro et al., 2016) to assess the effects of age and hearing sensitivity on each test measure, with $\alpha = 0.05$. Either participant age or hearing sensitivity were added as a fixed effect while separate models were used to evaluate the various outcome measures: IPM-FR magnitude or SNR in each AM rate condition, diotic FM or dichotic FM detection threshold and binaural FM difference score, and target-to-masker ratio threshold and spatial release-from-masking threshold. The effects of age and hearing sensitivity on the ASSR magnitude in each recording condition were also assessed to determine the potential effects of each factor on this neural response. Hearing sensitivity was estimated by averaging hearing thresholds across frequencies and ears to account for potential effects of variations in high frequency hearing sensitivity across participants.

Similar linear regression models were created to assess relationships between the neural measures and performance on the binaural FM and spatial release-from-masking measures. In these models, either IPM-FR magnitude or SNR from each AM rate condition was added as a fixed effect to predict each behavioral outcome measure. Paired *t*-tests, with $\alpha = 0.05$, were used to assess potential effects of AM rate on the IPM-FR magnitude and SNR. Similarly, paired *t*-tests were used to assess the effects of stimulus condition (diotic vs. dichotic) on ASSR magnitude. Since the AM remained diotic in each stimulus condition, a comparison of ASSR magnitude to the dichotic test stimulus which contained IPDs and the diotic test stimulus which did not contain IPDs would assess the degree to which the recording and response quality remained comparable across alternating stimulus repetitions.

RESULTS

Effects of Age and Hearing Sensitivity on Neural and Behavioral Measures

Participant age was not significantly associated with hearing sensitivity as measured by an average of pure-tone thresholds across frequency and ears [$F(1,28) = 2.79$, $p = 0.11$, $R^2 = 0.09$] or as measured by 500-Hz thresholds averaged across ears [$F(1,28) = 1.13$, $p = 0.30$, $R^2 = 0.04$].

Neural Measures

For visualization purposes, grand mean time-domain response waveforms are shown in **Figure 4** for each stimulus condition. The IPM-FR is observed as the steady-state neural response that follows the 6.8-Hz IPM rate in the dichotic test conditions. As expected, this steady-state response is absent in the diotic control conditions which contained no IPDs. While the IPM-FR can be observed in the time-domain, it is more easily examined and measured in the frequency domain. Grand mean response spectra averaged across each participant are provided in **Figure 5**. **Figure 5** also includes individual response spectra to illustrate the range of IPM-FR and ASSR magnitudes recorded across participants[1]. The IPM-FR is clearly observed as a response peak at 6.8 Hz in each dichotic test condition (indicated by the arrows in **Figures 5A,B**). Harmonics of the 6.8-Hz IPM rate can also be observed in each dichotic test condition. In addition, the ASSR can be observed as a response peak corresponding to the AM rate of each stimulus. The IPM-FR response peak at 6.8 Hz and subsequent harmonics are absent in the diotic control conditions for both stimuli, but ASSRs are still observed at the 40.8- and 81.6-Hz AM rates (**Figures 5C,D**, respectively). **Figures 4**, **5** are provided for the visualization of example IPM-FRs and ASSRs and to illustrate general response trends for each stimulus condition. Given that the main purpose of this study was to examine the effects of age on binaural temporal processing using the IPM-FR and to examine relationships between neural responses and behavioral performance, individual neural responses were used for all statistical analyses.

Results from the linear regression models that were used to assess the effects of age or hearing sensitivity on each neural measure are provided in **Table 1**. Analysis indicated significant effects of age on IPM-FR magnitude in the 40.8-Hz AM rate condition [$F_{(1,28)} = 7.17$, $p = 0.01$, $R^2 = 0.20$] and significant effects of age on both IPM-FR magnitude [$F_{(1,28)} = 14.67$, $p < 0.001$, $R^2 = 0.34$] and SNR [$F_{(1,28)} = 7.92$, $p = 0.009$, $R^2 = 0.22$] in the 81.6 Hz-AM rate condition. For these measures, older individuals tended to have weaker IPM-FRs compared to younger individuals (**Figure 6**). The lack of an age-related effect on response magnitude in the 6.8-Hz bins in the diotic control conditions suggests that age did not impact estimates of background noise. Variability in hearing sensitivity did not have a significant effect on any neural measure.

Additional analyses were completed to assess the potential effect of age on the ASSR in each test condition. For the 40.8-Hz AM rate condition, age had a significant effect on ASSR

[1]While neural responses from all 30 participants were used to calculate the grand mean spectra illustrated in **Figure 5**, to facilitate ease of visualization of the IPM-FR and ASSR in each panel, individual responses from one participant were removed from **Figures 5B,D** due to the presence of 60 Hz electrical noise that was introduced in one recording block when sound booth power was not turned off prior to recording.

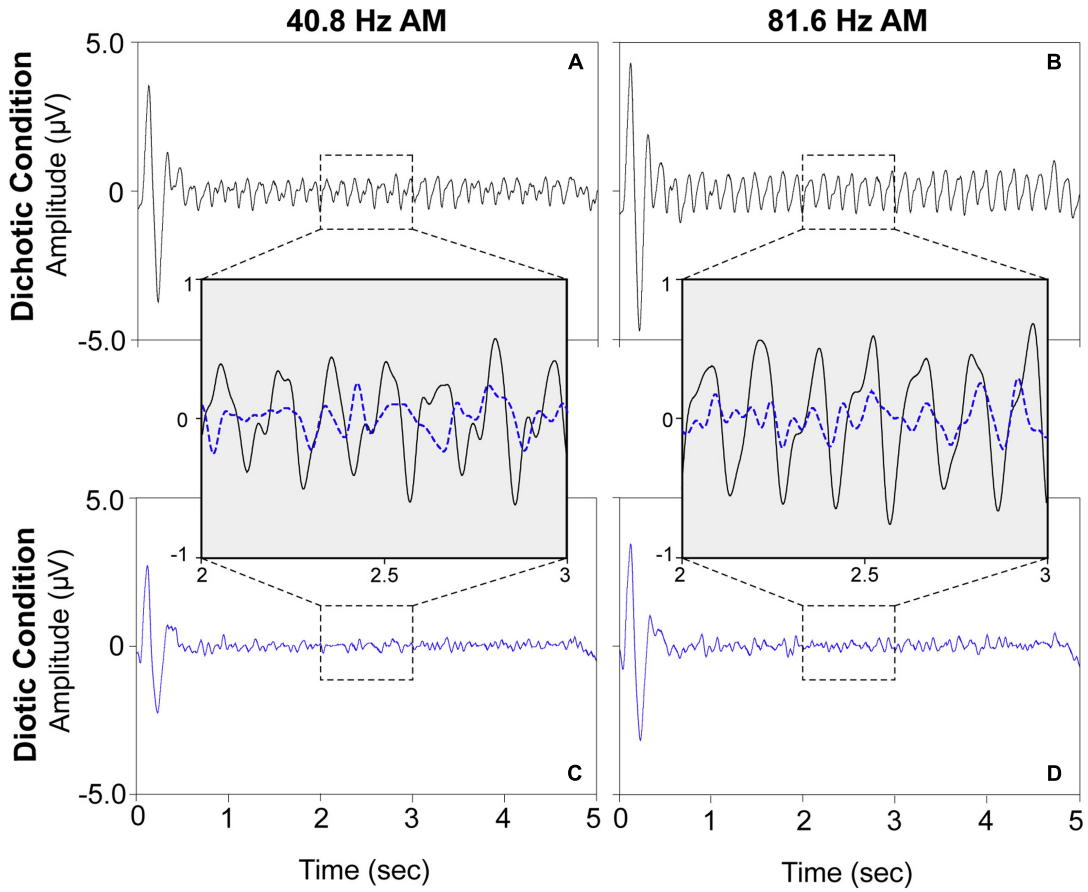

FIGURE 4 | Grand mean time-domain response waveforms for each stimulus condition averaged across all 30 participants at electrode M2. The dichotic test stimuli elicited steady-state neural responses at the IPM rate of 6.8 Hz (panels **A,B**) which are not observed in response to the diotic control stimuli (panels **C,D**). A direct comparison of a 1 s segment of the neural responses to the dichotic (black) and diotic (dotted blue) stimuli are provided in inset panels for the 40.8-Hz AM stimuli (left) and the 81.6-Hz AM stimuli (right). For visualization of the IPM-FR in the time domain, the ongoing EEG signal was bandpass filtered from 2–20 Hz.

magnitude in both the dichotic [$F_{(1,28)}$ = 10.85, p = 0.003, R^2 = 0.28] and diotic [$F_{(1,28)}$ = 12.26, p = 0.002, R^2 = 0.30] test conditions, such that older individuals had smaller ASSR magnitudes compared to younger participants. These age-related changes in ASSR magnitudes were not present in either the dichotic or diotic test conditions when the higher 81.6-Hz AM stimulus was used.

Behavioral Measures

Results from the linear regression models that were designed to examine the effects of age and hearing sensitivity on each behavioral measure are provided in **Table 2**. This analysis revealed a significant effect of age on diotic FM detection [$F_{(1,28)}$ = 4.40, p = 0.04, R^2 = 0.14] such that older individuals had higher FM detection thresholds, and therefore poorer performance on this task compared to younger individuals. However, results did not indicate that age had a significant effect on performance in the dichotic FM condition or on the FM difference score. Both age and hearing sensitivity had significant effects on target-to-masker ratio thresholds in the separated listening condition of the spatial release-from-masking task [Age: $F_{(1,28)}$ = 8.44, p = 0.007, R^2 = 0.23; Hearing Sensitivity: $F_{(1,28)}$ = 4.79, p = 0.04, R^2 = 0.15]. Neither age nor hearing sensitivity had a significant impact on performance in the more difficult colocated condition of this task where the distractor speakers were located at the same azimuth as the target speaker. When performance on these two listening conditions were compared, analyses revealed a significant impact of both age and hearing sensitivity on spatial release-from-masking thresholds [Age: $F_{(1,28)}$ = 7.04, p = 0.01, R^2 = 0.20; Hearing Sensitivity: $F_{(1,28)}$ = 4.45, p = 0.04, R^2 = 0.14]. For these measures, participants who were older or who had poorer average across-frequency hearing thresholds required more favorable target-to-masker ratios to obtain 50% performance (**Figure 7**).

Relationships Between Neural Responses and Behavioral Performance

Linear regression models were used to test relationships between the IPM-FR and performance on behavioral measures of binaural processing. Results from these regression models are provided in **Tables 3**, **4**.

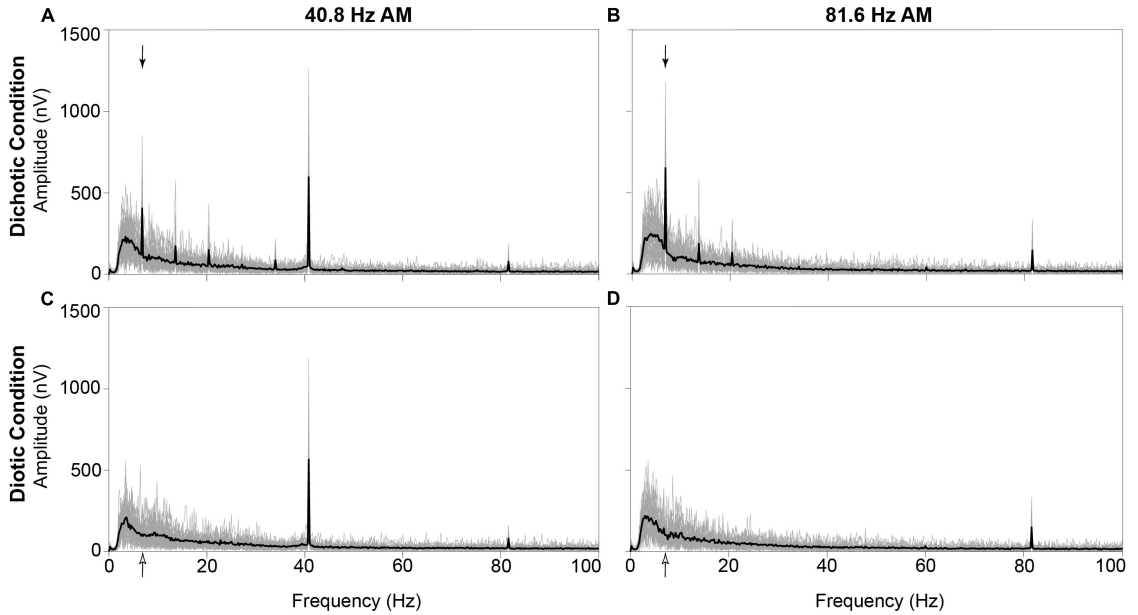

FIGURE 5 | Grand mean neural responses in the frequency domain averaged across all 30 participants (black line) and individual neural response spectra from each participant (gray lines) at electrode M2. The IPM-FR is clearly observed as a peak at 6.8 Hz in response to the dichotic test stimuli (denoted by filled arrows in panels **A,B**). As expected, this response peak is not observed at 6.8 Hz in response to the diotic control stimuli (6.8-Hz bin denoted by unfilled arrows in panels **C,D**). In addition, ASSRs are observed at 40.8 Hz (panels **A,C**) and 81.6 Hz (panels **B,D**) in response to both the dichotic and diotic test stimuli.

Binaural FM Detection

Analysis showed that the IPM-FR SNR in the 81.6-Hz AM condition was significantly predictive of dichotic FM detection thresholds [$F_{(1,28)}$ = 16.55, p < 0.001, R^2 = 0.37] and the FM difference score [$F_{(1,28)}$ = 18.76, p < 0.001, R^2 = 0.40]. Relationships between the IPM-FR SNR and these binaural FM measures are depicted in **Figure 8**, which shows that weaker IPM-FRs are associated with poorer behavioral IPD sensitivity as well as less benefit from the addition of binaural information to the FM detection task. Adding either participant age or estimated hearing sensitivity to these models did not predict any additional variance in performance. The IPM-FR SNR in the 81.6-Hz AM rate condition was also predictive of diotic FM detection thresholds [$F_{(1,28)}$ = 4.34, p = 0.046, R^2 = 0.13]. However, further analyses revealed that this relationship was primarily mediated by the effect of age. When the effects of age were accounted for in the model, the relationship between the IPM-FR SNR and diotic FM thresholds was no longer significant [$F_{(1,27)}$ = 1.58, p = 0.22]. IPM-FR magnitudes in either AM rate condition and IPM-FR SNRs in the 40.8-Hz AM condition were not related to performance on any binaural FM measure.

TABLE 1 | Test statistics from linear regression models that examined the effects of participant age and average hearing sensitivity across frequency on each neural measure.

Neural measure	AM Rate (Hz)	Age			Hearing sensitivity		
		F	*p*	R^2	*F*	*p*	R^2
IPM-FR Magnitude (nV): Dichotic Condition	40.8	7.17	0.01	0.20	0.05	0.82	0.00
	81.6	14.67	<0.001	0.34	1.05	0.31	0.04
Control Magnitude (nV): Diotic Condition	40.8	0.15	0.70	0.00	0.01	0.92	0.00
	81.6	0.02	0.89	0.00	1.37	0.25	0.05
IPM-FR SNR (dB)	40.8	3.48	0.07	0.11	0.27	0.61	0.01
	81.6	7.92	0.009	0.22	0.43	0.52	0.01

Spatial Release From Masking

Linear regression analyses revealed that IPM-FR magnitude in the 81.6-Hz AM condition was significantly predictive of target-to-masker ratio thresholds in the separated listening condition [$F_{(1,28)}$ = 7.43, p = 0.01; R^2 = 0.21] and spatial release-from-masking thresholds [$F_{(1,28)}$ = 4.98, p = 0.03, R^2 = 0.15]. The combination of average hearing threshold and IPM-FR magnitude as fixed effects in the model accounted for even more variability in target-to-masker ratio thresholds in the separated listening condition [$F_{(2,27)}$ = 5.79, p = 0.008, adjusted R^2 = 0.25] than IPM-FR magnitude alone (**Figure 9**). However, further analysis revealed that IPM-FR magnitude was not a significant predictor of these target-to-masker ratio thresholds when participant age was included in the model. Similarly, IPM-FR magnitude did not explain any additional variance in spatial release-from-masking thresholds than the addition of participant age and average hearing thresholds alone [$F_{(2,27)}$ = 4.79, p = 0.02, adjusted R^2 = 0.21]. In other words, the significant relationships between IPM-FR magnitude and performance on these spatial release-from-masking tasks may be primarily mediated by the effects of age. No other IPM-FR

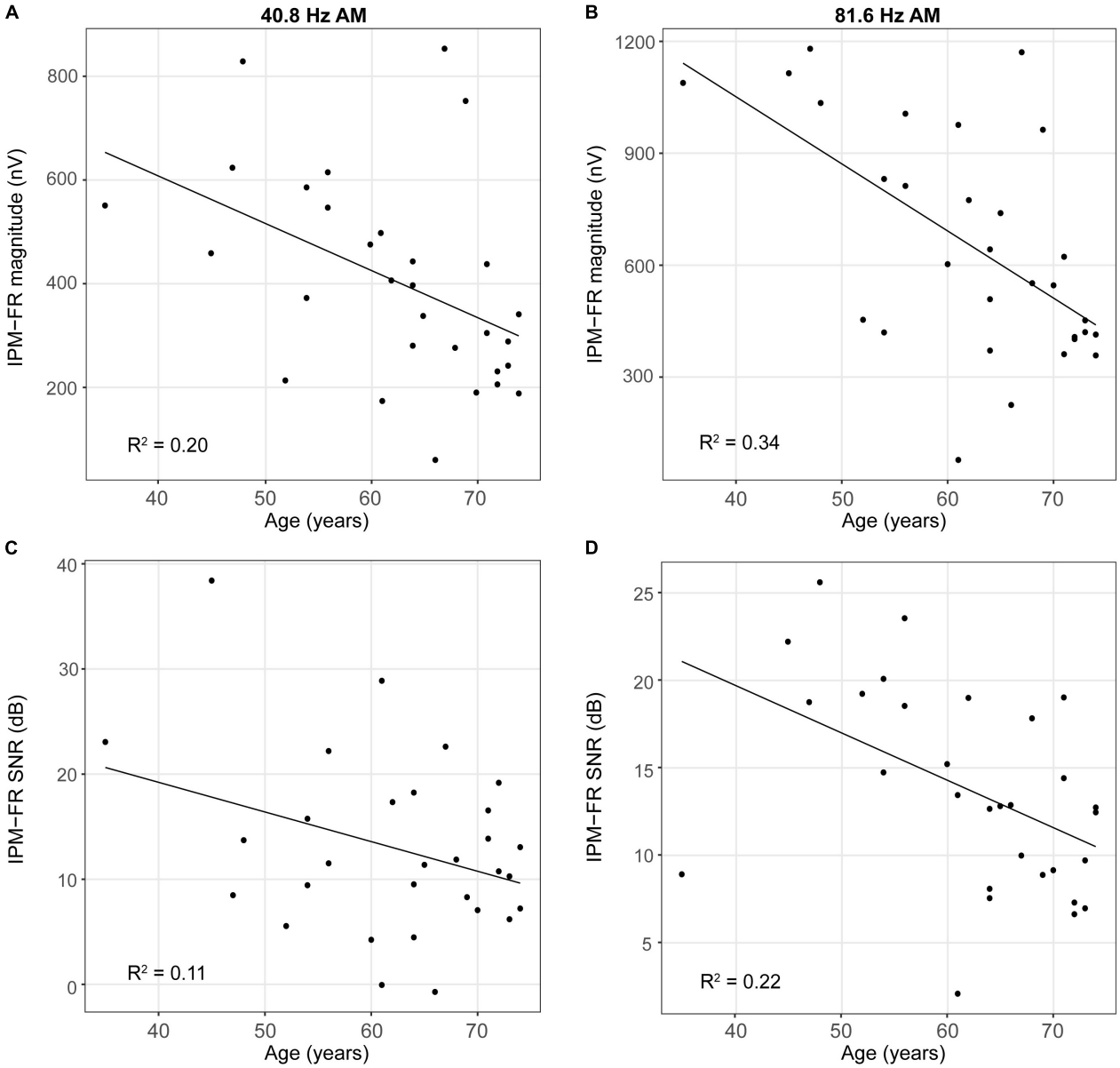

FIGURE 6 | Scatterplots depicting relationships between participant age and IPM-FR magnitude (panels **A,B**) and between participant age and IPM-FR SNR (panels **C,D**). Please note differences in the range of IPM-FR magnitudes displayed in Panels **(A)** and **(B)** and differences in the range of IPM-FR SNR values displayed in Panels **(C)** and **(D)**.

measures were significantly associated with performance on the spatial release-from-masking tasks.

Comparison of IPM-FRs by AM Rate

Mean neural response magnitudes and standard deviations for each electrophysiological measure are plotted in **Figure 10**. Paired *t*-tests were completed to determine whether IPM-FR magnitude and SNR were significantly different in response to different AM rates. Analysis showed that the magnitude of the IPM-FR was significantly larger in the 81.6-Hz AM rate condition compared to the 40.8-Hz AM rate condition [$t(1,29) = 7.60$, $p < 0.001$]. However, there was no significant difference in SNRs between the two stimulus conditions [$t(1,29) = 0.46$, $p = 0.65$]. In addition, there was no significant difference between neural responses in the diotic control condition at 6.8 Hz across the two AM rates [$t(1,29) = 1.65$, $p = 0.11$].

Comparison of ASSRs by Test Condition

Paired *t*-tests were also completed to determine whether ASSR magnitudes were significantly different across dichotic and diotic recording conditions. This analysis revealed that there was a significant effect of test condition (diotic vs. dichotic) on the

TABLE 2 | Test statistics from linear regression models that examined the effects of participant age and average hearing sensitivity across frequency on each behavioral measure.

Behavioral measure	Age			Hearing sensitivity		
	F	*p*	R^2	*F*	*p*	R^2
Binaural FM Detection Thresholds						
Dichotic Condition (log Hz)	2.03	0.16	0.07	0.07	0.79	0.00
Diotic Condition (log Hz)	4.40	0.04	0.14	0.03	0.87	0.00
Difference Score	1.83	0.19	0.06	0.00	0.99	0.00
Spatial Release-from-Masking Thresholds (dB)						
Separated Condition	8.44	0.007	0.23	4.79	0.04	0.15
Colocated Condition	0.01	0.93	0.00	0.01	0.91	0.00
Spatial Release from Masking	7.04	0.01	0.20	4.45	0.04	0.14

40.8-Hz ASSR such that ASSR magnitude was significantly higher in the dichotic condition compared to the diotic control condition [$t(1,29) = 2.63$, $p = 0.01$]. In contrast, the magnitude of the 81.6-Hz ASSR was not significantly impacted by test condition [$t(1,29) = -0.65$, $p = 0.52$].

DISCUSSION

The current study aimed to confirm the effects of age on the electrophysiological IPM-FR and behavioral measures of binaural temporal processing (binaural FM detection and spatial release from masking). In addition, this work was designed to determine whether age-related variability in the neural encoding of IPD cues, as measured by the IPM-FR, is related to performance on each behavioral task. Finally, the current study also aimed to explore potential differences in IPM-FRs measured using stimuli with different AM rates.

Effects of Age and Hearing Sensitivity on Binaural Processing

This study showed that age had a significant effect on IPM-FR magnitude and SNR. These findings are consistent with previous work (Vercammen et al., 2018) and provide additional evidence that age can impact the neural encoding of IPD cues (Ross et al., 2007a; Grose and Mamo, 2012a; Ozmeral et al., 2016; Papesh et al., 2017; Eddins and Eddins, 2018; Ungan et al., 2020). The IPM-FR is measured in response to IPD cues from temporal fine structure differences in an ongoing stimulus and is likely impacted by deficits in the extraction and integration of IPD information at the level of the brainstem. However, the IPM-FR is thought to be generated from neurons in the auditory cortex (Dajani and Picton, 2006; Undurraga et al., 2016). Therefore, it is likely that this response also reflects the cortical encoding of IPD cues. Although this work is unable to disentangle the potential effects of age on these subcortical and/or cortical processes, the current results do suggest that the IPM-FR represents a robust tool for the assessment of age-related declines in the neural processing of IPD cues.

Interestingly, while age did have a significant impact on diotic FM detection thresholds, it did not have a significant effect on dichotic FM detection thresholds or FM difference scores, which are thought to reflect IPD processing abilities. This is inconsistent with findings from Grose and Mamo (2012b), who showed that dichotic FM detection thresholds were better able differentiate between participant age groups than diotic FM detection thresholds. In addition, this previous work showed that younger and middle-aged participants were able to obtain more benefit from the addition of binaural information provided in the dichotic FM condition compared to older participants (Grose and Mamo, 2012b). Discrepancies between findings from the current study and those of Grose and Mamo (2012b) may be related to differences in stimulus modulation rates and durations. The current study used a 400-ms stimulus with a 6.8-Hz modulation frequency to match the IPM rate used in the electrophysiological measure, while Grose and Mamo (2012b) used a 1025-ms stimulus with a lower 2-Hz modulation frequency. The higher rate used in the current study resulted in more cycles of modulation per stimulus presentation over a shorter period of time compared to the stimulus used by Grose and Mamo (2012b). Increasing the number of modulation cycles has been shown to improve FM detection thresholds in both monaural (Hartmann and Klein, 1980; Wallaert et al., 2018; Palandrani et al., 2020) and dichotic listening conditions (Palandrani et al., 2020). Grose and Mamo (2012b) reported mean dichotic FM detection thresholds of approximately 2 Hz for their group of older (65–77 years) listeners and approximately 0.8 Hz for their group of middle aged (43–57 years) listeners, for an estimated average threshold across groups of approximately 1.4 Hz. This estimated threshold is slightly better than the mean dichotic FM detection threshold of 1.6 Hz in the current study, which tested a comparable group of listeners who ranged in age from 35–74 years. While it appears as if participants from Grose and Mamo (2012b) and those tested in the current study performed similarly on this task, differences in modulation rates used across studies makes these threshold comparisons difficult. Instead, as shown in Witton et al. (2000), thresholds can be converted to ITDs for a given modulation depth (or FM detection threshold), modulation rate, and center frequency to directly compare performance across studies. When compared in this way, the 1.4-Hz detection threshold from Grose and Mamo (2012b) corresponds to a maximum difference ITD of approximately 891.3 μs, while the 1.6-Hz detection threshold found in the current study corresponds to a maximum difference ITD of 299.6 μs, representing better performance on this task. It is possible that increasing the modulation rate reduces the difficulty of the dichotic FM detection task and consequently impacts this measure's sensitivity to the effects of age on binaural processing, giving rise to the discrepancy between the current results and those of Grose and Mamo (2012b). Future research should focus on further exploring the effects of age and modulation rate on IPD sensitivity using this dichotic FM detection task.

Participants in the current study were required to have hearing thresholds within normal limits (i.e., ≤25 dB HL) at 500 Hz. Variations in hearing sensitivity at higher frequencies were not expected to have any impact on the binaural FM measures or the IPM-FR, as a 500 Hz carrier tone was used as a stimulus for each of these measures. However, variability in higher frequency

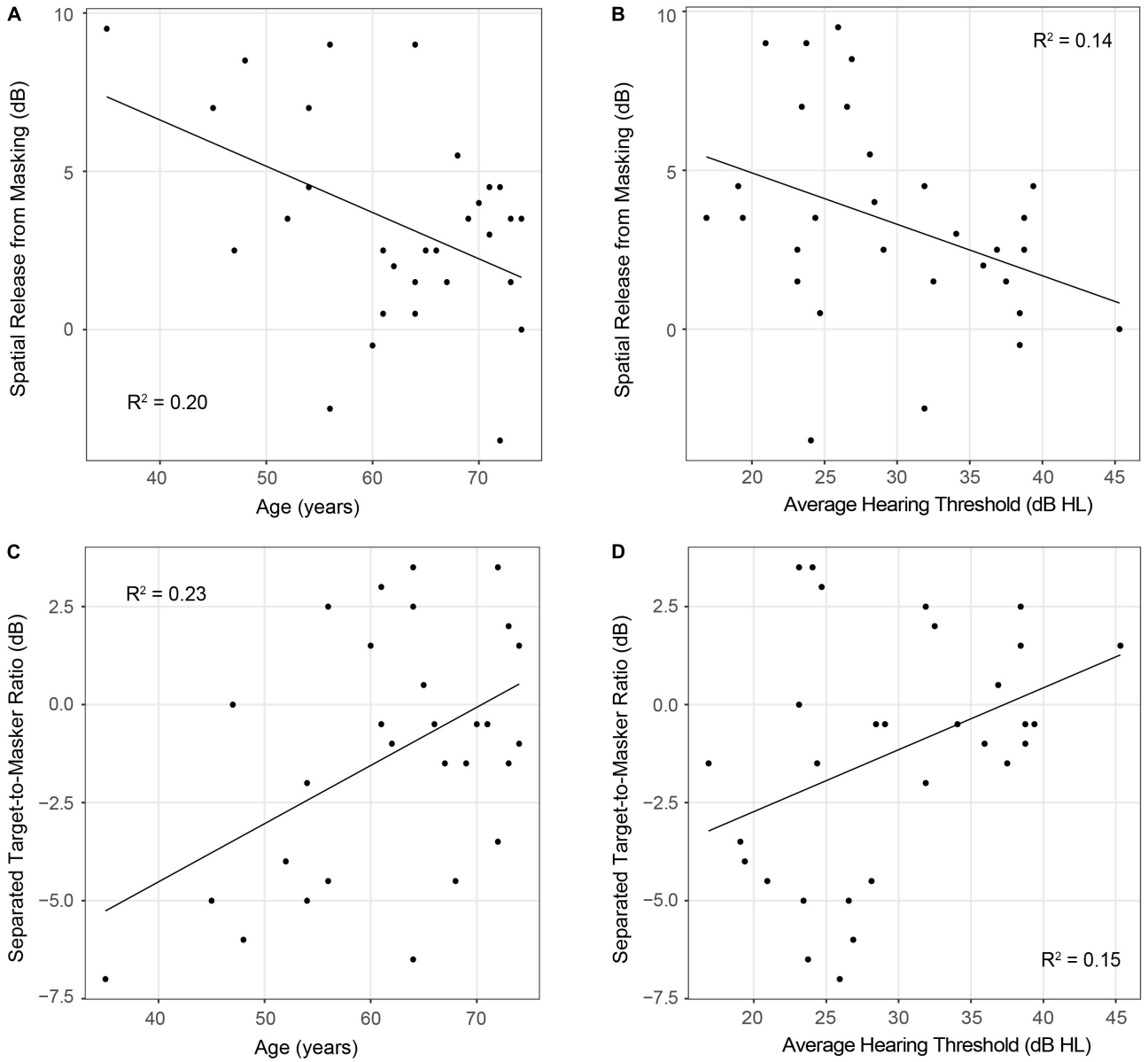

FIGURE 7 | Scatterplots depicting relationships between participant age (panels **A,C**) or average hearing thresholds (panels **B,D**) and performance on spatial release-from-masking measures. Note that lower spatial release-from-masking thresholds represent poorer performance on this measure and indicate that participants had smaller differences in performance between the colocated and separated speech-on-speech masking conditions. On the other hand, in the spatially separated speech-on-speech masking condition, lower thresholds represent better performance and indicate that participants were able to identify the target sentence at more difficult target-to-masker ratios than participants with higher thresholds.

TABLE 3 | Test statistics from linear regression models that examined relationships between each IPM-FR measure and each binaural FM measure.

		Dichotic FM (log Hz)			Diotic FM (log Hz)			FM Difference Score		
		F	*p*	R^2	*F*	*p*	R^2	*F*	*p*	R^2
IPM-FR Magnitude (nV)	40.8	1.91	0.18	0.06	4.12	0.05	0.13	2.64	0.11	0.09
	81.6	3.72	0.06	0.12	4.05	0.05	0.13	3.11	0.09	0.10
SNR (dB)	40.8	1.79	0.20	0.06	1.86	0.18	0.06	2.06	0.16	0.07
	81.6	16.55	<0.001	0.37	4.34*	0.05	0.13	18.76	<0.001	0.40

**Denotes relationships that were no longer statistically significant after accounting for the effects of age or average hearing sensitivity on behavioral performance.*

TABLE 4 | Test statistics from linear regression models that examined relationships between each IPM-FR measure and target-to-masker ratio thresholds in the separated and colocated speech-on-speech masking conditions as well as spatial release-from-masking thresholds.

		Separated Condition (dB)			Colocated Condition (dB)			Spatial Release from Masking (dB)		
		F	*p*	R^2	*F*	*p*	R^2	*F*	*p*	R^2
IPM-FR Magnitude (nV)	40.8	2.90	0.10	0.09	0.36	0.55	0.01	1.76	0.19	0.06
	81.6	7.43*	0.01	0.21	0.33	0.57	0.01	4.98*	0.03	0.15
SNR (dB)	40.8	1.74	0.20	0.06	0.80	0.38	0.03	0.73	0.40	0.03
	81.6	2.73	0.11	0.09	1.36	0.25	0.05	1.08	0.31	0.04

**Denotes relationships that were no longer statistically significant after accounting for the effects of age or average hearing sensitivity on behavioral performance.*

hearing thresholds may indicate differences in the overall health of the auditory system that could impact auditory processing, including the encoding and detection of low-frequency IPD cues. In order to account for this possibility, hearing sensitivity was characterized as the average of hearing thresholds from 250 to 8000 Hz across ears. This metric accounts for variability in low-frequency hearing thresholds as well as variability in high-frequency hearing thresholds across participants. The current results did not show any effect of hearing sensitivity on the IPM-FR or the binaural FM detection measures. Previous studies have not shown significant effects of high-frequency hearing sensitivity on low-frequency IPD discrimination using a variety of behavioral measures (Strelcyk and Dau, 2009; Grose and Mamo, 2012b; Moore et al., 2012; Eddins and Eddins, 2018), which is consistent with the current results. However, Vercammen et al. (2018) did show that hearing sensitivity impacted the neural encoding of IPDs as measured by IPM-FRs elicited by a 492-Hz carrier tone, even after presentation levels were adjusted for audibility. Unlike the current study that treated age as a continuous variable, Vercammen et al. (2018) separated participants into younger-, middle-, and older-aged normal hearing and hearing impaired participant groups, which differed in average hearing thresholds at 500 Hz. This may have resulted in different levels of stimulus audibility and therefore different presentation levels across these participant groups, which may have contributed to the main effect of hearing sensitivity reported by Vercammen et al. (2018) that was not observed in the current study.

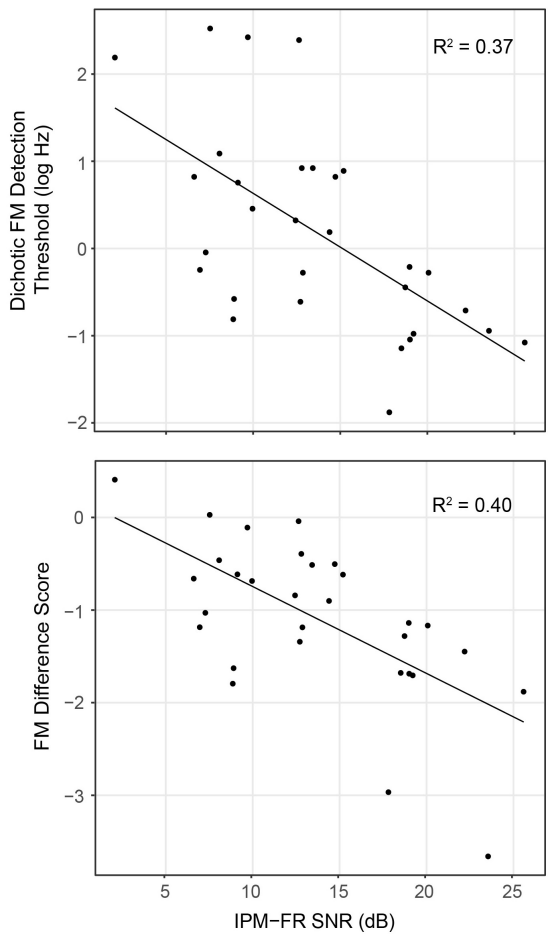

FIGURE 8 | Scatterplots depicting relationships between the IPM-FR SNR recorded in response to the 81.6-Hz AM stimulus for each participant and individual dichotic FM detection thresholds (top panel) and FM difference score estimates (bottom panel). Note that lower dichotic FM thresholds represent better performance on this task, while lower (more negative) FM difference scores indicate a greater difference between diotic and dichotic FM thresholds, representing a larger benefit from the addition of binaural information in this task.

Even though variability in high-frequency hearing thresholds was not expected to impact low-frequency IPD sensitivity, results showed that hearing sensitivity did have a significant effect on spatial release-from-masking thresholds and performance in the spatially separated speech-on-speech listening condition. It is likely that poorer high-frequency hearing sensitivity impacted the audibility of certain speech cues necessary for this behavioral task. Results showed that age also had a significant effect on these measures and had a stronger relationship with speech understanding than average hearing sensitivity. This is consistent with previous findings that have shown that both age and hearing sensitivity can independently impact performance on these spatial release-from-masking measures (Gallun et al., 2014; Papesh et al., 2017; Jakien and Gallun, 2018).

Since the ASSR reflects phase locking to the temporal envelope of the amplitude modulated stimulus, this electrophysiological measure can represent an index of temporal processing abilities. The current study showed that age had a significant impact on the 40.8 Hz ASSR recorded in each test condition, but did not significantly impact the 81.6 Hz ASSR. This finding suggests that in addition to impacting binaural temporal fine

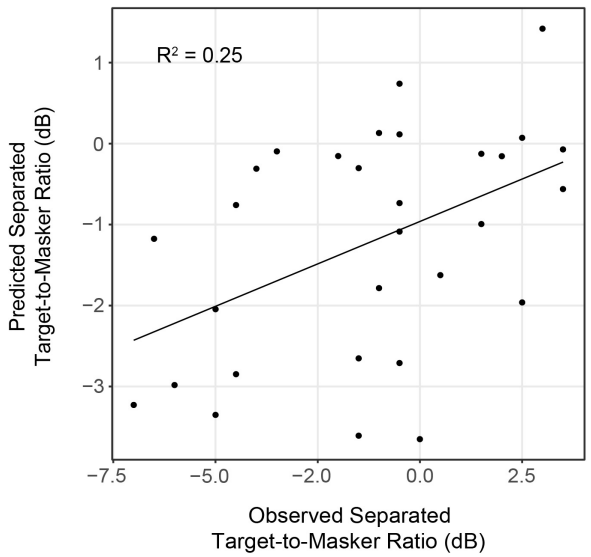

FIGURE 9 | Observed target-to-masker ratio thresholds in the spatially separated listening condition plotted as a function of predicted thresholds from a linear regression model that included average hearing threshold and IPM-FR magnitude to the 81.6-Hz AM stimulus condition as fixed effects.

structure processing, age may also affect temporal envelope processing abilities. This finding is consistent with results from Ungan et al. (2020), who used a binaural beat stimulus and an amplitude modulated stimulus to examine the effects of age on the neural encoding of binaural temporal fine structure information and on the neural encoding of temporal envelope information, respectively. As discussed previously, it is thought that higher amplitude modulation rates elicit ASSRs from brainstem structures, while lower amplitude modulation rates elicit ASSRs that may be generated by contributions from overlapping brainstem as well as cortical structures (Giraud et al., 2000; Herdman et al., 2002; Korczak et al., 2012). Therefore, the current results may reflect age-related declines in the cortical processing of temporal envelope information that is not reflected at the level of the brainstem. However, this pattern of results is inconsistent with previous literature that has not shown any significant effects of age on 40 Hz ASSRs (Johnson et al., 1988; Boettcher et al., 2001; Picton et al., 2003; Ross, 2008). In fact, several studies have actually shown that age tends to have a greater impact on ASSRs elicited by higher AM rates compared to lower AM rates (Leigh-Paffenroth and Fowler, 2006; Grose et al., 2009; Goossens et al., 2016). It is unknown why the current ASSR results followed an opposite pattern.

Relationships Between Neural Responses and Behavioral Performance

Relationships between binaural FM thresholds and the IPM-FR were examined in order to better understand whether this electrophysiological measure is reflective of neural processes underlying behavioral measures of IPD sensitivity. Results showed that the IPM-FR SNR in the 81.6-Hz AM condition

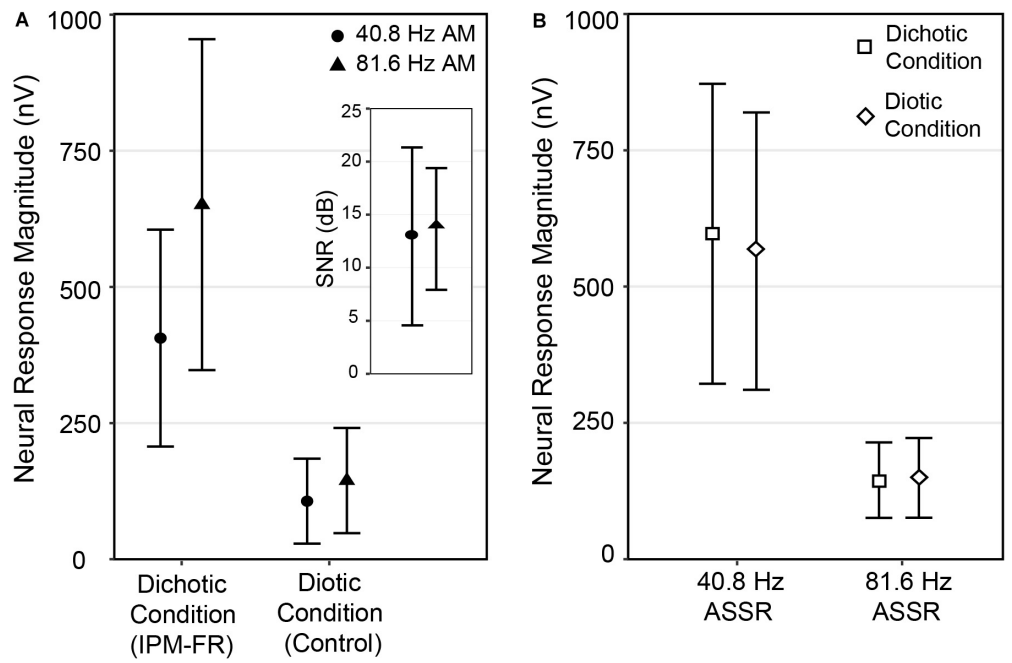

FIGURE 10 | Mean neural response magnitudes in the dichotic (IPM-FR) and diotic (control) conditions for the 40.8-Hz AM (filled circles) and the 81.6 Hz AM (filled triangles) stimuli (panel **A**) averaged across all 30 participants at electrode M2. The inset figure shows mean IPM-FR SNRs for the 40.8-Hz AM stimulus (filled circles) and the 81.6-Hz AM stimulus (filled triangles) calculated using individual neural responses to the dichotic and diotic stimuli. Mean ASSR magnitudes recorded in the dichotic (open square) and diotic (open diamond) stimulus are also provided (panel **B**). Error bars represent 1 standard deviation from the mean.

was associated with dichotic FM detection thresholds as well as the FM difference score that estimated benefit received from the addition of binaural information in the dichotic FM compared to the diotic FM task. The IPM-FR was able to account for 37% and 40% of the variability in these binaural FM detection measures, respectively. Unlike the IPM-FR, these FM detection tasks were not sensitive to the effects of age in the current study. However, the relationship between the IPM-FR and these binaural FM measures suggests that age-related variability the neural encoding of IPDs may be reflected by performance on these behavioral tasks. Previous studies that established links between the IPM-FR and behavioral IPD sensitivity used stimuli analogous to those used to the elicit the IPM-FR (Haywood et al., 2015; Undurraga et al., 2016; Vercammen et al., 2018). This work expanded on this existing literature by providing evidence that the IPM-FR is also reflective of individual variability in the processing of IPDs produced by dichotic FM.

Speech understanding abilities were assessed using measures of spatial release from masking. The current results showed that IPM-FR magnitude in response to the higher AM stimulus was significantly associated with target-to-masker ratio thresholds in the spatially separated speech-on-speech masking task. Linear regression model predictions of individual performance in this condition were further improved by including average hearing threshold estimates as an additional predictor in the model, such that the model was able to account for approximately 25% of the variance in performance on this task. However, this was not the case when participant age was added as an additional predictor in the model. Similarly, analyses revealed that the relationship between IPM-FR magnitude and spatial release-from-masking thresholds were also likely mediated by the effects of age. In other words, while the variability in the neural encoding of IPD cues was associated with performance on these behavioral speech perception measures, it was not able to account for a substantial amount of variability in performance over what was already accounted for by participant age.

These results are inconsistent with those from Papesh et al. (2017), who assessed relationships between the same spatial release-from-masking measures and the neural sensitivity to changes in IPDs measured using the acoustic change complex. In that study, neural responses were better predictors of spatial release-from-masking thresholds and target-to-masker ratio thresholds in the spatially separated listening condition than participant age or hearing sensitivity (Papesh et al., 2017). Although the IPM-FR and this acoustic change complex both reflect the neural encoding of IPD cues embedded within the temporal fine structure of an AM stimulus, it is possible that differences in the nature of each electrophysiological response may have contributed to these conflicting results. The periodic ±90° IPM used in the current study created a percept of the signal moving from one side to the other. As discussed by Haywood et al. (2015) and Undurraga et al. (2016), these periodic shifts between IPDs leading in the right and left ears in the ongoing stimulus result in modulation of activity in the right and left hemispheres. In contrast, the stimulus used by Papesh et al. (2017) consisted of a single phase shift from a zero to a completely anti-phasic 180° IPD. This stimulus would not periodically modulate the activity of right and left brain hemispheres and would result in a more diffuse intracranial stimulus percept. In addition, Undurraga et al. (2016) argued that a stimulus with a 180° IPD shift like that used by Papesh et al. (2017) may activate neurons in the lateral superior olivary complex of the brainstem that are responsible for processing interaural level differences. These neurons are less likely to be activated by the IPM-FR stimulus used in the current study, which is primarily thought to reflect activity of medial superior olivary complex neurons that are sensitive to IPDs (Undurraga et al., 2016). Therefore, it is also possible that differences in neural activation patterns may have contributed to discrepancies in results between the current study and those of Papesh et al. (2017). Finally, differences in the distribution of participant age as well as the higher 750-Hz carrier frequency used by Papesh et al. (2017) cannot be ruled out as additional factors that may have had an effect on results between these studies. Future research should focus on examining how neurophysiological links between the IPM-FR and speech understanding in spatialized noise are impacted by these factors.

Effects of Stimulus Parameters on Electrophysiological Responses

Effects of AM Rate on the IPM-FR

The current results suggest that changes in stimulus AM rates can impact the IPM-FR. These findings have implications for how stimulus parameters may be optimized to improve IPM-FR measurement reliability as well as to increase sensitivity to participant factors that may contribute to hearing difficulties. The lower AM rate used in the current study is similar to that used in initial studies on the IPM-FR (Haywood et al., 2015; McAlpine et al., 2016; Undurraga et al., 2016), and the higher AM rate is similar to that used in a recent study by Vercammen et al. (2018). However, the current work is the first to directly compare the effects of AM rate on IPM-FRs across the same individuals. Results showed that IPM-FR magnitude was larger in response to the stimulus that was amplitude modulated at 81.6 Hz compared to the stimulus that was amplitude modulated at 40.8 Hz. In addition, while SNRs were not significantly different across AM rate conditions, SNRs calculated from IPM-FRs elicited using the higher 81.6-Hz AM rate tended to be better predictors of behavior than those elicited using the lower 40.8-Hz AM rate. IPM-FRs elicited with the higher AM rate also tended to be more sensitive to the effects of age than those elicited with the lower AM rate. There are several potential explanations for our observed pattern of results. First, the use of a higher AM rate may improve the magnitude of the IPM-FR because it simply contains more AM cycles in the ongoing stimulus than a lower AM rate. This may create more neural responses, or looks, at the ongoing stimulus, which would be expected to result in an increase in neural response strength. A second explanation may be that a steeper modulation slope resulting from the faster 81.6-Hz AM increases neural synchrony, and therefore increases response strength compared to shallower slopes that would occur at lower AM rates. An additional explanation may

be that the stronger IPM-FR in the higher AM rate condition results from additional neural responses to energy contained in sidebands that result from the AM of the 500-Hz signal. If these sidebands occur in separate auditory filters and contain IPMs, then participants may essentially be benefiting by an increased number of available stimuli that each contain IPD cues. Finally, while it is known that changes in AM rate impact the activation of ASSR neural generator sites (Giraud et al., 2000; Herdman et al., 2002; Korczak et al., 2012), it is difficult to determine how changes in the activation of these different neural generators with AM rate may also impact the IPM-FR. Future work will attempt to further examine and test these potential explanations to better understand the effects of AM rate on the IPM-FR.

Effects of Recording Condition on the ASSR

Auditory steady-state responses were compared to assess recording quality between the dichotic test condition and the diotic control condition that alternated within each recording block. This data quality check is important for the current study because neural responses to the diotic control stimulus were used to calculate IPM-FR SNRs, and any systematic contamination of responses to a particular stimulus would compromise SNR estimations. Results from the current work showed that the 40.8-Hz ASSR was larger in the dichotic test condition compared to the diotic control condition. This was an unexpected finding, given that the only difference between the two stimuli was the addition of periodic IPMs in the temporal fine structure of the ongoing AM stimulus in the dichotic test condition. One possibility is that the higher 40.8-Hz ASSR magnitude observed in the dichotic condition is reflecting the presence of harmonics in the neural response to the IPMs in the stimulus, and do not reflect actual changes in the ASSR. As can be seen in **Figure 5A**, neural response peaks can be observed at multiples of the 6.8-Hz IPM rate. While these harmonics are reduced in amplitude as frequency increases, it is possible that the 6th harmonic, which would be equivalent to 40.8 Hz, is contributing to the magnitude of the ASSR measured at this frequency. Therefore, the specific stimulus parameters used in the current study may preclude this type of data quality check for the lower amplitude modulated stimulus. This issue is less likely to occur in the higher AM rate condition, as response harmonics that high in frequency are expected to be negligible, as can be observed in **Figure 5B**. Indeed, 81.6-Hz ASSR magnitudes were not significantly different across the diotic and dichotic stimuli, which suggests that recording quality was comparable across these two conditions.

CONCLUSION

The current work confirmed that the IPM-FR is sensitive to the effects of age on the neural encoding of IPD cues. In addition, this study verified that the IPM-FR is reflective of neural processes underlying behavioral IPD discrimination using tests of binaural FM sensitivity. Therefore, these results confirm that the IPM-FR represents a robust tool for the objective assessment of IPD sensitivity. However, further work is required to better understand links between the neural encoding of IPD cues as measured by the IPM-FR and behavioral measures of binaural temporal processing, especially those that assess speech understanding abilities. In addition, future research should continue to investigate the effects of different stimulus parameters on neural and behavioral measures of IPD sensitivity to better understand the effects of age on these responses. The continued development of measures that are sensitive to participant factors that are thought to impact binaural temporal processing and that are also reflective of functional auditory abilities will be integral to the clinical identification and management of auditory difficulties, especially in patients with normal hearing sensitivity.

DATA AVAILABILITY STATEMENT

The data supporting the conclusions of this article may be made available by the authors upon request.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the joint institutional review board of the Department of Veterans Affairs Portland Health Care System and Oregon Health & Science University. The participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

CB and FG conceived the study. CB, FG, TK, and RM designed the experiments. TK and RM collected and analyzed the data. CB, FG, TK, and RM interpreted the results. TK drafted the manuscript and revisions. CB, FG, and RM provided critical reviews of the manuscript and approved the final version. All authors contributed to the article and approved the submitted version.

FUNDING

This project was supported by funding from R01 DC015240 to CB and the VA Advanced Fellowship in Polytrauma/Traumatic Brain Injury Rehabilitation-Research to TK. Development of Portable Automated Rapid Testing (PART) is supported by R01 DC015051.

ACKNOWLEDGMENTS

Special thanks to Jaime Undurraga, David McAlpine, and Leslie Grush for assistance with stimulus design. Additional thanks to Alyse Gulack and Jay Vachhani for assistance with participant recruitment and data collection and to Melissa Papesh for input on experimental design and data interpretation.

REFERENCES

Boettcher, F. A., Poth, E. A., Mills, J. H., and Dubno, J. R. (2001). The amplitude modulation following response in young and aged human subjects. *Hear. Res.* 153, 32–42. doi: 10.1016/S0378-5955(00)00255-0

Bolia, R. S., Nelson, W. T., Ericson, M. A., and Simpson, B. D. (2000). A speech corpus for multitalker communications research. *J. Acoust. Soc. Am.* 107, 1065–1066.

Coffey, E. B. J., Nicol, T., White-Schwoch, T., Chandrasekaran, B., Krizman, J., Skoe, E., et al. (2019). Evolving perspectives on the sources of the frequency-following response. *Nat. Commun.* 10:5036. doi: 10.1038/s41467-019-13003-w

Dajani, H. R., and Picton, T. W. (2006). Human auditory steady-state responses to changes in interaural correlation. *Hear. Res.* 219, 85–100. doi: 10.1016/j.heares.2006.06.003

Dimitrijevic, A., John, M. S., Van Roon, P., Purcell, D., Adamonis, J., Ostroff, J., et al. (2002). Estimating the audiogram using multiple auditory steady-state responses. *J. Am. Acad. Audiol.* 13, 205–224.

Eddins, A. C., and Eddins, D. A. (2018). Cortical correlates of binaural temporal processing deficits in older adults. *Ear Hear.* 39, 594–604. doi: 10.1097/AUD.0000000000000518

Füllgrabe, C., and Moore, B. C. J. (2017). Evaluation of a method for determining binaural sensitivity to temporal fine structure (TFS-AF Test) for older listeners with normal and impaired low-frequency hearing. *Trends Hear.* 21, 1–14. doi: 10.1177/2331216517737230

Füllgrabe, C., and Moore, B. C. J. (2018). The association between the processing of binaural temporal-fine-structure information and audiometric threshold and age: a meta-analysis. *Trends Hear.* 22, 1–14. doi: 10.1177/2331216518797259

Füllgrabe, C., Moore, B. C. J., and Stone, M. A. (2015). Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition. *Front. Aging Neurosci.* 6:347. doi: 10.3389/fnagi.2014.00347

Füllgrabe, C., Sęk, A. P., and Moore, B. C. J. (2018). Senescent changes in sensitivity to binaural temporal fine structure. *Trends Hear.* 22, 1–16. doi: 10.1177/2331216518788224

Gallun, F. J., Diedesch, A. C., Kampel, S. D., and Jakien, K. M. (2013). Independent impacts of age and hearing loss on spatial release in a complex auditory environment. *Front. Neurosci.* 7:252. doi: 10.3389/fnins.2013.00252

Gallun, F. J., Mason, C. R., and Kidd, G. (2005). Binaural release from informational masking: results from a speech identification task. *J. Acoust. Soc. Am.* 118, 1614–1625. doi: 10.1121/1.4782307

Gallun, F. J., McMillan, G. P., Molis, M. R., Kampel, S. D., Dann, S. M., and Konrad-Martin, D. L. (2014). Relating age and hearing loss to monaural, bilateral, and binaural temporal sensitivity. *Front. Neurosci.* 8:172. doi: 10.3389/fnins.2014.00172

Gallun, F. J., Seitz, A., Eddins, D. A., Molis, M. R., Stavropoulos, T., Jakien, K. M., et al. (2018). Development and validation of Portable Automated Rapid Testing (PART) measures for auditory research. *Proc. Meet. Acoust.* 33:050002. doi: 10.1121/2.0000878

Giraud, A.-L., Lorenzi, C., Ashburner, J., Wable, J., Johnsrude, I., Frackowiak, R., et al. (2000). Representation of the temporal envelope of sounds in the human brain. *J. Neurophysiol.* 84, 1588–1598. doi: 10.1152/jn.2000.84.3.1588

Goossens, T., Vercammen, C., Wouters, J., and van Wieringen, A. (2016). Aging affects neural synchronization to speech-related acoustic modulations. *Front. Aging Neurosci.* 8:133. doi: 10.3389/fnagi.2016.00133

Grose, J. H., and Mamo, S. K. (2010). Processing of temporal fine structure as a function of age. *Ear Hear.* 31, 755–760. doi: 10.1097/AUD.0b013e3181e627e7

Grose, J. H., and Mamo, S. K. (2012a). Electrophysiological measurement of binaural beats: effects of primary tone frequency and observer age. *Ear Hear.* 33, 187–194. doi: 10.1097/AUD.0b013e318230bbbd

Grose, J. H., and Mamo, S. K. (2012b). Frequency modulation detection as a measure of temporal processing: age-related monaural and binaural effects. *Hear. Res.* 294, 49–54. doi: 10.1016/j.heares.2012.09.007

Grose, J. H., Mamo, S. K., and Hall, J. W. (2009). Age effects in temporal envelope processing: speech unmasking and auditory steady state responses. *Ear Hear.* 30, 568–575. doi: 10.1097/AUD.0b013e3181ac128f

Hartmann, W. M., and Klein, M. A. (1980). Theory of frequency modulation detection for low modulation frequencies. *J. Acoust. Soc. Am.* 67, 935–946. doi: 10.1121/1.383972

Hawley, M. L., Litovsky, R. Y., and Culling, J.F. (2004). The benefit of binaural hearing in a cocktail party: effect of location and type of interferer. *J. Acoust. Soc. Am.* 115, 833–843.

Haywood, N. R., Undurraga, J. A., Marquardt, T., and McAlpine, D. (2015). A comparison of two objective measures of binaural processing: the interaural phase modulation following response and the binaural interaction component. *Trends Hear.* 19, 1–17. doi: 10.1177/2331216515619039

He, N., Mills, J. H., Ahlstrom, J. B., and Dubno, J. R. (2008). Age-related differences in the temporal modulation transfer function with pure-tone carriers. *J. Acoust. Soc. Am.* 124, 3841–3849. doi: 10.1121/1.2998779

Herdman, A. T., Lins, O., Roon, P. V., Stapells, D. R., Scherg, M., and Picton, T. W. (2002). Intracerebral sources of human auditory steady-state responses. *Brain Topogr* 15, 69–86.

Hoover, E. C., Kinney, B. N., Bell, K. L., Gallun, F. J., and Eddins, D. A. (2019). A comparison of behavioral methods for indexing the auditory processing of temporal fine structure cues. *J. Speech Lang. Hear. Res.* 62, 2018–2034. doi: 10.1044/2019_JSLHR-H-18-0217

Hopkins, K., and Moore, B. C. J. (2011). The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. *J. Acoust. Soc. Am.* 130, 334–349. doi: 10.1121/1.3585848

Jakien, K. M., and Gallun, F. J. (2018). Normative data for a rapid, automated test of spatial release from masking. *Am. J. Audiol.* 27, 529–538. doi: 10.1044/2018_AJA-17-0069

Jakien, K. M., Kampel, S. D., Stansell, M. M., and Gallun, F. J. (2017). Validating a rapid, automated test of spatial release from masking. *Am. J. Audiol.* 26, 507–518. doi: 10.1044/2017_AJA-17-0013

Johnson, B. W., Weinberg, H., Ribary, U., Cheyne, D. O., and Ancill, R. (1988). Topographic distribution of the 40 Hz auditory evoked-related potential in normal and aged subjects. *Brain Topogr.* 1, 117–121. doi: 10.1007/BF01129176

King, A., Hopkins, K., and Plack, C. J. (2014). The effects of age and hearing loss on interaural phase difference discrimination. *J. Acoust. Soc. Am.* 135, 342–351.

Korczak, P., Smart, J., Delgado, R., Strobel, T. M., and Bradford, C. (2012). Auditory steady-state responses. *J. Am. Acad. Audiol.* 23, 146–170. doi: 10.3766/jaaa.23.3.3

Leigh-Paffenroth, E. D., and Fowler, C. G. (2006). Amplitude-modulated auditory steady-state responses in younger and older listeners. *J. Am. Acad. Audiol.* 17, 582–597. doi: 10.3766/jaaa.17.8.5

Lelo de Larrea-Mancera, E. S., Stavropoulos, T., Hoover, E. C., Eddins, D. A., Gallun, F. J., and Seitz, A. R. (2020). Portable Automated Rapid Testing (PART) for auditory research: validation in a normal hearing population. *J. Acoust. Soc. Am.* 148, 1831–1851. doi: 10.1121/10.0002108

Levi, E. C., Folsom, R. C., and Dobie, R. A. (1993). Amplitude-modulation following response (AMFR): effects of modulation rate, carrier frequency, age, and state. *Hear, Res.* 68, 42–52. doi: 10.1016/0378-5955(93)90063-7

McAlpine, D., Haywood, N., Undurraga, J., and Marquardt, T. (2016). "Objective measures of neural processing of interaural time differences," in *Physiology, Psychoacoustics and Cognition in normal and impaired hearing*, eds P. van Dijk, D. Baskent, E. Gaudrain, E. de Kleine, A. Wagner, and C. Lanting (New York: Springer International Publishing), 197–205

Moore, B. C. J., Glasberg, B. R., Stoev, M., Füllgrabe, C., and Hopkins, K. (2012). The influence of age and high-frequency hearing loss on sensitivity to temporal fine structure at low frequencies (L). *J. Acoust. Soc. Am.* 131, 1003–1006. doi: 10.1121/1.3672808

Ozmeral, E. J., Eddins, D. A., and Eddins, A. C. (2016). Reduced temporal processing in older, normal-hearing listeners evident from electrophysiological responses to shifts in interaural time difference. *J. Neurophysiol.* 116, 2720–2729. doi: 10.1152/jn.00560.2016

Palandrani, K. N., Hoover, E., Stavropoulos, T., Seitz, A. R., Isarangura, S., Gallun, F. J., et al. (2020). Temporal integration of monaural and dichotic frequency modulation. *PsyArXiv* [Preprint]. doi: 10.31234/osf.io/269gp

Papesh, M. A., Folmer, R. L., and Gallun, F. J. (2017). Cortical measures of binaural processing predict spatial release from masking performance. *Front. Hum, Neurosci.* 11:124. doi: 10.3389/fnhum.2017.00124

Parthasarathy, A., Hancock, K. E., Bennett, K., DeGruttola, V., and Polley, D. B. (2020). Bottom-up and top-down neural signatures of disordered multi-talker speech perception in adults with normal hearing. *eLife* 9:e51419. doi: 10.7554/eLife.51419

Parthasarathy, A., and Kujawa, S. G. (2018). Synaptopathy in the aging cochlea: characterizing early-neural deficits in auditory temporal envelope processing. *J. Neurosci.* 38, 7108–7119. doi: 10.1523/JNEUROSCI.3240-17.2018

Picton, T. W., John, M. S., Dimitrijevic, A., and Purcell, D. (2003). Human auditory steady-state responses. *Int. J. Audiol.* 42, 177–219. doi: 10.3109/14992020309101316

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., and R Core Team. (2016). *nlme: Linear and Nonlinear Mixed Effects Models*.

Ross, B. (2008). A novel type of auditory responses: temporal dynamics of 40-Hz steady-state responses induced by changes in sound localization. *J. Neurophysiol.* 100, 1265–1277. doi: 10.1152/jn.00048.2008

Ross, B., Fujioka, T., Tremblay, K. L., and Picton, T. W. (2007a). Aging in binaural hearing begins in mid-life: evidence from cortical auditory-evoked responses to changes in interaural phase. *J. Neurosci.* 27, 11172–11178. doi: 10.1523/JNEUROSCI.1813-07.2007

Ross, B., Tremblay, K. L., and Picton, T. W. (2007b). Physiological detection of interaural phase differences. *J. Acoust. Soc. Am.* 121, 1017–1027. doi: 10.1121/1.2404915

Ruggles, D., Bharadwaj, H., and Shinn-Cunningham, B. G. (2012). Why middle-aged listeners have trouble hearing in everyday settings. *Curr. Biol.* 22, 1417–1422. doi: 10.1016/j.cub.2012.05.025

Shaheen, L. A., Valero, M. D., and Liberman, M. C. (2015). Towards a diagnosis of cochlear neuropathy with envelope following responses. *J. Assoc. Res. Otolaryngol.* 16, 727–745. doi: 10.1007/s10162-015-0539-3

Stecker, G. C., and Gallun, F. J. (2012). *Binaural Hearing, Sound Localization, and Spatial Hearing*. San Diego, CA: Plural Publishing, Inc.

Strelcyk, O., and Dau, T. (2009). Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing. *J. Acoust. Soc. Am.* 125, 3328–3345. doi: 10.1121/1.3097469

Undurraga, J. A., Haywood, N. R., Marquardt, T., and McAlpine, D. (2016). Neural representation of interaural time differences in humans—An objective measure that matches behavioural performance. *J. Assoc. Res. Otolaryngol.* 17, 591–607. doi: 10.1007/s10162-016-0584-6

Ungan, P., Yagcioglu, S., and Ayik, E. (2020). Effects of aging on event-related potentials to single-cycle binaural beats and diotic amplitude modulation of a tone. *Brain Res.* 1740:46849. doi: 10.1016/j.brainres.2020.146849

Vercammen, C., Goossens, T., Undurraga, J., Wouters, J., and van Wieringen, A. (2018). Electrophysiological and behavioral evidence of reduced binaural temporal processing in the aging and hearing impaired human auditory system. *Trends Hear.* 22, 1–12. doi: 10.1177/2331216518785733

Wallaert, N., Varnet, L., Moore, B. C. J., and Lorenzi, C. (2018). Sensorineural hearing loss impairs sensitivity but spares temporal integration for detection of frequency modulation. *J. Acoust. Soc. Am.* 144, 720–733. doi: 10.1121/1.5049364

Wambacq, I. J. A., Koehnke, J., Besing, J., Romei, L. L., DePierro, A., and Cooper, D. (2009). Processing interaural cues in sound segregation by young and middle-aged brains. *J. Am. Acad. Audiol.* 20, 453–458. doi: 10.3766/jaaa.20.7.6

Whiteford, K. L., Kreft, H. A., and Oxenham, A. J. (2017). Assessing the role of place and timing cues in coding frequency and amplitude modulation as a function of age. *J. Assoc. Res. Otolaryngol.* 18, 619–633. doi: 10.1007/s10162-017-0624-x

Whiteford, K. L., and Oxenham, A. J. (2015). Using individual differences to test the role of temporal and place cues in coding frequency modulation. *J. Acoust. Soc. Am.* 138, 3093–3104. doi: 10.1121/1.4935018

Witton, C., Green, G. G. R., Rees, A., and Henning, G. B. (2000). Monaural and binaural detection of sinusoidal phase modulation of a 500-Hz tone. *J. Acoust. Soc. Am.* 108, 1826–1833. doi: 10.1121/1.1310195

Wu, P. Z., Liberman, L. D., Bennett, K., de Gruttola, V., O'Malley, J. T., and Liberman, M. C. (2019). Primary neural degeneration in the human cochlea: evidence for hidden hearing loss in the aging ear. *Neuroscience* 407, 8–20. doi: 10.1016/j.neuroscience.2018.07.053

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

REVIEW
published: 03 February 2021
doi: 10.3389/fnins.2021.596478

Modulation of Neuronal Potassium Channels During Auditory Processing

*Jing Wu and Leonard K. Kaczmarek**

Department of Pharmacology, Yale University School of Medicine, New Haven, CT, United States

The extraction and localization of an auditory stimulus of interest from among multiple other sounds, as in the 'cocktail-party' situation, requires neurons in auditory brainstem nuclei to encode the timing, frequency, and intensity of sounds with high fidelity, and to compare inputs coming from the two cochleae. Accurate localization of sounds requires certain neurons to fire at high rates with high temporal accuracy, a process that depends heavily on their intrinsic electrical properties. Studies have shown that the membrane properties of auditory brainstem neurons, particularly their potassium currents, are not fixed but are modulated in response to changes in the auditory environment. Here, we review work focusing on how such modulation of potassium channels is critical to shaping the firing pattern and accuracy of these neurons. We describe how insights into the role of specific channels have come from human gene mutations that impair localization of sounds in space. We also review how short-term and long-term modulation of these channels maximizes the extraction of auditory information, and how errors in the regulation of these channels contribute to deficits in decoding complex auditory information.

Keywords: cocktail party effect, sound localization, MNTB, potassium channels, firing pattern

OPEN ACCESS

Edited by:
Elizabeth Anne McCullagh, Oklahoma State University, United States

Reviewed by:
Felix Felmy, University of Veterinary Medicine Hannover, Germany
Go Ashida, University of Oldenburg, Germany
Matthew J. Fischl, National Institute on Deafness and Other Communication Disorders, United States

***Correspondence:**
Leonard K. Kaczmarek
leonard.kaczmarek@yale.edu

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience

Received: *19 August 2020*
Accepted: *11 January 2021*
Published: *03 February 2021*

Citation:
Wu J and Kaczmarek LK (2021) Modulation of Neuronal Potassium Channels During Auditory Processing. Front. Neurosci. 15:596478. doi: 10.3389/fnins.2021.596478

INTRODUCTION

The ability to discriminate and isolate a particular source of sound in a complex auditory environment, also referred to as the *cocktail party effect*, is a remarkable feature of the human auditory system (Haykin and Chen, 2005; McDermott, 2009). It requires neurons in auditory brainstem nuclei to encode aspects of a sound, such as its timing, frequency, and intensity, and then to compare differences in these characteristics in the inputs coming from the two ears. A major cue that is used to discriminate the location of a sound in space is its time of arrival at the two ears. A sound that arrives at the right ear earlier than at the left will be perceived as coming from the right, while one that arrives at both ears simultaneously will appear to originate in front of (or directly behind) the listener (**Figure 1A**). Similarly, a higher intensity of sound at the right ear will promote the impression that the sound originates on the right. Other cues, such as the frequency distribution within the sound, contribute to detection of sound location, particularly in distinguishing sounds coming from above, below or behind the listener (Tollin, 2009; Grothe et al., 2010; Grothe and Pecka, 2014; Yin et al., 2019). Nevertheless, the time of arrival appears to be an essential cue in distinguishing sound location and is essential to a person's ability to

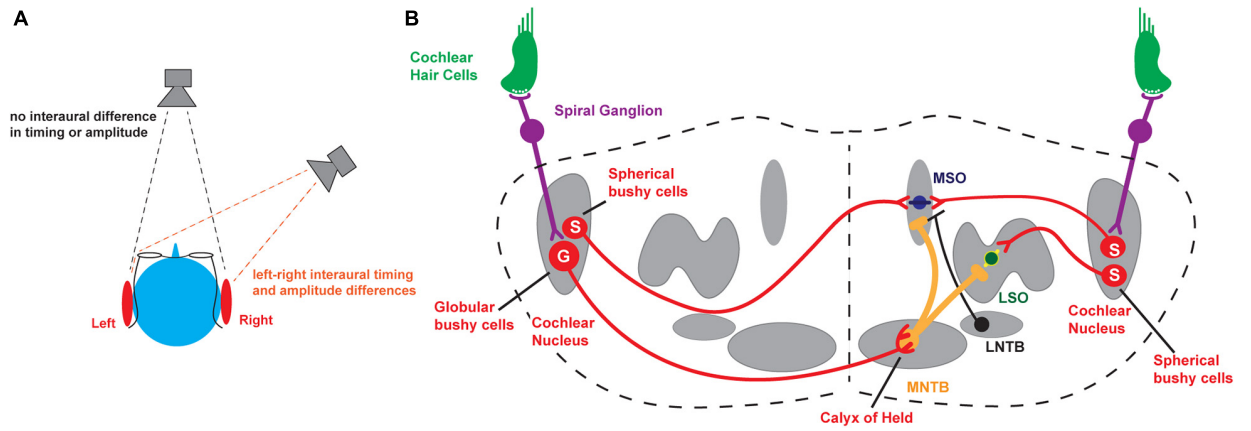

FIGURE 1 | Binaural hearing. **(A)** Schematic illustrating interaural comparisons. **(B)** Brainstem circuits underlying the detection of interaural timing and intensity differences. MNTB, medial nucleus of the trapezoid body; MSO, medial superior olive; LSO, lateral superior olive; S, Spherical bushy cells; G, Globular bushy cells. For completeness, the lateral nucleus of the trapezoid body (LNTB), which sends inhibitory input to the MSO, is also depicted.

focus attention to content originating in one specific location and to ignore multitudes of sounds originating elsewhere (i.e., the cocktail party effect). Humans can readily detect interaural time differences of tens of microseconds, much less than the duration of a neuronal action potential (Carr and MacLeod, 2010).

The major neuronal circuits that detect interaural differences in timing and intensity of sounds are located in the brainstem and are illustrated in **Figure 1B**. Along with other aspects of an auditory stimulus, the information that is used for the localization of sounds is first detected and transduced by auditory hair cells and then transmitted through neurons in the spiral ganglion to bushy cells in the anteroventral cochlear nucleus (AVCN), faithfully preserving information about the frequency, intensity, and timing of each stimulus. In turn, spherical bushy cells relay this information to the medial and lateral superior olivary nuclei (MSO and LSO) where interaural time and intensity differences from the two cochleae are computed and compared directly (Tollin, 2003, 2009; Grothe et al., 2010; Grothe and Pecka, 2014; Yin et al., 2019). An intermediate nucleus of this circuit is the medial nucleus of the trapezoid body (MNTB). The MNTB is an inverting relay, receiving excitatory inputs from the contralateral globular bushy cells and providing ipsilateral glycinergic inhibition to both the MSO and LSO (Moore and Caspary, 1983). The firing pattern of each MNTB neuron is dominated by an excitatory synaptic input from a giant presynaptic terminal, called the *calyx of Held*, located at the end of the axons of the globular bushy cells of the AVCN (Held, 1893; Banks and Smith, 1992; Forsythe and Barnes-Davies, 1993). The calyx of Held synapse targets most of the cell body, providing very secure and accurately timed excitation of the MNTB neurons that *in vivo* may enable spatial localization of sound transients (Joris and Trussell, 2018). The large calyx of Held has been used widely to investigate presynaptic ion channels, neurotransmitter release, and synaptic plasticity (for reviews see Schneggenburger and Forsythe, 2006; Borst and Soria van Hoeve, 2012; Baydyuk et al., 2016). Many of the studies of how ion channels become modified by stimulation in the auditory brainstem have been carried out using the principal neurons of the MNTB, as well as these specialized presynaptic terminals.

POTASSIUM CHANNELS IN AUDITORY BRAINSTEM NEURONS

Most of the neurons in pathways that provide inputs to the MSO and LSO have two critical features that distinguish them from most other neurons in the nervous system. First, they are capable of firing at very high rates of up to 800 action potentials/second or more (Taschenberger and von Gersdorff, 2000; Kopp-Scheinpflug et al., 2003b). Second, they lock their action potentials with microsecond precision to the specific phase of a sound with a frequency of <2 kHz or so. Neurons with characteristic frequencies >2 kHz lock their action potentials to the envelope of amplitude-modulated high frequency sounds. These high frequency neurons are also typically able to phase lock to lower frequencies even better than neurons with lower characteristic frequencies (Joris et al., 1994).

Another feature of neurons such as those in the MNTB is their ability to extract auditory information even with little or no change in overall firing rate. Even in silence, MNTB neurons are driven to fire at rates from 10 to ∼ 200 Hz by afferent activity that originates as spontaneous release of transmitter from hair cells (Brownell, 1975; Hermann et al., 2007; Kopp-Scheinpflug et al., 2008). At low intensities of sound stimulation, however, the action potentials become entrained to the auditory stimulus (**Figure 2**). As the intensity of sounds is increased, the firing rate of MNTB neurons can be pushed to over 800 Hz. Thus, these neurons respond to changes in both the timing and intensity of auditory stimuli.

These ability of auditory brainstem neurons to transmit auditory information at these high rates requires them to have membrane properties appropriate for rapid transduction of synaptic inputs into outgoing trains of action potentials. The proteins that determine the intrinsic excitability of neurons are

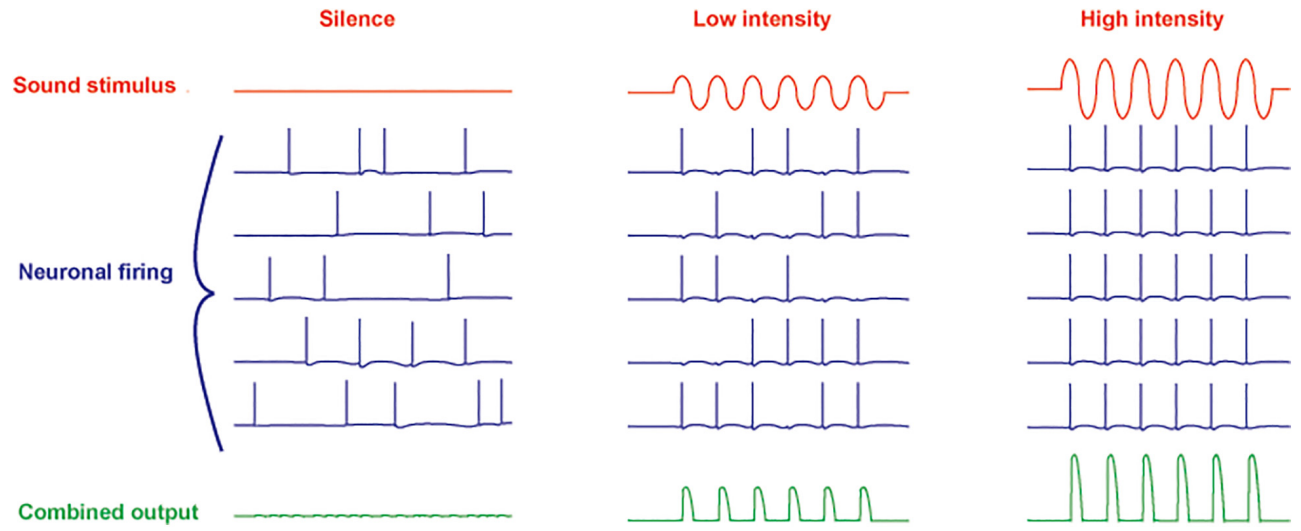

FIGURE 2 | Schematic diagram of responses of five neurons to increasing intensity of a sinusoidal sound stimulus. A combined ensemble average of the output is illustrated in the bottom traces. In silence, spontaneous activity (originating in spontaneous release of neurotransmitter from cochlear hair cells) in all neurons is uncorrelated (*left*). At low sound intensities, the action potentials become entrained to the stimulus with no overall change in firing rate (*center*). At high intensity, the firing rate of the neurons is increased with all action potentials phase-locked to the stimulus (*right*).

ion channel proteins and the auxiliary proteins and enzymes that directly determine their biophysical properties. The major pore-forming α-subunits that constitute the core of each ion channel are selective for sodium, calcium, potassium or chloride ions, and each of these is critical to defining the way a neuron responds to stimulation (Leao, 2019; Rasmussen and Trimmer, 2019; Kaczmarek, 2020). Among these pore-forming subunits, the most diverse group is that of the α-subunits for channels selective for potassium ions. The number of known genes that encode potassium channel α-subunits (77 genes) is greater than that for all the other subunits combined (Kaczmarek, 2020). Moreover, in most cases, the pore of a potassium channel is formed by a tetramer of α-subunits, which can be the product of the same or a different gene (**Figures 3A,B**). Alternative splicing of the mRNAs for most of the α-subunits, coupled with the fact that each subunit can be modified by posttranslational events such as phosphorylation or assembly with auxiliary subunits, provides a near infinite number of possibilities to regulate properties of potassium channels.

Potassium channels can be categorized into five groups, each of which has different gating and pharmacological properties and contributes in a different way to aspects of excitability (Alexander et al., 2019). These are (1) voltage-dependent channels of the Kv family, (2) calcium-activated channels (K_{Ca} channels), (3) sodium-activated channels (K_{Na} channels), (4) inwardly rectifying channels (K_{ir} channels), and (5) two pore domain channels (K_{2P} channels). Recent work has also indicated that some of these channels regulate other aspects of cell biology beyond membrane excitability, so called non-conducting functions (Kaczmarek, 2006; Lee et al., 2014; Rasmussen and Trimmer, 2019). The channels that have received the most amount of experimental attention in auditory brainstem neurons, particularly in the MNTB, are depicted in **Figure 3A**. These are the Kv1.1, Kv1.2, Kv2.2, Kv3.1, and Kv3.3 channel subunits of the Kv family and the sodium dependent K_{Na} channels. Several other channels, whose functions are less well understood, including Kv1.6, Kv3.4, Kv4.3, Kv11 channels, and the two-pore domain subunits $K_{2P}1$ and $K_{2P}15$ will also be discussed.

HUMAN MUTATIONS THAT IMPACT SOUND LOCALIZATION

Some key insights into which potassium channels play key roles in auditory neurons have come from the study of neurological disorders that are associated with deficits in processing auditory information, particularly the localization of sounds in space. These include autism, Fragile X syndrome, and certain ataxias (Middlebrooks et al., 2013; Ferron, 2016; McCullagh et al., 2020). Fragile X syndrome, the leading known cause of inherited intellectual disability, results from mutations that suppress the expression of Fragile X Mental Retardation Protein (FMRP), a mRNA binding protein that controls the function and the expression level of a variety of proteins including several ion channels such as the Kv3.1, Kv3.3, Kv1.2 and K_{Na}1.1channels (Darnell et al., 2001; Brown et al., 2010; Strumbos et al., 2010a; Zhang et al., 2012; Yang et al., 2018). The characteristics of these channels will be described later. Fragile X patients are hypersensitive to auditory stimuli (St Clair et al., 1987; Arinami et al., 1988; Rojas et al., 2001; Castren et al., 2003; Roberts et al., 2005; Van der Molen et al., 2012) and are impaired in their ability to discriminate interaural timing, rendering them unable to localize sounds (Hall et al., 2009; Rotschafer and Razak, 2014). A second genetic condition is Spinocerebellar Ataxia type 13 (SCA13), a movement disorder caused by mutations in the gene encoding the Kv3.3 channel (Zhang and Kaczmarek, 2016). These

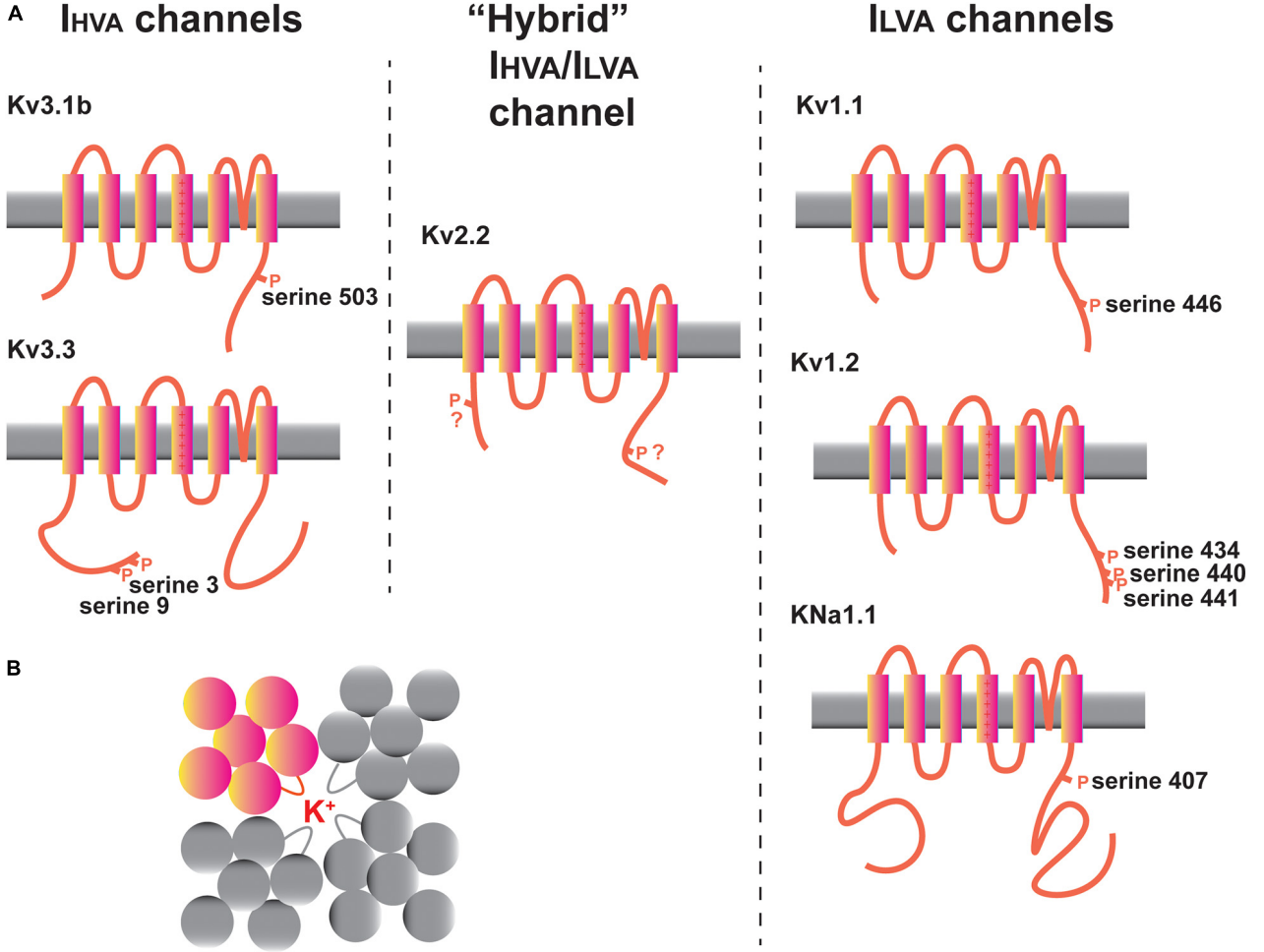

FIGURE 3 | Potassium channels in the MNTB. **(A)** Diagrams illustrating transmembrane topology and some of the regulatory phosphorylation sites in I_{HVA} (Kv3 family) and I_{LVA} (Kv1 and K_{Na} families) channels with documented functions in the MNTB. The intermediate I_{HVA}/I_{LVA} Kv2 family Kv2.2 channel is also depicted. **(B)** "Top down" schematic diagram illustrating the assembly of four Kv family subunits into a tetrameric channel with the potassium ion conduction pathway in the center.

mutations cause either early-onset or late-onset motor disabilities and different mutations result in either a loss of Kv3.3 current or a change in its voltage-dependence or kinetics. Nevertheless, even in the case of the late-onset disease, younger SCA13 patients with no motor symptoms are unable to discriminate interaural timing differences of up to a millisecond (Middlebrooks et al., 2013).

VOLTAGE-DEPENDENCE AND KINETICS OF A POTASSIUM CHANNEL DETERMINES ITS EFFECT ON EXCITABILITY: I_{HVA} AND I_{LVA} CHANNELS

In auditory neurons as in other excitable cells, potassium channels regulate several key aspects of intrinsic excitability (**Figure 4**). These aspects include: (1) the resting membrane potential, (2) the width of an action potential (**Figure 4A**), (3) the threshold for generation of an action potential (**Figures 4A,D**), (4) the timing of an action potential following a synaptic input (**Figures 4A,B**), (5) whether the neuron fires multiple action potentials in response to stimulation, and, if so, the frequency of firing (**Figure 4C**), and (6) the amount of neurotransmitter released at a synaptic ending following a presynaptic action potential.

Because the efflux of potassium ions during the opening of a channel hyperpolarizes neurons, a common perception is that increases in potassium currents reduce excitability. This is, however, not always the case. The opening of many potassium channels, particularly those in the Kv family, depends on the membrane potential. If a potassium channel opens at the resting potential, this will indeed hyperpolarize the cell and reduce excitability. A subset of channels, however, open rapidly only at very positive potentials, i.e., only during an action potential as the membrane potential approaches 0 mV. An increase in the number or activity of such channels, which speeds repolarization of action potentials without affecting threshold, actually increases neuronal excitability because it reduces the inactivation of sodium channels. The duration of action potentials ($\sim 200\ \mu$ s) in

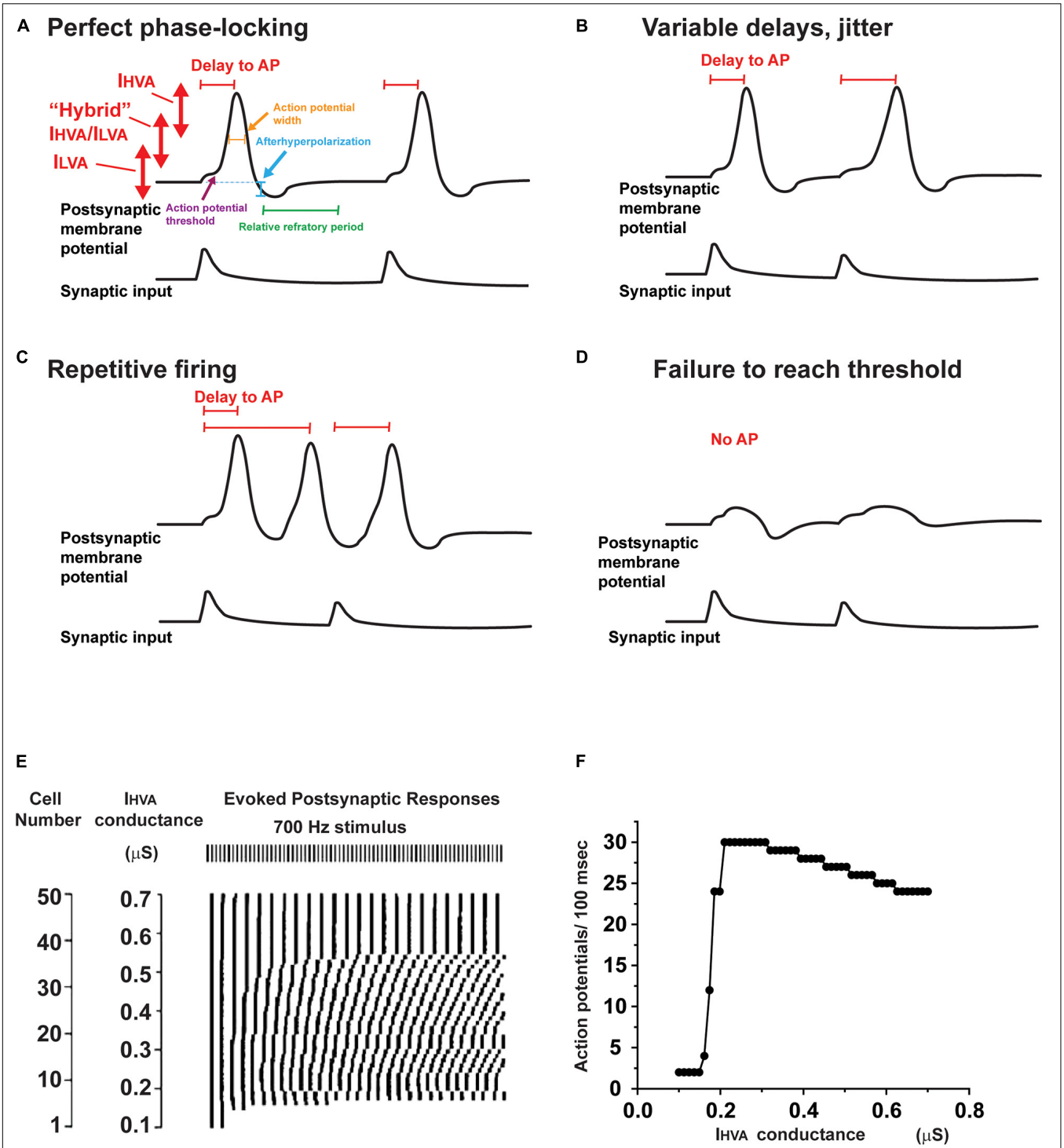

FIGURE 4 | Effects of altering potassium conductances on firing patterns. **(A)** Illustration of the voltage ranges in which I_{HVA} and I_{LVA}, as well as intermediate potassium channels operate during the response of a neuron to two consecutive synaptic inputs. Changes in these potassium currents are able to alter the resting potential, action potential threshold, height and width, as well as the afterhyperpolarization and refractory period that follows each action potential. **(B–D)** A change in potassium currents alter the response of a neuron from a response like that in panel **(A)** to one with variable delays in response to each synaptic input **(B)**, repetitive firing that is not locked to the synaptic input **(C)** or failure to evoke an action potential **(D)**. **(E)** Results of numerical simulations showing that the timing of action potentials evoked by a regular 700 Hz stimulus differs in neurons that have different levels of I_{HVA} currents. The timing of the evoked action potentials is presented in raster plot under the stimulus pulses (Kaczmarek, 2012). **(F)** A plot of firing rate as a function of the different levels of I_{HVA} currents in panel **(E)**. Note that, although the firing rate changes by only a small amount over wide ranges of I_{HVA}, the timing of the action potentials with respect to the stimuli (shown in panel **E**) is very sensitive to small changes in I_{HVA}.

many auditory brainstem neurons is less than the time constant for inactivation of sodium current in the same cells (Leao et al., 2006a; Scott et al., 2010). Thus, limiting the height and width of action potentials reduces the amount of sodium current inactivation with a single action potential. Moreover, rapid repolarization by potassium currents increases the proportion of time spent at negative potentials, accelerating recovery from inactivation. Thus, in both numerical simulations and experiments with transfected neurons, increasing the number of potassium channels that open rapidly at positive potentials markedly enhances the ability to fire at high rates (Kaczmarek et al., 2005; Song et al., 2005).

Potassium currents that activate near the resting potentials have been termed "low-voltage activated" or I_{LVA} currents, while those that activate selectively during action potentials have been termed "high-voltage activated" or I_{HVA} currents. This distinction is particularly important for auditory brainstem neurons, which fire at hundreds of Hz. **Figure 3A** separates some of the channels expressed in MNTB neurons into these two groups, together with one channel (Kv2.2) that falls between these two extremes. The approximate voltage range in which these classes of channel begin to activate is depicted in **Figure 4A**.

ION CHANNELS CAN BE RAPIDLY MODIFIED BY SECOND MESSENGER PATHWAYS

All ion channel proteins can be modified by posttranslational mechanisms, such as protein phosphorylation. Some modifications occur after a channel has been synthesized and may influence its stability to proteases, its location and the rate at which it is trafficked or inserted into the plasma membrane (Yang et al., 2007b; Vacher and Trimmer, 2011). Other processes can rapidly modify either the current amplitude or kinetics of a channel once it has been in placed in the membrane. **Figure 3A** indicates some of the phosphorylation sites on the I_{HVA} and I_{LVA} channels present in the MNTB which have been demonstrated to modify the currents that flow through these channels. Clearly, modulation of potassium currents has the potential to alter the way a neuron responds to precisely timed input (Kaczmarek, 2012). For example, in **Figure 4A**, a single action potential is accurately locked to each synaptic depolarization. An alteration in potassium currents could change a response like that in **Figure 4A** to one with a variable delay from stimulation to the onset of an action potential (jitter, **Figure 4B**), an overabundance of evoked action potentials (**Figure 4C**) or failure to evoke an action potential in response to synaptic depolarization (**Figure 4D**).

At first, one might assume that the ability to fire at a high rate with high temporal accuracy might require an invariant set of potassium currents that cannot be modulated. However, a fixed set of potassium currents that provide optimal locking to one pattern of stimulation (e.g., that evoked by low frequency sounds at low intensity) may fail to generate an adequate response to another pattern (e.g., high frequency sounds at high intensity). This is evident in numerical simulations such as those in **Figure 4E**, which depicts a raster plot of the timing of action potentials evoked by a 700 Hz stimulus in 50 different neurons, which all have the same amplitude of I_{LVA} but have different levels of I_{HVA} currents (Kaczmarek, 2012). The timing of the responses varies substantially such that at low levels of I_{HVA}, firing in response to the stimulus train cannot be sustained, while regular responses are evoked only at much higher levels of I_{HVA} (**Figure 4F**). The specific levels of I_{HVA} or I_{LVA} required for optimal locking of action potentials to synaptic inputs depend, however, on the intensity and frequency of the stimulus, such that no one set of conductances is optimal for all conditions. In addition, at synapses such as the calyx of Held, the amount of neurotransmitter release and the recovery of the readily releasable pool of neurotransmitter change as a function of firing rate of the presynaptic AVCN cell (Wang and Kaczmarek, 1998; Hermann et al., 2007; Lorteije et al., 2009). Thus, altering potassium current by protein phosphorylation of other second messenger pathways may allow a neuron to adapt appropriately to different patterns and amplitudes of synaptic inputs.

Kv3 AND Kv2 CHANNELS ARE REQUIRED FOR HIGH RATES OF FIRING

The canonical high-voltage activated potassium channels that are required for rapid firing in many auditory brainstem neurons are Kv3 channels, particularly Kv3.1 and Kv3.3 (Kaczmarek and Zhang, 2017). These are expressed in the cochlear nucleus, MNTB, superior olive, and inferior colliculus, as well as many fast-firing neurons in other parts of the adult nervous system (Perney et al., 1992; Perney and Kaczmarek, 1997; Grigg et al., 2000). Under certain circumstances, Kv2.2 can also adopt this role to permit neurons such as those of the MNTB to fire at rates greater than 100 Hz (Steinert et al., 2008, 2011).

The best-studied Kv3 family member is Kv3.1b, which is expressed in the soma of the postsynaptic principal MNTB neurons as well as in the calyx of Held presynaptic terminals (Li et al., 2001; Macica and Kaczmarek, 2001; Elezgarai et al., 2003; Song et al., 2005; Choudhury et al., 2020). Under normal conditions, the voltage-dependence of Kv3.1b channels matches that described above for I_{HVA} currents. As a cell is progressively depolarized, currents begin to appear at potentials of ~ -15 to -10 mV and 50% activation occurs at $\sim +15$ mV (Luneau et al., 1991; Brown et al., 2016). The Kv3.1b current activates very rapidly during the upstroke of an action potential. A current with these characteristics is present in patch clamp recordings of MNTB neurons (Wang et al., 1998).

Genetic and pharmacological experiments, as well as numerical simulations, using a variety of neurons have shown that Kv3.1b contributes to the rapid repolarization of action potentials that allows them to fire at rates of several hundred Hz (Kaczmarek and Zhang, 2017). In MNTB neurons, genetic deletion of the *Kv3.1* gene does not alter total levels of potassium current (Macica et al., 2003; Choudhury et al., 2020), because of a compensatory increase in Kv3.3 current, with no change in levels of Kv3.3 protein (Choudhury et al., 2020). As a result, the extent to which Kv3.1 knockout alters characteristics of the currents and

the ability of MNTB neurons to be driven at high rates depends on experimental factors such as animal strain and recording temperature (Macica et al., 2003; Choudhury et al., 2020).

Kv3.3 channels are also widely expressed in the auditory brainstem (Li et al., 2001). Their conducting properties are in general similar to those of Kv3.1 channels in that they produce I_{HVA} currents that repolarize action potentials (Kaczmarek and Zhang, 2017). Unlike Kv3.1, however, Kv3.3 channels inactivate during sustained depolarization lasting tens to hundreds of milliseconds. Moreover, the cytoplasmic C-terminus of Kv3.3 is larger than that of other members of the Kv3 family. This cytoplasmic region binds several proteins that directly nucleate actin filaments, including Hax-1 and the Arp2/3 complex. As a result, when Kv3.3 channels are inserted into the plasma membrane, they are capable of triggering a dense subcortical actin network (Zhang et al., 2016). Both the cellular and subcellular distribution of Kv3.3 is distinct from that of Kv3.1. At the cellular level, both Kv3.3 and Kv3.1 are expressed in neurons of the AVCN and the MNTB, but only Kv3.3 is found in neurons of the LSO and MSO (Perney and Kaczmarek, 1997; Li et al., 2001). At the subcellular level, Kv3.1 localizes to the "back" face of the terminals (calyces of Held) of AVCN globular bushy cells (Elezgarai et al., 2003), while Kv3.3 localizes to the presynaptic membrane facing the postsynaptic neurons, which is characterized by a dense subcortical actin network (Zhang et al., 2016; Kaczmarek et al., 2019). While immunostaining suggests that, within the MNTB itself, Kv3.3 is largely confined to the presynaptic terminals of AVCN neurons (Zhang et al., 2016), genetic deletion of either Kv3.1 or Kv3.3 does not reduce total potassium current in MNTB neurons, presumably because of compensatory changes in expression of the other subunit (Macica et al., 2003; Choudhury et al., 2020). In contrast, genetic knockout of Kv3.3 reduces potassium current in LSO neurons and severely impairs their ability to fire at high rates (Choudhury et al., 2020).

The importance of Kv3.3 for the discrimination of the source of a sound that is required for the cocktail party effect has come from studies of patients with SCA13, which is caused by mutations in in KCNC3, the human gene encoding Kv3.3 (Zhang and Kaczmarek, 2016). Kv3.3 channels are particularly abundant in cerebellar Purkinje cells and these mutations produce either early-onset or late-onset cerebellar degeneration. As noted earlier, however, late-onset SCA13 patients are completely unable to resolve interaural timing or intensity differences, even decades before they develop any detectable motor symptoms (Middlebrooks et al., 2013).

Kv2.2 channels, unlike Kv3.1 and Kv3.3, do not produce true I_{HVA} currents in that they begin to activate even with small depolarizations from the resting potential and, in MNTB neurons, are already half-activated at ~ -10 mV (Johnston et al., 2008b; Tong et al., 2013). At this potential, the Kv3 channels are just beginning to activate during an action potential (Kanemasa et al., 1995). Although they activate more slowly than Kv3 channels, under appropriate conditions, Kv2.2 can contribute to the repolarization of action potentials (Johnston et al., 2008b; Tong et al., 2013). They can therefore be considered to be hybrid I_{HVA}-I_{LVA} channels, and they will be discussed again in a later section.

SHORT-TERM MODULATION OF I_{HVA} CHANNELS

The relative contribution of Kv2 and Kv3 family channels to overall current is subject to ongoing modulation by the auditory environment. A change in potassium currents can adapt a neuron to respond appropriately to different frequencies, intensities or patterns of stimulation. The matching of potassium currents to patterns of synaptic inputs may provide an explanation for why in almost every auditory brainstem nucleus, the pattern of expression of potassium channels differs from cell-to-cell and that the level of potassium currents is subject to continual modification by the auditory environment. For example, in common with a subset of other channels in other auditory nuclei, Kv3.1b is expressed along the tonotopic gradient in the MNTB (**Figure 5A**). Highest levels are found in neurons in the medial, high frequency aspect of the MNTB (Li et al., 2001; von Hehn et al., 2004; Brew and Forsythe, 2005; Strumbos et al., 2010a).

Auditory brainstem neurons and, in particular, neurons of the MNTB and AVCN, have provided key findings on how potassium channels are regulated by incoming stimuli. For example, it is now apparent that Kv3.1 channels in such neurons are regulated over multiple time scales, ranging from tens of seconds to months following changes in auditory inputs, and this will be described in more detail below. There are two ways in which current amplitude can be modulated within a cell. First, rapid changes in potassium currents (occurring in seconds to tens of seconds) can be produced by posttranslational modifications such as phosphorylation of the channel protein. Second, the amount of channel protein in the plasma membrane can be changed, as a result of changes in transcription, translation and/or trafficking into the plasma membrane.

PHOSPHORYLATION OF Kv3.1b CHANNELS

Even fundamental aspects of a channel such as its voltage-dependence can be altered by phosphorylation. For example, casein kinase 2, often considered a constitutively active kinase, can adjust the voltage dependence of Kv3.1 current in MNTB neurons (Macica and Kaczmarek, 2001). As stated earlier, Kv3.1 channels are I_{HVA} channels that normally generate significant current at potentials positive to ~ -10 mV. In response to inhibitors of casein kinase 2, however, Kv3.1 currents begin to activate at ~ -40 mV, much closer to the resting potential, effectively turning them into I_{LVA} channels (Kanemasa et al., 1995; Macica et al., 2003). The specific sites on Kv3.1 that are modified by this enzyme are, however, not yet known, nor is it understood under what biological conditions Kv3.1 phosphorylation by casein kinase 2 is altered. The actions of casein kinase 2 can also be mimicked by a novel class of imidazolinedione compounds that convert Kv3.1 currents into I_{LVA}-like currents (Taskin et al., 2015; Brown et al., 2016; El-Hassar et al., 2019).

A much clearer biological role for phosphorylation of Kv3.1 has been found for protein kinase C (PKC). There exist two

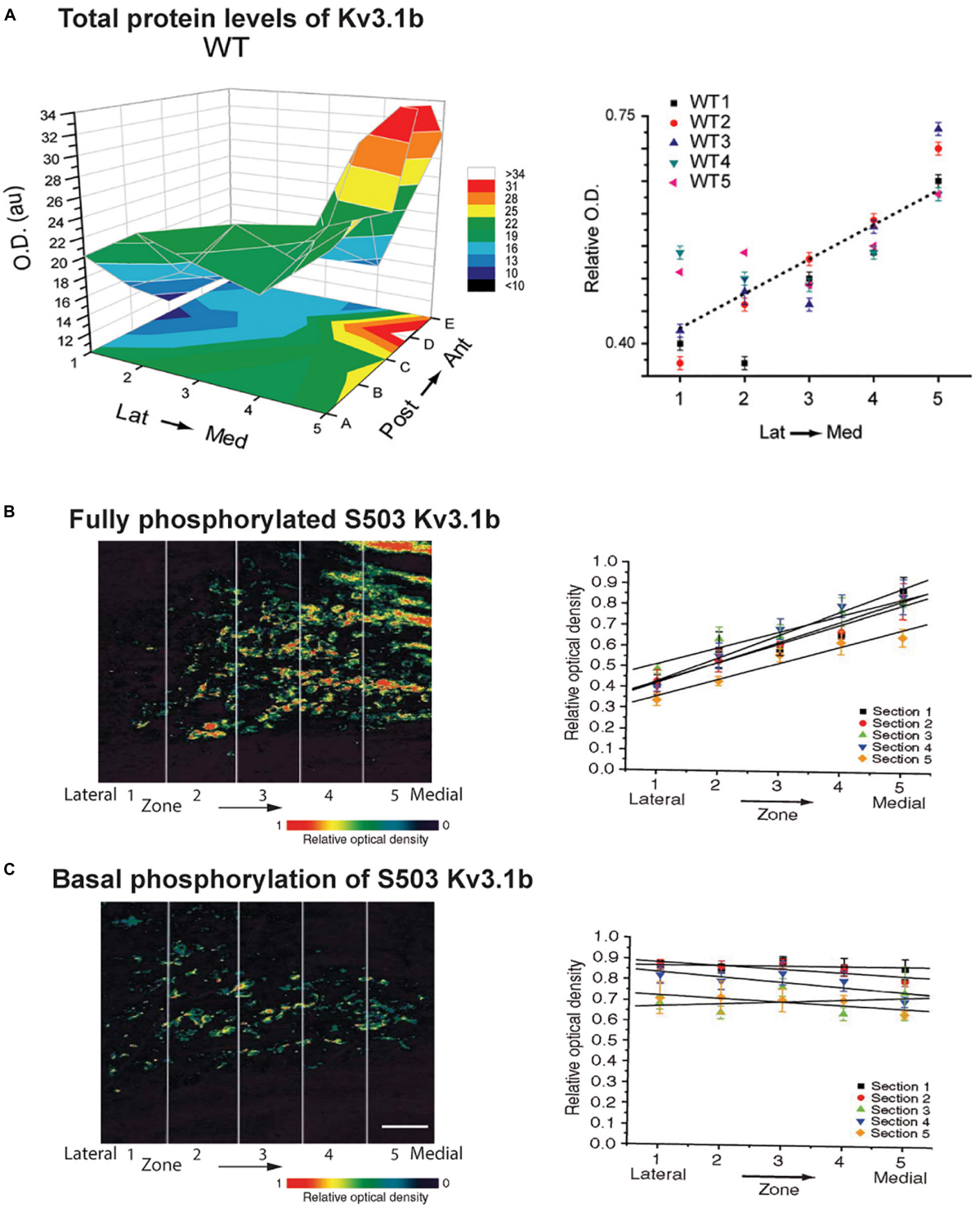

FIGURE 5 | Tonotopic gradients of Kv3.1b and phosphorylated Kv3.1b in the MNTB. **(A)** Representative three-dimensional plot of average Kv3.1b immunoreactivity (OD) in each of 25 stereotaxic zones along the lateral to medial and posterior to anterior axes in a mouse MNTB. Graph at right plots relative Kv3.1b immunoreactivity across the medial-lateral axis for five different animals (Strumbos et al., 2010a). **(B,C)** Tonotopic gradients of Kv3.1b phosphorylation in MNTB treated with a PKC activator to maximally stimulate phosphorylation of the channel **(B)** and in resting unstimulated MNTB slices **(C)**. Left panels show pseudocolor images of serine 503 phosphorylation of Kv3.1b detected with a phospho-specific antibody. Right panels show quantification of phosphorylation in five tonotopic zones along the lateral to medial axis (n = 5 sections in each case). A clear tonotopic gradient, matching that of total Kv3.1b protein is seen when channels are all maximally phosphorylated **(B)**. No gradient is, however, detected under resting conditions, indicating that the degree of phosphorylation is reduced in the medial high frequency MNTB neurons (Song et al., 2005). Scale bar, 200 μm.

splice variants of the Kv3.1 gene, Kv3.1a and Kv3.1b, which differ only in the presence of a long cytoplasmic C-terminal domain in the Kv3.1b isoform. This longer cytoplasmic domain provides an additional phosphorylation site (serine 503) for PKC. When this site is phosphorylated, the amplitude of Kv3.1b currents is partially suppressed (**Figure 6A**; Kanemasa et al., 1995; Macica et al., 2003). In auditory neurons, Kv3.1a dominates early in development and Kv3.1b becomes the dominant form after the onset of hearing (Perney et al., 1992; Perney and Kaczmarek, 1997; Liu S.J. and Kaczmarek, 1998a).

A change in the ambient auditory environment, as occurs when a person (or a rat) moves from a quiet setting to a cocktail party situation, produces a change in the phosphorylation of Kv3.1 at residue serine 503 and in I_{HVA} current amplitude. When rats are maintained in a quiet environment, Kv3.1b channels in AVCN and MNTB neurons, as well as in the calyxes of Held, are highly phosphorylated at this site (Song et al., 2005). In response to a physiological increase in sound levels (free-field click trains at 600 Hz, 70 dB sound pressure level (SPL) for 5 min), comparable to a "cocktail party" sound environment, Kv3.1b undergoes dephosphorylation at serine 503 (**Figures 6B,C**).

Experiments *in vitro* with MNTB brain slices confirmed that Kv3.1b channels are highly phosphorylated in the absence of stimulation, but are dephosphorylated within seconds upon stimulation of the input from the AVCN at 600 Hz (**Figures 6D,E**; Song et al., 2005). Consistent with the fact that phosphorylation at serine 503 suppresses Kv3.1b current, such stimulation increased I_{HVA} current and increased the ability of the principal neurons of the MNTB to fire action potentials at higher rates of stimulation by intracellular current pulses. Pharmacological and co-immunoprecipitation experiments revealed that the PKC-δ isoform selectively contributes to the basal phosphorylation of the serine 503 site, and that its dephosphorylation during stimulation of the input to the MNTB is mediated by protein phosphatases PP1/PP2A (Song et al., 2005; Song and Kaczmarek, 2006).

As stated above, levels of Kv3.1b channels vary along the tonotopic axis of the MNTB with highest levels in neurons in the medial aspect of the MNTB where neurons preferentially respond to high frequency sounds (Li et al., 2001; von Hehn et al., 2004; Brew and Forsythe, 2005; Strumbos et al., 2010a). It appears that the effect of this gradient of Kv3.1b protein on I_{HVA} current may be further enhanced by phosphorylation. When pharmacological agents are used to maximally stimulate serine 503 phosphorylation, a clear tonotopic gradient of phosphorylation is observed along the lateral-medial axis of the MNTB, exactly matching that of total Kv3.1b protein (**Figure 5B**; Song et al., 2005). In the absence of stimulation, however, levels of serine 503-phosphorylated Kv3.1b are uniform across this axis. Thus, the proportion of phosphorylated (i.e., suppressed) Kv3.1 channels is greatest in neurons at the lateral low-frequency end of the nucleus and lowest at the medial high-frequency end (**Figure 5C**; Song et al., 2005). Whether this difference in phosphorylation at different ends of the MNTB reflects a gradient of expression of PKC, PP1/P2A or some other regulator of signaling pathways is not yet known. Nevertheless, this finding suggests that under such basal conditions the gradient in I_{HVA} current is greater than that of levels of Kv3.1b itself, but that this can be modified by auditory inputs over a time course of seconds to minutes.

PHOSPHORYLATION OF Kv3.3 CHANNELS

The Kv3.3 channel, which is co-expressed with Kv3.1 in the presynaptic calyx of Held, is also regulated by PKC (Desai et al., 2008; Zhang et al., 2016). In contrast to Kv3.1b, however, the major phosphorylated residues are located on the cytoplasmic N-terminus, and phosphorylation of these sites increases rather than decreases current (Desai et al., 2008). Whether these sites on Kv3.3 are modified during changes in the auditory environment and how they impact the function of the MNTB are not yet known.

RAPID REGULATION OF Kv2.2 CHANNELS IN MNTB NEURONS

Numerous numerical computations of the gating of Kv3 channels have shown that these channels are essential for neurons to fire at high rates. This is because their rapid deactivation minimizes the relative refractory period that follows an action potential. Indeed, deactivation is so rapid that a unique gating process ensures complete repolarization; Kv3.1 generates a resurgent potassium current during the falling phase of an action potential that provides the repolarization drive to terminate each spike in a train (Labro et al., 2015). Nevertheless, it appears that Kv3 channels sometimes delegate some of their role in repolarization to Kv2.2 channels, perhaps when firing rates are low for a sustained period. Like Kv3.1, the "hybrid I_{HVA}-I_{LVA}" Kv2.2 channels are expressed in a gradient along the tonotopic axis of the MNTB, but this gradient is in the opposite direction from that of Kv3.1, with highest levels in the lateral low-frequency neurons (Johnston et al., 2008b; Tong et al., 2013). They contribute to the hyperpolarization between action potentials during repetitive firing such that genetic elimination of Kv2.2 reduces the number of action potentials that can be evoked by repetitive stimulation (Johnston et al., 2008b; Tong et al., 2013). Stimulation of presynaptic inputs to MNTB at low rates, at or below those encountered *in vivo* in silence (10 Hz-150 Hz stimulation Brownell, 1975; Hermann et al., 2007; Kopp-Scheinpflug et al., 2008), suppresses Kv3.1 currents (Steinert et al., 2008) while increasing the amplitude of Kv2.2 currents (Steinert et al., 2011). The mechanism of this increase has been shown to require the release of nitric oxide (NO), and activation of the cyclic GMP-dependent protein kinase and PKC (Steinert et al., 2008, 2011). Although it is known that phosphorylation of specific sites on the closely related Kv2.1 potassium channel in other cells influences its biophysical properties (Ikematsu et al., 2011), and its insertion into the plasma membrane (He et al., 2015), the specific sites on Kv2.2 required for its recruitment in MNTB neurons are yet known.

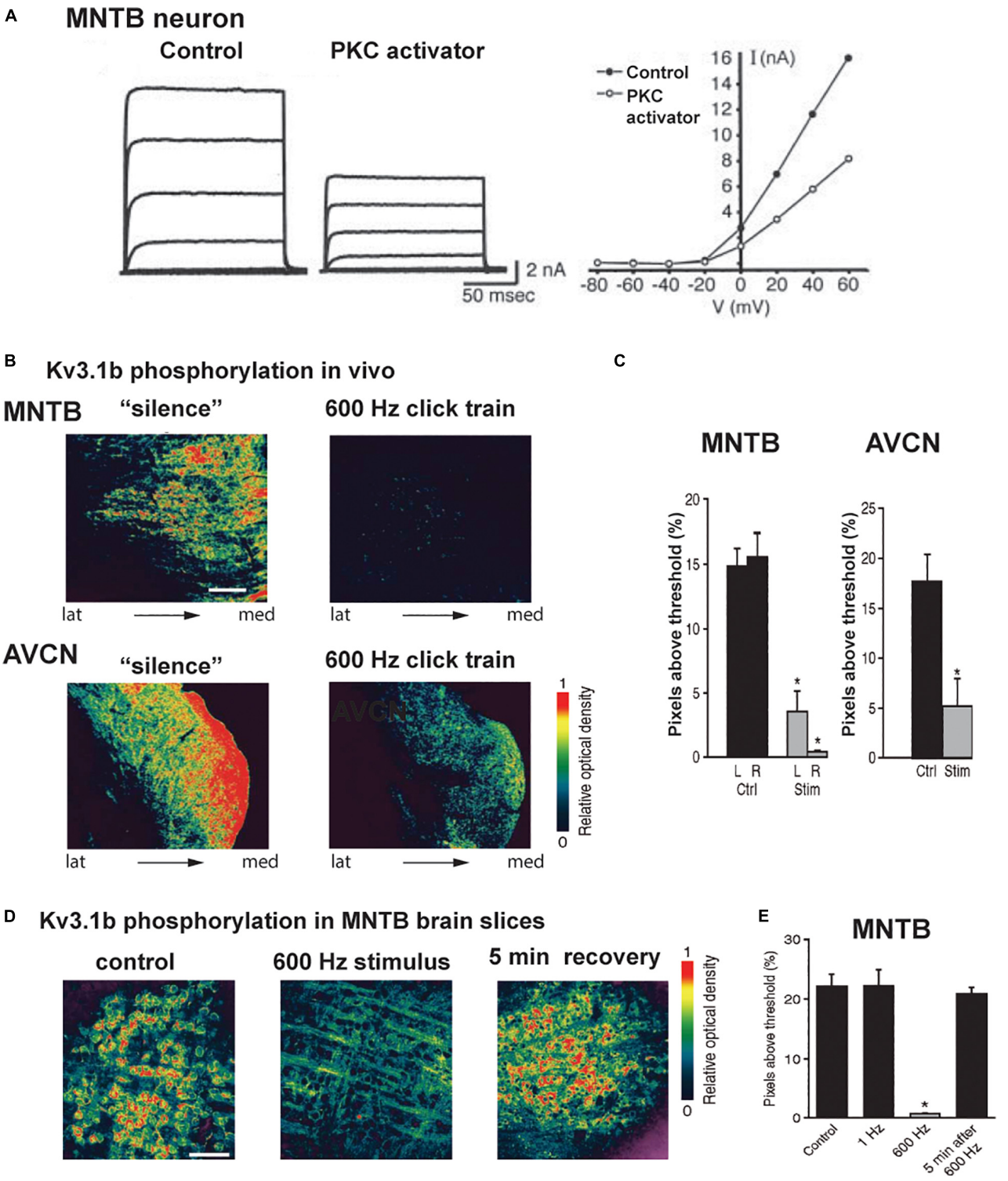

FIGURE 6 | Kv3.1b phosphorylation is reduced *in vivo* and *in vitro* after auditory or synaptic stimulation. **(A)** Voltage clamp traces and current-voltage relations of isolated I_{HVA} current in an MNTB neuron showing that current is reduced after exposure to an activator of PKC (Macica et al., 2003). **(B)** Pseudocolor images of phospho-Kv3.1b immunostaining in an MNTB (*top*) and AVCN (*bottom*) from animals kept in a soundproof room or exposed to stimulation with a 600-Hz click train (70 dB SPL, sound pressure level for 5 min). **(C)** Quantification of the phospho-Kv3.1 immunofluorescence in MNTB and AVCN (Ctrl, no stimulation; Stim, 600 Hz click train). **(D)** Pseudocolor images of phospho-Kv3.1b immunostaining in MNTB principal neurons in brain slices, under control conditions (*left*) and after stimulation at 600 Hz for 20 s (*center*) and 5 min after 600-Hz stimulation (*right*). Scale bar, 100 μm. **(E)** Quantification of change in phosphorylation for the three conditions in panel **(C)**, as well as for a low rate of stimulation (1 Hz for 20 s). In panels **(C,E)**, $^*p < 0.05$ when compared to control (Song et al., 2005).

ACTIVITY-DEPENDENT CHANGES IN EXPRESSION OF POTASSIUM CHANNEL PROTEINS

The phospho-specific immunostaining techniques used to examine Kv3.1b in the experiments described above allowed changes to be detected within 60 s of stimulation (Song and Kaczmarek, 2006). Changes in phosphorylation state can, however, occur within less than a second of stimulation, so changes in I_{HVA} current in response to auditory input may occur more rapidly than was detected in those experiments. In addition to these rapid changes, there is also a much slower change in the levels of Kv3.1b channel protein in the plasma membrane that is also brought about by incoming sounds. The difference in levels of Kv3.1b protein between the medial high-frequency and the lateral low-frequency MNTB neurons is maintained by ongoing auditory stimulation. This gradient is absent in mice that undergo hearing loss because of cochlear hair cell degeneration (von Hehn et al., 2004; Leao et al., 2006b). In the C57BL/6 strain of mice, which undergoes age-related hearing loss, the Kv3.1b gradient in the MNTB is present for the first few months after birth, but is lost by 6 months of age (von Hehn et al., 2004).

Exposure of normal rats or mice to sound stimuli similar to those that produce Kv3.1b dephosphorylation also triggers the synthesis of new Kv3.1b protein and, depending on the frequency distribution of the sound, can alter the tonotopic gradient (Leao et al., 2010; Strumbos et al., 2010a,b). Changes in levels of Kv3.1b protein in MNTB neurons can be detected within 20 min to one hour after the onset of the sound stimulus. One interesting aspect of these experiments is that levels of Kv3.1b in neurons outside of the tonotopic region targeted by the auditory stimulus may decrease during the same time period (Strumbos et al., 2010b). This observation suggests the existence of mechanisms that regulate the response of the MNTB globally, perhaps by sideband inhibition or other mechanisms that govern lateral interactions among neurons in this nucleus (Kaczmarek, 2019).

THE FMRP PROTEIN REGULATES LEVELS OF Kv3.1 CHANNELS IN RESPONSE TO STIMULATION

A key mechanism that plays a role in stimulating the production of new Kv3.1b channels in response to auditory stimulation is the FMRP pathway. FMRP is an mRNA-binding protein that is required for activity-dependent translation of mRNAs for a large subset of proteins (De Rubeis and Bagni, 2010; Darnell et al., 2011; Richter et al., 2015). Human patients lacking FMRP have Fragile X syndrome, the leading known inherited cause of autism, and loss of FMRP leads to hypersensitivity to auditory stimuli and impairs the ability to localize sounds in space (Hall et al., 2009; Rotschafer and Razak, 2014).

While FMRP is now known to control a variety of biological processes, one of the ways it influences protein synthesis is by binding its target mRNAs and suppressing their translation. Subsequent neuronal stimulation may alleviate this block of translation, leading to enhanced synthesis of the protein. Messenger RNA for Kv3.1 was one of the very first described targets for FMRP (Darnell et al., 2001, 2011; Strumbos et al., 2010a). Consistent with this "canonical" role for FMRP in regulating the expression of this channel, mice lacking FMRP have elevated levels of Kv3.1 protein and I_{HVA} currents in MNTB neurons (Strumbos et al., 2010a). The tonotopic gradient of Kv3.1 is absent in these animals (**Figure 7A**), and auditory stimulation has no effect in further enhancing levels of Kv3.1 in either the MNTB or AVCN (**Figure 7B**), suggesting that rates of translation are maximal in the absence of FMRP (Strumbos et al., 2010a). While it is likely that lack of FMRP results in loss of tonotopy and activity-dependent translation for many of the other proteins whose mRNAs are targets of this mRNA-binding protein, this has not been tested directly (McCullagh et al., 2020).

REGULATION OF TRANSCRIPTION OF I_{HVA} CHANNEL mRNAs BY CREB

Rapid adjustment of the levels of I_{HVA} currents by activation of protein kinases, and potentially the slower changes produced by activity-dependent synthesis or insertion into the plasma membrane may play a role in the cocktail party effect by adjusting firing patterns to maximize the extraction of aspects of a sound required for localization in space. The fact that individuals vary considerably in their ability to localize sounds may result, however, from differences in even longer-term mechanisms, specifically regulation of transcription, which influences levels of ion channel mRNA in each neuron. The gene for Kv3.1 has a cyclic AMP/Ca2^{+}-response element (CRE) upstream of the start of transcription (Gan et al., 1996; Gan and Kaczmarek, 1998). Transcription of the Kv3.1 gene is triggered when stimulation of neurons elevates cytoplasmic cyclic AMP or Ca^{2+} levels. These cause the phosphorylation of the transcription factor CREB (Cyclic AMP/Ca^{2+}-Responsive Element-Binding protein) (Gonzalez and Montminy, 1989; Dash et al., 1991), which then binds the CRE and activates synthesis of Kv3.1 mRNA (Gan et al., 1996). Depolarization of inferior colliculus neurons for six hours (Liu S.Q. and Kaczmarek, 1998b) or of long-term cultures of MNTB neurons for several days (Tong et al., 2010), using a high potassium external solution, has been shown to increase I_{HVA} currents and levels of mRNA for Kv3.1 and Kv3.3 channel subunits, respectively.

To examine the role of CREB phosphorylation *in vivo*, immunostaining for phosphorylated CREB (pCREB) in the MNTB has been used to determine which neurons are likely to be actively transcribing mRNA for Kv3.1 and other genes regulated by a CRE (von Hehn et al., 2004; **Figure 8A**). These experiments have demonstrated that transcription, like Kv3.1 protein synthesis and phosphorylation, is actively controlled by auditory inputs. All neurons in the MNTB appear to have similar levels of CREB itself. The phosphorylation of CREB, however, appears to be an "all or none" event that occurs in subsets of neurons clustered at different locations along the tonotopic axis. Because the location of these clusters along the axis varies from animal to animal in fixed tissue, it is likely that these patterns

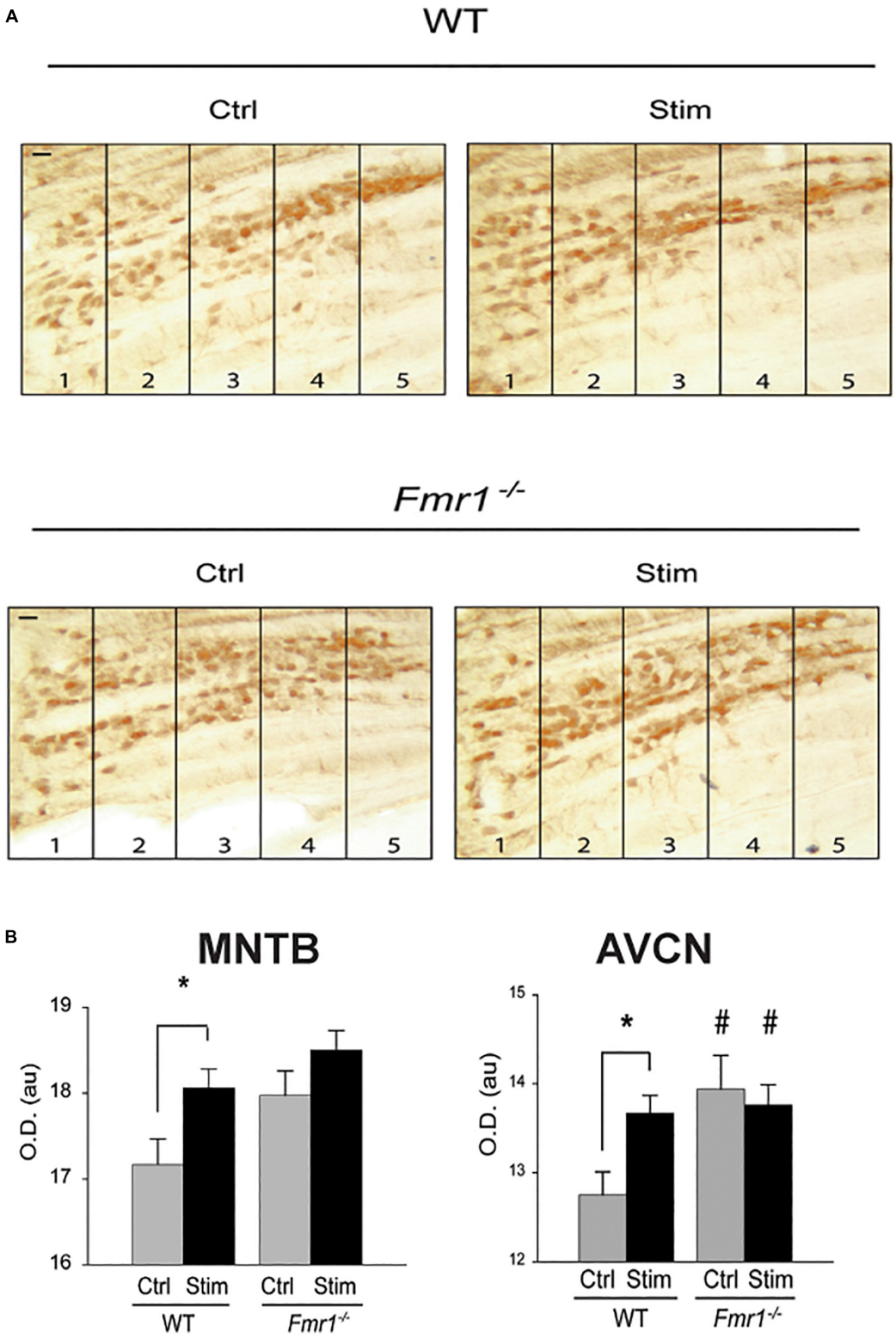

FIGURE 7 | FMRP is required for sound-induced increases in Kv3.1b protein in MNTB. **(A)** Representative sections immunolabeled for total Kv3.1b protein in wild type mice or those in which the gene for FMRP was deleted (*Fmr1*$^{-/-}$). Animals were maintained in silence (Ctrl) or exposed to acoustic stimulation for 30 min (Stim, 32 kHz tones amplitude-modulated at 380-420 Hz, 65 dB). Numbers refer to lateral-to-medial tonotopic zones 1-5 as in **Figure 5**. **(B)** Quantification of changes in Kv3.1b in MNTB and AVCN demonstrates that stimulation increases levels of protein in the wild type mice but not those lacking FMRP, which have uniformly higher levels of Kv3.1b in the absence of stimulation (*p = 0.02; $^{\#}p$ = 0.04 compared to wild type controls) (Strumbos et al., 2010a).

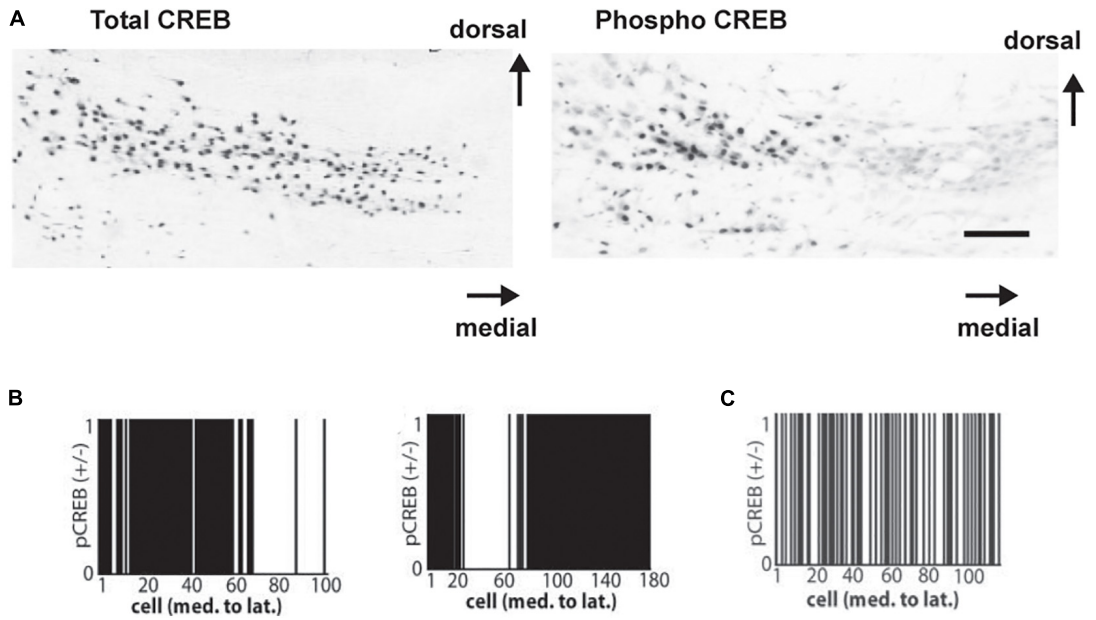

FIGURE 8 | (A) Coronal sections showing immunoreactivity for CREB (*left*) and for pCREB (*right*) in mice with good hearing (8-month-old CBA/J strain). Scale bar, 250 μm. **(B)** Clustered distribution of nuclear pCREB immunoreactivity in all cells across the medial-lateral axis for two normal hearing mice. Cells positive for pCREB were assigned a value of 1, and those lacking pCREB were assigned a value of 0. **(C)** As **(B)** but showing a random distribution of pCREB immunoreactivity in a hearing-impaired mouse (6 month old C57BL/6 strain is shown) (von Hehn et al., 2004).

reflect differences in incoming auditory inputs at the time of fixation. More significantly, these clusters are completely absent in hearing impaired mice (von Hehn et al., 2004; **Figures 8B,C**).

Kv1 FAMILY I_{LVA} CHANNELS ACTIVATE CLOSE TO THE RESTING MEMBRANE POTENTIAL

Several different classes of potassium channels contribute to I_{LVA} currents in auditory brainstem neurons. In general, I_{LVA} potassium currents endow neurons with a low input resistance and fast time constant and, in the auditory brainstem, are prominent in neurons that lock their action potential precisely to incoming auditory stimuli, such as bushy cells of the AVCN, MNTB neurons and octopus cells in the posteroventral cochlear nucleus (Oertel, 1983; Smith and Rhode, 1987; Manis and Marx, 1991; Brew and Forsythe, 1995; Oertel et al., 2008; Rusznak et al., 2008). Because these currents start to activate significantly with small depolarizations from the resting potential, they limit the firing of action potentials in response to sustained depolarization or prolonged postsynaptic potentials (Manis and Marx, 1991; Schwarz and Puil, 1997; Rothman and Manis, 2003; Cao et al., 2007).

I_{LVA} currents also play a key role in neurons of the LSO and the MSO (Barnes-Davies et al., 2004; Svirskis et al., 2004; Mathews et al., 2010; Fischer et al., 2018; Nabel et al., 2019). In the MSO, I_{LVA} currents are present in proximal dendrites and the soma, and their activation by excitatory synaptic inputs in the dendrites shortens the duration of the excitatory postsynaptic potentials (EPSPs) as they propagate toward the soma (Mathews et al., 2010). This feature contributes to the ability of MSO neurons to resolve differences of the order of tens of microseconds in the time or arrival of binaural inputs. In the LSO, there is a tonotopic gradient of I_{LVA} currents, with high levels of I_{LVA} in neurons in the lateral, low frequency limb of this nucleus (Barnes-Davies et al., 2004). As in the AVCN and MNTB, this limits the firing of action potentials in response to sustained depolarization and serves to preserve timing information (Remme et al., 2014).

The dominant and best studied I_{LVA} currents in MNTB and LSO neurons are produced by the Kv1.1 and Kv1.2 voltage-dependent subunits (Dodson et al., 2002, 2003; Barnes-Davies et al., 2004; Gittelman and Tempel, 2006). These are expressed ubiquitously throughout the central and peripheral nervous system. In principal neurons of the MNTB, heteromeric Kv1.1/Kv1.2 channels are localized to the initial segment of the axon of postsynaptic neurons, where they provide a dominant component of the I_{LVA} current (Dodson et al., 2002). The activation of these channels ensures that a synaptic input or a sustained depolarization produces only a single action potential that is faithfully locked to the onset of the stimulus (Oertel, 1983; Smith and Rhode, 1987). This mode of response is essential for accurate localization of sounds in space. Thus, knockout of the gene for the Kv1.1 channel in mice increases the latency and jitter of sound-evoked action potentials in MNTB neurons (Kopp-Scheinpflug et al., 2003a; Gittelman and Tempel, 2006) and renders the mice unable to localize sounds (Karcz et al., 2011; Robbins and Tempel, 2012).

There is evidence that several other potassium channels may contribute to the I_{LVA} currents of MNTB postsynaptic

neurons. These include the voltage-dependent channel subunits Kv1.6 (Dodson et al., 2002) and Kv11 (Hardman and Forsythe, 2009), as well as $K_{Na}1.1$ and $K_{Na}1.2$ (also termed Slack and Slick), potassium channels activated by elevations of intracellular sodium (K_{Na} channels) (Bhattacharjee et al., 2002, 2005; Yang et al., 2007a).

The complement of I_{LVA} channels in the presynaptic calyx of Held terminals is slightly different from those in the postsynaptic cells. Kv1.2 homomers have been documented on the membrane of the presynaptic axons (Dodson et al., 2003). Kv1.3 channels are also localized to the plasma membrane of the presynaptic terminal and to small intracellular vesicles in the cytoplasm of the terminal (Gazula et al., 2010). Within this terminal, the role of I_{LVA} currents is likely to differ from in the postsynaptic cells because the firing pattern of the terminals is driven by action potentials generated at the cell body of the AVCN bushy cell. One such role for the Kv1.2 channels is to ensure that each incoming action potential from the axon triggers only a single spike at the terminal itself (Dodson et al., 2003).

Compared to the I_{HVA} channels, much less is known about factors that modulate Kv1 family channels in response to changes in the sound environment. These channels all have documented phosphorylation sites and experiments carried out largely using non-excitable cells suggest that their current amplitude and trafficking can be regulated by second messenger pathways (Vacher and Trimmer, 2011). For example, the amplitude of Kv1.1 currents has been reported to be regulated by the cyclic-AMP-dependent protein kinase (Winklhofer et al., 2003). Phosphorylation of serine 446 in Kv1.1 modulates its association with auxiliary subunits (Singer-Lahat et al., 1999). Phosphorylation of Kv1.2 regulates its intracellular trafficking (Yang et al., 2007b). PKC also regulates Kv1.1 through a mechanism that does not require its consensus phosphorylation sites (Boland and Jackson, 1999). Whether any of these mechanisms are engaged by stimulation of auditory neurons is, however, unknown.

POTENTIAL CONTRIBUTIONS OF K_{Na} CHANNELS TO I_{LVA} CURRENTS

The $K_{Na}1.1$ and $K_{Na}1.2$ channels are widely expressed in the nervous system, including auditory brainstem neurons (Bhattacharjee et al., 2002, 2003). Based on what is known about their biochemical and biophysical properties, it is possible to make some tentative predictions about how their regulation contributes to I_{LVA} currents in MNTB neurons and related cells. For example, the higher levels of firing induced by high intensity sounds will elevate intracellular sodium levels by sodium entry both through voltage-dependent sodium channels and ionotropic glutamate receptors on MNTB neurons. These would be expected to increase the activity of the sodium sensitive K_{Na} channels (Kaczmarek, 2013; Kaczmarek et al., 2017). Using brain slice preparations, it has been shown that drugs that promote the opening of K_{Na} channels increase the fidelity with which MNTB neurons lock their action potentials to a stimulus train (Yang et al., 2007a).

Phosphorylation also plays a direct role in regulating the amplitude of K_{Na} currents. For example, the $K_{Na}1.1$ channel has a serine residue (serine 407 in the rat channel) that, when phosphorylated by PKC, increases current amplitude (Santi et al., 2006; Barcia et al., 2012). If an increase in sound intensity in the physiological range causes dephosphorylation of this $K_{Na}1.1$ residue, as occurs for serine 503 in Kv3.1, this would enhance I_{LVA} current and aid temporal accuracy of firing. Such short term increases in I_{LVA} current in brainstem neurons could potentially contribute to improving the discrimination of sounds in a noisy environment, as occurs in the cocktail party effect (Zion Golumbic et al., 2013). Speculations on increases in I_{LVA} are limited by the fact that the effects of PKC activation on $K_{Na}1.1$ channels differ from those on $K_{Na}1.2$ channels or $K_{Na}1.1/K_{Na}1.2$ heteromeric channels (Santi et al., 2006). Human mutations in $K_{Na}1.1$ that increase current amplitude result in intellectual deficits so severe that they preclude the characterization of auditory function (Kim and Kaczmarek, 2014).

DIRECT EFFECTS OF FMRP BINDING TO I_{LVA} CHANNEL SUBUNITS

As described earlier, patients lacking the RNA-binding protein FMRP, are impaired in their ability to localize sounds and suffer from hyperacusis. At least part of this disability may result not only from the effects of FMRP on translation of mRNAs but because FMRP binds directly to some of the ion channels that provide I_{LVA} currents (McCullagh et al., 2020). Both Kv1.2 and $K_{Na}1.1$ channels have been found to exist in a protein complex with FMRP (Brown et al., 2010; Zhang et al., 2012; Yang et al., 2018). Interestingly, binding of FMRP to Kv1.2 has been found to require prior phosphorylation of the channel at serine residues in its cytoplasmic C-terminal domain (Yang et al., 2018). This interaction enhances the activity of Kv1.2 channels. Similarly, the interaction of FMRP with the cytoplasmic C-terminus of $K_{Na}1.1$ potently stimulates channel activity (Brown et al., 2010; Zhang et al., 2012). Consistent with these findings, the I_{LVA} component of potassium current is substantially reduced in MNTB neurons in a mouse model of Fragile X syndrome in which the gene for FMRP has been deleted (Brown et al., 2010; El-Hassar et al., 2019).

Like the mRNAs for the I_{HVA} channels Kv3.1 and Kv3.3, the mRNAs for Kv1.2 and $K_{Na}1.1$ are targets of FMRP (Darnell et al., 2011). One might expect, therefore, that in addition to the direct effects of loss of FMRP-binding on channel activity, the rates of synthesis of these two proteins in response to changes in the auditory environment may be compromised in animals and humans lacking FMRP. Nevertheless, whether I_{LVA} currents are altered by auditory stimuli, and the full time-course and mechanisms of such changes, are yet to be investigated.

A study has compared the relative amplitudes of I_{LVA} and I_{HVA} components of potassium current in MNTB neurons from wild type mice and those lacking FMRP (El-Hassar et al., 2019). Consistent with the data described in the preceding sections, the ratio of I_{HVA} to I_{LVA} is enhanced in the FMRP knockout animals.

Because increases in I_{HVA} enhance high frequency firing while I_{LVA} currents are required for temporal accuracy, this change in both components is consistent with both the hypersensitivity of Fragile X patients to loud sounds and their inability to localize sounds in space. Pharmacological agents that alter the voltage dependence of Kv3.1 channels, converting them from I_{HVA} to I_{LVA} channels, may provide a potential direction for helping to ameliorate the auditory phenotype of Fragile X patients (El-Hassar et al., 2019).

OTHER CHANNELS THAT MODULATE RESPONSE PROPERTIES

Several other types of potassium channels are expressed in the MNTB of different species and at different times in development. These include Kv4.3 subunits, which have been reported in the MNTB of mice but not that of rats (Johnston et al., 2008a). These subunits generate rapidly inactivating "A-type" potassium currents, which could potentially influence the timing of neuronal responses, but their precise role is not understood. Immunoreactivity for a second inactivating potassium channel subunit, Kv3.4, has been reported in the calyx of Held presynaptic terminals (Ishikawa et al., 2003), but mRNA for this subunit is absent in the MNTB (Choudhury et al., 2020), suggesting that Kv3.4 could be synthesized in the AVCN and then transported to the calyces. Because Kv3.4 is expressed in all major fiber tracts in the developing brain, it is possible that this subunit has a role in the navigation of the fibers of AVCN neurons early in development (Huang et al., 2012).

Finally, there exists a class of potassium currents that maintain the membrane potential of a neuron close its normal negative value (~ -45 to -75 mV). The basic concept of a leak potassium channel arose from finding that pharmacological block of most of the known categories of potassium channels does not abolish this resting potential. Thus, leak channels have a fixed open probability that does not change with voltage or rapid changes in intracellular Ca^{2+} or Na^{+}. Nearly all numerical simulations of the firing patterns of neurons incorporate such a voltage-independent leak current that reverses at a negative potential (Wang et al., 1998; Rothman and Manis, 2003; Song et al., 2005; Kaczmarek, 2012). Because leak potassium channels are open at all voltages, changes in such leak currents could potentially alter all aspects of neuronal firing, including the resting potential, the time constant and the threshold and height of action potentials. In many cases, however, the magnitude of leak currents is smaller than that of the other "active" potassium currents, which dominate these aspects. The subfamily of two pore domain potassium channels (K_{2P}, also termed KCNK) contains fifteen members, and has been found to give rise to leak-type K^{+} channels. In contrast to other potassium channels, in which four α-subunits assemble to form a functional channel, K_{2P} channels assemble as dimers (Niemeyer et al., 2016; Zuniga and Zuniga, 2016).

Expression of the mRNAs for several K_{2P} subunits has been documented in the auditory brainstem. These include $K_{2P}15$ (also termed TASK5), $K_{2P}1$ (TWIK-1), $K_{2P}12$ (THIK-2), $K_{2P}6$ (TWIK-2), $K_{2P}2$ (TREK-1), $K_{2P}10$(TREK-2), $K_{2P}4$ (TRAAK), and $K_{2P}9$ (TASK-3) in the rat cochlear nucleus (Holt et al., 2006). MNTB neurons express mRNA for $K_{2P}1$ (Kaczmarek et al., 2005), and the properties of the leak currents in MNTB neurons have been found to match that of $K_{2P}1$ (Berntson and Walmsley, 2008). Of particular interest is the $K_{2P}15$ (TASK5) subunit, which, except for a subset of cerebellar and olfactory neurons, is expressed almost exclusively in central auditory pathways (Karschin et al., 2001; Ehmann et al., 2013). When expressed by itself in heterologous expression systems, $K_{2P}15$ fails to generate currents (Karschin et al., 2001), suggesting functional channels in neurons represent a heteromer of $K_{2P}15$ with another K_{2P} subunit. Deafening by cochlear ablation produces a large sustained reduction in expression of $K_{2P}15$ in brainstem neurons and inferior colliculus, and also decreases expression of several other K_{2P} subunits (Holt et al., 2006; Cui et al., 2007; Dong et al., 2009). Acute shRNA-mediated knockdown of $K_{2P}15$ in auditory produces a depolarization of the resting membrane potential, increasing the width and latency of action potentials and enhanced firing in globular bushy cells of the cochlear nucleus, with smaller effects in MNTB neurons, which have lower $K_{2P}15$ expression (Saher, 2020). These findings suggest that signaling pathways that modify K_{2P} currents in auditory neurons could have major effects on their intrinsic excitability.

CONCLUSION

Research over the past two decades has amply demonstrated that the intrinsic excitability of neurons responsible for the very early stages of auditory processing are not fixed. The importance of the correct balance of ion channels is evident from human conditions such as SCA13 and Fragile X syndrome, which do not result in deafness but severely impair the ability to distinguish sounds in a noisy environment. Even rapid changes in the auditory environment can produce rapid and reversible posttranslational modifications of channels expressed in auditory brainstem nuclei such as the MNTB. Such rapid changes in excitability are certain to contribute to an alteration in the way in which sounds are processed when a person moves from a quiet environment to a noisy one, such as one in which the cocktail party effect is manifest. Longer-term mechanisms that depend on incoming sounds are required to maintain the appropriate balance of different ion channels. Some of these mechanisms involve the transcription and translation of mRNAs for channel subunits and are required for the correct distribution of several ion channels along the tonotopic axis of auditory nuclei. Such longer-term changes in ion channels may also contribute to experience-dependent differences in the auditory abilities of individuals, for example, the superior ability of orchestra conductors to localize sounds (Münte et al., 2001; Nager et al., 2003). Nevertheless, because of their experimental tractability, rats and mice have been the major species used for the study of potassium channels properties in auditory neurons. Future work will require the determination of how many of the findings can be generalized to other species, particularly those with hearing that more closely matches that of humans.

The correction of the abnormalities in auditory processing related to ion channels in humans may involve both pharmacological approaches to alter channel activity and genetic approaches that correct the underlying defects. For example, pharmacological agents that alter the gating of Kv3 family channels have the potential to correct the abnormal firing patters of auditory brainstem neurons in Fragile X syndrome (El-Hassar et al., 2019). Changing the levels of expression of either wild type or mutant ion channels can be achieved using antisense oligonucleotides, and these may come to be used as a therapy for genetic diseases such as SCA13 (Bushart and Shakkottai, 2019).

AUTHOR CONTRIBUTIONS

Both authors equally contributed to the drafting and writing of the final version of the manuscript and read and approved the final version of the manuscript.

FUNDING

Work in the authors' laboratory was supported by the NIH grants DC01919, NS102239, and NS111242 (to LK).

REFERENCES

Alexander, S. P. H., Mathie, A., Peters, J. A., Veale, E. L., Striessnig, J., Kelly, E., et al. (2019). THE CONCISE GUIDE TO PHARMACOLOGY 2019/20: ion channels. *Br. J. Pharmacol.* 176(Suppl. 1), S142–S228.

Arinami, T., Sato, M., Nakajima, S., and Kondo, I. (1988). Auditory brain-stem responses in the fragile X syndrome. *Am. J. Hum. Genet.* 43, 46–51.

Banks, M. I., and Smith, P. H. (1992). Intracellular recordings from neurobiotin-labeled cells in brain slices of the rat medial nucleus of the trapezoid body. *J. Neurosci.* 12, 2819–2837. doi: 10.1523/jneurosci.12-07-02819.1992

Barcia, G., Fleming, M. R., Deligniere, A., Gazula, V. R., Brown, M. R., Langouet, M., et al. (2012). De novo gain-of-function KCNT1 channel mutations cause malignant migrating partial seizures of infancy. *Nat. Genet.* 44, 1255–1259. doi: 10.1038/ng.2441

Barnes-Davies, M., Barker, M. C., Osmani, F., and Forsythe, I. D. (2004). Kv1 currents mediate a gradient of principal neuron excitability across the tonotopic axis in the rat lateral superior olive. *Eur. J. Neurosci.* 19, 325–333. doi: 10.1111/j.0953-816x.2003.03133.x

Baydyuk, M., Xu, J., and Wu, L. G. (2016). The calyx of Held in the auditory system: structure, function, and development. *Hear. Res.* 338, 22–31. doi: 10.1016/j.heares.2016.03.009

Berntson, A. K., and Walmsley, B. (2008). Characterization of a potassium-based leak conductance in the medial nucleus of the trapezoid body. *Hear. Res.* 244, 98–106. doi: 10.1016/j.heares.2008.08.003

Bhattacharjee, A., Gan, L., and Kaczmarek, L. K. (2002). Localization of the slack potassium channel in the rat central nervous system. *J. Comp. Neurol.* 454, 241–254. doi: 10.1002/cne.10439

Bhattacharjee, A., Joiner, W. J., Wu, M., Yang, Y., Sigworth, F. J., and Kaczmarek, L. K. (2003). Slick (Slo2.1), a rapidly-gating sodium-activated potassium channel inhibited by ATP. *J. Neurosci.* 23, 11681–11691. doi: 10.1523/jneurosci.23-37-11681.2003

Bhattacharjee, A., Von Hehn, C. A., Mei, X., and Kaczmarek, L. K. (2005). Localization of the Na+-activated K+ channel Slick in the rat central nervous system. *J. Comp. Neurol.* 484, 80–92. doi: 10.1002/cne.20462

Boland, L. M., and Jackson, K. A. (1999). Protein kinase C inhibits Kv1.1 potassium channel function. *Am. J. Physiol.* 277, C100–C110.

Borst, J. G., and Soria van Hoeve, J. (2012). The calyx of Held synapse: from model synapse to auditory relay. *Annu. Rev. Physiol.* 74, 199–224. doi: 10.1146/annurev-physiol-020911-153236

Brew, H. M., and Forsythe, I. D. (1995). Two voltage-dependent K+ conductances with complementary functions in postsynaptic integration at a central auditory synapse. *J. Neurosci.* 15, 8011–8022. doi: 10.1523/jneurosci.15-12-08011.1995

Brew, H. M., and Forsythe, I. D. (2005). Systematic variation of potassium current amplitudes across the tonotopic axis of the rat medial nucleus of the trapezoid body. *Hear. Res.* 206, 116–132. doi: 10.1016/j.heares.2004.12.012

Brown, M. R., El-Hassar, L., Zhang, Y., Alvaro, G., Large, C. H., and Kaczmarek, L. K. (2016). Physiological modulators of Kv3.1 channels adjust firing patterns of auditory brain stem neurons. *J. Neurophysiol.* 116, 106–121. doi: 10.1152/jn.00174.2016

Brown, M. R., Kronengold, J., Gazula, V. R., Chen, Y., Strumbos, J. G., Sigworth, F. J., et al. (2010). Fragile X mental retardation protein controls gating of the sodium-activated potassium channel Slack. *Nat. Neurosci.* 13, 819–821. doi: 10.1038/nn.2563

Brownell, W. E. (1975). Organization of the cat trapezoid body and the discharge characteristics of its fibers. *Brain Res.* 94, 413–433. doi: 10.1016/0006-8993(75)90226-7

Bushart, D. D., and Shakkottai, V. G. (2019). Ion channel dysfunction in cerebellar ataxia. *Neurosci. Lett.* 688, 41–48. doi: 10.1016/j.neulet.2018.02.005

Cao, X. J., Shatadal, S., and Oertel, D. (2007). Voltage-sensitive conductances of bushy cells of the Mammalian ventral cochlear nucleus. *J. Neurophysiol.* 97, 3961–3975. doi: 10.1152/jn.00052.2007

Carr, C. E., and MacLeod, K. M. (2010). Microseconds matter. *PLoS Biol.* 8:e1000405. doi: 10.1371/journal.pbio.1000405

Castren, M., Paakkonen, A., Tarkka, I. M., Ryynanen, M., and Partanen, J. (2003). Augmentation of auditory N1 in children with fragile X syndrome. *Brain Topogr.* 15, 165–171.

Choudhury, N., Linley, D., Richardson, A., Anderson, M., Robinson, S. W., Marra, V., et al. (2020). Kv3.1 and Kv3.3 subunits differentially contribute to Kv3 channels and action potential repolarization in principal neurons of the auditory brainstem. *J. Physiol.* 598, 2199–2222. doi: 10.1113/jp279668

Cui, Y. L., Holt, A. G., Lomax, C. A., and Altschuler, R. A. (2007). Deafness associated changes in two-pore domain potassium channels in the rat inferior colliculus. *Neuroscience* 149, 421–433. doi: 10.1016/j.neuroscience.2007.05.054

Darnell, J. C., Jensen, K. B., Jin, P., Brown, V., Warren, S. T., and Darnell, R. B. (2001). Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function. *Cell* 107, 489–499. doi: 10.1016/s0092-8674(01)00566-9

Darnell, J. C., Van Driesche, S. J., Zhang, C., Hung, K. Y., Mele, A., Fraser, C. E., et al. (2011). FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell* 146, 247–261. doi: 10.1016/j.cell.2011.06.013

Dash, P. K., Karl, K. A., Colicos, M. A., Prywes, R., and Kandel, E. R. (1991). cAMP response element-binding protein is activated by Ca2+/calmodulin- as well as cAMP-dependent protein kinase. *Proc. Natl. Acad. Sci. U.S.A.* 88, 5061–5065. doi: 10.1073/pnas.88.11.5061

De Rubeis, S., and Bagni, C. (2010). Fragile X mental retardation protein control of neuronal mRNA metabolism: Insights into mRNA stability. *Mol. Cell. Neurosci.* 43, 43–50. doi: 10.1016/j.mcn.2009.09.013

Desai, R., Kronengold, J., Mei, J., Forman, S. A., and Kaczmarek, L. K. (2008). Protein kinase C modulates inactivation of Kv3.3 channels. *J. Biol. Chem.* 283, 22283–22294. doi: 10.1074/jbc.m801663200

Dodson, P. D., Barker, M. C., and Forsythe, I. D. (2002). Two heteromeric Kv1 potassium channels differentially regulate action potential firing. *J. Neurosci.* 22, 6953–6961. doi: 10.1523/jneurosci.22-16-06953.2002

Dodson, P. D., Billups, B., Rusznak, Z., Szucs, G., Barker, M. C., and Forsythe, I. D. (2003). Presynaptic rat Kv1.2 channels suppress synaptic terminal hyperexcitability following action potential invasion. *J. Physiol.* 550, 27–33. doi: 10.1113/jphysiol.2003.046250

Dong, S., Mulders, W. H., Rodger, J., and Robertson, D. (2009). Changes in neuronal activity and gene expression in guinea-pig auditory brainstem after unilateral partial hearing loss. *Neuroscience* 159, 1164–1174. doi: 10.1016/j.neuroscience.2009.01.043

Ehmann, H., Hartwich, H., Salzig, C., Hartmann, N., Clement-Ziza, M., Ushakov, K., et al. (2013). Time-dependent gene expression analysis of the developing superior olivary complex. *J. Biol. Chem.* 288, 25865–25879. doi: 10.1074/jbc.m113.490508

Elezgarai, I., Diez, J., Puente, N., Azkue, J. J., Benitez, R., Bilbao, A., et al. (2003). Subcellular localization of the voltage-dependent potassium channel Kv3.1b in postnatal and adult rat medial nucleus of the trapezoid body. *Neuroscience* 118, 889–898. doi: 10.1016/s0306-4522(03)00068-x

El-Hassar, L., Song, L., Tan, W. J. T., Large, C. H., Alvaro, G., Santos-Sacchi, J., et al. (2019). Modulators of Kv3 potassium channels rescue the auditory function of fragile X mice. *J. Neurosci.* 39, 4797–4813. doi: 10.1523/jneurosci.0839-18.2019

Ferron, L. (2016). Fragile X mental retardation protein controls ion channel expression and activity. *J. Physiol.* 594, 5861–5867. doi: 10.1113/jp270675

Fischer, L., Leibold, C., and Felmy, F. (2018). Resonance properties in auditory brainstem neurons. *Front. Cell. Neurosci.* 12:8. doi: 10.3389/fncel.2018.00008

Forsythe, I. D., and Barnes-Davies, M. (1993). The binaural auditory pathway: membrane currents limiting multiple action potential generation in the rat medial nucleus of the trapezoid body. *Proc. Biol. Sci.* 251, 143–150. doi: 10.1098/rspb.1993.0021

Gan, L., and Kaczmarek, L. K. (1998). When, where, and how much? Expression of the Kv3.1 potassium channel in high-frequency firing neurons. *J. Neurobiol.* 37, 69–79. doi: 10.1002/(sici)1097-4695(199810)37:1<69::aid-neu6>3.0.co;2-6

Gan, L., Perney, T. M., and Kaczmarek, L. K. (1996). Cloning and characterization of the promoter for a potassium channel expressed in high frequency firing neurons. *J. Biol. Chem.* 271, 5859–5865. doi: 10.1074/jbc.271.10.5859

Gazula, V. R., Strumbos, J. G., Mei, X., Chen, H., Rahner, C., and Kaczmarek, L. K. (2010). Localization of Kv1.3 channels in presynaptic terminals of brainstem auditory neurons. *J. Comp. Neurol.* 518, 3205–3220. doi: 10.1002/cne.22393

Gittelman, J. X., and Tempel, B. L. (2006). Kv1.1-containing channels are critical for temporal precision during spike initiation. *J. Neurophysiol.* 96, 1203–1214. doi: 10.1152/jn.00092.2005

Gonzalez, G. A., and Montminy, M. R. (1989). Cyclic AMP stimulates somatostatin gene transcription by phosphorylation of CREB at serine 133. *Cell* 59, 675–680. doi: 10.1016/0092-8674(89)90013-5

Grigg, J. J., Brew, H. M., and Tempel, B. L. (2000). Differential expression of voltage-gated potassium channel genes in auditory nuclei of the mouse brainstem. *Hear. Res.* 140, 77–90. doi: 10.1016/s0378-5955(99)00187-2

Grothe, B., and Pecka, M. (2014). The natural history of sound localization in mammals–a story of neuronal inhibition. *Front. Neural Circuits* 8:116. doi: 10.3389/fncir.2014.00116

Grothe, B., Pecka, M., and Mcalpine, D. (2010). Mechanisms of sound localization in mammals. *Physiol. Rev.* 90, 983–1012. doi: 10.1152/physrev.00026.2009

Hall, S. S., Walter, E., Sherman, E., Hoeft, F., and Reiss, A. L. (2009). The neural basis of auditory temporal discrimination in girls with fragile X syndrome. *J. Neurodev. Disord.* 1, 91–99. doi: 10.1007/s11689-009-9007-x

Hardman, R. M., and Forsythe, I. D. (2009). Ether-a-go-go-related gene K+ channels contribute to threshold excitability of mouse auditory brainstem neurons. *J. Physiol.* 587, 2487–2497. doi: 10.1113/jphysiol.2009.170548

Haykin, S., and Chen, Z. (2005). The cocktail party problem. *Neural Comput.* 17, 1875–1902.

He, K., McCord, M. C., Hartnett, K. A., and Aizenman, E. (2015). Regulation of pro-apoptotic phosphorylation of Kv2.1 K+ channels. *PLoS One* 10:e0129498. doi: 10.1371/journal.pone.0129498

Held, H. (1893). Die centrale Gehörleitung. *Arch. Anat. Physiol. Anat. Abt.* 17, 201–248.

Hermann, J., Pecka, M., Von Gersdorff, H., Grothe, B., and Klug, A. (2007). Synaptic transmission at the calyx of Held under in vivo like activity levels. *J. Neurophysiol.* 98, 807–820. doi: 10.1152/jn.00355.2007

Holt, A. G., Asako, M., Duncan, R. K., Lomax, C. A., Juiz, J. M., and Altschuler, R. A. (2006). Deafness associated changes in expression of two-pore domain potassium channels in the rat cochlear nucleus. *Hear. Res.* 21, 146–153. doi: 10.1016/j.heares.2006.03.009

Huang, C. Y., Chu, D., Hwang, W. C., and Tsaur, M. L. (2012). Coexpression of high-voltage-activated ion channels Kv3.4 and Cav1.2 in pioneer axons during pathfinding in the developing rat forebrain. *J. Comp. Neurol.* 520, 3650–3672. doi: 10.1002/cne.23119

Ikematsu, N., Dallas, M. L., Ross, F. A., Lewis, R. W., Rafferty, J. N., David, J. A., et al. (2011). Phosphorylation of the voltage-gated potassium channel Kv2.1 by AMP-activated protein kinase regulates membrane excitability. *Proc. Natl. Acad. Sci. U.S.A.* 108, 18132–18137. doi: 10.1073/pnas.1106201108

Ishikawa, T., Nakamura, Y., Saitoh, N., Li, W. B., Iwasaki, S., and Takahashi, T. (2003). Distinct roles of Kv1 and Kv3 potassium channels at the calyx of Held presynaptic terminal. *J. Neurosci.* 23, 10445–10453. doi: 10.1523/jneurosci.23-32-10445.2003

Johnston, J., Griffin, S. J., Baker, C., and Forsythe, I. D. (2008a). Kv4 (A-type) potassium currents in the mouse medial nucleus of the trapezoid body. *Eur. J. Neurosci.* 27, 1391–1399. doi: 10.1111/j.1460-9568.2008.06116.x

Johnston, J., Griffin, S. J., Baker, C., Skrzypiec, A., Chernova, T., and Forsythe, I. D. (2008b). Initial segment Kv2.2 channels mediate a slow delayed rectifier and maintain high frequency action potential firing in medial nucleus of the trapezoid body neurons. *J. Physiol.* 586, 3493–3509. doi: 10.1113/jphysiol.2008.153734

Joris, P. X., and Trussell, L. O. (2018). The calyx of Held: a hypothesis on the need for reliable timing in an intensity-difference encoder. *Neuron* 100, 534–549. doi: 10.1016/j.neuron.2018.10.026

Joris, P. X., Smith, P. H., and Yin, T. C. (1994). Enhancement of neural synchronization in the anteroventral cochlear nucleus. II. Responses in the tuning curve tail. *J. Neurophysiol.* 71, 1037–1051. doi: 10.1152/jn.1994.71.3.1037

Kaczmarek, L. K. (2006). Non-conducting functions of voltage-gated ion channels. *Nat. Rev. Neurosci.* 7, 761–771. doi: 10.1038/nrn1988

Kaczmarek, L. K. (2012). Gradients and modulation of K(+) channels optimize temporal accuracy in networks of auditory neurons. *PLoS Comput. Biol.* 8:e1002424. doi: 10.1371/journal.pcbi.1002424

Kaczmarek, L. K. (2013). Slack, slick and sodium-activated potassium channels. *ISRN Neurosci.* 2013:354262.

Kaczmarek, L. K. (2019). "Extraction of auditory information by modulation of neuronal ion channels," in *The Oxford Handbook of the Auditory Brainstem*, ed. K. Kandler (Oxford: OUP), doi: 10.1093/ oxfordhb/ 9780190849061.013.23

Kaczmarek, L. K. (2020). "Excitable membrane properties of neurons," in *The Oxford Handbook of Neuronal Ion Channels*, ed. A. Bhattacharjee (Oxford: OUP), doi: 10.1093/oxfordhb/ 9780190669164.013.20

Kaczmarek, L. K., Aldrich, R. W., Chandy, K. G., Grissmer, S., Wei, A. D., and Wulff, H. (2017). International union of basic and clinical pharmacology. C. Nomenclature and properties of calcium-activated and sodium-activated potassium channels. *Pharmacol. Rev.* 69, 1–11. doi: 10.1124/pr.116.012864

Kaczmarek, L. K., and Zhang, Y. (2017). Kv3 channels: enablers of rapid firing, neurotransmitter release, and neuronal endurance. *Physiol. Rev.* 97, 1431–1468. doi: 10.1152/physrev.00002.2017

Kaczmarek, L. K., Bhattacharjee, A., Desai, R., Gan, L., Song, P., Von Hehn, C. A., et al. (2005). Regulation of the timing of MNTB neurons by short-term and long-term modulation of potassium channels. *Hear. Res.* 206, 133–145. doi: 10.1016/j.heares.2004.11.023

Kaczmarek, L. K., Wu, X.-S., Subramanian, S., Xia, J., Zhang, Y., El-Hassar, L., et al. (2019). The Kv3.3 potassium channel controls endocytosis by organizing the actin cytoskeleton at nerve terminals. *Neurosci. Abstr.* (CD ROM), 121.29.

Kanemasa, T., Gan, L., Perney, T. M., Wang, L. Y., and Kaczmarek, L. K. (1995). Electrophysiological and pharmacological characterization of a mammalian Shaw channel expressed in NIH 3T3 fibroblasts. *J. Neurophysiol.* 74, 207–217. doi: 10.1152/jn.1995.74.1.207

Karcz, A., Hennig, M. H., Robbins, C. A., Tempel, B. L., Rubsamen, R., and Kopp-Scheinpflug, C. (2011). Low-voltage activated Kv1.1 subunits are crucial for the processing of sound source location in the lateral superior olive in mice. *J. Physiol.* 589, 1143–1157. doi: 10.1113/jphysiol.2010.203331

Karschin, C., Wischmeyer, E., Preisig-Muller, R., Rajan, S., Derst, C., Grzeschik, K. H., et al. (2001). Expression pattern in brain of TASK-1, TASK-3, and a tandem pore domain K(+) channel subunit, TASK-5, associated with the central auditory nervous system. *Mol. Cell. Neurosci.* 18, 632–648. doi: 10.1006/mcne.2001.1045

Kim, G. E., and Kaczmarek, L. K. (2014). Emerging role of the KCNT1 Slack channel in intellectual disability. *Front. Cell. Neurosci.* 8:209. doi: 10.3389/fncel.2014.00209

Kopp-Scheinpflug, C., Fuchs, K., Lippe, W. R., Tempel, B. L., and Rubsamen, R. (2003a). Decreased temporal precision of auditory signaling in Kcna1-null mice: an electrophysiological study in vivo. *J. Neurosci.* 23, 9199–9207. doi: 10.1523/jneurosci.23-27-09199.2003

Kopp-Scheinpflug, C., Lippe, W. R., Dorrscheidt, G. J., and Rubsamen, R. (2003b). The medial nucleus of the trapezoid body in the gerbil is more than a relay: comparison of pre- and postsynaptic activity. *J. Assoc. Res. Otolaryngol.* 4, 1–23. doi: 10.1007/s10162-002-2010-5

Kopp-Scheinpflug, C., Tolnai, S., Malmierca, M. S., and Rubsamen, R. (2008). The medial nucleus of the trapezoid body: comparative physiology. *Neuroscience* 154, 160–170. doi: 10.1016/j.neuroscience.2008.01.088

Labro, A. J., Priest, M. F., Lacroix, J. J., Snyders, D. J., and Bezanilla, F. (2015). Kv3.1 uses a timely resurgent K(+) current to secure action potential repolarization. *Nat. Commun.* 6:10173.

Leao, K. E., Leao, R. N., Deardorff, A. S., Garrett, A., Fyffe, R., and Walmsley, B. (2010). Sound stimulation modulates high-threshold K(+) currents in mouse auditory brainstem neurons. *Eur. J. Neurosci.* 32, 1658–1667. doi: 10.1111/j.1460-9568.2010.07437.x

Leao, R. M. (2019). The ion channels and synapses responsible for the physiological diversity of mammalian lower brainstem auditory neurons. *Hear. Res.* 376, 33–46. doi: 10.1016/j.heares.2018.12.011

Leao, R. N., Naves, M. M., Leao, K. E., and Walmsley, B. (2006a). Altered sodium currents in auditory neurons of congenitally deaf mice. *Eur. J. Neurosci.* 24, 1137–1146. doi: 10.1111/j.1460-9568.2006.04982.x

Leao, R. N., Sun, H., Svahn, K., Berntson, A., Youssoufian, M., Paolini, A. G., et al. (2006b). Topographic organization in the auditory brainstem of juvenile mice is disrupted in congenital deafness. *J. Physiol.* 571, 563–578. doi: 10.1113/jphysiol.2005.098780

Lee, A., Fakler, B., Kaczmarek, L. K., and Isom, L. L. (2014). More than a pore: ion channel signaling complexes. *J. Neurosci.* 34, 15159–15169. doi: 10.1523/jneurosci.3275-14.2014

Li, W., Kaczmarek, L. K., and Perney, T. M. (2001). Localization of two high-threshold potassium channel subunits in the rat central auditory system. *J. Comp. Neurol.* 437, 196–218. doi: 10.1002/cne.1279

Liu, S. J., and Kaczmarek, L. K. (1998). The expression of two splice variants of the Kv3.1 potassium channel gene is regulated by different signaling pathways. *J. Neurosci.* 18, 2881–2890. doi: 10.1523/jneurosci.18-08-02881.1998

Liu, S. Q., and Kaczmarek, L. K. (1998). Depolarization selectively increases the expression of the Kv3.1 potassium channel in developing inferior colliculus neurons. *J. Neurosci.* 18, 8758–8769. doi: 10.1523/jneurosci.18-21-08758.1998

Lorteije, J. A., Rusu, S. I., Kushmerick, C., and Borst, J. G. (2009). Reliability and precision of the mouse calyx of Held synapse. *J. Neurosci.* 29, 13770–13784. doi: 10.1523/jneurosci.3285-09.2009

Luneau, C. J., Williams, J. B., Marshall, J., Levitan, E. S., Oliva, C., Smith, J. S., et al. (1991). Alternative splicing contributes to K+ channel diversity in the mammalian central nervous system. *Proc. Natl. Acad. Sci. U.S.A.* 88, 3932–3936. doi: 10.1073/pnas.88.9.3932

Macica, C. M., and Kaczmarek, L. K. (2001). Casein kinase 2 determines the voltage dependence of the Kv3.1 channel in auditory neurons and transfected cells. *J. Neurosci.* 21, 1160–1168. doi: 10.1523/jneurosci.21-04-01160.2001

Macica, C. M., Von Hehn, C. A., Wang, L. Y., Ho, C. S., Yokoyama, S., Joho, R. H., et al. (2003). Modulation of the kv3.1b potassium channel isoform adjusts the fidelity of the firing pattern of auditory neurons. *J. Neurosci.* 23, 1133–1141. doi: 10.1523/jneurosci.23-04-01133.2003

Manis, P. B., and Marx, S. O. (1991). Outward currents in isolated ventral cochlear nucleus neurons. *J. Neurosci.* 11, 2865–2880. doi: 10.1523/jneurosci.11-09-02865.1991

Mathews, P. J., Jercog, P. E., Rinzel, J., Scott, L. L., and Golding, N. L. (2010). Control of submillisecond synaptic timing in binaural coincidence detectors by K(v)1 channels. *Nat. Neurosci.* 13, 601–609. doi: 10.1038/nn.2530

McCullagh, E. A., Rotschafer, S. E., Auerbach, B. D., Klug, A., Kaczmarek, L. K., Cramer, K. S., et al. (2020). Mechanisms underlying auditory processing deficits in Fragile X syndrome. *FASEB J.* 34, 3501–3518. doi: 10.1096/fj.201902435r

McDermott, J. H. (2009). The cocktail party problem. *Curr. Biol.* 19, R1024–R1027.

Middlebrooks, J. C., Nick, H. S., Subramony, S. H., Advincula, J., Rosales, R. L., Lee, L. V., et al. (2013). Mutation in the Kv3.3 voltage-gated potassium channel causing spinocerebellar ataxia 13 disrupts sound-localization mechanisms. *PLoS One* 8:e76749. doi: 10.1371/journal.pone.0076749

Moore, M. J., and Caspary, D. M. (1983). Strychnine blocks binaural inhibition in lateral superior olivary neurons. *J. Neurosci.* 3, 237–242. doi: 10.1523/jneurosci.03-01-00237.1983

Münte, T. F., Kohlmetz, C., Nager, W., and Altenmuller, E. (2001). Superior auditory spatial tuning in conductors. *Nature* 409:580. doi: 10.1038/35054668

Nabel, A. L., Callan, A. R., Gleiss, S. A., Kladisios, N., Leibold, C., and Felmy, F. (2019). Distinct distribution patterns of potassium channel sub-units in somato-dendritic compartments of neurons of the medial superior olive. *Front. Cell. Neurosci.* 13:38. doi: 10.3389/fncel.2019.00038

Nager, W., Kohlmetz, C., Altenmuller, E., Rodriguez-Fornells, A., and Münte, T. F. (2003). The fate of sounds in conductors' brains: an ERP study. *Brain Res. Cogn. Brain Res.* 17, 83–93. doi: 10.1016/s0926-6410(03)00083-1

Niemeyer, M. I., Cid, L. P., Gonzalez, W., and Sepulveda, F. V. (2016). Gating, regulation, and structure in K2P K+ channels: in varietate concordia? *Mol. Pharmacol.* 90, 309–317. doi: 10.1124/mol.116.103895

Oertel, D. (1983). Synaptic responses and electrical properties of cells in brain slices of the mouse anteroventral cochlear nucleus. *J. Neurosci.* 3, 2043–2053. doi: 10.1523/jneurosci.03-10-02043.1983

Oertel, D., Shatadal, S., and Cao, X. J. (2008). In the ventral cochlear nucleus Kv1.1 and subunits of HCN1 are colocalized at surfaces of neurons that have low-voltage-activated and hyperpolarization-activated conductances. *Neuroscience* 154, 77–86. doi: 10.1016/j.neuroscience.2008.01.085

Perney, T. M., and Kaczmarek, L. K. (1997). Localization of a high threshold potassium channel in the rat cochlear nucleus. *J. Comp. Neurol.* 386, 178–202. doi: 10.1002/(sici)1096-9861(19970922)386:2<178::aid-cne2>3.0.co;2-z

Perney, T. M., Marshall, J., Martin, K. A., Hockfield, S., and Kaczmarek, L. K. (1992). Expression of the mRNAs for the Kv3.1 potassium channel gene in the adult and developing rat brain. *J. Neurophysiol.* 68, 756–766. doi: 10.1152/jn.1992.68.3.756

Rasmussen, H. B., and Trimmer, J. S. (2019). "The voltage-dependent K+ channel family," in *The Oxford Handbook of Neuronal Ion Channels*, ed. A. Bhattacharjee (Oxford: OUP), doi: 10.1093/oxfordhb/ 9780190669164.013.1

Remme, M. W., Donato, R., Mikiel-Hunter, J., Ballestero, J. A., Foster, S., Rinzel, J., et al. (2014). Subthreshold resonance properties contribute to the efficient coding of auditory spatial cues. *Proc. Natl. Acad. Sci. U.S.A.* 111, E2339–E2348.

Richter, J. D., Bassell, G. J., and Klann, E. (2015). Dysregulation and restoration of translational homeostasis in fragile X syndrome. *Nat. Rev. Neurosci.* 16, 595–605. doi: 10.1038/nrn4001

Robbins, C. A., and Tempel, B. L. (2012). Kv1.1 and Kv1.2: similar channels, different seizure models. *Epilepsia* 53(Suppl. 1), 134–141. doi: 10.1111/j.1528-1167.2012.03484.x

Roberts, J., Hennon, E. A., Anderson, K., Roush, J., Gravel, J., Skinner, M., et al. (2005). Auditory brainstem responses in young males with Fragile X syndrome. *J. Speech Lang. Hear. Res.* 48, 494–500. doi: 10.1044/1092-4388(2005/034)

Rojas, D. C., Benkers, T. L., Rogers, S. J., Teale, P. D., Reite, M. L., and Hagerman, R. J. (2001). Auditory evoked magnetic fields in adults with fragile X syndrome. *Neuroreport* 12, 2573–2576. doi: 10.1097/00001756-200108080-00056

Rothman, J. S., and Manis, P. B. (2003). The roles potassium currents play in regulating the electrical activity of ventral cochlear nucleus neurons. *J. Neurophysiol.* 89, 3097–3113. doi: 10.1152/jn.00127.2002

Rotschafer, S. E., and Razak, K. A. (2014). Auditory processing in fragile x syndrome. *Front. Cell. Neurosci.* 8:19. doi: 10.3389/fncel.2014.00019

Rusznak, Z., Bakondi, G., Pocsai, K., Por, A., Kosztka, L., Pal, B., et al. (2008). Voltage-gated potassium channel (Kv) subunits expressed in the rat cochlear nucleus. *J. Histochem. Cytochem.* 56, 443–465. doi: 10.1369/jhc.2008.950303

Saher, M. H. (2020). *Investigation of The Physiological Role Of The Electrically Silent K2P Subunit Task5 in The Auditory Brainstem*. Ph.D. thesis, University of Heidelberg, Germany.

Santi, C. M., Ferreira, G., Yang, B., Gazula, V. R., Butler, A., Wei, A., et al. (2006). Opposite regulation of Slick and Slack K+ channels by neuromodulators. *J. Neurosci.* 26, 5059–5068. doi: 10.1523/jneurosci.3372-05.2006

Schneggenburger, R., and Forsythe, I. D. (2006). The calyx of Held. *Cell Tissue Res.* 326, 311–337.

Schwarz, D. W., and Puil, E. (1997). Firing properties of spherical bushy cells in the anteroventral cochlear nucleus of the gerbil. *Hear. Res.* 114, 127–138. doi: 10.1016/s0378-5955(97)00162-7

Scott, L. L., Mathews, P. J., and Golding, N. L. (2010). Perisomatic voltage-gated sodium channels actively maintain linear synaptic integration in principal neurons of the medial superior olive. *J. Neurosci.* 30, 2039–2050. doi: 10.1523/jneurosci.2385-09.2010

Singer-Lahat, D., Dascal, N., and Lotan, I. (1999). Modal behavior of the Kv1.1 channel conferred by the Kvbeta1.1 subunit and its regulation by dephosphorylation of Kv1.1. *Pflugers Arch.* 439, 18–26. doi: 10.1007/s004240051123

Smith, P. H., and Rhode, W. S. (1987). Characterization of HRP-labeled globular bushy cells in the cat anteroventral cochlear nucleus. *J. Comp. Neurol.* 266, 360–375. doi: 10.1002/cne.902660305

Song, P., and Kaczmarek, L. K. (2006). Modulation of Kv3.1b potassium channel phosphorylation in auditory neurons by conventional and novel protein kinase C isozymes. *J. Biol. Chem.* 281, 15582–15591. doi: 10.1074/jbc.m512866200

Song, P., Yang, Y., Barnes-Davies, M., Bhattacharjee, A., Hamann, M., Forsythe, I. D., et al. (2005). Acoustic environment determines phosphorylation state of the Kv3.1 potassium channel in auditory neurons. *Nat. Neurosci.* 8, 1335–1342. doi: 10.1038/nn1533

St Clair, D. M., Blackwood, D. H., Oliver, C. J., and Dickens, P. (1987). P3 abnormality in fragile X syndrome. *Biol. Psychiatry* 22, 303–312. doi: 10.1016/0006-3223(87)90148-x

Steinert, J. R., Kopp-Scheinpflug, C., Baker, C., Challiss, R. A. J., Mistry, R., Haustein, M. D., et al. (2008). Nitric oxide is a volume transmitter regulating postsynaptic excitability at a glutamatergic synapse. *Neuron* 60, 642–656.

Steinert, J. R., Robinson, S. W., Tong, H., Haustein, M. D., Kopp-Scheinpflug, C., and Forsythe, I. D. (2011). Nitric oxide is an activity-dependent regulator of target neuron intrinsic excitability. *Neuron* 71, 291–305.

Strumbos, J. G., Brown, M. R., Kronengold, J., Polley, D. B., and Kaczmarek, L. K. (2010a). Fragile X mental retardation protein is required for rapid experience-dependent regulation of the potassium channel Kv3.1b. *J. Neurosci.* 30, 10263–10271.

Strumbos, J. G., Polley, D. B., and Kaczmarek, L. K. (2010b). Specific and rapid effects of acoustic stimulation on the tonotopic distribution of Kv3.1b potassium channels in the adult rat. *Neuroscience* 167, 567–572.

Svirskis, G., Kotak, V., Sanes, D. H., and Rinzel, J. (2004). Sodium along with low-threshold potassium currents enhance coincidence detection of subthreshold noisy signals in MSO neurons. *J. Neurophysiol.* 91, 2465–2473.

Taschenberger, H., and von Gersdorff, H. (2000). Fine-tuning an auditory synapse for speed and fidelity: developmental changes in presynaptic waveform, EPSC kinetics, and synaptic plasticity. *J. Neurosci.* 20, 9162–9173. doi: 10.1523/JNEUROSCI.20-24-09162.2000

Taskin, B., Von Schoubye, N. L., Sheykhzade, M., Bastlund, J. F., Grunnet, M., and Jespersen, T. (2015). Biophysical characterization of KV3.1 potassium channel activating compounds. *Eur. J. Pharmacol.* 758, 164–170. doi: 10.1016/j.ejphar.2015.03.061

Tollin, D. J. (2003). The lateral superior olive: a functional role in sound source localization. *Neuroscientist* 9, 127–143. doi: 10.1177/1073858403252228

Tollin, D. J. (2009). "Development of sound localization," in *Oxford Handbook of Developmental Behavioral Neuroscience*, eds M. S. Blumberg, J. H. Freeman, and S. R. Robinson (Oxford: OUP)

Tong, H., Kopp-Scheinpflug, C., Pilati, N., Robinson, S. W., Sinclair, J. L., Steinert, J. R., et al. (2013). Protection from noise-induced hearing loss by Kv2.2 potassium currents in the central medial olivocochlear system. *J. Neurosci.* 33, 9113–9121. doi: 10.1523/JNEUROSCI.5043-12.2013

Tong, H., Steinert, J. R., Robinson, S. W., Chernova, T., Read, D. J., Oliver, D. L., et al. (2010). Regulation of Kv channel expression and neuronal excitability in rat medial nucleus of the trapezoid body maintained in organotypic culture. *J. Physiol.* 588, 1451–1468. doi: 10.1113/jphysiol.2009.186676

Vacher, H., and Trimmer, J. S. (2011). Diverse roles for auxiliary subunits in phosphorylation-dependent regulation of mammalian brain voltage-gated potassium channels. *Pflugers Arch.* 462, 631–643. doi: 10.1007/s00424-011-1004-8

Van der Molen, M. J., Van Der Molen, M. W., Ridderinkhof, K. R., Hamel, B. C., Curfs, L. M., and Ramakers, G. J. (2012). Auditory and visual cortical activity during selective attention in fragile X syndrome: a cascade of processing deficiencies. *Clin. Neurophysiol.* 123, 720–729. doi: 10.1016/j.clinph.2011.08.023

von Hehn, C. A., Bhattacharjee, A., and Kaczmarek, L. K. (2004). Loss of Kv3.1 tonotopicity and alterations in cAMP response element-binding protein signaling in central auditory neurons of hearing impaired mice. *J. Neurosci.* 24, 1936–1940. doi: 10.1523/JNEUROSCI.4554-03.2004

Wang, L. Y., and Kaczmarek, L. K. (1998). High-frequency firing helps replenish the readily releasable pool of synaptic vesicles. *Nature* 394, 384–388. doi: doi.org/10.1038/28645

Wang, L. Y., Gan, L., Forsythe, I. D., and Kaczmarek, L. K. (1998). Contribution of the Kv3.1 potassium channel to high-frequency firing in mouse auditory neurones. *J. Physiol.* 509(Pt 1), 183–194. doi: 0.1111/j.1469-7793.1998.183bo.x

Winklhofer, M., Matthias, K., Seifert, G., Stocker, M., Sewing, S., Herget, T., et al. (2003). Analysis of phosphorylation-dependent modulation of Kv1.1 potassium channels. *Neuropharmacology* 44, 829–842. doi: 10.1016/S0028-3908(03)00070-4

Yang, B., Desai, R., and Kaczmarek, L. K. (2007a). Slack and Slick K(Na) channels regulate the accuracy of timing of auditory neurons. *J. Neurosci.* 27, 2617–2627. doi: 10.1523/JNEUROSCI.5308-06.2007

Yang, J. W., Vacher, H., Park, K. S., Clark, E., and Trimmer, J. S. (2007b). Trafficking-dependent phosphorylation of Kv1.2 regulates voltage-gated potassium channel cell surface expression. *Proc. Natl. Acad. Sci. U.S.A.* 104, 20055–20060. doi: 10.1073/pnas.0708574104

Yang, Y. M., Arsenault, J., Bah, A., Krzeminski, M., Fekete, A., Chao, O. Y., et al. (2018). Identification of a molecular locus for normalizing dysregulated GABA release from interneurons in the Fragile X brain. *Mol. Psychiatry* 25, 2017–2035. doi: 10.1038/s41380-018-0240-0

Yin, T. C. T., Smith, P. H., and Joris, P. X. (2019). Neural mechanisms of binaural processing in the auditory brainstem. *Compr. Physiol.* 9, 1503–1575. doi: 10.1002/cphy.c180036

Zhang, Y., and Kaczmarek, L. K. (2016). Kv3.3 potassium channels and spinocerebellar ataxia. *J. Physiol.* 594, 4677–4684. doi: 10.1113/JP271343

Zhang, Y., Brown, M. R., Hyland, C., Chen, Y., Kronengold, J., Fleming, M. R., et al. (2012). Regulation of neuronal excitability by interaction of fragile X mental retardation protein with slack potassium channels. *J. Neurosci.* 32, 15318–15327. doi: 10.1523/JNEUROSCI.2162-12.2012

Zhang, Y., Zhang, X. F., Fleming, M. R., Amiri, A., El-Hassar, L., Surguchev, A. A., et al. (2016). Kv3.3 channels bind Hax-1 and Arp2/3 to assemble a stable local actin network that regulates channel gating. *Cell* 165, 434–448. doi: 10.1016/j.cell.2016.02.009

Zion Golumbic, E. M., Ding, N., Bickel, S., Lakatos, P., Schevon, C. A., Mckhann, G. M., et al. (2013). Mechanisms underlying selective neuronal tracking of attended speech at a "cocktail party". *Neuron* 77, 980–991. doi: 10.1016/j.neuron.2012.12.037

Zuniga, L., and Zuniga, R. (2016). Understanding the Cap Structure in K2P Channels. *Front. Physiol.* 7:228. doi: 10.3389/fphys.2016.00228

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor is currently organizing a Research Topic with one of the authors LK.

frontiers in Neuroscience

REVIEW
published: 19 March 2021
doi: 10.3389/fnins.2021.610957

Impaired Binaural Hearing in Adults: A Selected Review of the Literature

*Frederick J. Gallun**

Oregon Hearing Research Center, Oregon Health and Science University, Portland, OR, United States

Despite over 100 years of study, there are still many fundamental questions about binaural hearing that remain unanswered, including how impairments of binaural function are related to the mechanisms of binaural hearing. This review focuses on a number of studies that are fundamental to understanding what is known about the effects of peripheral hearing loss, aging, traumatic brain injury, strokes, brain tumors, and multiple sclerosis (MS) on binaural function. The literature reviewed makes clear that while each of these conditions has the potential to impair the binaural system, the specific abilities of a given patient cannot be known without performing multiple behavioral and/or neurophysiological measurements of binaural sensitivity. Future work in this area has the potential to bring awareness of binaural dysfunction to patients and clinicians as well as a deeper understanding of the mechanisms of binaural hearing, but it will require the integration of clinical research with animal and computational modeling approaches.

Keywords: lateralization, localization, binaural, spatial hearing, impairment, auditory

OPEN ACCESS

Edited by:
Elizabeth Anne McCullagh, Oklahoma State University, United States

Reviewed by:
Tobias Neher, University of Southern Denmark, Denmark
Takako Fujioka, Stanford University, United States

***Correspondence:**
Frederick J. Gallun
gallunf@ohsu.edu

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience

Received: *28 September 2020*
Accepted: *19 February 2021*
Published: *19 March 2021*

Citation:
Gallun FJ (2021) Impaired Binaural Hearing in Adults: A Selected Review of the Literature. Front. Neurosci. 15:610957. doi: 10.3389/fnins.2021.610957

INTRODUCTION

The ability to process the information available in pressure waves arriving at the two ears ("binaural hearing") is available to living creatures ranging from insects (Hedwig and Stumpner, 2016) to humans (for review see Stecker and Gallun, 2012). Binaural hearing has obvious defensive and predatory advantages, as well as serving an important communicative function. Consequently, dysfunction of the binaural system can reduce the ability to navigate the auditory scene (Gallun and Best, 2020). This review of binaural impairment in adult human listeners will start with an overview of the history of the area and the current model of the binaural system. After surveying a variety of methods of characterizing binaural impairment, the literature on patient groups will be selectively reviewed. This review of the patient literature will be divided into two main sections. The first will focus on patients with conductive hearing loss (CHL) and sensorineural hearing loss (SNHL). The second will turn to those patients with central dysfunction, for whom detection of pure tones is often normal or near normal but for whom binaural sensitivity has been shown to be impaired. The final section will address the many opportunities for additional studies that are made clear by what this review is and is not able to tell us about the mechanisms of binaural impairment and about the abilities of various patient groups to make use of the auditory spatial cues available in the environment.

Binaural function has been studied clinically from as far back as 1876 (Pierce, 1901). The importance of studying abnormal auditory function has been known from the very first studies of binaural hearing. The work of Venturi (1796) was described by Wade and Deutsch (2008) and Stecker and Gallun (2012). Venturi hypothesized, based on his comparisons of monaural and binaural listening, that the relative intensities at the two ears (the "interaural level difference"; ILD),

give rise to the ability to localize sounds in space. Rayleigh (1907) was the first to go beyond the ILD and show that differences in the time of arrival of a sound at the two ears (the "interaural time difference"; ITD) is also a potent binaural cue for localization of sound sources. Even before the neural sites of binaural interaction had been identified, clinician scientists such as Greene (1929) and Walsh (1957) were applying psychoacoustical techniques to study the binaural abilities of their patients and using those results to form hypotheses about the underlying anatomy and physiology.

Since these early clinical studies, much has been revealed about the anatomy and physiology of the binaural system. **Figure 1** provides a schematic diagram of some of the main aspects of what Stecker and Gallun (2012) determined to be the currently accepted model of how the binaural system is connected physiologically, with an emphasis on the ascending binaural system. For further details of the basic architecture of this system see Yin (2002). Essential to the functioning of these pathways is high-fidelity transduction from the outer ear to the lateral and medial nuclei of the superior olivary complex (LSO; MSO) via the medial and lateral nuclei of the trapezoidal body (MNTB; LNTB). The reason for this is that we now know that both nuclei depend on microsecond (μs) temporal precision for the comparison of neural impulses from the two ears. As described in detail in Stecker and Gallun (2012), the discharge rates of LSO and MSO neurons can vary across nearly their entire response range when presented with only 1 millisecond (ms) of delay between the left and right ear inputs. Such precision, which is essential for allowing the system to be sensitive to ITDs as low as 10 μs (see Stecker and Gallun, 2012, Table 14-1) depends critically on the transformation of instantaneous pressure levels at the eardrum to spikes on the auditory nerve (Carr and Konishi, 1990). This transformation, which is known as neural phase-locking, allows the system to encode both the temporal fine structure (TFS) and the temporal envelope structure (TES) of the stimuli at each ear (Dreyer and Delgutte, 2006), from which LSO and MSO are able to accurately extract ILD and ITD information. For this reason, diseases that degrade neural transduction, such as multiple sclerosis (MS), are particularly likely to result in binaural impairment even when pure-tone thresholds are within the normal range.

Degradation of the transduction of pressure waves in the air to fluid motion in the cochlea, known as "conductive hearing loss" (CHL), can arise from blockages in the ear canal, perforations of the tympanic membrane, blockages of the middle ear, or dysfunction of the bones of the middle ear (the "ossicles") due to breakage or stiffening. All of these could delay or distort neural phase-locking and introduce either systematic or random error into the binaural analysis, especially if there was a substantial asymmetry between the two ears. The next site of potential dysfunction is the cochlea itself, impairment of which is known as SNHL. SNHL is defined as involving the basilar membrane, the fluids of the cochlear ducts, the organ of Corti, inner, and outer hair cells, and the stria vascularis, which is responsible for the blood supply to the cochlea. Damage or dysfunction at any of these sites has the potential to reduce both the amount of neural signaling as well as the accuracy of phase-locking to the pressure waves reaching the outer ear.

While the clinical diagnosis of CHL and SNHL relies upon the detection of pure tones presented via earphones (air-conduction, AC) or bone-vibration (bone-conduction, BC), it does not involve any measure of phase-locking. If AC and BC measurements reveal similar thresholds, but detection requires higher tone levels than is normal for a tone of that frequency, as specified by international standards (e.g., ANSI/ASA S3.6, 2018), then CHL is ruled out and SNHL is suspected. This renders diagnosis of SNHL and CHL insufficiently precise for assessing the potential impact on the binaural system. All impairment beyond the cochlea, including dysfunction of the auditory nerve, is known as retrocochlear hearing loss (RHL), the modern diagnosis of which depends largely on imaging, although it can also involve the auditory brainstem response (ABR), which assesses neural timing. The vast majority of RHL diagnoses are due to vestibular schwannoma (also called acoustic neuroma), which is a benign tumor that grows on the vestibular portion of the eighth nerve. As it grows, this tumor damages the nerve and, if allowed to grow large enough, can damage the cochlear nucleus (CN) and other brainstem structures. For further details on the causes, diagnosis, and categorization of the types and severity of hearing losses, see Katz (2014).

As can be seen in **Figure 1**, there are multiple sites along the neural auditory pathway where damage or disease could result in binaural impairment. The first is on the auditory nerve itself, where signal transduction and transmission can be reduced by loss or dysfunction of synapses and/or auditory nerve fibers. Even if the signals arriving at the synapse of the auditory nerve with the CN are not degraded or reduced, there is the possibility of dysfunction within the CN or in the signaling pathway to the trapezoidal body, either due to damage to those pathways or demyelination, which could impede or delay neural impulses and introduce random errors into binaural comparisons. In addition to the trapezoidal body and the superior olivary complex, the dorsal nucleus of the lateral lemniscus (DNLL), the inferior colliculus (IC), the medial geniculate body of the thalamus (MGB), and auditory cortex (AC) all are involved in conveying and processing binaural information. As such, damage or delays at any of these sites could degrade the transformation of binaural information into spatial maps and the ability to assign spatial locations to perceived objects.

METHODS OF CHARACTERIZING BINAURAL IMPAIRMENT

All of the earliest work on impaired localization (reviewed in Pierce, 1901) appears to have relied upon either anecdotal reports or examinations of the ability to identify the location of sounds in a test room. One of the earliest methods of quantifying localization ability beyond simply determining whether it was accurate or inaccurate, was that of Jongkees and Van der Veer (1957), who presented sounds from one of eighteen loudspeakers and participants pointed to the perceived location with their eyes closed. Response deviations were compared to those of a group of participants with no known peripheral or central pathology. Based on whether the deviations were within this

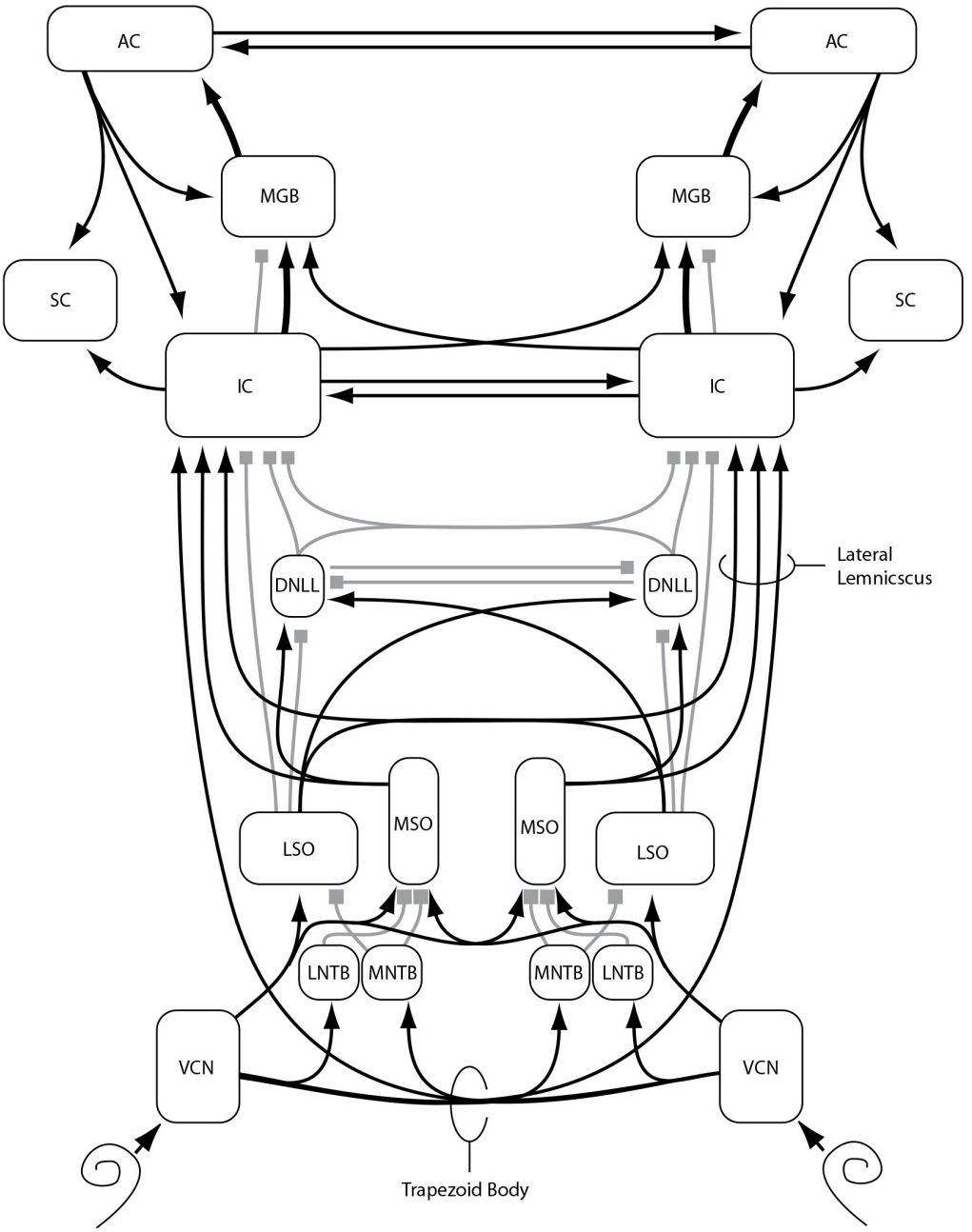

FIGURE 1 | Major nuclei (boxes) and primarily excitatory (lines with arrows) or inhibitory (lines with squares) interconnections of the ascending auditory pathway, with an emphasis on binaural function and connectivity, as described in the text: VCN: ventral cochlear nucleus, MNTB: medial nucleus of the trapezoid body, LNTB: lateral nucleus of the trapezoid body, MSO: medial nucleus of the superior olivary complex, LSO: lateral nucleus of the superior olivary complex, DNLL: dorsal nucleus of the lateral lemniscus, IC: inferior colliculus, SC: superior colliculus, MGB: medial geniculate body of the thalamus, AC: auditory cortex. Additional nuclei, projections, and subdivisions are omitted for clarity. Reproduced from Stecker and Gallun (2012), in Translational Perspectives in Auditory Neuroscience: Normal Aspects of Hearing (p. 395) by Tremblay, K., & Burkard, R. Copyright© 2012 Plural Publishing, Inc. All rights reserved. Used with permission.

range or not, the patients were categorized as having "normal" or "pathological" localization. Häusler et al. (1983) conducted the first localization experiments using discrimination tasks, in which two intervals are presented, only one of which contains a target. Discrimination is preferable to methods of adjustment or identification because it allows sensitivity to a stimulus to be measured independently of the expectations or willingness of the observer to make a particular response (Green and Swets, 1974). Using a loudspeaker array in an anechoic chamber, a two-interval forced-choice procedure was used to test the minimum audible angle (MAA; Mills, 1958) in the horizontal and vertical planes. Akeroyd and Whitmer (2016) reviewed 29 studies that used a

variety of localization tasks, most of which involved pointing or identifying source locations rather than discrimination. Both types of tasks can be done either with real speakers (either in an array with fixed locations or on a movable boom) or using a virtual acoustical simulation (VAS; Brungart et al., 2017). To create a VAS, head-related transfer functions (HRTFs) are imposed on the test stimuli, allowing the acoustic cues associated with a given spatial location of the stimulus to be presented over headphones (Wightman and Kistler, 1989). An example of a speaker array that was used for localization experiments with a closed set of fixed loudspeakers is shown in **Figure 2**.

In addition to tests of localization, it is also desirable to separate out the various acoustical cues and test sensitivity to each independently. The first tests of binaural sensitivity to ITD and ILD were conducted by Greene (1929), using a custom-made device that allowed ITD to be manipulated by changing the length of a column of water through which the sound was conducted, and thus the speed of travel through the fixed-length tube. ITD was not measurable with this device, but deviations from normal ITD sensitivity could be detected. Similarly, attenuation of the signal reaching one ear or the other allowed ILD to be manipulated in a manner that allowed abnormal sensitivity to be detected. Using headphones and modern electronics, Häusler et al. (1983) estimated the Just Noticeable Difference (JND) in ITD and ILD. The JNDs were measured with stimuli set to equal "sensation level" (SL), which is defined as a given level above detection threshold, thereby controlling for differences in hearing thresholds between ears. This basic approach has been used many times since (Hawkins and Wightman, 1980; Smoski and Trahiotis, 1986).

Another approach to measuring binaural ability, inspired by the work of Licklider (1948), is the binaural masking level difference (MLD). The MLD is defined as improved performance on a masked tone detection or speech identification task that is associated with the imposition of interaural differences on either the target or the masker. Melnick and Bilger (1965) conducted the first study of the MLD in patients, using a speech stimulus. Olsen et al. (1976) and Olsen and Noffsinger (1976), were the first to measure the MLD for a 500 Hz pure tone, which is more common in modern studies than is the use of speech targets.

FIGURE 2 | Example of a loudspeaker array used to test localization ability by identification of the loudspeaker from which a test signal has been presented. Such arrays can also be used to test spatial release from masking with speech or other stimuli. See text for experimental details. Reproduced with permission from Brungart et al. (2017). Copyright 2017, Acoustical Society of America.

A more recent approach to behavioral testing of binaural function involves spatial release from masking (SRM; Marrone et al., 2008), which is similar to the speech MLD but involves a loudspeaker array such as that shown in **Figure 2**, or a VAS. The first approaches (Duquesnoy, 1983; Bronkhorst and Plomp, 1989; Peissig and Kollmeier, 1997; Arbogast et al., 2005) presented a target sound from one location and a masking sound from another location, as well as varying the interfering sounds to include competing speech. However, these studies suffered from the confound that when the masker was presented at a single location, different signal to noise ratios (SNRs) were available at the two ears. This in turn leads to the availability of a "better-ear" listening strategy, which means that performance may improve even if the listener has no binaural sensitivity whatsoever. Using a method suggested by the manipulations and better-ear calculations of Hawley et al. (2004); Marrone et al. (2008) demonstrated that the better-ear effect can be eliminated (at least on a long-term basis) by displacing two maskers symmetrically to the left and right of the target. Gallun et al. (2013) introduced a VAS version of this test and a testing procedure based on the single descending track or "progressive tracking" method. This procedure is very fast (under 10 min) but is better suited for detecting abnormal performance than for obtaining precise measurements of threshold (Gallun et al., 2015). Ellinger et al. (2017) took advantage of the VAS to present processed speech signals in which ITDs and ILDs were manipulated independently and thus could be either reinforcing or conflicting, as Colburn and Durlach (1965) had done using an MLD paradigm.

An additional method for testing binaural sensitivity that has gained popularity recently is similar to the MLD, in that it involves the detection of interaural phase differences (IPDs), but instead of detecting a signal in noise, the task asks the listener to report directly on their binaural percept. The earliest work (Green et al., 1976; Witton et al., 2000) involved presenting a static tone to one ear and a tone to which frequency modulation (FM) had been applied to the other ear. It was found that for people with normal hearing, the presence of FM in the target stimulus was detectable at lower modulation depth when the FM resulted in IPDs than when the same FM was presented monaurally. Grose and Mamo (2012) extended this paradigm by presenting FM diotically (same FM at both ears) or dichotically (FM reversed in phase at the two ears). An example of a dichotic FM stimulus is shown in **Figure 3**.

In addition to behavioral methods, researchers have also used neurophysiological responses from cortical neurons to compare binaural responses in patients to those found in control participants. That binaural signals produce different neurophysiological responses than do diotic signals has been known for many years (Butler and Kluskens, 1971; Fowler and Mikami, 1992), but only recently have these responses been measured in patient groups. Ross et al. (2007), using magnetoencephalography (MEG), measured the auditory evoked responses (P1-N1-P2 complex) in response to binaural changes in the middle of ongoing signals, and later researchers

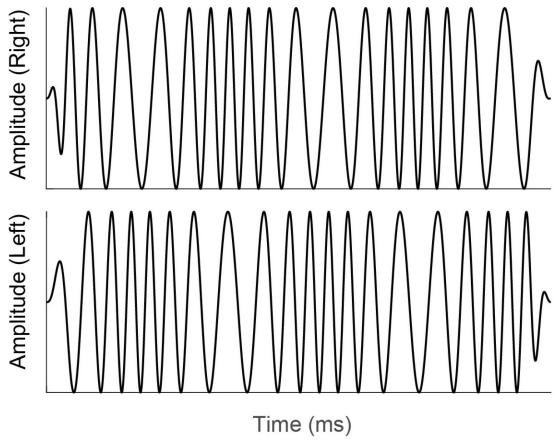

FIGURE 3 | Schematic diagram of the time-amplitude waveforms of a dichotic FM stimulus. For illustration purposes, the carrier frequency has been reduced and modulation depth increased from the values that would be used experimentally. Note that while there is no onset or offset difference in amplitude or phase, the phases of the signals at the left and right ears (top and bottom waveforms) are continually changing, resulting in an interaural phase difference that changes over time and an interaural time difference that shifts from left-leading to right-leading and back again. Reproduced under Creative Commons reuse license from Koerner et al. (2020).

(Papesh et al., 2017; Eddins and Eddins, 2018) extended this to electroencephalography (EEG). The P1-N1-P2 complex, which is measured in the time domain, arises at multiple levels of the thalamus and AC (Haywood et al., 2015). More recently (Haywood et al., 2015; Anderson et al., 2018; Vercammen et al., 2018; Koerner et al., 2020) researchers have used sequences of stimuli that changed rapidly (<7 Hz) in binaural configuration and measured the interaural phase modulation following response (IPM-FR). These responses are believed to be generated from the same sites as the P1-N1-P2 complex, but the steady-state nature allows them to be examined in the frequency domain. The amplitude of the response is used to quantify the degree to which changes in neurophysiological responses are correlated with the changes in binaural configuration. **Figure 4A** shows the stimuli used to generate the IPM-FR, which is shown in **Figure 4C**. The stimuli shown in **Figure 4A** are also similar to those used to generate the P1-N1-P2 complex, which is shown in **Figure 4B**. **Figure 4D** shows the response to a diotic stimulus, which does not generate the IPM-FR. Thus, the comparison of **Figures 4C,D** shows how to identify the IPM-FR peak in **Figure 4C**. Further examples of the P1-N1-P2 complex can be seen in **Figure 12**.

STUDIES ON BINAURAL FUNCTION WITH PARTICIPANTS WITH PERIPHERAL HEARING LOSS

The section on methods of characterizing binaural impairment describes a range of techniques that have all been applied to listeners with peripheral hearing loss (CHL and/or SNHL). This section will review a selection of those studies, focusing on the earliest studies that pioneered the methods and then moving on to those studies with the largest number of participants and the most comprehensive methods. While tentative conclusions are drawn in some cases, the field of clinical research in the area of binaural dysfunction is still developing its evidence base. For this reason, the emphasis of this, as of the later sections, is as much on what has been done as on what has been learned.

Peripheral Loss and Localization and Lateralization

Pierce (1901) describes a number of experiments using sophisticated equipment for testing the localization abilities of those with normal hearing, but only anecdotal evidence and simple localization tasks are described in the sections on people with impaired hearing. In one of the earliest descriptions of a localization experiment in people with hearing loss, Greene (1929) conducted experiments on eight patients with unilateral or bilateral CHL caused by otitis media. Using a "short-circuited binaural stethoscope" composed of section of rubber tubing connected to the two ear-pieces of a stethoscope, he measured how far from midline he had to tap a pencil on the tubing before patients could detect the displacement of the taps. All of these patients were able to detect an average displacement from midline of only 2.6 cm, which was not different from the detection thresholds of a normal-hearing control group.

Jongkees and Van der Veer (1957) tested 61 patients with hearing loss using the loudspeaker-based localization methods described in section "Methods of Characterizing Binaural Impairment." Listeners were divided into seven groups based on etiology of hearing impairment [chronic otitis media, atresia, SNHL, otosclerosis (unoperated and operated), patients after "reconstructive radical mastoid surgery" or tympanoplasty, unilateral total deafness], and compared the results to the localization abilities of 40 people with normal hearing. Across multiple groups of patients, 72% of those patients tested (44/61) had "pathological" localization, while only 18% (11/61) reported difficulties when asked about localization in their daily lives. Otosclerosis, with or without surgery, and atresia were both associated with abnormal directional hearing in all patients, while in every other group at least a third of the patients had localization functions that were statistically indistinguishable from those in the normal hearing group.

Jongkees and Van der Veer (1957) also measured pure-tone detection thresholds by employing both air- and bone-conducted audiometry at 1,000 Hz, but found that this was not a very reliable indicator of who would be able to perform their localization task. As discussed by Durlach et al. (1981), there is no discussion of whether or not head movements were allowed, and duration of disease is not reported. Indeed, the discussion of the role of head movements as a cue to localization for people with unilateral deafness in Jongkees and Van der Veer (1958) suggests that this cue was available and may explain some of the variability in performance among the patients tested. So, it is possible that some of the patients with better localization had been living with the disease for years and had learned to localize using

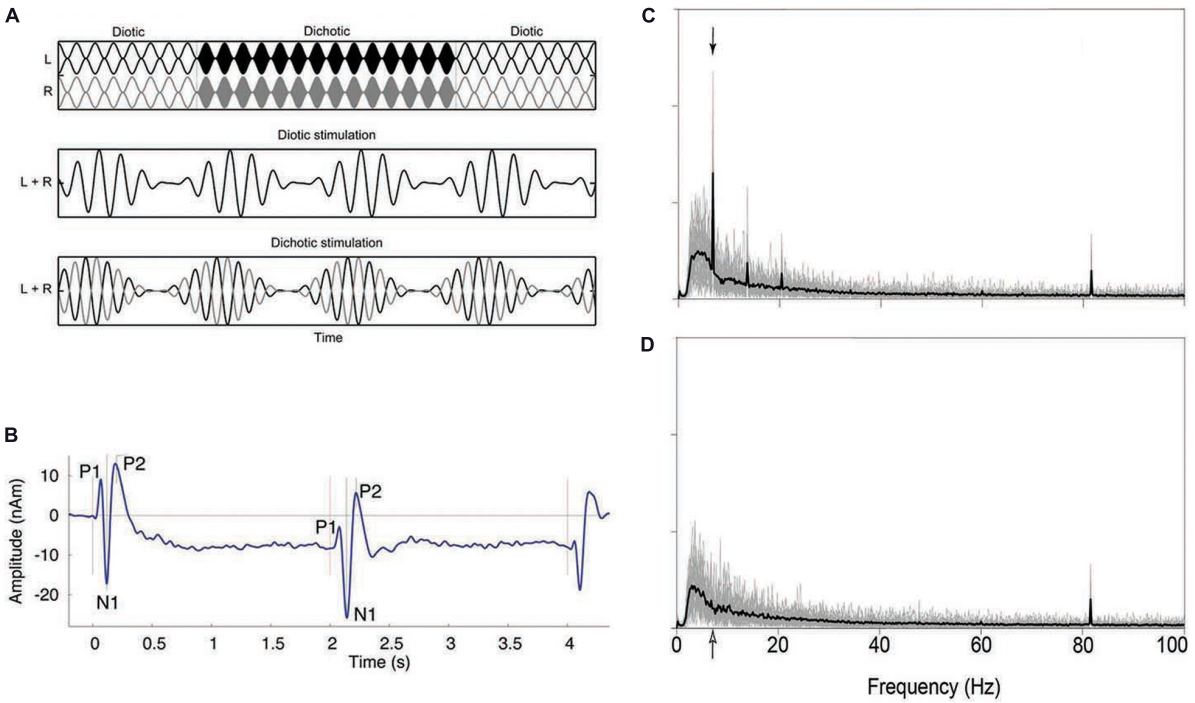

FIGURE 4 | Example of a stimulus that shifts from diotic to dichotic **(A)** and the evoked response generated in the time domain (P1-N1-P2 complex; **B**) and frequency domain (IPM-FR; **C,D**). See text for further details. Note that the stimulus shown shifts from diotic to dichotic and back to diotic, while the stimulus that would be used to evoke the P1-N1-P2 complex shown in panel **(B)** would only shift once, from diotic to dichotic, at the temporal midpoint of the stimulus. The stimulus used to generate the IPM-FR would contain many such alternations, at a characteristic rate, usually between 5 and 10 Hz. The arrows in panels **(C,D)** indicate the frequency at which the stimulus used to generate the IPM-FR alternated from diotic to dichotic, which in this case was 6.8 Hz. Additional evoked responses shown in panel **(B)** indicate the onset and the offset of the signal, while those in panels **(C,D)** indicate the response to the amplitude modulation rate (81.6 Hz) of the 500 Hz carrier. Additional low-frequency peaks in panel **(C)** represent aliasing at integer multiples of the IPM rate of 6.8 Hz. Panel **(D)** shows the response to a diotic stimulus and thus does not contain the peaks indicating the presence of the IPM-FR but does show the response to the modulation of the carrier amplitude at 81.6 Hz. Reproduced with permission from Ross et al. (2007); Vercammen et al. (2018), and Koerner et al. (2020). Vercammen et al. copyright 2018, Sage Publications. Ross et al. copyright Journal of Neuroscience. Koerner et al. reused under Creative Commons license.

monaural cues and/or head movements, while some of the patients with worse localization were newly suffering and had not developed these skills.

The first tests of clinical patients' binaural function using modern methods, in which sensitivity is dissociated from response bias, occurred in the 1980s. Prior to this time, all of the tests involved methods that were susceptible to response bias. For example, the localization of single sounds through identifying a location or adjusting knob can be influenced by expectations of where the sounds are likely to appear. Similarly, pressing a button when a sound is presented or repeating a spoken word depends on the willingness of the listener to report what they experienced. Discrimination tasks, in which two intervals are presented, only one of which contains a target, allows sensitivity to a stimulus to be measured independently of the expectations or willingness of the observer with regard to making a response (Green and Swets, 1974). Häusler et al. (1983) used discrimination tasks to measure both localization and sensitivity to interaural differences. For the localization task, which measured the MAA, or the ability to distinguish two loudspeaker locations, the participant was asked to discriminate two 1-s bursts of a broadband noise at a test location from two 1-s bursts presented at a reference location. Reference locations were either in front ("front-referenced MAA") or to the side ("side-referenced MAA"), and test locations were either displaced horizontally or vertically. The experimenter varied the size of the differences adaptively, based on past performance, in an attempt to find the value that led to 80% correct performance. No feedback was given to the participant.

Häusler et al. (1983) reported data from 49 patients with peripheral hearing loss and 39 normal-hearing control listeners. Of the patients, 14 had bilateral SNHL, 17 had conductive losses, 9 had one deaf ear (no behavioral response to sound), and 9 had unilateral loss due to Ménière's disease (for details on the audiological configurations associated with the disease, see Belinchon et al., 2011). A main finding of this study in terms of localization abilities of these patients was that, as had been found with earlier work, even in participants with similar etiologies, the ability to detect pure-tones was not a useful predictor of binaural function. It was revealed, however, that high-frequency hearing loss was associated with impairments in the ability to make vertical discriminations. In addition, the patients with SNHL and poor speech understanding were impaired on the vertical MAA and the side-referenced horizontal

MAA, while the participants with similar audiograms and good speech understanding had good performance on all the tasks. The authors speculated that spectral discrimination deficits were probably responsible for the relationship between localization and speech discrimination, but as only broadband noise was tested, it was impossible to know how participants would have performed with narrowband stimuli.

Noble et al. (1994, 1997) also found only weak relationships between pure-tone detection ability and horizontal localization and replicated the finding that the lack of audibility associated with high-frequency hearing loss reduces access to spectral cues important for vertical localization and front-back discrimination. However, they were unable to replicate the relationship observed by Häusler et al. (1983) between localization and speech in noise ability. Abel et al. (2000) and Dobreva et al. (2011), although focused on the issue of aging, also reported that even relatively mild high-frequency hearing loss interferes with vertical localization. Neither study was able to clearly show how the detection of pure-tones or aging is related to horizontal localization, despite the fact that both lead to increases in between-subject variability.

To better understand the sources of between-subject variance in localization, Neher et al. (2011) tested the role of cognition and two measures of auditory processing ability (monaural spectral ripple discrimination and binaural TFS sensitivity) in a group of 23 older listeners (aged 60–78 years; mean of 67 years) with audiometric thresholds outside the normal range (pure tone average, "PTA" of 27–53 dB; mean of 41 dB HL) and a group of 8 younger listeners (aged 20–44 years; mean of 35 years; thresholds of 20 dB HL or better below 6 kHz). While the older listeners did more poorly than did the younger listeners on a loudspeaker identification task in an anechoic chamber with a speaker array with 15° horizontal separations between speakers, none of the other tests predicted performance. Neither age nor PTA was significantly correlated with localization performance.

Brungart et al. (2017) compared localization in anechoic and virtual conditions and found that those with hearing loss suffered an overall reduction in performance, which was associated both with increased age and pure-tone thresholds. Best et al. (2011) tested hearing-impaired listeners in both a quiet condition and in a condition with interfering sounds and found that while they performed similarly to normally hearing controls in quiet, performance was worse by about 7° in the presence of interference. Buchholz and Best (2020) followed up on this experiment using a more realistic listening environment and found that four of the fifteen subjects with hearing loss had particularly poor localization in quiet and in the presence of masking sounds, and that the worst performers were those with low-frequency hearing loss. This is an important caveat to the general finding of a lack of a relationship between pure-tone thresholds and localization, which should be specifically examined in future work.

While this is only a sample of the data that have been collected on localization abilities of those with peripheral impairment, the interested reader is referred to Akeroyd and Whitmer (2016), who reviewed 29 studies of bilateral SNHL conducted between 1983 and 2014. **Figure 5**, which is reproduced from that chapter, summarizes the results of those studies by plotting the within-study differences in localization accuracy between the normally hearing and hearing impaired groups. Positive values, which represent the majority of the data, indicate worse acuity for those with peripheral impairment. Akeroyd and Whitmer (2016) concluded that, while the effects of age and hearing loss are difficult to separate, it seems likely that hearing loss results in about a 5° decrease in left–right localization accuracy. Furthermore, the size of the relationship between localization accuracy and hearing loss is fairly weak, amounting to a correlation of about 0.40.

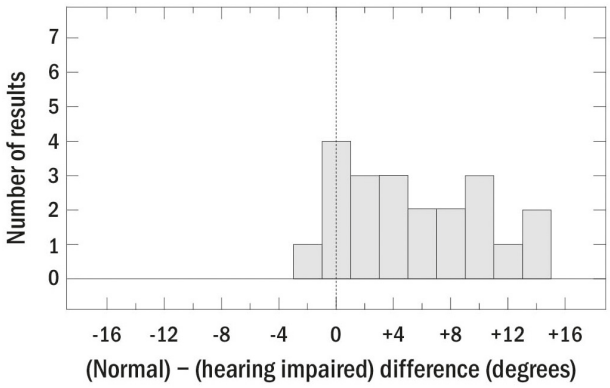

FIGURE 5 | Summary plot of the differences between localization accuracy for normally hearing and hearing-impaired listeners for 29 studies review by Akeroyd and Whitmer (2016). The dashed line indicates similar acuity for those with and without peripheral impairment. Reproduced with permission from Akeroyd and Whitmer (2016). Copyright 2016, Springer Nature.

Overall, the studies described above, as well as those reviewed by Akeroyd and Whitmer (2016), emphasize that while it is clear that hearing loss has a negative effect on localization, the audiogram is not sufficient for characterizing the mechanisms underlying the effect, at least for those with high-frequency hearing loss. Further work on these relationships, especially for those with low-frequency hearing loss and asymmetrical losses, is likely to be most productive when performed in combination with, or when informed by, animal and/or computational modeling. While low correlations and small effect sizes can be important for understanding the mechanisms and effects of peripheral hearing loss on localization ability, it is difficult to obtain significant results without running very large samples. Furthermore, the low correlations suggest that a predictive model of individual localization performance relevant for diagnosis and rehabilitation in a clinical setting will need to consider the influence of other factors such as age, cognition, and the integrity of the brainstem and auditory cortical pathways. For example, and as mentioned above, it is quite likely that for a patient with a particular hearing loss, localization ability may improve over time as they gain more awareness of localization cues and strategies that are specific to their binaural abilities. One way of getting more mechanistic insight into the ability of individual listeners to use spatial cues is to test sensitivity to ITD and ILD independently rather than through localization tasks.

Peripheral Loss and Binaural Sensitivity to Interaural Differences

Häusler et al. (1983), in addition to measuring the MAA, measured the JND for ITDs and ILDs. For the interaural JND task, the participant was asked to discriminate two 1-s bursts of a broadband noise to which an ITD or ILD had been applied from two 1-s bursts of a diotic noise. A main finding was that interaural differences in time and level were independently impaired by various types of loss and, as with localization, the most severe binaural impairment was observed in those with extreme unilateral losses. Furthermore, those people with SNHL and poor speech understanding (who were impaired on the vertical MAA and the side-referenced horizontal MAA) had normal JNDs for time and intensity, as did the participants with similarly impaired audiograms and good speech understanding.

Hawkins and Wightman (1980); Koehnke et al. (1986), and Smoski and Trahiotis (1986) also used forced-choice methods to test sensitivity to a set of different binaural cues and/or different stimuli for small groups of people with hearing loss. In each case, despite careful training, large numbers of trials, and careful control of stimulus and response variables, there was a trend toward worse performance in the group with hearing loss, but invariably there were some who still performed in or near the normal range. Gabriel et al. (1992); Koehnke et al. (1995), and Smith-Olinde et al. (2004) measured sensitivity to ITD and ILD with a variety of carriers and reference conditions and found substantial binaural impairment in their participants with reduced sensitivity to pure tones, but were unable to find specific relationships between interaural sensitivity and values of the audiogram. **Figure 6** reproduces data from Smith-Olinde et al. (2004) for ITD and ILD sensitivity as a function of SL. Neither the normal-hearing nor the hearing-impaired listeners appear to have JNDs that are predictable from SL alone.

While asking listeners to report the location of sounds presented over headphones is a direct method of measuring binaural sensitivity, the MLD has the advantage of simplifying the task to one of detection of a signal in noise, which does not require introspection about binaural percepts and thus may lead to more reliable performance. Melnick and Bilger (1965) conducted one of the first studies of sensitivity to interaural differences in people with peripheral hearing loss and did so by measuring the MLD for a speech target. Inspired by the work of Licklider (1948), they conducted a detailed investigation of the MLD with 61 patients and 14 normally hearing listeners. Despite the presence of hearing losses that ranged from mild to severe and symmetrical to asymmetrical, all of those tested were able to make use of the IPD to obtain better speech intelligibility. Unfortunately for later researchers, the data were only compared in terms of group means (which showed no differences) and rather than reporting the MLD for each listener, patients were simply categorized as "normal" or "abnormal" with reference to normal performance.

Bocca and Antonelli (1976) performed a similar study, but rather than categorical reports of performance, they compared MLDs across groups of patients and control listeners. The MLD was determined for speech signals presented in white noise. Interaural conditions were tested in which the speech and noise were both diotic (N_0S_0), or the noise was diotic and speech was delayed to the left or right ear by 0.8 ms (N_0S_T). The MLD was defined as the difference between the levels that produced the same percentage of correct responses for N_0S_0 and N_0S_T, based on the psychometric functions obtained. Data were compared to that of a control group of 20 listeners with normal hearing thresholds, who had an average MLD of roughly 7.5 dB.

For ten listeners with symmetrical CHL (PTA of at least 60 dB HL in both ears), the MLD was only slightly lower than that of the controls (7 dB). Ten listeners with asymmetrical conductive loss (PTA of 60 dB HL in one ear, PTA of 0–20 dB HL in the other) had an average MLD of 4.5 dB when the poorer ear was leading in time, which was further reduced to 2.5 dB when the better ear was leading. For those with Ménière's disease (essentially flat unilateral losses of at least 50 dB HL), the MLD was roughly –0.5 dB when the signal to the poorer ear was leading in time and was 3.5 dB when the better ear was leading in time. For those with presbycusis, the MLD was 6 dB. These results support the findings of Melnick and Bilger (1965) in that MLDs were obtained for all the patient groups in at least one condition, showing that all had the ability to benefit from interaural differences. The differences in MLD based on the ear leading in time suggests that there may be important interactions between the damage and the stimuli, but, as the authors note, the experiments conducted are insufficient to provide insight into all of the issues that were uncovered. It was also unclear why the asymmetrical conductive loss patients had higher MLDs when the poorer ear was leading in time, but the effect was reversed for those with Ménière's disease. In general, there has been very little work in the past 20 years either on the effects of asymmetrical losses on sensitivity to ITD or on the effects of Ménière's disease on binaural hearing.

Olsen et al. (1976) also tested the MLD in a range of patients, and were seemingly unaware of the work of Melnick and Bilger (1965). Their study was immediately replicated (Olsen and Noffsinger, 1976) using the same methods. Three conditions were tested: N_0S_0, where both signals were diotic, and N_0S_π and $N_\pi S_0$ and where one signal was diotic and the other was reversed in phase at the two ears. As the methods from the two studies were essentially identical, the results have been combined here. Additional data from this study will be described in the section titled "Studies on Binaural Function With Participants With Central Dysfunction" as well. Combining the patients with peripheral loss across both studies results in a group of 124 patients: 62 with high-frequency noise-induced hearing loss, a group of 32 people with unilateral losses due to Ménière's disease, 10 patients with unilateral conductive loss, and 20 patients with presbycusis. Results were compared to data from 62 control participants with normal hearing. Ninety-one percent of those with Ménière's disease showed abnormal performance, as opposed to only 50% of those patients with presbycusis or conductive losses. Thirty-seven percent of those with noise trauma performed abnormally. These results support the conclusion that low-frequency hearing loss, which is more common in Ménière's disease than in the other groups, is one

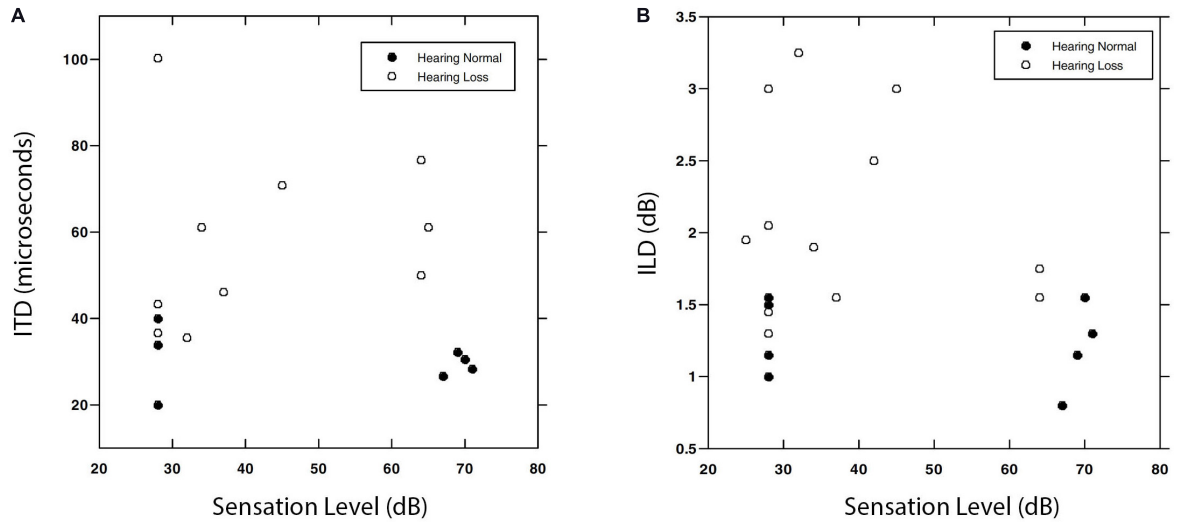

FIGURE 6 | Interaural time difference **(A)** and ILD **(B)** values associated with JND thresholds for a 0.5 kHz narrowband noise. Data are shown for normally hearing (filled symbols) and hearing-impaired (open symbols) listeners as a function of SL. Note that the normally hearing listeners repeated the task at both a high and a low SL, while the impaired group were tested at a single level from which SL was calculated. Reproduced with permission from Smith-Olinde et al. (2004). Copyright 2004, American Speech-Language-Hearing Association.

of the aspects of peripheral hearing loss most likely to result in binaural dysfunction.

Jerger et al. (1984) measured the MLD for pure tones in a very large group of people with a wide range of symmetrical and asymmetrical conductive and SNHLs (n = 651). The main result, consistent with the data just reviewed, was that the MLD was found to depend on hearing thresholds at 500 Hz and to follow the same audibility function measured by McFadden (1968). Data from that study are reproduced in **Figure 7**. Jerger et al. (1984) suggested that these curves could be used to adjust the expected MLD values as a function of 500-Hz detection threshold. Adjustments for higher frequency losses and asymmetrical losses were also described. Jerger et al. (1984) suggested that these corrections could be used to identify patients with abnormally small MLDs, and thus the MLD be used to screen for retrocochlear pathologies, as is discussed in the section titled "Studies on Binaural Function With Participants With Central Dysfunction."

One of the essential insights into both peripheral loss and binaural impairment that has emerged in the past two decades is the idea that peripheral loss can lead not only to reduced sensitivity to pure-tones, but can also interfere with phase-locking at the level of the auditory nerve. Recognition of the importance of sensitivity to both the TFS and the TES of the binaural stimulus has led to significant advancements in both the testing and modeling of binaural impairment. For example, Lacher-Fougère and Demany (2005) showed that hearing impairment reduces sensitivity to ITD in the carrier but not the envelope of modulated tones, implying impaired TFS processing but preserved TES processing, and only a weak relationship with the audiogram. Similarly, the influence of hearing loss on the MLD (as shown in **Figure 7**) has recently been replicated and expanded with extensive modeling by Bernstein and Trahiotis (2016, 2018, 2019), who demonstrated that even very small changes in pure-tone detection threshold can result in reliable reductions in the MLD that can be modeled by reduced encoding of TFS. Moore (2020) provides a comprehensive overview of these results as well as a range of other recent studies of the role of TFS in binaural sensitivity (e.g., Ross et al., 2007; Hopkins and Moore, 2009; Grose and Mamo, 2010; King et al., 2014; Füllgrabe and Moore, 2018).

Peripheral Loss and SRM

Another area in which the effects of peripheral loss have been studied extensively but are still poorly understood mechanistically is that of SRM, which is defined as any improvement in target detection or recognition that accompanies the introduction of spatial cues that differ between a stimulus to be detected or identified (the target) and a stimulus to be ignored (the masker). Binaural release from masking (BRM) refers to the improvements that occur when a binaural cue is provided, which may or may not result in a spatial percept. It is worth distinguishing the two, especially when the goal is to understand the underlying mechanisms.

Early work, described in the section on peripheral loss and binaural sensitivity to interaural differences, focused on the MLD, using headphone presentation in the presence of noise (Olsen et al., 1976; Olsen and Noffsinger, 1976). Much of the later work (Duquesnoy, 1983; Bronkhorst and Plomp, 1989; Peissig and Kollmeier, 1997; Arbogast et al., 2005) moved to loudspeaker presentation of target and masker in various spatial configurations, as well as introducing the use of speech as a masker. While these modifications increased the realism of the testing scenarios, in these studies the maskers were generally presented from a single location, resulting in different SNRs at the two ears. This in turn leads to the availability of a

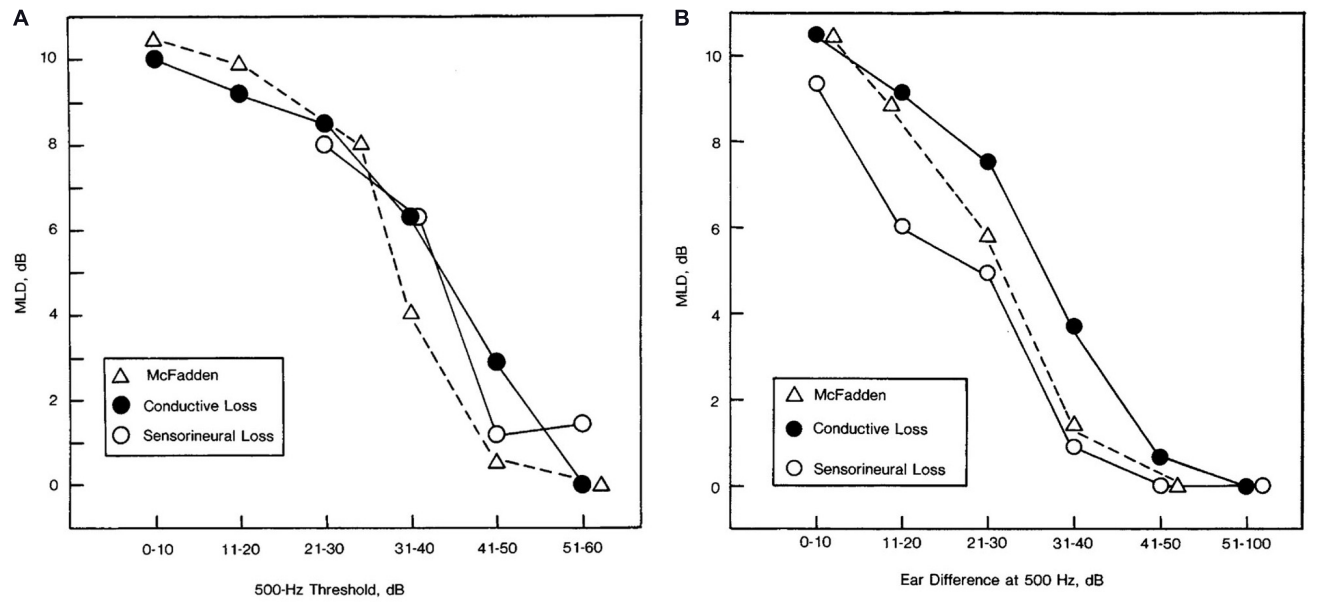

FIGURE 7 | Masking level difference as a function of pure-tone detection threshold at 500 Hz for listeners with symmetrical hearing losses across ears **(A)** and the amount of asymmetry in pure-tone detection thresholds between ears at 500 Hz for listeners with unilateral losses **(B)**. Data are shown for 28 participants with conductive loss in panel **(A)**, 48 participants with conductive loss in panel **(B)**, 71 participants with sensorineural hearing loss in panel **(A)**, and 55 participants with sensorineural hearing loss in panel **(B)**. For comparison purposes, the relevant data from McFadden (1968), in which normally hearing listeners detected signals of various levels are also plotted in each panel. Reproduced with permission from Jerger et al. (1984). Copyright 1984, American Medical Association. All rights reserved.

"better-ear" listening strategy, which means that performance may improve even if the listener has no binaural sensitivity whatsoever. Using a method suggested by the manipulations and better-ear calculations of Hawley et al. (2004); Marrone et al. (2008) demonstrated that the better-ear effect can be significantly reduced (at least in terms of the long-term spectrum) by displacing two maskers symmetrically to the left and right of the target. In these conditions, people with higher pure-tone detection thresholds still exhibit less SRM than do people with thresholds in the normal range. Gallun et al. (2013) introduced a rapid version of this test and observed independent effects of age and hearing loss on performance, but testing was limited to those with SNHL in the mild to moderate range. The test can also be performed using a virtual loudspeaker array presented over headphones with similar results (Jakien et al., 2017; Srinivasan et al., 2020), which allowed Ellinger et al. (2017) to use the VAS to present processed speech signals that included ITD, ILD, or both. Ellinger et al. (2017) found that people with higher pure-tone detection thresholds due to SNHL obtained spatial benefit in all of the conditions tested. Jakien and Gallun (2018) used symmetrical maskers to test a large sample of listeners (n = 82) varying in age, with and without mild-to-moderate SNHL. The results were used to make a predictive linear regression model of performance which was able to account for 38% of the variance in SRM with only the audiogram. Kubiak et al. (2020) used a more sophisticated model, incorporating speech intelligibility measured with standard diagnostic tests, and were able to account for up to 80% of the variance in speech intelligibility among a group of 23 participants with impaired hearing due to SNHL and 7 with normal hearing. These results suggest that while there is significant variance unaccounted for by the pure-tone detection thresholds, perhaps some of the equivocal results of earlier work were due to insufficiently large sample sizes or insufficient model complexity.

Other studies with symmetrically placed maskers (Glyde et al., 2013; Besser et al., 2015) have also found a relationship with pure-tone detection thresholds. Best et al. (2013, 2017) have argued, however, that this relationship may be due to an "energetic limit" on spatial release, where people with more impaired hearing require higher SNRs to understand the target. Based on this argument, the relationship with pure-tone detection threshold may be epiphenomenal, and due to the reduced performance at low SNRs, rather than a binaural deficit. This again argues for the importance of relying less on indirect measures of impairment, such as pure-tone detection thresholds, and relating performance in real-world environments to specific tests of binaural sensitivity.

One study that has applied such a direct approach is Baltzell et al. (2020), who used a headphone test to measure ITD sensitivity and BRM with only an ITD cue for 11 normally hearing listeners and 9 listeners with impaired hearing. By manipulating the interaural correlation of the stimuli for both tests, they were able to obtain a range of ITD thresholds and a range of BRM values for all participants. The relationships between ITD threshold and BRM that were observed are plotted in **Figure 8A**. While BRM was well-predicted by ITD sensitivity in listeners with normal hearing, there was great variability among those with impaired pure-tone thresholds, and many had

BRM values that were worse than were predicted by the best-fitting line for the control participants. **Figure 8B** shows the deviation from the predictions for the listeners with impaired hearing. In addition, as has been found by others (Jerger et al., 1984; Neher et al., 2011, 2017; Bernstein and Trahiotis, 2016, 2018, 2019; King et al., 2017), low-frequency PTA was a significant predictor both of ITD sensitivity and of BRM.

From the data presented in sections "Peripheral Loss and Binaural Sensitivity to Interaural Differences" and "Peripheral Loss and SRM," it is clear that at least some listeners with poorer pure-tone detection thresholds are impaired on tasks requiring good TFS sensitivity, especially for narrowband stimuli and those in which the ITD is not present in the TES (Grose and Mamo, 2012; Gallun et al., 2014; Spencer et al., 2016; Best and Swaminathan, 2019; Baltzell et al., 2020). On the other hand, there are many participants described in the literature with binaural sensitivity in the normal range, despite poor ability to detect pure tones at low levels. As some of the early literature suffers from methodological issues, it might be tempting to argue that the differences are based on methodology of the experiments or perhaps motivation and/or training of the listeners. In order to address this Spencer et al. (2016) went to great lengths to collect a data set using the strongest methods, including training the listeners extensively and repeating the measures to ensure that only data with strong internal consistency were included. Indeed, internal consistency was high across the full data set, suggesting that the values were measured reliably. Nonetheless, substantial heterogeneity was observed among the younger hearing-impaired listeners tested, with many performing in the same range as the normal-hearing control participants. The results of Spencer et al. (2016) underscore the message of nearly 100 years of research on hearing impairment and sensitivity to interaural differences: the ability to detect pure-tones is an indicator of who may be suffering from binaural impairment, but it is not sufficient for strong predictions, especially when the stimuli are broadband and the hearing loss is worst in the high frequencies.

The modeling approaches that have been most successful in predicting the binaural abilities of individual listeners (rather than group differences) have combined pure-tone detection thresholds with metrics unrelated to spatial cues such as age, measures of speech understanding (Kubiak et al., 2020), and/or measures of cognitive function (Gallun and Jakien, 2019). In addition, there are computational modeling approaches that show great promise in helping identify the specific mechanisms responsible for binaural impairment (Le Goff et al., 2013; Mao et al., 2015; Moncada-Torres et al., 2018). The most promising opportunities for future research are those that involve a process of informational feedback between human patient research and targeted animal and computational models. Ideally, such a program would start by developing models based on the existing human data, which in turn would predict the factors that are most important for measuring in the patients. Then the human experiments could be developed to measure and control those factors as a way of testing the models. To date, few binaural clinical research programs have followed this process, but some of the most successful (e.g., Baltzell et al., 2020) are definitely moving in this direction.

STUDIES ON BINAURAL FUNCTION WITH PARTICIPANTS WITH CENTRAL DYSFUNCTION

While many researchers have focused their studies of binaural impairment on the relationship with pure-tone detection thresholds, as audiological practice might suggest, there are a number of other groups that have conducted investigations of some of the other ways in which binaural function could be impaired. Much of the earliest work on binaural impairment was conducted with patients with normal or near-normal pure-tone detection thresholds who were diagnosed with MS, strokes, brain tumors, or traumatic brain injury. More recently, there has been considerable interest in the effects of aging, alone and in combination with hearing loss, on binaural function. The results of such studies are extremely important for connecting the animal literature on binaural processing to the human literature on binaural function. In animals, detailed information can be gained about the pathways and signal processing associated with the binaural system, but data on complex behavioral tasks are very difficult to obtain. With humans, the opposite is generally true. The exception is when imaging data are available showing precise lesion locations for a patient who has also performed spatial hearing tasks. As imaging techniques improve and our understanding of neurological disease progresses, there is great opportunity for our knowledge of the mechanisms of binaural hearing to improve as well.

Aging

One of the most difficult issues associated with the study of hearing loss, especially in the high frequencies, is the comorbidity with aging. For binaural impairment, this issue is especially important to address, as there is considerable reason to believe that those with aging auditory systems can potentially suffer from a wide range of monaural and cognitive impairments that are likely to influence performance on tests of binaural function (reviewed in Gallun and Best, 2020). In addition, it is increasingly clear that aging itself can impair the functioning of the binaural system. One of the major obstacles to studying the effects of aging on binaural processing is the difficulty of comparing younger and older listeners independently of differences in peripheral hearing. Either the researcher must limit participants to a very specific audiogram, which may limit the generalizability of the results, or statistical approaches must be used to factor out the effects of peripheral hearing loss. Statistical separation of the influences of the two factors requires testing larger samples and recruiting listeners in a manner that age and hearing loss can be statistically dissociated.

Aging and Localization

Abel et al. (2000) tested the horizontal localization abilities of 112 participants aged 10–81, divided into seven age groups, each with 16 participants within the same decade of life. Both

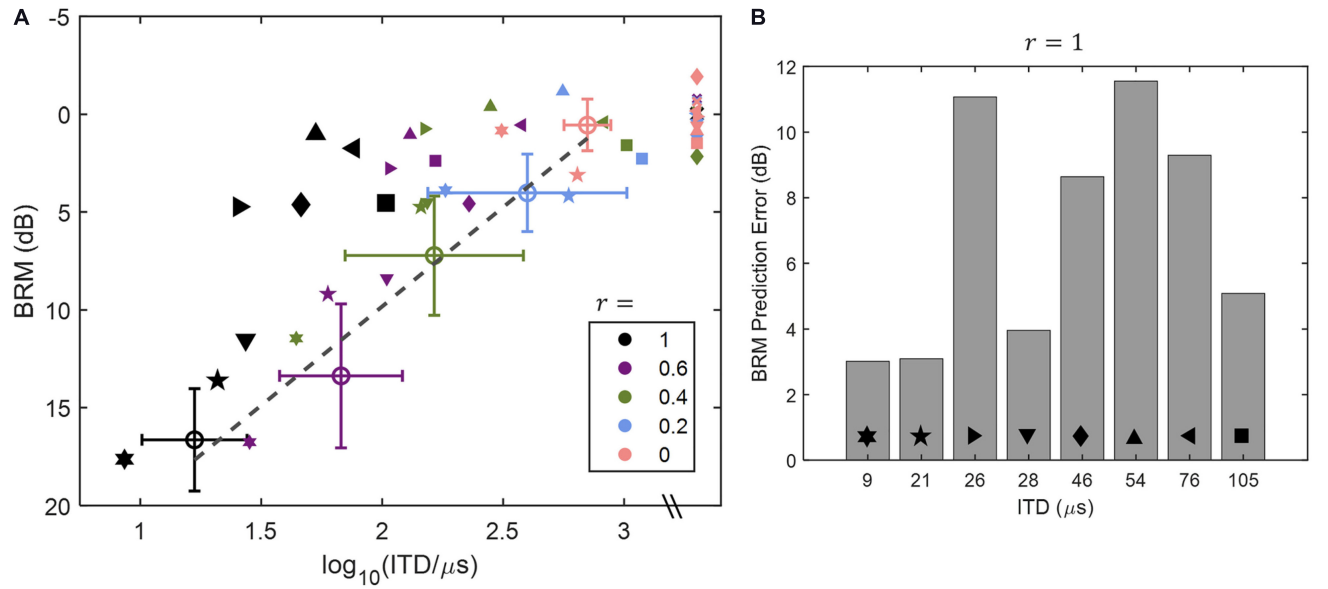

FIGURE 8 | Data and modeling reproduced with permission from Baltzell et al. (2020). Copyright 2020, Acoustical Society of America. Relationships between ITD threshold and BRM as a function of interaural correlation (*r*) of the stimuli tested are shown in panel **(A)**, and the deviation in dB between the best-fitting line for the listeners with normal hearing [open circles in panel **(A)**] and the values for the hearing-impaired listeners [filled symbols in panel **(A)**] are plotted in panel **(B)**. Values of *r* are indicated by color of the symbols in panel **(A)**, and gray symbols indicate points for which the ITD was unmeasurable. One listener, who had an unmeasurable threshold for an interaural correlation of 1, was excluded from the analysis and is not included in the figure. See text for further details.

of the youngest groups had hearing thresholds no greater than 12 dB HL from 0.5 to 4 kHz, but those aged 30–59 were allowed to have thresholds up to 22 dB HL and those 60–81 were allowed to have thresholds has high as 37 dB HL. Thus, there was a systematic increase in average thresholds with age, reaching a maximum difference of 25 dB at 4 kHz between the youngest and oldest groups. Horizontal localization performance declined significantly by the third decade and overall was reduced by 12–15% between the youngest and the oldest listeners. Much of the error was attributable to front-back confusions, however. As these errors are likely to be associated with loss of spectral cues due to high-frequency audibility differences between the groups, it is not possible to say definitively whether these data represent an aging effect independent of the age-related changes in hearing thresholds. Similar studies by Dobreva et al. (2011) and those reviewed by Freigang et al. (2015) also found that the accuracy and precision of localization is reduced in older listeners, but it was not possible to definitively separate the effects of aging from slight age-related declines in peripheral function. The effect of age on localization is shown in **Figure 9**, reproduced from Dobreva et al. (2011). While these results show a strong effect of aging on horizontal localization for mid-frequency signals, but not low-frequency signals, they are also consistent with the results of Bernstein and Trahiotis (2016, 2018, 2019), who suggested that horizontal localization ability may be reduced by hearing loss within the normal range. In this case it is difficult to determine whether these changes in localization ability are truly age effects or whether they might be the effects of small differences in detection ability.

Aging and Binaural Sensitivity

Ross et al. (2007) measured the effects of aging on binaural function using both behavioral and magnetoencephalographic measures. They controlled for hearing effects by ensuring that all listeners had thresholds below 20 dB HL between 0.5 and 2 kHz and no higher than 40 dB HL at 4 kHz. As mentioned in the section on aging and localization, however, this does not

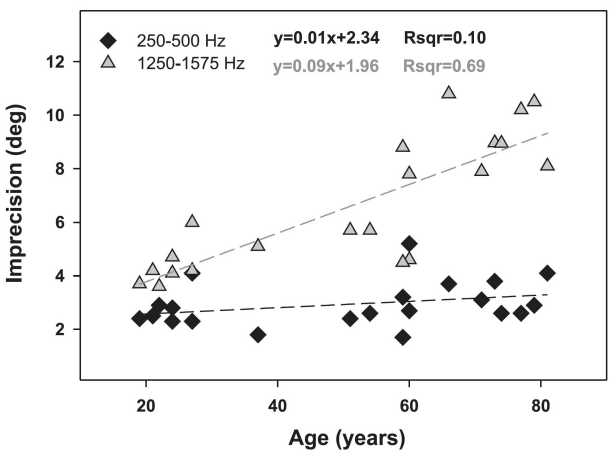

FIGURE 9 | The relationship between age and localization accuracy for two narrowband signals: 250–500 Hz (black diamonds) and 1,250–1,575 Hz (gray triangles). Dotted lines indicate linear regressions associated with the equations shown. Data and modeling reproduced with permission from Dobreva et al. (2011). Copyright 2011, American Physiological Society.

allow us to rule out the influence of small differences in audibility. Their behavioral measure showed that the middle-aged and older listeners were less able to detect an interaural phase shift in an ongoing amplitude-modulated tone than were the younger participants in the study and, as will be discussed in the section on neural measures of binaural sensitivity in older listeners, that this difference was reflected in recording of their brain activity. Their behavioral results were replicated and extended by Grose and Mamo (2010, 2012). The results of Grose and Mamo (2010; **Figure 10**) showed that while 60% of the listeners younger than 27 years were able to detect phase shifts for stimuli with carrier frequencies as high as 1.25 kHz, only 15% those 40–55 years of age could do the task at this frequency, and only about 5% of the listeners who were aged 63–75. Grose and Mamo (2012) extended these results by using the dichotic FM task described in the section on methods of characterizing binaural impairment, in which a 500-Hz tone was modulated in frequency in different directions in the two ears, creating a fluctuating IPD. The same range of hearing losses were present in these listeners. While the younger group (aged 19–29 years, n = 12) could detect binaural FM of 0.4 Hz on average, the middle-aged group (43–57 years) could only detect binaural FM of 0.8 Hz, and the older group (aged 65–77 years) needed almost 2 Hz of modulation before they could perform the task. The younger group also performed better than the other two groups when the stimuli were presented diotically, but even the younger listeners still needed more than 2 Hz of modulation in order to perform the task with no binaural difference. When the stimuli were presented diotically to the middle-aged and older groups, they needed 3 and 3.5 Hz, respectively.

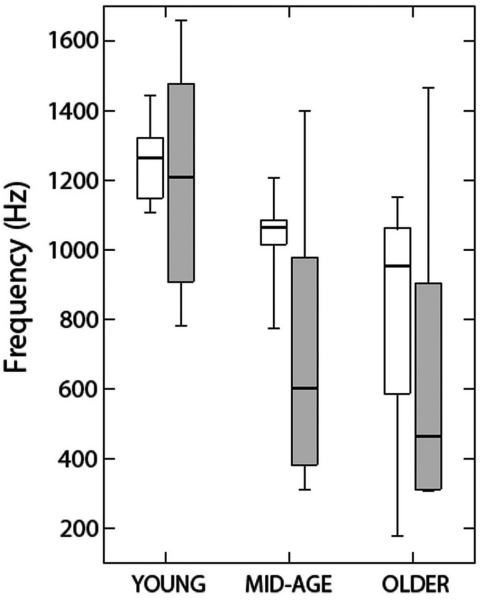

FIGURE 10 | Data indicate the highest carrier frequency at which listeners in three age groups could discriminate diotic from dichotic stimuli. Open rectangles show data from Grose and Mamo (2010) and shaded rectangles show data from Ross et al. (2007). See text for further details. Reproduced with permission from Grose and Mamo (2010). Copyright 2010, Wolters Kluwer Health, Inc.

King et al. (2014) tested 46 listeners varying in age from 18 to 83 with a wide range of pure-tone detection thresholds (−1 to 68 dB SPL at 1 kHz) for whom age and pure-tone sensitivity was uncorrelated at 0.5 and 1 kHz (r = 0.08) but more strongly correlated at higher frequencies (r = 0.439). Listeners were asked to detect IPDs in low frequency (250 or 500 Hz) tones amplitude-modulated at a rate of 20 Hz. IPDs were applied to either the carrier or the modulator in order to test the hypothesis that there is an age-related deficit in TFS processing that is independent of a binaural impairment. Presentation levels were set by first measuring the detection threshold for the stimuli and presenting all stimuli at a minimum of 30 dB SL. The data revealed age-related deficits in binaural processing for the 500 Hz tones, whether the IPD was applied to the carrier or the modulator, and for the 250 Hz tones, but only when the IPD was applied to the modulator. These data were taken to support a generalized age-related decline in temporal processing rather than a specifically binaural impairment.

Füllgrabe (2013) also found evidence for a general temporal processing deficit in a sample of 102 participants with normal audiometric thresholds varying in age from 18 to 90 years. Both a monaural TFS test and a binaural TFS test showed systematic declines in performance as age increased. Füllgrabe and Moore (2018) performed a meta-analysis of 19 studies that used the same binaural TFS test and found that while age and pure-tone detection thresholds were both significant predictors of performance, age accounted for more variance in every comparison conducted. The total amount of variance accounted for by both factors was never more than 42%, however, suggesting that these two variables alone are insufficient to account for performance on even a very specific psychophysical task.

Whiteford et al. (2017) also measured the effects of age on monaural and binaural temporal sensitivity by comparing the detection of slow-rate (1 Hz) AM and FM with the detection of fast-rate (20 Hz) AM and FM. Both AM and FM were tested diotically and dichotically. Dichotic AM results in time-varying ILDs and dichotic FM results in time-varying ITDs, as described above for the experiments of Grose and Mamo (2012). Whiteford et al. (2017) tested 85 listeners aged 20–80 years, with pure-tone average thresholds (0.5, 1, and 2 kHz) that were all in the normal range (no greater than 20 dB HL). Average thresholds and age were correlated (r = 0.56). All stimuli were presented at 60 dB SPL. The hypothesis tested was that there would be at most small effects of age on AM detection at either rate, but that FM detection would be more impaired with aging for the slow-rate stimulus, where only TFS cues were available. Contrary to expectations, however, age effects were observed not only for both slow and fast FM presented diotically or dichotically, but also for fast dichotic AM. These results also support the idea that impairments in temporal processing associated with aging are likely to involve a variety of processes, including, but not limited to, TFS sensitivity.

Gallun et al. (2014) reached similar conclusions when they measured the temporal processing abilities of a large group of participants (n = 78) varying in age (18–75 years) and hearing

loss (0–40 dB HL at 2 kHz). Listeners were tested on monaural, binaural, and bilateral timing discrimination tasks with brief (4 ms) stimuli with narrow or broad frequency content, all centered at 2 kHz. In the monaural task, listeners were asked to detect a gap between two stimuli, in the binaural task they were asked to detect an ITD, and in the bilateral task, they detected delays between the presentation of a signal to the left and right ear. Gaps, ITDs, and bilateral delays were adaptively varied to determine threshold. Presentation level was set to 30 dB above detection threshold for each stimulus. Rather than relying upon group differences or partial correlations, a linear mixed model was developed in which performance across all stimuli and tasks was modeled, taking into account each individual's age and stimulus detection thresholds. One advantage of this approach is that each individual can be assigned an intercept value for their function, reflecting that individual's ability to perform psychophysical tasks. The model predicted 20–40% increases in monaural gap detection thresholds per decade of aging, 15–20% increases in ITD discrimination thresholds with every decade, and 0.9–10% increases in bilateral delay sensitivity, all independently of increases in temporal processing ability with increases in signal detection thresholds. **Figure 11** shows the model predictions from Gallun et al. (2014), where the age effects are indicated by the predictions for a 20 year old (black lines) and a 60 year old (gray lines). The top row shows the increases in threshold in the three conditions for a tonal stimulus as a function of increases in stimulus detection thresholds, while the bottom row shows the changes predicted for the same listeners and conditions for a broadband stimulus. **Figure 11** demonstrates that while there are indeed timing and/or binaural impairments associated with even slight hearing loss, statistical modeling can be used to more clearly quantify the independent effects of aging and pure-tone sensitivity and their interactions with stimuli and tasks.

Aging and SRM

Many of the studies examining the effects of hearing loss on SRM, especially with speech stimuli, have also focused on the effects of aging. Early work measuring the MLD with tones (e.g., Olsen et al., 1976) was replicated by Pichora-Fuller and Schneider (1991, 1992), who found that older listeners with slight hearing loss had significantly impaired MLDs and were able to account for this by applying a computational model in which TFS sensitivity varied for younger and older listeners. Duquesnoy (1983) and Gelfand et al. (1988) both noted reduced spatial benefit in their older participants with normal hearing relative to their younger participants. However, the effects were smaller than were the differences between the participants with normal and impaired hearing.

Dubno et al. (2008) were the first to use symmetrically placed maskers and compare younger and older listeners with normal hearing (n = 30). While SRM did not differ between the groups, performance for the older listeners was worse than a model based on the audiogram predicted, and there was a strong correlation between age and performance in the spatially separated condition. Marrone et al. (2008) also tested younger and older listeners with and without hearing impairment (n = 40) and observed relatively small effects of age independent of hearing loss. Glyde et al. (2013) measured SRM in a speech-on-speech masking task with a group of 80 listeners varying in age (7–89 years) and hearing loss. While there were substantial effects of age in the simple correlations, partial correlations taking into account hearing loss did not reveal significant relationships between age and performance. One possible reason for this is the inclusion of children, for whom spatial benefit increases with age, rather than declining as it does for adults.

Gallun et al. (2013) used a version of the Marrone et al. (2008) task and tested 52 listeners across three experiments and showed stronger effects of aging than of hearing loss on SRM. One possible reason for the difference between these results and those of Glyde et al. (2013) was the use of 45° of separation between target and each masker as opposed to the 90° used in Glyde et al. (2013). Jakien et al. (2017) verified that the SRM observed in the Marrone et al. (2008) task is maximal at about 45°, suggesting that even an impairment that reduces the "effective" spatial separation from 90° to 45° would likely have a minimal effect on SRM. Srinivasan et al. (2016) explored the effect of spatial separation in greater detail by examining SRM with small separations and discovered that the effects of aging are the most apparent with separations less than 15°.

These results with SRM support the findings of age-related declines in binaural sensitivity and localization accuracy. The primary difficulty with interpreting all of these results, however, is that it is unclear whether the mechanisms by which aging and hearing loss cause binaural impairment are fundamentally different. Future work combining human studies with computational models of the binaural system can shed light on this, especially if informed by animal models of binaural impairments associated with age and hearing loss.

Neural Measures of Binaural Sensitivity in Older Listeners

As mentioned in the section on aging and binaural sensitivity, Ross et al. (2007) conducted the first study to compare older and younger listeners on both behavioral and neural measures of binaural sensitivity. Using the P1-N1-P2 complex responses as measured with MEG (see section "Methods of Characterizing Binaural Impairment" for details) as well as a behavioral detection task using the same stimulus, they were able to compare the highest carrier frequency at which an interaural phase reversal in the carrier frequency of an amplitude-modulated tone was detectable by a human observer and at which the P1-N1-P2 complex was detectable. Both the maximum frequency at which the P1-N1-P2 complex was present and the maximum frequency at which the older listener could detect the binaural change was lower for the older listeners than for the younger, and even the middle-aged listeners differed from the younger listeners. Data from Ross et al. (2007) are shown in **Figure 12**. These results led to a substantial increase in the number of researchers interested in the effects of aging on binaural sensitivity. Many of the studies that followed (e.g., Papesh et al., 2017; Eddins and Eddins, 2018; Eddins et al., 2018; Vercammen et al., 2018) suggested that age-related binaural impairment on behavioral tasks in humans is related to reductions in TFS encoding. Eddins and Eddins (2018),

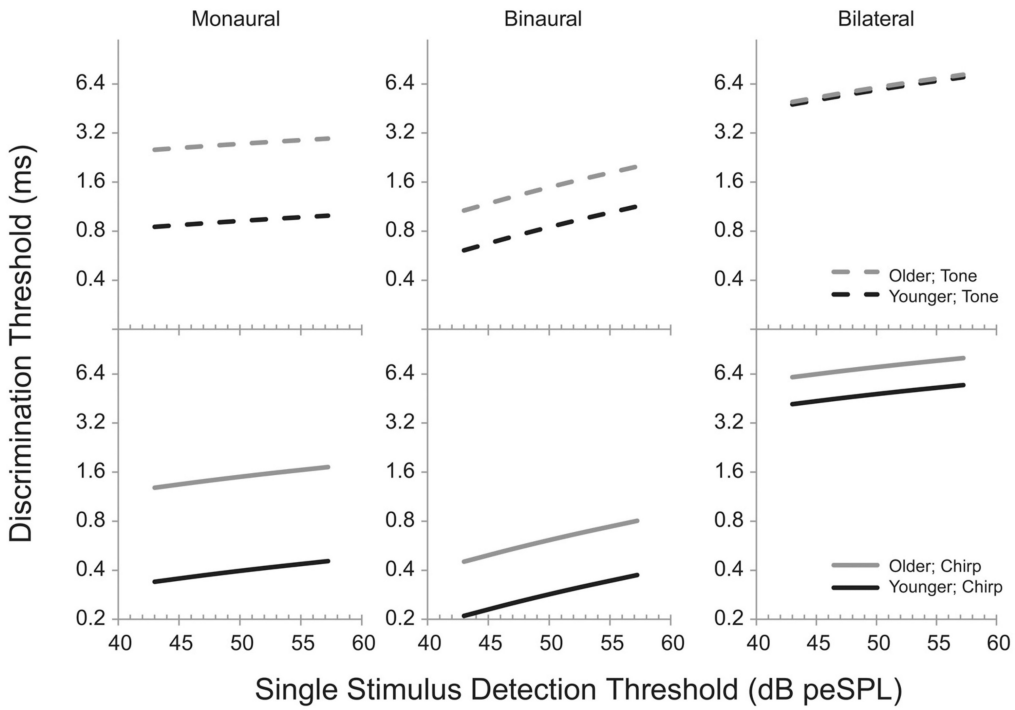

FIGURE 11 | Model predictions from Gallun et al. (2014) showing the effects of detection threshold for a brief (4 ms) stimulus in peak-equivalent units (peSPL; equivalent level for a 1-s pure tone with the same peak level) on discrimination thresholds. Predictions are shown for a modeled younger listener (20 years; black lines) and a modeled older listener (gray lines). Solid lines are for a broadband stimulus ("chirp") and dashed lines are for a narrowband stimulus ("tone"). See text for a description of the monaural, binaural, and bilateral tasks. Reproduced with permission from Gallun et al. (2014) under Creative Commons license.

based on the finding that binaural encoding is reduced for a 500 Hz stimulus but not a 4,000 Hz stimulus, suggested that TFS encoding at the level of the cortex is reduced by aging but that TES sensitivity is not. While the electrophysiological and behavioral responses at the level of the cortex are fairly strong for these and other binaural tasks (Papesh et al., 2017), some electrophysiological measures (Anderson et al., 2018) show poor relationships with binaural sensitivity. In addition, some research (Koerner et al., 2020) has found a diversity of relationships between neurophysiological responses and different behavioral tasks in the same listeners. Koerner et al. (2020) were the first to use a behavioral task that used a stimulus that was directly comparable to the IPM-FR stimulus and found no relationship between the behavioral results and aging or behavioral thresholds and neural amplitudes. In the same listeners, however, the SRM task used by Papesh et al. (2017) was correlated with the amplitude of the neural response at an individual level.

Anderson et al. (2018) interpreted the lack of a relationship between the behavioral and neural data in their study as suggesting that the relationship between encoding and behavior may be more complicated than simply that aging leads to increased variability in TFS encoding at the level of the auditory nerve, which should lead to decreased behavioral thresholds that are related to reduced neural amplitudes. Consistent with this interpretation, electrophysiological recordings in aged monkeys (Juarez-Salinas et al., 2010; Engle and Recanzone, 2013) revealed age-related degradations of the inhibitory connections between cortical and subcortical brain areas, leading to reductions in the tuning of cortical areas specialized for spatial hearing. If this occurs in the brains of older humans as well, then the brainstem encoding could be reduced by one mechanism (perhaps related to auditory nerve phase locking) while the cortical representations of space could be reduced by a separate mechanism (perhaps related to reduced inhibitory connections between brain areas). On the other hand, Maier et al. (2008) tested older gerbils on a spatial discrimination task and found reduced performance, but in this case accompanied by reduced inhibition within the brainstem structures essential for processing interaural timing differences.

In summary, our understanding of the mechanisms relating aging to binaural sensitivity is far from complete. What is clear is that there are aging effects that can be observed both behaviorally and neurally, and that animal and computational models have the potential to both clarify existing data and point the way toward new approaches to the study of aging and binaural function.

Multiple Sclerosis and Binaural Impairment

Multiple sclerosis is a disease of the nervous system that leads to lesions both at the level of the brainstem and at the cortex. While there have been a small number of studies of binaural sensitivity in MS patients, the promise of modern imaging technologies has only begun to be explored. With the ability to identify specific

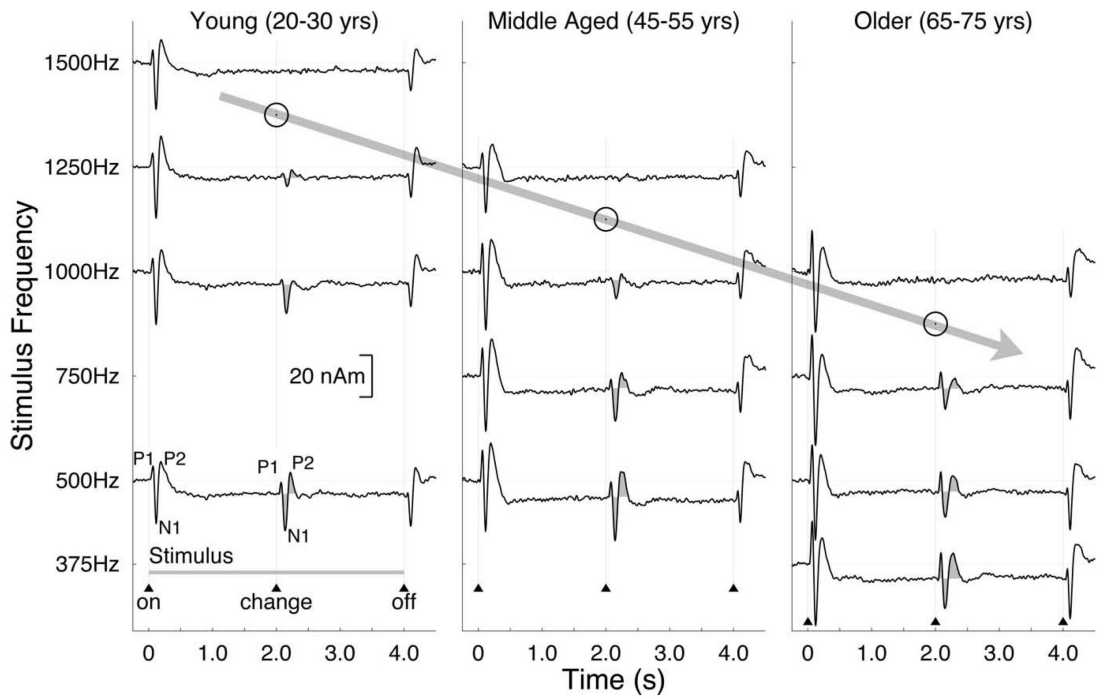

FIGURE 12 | Auditory evoked responses measured by MEG showing the P1-N1-P2 complex to a change from a diotic to a dichotic stimulus as a function of carrier frequency and age group. Data reproduced with permission from Ross et al. (2007). Copyright 2007, Society for Neuroscience.

lesions and relate them to binaural function, it may be possible to obtain evidence in humans of the specific roles of various brain areas; evidence that only animal models could previously provide. This is also the promise of the work detailed in the section on the effects of brain tumors and lesions on binaural sensitivity.

The existing literature shows clearly that MS patients are likely to have binaural deficits. One of the earliest reports to show this was the MLD data on 100 patients with MS reported by Olsen et al. (1976), "almost all of whom" had normal audiograms and speech reception thresholds in quiet. Using both speech and tonal targets and comparing to the MLD from a group of control subjects, at least 41% percent of the MS patients were in the abnormal region for one or more stimuli. Jerger et al. (1986), using the correction to the MLD for hearing thresholds developed in Jerger et al. (1984), found the MLD to be abnormal in 45% of a group of 62 MS patients. Similarly, Musiek et al. (1989) found that 50% of a group of 26 MS patients (all with normal audiograms) had abnormal MLDs, as compared with 20 control subjects.

In addition to the MLD, some studies have also examined sensitivity to ITD and ILD directly, such as Häusler and Levine (1980), who measured JNDs in interaural time and intensity for 29 patients with MS. Results were compared to those of 36 normal-hearing controls with no known neurological impairment. All of the controls had ITD JNDs between 10 and 40 μs, while 13 of the MS patients had ITD JNDs of 50 μs or greater. All of the controls had ILD JNDs of 0.5 – 2.0 dB, while 6 of the MS patients had thresholds of 2 dB or greater. Similarly, Häusler et al. (1983) included 26 of the patients with MS in their study. Of these, two-thirds had abnormal performance with the side referenced MAA task but did well on the center-referenced MAA task. Colburn (1982), in reviewing these data stated:

> *"The fact that different tests (particularly interaural time and interaural intensity discrimination) give independently normal and abnormal performance, even with the same stimulus, suggests that different regions of the brain are involved in processing the different aspects of the stimulus, such as interaural time delay versus interaural intensity differences" (p. 42).*

This statement reveals how useful it can be to test patients with neurological disorders and how essential it is to pair these tests with animal models and the results of neurophysiological experiments. While it is well-accepted today that ITD and ILD are processed by independent brain regions, these patient data were early indicators that preceded the more definitive neurophysiological studies to come.

In the 1990s, several additional studies of lateralization and sensitivity to interaural differences were conducted with patients suffering from MS (Levine et al., 1993a,b, 1994; Furst et al., 1995; Häusler and Levine, 2000). Due to improvements in the ability to image the central auditory system, these studies were able to shed light on how the MS lesions related to binaural impairment. In patients with lesions of the brainstem, binaural function was found to be impaired on both lateralization and detection of interaural differences (Furst et al., 2000). Based on the patterns of MS lesions and their relationships to binaural dysfunction, Furst et al. (2000) proposed a model in which MS impairs ITD-based detection and lateralization due to the imposition of additional conduction delays on the neural

networks underpinning ITD sensitivity. Further work in this area would be tremendously informative.

Binaural Impairment in Patients With Brain Tumors and Lesions

While it is tempting to imagine that the earliest work on binaural impairment focused primarily on peripheral hearing loss, the reality is that Greene (1929) also used his "short-circuited" binaural stethoscope to study fifteen patients with brain tumors. Five of the fifteen were abnormal, as compared to the control group, which distinguished them from those with peripheral loss, none of whom were abnormal in their sensitivity to binaural differences measured in this manner. It was these data that convinced Greene (1929) find collaborators to help him develop the innovative equipment for distinguishing ITD and ILD sensitivity described in the section on methods of characterizing binaural impairment. Using this apparatus, Greene, who was a neurosurgeon, tested 51 of his patients with neurological disease and compared their performance to that of control participants with no known neurological disease and normal peripheral hearing. Only three of the patients showed localization that differed from the control group when they were asked to identify the location of a ticking watch, but fifteen of those with normal localization in the ticking watch task, where both ITD and ILD cues were present, had abnormal lateralization perception with ITD or ILD alone. For ten of these patients, perception of only one cue was abnormal, while for five of them, perception of both cues were abnormal in isolation, despite no impairment when asked to localize a ticking watch.

Greene (1929) used these clinical data to form hypotheses about the underlying physiology, noting that at that time it was unknown whether or not the auditory nerves from the two ears connect to one or both temporal lobes. Based on the observation that the majority of those with impaired localization had temporal lobe tumors, he concluded that it is likely that the monaural auditory pathways extended to both temporal lobes. Furthermore, he mentioned that it was unknown where in the auditory system sound localization occurs. Based again on his finding that the majority of the lesions in those with impairment were in the temporal lobe, he concluded that it is likely that this is where the localization ability resides. He acknowledged, however, that his sample size was too small for this to be more than speculation.

Yet, despite these remarkable aspects, some important elements are missing from Greene (1929). Durlach et al. (1981) pointed out that the data are all reported entirely in terms of average response, with no indication of within-subject variability. In addition, the between-subject information is reported in tables where each listener's ability to localize is categorized on a six-element scale ranging from "normal" to "completely absent." It is difficult to read the report and not wish one had access to the full data set upon which these categorical judgments are based.

Walsh (1957) conducted a series of experiments following up on the work of Greene (1929) almost 30 years later, but that took advantage of the advances in electronic devices and in psychophysical techniques that had occurred during the intervening years. The work was inspired by Wallach et al. (1949) and used electrical pulses delivered either to the two ears or to two loudspeakers. A phase-delay circuit was also used to deliver a 250-Hz tone to the two ears either delayed by 90° or undelayed (presented "diotically"). An unspecified "small number" of normal hearing controls were used to establish that the clicks sounded like a single click when the delay was less than 2.5 ms ("precedence threshold"), and that ITDs of 100 μs were about the smallest that could be distinguished from a diotic signal ("ITD threshold"). Twenty-one patients with cerebral lesions were tested, and all could detect ITDs, but roughly half had detection thresholds greater than 300 μs. Several patients were able to distinguish ITDs below 300 μs despite significant lesions to the auditory processing areas of one hemisphere, leading Walsh (1957) to conclude that binaural detection does not require both hemispheres. Of the 15 cerebral lesion patients who completed precedence-effect testing, all reported that for some portion of the range of delay they experienced a single sound ("fusion") and only four had precedence thresholds exceeding 4 ms. Only one of the 12 cerebral lesion patients who performed the task of detecting a phase-delay was unable to do so successfully. From these studies, Walsh (1957) concluded that it is likely that binaural comparison occurs at the level of the brainstem, the output of which is sent to both hemispheres.

Häusler et al. (1983) did not measure the abilities of people with cerebral tumors, but they did report data from 7 patients who had a tumor on the eighth nerve called a vestibular schwannoma (referred to in the text as a "neurinoma"). Quite uniformly, it was found that these patients were among the most impaired of all the groups tested in that study. All were unable to perform the MAA task in the normal range, in either the center- or side-referenced condition. ITD discrimination thresholds were in the abnormal range for the majority (but not all) of the patients with tumors and only one patient with an eighth-nerve tumor had normal ILD sensitivity. It should be noted that this was the most recent study that could be found in which patients with eighth-nerve tumors were included in a binaural experiment.

Colburn (1982) concluded that for the patients with eighth-nerve tumors,

> *"information flow on the auditory nerve is extremely disrupted by some auditory nerve lesions. The timing and intensity information can be essentially eliminated at supra-threshold levels, even when the threshold value is only slightly affected. If one postulates a tumor pressing against nerve fibers, it is easy to imagine not only a disruption of the timing of individual action potentials but also an interference with the number of firings (e.g., by an increase in the refractory period) without an associated change in the absolute detection threshold. I am not aware of an animal model for this condition" (p. 41).*

It is probably fair to say that little has changed in terms of our understanding of the effects of eighth-nerve tumors on binaural hearing in the intervening decades. On the other hand, one area that has been very informative is the study of patients with strokes that impinge on the auditory brain regions. Furst et al. (2000) reported results from patients with strokes in the auditory brainstem areas and that both lateralization and sensitivity to

interaural differences were impaired. From this, they proposed a model in which strokes lead to diminished ITD and ILD, as well as lateralization, by damaging the connections in the brainstem.

Horizontal localization in patients with cortical and brainstem damage from strokes was also studied by Sonoda et al. (2001) and Przewoźny et al. (2015a,2015b), who both found accuracy of sound source identification to be reduced in some of their patients. An important consideration in studying stroke patients is the potential for damage to non-auditory areas to result in difficulties with a localization task. For example, recent work on the relationships between auditory neglect due to stroke and spatial hearing is reviewed in Gutschalk and Dykstra (2015), who conclude that more work is needed to develop clinical protocols that can clearly distinguish localization deficits from disorders of spatial cognition.

Binaural Impairment in Patients With Traumatic Brain Injury

The work of the Vietnam Head Injury Study (VHIS; Sedge, 1987) stands essentially alone in the study of binaural impairment in those with penetrating head wounds. Phase 2 of the VHIS started in 1980 and tested 482 head injured patients and 82 controls. Most of the tests took place 14 years after injury. Mueller and Beck (1987) reported on the MLD thresholds for a 500 Hz tone presented in narrowband noise obtained from 55 control subjects and 92 Veterans with a history of penetrating head wounds. There was a small difference between the MLD for the controls and those with a history of brain injury, but it was less than 1 dB and was non-significant. These results were interpreted as consistent with the fact that the Veterans in this study had primarily cortical injuries, as the authors believed the MLD to be a measure of brainstem integrity. This is an important finding and one on which it would be useful to have more data.

Another population found to have difficulties with binaural tasks is those who have experienced head trauma (Gallun et al., 2012; Saunders et al., 2015; Roup and Powell, 2016; Hoover et al., 2017; Kubli et al., 2018). Binaural dysfunction in this group is particularly difficult to characterize due to both the heterogeneity of the injuries and the increased likelihood of impairment on complex tasks. The issue of heterogeneity derives in part from the diversity of physical events that can cause even mild Traumatic Brain Injury (mTBI). The literature includes both studies of patients with a history of exposure to high-intensity blasts during their military exposure ("blast exposure"; Gallun et al., 2012, 2016; Saunders et al., 2015; Kubli et al., 2018) and studies of patients with mTBI following non-military events such as falls, sports injuries, and motor vehicle accidents. Even within these two categories, however, there is very little reason to believe that damage to the same brain areas would occur under different physical conditions. Unfortunately, unlike with penetrating head wounds, strokes, or even MS, it is a hallmark of mTBI that only rarely does it result in injuries that can be revealed by current clinical imaging approaches. As reviewed in Gutschalk and Dykstra (2015) and Felix et al. (2018), there is a wide range of ways in which injury to the brain could result in impaired performance on tests of sensitivity to binaural and spatial information. For these reasons, it is essential to interpret the mTBI literature with care and consider carefully the possibility that group differences may not be reliable predictors of what an individual patient may experience.

This heterogeneity across patients may explain why Gallun et al. (2012) found that there was a small, but statistically significant, subset of their injured patients who had abnormally poor MLD scores, but Gallun et al. (2016) were not able to replicate this finding. Instead Gallun et al. (2016) observed data more similar to those of Mueller and Beck (1987), where patients with a history of blast exposure (only some of whom had an mTBI diagnosis) were more likely to show auditory processing difficulties on complex tests but no difficulties on detection tasks such as the MLD. Saunders et al. (2015) also observed abnormal SRM in their larger sample of blast-exposed Veterans, but tested SRM for speech-on-speech masking rather than a detection task. Kubli et al. (2018) explored task complexity explicitly in an SRM task by asking blast-exposed Veterans to localize the voice of a person talking about a specific topic ("sports," "food," etc.), either in a quiet room or in the presence of one or more competing speakers talking about other topics. While the injured Veterans performed similarly to a non-blast-exposed control group in quiet, there were significant increases in the group differences in SRM as the complexity of the acoustical environment increased.

Roup and Powell (2016) reported binaural impairment in a group of people who had suffered mTBIs from non-military causes, as did Hoover et al. (2017), who used a wide range of monaural and binaural tests and found significant impairment in the mTBI group. For Hoover et al. (2017), it was impossible to identify a specific monaural or binaural deficit common to all of those with mTBI. These results suggest that both military and non-military brain injury can impair the binaural system, but that knowing the details of the injury, the tests used, and the types of binaural impairment revealed are essential for drawing conclusions that can be used to generalize the results to beyond those patients included in the study.

These results, like the data reviewed from patients who have suffered strokes, developed MS, or are undergoing the normal aging process, reveal the complexity of doing clinical research with patient populations. Nonetheless, shedding light on binaural dysfunction is of great benefit to the patients and to the clinicians who treat them. The binaural system is poorly understood by the general public, and even by most clinical specialists. By clarifying the dysfunction likely to occur among patients with various diseases, it becomes possible to develop new clinical tests as well as to train clinicians in how to counsel those with binaural impairment. In addition, there is potential to deepen our understanding of how the binaural system functions by learning what mechanisms can be impaired and how such mechanisms can change binaural processing.

FUTURE DIRECTIONS

The work reviewed in the sections above suggests that, while much has been learned about binaural impairment since the first reports over 100 years ago, there is still much to be discovered. The field would benefit from further research in a number of areas, including animal models and computational modeling.

As is clear from this review, the approach of testing participants who have diseases known (or suspected) to affect the binaural system has substantial potential benefit in two different ways. The first is that we are likely to learn more about those diseases and what abilities and difficulties people with those diseases are likely to experience. The second is that these diseases allow scientific investigation of auditory processing that has been perturbed in ways that are otherwise only possible to do in animal research.

It should be noted that despite the many studies and conditions discussed, this review has not been comprehensive, as there has been no discussion of binaural development in children, nor of the effects on binaural function of many additional auditory and brain diseases. In some cases, such as with development, this was due to a need to limit the scope, and in many others it was due to the lack of a well-developed literature. There are many reasons for the limited literature on binaural impairment, both for some of the conditions discussed and for many of those not discussed. The most important is that there are substantial challenges associated with analyzing "nature's experiments." The most difficult obstacle is that, unlike in the laboratory, the perturbations of the system are not uniform and are not easily documented. This is why it is of great benefit to develop animal and computational models, where clear relationships can be established between internal modifications of the system and externally measured values. In addition, it is of great value to take advantage of existing human brain imaging technologies and push for the development of new methods that will allow the binaural system to be more clearly revealed.

The other significant obstacle to taking advantage of the binaural impairments imposed by disease, injury, and natural biological processes is one of scientific and clinical culture, rather than techniques or knowledge. While some of the literature cited above reveals collaborations among clinicians and scientists and publications of clinical and basic research in the same journals, much more of it does not. The early literature is striking for a number of reasons, including the creativity and innovation shown in the methods and the insight revealed by the scientists. One aspect that should not be overlooked, however, is the degree to which the work was being done by clinician scientists, testing their own patients using cutting-edge methods. Progress in clinical research on binaural impairment depends on using the newest approaches in testing binaural hearing to better understand the abilities of large numbers of patients with similar disease states, as revealed by the best clinical metrics available. To do this requires us to attend the same conferences, publish in the same journals, and collaborate on grant applications together. Only in this way can new approaches for clinical care be developed and new insight gained into the ways that the binaural system can change its functioning in the course of a human lifetime.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

This work was supported by the NIH NIDCD R01 DC 015051 to FG.

ACKNOWLEDGMENTS

The author is grateful for the help of Virginia Best, Steve Colburn, Anna Diedesch, Nathan Spencer, Michelle Molis, and Nirmal Srinivasan in thinking through the issues covered in this review.

REFERENCES

Abel, S. M., Giguère, C., Consoli, A., and Papsin, B. C. (2000). The effect of aging on horizontal plane sound localization. *J. Acoust. Soc. Am.* 108, 743–752.

Akeroyd, M. A., and Whitmer, W. M. (2016). "Spatial hearing and hearing aids," in *Hearing Aids*, eds G. R. Popelka, B. C. J. Moore, R. R. Fay, and A. N. Popper (Cham: Springer International Publishing), 181–215. doi: 10.1007/978-3-319-33036-5_7

Anderson, S., Ellis, R., Mehta, J., and Goupell, M. J. (2018). Age-related differences in binaural masking level differences: behavioral and electrophysiological evidence. *J. Neurophysiol.* 120, 2939–2952. doi: 10.1152/jn.00255.2018

ANSI/ASA S3.6 (2018). *Specification for Audiometers*. Melville, NY: Acoustical Society of America.

Arbogast, T. L., Mason, C. R., and Kidd, G. (2005). The effect of spatial separation on informational masking of speech in normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.* 117, 2169–2180. doi: 10.1121/1.1861598

Baltzell, L. S., Swaminathan, J., Cho, A. Y., Lavandier, M., and Best, V. (2020). Binaural sensitivity and release from speech-on-speech masking in listeners with and without hearing loss. *J. Acoust. Soc. Am.* 147, 1546–1561. doi: 10.1121/10.0000812

Belinchon, A., Garrigues, H. P., Tenias, J. M., and Lopez, A. (2011). Hearing assessment in Menière's disease. *Laryngoscope* 121, 622–626. doi: 10.1002/lary.21335

Bernstein, L. R., and Trahiotis, C. (2016). Behavioral manifestations of audiometrically-defined "slight" or "hidden" hearing loss revealed by measures of binaural detection. *J Acoust. Soc. Am.* 140, 3540–3548. doi: 10.1121/1.4966113

Bernstein, L. R., and Trahiotis, C. (2018). Effects of interaural delay, center frequency, and no more than "slight" hearing loss on precision of binaural processing: empirical data and quantitative modeling. *J. Acoust. Soc. Am.* 144, 292–307. doi: 10.1121/1.5046515

Bernstein, L. R., and Trahiotis, C. (2019). No more than "slight" hearing loss and degradations in binaural processing. *J. Acoust. Soc. Am.* 145, 2094–2102. doi: 10.1121/1.5096652

Besser, J., Festen, J. M., Goverts, S. T., Kramer, S. E., and Pichora-Fuller, M. K. (2015). Speech-in-Speech listening on the LiSN-S test by older adults with good audiograms depends on cognition and hearing acuity at high frequencies. *Ear Hear.* 36, 24–41. doi: 10.1097/AUD.0000000000000096

Best, V., Carlile, S., Kopèo, N., and van Schaik, A. (2011). Localization in speech mixtures by listeners with hearing loss. *J. Acoust. Soc. Am.* 129, EL210–EL215. doi: 10.1121/1.3571534

Best, V., Mason, C. R., Swaminathan, J., Roverud, E., and Kidd, G. (2017). Use of a glimpsing model to understand the performance of listeners with and without hearing loss in spatialized speech mixtures. *J. Acoust. Soc. Am.* 141, 81–91. doi: 10.1121/1.4973620

Best, V., and Swaminathan, J. (2019). Revisiting the detection of interaural time differences in listeners with hearing loss. *J. Acoust. Soc. Am.* 145, EL508–EL513. doi: 10.1121/1.5111065

Best, V., Thompson, E. R., Mason, C. R., and Kidd, G. (2013). An energetic limit on spatial release from masking. *J. Assoc. Res. Otolaryngol.* 14, 603–610. doi: 10.1007/s10162-013-0392-1

Bocca, E., and Antonelli, A. R. (1976). Masking level difference: another tool for the evaluation of peripheral and cortical defects. *Audiology* 15, 480–487. doi: 10.3109/00206097609071809

Bronkhorst, A. W., and Plomp, R. (1989). Binaural speech intelligibility in noise for hearing-impaired listeners. *J. Acoust. Soc. Am.* 86, 1374–1383. doi: 10.1121/1.398697

Brungart, D. S., Cohen, J. I., Zion, D., and Romigh, G. (2017). The localization of non-individualized virtual sounds by hearing impaired listeners. *J. Acoust. Soc. Am.* 141, 2870–2881. doi: 10.1121/1.4979462

Buchholz, J. M., and Best, V. (2020). Speech detection and localization in a reverberant multitalker environment by normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.* 147, 1469–1477. doi: 10.1121/10.0000844

Butler, R. A., and Kluskens, L. (1971). The influence of phase inversion on the auditory evoked response. *Audiology* 10, 353–357. doi: 10.3109/00206097109072572

Carr, C. E., and Konishi, M. (1990). A circuit for detection of interaural time differences in the brain stem of the barn owl. *J. Neurosci.* 10, 3227–3246. doi: 10.1523/JNEUROSCI.10-10-03227.1990

Colburn, H. S. (1982). Binaural interaction and localization with various hearing impairments. *Scand. Audiol. Suppl.* 15, 27–45.

Colburn, H. S., and Durlach, N. I. (1965). Time-Intensity relations in binaural unmasking. *J. Acoust. Soc. Am.* 38, 93–103. doi: 10.1121/1.1909625

Dobreva, M. S., O'Neill, W. E., and Paige, G. D. (2011). Influence of aging on human sound localization. *J. Neurophysiol.* 105, 2471–2486.

Dreyer, A., and Delgutte, B. (2006). Phase locking of auditory-nerve fibers to the envelopes of high-frequency sounds: implications for sound localization. *J. Neurophysiol.* 96, 2327–2341. doi: 10.1152/jn.00326.2006

Dubno, J. R., Ahlstrom, J. B., and Horwitz, A. R. (2008). Binaural advantage for younger and older adults with normal hearing. *J.Speech Lang. Hear. Res.* 51, 539–556. doi: 10.1044/1092-4388(2008/039)

Duquesnoy, A. J. (1983). Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons. *J. Acoust. Soc. Am.* 74, 739–743. doi: 10.1121/1.389859

Durlach, N. I., Thompson, C. L., and Colburn, H. S. (1981). Binaural interaction in impaired listeners: a review of past research. *Audiology* 20, 181–211. doi: 10.3109/00206098109072694

Eddins, A. C., and Eddins, D. A. (2018). Cortical correlates of binaural temporal processing deficits in older adults. *Ear Hear.* 39, 594–604. doi: 10.1097/AUD.0000000000000518

Eddins, A. C., Ozmeral, E. J., and Eddins, D. A. (2018). How aging impacts the encoding of binaural cues and the perception of auditory space. *Hear. Res.* 369, 79–89. doi: 10.1016/j.heares.2018.05.001

Ellinger, R. L., Jakien, K. M., and Gallun, F. J. (2017). The role of interaural differences on speech intelligibility in complex multi-talker environments. *J. Acoust. Soc. Am.* 141, EL170–EL176. doi: 10.1121/1.4976113

Engle, J., and Recanzone, G. H. (2013). Characterizing spatial tuning functions of neurons in the auditory cortex of young and aged monkeys: a new perspective on old data. *Front. Aging Neurosci.* 4:36. doi: 10.3389/fnagi.2012.00036

Felix, R. A., Gourévitch, B., and Portfors, C. V. (2018). Subcortical pathways: towards a better understanding of auditory disorders. *Hear. Res.* 362, 48–60. doi: 10.1016/j.heares.2018.01.008

Fowler, C. G., and Mikami, C. M. (1992). The late auditory evoked potential masking-level difference as a function of noise level. *J. Speech Lang. Hear. Res.* 35, 216–221. doi: 10.1044/jshr.3501.216

Freigang, C., Richter, N., Rübsamen, R., and Ludwig, A. A. (2015). Age-related changes in sound localisation ability. *Cell Tissue Res.* 361, 371–386. doi: 10.1007/s00441-015-2230-8

Füllgrabe, C. (2013). Age-Dependent changes in temporal-fine-structure processing in the absence of peripheral hearing loss. *Am. J. Audiol.* 22, 313–315. doi: 10.1044/1059-0889(2013/12-0070)

Füllgrabe, C., and Moore, B. C. J. (2018). The association between the processing of binaural temporal-fine-structure information and audiometric threshold and age: a meta-analysis. *Trends Hear.* 22:2331216518797259. doi: 10.1177/2331216518797259

Furst, M., Aharonson, V., Levine, R. A., Fullerton, B. C., Tadmor, R., Pratt, H., et al. (2000). Sound lateralization and interaural discrimination.: effects of brainstem infarcts and multiple sclerosis lesions. *Hear. Res.* 143, 29–42. doi: 10.1016/S0378-5955(00)00019-8

Furst, M., Levine, R. A., Korczyn, A. D., Fullerton, B. C., Tadmor, R., and Algom, D. (1995). Brainstem lesions and click lateralization in patients with multiple sclerosis. *Hear. Res.* 82, 109–124. doi: 10.1016/0378-5955(94)00170-U

Gabriel, K. J., Koehnke, J., and Colburn, H. S. (1992). Frequency dependence of binaural performance in listeners with impaired binaural hearing. *J. Acoust. Soc. Am.* 91, 336–347. doi: 10.1121/1.402776

Gallun, F. J., and Best, V. (2020). "Age-Related changes in segregation of sound sources," in *Aging and Hearing: Causes and Consequences*, eds K. S. Helfer, E. L. Bartlett, A. N. Popper, and R. R. Fay (Cham: Springer International Publishing), 143–171. doi: 10.1007/978-3-030-49367-7_7

Gallun, F. J., Diedesch, A. C., Kampel, S. D., and Jakien, K. M. (2013). Independent impacts of age and hearing loss on spatial release in a complex auditory environment. *Front. Neurosci.* 7:252. doi: 10.3389/fnins.2013.00252

Gallun, F. J., Diedesch, A. C., Kubli, L. R., Walden, T. C., Folmer, R. L., Lewis, M. S., et al. (2012). Performance on tests of central auditory processing by individuals exposed to high-intensity blasts. *J. Rehabil. Res. Dev.* 49:1005. doi: 10.1682/JRRD.2012.03.0038

Gallun, F. J., and Jakien, K. M. (2019). "The ability to allocate attentional resources to a memory task predicts speech-on-speech masking for older listeners," in *Proceedings of the 23rd International Congress on Acoustics*, Aachen.

Gallun, F. J., Lewis, M. S., Folmer, R. L., Hutter, M., Papesh, M. A., Belding, H., et al. (2016). Chronic effects of exposure to high-intensity blasts: results of tests of central auditory processing. *J. Rehabil. Res. Dev.* 53, 705–720.

Gallun, F. J., McMillan, G. P., Kampel, S. D., Jakien, K. M., Srinivasan, N. K., Stansell, M. M., et al. (2015). Verification of an automated headphone-based test of spatial release from masking. *Proc. Meet. Acoust.* 25:050001. doi: 10.1121/2.0000165

Gallun, F. J., McMillan, G. P., Molis, M. R., Kampel, S. D., Dann, S. M., and Konrad-Martin, D. L. (2014). Relating age and hearing loss to monaural, bilateral, and binaural temporal sensitivity1. *Front. Neurosci.* 8:172. doi: 10.3389/fnins.2014.00172

Gelfand, S. A., Ross, L., and Miller, S. (1988). Sentence reception in noise from one versus two sources: effects of aging and hearing loss. *J. Acoust. Soc. Am.* 83, 248–256. doi: 10.1121/1.396426

Glyde, H., Cameron, S., Dillon, H., Hickson, L., and Seeto, M. (2013). The effects of hearing impairment and aging on spatial processing. *Ear Hear.* 34:15. doi: 10.1097/AUD.0b013e3182617f94

Green, D. M., and Swets, J. A. (1974). *Signal Detection Theory and Psychophysics*. New York, NY: Wiley.

Green, G. G. R., Heffer, J. S., and Ross, D. A. (1976). The dectability of apparent source movement effected by interaural phase modulation. *J. Physiol.* 260:49.

Greene, T. C. (1929). The ability to localize sound: a study of binaural hearing in patients with tumor of the brain. *Arch. Surg.* 18:1825. doi: 10.1001/archsurg.1929.01140130927061

Grose, J. H., and Mamo, S. K. (2010). Processing of temporal fine structure as a function of age. *Ear Hear.* 31, 755–760. doi: 10.1097/AUD.0b013e3181e627e7

Grose, J. H., and Mamo, S. K. (2012). Frequency modulation detection as a measure of temporal processing: age-related monaural and binaural effects. *Hear. Res.* 294, 49–54. doi: 10.1016/j.heares.2012.09.007

Gutschalk, A., and Dykstra, A. R. (2015). "Chapter 31—Auditory neglect and related disorders," in *Handbook of Clinical Neurology*, Vol. 129, eds M. J. Aminoff, F. Boller, and D. F. Swaab (Amsterdam: Elsevier), 557–571. doi: 10.1016/B978-0-444-62630-1.00031-7

Häusler, R., Colburn, S., and Marr, E. (1983). Sound localization in subjects with impaired hearing: spatial-discrimination and interaural-discrimination tests. *Acta Otolaryngol.* 96(Suppl. 400), 1–62.

Häusler, R., and Levine, R. A. (1980). Brain stem auditory evoked potentials are related to interaural time discrimination in patients with multiple sclerosis. *Brain Res.* 191, 589–594. doi: 10.1016/0006-8993(80)91312-8

Häusler, R., and Levine, R. A. (2000). Auditory dysfunction in stroke. *Acta Otolaryngol.* 120, 689–703. doi: 10.1080/000164800750000207

Hawkins, D. B., and Wightman, F. L. (1980). Interaural time discrimination ability of listeners with sensorineural hearing loss. *Audiology* 19, 495–507. doi: 10.3109/00206098009070081

Hawley, M. L., Litovsky, R. Y., and Culling, J. F. (2004). The benefit of binaural hearing in a cocktail party: effect of location and type of interferer. *J. Acoust. Soc. Am.* 115, 833–843. doi: 10.1121/1.1639908

Haywood, N. R., Undurraga, J. A., Marquardt, T., and McAlpine, D. (2015). A comparison of two objective measures of binaural processing: the interaural phase modulation following response and the binaural interaction component. *Trends Hear.* 19, 1–17. doi: 10.1177/2331216515619039

Hedwig, B., and Stumpner, A. (2016). "Central neural processing of sound signals in insects," in *Insect Hearing*, eds G. S. Pollack, A. C. Mason, A. N. Popper, and R. R. Fay (Cham: Springer International Publishing), 177–214. doi: 10.1007/978-3-319-28890-1_8

Hoover, E. C., Souza, P. E., and Gallun, F. J. (2017). Auditory and cognitive factors associated with speech-in-noise complaints following mild traumatic brain injury. *J. Am. Acad. Audiol.* 28, 325–339.

Hopkins, K., and Moore, B. C. J. (2009). The contribution of temporal fine structure to the intelligibility of speech in steady and modulated noise. *J. Acoust. Soc. Am.* 125, 442–446. doi: 10.1121/1.3037233

Jakien, K. M., and Gallun, F. J. (2018). Normative data for a rapid, automated test of spatial release from masking. *Am. J. Audiol.* 27, 529–538. doi: 10.1044/2018_AJA-17-0069

Jakien, K. M., Kampel, S. D., Stansell, M. M., and Gallun, F. J. (2017). Validating a rapid, automated test of spatial release from masking. *Am. J. Audiol.* 26, 507–518. doi: 10.1044/2017_AJA-17-0013

Jerger, J., Brown, D., and Smith, S. (1984). Effect of peripheral hearing loss on the masking level difference. *Arch. Otolaryngol.* 110, 290–296. doi: 10.1001/archotol.1984.00800310014003

Jerger, J. F., Oliver, T. A., Chmiel, R. A., and Rivera, V. M. (1986). Patterns of auditory abnormality in multiple sclerosis: exemples de cas auditifs anormaux dans les scléroses en plaques. *Audiology* 25, 193–209. doi: 10.3109/00206098609078386

Jongkees, L. B., and Van der Veer, R. A. (1957). Directional hearing capacity in hearing disorders. *Acta Otolaryngol.* 48, 465–474.

Jongkees, L. B. W., and Van der Veer, R. A. D. (1958). On directional sound localization in unilateral deafness and its explanation. *Acta Otolaryngol.* 49, 119–131. doi: 10.3109/00016485809134735

Juarez-Salinas, D. L., Engle, J. R., Navarro, X. O., and Recanzone, G. H. (2010). Hierarchical and serial processing in the spatial auditory cortical pathway is degraded by natural aging. *J. Neurosci.* 30, 14795–14804. doi: 10.1523/JNEUROSCI.3393-10.2010

Katz, J. (2014). *Handbook of Clinical Audiology*, 7th Edn. Philadelphia, PA: Lippincott William &Wilkins.

King, A., Hopkins, K., and Plack, C. J. (2014). The effects of age and hearing loss on interaural phase difference discrimination. *J. Acoust. Soc. Am.* 135, 342–351. doi: 10.1121/1.4838995

King, A., Hopkins, K., Plack, C. J., Pontoppidan, N. H., Bramsløw, L., Hietkamp, R. K., et al. (2017). The effect of tone-vocoding on spatial release from masking for old, hearing-impaired listeners. *J. Acoust. Soc. Am.* 141, 2591–2603. doi: 10.1121/1.4979593

Koehnke, J., Colburn, H. S., and Durlach, N. I. (1986). Performance in several binaural-interaction experiments. *J. Acoust. Soc. Am.* 79, 1558–1562. doi: 10.1121/1.393682

Koehnke, J., Culotta, C. P., Hawley, M. L., and Colburn, H. S. (1995). Effects of reference interaural time and intensity differences on binaural performance in listeners with normal and impaired hearing. *Ear Hear.* 16, 331–353.

Koerner, T. K., Muralimanohar, R. K., Gallun, F. J., and Billings, C. J. (2020). Age-Related deficits in electrophysiological and behavioral measures of binaural temporal processing. *Front. Audit. Neurosci.* 14:578566. doi: 10.3389/fnins.2020.578566

Kubiak, A. M., Rennies, J., Ewert, S. D., and Kollmeier, B. (2020). Prediction of individual speech recognition performance in complex listening conditions. *J. Acoust. Soc. Am.* 147, 1379–1391. doi: 10.1121/10.0000759

Kubli, L. R., Brungart, D., and Northern, J. (2018). Effect of blast injury on auditory localization in military service members. *Ear Hear.* 39:457. doi: 10.1097/AUD.0000000000000517

Lacher-Fougère, S., and Demany, L. (2005). Consequences of cochlear damage for the detection of interaural phase differences. *J. Acoust. Soc. Am.* 118, 2519–2526. doi: 10.1121/1.2032747

Le Goff, N. L., Buchholz, J. M., and Dau, T. (2013). "Modeling horizontal localization of complex sounds in the impaired and aided impaired auditory system," in *The Technology of Binaural Listening*, ed. J. Blauert (Berlin: Springer), 121–144. doi: 10.1007/978-3-642-37762-4_5

Levine, R. A., Gardner, J. C., Fullerton, B. C., Stufflebeam, S. M., Carlisle, E. W., Furst, M., et al. (1993a). Effects of multiple sclerosis brainstem lesions on sound lateralization and brainstem auditory evoked potentials. *Hear. Res.* 68, 73–88. doi: 10.1016/0378-5955(93)90066-A

Levine, R. A., Gardner, J. C., Fullerton, B. C., Stufflebeam, S. M., Furst, M., and Rosen, B. R. (1994). Multiple sclerosis lesions of the auditory pons are not silent. *Brain* 117, 1127–1141. doi: 10.1093/brain/117.5.1127

Levine, R. A., Gardner, J. C., Stufflebeam, S. M., Fullerton, B. C., Carlisle, E. W., Furst, M., et al. (1993b). Binaural auditory processing in multiple sclerosis subjects. *Hear. Res.* 68, 59–72. doi: 10.1016/0378-5955(93)90065-9

Licklider, J. (1948). The influence of interaural phase relations upon the masking of speech by white noise. *J. Acoust. Soc. Am.* 20, 150–159.

Maier, J. K., Kindermann, T., Grothe, B., and Klump, G. M. (2008). Effects of omni-directional noise-exposure during hearing onset and age on auditory spatial resolution in the Mongolian gerbil (*Meriones unguiculatus*)—A behavioral approach. *Brain Res.* 1220, 47–57. Scopus. doi: 10.1016/j.brainres.2008.01.083

Mao, J., Koch, K.-J., Doherty, K. A., and Carney, L. H. (2015). Cues for diotic and dichotic detection of a 500-Hz tone in noise vary with hearing loss. *J. Assoc. Res. Otolaryngol.* 16, 507–521. doi: 10.1007/s10162-015-0518-8

Marrone, N., Mason, C. R., and Kidd, G. (2008). The effects of hearing loss and age on the benefit of spatial separation between multiple talkers in reverberant rooms. *J. Acoust. Soc. Am.* 124, 3064–3075. doi: 10.1121/1.2980441

McFadden, D. (1968). Masking-level differences determined with and without interaural disparities in masker intensity. *J. Acoust. Soc. Am.* 44, 212–223. doi: 10.1121/1.1911057

Melnick, W., and Bilger, R. C. (1965). Hearing loss and auditory lateralization. *J. Speech Hear. Res.* 8, 3–12. doi: 10.1044/jshr.0801.03

Mills, A. W. (1958). On the minimum audible angle. *J. Acoust. Soc. Am.* 30, 237–246. doi: 10.1121/1.1909553

Moncada-Torres, A., Joshi, S. N., Prokopiou, A., Wouters, J., Epp, B., and Francart, T. (2018). A framework for computational modelling of interaural time difference discrimination of normal and hearing-impaired listeners. *J. Acoust. Soc. Am.* 144, 940–954. doi: 10.1121/1.5051322

Moore, B. C. J. (2020). Effects of hearing loss and age on the binaural processing of temporal envelope and temporal fine structure information. *Hear. Res.* 402:107991. doi: 10.1016/j.heares.2020.107991

Mueller, H. G., and Beck, W. G. (1987). Brainstem level test results following head injury. *Semin. Hear.* 8, 253–260.

Musiek, F. E., Gollegly, K. M., Kibbe, K. S., and Reeves, A. G. (1989). Electrophysiologic and behavioral auditory findings in multiple sclerosis. *Otol. Neurotol.* 10, 343–344.

Neher, T., Laugesen, S., Søgaard Jensen, N., and Kragelund, L. (2011). Can basic auditory and cognitive measures predict hearing-impaired listeners' localization and spatial speech recognition abilities? *J. Acoust. Soc. Am.* 130, 1542–1558. doi: 10.1121/1.3608122

Neher, T., Wagener, K. C., and Latzel, M. (2017). Speech reception with different bilateral directional processing schemes: influence of binaural hearing, audiometric asymmetry, and acoustic scenario. *Hear. Res.* 353, 36–48. doi: 10.1016/j.heares.2017.07.014

Noble, W., Byrne, D., and Lepage, B. (1994). Effects on sound localization of configuration and type of hearing impairment. *J. Acoust. Soc. Am.* 95, 992–1005. doi: 10.1121/1.408404

Noble, W., Byrne, D., and Ter-Horst, K. (1997). Auditory localization, detection of spatial separateness, and speech hearing in noise by hearing impaired listeners. *J. Acoust. Soc. Am.* 102, 2343–2352. doi: 10.1121/1.419618

Olsen, W. O., and Noffsinger, D. (1976). Masking level differences for cochlear and brain stem lesions. *Ann. Otol. Rhinol. Laryngol.* 85, 820–825. doi: 10.1177/000348947608500611

Olsen, W. O., Noffsinger, D., and Carhart, R. (1976). Masking level differences encountered in clinical populations. *Audiology* 15, 287–301. doi: 10.3109/00206097609071789

Papesh, M. A., Folmer, R. L., and Gallun, F. J. (2017). Cortical measures of binaural processing predict spatial release from masking performance. *Front. Hum. Neurosci.* 11:124. doi: 10.3389/fnhum.2017.00124

Peissig, J., and Kollmeier, B. (1997). Directivity of binaural noise reduction in spatial multiple noise-source arrangements for normal and impaired listeners. *J. Acoust. Soc. Am.* 101, 1660–1670. doi: 10.1121/1.418150

Pichora-Fuller, M. K., and Schneider, B. A. (1991). Masking-Level differences in the elderly. *J. Speech Lang. Hear. Res.* 34, 1410–1422. doi: 10.1044/jshr.3406.1410

Pichora-Fuller, M. K., and Schneider, B. A. (1992). The effect of interaural delay of the masker on masking-level differences in young and old adults. *J. Acoust. Soc. Am.* 91, 2129–2135. doi: 10.1121/1.403673

Pierce, A. H. (1901). *Studies in Auditory and Visual Space Perception*. New York, NY: Longmans, Green, and Company.

Przewoźny, T., Gójska-Grymajło, A., and Gąsecki, D. (2015a). Auditory spatial deficits in the early stage of ischemic cerebral stroke. *J. Stroke Cerebrovasc. Dis.* 24, 1905–1916. doi: 10.1016/j.jstrokecerebrovasdis.2015.05.001

Przewoźny, T., Gójska-Grymajło, A., Szmuda, T., and Markiet, K. (2015b). Auditory spatial deficits in brainstem disorders. *Neurol. Neurochir. Pol.* 49, 401–411. doi: 10.1016/j.pjnns.2015.10.001

Rayleigh, L. (1907). On our perception of sound direction. *Philos. Mag.* 13, 214–232.

Ross, B., Fujioka, T., Tremblay, K. L., and Picton, T. W. (2007). Aging in binaural hearing begins in mid-life: evidence from cortical auditory-evoked responses to changes in interaural phase. *J. Neurosci.* 27, 11172–11178. doi: 10.1523/JNEUROSCI.1813-07.2007

Roup, C. M., and Powell, J. (2016). Binaural processing in adults with a history of traumatic brain injury. *J. Acoust. Soc. Am.* 140, 3268–3268. doi: 10.1121/1.4970370

Saunders, G. H., Frederick, M. T., Arnold, M., Silverman, S., Chisolm, T. H., Myers, P. et al. (2015). Auditory difficulties in blast-exposed Veterans with clinically normal hearing. *J. Rehabil. Res. Dev.* 52, 343–360. doi: 10.1682/JRRD.2014.11.0275

Sedge, R. K. (1987). Introduction to the vietnam head injury study. *Semin. Hear.* 8, 191–213.

Smith-Olinde, L., Besing, J., and Koehnke, J. (2004). Interference and enhancement effects on interaural time discrimination and level discrimination in listeners with normal hearing and those with hearing loss. *Am. J. Audiol.* 13, 80–95. doi: 10.1044/1059-0889(2004/011)

Smoski, W. J., and Trahiotis, C. (1986). Discrimination of interaural temporal disparities by normal-hearing listeners and listeners with high-frequency sensorineural hearing loss. *J. Acoust. Soc. Am.* 79, 1541–1547. doi: 10.1121/1.393680

Sonoda, S., Mori, M., and Goishi, A. (2001). Pattern of localisation error in patients with stroke to sound processed by a binaural sound space processor. *J. Neurol. Neurosurg. Psychiatry* 70, 43–49. doi: 10.1136/jnnp.70.1.43

Spencer, N. J., Hawley, M. L., and Colburn, H. S. (2016). Relating interaural difference sensitivities for several parameters measured in normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.* 140, 1783–1799. doi: 10.1121/1.4962444

Srinivasan, N. K., Holtz, A., and Gallun, F. J. (2020). Comparing spatial release from masking using traditional methods and portable automated rapid testing iPad App. *Am. J. Audiol.* 29, 907–915. doi: 10.1044/2020_AJA-20-00078

Srinivasan, N. K., Jakien, K. M., and Gallun, F. J. (2016). Release from masking for small spatial separations: effects of age and hearing loss. *J. Acoust. Soc. Am.* 140, EL73–EL78. doi: 10.1121/1.4954386

Stecker, G. C., and Gallun, F. J. (2012). "Binaural hearing, sound localization, and spatial hearing," in *Translational Perspectives in Auditory Neuroscience: Normal Aspects of Hearing*, Vol. 383, eds K. L. Tremblay and R. F. Burkard (San Diego, CA: Plural Publishing, Inc), 383–433.

Venturi, J. B. (1796). Considérations sur la connaissance de l'étendue que nous donne le sens de l'ouïe. *Magasin Encyclopédique J. Sci. Lett. Arts* 3, 29–37.

Vercammen, C., Goossens, T., Undurraga, J., Wouters, J., and van Wieringen, A. (2018). Electrophysiological and behavioral evidence of reduced binaural temporal processing in the aging and hearing impaired human auditory system. *Trends Hear.* 22:2331216518785733. doi: 10.1177/2331216518785733

Wade, N., and Deutsch, D. (2008). Binaural hearing: before and after the stethophone. *Acoust. Today* 4, 16–27.

Wallach, H., Newman, E. B., and Rosenzweig, M. R. (1949). A precedence effect in sound localization. *J. Acoust. Soc. Am.* 21, 468–468. doi: 10.1121/1.1917119

Walsh, E. G. (1957). An investigation of sound localization in patients with neurological abnormalities. *Brain* 80, 222–250. doi: 10.1093/brain/80.2.222

Whiteford, K. L., Kreft, H. A., and Oxenham, A. J. (2017). Assessing the role of place and timing cues in coding frequency and amplitude modulation as a function of age. *J. Assoc. Res. Otolaryngol.* 18, 619–633. doi: 10.1007/s10162-017-0624-x

Wightman, F. L., and Kistler, D. J. (1989). Headphone simulation of free-field listening. I: stimulus synthesis. *J. Acoust. Soc. Am.* 85, 858–867. doi: 10.1121/1.397557

Witton, C., Green, G. G., Rees, A., and Henning, G. B. (2000). Monaural and binaural detection of sinusoidal phase modulation of a 500-Hz tone. *J. Acoust. Soc. Am.* 108, 1826–1833.

Yin, T. C. T. (2002). "Neural mechanisms of encoding binaural localization cues in the auditory brainstem," in *Integrative Functions in the Mammalian Auditory Pathway*, eds D. Oertel, R. R. Fay, and A. N. Popper (New York, NY: Springer), 99–159. doi: 10.1007/978-1-4757-3654-0_4

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

REVIEW
published: 15 June 2021
doi: 10.3389/fnins.2021.651209

Auditory Discrimination in Autism Spectrum Disorder

*Sarah Elizabeth Rotschafer**

Department of Biomedical Sciences, Mercer University School of Medicine, Savannah, GA, United States

Autism spectrum disorder (ASD) is increasingly common with 1 in 59 children in the United States currently meeting the diagnostic criteria. Altered sensory processing is typical in ASD, with auditory sensitivities being especially common; in particular, people with ASD frequently show heightened sensitivity to environmental sounds and a poor ability to tolerate loud sounds. These sensitivities may contribute to impairments in language comprehension and to a worsened ability to distinguish relevant sounds from background noise. Event-related potential tests have found that individuals with ASD show altered cortical activity to both simple and speech-like sounds, which likely contribute to the observed processing impairments. Our goal in this review is to provide a description of ASD-related changes to the auditory system and how those changes contribute to the impairments seen in sound discrimination, sound-in-noise performance, and language processing. In particular, we emphasize how differences in the degree of cortical activation and in temporal processing may contribute to errors in sound discrimination.

Keywords: autism, event-related potentials, sound discrimination, language, cortex

OPEN ACCESS

Edited by:
Dan Tollin,
University of Colorado, United States

Reviewed by:
Anu Sharma,
University of Colorado Boulder, United States
Ursula Koch,
Freie Universität Berlin, Germany

***Correspondence:**
Sarah Elizabeth Rotschafer
rotschafer_se@mercer.edu

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience

Received: *08 January 2021*
Accepted: *23 March 2021*
Published: *15 June 2021*

Citation:
Rotschafer SE (2021) Auditory Discrimination in Autism Spectrum Disorder. Front. Neurosci. 15:651209. doi: 10.3389/fnins.2021.651209

INTRODUCTION

Autism spectrum disorder (ASD) is a developmental condition that is characterized by abnormalities in social communication, restricted behavior, and repetitive behavior. People with ASD also experience altered sensory processing and show both hyper- and hypo-reactivity to sensory input (Heaton et al., 2008; American Psychiatric Association, 2013; Kujala et al., 2013). For reference, between 60 and 96% of people diagnosed with ASD report sensory sensitivities (Schauder and Bennetto, 2016; Kuiper et al., 2019). Responses to auditory stimuli are especially impacted in ASD, with increased sensitivity to noise and difficulty filtering sound from background noise as cardinal features of ASD (Jones et al., 2009; DePape et al., 2012). Behaviorally, individuals with ASD may also be hypersensitive to certain environmental noises, show a decreased tolerance of loud noises, and have a reduced ability to habituate to auditory stimuli (Rosenhall et al., 1999; Khalfa et al., 2004; O'Connor, 2012; Lawson et al., 2015; Bidet-Caulet et al., 2017; Hudac et al., 2018; Ruiz-Martinez et al., 2020). Changes in how sound is received in ASD can also manifest as atypical linguistic processing and comprehension. In addition to delayed language acquisition, individuals with ASD may also show an impaired ability to understand phrases or comprehend single words (Mitchell et al., 2006; Hudry et al., 2010). Ultimately, a significant portion of children with ASD are diagnosed as minimally verbal, and atypical auditory processing may be a contributing factor (Tager-Flusberg et al., 2009; Tager-Flusberg and Kasari, 2013).

These ASD-related changes in auditory behavior and linguistic processing are reflected in dysfunction at the cortical level. Cortical function in ASD is hypothesized to be impacted by an increase in endogenous cortical "noise," resulting from altered ratios of excitation to

inhibition within neural circuits (Rubenstein and Merzenich, 2003; Simmons et al., 2009; Sohal and Rubenstein, 2019). This review therefore will consider the possible causes of ASD-related irregularities in language processing with special attention to how people with ASD process language in natural environmental conditions. As an attempt to understand how impairments in language processing arise in ASD, we will begin by examining errors associated with ASD in processing simple stimuli at the cortical level. In particular, we will focus on the early cortical responses to simple stimuli. Next, we will compare how simple sounds and simple linguistic stimuli are processed, with emphasis placed on cortical responses that track changes in sound stimuli (the mismatch negativity and P300). Lastly, we will review how speech in noise stimuli is represented in ASD. We will discuss aspects of background sounds that make extracting language more difficult in ASD and features that impair linguistic targets' detection.

SIMPLE SOUNDS

A meta-analysis determined that 90% of individuals with ASD experience sensory abnormalities, with auditory hypersensitivity as the most common modality (Gomes et al., 2008). One way in which heightened auditory sensitivity to sound manifests is through pitch differentiation ability as shown through both electrophysiological and behavioral measures. Individuals with ASD perform better than typically developing participants on pitch discrimination tasks, as shown using same-same or same-different testing paradigms (Bonnel et al., 2010). Enhanced pitch discrimination, however, does not necessarily translate into superior linguistic processing ability. Reponses to simple sounds in ASD show atypical sensory peak amplitude and variable peak latencies that may ultimately impair how well individuals with ASD are able to decode language and process environmental noise (Oram Cardy et al., 2008; Port et al., 2016).

To probe cortical responses to sound in ASD, studies generally present participants with various simple sound targets and then measure cortical response using electroencephalograms (EEG) or magnetoencephalograms (MEG). Both EEGs and MEGs represent cortical auditory processing as a series of waveforms (positivities and negativities) that occur in a stereotyped sequential manner as different portions of the cortex become active in response to sound. Generally, EEGs confer greater temporal resolution of responses, while MEGs have superior spatial resolution. Here, we will discuss how "sensory peaks" (P1, N1/M100) respond to simple sounds in ASD, and in how changes in those sounds are represented by the mismatch negativity (MMN).

P1

P1 is an early cortical response to sound and is thought to reflect thalamocortical transmission along the ascending auditory pathway (Eggermont et al., 1997). In typically developing children, P1 amplitude can track stimulus complexity (amplitude increases with stimulus complexity) (Ceponiene et al., 2001), and can be modulated by arousal (Pratt et al., 2012). Several studies reported reduced P1 amplitude in ASD (Buchwald et al., 1992; Ceponiene et al., 2003a,b; Orekhova et al., 2008; Donkers et al., 2015). Poor early representation of auditory stimuli may subsequently impair the ability of participants with ASD to discriminate between sounds (Ruiz-Martinez et al., 2020). Moreover, in ASD, P1 shows an abnormal lack of modulation in response to changes in the temporal features of sound. In typically developing participants, increasing the presentation rate of a stimulus resulted in attenuation of the P1 amplitude. By contrast, individuals with ASD did not show any change in amplitude as the stimulus presentation rate was modulated, which suggests reduced sensitivity to changes in how sound is represented temporally (Buchwald et al., 1992). Ruiz-Martinez et al. (2020) found a lack of P1 response habituation in ASD to repetitive stimuli, which may indicate a reduced ability to predict and adapt to incoming stimuli, thus contributing to the enhanced auditory sensitivity seen in people with ASD (**Table 1**).

N1 and M100

When processing sound, EEGs and MEGs show an event approximately 100 ms after stimulus presentation known as the N1 (EEG) or M100 (MEG). This response is thought to represent activity at the auditory cortex, superior temporal gyrus, and auditory association areas. There are also data suggesting parietal and frontal cortex involvement (Naatanen and Picton, 1987). In typically developing people, the N1/M100 latency is longer in children and shortens as they age. In children with ASD however, M100 latencies are more variable. With regard to auditory response development, peak latencies in children with ASD were reported to change with age in a manner similar to results seen in typically developing children in the left hemisphere of the brain, but M100 latencies in the right hemisphere did not change with age (Gage et al., 2003b). Other studies cataloged unusually long N1/M100 latencies in a variety of testing paradigms (Oades et al., 1988; Bruneau et al., 1999; Seri et al., 1999; Korpilahti et al., 2007; Sokhadze et al., 2009; Roberts et al., 2010; Port et al., 2016; **Table 2**). Port et al. (2016) tested children with ASD once when they were between the ages of 6 and 11 years, then again 5 years later. They had participants passively listen to a series of pure tones and found delays in the M100 response latency at both time points and discovered a relationship between M100 delay and clinical ASD severity. Seri et al. (1999) specifically tested children with tuberous sclerosis and found delayed peak latency, showing that at least for tuberous sclerosis, the results from a specific sub-diagnosis of ASD were consistent with ASD results on the whole. Similarly, longer N1 latency was present in children with Asperger syndrome (Korpilahti et al., 2007). Oades et al.'s (1988) approach was slightly different from the other listed studies in that they asked children with ASD to perform a task while listening to test stimuli—the participants were instructed to press a button when they heard a target tone stimulus and ignore non-target distractor tones. Interestingly, Oades et al. (1988) found that N1 latencies were longer in response to non-target stimuli but shorter in response to target stimuli. It is possible then that variation in N1 latencies may be a correlate of abnormalities in how children with ASD direct their attention when performing auditory tasks (Oades et al., 1988).

TABLE 1 | P1 peak amplitude in response to simple sound stimuli.

P1 Simple Sounds	Research	Participants
Reduced amplitude	Buchwald et al., 1992	Adults with ASD
	Donkers et al., 2015	4–12 year old males with ASD
	Orekhova et al., 2008	4–8 year old males and females with ASD
	Ceponiene et al., 2003b	6–12 year old males with high functioning ASD
	Ruiz-Martinez et al., 2020	5–11 year old males and one female with ASD; included minimally verbal children

Notable participant features and sub-diagnoses are underlined.

TABLE 2 | N1 peak in response to simple sound stimuli.

N1 simple sounds	Research	Participants
Greater Amplitude	Flagg et al., 2005*	8–17 year old males with ASD
	Gage et al., 2003a*	8–14 year old males with ASD
	Dawson et al., 1986*	6–18 year old males with ASD; some with intellectual impairment
	Rojas et al., 2001	Adults with fragile X syndrome
	Van der Molen et al., 2012a,b	18–42 year old males with fragile X syndrome
	Castren et al., 2003	7–13 year old males with fragile X syndrome
Longer Latency	Port et al., 2016	Mean age 8 years old at initial recruitment; males
	Roberts et al., 2010*	Mean age 10 years old; sex not reported
	Sokhadze et al., 2009*	9–27 year old males and one female with high functioning ASD
	Bruneau et al., 1999*	4–8 year old males and females with intellectual impairment and ASD
	Seri et al., 1999	7–10 year old with tuberous sclerosis; sex not reported
	Korpilahti et al., 2007	9–12 year old males with Asperger syndrome
	Oram Cardy et al., 2008*	7–18 year old males and females with ASD and/or Asperger syndrome
Shorter/Atypical Latency	Gage et al., 2003b	8–14 year old males with ASD
	Oades et al., 1988	6–18 year old males and one female
	Ferri et al., 2003	6–19 year old males with ASD and intellectual impairment

Specific sub-diagnoses and notable features of participants are underlined. Asterisks indicate that the reported result was only seen in the right hemisphere.

Ferri et al. (2003) reported shorter peak latencies in children with ASD, though they noted that their study used children diagnosed with ASD and intellectual impairment as participants, while similar research tended to base their results on high functioning children with ASD. As such, they raise the possibility that the degree of intellectual impairment present in children with ASD may impact response latencies (Ferri et al., 2003). However, Bruneau et al. (2003) also recruited children with ASD and intellectual impairment but found the opposite result—longer peak latencies. As a point of differentiation, Bruneau et al. (2003) and Ferri et al. (2003) used distinctly different age groups, meaning that multiple factors may act in combination to influence auditory responsiveness in ASD.

Work examining N1/M100 in people with ASD frequently reported abnormalities in lateralization. In typically developing people, the left temporal cortex response to sound is generally greater than that of the right temporal cortex (Eyler et al., 2012). However, several studies that included participants with ASD reported prolonged latencies specific to the right hemisphere (**Table 2**; Bruneau et al., 2003; Gage et al., 2003b; Oram Cardy et al., 2008; Sokhadze et al., 2009; Roberts et al., 2010), and an overall increase in right hemisphere responsiveness to sound (Dawson et al., 1986; Gage et al., 2003a,b; Flagg et al., 2005). These findings correlate well with MRI work showing superior temporal gyrus activity to be symmetrical (as opposed to showing a leftward bias) in adults with ASD as a result of increased right hemisphere superior temporal gyrus volume (Jou et al., 2010). Prolonged right hemisphere latencies are consistent with results showing developmental delays in ASD and suggest that some of the errors seen ASD in sound processing may stem from abnormalities in gross neuroanatomy.

As a general statement, people with ASD seem to show greater peak amplitudes and longer peak latencies for N1/M100. This trend is informative with regard to how auditory function may be fundamentally altered in ASD at early stages of cortical processing. Delays in peak latency correlate with stimulus complexity and the recruitment of neural resources. In which case, longer peak latencies and broader recruitment of neural resources could mean that individuals with ASD find simple stimuli to be more complex than control participants generally do (Lepisto et al., 2008). These findings may also be a byproduct of a loss of long-range connections in the ASD brain, causing people with ASD to rely more heavily on local connections to process sound stimuli (Jou et al., 2010).

MMN and MMF

The mismatch negativity is a waveform that reflects changes in stimuli; generally, it is considered to act as an automatic orienting reflex that marks changes in an environment. The MMN can be modulated by participants focusing their attention

on a stimulus, but can still be elicited when attention is not directed at stimuli (Alho, 1995; Naatanen and Alho, 1995). The MMN is thought to represent activity at the auditory cortex that has been supplemented by inputs from the frontal lobe. It may also reflect activity at the hippocampus and thalamus (Alho, 1995; Garrido et al., 2009). In studies using magnetoencephalography, this waveform is referred to as the magnetic mismatch field latency (MMF). Because MMN/MMF tracks changes in stimuli, it is often studied using some variation of an "oddball" task, where participants listen to a stream of identical standard sounds that have a target stimulus (a stimulus that deviates from the standard sounds in some metric) or novel stimulus interleaved.

Research that tested the MMN amplitude in response to pure tone stimuli in ASD reported a range of findings that may reflect variation in participants' sub-diagnosis, degree of intellectual impairment, tolerance of change, or age (Gomot et al., 2002, 2011; Ceponiene et al., 2003b; Ferri et al., 2003; Tecchio et al., 2003; Vlaskamp et al., 2017; **Table 3**). Ferri et al. (2003) and Gomot et al. (2011) found that children with ASD and intellectual impairments had larger MMN responses to deviant stimuli, although a MEG study that also focused on low functioning individuals with ASD found reduced MMF amplitude (Tecchio et al., 2003). While both Ferri et al. (2003) and Tecchio et al. (2003) surveyed people with ASD and intellectual impairment, Ferri et al. (2003) put more focus on younger individuals (6–19 years) and excluded participants with known chromosomal abnormalities (fragile X syndrome, tuberous sclerosis, etc.). Tecchio et al. (2003) drew participants from a greater range of ages (8–32 years) and did not seek to exclude certain sub-diagnoses. Gomot et al. (2011) also saw heighted MMN responses in ASD and found that MMN amplitude was associated with participants' ability to tolerate change. As such, the variability in MMN amplitude results may reflect age-related differences and/or differences in how certain sensory impairments manifest in ASD. Consistent with this line of thought, work that tested children with high functioning children ASD generally did not see any significant MMN differences, suggesting that the sub-diagnosis and the degree of intellectual impairment of participants could significantly impact the MMN profile (Gomot et al., 2002; Ceponiene et al., 2003b). Notably, when specific conditions on the autism spectrum were considered, the results were uneven. MMN peak amplitude was reduced in fragile X syndrome (Van der Molen et al., 2012b), and abnormally prolonged Rett syndrome (Stauder et al., 2006; Foxe et al., 2016), lending credibility to the notion that subtle variations sub-diagnosis phenotype may drive some of the inconsistencies found in MMN results. Vlaskamp et al. (2017) also sought to explain some of the variability in ASD MMN findings by testing the ASD response to frequency and duration deviants in an oddball task using a relatively large number of participants in a more discrete age range (8–12 years). That work found that MMN amplitude was reduced in individuals with ASD to both frequency and duration deviants (Vlaskamp et al., 2017), which may reflect a lessened ability to track certain types of changes in auditory stimuli in ASD.

Similar to research describing MMN/MMF peak amplitude, data describing MMN/MMF latency vary by the study design and the participant composition. Consistent with the superior pitch discrimination observed in ASD, Gomot et al. (2002) found that MMN latency was shorter in children with ASD. There was also evidence suggesting that the MMN response in ASD may receive contributions from generators other than those used in typically developing controls, which may partially explain the observed change in MMN latency (Gomot et al., 2002). By contrast, other work that specifically examined the relationship between auditory sensitivity (as determined by performing a sensory profile assessment) and MMF latency found that participants who had ASD and atypical auditory sensitivity also tended to have prolonged MMF latencies (Matsuzaki et al., 2017). Studies that tested children with ASD using pure tones in an oddball paradigm also found longer MMN/MMF latencies (Seri et al., 1999; Jansson-Verkasalo et al., 2005; Oram Cardy et al., 2005b; Matsuzaki et al., 2017). Seri et al. (1999) reported longer MMN latencies in children diagnosed with ASD and tuberous sclerosis. Similarly, research studying Asperger syndrome, a condition that has subsequently been categorized as ASD by the DSM-5, also found delayed MMN responses (Jansson-Verkasalo et al., 2005). These instances then provide additional support for the idea that variability in processing simple sounds may be tied to specific impairments in sub-diagnoses and/or auditory sensitivity (**Table 3**). Moreover, changes in MMN latency track task difficulty. Therefore, increased peak latencies could also be an artifact of participants with ASD finding auditory tasks more difficult than did typically developing participants (Garrido et al., 2009).

SPEECH SOUNDS

Vowels and Phonemes

People with ASD demonstrate abnormal responses to phonemes and other speech-like sounds, which may stem from an altered perception of less complex stimuli or from impairments in auditory attention that are specific to linguistic components. Phonemes are the units of sound that comprise a language. Therefore, while phonemes have simple and complex tonal components, they have additional meaning in that they are used in an inherently social context. Aberrations in how "simple" stimuli are processed may hinder people with ASD in processing phonemes to some extent; however, there is also evidence that processing difficulties may be unique to linguistic elements.

MMN and MMF

The MMN response to phonemes is somewhat varied, though some trends have emerged. Some studies reported that MMN/MMF peaks from participants with ASD were not different from those seen in typically developing children when presented with vowel sounds (Kemner et al., 1995; Ceponiene et al., 2003b; O'Brien et al., 2020). However, Kuhl et al. (2005) studied low-functioning pre-school aged children using a speech versus computer synthesized non-speech oddball task and found that although the MMN generated in response to non-speech sounds was intact, the MMN peak was lost in children with ASD in response to syllable changes (**Table 4**)

TABLE 3 | MMN peak amplitude and latency in response to simple sound stimuli.

MMN simple sounds	Research	Participants
Greater Amplitude	Gomot et al., 2011	5–11 year old males and females
	Ferri et al., 2003	6–19 year old males with ASD and intellectual impairment
No difference in Amplitude	Gomot et al., 2002	5–9 year old males and females with ASD
	Ceponiene et al., 2003b	6–12 year old males with high functioning ASD
Abnormal waveform	Foxe et al., 2016	4–21 year old girls with Rett syndrome
Reduced Amplitude	Tecchio et al., 2003	8–32 year olds; ASD and intellectual impairment
	Vlaskamp et al., 2017	8–12 year old males and females with ASD and/or Asperger syndrome
	Van der Molen et al., 2012a	18–42 year old males with fragile X syndrome
Longer Latency	Matsuzaki et al., 2017	Mean age 9.5 year old males with ASD and auditory sensitivity
	Jansson-Verkasalo et al., 2005	Mean age 11 years old; males with Asperger syndrome
	Seri et al., 1999	7–10 year old with tuberous sclerosis; sex not reported
	Foxe et al., 2016	4–21 year old girls with Rett syndrome
Shorter latency	Gomot et al., 2002	5–9 year old males and females with ASD

Notable participant features and sub-diagnoses are underlined.

(Kuhl et al., 2005). Interestingly, an fMRI study that examined voice active portions of the brain (superior temporal sulcus) found a similar result—participants with ASD showed a lack of activity when they were presented with vocal sounds, but showed no difference from controls when listening to non-vocal stimuli (Gervais et al., 2004).

Magnetic mismatch field latency was generally delayed in vowel discrimination oddball tasks for individuals with ASD (Oram Cardy et al., 2005b; Roberts et al., 2011; Matsuzaki et al., 2019). When tested with across-phoneme changes (switching between the vowels /a/ and /o/), individuals with ASD showed prolonged MMF latencies over the left hemisphere which were correlated with symptom severity. Varying pure tone or vowel duration, however, did not reveal any ASD-related differences (Kasai et al., 2005). Therefore, while changing the physical aspects of stimuli did not seem to elicit changes in MMN, rapid switching between vowel stimuli did, which may mean that the difficulty that people with ASD experience in processing language may be partially rooted in a poor ability to follow temporal cues in a linguistic context (Kasai et al., 2005).

Lepisto et al. (2005, 2006) designed a task that introduced pitch, duration, and vowel changes to speech and non-speech stimuli and tested children with ASD and Asperger syndrome. Results showed MMN enhancement in response to speech and non-speech pitch changes in both groups (Lepisto et al., 2005, 2006). To elaborate on these results, Lepisto et al. (2008) tested for differences in MMN peaks elicited by changes in pitch or phoneme-type in speech stimuli. To that end, they created a paradigm where either pitch stimuli or phoneme stimuli were presented, and either (1) the features of the standard and deviant stimuli were unaltered, or where (2) irrelevant variations were introduced to the standard and deviant stimuli (stimuli were varied with regard to pitch in the phoneme deviant category or with regard to phoneme in the pitch deviant category). Children with ASD had elevated MMN amplitude in response to pitch changes in both conditions, and for phoneme changes in the first condition. However, the MMN enhancement seen in response to phoneme changes in the ASD group was lost in the second, more speech-like condition (Lepisto et al., 2008). The authors suggest that the enhanced simple sound processing typical of children with ASD may make processing linguistic stimuli more difficult because they may be less able to ignore irrelevant cues (Lepisto et al., 2008; Cheng et al., 2017). Or, put another way, in the context of speech, people with ASD may be too overly focused on "low level" acoustic cue variation to track phoneme changes (Lepisto et al., 2008; Cheng et al., 2017).

In summary, work that studied the ASD MMN/MMF in response to speech stimuli varied with intellectual impairment, ASD subtype, and the nature of the auditory task. MMN did not seem to be especially sensitive to changes in the physical features of simple sounds (pitch and duration); however, it was markedly altered in response to vowel changes. This may imply (1) that there are stimulus features specific to language that slow auditory processing and (2) that impaired ability to rapidly detect changes in incoming speech stimuli may be fundamental to ASD language processing deficits. Consistent with this line of thought, Lepisto et al. (2005, 2006, 2008) saw enhanced MMNs in response to changes in speech stimuli suggesting that individuals with ASD devote more processing power to encoding low level features of linguistic stimuli.

P3

The P3a is a response to unexpected or surprising stimuli and is thought to reflect activity in the frontal, temporal, and parietal lobes (Polich, 2007). As such, it is often tested using an oddball paradigm. The P3a in particular is associated with the detection of novelty and the ability to orient to a stimulus (Yamaguchi and Knight, 1991; Verleger et al., 1994). This is significant as children with ASD routinely show impaired sound orientation for both social and non-social stimuli (Dawson et al., 1989).

In two different studies, reduced P300 amplitude was seen in response to phonemes but not to tonal stimulus. In the first, Dawson et al. (1988) presented a phonetic "da" and a piano chord as stimuli to children with ASD (**Table 5**). While there was no apparent change in how the piano chord was processed

TABLE 4 | MMN in response to vowel and phoneme stimuli.

MMN vowels and phonemes	Research	Test	Participants
Longer latency	Kasai et al., 2005	Phoneme changes	Mean age 27 year old males and females with ASD
	Matsuzaki et al., 2019	Oddball using vowels	Mean age 22 year old males with ASD
	Galilee et al., 2017	Speech vs non-speech	4–6 year old males and females with high functioning ASD
	Roberts et al., 2011	Vowel vs tone oddball	6–15 year old males and females with ASD (Asperger syndrome included) and language impairment
No latency or amplitude differences	Ceponiene et al., 2003b	Vowel sounds	6–12 year old males with high functioning ASD
	Kemner et al., 1995	Vowel sound oddball	7–13 year old males and female; ASD with intellectual impairment
	O'Brien et al., 2020	Vowel and tone oddball	5–15 year old males and females with tuberous sclerosis
Greater amplitude	Lepisto et al., 2008	Pitch or phoneme-type changes in speech stimuli	7–11 year old boys with ASD
	Lepisto et al., 2005	Pitch, duration, and vowel changes in speech and non-speech stimuli oddball	7–11 year old males with ASD; lower verbal IQ in ASD group
	Lepisto et al., 2006	Pitch, duration, and vowel changes in speech and non-speech stimuli oddball	8–11 year old boys with Asperger syndrome
Reduced amplitude	Kuhl et al., 2005	Speech vs computer synthesized non-speech oddball task	3–4 year old males and females with ASD; low functioning

Specific sub-diagnoses and notable features of participants are underlined.

between groups, the children with ASD showed a significant reduction in P3 size in the left hemisphere in response to the "da" phoneme (Dawson et al., 1988). This finding was consistent with later work showing P3a reductions in ASD participants on oddball tasks that used speech stimuli (Kemner et al., 1995; Lepisto et al., 2005, 2006), and work showing reduced amplitude in fragile X syndrome, a condition on the autism spectrum (St Clair et al., 1987; Van der Molen et al., 2012a,b). In the second study, Ceponiene et al. (2003b) had high functioning children with ASD undergo an oddball task using simple tones, complex tones, and vowel sounds as stimuli. They found that there were no differences in the P3a response elicited by simple and complex stimuli between the ASD and typically developing groups, but when presented with vowels, the P3a disappeared in the ASD group (Ceponiene et al., 2003b). Comparable results were also found in individuals with Rett syndrome, a condition on the autism spectrum (Stauder et al., 2006). Taken together, these results indicate that vowel and phoneme processing is uniquely impacted in ASD and specific genetic conditions on the autism spectrum.

Whitehouse and Bishop (2008) expanded on these findings by testing the role of attention in speech sound processing. They saw a general reduction in peak amplitudes when children with ASD passively listened to speech sounds, but peak amplitudes were restored when they were required to actively attend to the speech stimuli (Whitehouse and Bishop, 2008). Interestingly, they also found that children with ASD were less likely to orient to novel tones that were embedded in a stream of speech sounds, but orienting was intact when speech sounds were embedded in tonal stimuli. These results show that, first, responses to speech can be modulated by attention in ASD, and second that participants with ASD are able to attend to speech stimuli depending on the context in which speech sounds are presented. The second finding runs somewhat contrary to work showing reduced orienting to speech in children with ASD (i.e., Dawson et al., 1998; Ceponiene et al., 2003b; Kuhl et al., 2005) in that orienting to speech was possible in ASD under specific conditions.

SPEECH IN NOISE

Research that studied how people with ASD process phonemes focused on how specific components of speech may impede processing in ASD. Those inquiries found that people with ASD had difficulty following phoneme changes, may become distracted by contextually irrelevant features of language, and may show attentional deficits with regard to linguistic stimuli. In addition to the challenges that seem to be intrinsic to speech processing in ASD, individuals with ASD experience particular difficulty processing sound when it is presented amid environmental noise. Background noise can impact simple sound detection, but can also impair linguistic processing to a greater or lesser degree depending on the features of the background noise stimuli. The presence of background stimuli could also contribute to the heightened cortical "noise" that is thought to interfere with auditory function in ASD (Rubenstein and Merzenich, 2003; Sohal and Rubenstein, 2019).

Simple Sounds and Phonemes in Noise in ASD

Even relatively simple background noise can increase the processing load in ASD to the point at which target identification is impaired. To investigate how the presence of background noise affected simple sound processing in ASD, Mamashli et al. (2017)

TABLE 5 | P3 latency and amplitude in response to vowel and phoneme stimuli research.

P3 vowels and phonemes	Research	Test	Participants
Missing	Ceponiene et al., 2003b	Oddball with simples tones, complex tones, and vowels	High functioning ASD
	Stauder et al., 2006	Audiovisual oddball	2–60 year old females with Rett syndrome
Reduced amplitude	Dawson et al., 1988*	Phoneme vs piano cord	6–18 year old males with ASD; some with intellectual impairment
	Kemner et al., 1995	Vowel sound oddball	7–13 year old males and female with ASD and intellectual impairment
	Lepisto et al., 2005	Pitch, duration, and vowel changes in speech and non-speech stimuli oddball	7–11 year old males with ASD; verbal IQ lower in ASD group
	Whitehouse and Bishop, 2008	Vowel and complex sound oddball	8–14 year old males with ASD; verbal IQ lower in ASD group
	Lepisto et al., 2006	Pitch, duration, and vowel changes in speech and non-speech stimuli oddball	8–11 year old boys with Asperger syndrome
	Van der Molen et al., 2012a,b	Auditory tonal stimuli	Mean age 29 years old; males with fragile X syndrome
	St Clair et al., 1987	Oddball task	Adults with fragile X syndrome

Notable features of participants and sub-diagnoses are underlined. Asterisk indicates results only applied to phoneme stimuli.

probed the neural generators of MMF using pure tones against either a quiet or a multi-speaker babble background. In the quiet condition, no difference was found in MMF between groups; however, activation of the inferior frontal gyrus, a generator of MMF, was reduced in the noise condition. Because the inferior frontal gyrus acts to evaluate syntax in incoming language (Tanaka et al., 2017), reduced inferior frontal gyrus activation is consistent with a processing impairment that is specific to language. Moreover, in examining how MMF generators coordinate activity, measures of functional connectivity revealed increased recruitment of neural resources in ASD for both the quiet and the noisy conditions (Mamashli et al., 2017). This suggests that the impact of background noise on speech perception may be a by-product of a general over-responsiveness to sound in ASD.

In this vein, Russo et al. (2009) tested cortical responses to phonemes in children with ASD in noisy conditions. Children either listened to the syllable /da/ with a silent background or while embedded in a white noise background. As might be expected, typically developing children tended to show delayed and reduced responses to stimuli presented in the background noise condition. The ASD group, however, showed only a very slight difference between their responses to the quiet and noisy background conditions. This implies that children with ASD experience a more involved cortical response when processing speech stimuli regardless of the background (Russo et al., 2009). As such, the additional demands of performing more complex speech discrimination tasks (such as identifying words and sentences) against a noisy background may exacerbate problems in what are already tenuous language processing skills.

Speech in Noise Detection in ASD

While background noise impairs speech detection in ASD generally, the features of the background noise in which those speech targets are presented can also impact how well participants are able to extract speech. For instance, people with ASD were found to perform poorly in speech discrimination tasks when the background noise contained temporal dips. Most often, tests aimed at determining how well individuals with ASD are able to extract speech sounds from background noise consist of having participants identify either a word or portions of a sentence while simultaneously presenting a background noise that competes with the target in some way. To test how spectral and temporal variations in background noise effected individuals with ASD's ability to detect speech, Groen et al. (2009) designed an experiment where participants were instructed to repeat back target words presented against a "pink noise" background (noise with spectral energy divided evenly across the frequency bands of the human auditory system). To specifically test the effect of temporal variation, Groen et al. (2009) created pink noise backgrounds that either had or did not have temporal dips. Performance in both the typically developing and ASD group suffered when temporal dips were introduced, but the degree of disruption was greater in the ASD group (Groen et al., 2009). Similarly, when asked to extract speech (whole sentences) from a sampling of different background noise conditions, the performance of participants with ASD was significantly worse in conditions where temporal or spectro-temporal dips were introduced into a speech-like background (Alcantara et al., 2004). It is hypothesized that during temporal dips in background noise, typically developing listeners can piece together the target speech using contextual cues, but listeners with ASD are less able to gather or use those cues (Alcantara et al., 2004; Qian and Lipkin, 2011). In addition, work that focused on speech detection against an "attention demanding" multi-talker background stimulus found that individuals with ASD performed significantly worse than their typically developing counterparts (Dunlop et al., 2016). Taken together, these results indicate that individuals with ASD may be less able to integrate information gathered during breaks in background stimuli, and, as difficulty extracting target

speech against multi-talker background suggests, individuals with ASD may also require a greater signal-to-noise ratio to discriminate spoken words.

Schelinski and von Kriegstein (2020) tested whether signal-to-noise ratio differences in ASD were a driving factor for speech in background noise discernment errors. In their experiment, typically developing adults and adults with ASD listened to speech presented against a continuous speech-shaped background noise. Results showed that typically developing participants could detect target stimuli at a significantly lower sound-in-noise ratio than the participants with ASD. Put another way, the difference in intensity between the signal and background noise had to be greater for the group with ASD to detect the target signal. These results are somewhat contrary to those of Alcantara et al. (2004) and Groen et al. (2009) in that they show target detection impairment in ASD with a continuous noise background rather than only in temporally shaped background noise. However, in the interest of allowing for greater expression of ASD symptom variability, Schelinski and von Kriegstein (2020) did use less challenging speech recognition thresholds in their research than those used in previous studies. This may mean that temporal processing deficits in ASD could be ameliorated by conditions with a more favorable signal-to-noise ratio (Schelinski and von Kriegstein, 2020).

As another possibility, individuals with ASD may find speech in noise conditions challenging as a result of altered voice perception. Schelinski and von Kriegstein (2020) also found that speech-in-noise recognition correlated with vocal pitch perception ability in typically developing adults, but not in adults with ASD. This effect seems to be limited to vocal pitch discrimination however, as testing vocal timbre discrimination did not reveal differences between typically developing and ASD groups (Schelinski et al., 2017). In keeping with the idea that poor voice recognition contributes to poor sound in noise performance, fMRI data also demonstrated that participants with ASD had an impaired ability to recognize a target speaker's voice as shown by reduced activity in the superior temporal sulcus and superior temporal gyrus (temporal voice areas) (Schelinski et al., 2014, 2016).

While individuals with ASD do seem to experience aberrations specific to processing speech and voices, given that speech is an inherently social activity, the desire to engage in social exchange must also be considered when testing speech-in-noise processing. When participants with ASD were presented with a conversation between two people that had competing "ecologically valid" background noise (noise typical of everyday social situations), typically developing and participants with ASD had comparable patterns of brain activity (Hernandez et al., 2020). However, the angular gyrus was relatively more active in participants with ASD and angular gyrus activity was correlated with social motivation (as measured by the Social Responsiveness Scale—2nd Edition, SRS-2). Therefore, Hernandez et al. (2020) speculate that participants with higher social motivation would be more likely to direct attention to the conversational sound in noise targets, and therefore may perform better than their less socially motivated peers in recalling conversational targets (Hernandez et al., 2020).

DISCUSSION

Developmental Delays and Simple Stimuli Processing in ASD

Delayed or atypical development of sensory systems is a common feature of ASD and may underlie the peak latency and amplitude changes observed in EEG/MEG testing. In typically developing individuals, waveforms become more complex as children age. Importantly, auditory development in early childhood sees reductions in P1 latency and amplitude as N1 becomes more prominent (Oades et al., 1997; Sharma et al., 1997; Ponton et al., 2000; Wunderlich et al., 2006). Early developmental changes also track reductions in N1 latency (Bruneau et al., 1997; Oades et al., 1997; Sharma et al., 1997; Ponton et al., 2000; McArthur and Bishop, 2002). Several of the studies reviewed here reported significant ASD-related delays in N1 latency results which could reflect under-development of the auditory system. Consistent with this finding, work that compared peak latency with participant's age failed to find any age-related change in M100 peak latency over the right hemisphere (Gage et al., 2003a), reinforcing the idea that the errors found at N1 may be developmental in nature.

The majority of studies that characterized N1 also found increases in peak latency and amplitude that were almost exclusively seen over the right hemisphere, suggesting poor response lateralization in ASD. In typically developing people, auditory areas of the brain (including the STG) undergo synaptic pruning events that cause decreases in size relative to brain volume between childhood and young adulthood. During this developmental period, a left-right hemisphere asymmetry in auditory area volume is established. Symmetry in auditory area size in ASD seems to be rooted in a failure to loose volume in the right (or non-dominant) hemisphere of the brain (Devous et al., 2006; Jou et al., 2010). As such, the larger responses seen over the right hemisphere in ASD may simply reflect larger right hemisphere EEG/MEG generators that are a consequence of impaired synaptic pruning during development.

With regard to P1, peak amplitude was reduced in ASD in all studies reviewed, including studies with participants as young as 4 years old. This shows that even from a young age, sensory processing is abnormal in children with ASD (Orekhova et al., 2008; Donkers et al., 2015). P1 also seemed insensitive to changes in the timing of stimuli and lacked a habituation response in ASD (Buchwald et al., 1992; Ruiz-Martinez et al., 2020); as such, changes in stimuli was not well represented by P1. Functionally, the inability to track changes in a stimulus may manifest as poor sensory gating in ASD. Similarly, a lack of habituation to stimuli could conceivably contribute to the heightened auditory sensitivity typical of ASD. Impaired representation of stimulus timing may ultimately play a role in the linguistic deficiencies found in ASD, as individual with ASD may have difficulty following spectro-temporal changes in words.

Arousal and Attention in ASD

At rest, individuals with ASD show differences in their states of arousal (as measured by skin conductance, body temperature,

and heart rate) relative to control groups (Schoen et al., 2009; Mathersul et al., 2013; Eilam-Stock et al., 2014; Prince et al., 2017). Participants' state of arousal is relevant to auditory testing because arousal impacts the size of the EEG/MEG peaks observed. Kozlowska et al. (2017) found that an individual's baseline state of arousal seems to act as a precondition to response magnitude. When tested with an auditory oddball task, children with neurologic disturbances that caused heightened arousal showed greater amplitudes in ERP components (Kozlowska et al., 2017). Additionally, children with higher arousal (as measured by heart rate) tended to show larger P3a peaks on an oddball task (Wass et al., 2019). Taken together, it is possible that changes in ASD participants' state of arousal during testing may drive some of the variability found in ASD auditory EEG/MEG data; therefore, future studies may consider including a control for arousal when testing auditory processing in ASD.

Differences in how people with ASD direct their attention when undergoing auditory EEG/MEG testing could also account for some of the observed inconsistencies in the literature. Individuals with ASD consistently show anomalies in joint attention and orienting to speech, and frequently show co-morbidity with attention deficit/hyperactivity disorder (reviewed in Mundy, 2018; Sharma et al., 2018). Auditory attention sharpens frequency tuning and can act to enhance gain for target stimuli (de Boer and Krumbholz, 2018). It would follow then that ASD-related anomalies in how attention is directed could drive atypical results in simple sound testing, as seen in Oades et al. (1988). Measures of attention have also been used to predict speech in noise ability (Moore et al., 2010), which is consistent with the attentional effects described in studies that asked participants with ASD to extract speech sounds from background noise (Whitehouse and Bishop, 2008; Dunlop et al., 2016; Hernandez et al., 2020).

ASD Sub-Diagnosis in Auditory Processing

Idiopathic ASD is the most common form of ASD; however, that diagnosis may describe a number of genetic conditions that potentially have slightly different presentations. Comparing non-syndromic ASD with known genetic conditions on the autism spectrum is useful in interpreting some of the variability found in ASD EEG/MEG data. For instance, Rett syndrome is a condition on the autism spectrum that presents with cognitive impairments. Consistent with ASD results, individuals with Rett syndrome show general increases in early sensory peak latencies (Stauder et al., 2006; Foxe et al., 2016), but divergent MMN results. Also, while the majority of ASD studies reported reduced P3 amplitudes, P3 was missing entirely in Rett syndrome (Stauder et al., 2006). Tuberous sclerosis, another autism spectrum condition, is characterized by unchecked protein synthesis and tumor growth that cause abnormal neural connections, longer peak latencies, and asymmetrical N1 response lateralization (Seri et al., 1999; Feliciano et al., 2013; O'Brien et al., 2020). Excess protein production is also found in fragile X syndrome, which renders EEG/MEG results that are mostly in line with typical ASD responses—N1 amplitude is greater, latency is longer, and P3 waveforms are abnormal (St Clair et al., 1987; Rojas et al., 2001; Van der Molen et al., 2012a,b). Unlike Rett syndrome patients however, participants with fragile X show ASD-like N1 lateralization (Rojas et al., 2001). Taken as a whole, fragile X syndrome and Rett syndrome data show that even conditions with similar etiologies can produce subtly different EEG/MEG results.

Temporal Integration in ASD

The capacity to distinguish speech sounds and parse speech depends on the ability to follow rapid temporal cues on the order of milliseconds. As such, it is conceivable that even small changes in temporal processing may ultimately have a significant impact on an individuals' ability to perceive and understand speech (Tallal et al., 1993). Gap detection tests measure auditory temporal processing ability by presenting listeners with a series of sounds and varying the interval of time between presentations. The goal of these tests is to determine the interval duration at which listeners are able to perceive the sounds as discrete stimuli rather than as a single continuous sound. Gap detection testing has consistently shown that children with ASD tend to need longer gaps in order to identify stimuli, which suggests that children with ASD experience impaired temporal resolution in processing sound (Kwakye et al., 2011; Bhatara et al., 2013; Foss-Feig et al., 2017). Recent literature also found that not only did children with ASD require longer gaps to parse sounds, but that gap detection ability was correlated with lessened phonological awareness and impaired speech-in-noise detection (Foss-Feig et al., 2017). Similarly, MEG work showed children with ASD failed to respond to the second stimulus when duos of pure tones were presented in rapid succession, supporting the idea that rapid temporal processing is impaired in ASD (Oram Cardy et al., 2005a). In point of fact, all EEG/MEG peaks but P1 are sensitive to changes in stimulus presentation rate (Dinces and Sussman, 2011). This suggests that at least some of the difficulty that children with ASD have in attending to syllables/words/sentences arise from temporal processing impairments (Bhatara et al., 2013; Foss-Feig et al., 2017).

FURTHER CONSIDERATIONS

In review, some consistent themes emerged with regard to ASD EEG/MEG data, but still many EEG/MEG components reported a range of responses. This is at least somewhat expected given the variety of sub-diagnoses, intellectual ability, and developmental delays represented under to ASD umbrella. In order to better understand the driving factors in ASD-related sensory disability, future studies may incorporate the following five considerations. First, studies that required participants to direct their attention toward or away from a stimulus found consistent changes in EEG/MEG responses (de Boer and Krumbholz, 2018). Given that people with ASD routinely exhibit difficulties around attentional focus, incorporating an assessment of attentional ability may be worthwhile. Second, because state of arousal does have an impact on EEG/MEG response magnitude, monitoring participants' state of arousal during testing may also aid in understanding

abnormal responses. Third, some studies made efforts to control for developmental delays by using not only age matched, but developmentally aged matched controls for participants with ASD. Such endeavors can be beneficial in interpreting aberrant data, but enforcing age range restrictions of participants included in a study may decrease ambiguity in the results. Fourth, the ASD umbrella represents a wide range of both known and yet-to-be-identified genetic conditions. Given that known genetic conditions give varying EEG/MEG results, it is not unreasonable to suppose that the conditions represented in idiopathic ASD may also provide idiosyncratic results. As such, including any efforts that have been made to identify the underlying cause of participants' ASD in reports would be beneficial. Lastly, the majority of the literature studying auditory processing in individuals with ASD used exclusively male participants. While ASD does seem to be more common in males, ASD is also thought to be underdiagnosed in females. In order to better understand and diagnose ASD in girls, a complete picture of the sensory issues they face is essential.

AUTHOR CONTRIBUTIONS

The author researched and wrote the entire manuscript.

FUNDING

This work was supported by a Seed Grant from Mercer University School of Medicine.

REFERENCES

Alcantara, J. I., Weisblatt, E. J., Moore, B. C., and Bolton, P. F. (2004). Speech-in-noise perception in high-functioning individuals with autism or Asperger's syndrome. *J. Child Psychol. Psychiatry* 45, 1107–1114.

Alho, K. (1995). Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes. *Ear Hear.* 16, 38–51. doi: 10.1097/00003446-199502000-00004

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders*, 5th Edn. Washington, DC: American Psychiatric Association.

Bhatara, A., Babikian, T., Laugeson, E., Tachdjian, R., and Sininger, Y. S. (2013). Impaired timing and frequency discrimination in high-functioning autism spectrum disorders. *J. Autism Dev. Disord.* 43, 2312–2328. doi: 10.1007/s10803-013-1778-y

Bidet-Caulet, A., Latinus, M., Roux, S., Malvy, J., Bonnet-Brilhault, F., and Bruneau, N. (2017). Atypical sound discrimination in children with ASD as indicated by cortical ERPs. *J. Neurodev. Disord.* 9:13.

Bonnel, A., McAdams, S., Smith, B., Berthiaume, C., Bertone, A., Ciocca, V., et al. (2010). Enhanced pure-tone pitch discrimination among persons with autism but not Asperger syndrome. *Neuropsychologia* 48, 2465–2475. doi: 10.1016/j.neuropsychologia.2010.04.020

Bruneau, N., Bonnet-Brilhault, F., Gomot, M., Adrien, J. L., and Barthelemy, C. (2003). Cortical auditory processing and communication in children with autism: electrophysiological/behavioral relations. *Int. J. Psychophysiol.* 51, 17–25. doi: 10.1016/s0167-8760(03)00149-1

Bruneau, N., Roux, S., Adrien, J. L., and Barthelemy, C. (1999). Auditory associative cortex dysfunction in children with autism: evidence from late auditory evoked potentials (N1 wave-T complex). *Clin. Neurophysiol.* 110, 1927–1934. doi: 10.1016/s1388-2457(99)00149-2

Bruneau, N., Roux, S., Guerin, P., Barthelemy, C., and Lelord, G. (1997). Temporal prominence of auditory evoked potentials (N1 wave) in 4-8-year-old children. *Psychophysiology* 34, 32–38. doi: 10.1111/j.1469-8986.1997.tb02413.x

Buchwald, J. S., Erwin, R., Van Lancker, D., Guthrie, D., Schwafel, J., and Tanguay, P. (1992). Midlatency auditory evoked responses: P1 abnormalities in adult autistic subjects. *Electroencephalogr. Clin. Neurophysiol.* 84, 164–171. doi: 10.1016/0168-5597(92)90021-3

Castren, M., Paakkonen, A., Tarkka, I. M., Ryynanen, M., and Partanen, J. (2003). Augmentation of auditory N1 in children with fragile X syndrome. *Brain Topogr.* 15, 165–171.

Ceponiene, R., Lepisto, T., Alku, P., Aro, H., and Naatanen, R. (2003a). Event-related potential indices of auditory vowel processing in 3-year-old children. *Clin. Neurophysiol.* 114, 652–661. doi: 10.1016/s1388-2457(02)00436-4

Ceponiene, R., Lepisto, T., Shestakova, A., Vanhala, R., Alku, P., Naatanen, R., et al. (2003b). Speech-sound-selective auditory impairment in children with autism: they can perceive but do not attend. *Proc. Natl. Acad. Sci. U.S.A.* 100, 5567–5572. doi: 10.1073/pnas.0835631100

Ceponiene, R., Shestakova, A., Balan, P., Alku, P., Yiaguchi, K., and Naatanen, R. (2001). Children's auditory event-related potentials index sound complexity and "speechness". *Int. J. Neurosci.* 109, 245–260. doi: 10.3109/00207450108986536

Cheng, S. T. T., Lam, G. Y. H., and To, C. K. S. (2017). Pitch perception in tone language-speaking adults with and without autism spectrum disorders. *Iperception* 8:2041669517711200.

Dawson, G., Finley, C., Phillips, S., and Galpert, L. (1986). Hemispheric specialization and the language abilities of autistic children. *Child Dev.* 57, 1440–1453. doi: 10.2307/1130422

Dawson, G., Finley, C., Phillips, S., Galpert, L., and Lewy, A. (1988). Reduced P3 amplitude of the event-related brain potential: its relationship to language ability in autism. *J. Autism Dev. Disord.* 18, 493–504. doi: 10.1007/bf02211869

Dawson, G., Finley, C., Phillips, S., and Lewy, A. (1989). A comparison of hemispheric asymmetries in speech-related brain potentials of autistic and dysphasic children. *Brain Lang.* 37, 26–41. doi: 10.1016/0093-934x(89)90099-0

Dawson, G., Meltzoff, A. N., Osterling, J., Rinaldi, J., and Brown, E. (1998). Children with autism fail to orient to naturally occurring social stimuli. *J. Autism Dev. Disord.* 28, 479–485.

de Boer, J., and Krumbholz, K. (2018). Auditory attention causes gain enhancement and frequency sharpening at successive stages of cortical processing-evidence from human electroencephalography. *J. Cogn. Neurosci.* 30, 785–798. doi: 10.1162/jocn_a_01245

DePape, A. M., Hall, G. B., Tillmann, B., and Trainor, L. J. (2012). Auditory processing in high-functioning adolescents with autism spectrum disorder. *PLoS One* 7:e44084. doi: 10.1371/journal.pone.0044084

Devous, M. D. Sr., Altuna, D., Furl, N., Cooper, W., Gabbert, G., Ngai, W. T., et al. (2006). Maturation of speech and language functional neuroanatomy in pediatric normal controls. *J. Speech Lang. Hear. Res.* 49, 856–866. doi: 10.1044/1092-4388(2006/061)

Dinces, E., and Sussman, E. (2011). Effects of acoustic complexity on processing sound intensity in 10- to 11-year-old children: evidence from cortical auditory evoked potentials. *Laryngoscope* 121, 1785–1793. doi: 10.1002/lary.21820

Donkers, F. C., Schipul, S. E., Baranek, G. T., Cleary, K. M., Willoughby, M. T., Evans, A. M., et al. (2015). Attenuated auditory event-related potentials and associations with atypical sensory response patterns in children with autism. *J. Autism Dev. Disord.* 45, 506–523. doi: 10.1007/s10803-013-1948-y

Dunlop, W. A., Enticott, P. G., and Rajan, R. (2016). Speech discrimination difficulties in high-functioning autism spectrum disorder are likely independent of auditory hypersensitivity. *Front. Hum. Neurosci.* 10:401. doi: 10.3389/fnhum.2016.00401

Eggermont, J. J., Ponton, C. W., Don, M., Waring, M. D., and Kwong, B. (1997). Maturational delays in cortical evoked potentials in cochlear implant users. *Acta Otolaryngol.* 117, 161–163. doi: 10.3109/00016489709117760

Eilam-Stock, T., Xu, P., Cao, M., Gu, X., Van Dam, N. T., Anagnostou, E., et al. (2014). Abnormal autonomic and associated brain activities during rest in autism spectrum disorder. *Brain* 137, 153–171. doi: 10.1093/brain/awt294

Eyler, L. T., Pierce, K., and Courchesne, E. (2012). A failure of left temporal cortex to specialize for language is an early emerging and fundamental property of autism. *Brain* 135, 949–960. doi: 10.1093/brain/awr364

Feliciano, D. M., Lin, T. V., Hartman, N. W., Bartley, C. M., Kubera, C., Hsieh, L., et al. (2013). A circuitry and biochemical basis for tuberous sclerosis symptoms: from epilepsy to neurocognitive deficits. *Int. J. Dev. Neurosci.* 31, 667–678. doi: 10.1016/j.ijdevneu.2013.02.008

Ferri, R., Elia, M., Agarwal, N., Lanuzza, B., Musumeci, S. A., and Pennisi, G. (2003). The mismatch negativity and the P3a components of the auditory event-related potentials in autistic low-functioning subjects. *Clin. Neurophysiol.* 114, 1671–1680. doi: 10.1016/s1388-2457(03)00153-6

Flagg, E. J., Cardy, J. E., Roberts, W., and Roberts, T. P. (2005). Language lateralization development in children with autism: insights from the late field magnetoencephalogram. *Neurosci. Lett.* 386, 82–87. doi: 10.1016/j.neulet.2005.05.037

Foss-Feig, J. H., Schauder, K. B., Key, A. P., Wallace, M. T., and Stone, W. L. (2017). Audition-specific temporal processing deficits associated with language function in children with autism spectrum disorder. *Autism Res.* 10, 1845–1856. doi: 10.1002/aur.1820

Foxe, J. J., Burke, K. M., Andrade, G. N., Djukic, A., Frey, H. P., and Molholm, S. (2016). Automatic cortical representation of auditory pitch changes in Rett syndrome. *J. Neurodev. Disord.* 8:34.

Gage, N. M., Siegel, B., Callen, M., and Roberts, T. P. (2003a). Cortical sound processing in children with autism disorder: an MEG investigation. *Neuroreport* 14, 2047–2051. doi: 10.1097/00001756-200311140-00008

Gage, N. M., Siegel, B., and Roberts, T. P. (2003b). Cortical auditory system maturational abnormalities in children with autism disorder: an MEG investigation. *Brain Res. Dev. Brain Res.* 144, 201–209. doi: 10.1016/s0165-3806(03)00172-x

Galilee, A., Stefanidou, C., and McCleery, J. P. (2017). Atypical speech versus non-speech detection and discrimination in 4- to 6- yr old children with autism spectrum disorder: An ERP study. *PLoS One* 12:e0181354. doi: 10.1371/journal.pone.0181354

Garrido, M. I., Kilner, J. M., Stephan, K. E., and Friston, K. J. (2009). The mismatch negativity: a review of underlying mechanisms. *Clin. Neurophysiol.* 120, 453–463. doi: 10.1016/j.clinph.2008.11.029

Gervais, H., Belin, P., Boddaert, N., Leboyer, M., Coez, A., Sfaello, I., et al. (2004). Abnormal cortical voice processing in autism. *Nat. Neurosci.* 7, 801–802. doi: 10.1038/nn1291

Gomes, E., Pedroso, F. S., and Wagner, M. B. (2008). Auditory hypersensitivity in the autistic spectrum disorder. *Pro Fono* 20, 279–284.

Gomot, M., Blanc, R., Clery, H., Roux, S., Barthelemy, C., and Bruneau, N. (2011). Candidate electrophysiological endophenotypes of hyper-reactivity to change in autism. *J. Autism Dev. Disord.* 41, 705–714. doi: 10.1007/s10803-010-1091-y

Gomot, M., Giard, M. H., Adrien, J. L., Barthelemy, C., and Bruneau, N. (2002). Hypersensitivity to acoustic change in children with autism: electrophysiological evidence of left frontal cortex dysfunctioning. *Psychophysiology* 39, 577–584. doi: 10.1111/1469-8986.3950577

Groen, W. B., van Orsouw, L., Huurne, N., Swinkels, S., van der Gaag, R. J., Buitelaar, J. K., et al. (2009). Intact spectral but abnormal temporal processing of auditory stimuli in autism. *J. Autism Dev. Disord.* 39, 742–750. doi: 10.1007/s10803-008-0682-3

Heaton, P., Davis, R. E., and Happe, F. G. (2008). Research note: exceptional absolute pitch perception for spoken words in an able adult with autism. *Neuropsychologia* 46, 2095–2098. doi: 10.1016/j.neuropsychologia.2008.02.006

Hernandez, L. M., Green, S. A., Lawrence, K. E., Inada, M., Liu, J., Bookheimer, S. Y., et al. (2020). Social attention in autism: neural sensitivity to speech over background noise predicts encoding of social information. *Front. Psychiatry* 11:343. doi: 10.3389/fpsyt.2020.00343

Hudac, C. M., DesChamps, T. D., Arnett, A. B., Cairney, B. E., Ma, R., Webb, S. J., et al. (2018). Early enhanced processing and delayed habituation to deviance sounds in autism spectrum disorder. *Brain Cogn.* 123, 110–119. doi: 10.1016/j.bandc.2018.03.004

Hudry, K., Leadbitter, K., Temple, K., Slonims, V., McConachie, H., Aldred, C., et al. (2010). Preschoolers with autism show greater impairment in receptive compared with expressive language abilities. *Int. J. Lang. Commun. Disord.* 45, 681–690. doi: 10.3109/13682820903461493

Jansson-Verkasalo, E., Kujala, T., Jussila, K., Mattila, M. L., Moilanen, I., Naatanen, R., et al. (2005). Similarities in the phenotype of the auditory neural substrate in children with Asperger syndrome and their parents. *Eur. J. Neurosci.* 22, 986–990. doi: 10.1111/j.1460-9568.2005.04216.x

Jones, C. R., Happe, F., Baird, G., Simonoff, E., Marsden, A. J., Tregay, J., et al. (2009). Auditory discrimination and auditory sensory behaviours in autism spectrum disorders. *Neuropsychologia* 47, 2850–2858. doi: 10.1016/j.neuropsychologia.2009.06.015

Jou, R. J., Minshew, N. J., Keshavan, M. S., Vitale, M. P., and Hardan, A. Y. (2010). Enlarged right superior temporal gyrus in children and adolescents with autism. *Brain Res.* 1360, 205–212. doi: 10.1016/j.brainres.2010.09.005

Kasai, K., Hashimoto, O., Kawakubo, Y., Yumoto, M., Kamio, S., Itoh, K., et al. (2005). Delayed automatic detection of change in speech sounds in adults with autism: a magnetoencephalographic study. *Clin. Neurophysiol.* 116, 1655–1664. doi: 10.1016/j.clinph.2005.03.007

Kemner, C., Verbaten, M. N., Cuperus, J. M., Camfferman, G., and van Engeland, H. (1995). Auditory event-related brain potentials in autistic children and three different control groups. *Biol. Psychiatry* 38, 150–165. doi: 10.1016/0006-3223(94)00247-z

Khalfa, S., Bruneau, N., Roge, B., Georgieff, N., Veuillet, E., Adrien, J. L., et al. (2004). Increased perception of loudness in autism. *Hear Res.* 198, 87–92. doi: 10.1016/j.heares.2004.07.006

Korpilahti, P., Jansson-Verkasalo, E., Mattila, M. L., Kuusikko, S., Suominen, K., Rytky, S., et al. (2007). Processing of affective speech prosody is impaired in Asperger syndrome. *J. Autism Dev. Disord.* 37, 1539–1549. doi: 10.1007/s10803-006-0271-2

Kozlowska, K., Melkonian, D., Spooner, C. J., Scher, S., and Meares, R. (2017). Cortical arousal in children and adolescents with functional neurological symptoms during the auditory oddball task. *Neuroimage Clin.* 13, 228–236. doi: 10.1016/j.nicl.2016.10.016

Kuhl, P. K., Coffey-Corina, S., Padden, D., and Dawson, G. (2005). Links between social and linguistic processing of speech in preschool children with autism: behavioral and electrophysiological measures. *Dev. Sci.* 8, F1–F12.

Kuiper, M. W. M., Verhoeven, E. W. M., and Geurts, H. M. (2019). Stop making noise! Auditory sensitivity in adults with an autism spectrum disorder diagnosis: physiological habituation and subjective detection thresholds. *J. Autism Dev. Disord.* 49, 2116–2128. doi: 10.1007/s10803-019-03890-9

Kujala, T., Lepisto, T., and Naatanen, R. (2013). The neural basis of aberrant speech and audition in autism spectrum disorders. *Neurosci. Biobehav. Rev.* 37, 697–704. doi: 10.1016/j.neubiorev.2013.01.006

Kwakye, L. D., Foss-Feig, J. H., Cascio, C. J., Stone, W. L., and Wallace, M. T. (2011). Altered auditory and multisensory temporal processing in autism spectrum disorders. *Front. Integr. Neurosci.* 4:129. doi: 10.3389/fnint.2010.00129

Lawson, R. P., Aylward, J., White, S., and Rees, G. (2015). A striking reduction of simple loudness adaptation in autism. *Sci. Rep.* 5:16157.

Lepisto, T., Kajander, M., Vanhala, R., Alku, P., Huotilainen, M., Naatanen, R., et al. (2008). The perception of invariant speech features in children with autism. *Biol. Psychol.* 77, 25–31. doi: 10.1016/j.biopsycho.2007.08.010

Lepisto, T., Kujala, T., Vanhala, R., Alku, P., Huotilainen, M., and Naatanen, R. (2005). The discrimination of and orienting to speech and non-speech sounds in children with autism. *Brain Res.* 1066, 147–157. doi: 10.1016/j.brainres.2005.10.052

Lepisto, T., Silokallio, S., Nieminen-von Wendt, T., Alku, P., Naatanen, R., and Kujala, T. (2006). Auditory perception and attention as reflected by the brain event-related potentials in children with Asperger syndrome. *Clin. Neurophysiol.* 117, 2161–2171. doi: 10.1016/j.clinph.2006.06.709

Mamashli, F., Khan, S., Bharadwaj, H., Michmizos, K., Ganesan, S., Garel, K. A., et al. (2017). Auditory processing in noise is associated with complex patterns of disrupted functional connectivity in autism spectrum disorder. *Autism Res.* 10, 631–647. doi: 10.1002/aur.1714

Mathersul, D., McDonald, S., and Rushby, J. A. (2013). Autonomic arousal explains social cognitive abilities in high-functioning adults with autism spectrum disorder. *Int. J. Psychophysiol.* 89, 475–482. doi: 10.1016/j.ijpsycho.2013.04.014

Matsuzaki, J., Kagitani-Shimono, K., Sugata, H., Hanaie, R., Nagatani, F., Yamamoto, T., et al. (2017). Delayed mismatch field latencies in autism spectrum disorder with abnormal auditory sensitivity:

a magnetoencephalographic study. *Front. Hum. Neurosci.* 11:446. doi: 10.3389/fnhum.2017.00446

Matsuzaki, J., Ku, M., Berman, J. I., Blaskey, L., Bloy, L., Chen, Y. H., et al. (2019). Abnormal auditory mismatch fields in adults with autism spectrum disorder. *Neurosci. Lett.* 698, 140–145. doi: 10.1016/j.neulet.2018.12.043

McArthur, G., and Bishop, D. (2002). Event-related potentials reflect individual differences in age-invariant auditory skills. *Neuroreport* 13, 1079–1082. doi: 10.1097/00001756-200206120-00021

Mitchell, S., Brian, J., Zwaigenbaum, L., Roberts, W., Szatmari, P., Smith, I., et al. (2006). Early language and communication development of infants later diagnosed with autism spectrum disorder. *J. Dev. Behav. Pediatr.* 27, S69–S78.

Moore, D. R., Ferguson, M. A., Edmondson-Jones, A. M., Ratib, S., and Riley, A. (2010). Nature of auditory processing disorder in children. *Pediatrics* 126, e382–e380. doi: 10.1542/peds.2009-2826

Mundy, P. (2018). A review of joint attention and social-cognitive brain systems in typical development and autism spectrum disorder. *Eur. J. Neurosci.* 47, 497–514. doi: 10.1111/ejn.13720

Naatanen, R., and Alho, K. (1995). Mismatch negativity–a unique measure of sensory processing in audition. *Int. J. Neurosci.* 80, 317–337. doi: 10.3109/00207459508986107

Naatanen, R., and Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology* 24, 375–425. doi: 10.1111/j.1469-8986.1987.tb00311.x

Oades, R. D., Dittmann-Balcar, A., and Zerbin, D. (1997). Development and topography of auditory event-related potentials (ERPs): mismatch and processing negativity in individuals 8-22 years of age. *Psychophysiology* 34, 677–693. doi: 10.1111/j.1469-8986.1997.tb02143.x

Oades, R. D., Walker, M. K., Geffen, L. B., and Stern, L. M. (1988). Event-related potentials in autistic and healthy children on an auditory choice reaction time task. *Int. J. Psychophysiol.* 6, 25–37. doi: 10.1016/0167-8760(88)90032-3

O'Brien, A. M., Bayet, L., Riley, K., Nelson, C. A., Sahin, M., and Modi, M. E. (2020). Auditory processing of speech and tones in children with tuberous sclerosis complex. *Front. Integr. Neurosci.* 14:14. doi: 10.3389/fnint.2020.00014

O'Connor, K. (2012). Auditory processing in autism spectrum disorder: a review. *Neurosci. Biobehav. Rev* 36, 836–854. doi: 10.1016/j.neubiorev.2011.11.008

Oram Cardy, J. E., Flagg, E. J., Roberts, W., Brian, J., and Roberts, T. P. (2005a). Magnetoencephalography identifies rapid temporal processing deficit in autism and language impairment. *Neuroreport* 16, 329–332. doi: 10.1097/00001756-200503150-00005

Oram Cardy, J. E., Flagg, E. J., Roberts, W., and Roberts, T. P. (2005b). Delayed mismatch field for speech and non-speech sounds in children with autism. *Neuroreport* 16, 521–525. doi: 10.1097/00001756-200504040-00021

Oram Cardy, J. E., Flagg, E. J., Roberts, W., and Roberts, T. P. (2008). Auditory evoked fields predict language ability and impairment in children. *Int. J. Psychophysiol.* 68, 170–175. doi: 10.1016/j.ijpsycho.2007.10.015

Orekhova, E. V., Stroganova, T. A., Prokofyev, A. O., Nygren, G., Gillberg, C., and Elam, M. (2008). Sensory gating in young children with autism: relation to age, IQ, and EEG gamma oscillations. *Neurosci. Lett.* 434, 218–223. doi: 10.1016/j.neulet.2008.01.066

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118, 2128–2148. doi: 10.1016/j.clinph.2007.04.019

Ponton, C. W., Eggermont, J. J., Kwong, B., and Don, M. (2000). Maturation of human central auditory system activity: evidence from multi-channel evoked potentials. *Clin. Neurophysiol.* 111, 220–236. doi: 10.1016/s1388-2457(99)00236-9

Port, R. G., Edgar, J. C., Ku, M., Bloy, L., Murray, R., Blaskey, L., et al. (2016). Maturation of auditory neural processes in autism spectrum disorder–a longitudinal MEG study. *Neuroimage Clin.* 11, 566–577. doi: 10.1016/j.nicl.2016.03.021

Pratt, K., Baird, G., and Gringras, P. (2012). Ensuring successful admission to hospital for young people with learning difficulties, autism and challenging behaviour: a continuous quality improvement and change management programme. *Child Care Health Dev.* 38, 789–797. doi: 10.1111/j.1365-2214.2011.01335.x

Prince, E. B., Kim, E. S., Wall, C. A., Gisin, E., Goodwin, M. S., Simmons, E. S., et al. (2017). The relationship between autism symptoms and arousal level in toddlers with autism spectrum disorder, as measured by electrodermal activity. *Autism* 21, 504–508. doi: 10.1177/1362361316648816

Qian, N., and Lipkin, R. M. (2011). A learning-style theory for understanding autistic behaviors. *Front. Hum. Neurosci.* 5:77. doi: 10.3389/fnhum.2011.00077

Roberts, T. P., Cannon, K. M., Tavabi, K., Blaskey, L., Khan, S. Y., Monroe, J. F., et al. (2011). Auditory magnetic mismatch field latency: a biomarker for language impairment in autism. *Biol. Psychiatry* 70, 263–269. doi: 10.1016/j.biopsych.2011.01.015

Roberts, T. P., Khan, S. Y., Rey, M., Monroe, J. F., Cannon, K., Blaskey, L., et al. (2010). MEG detection of delayed auditory evoked responses in autism spectrum disorders: towards an imaging biomarker for autism. *Autism Res.* 3, 8–18.

Rojas, D. C., Benkers, T. L., Rogers, S. J., Teale, P. D., Reite, M. L., and Hagerman, R. J. (2001). Auditory evoked magnetic fields in adults with fragile X syndrome. *Neuroreport* 12, 2573–2576. doi: 10.1097/00001756-200108080-00056

Rosenhall, U., Nordin, V., Sandstrom, M., Ahlsen, G., and Gillberg, C. (1999). Autism and hearing loss. *J. Autism Dev. Disord.* 29, 349–357.

Rubenstein, J. L., and Merzenich, M. M. (2003). Model of autism: increased ratio of excitation/inhibition in key neural systems. *Genes Brain Behav.* 2, 255–267. doi: 10.1034/j.1601-183x.2003.00037.x

Ruiz-Martinez, F. J., Rodriguez-Martinez, E. I., Wilson, C. E., Yau, S., Saldana, D., and Gomez, C. M. (2020). Impaired P1 Habituation and mismatch negativity in children with autism spectrum disorder. *J. Autism Dev. Disord.* 50, 603–616. doi: 10.1007/s10803-019-04299-0

Russo, N., Zecker, S., Trommer, B., Chen, J., and Kraus, N. (2009). Effects of background noise on cortical encoding of speech in autism spectrum disorders. *J. Autism Dev. Disord.* 39, 1185–1196. doi: 10.1007/s10803-009-0737-0

Schauder, K. B., and Bennetto, L. (2016). Toward an interdisciplinary understanding of sensory dysfunction in autism spectrum disorder: an integration of the neural and symptom literatures. *Front. Neurosci.* 10:268. doi: 10.3389/fnins.2016.00268

Schelinski, S., Borowiak, K., and von Kriegstein, K. (2016). Temporal voice areas exist in autism spectrum disorder but are dysfunctional for voice identity recognition. *Soc. Cogn. Affect. Neurosci.* 11, 1812–1822. doi: 10.1093/scan/nsw089

Schelinski, S., Riedel, P., and von Kriegstein, K. (2014). Visual abilities are important for auditory-only speech recognition: evidence from autism spectrum disorder. *Neuropsychologia* 65, 1–11. doi: 10.1016/j.neuropsychologia.2014.09.031

Schelinski, S., Roswandowitz, C., and von Kriegstein, K. (2017). Voice identity processing in autism spectrum disorder. *Autism Res.* 10, 155–168. doi: 10.1002/aur.1639

Schelinski, S., and von Kriegstein, K. (2020). Brief report: speech-in-noise recognition and the relation to vocal pitch perception in adults with autism spectrum disorder and typical development. *J Autism Dev. Disord.* 50, 356–363. doi: 10.1007/s10803-019-04244-1

Schoen, S. A., Miller, L. J., Brett-Green, B. A., and Nielsen, D. M. (2009). Physiological and behavioral differences in sensory processing: a comparison of children with autism spectrum disorder and sensory modulation disorder. *Front. Integr. Neurosci.* 3:29. doi: 10.3389/neuro.07.029.2009

Seri, S., Cerquiglini, A., Pisani, F., and Curatolo, P. (1999). Autism in tuberous sclerosis: evoked potential evidence for a deficit in auditory sensory processing. *Clin. Neurophysiol.* 110, 1825–1830. doi: 10.1016/s1388-2457(99)00137-6

Sharma, A., Kraus, N., McGee, T. J., and Nicol, T. G. (1997). Developmental changes in P1 and N1 central auditory responses elicited by consonant-vowel syllables. *Electroencephalogr. Clin. Neurophysiol.* 104, 540–545. doi: 10.1016/s0168-5597(97)00050-6

Sharma, S. R., Gonda, X., and Tarazi, F. I. (2018). Autism spectrum disorder: classification, diagnosis and therapy. *Pharmacol. Ther.* 190, 91–104. doi: 10.1016/j.pharmthera.2018.05.007

Simmons, D. R., Robertson, A. E., McKay, L. S., Toal, E., McAleer, P., and Pollick, F. E. (2009). Vision in autism spectrum disorders. *Vision Res.* 49, 2705–2739.

Sohal, V. S., and Rubenstein, J. L. R. (2019). Excitation-inhibition balance as a framework for investigating mechanisms in neuropsychiatric disorders. *Mol. Psychiatry* 24, 1248–1257. doi: 10.1038/s41380-019-0426-0

Sokhadze, E., Baruth, J., Tasman, A., Sears, L., Mathai, G., El-Baz, A., et al. (2009). Event-related potential study of novelty processing abnormalities in autism. *Appl. Psychophysiol. Biofeedback* 34, 37–51. doi: 10.1007/s10484-009-9074-5

St Clair, D. M., Blackwood, D. H., Oliver, C. J., and Dickens, P. (1987). P3 abnormality in fragile X syndrome. *Biol. Psychiatry* 22, 303–312. doi: 10.1016/0006-3223(87)90148-x

Stauder, J. E., Smeets, E. E., van Mil, S. G., and Curfs, L. G. (2006). The development of visual- and auditory processing in Rett syndrome: an ERP study. *Brain Dev.* 28, 487–494. doi: 10.1016/j.braindev.2006.02.011

Tager-Flusberg, H., and Kasari, C. (2013). Minimally verbal school-aged children with autism spectrum disorder: the neglected end of the spectrum. *Autism Res.* 6, 468–478. doi: 10.1002/aur.1329

Tager-Flusberg, H., Rogers, S., Cooper, J., Landa, R., Lord, C., Paul, R., et al. (2009). Defining spoken language benchmarks and selecting measures of expressive language development for young children with autism spectrum disorders. *J. Speech Lang. Hear. Res.* 52, 643–652. doi: 10.1044/1092-4388(2009/08-0136)

Tallal, P., Miller, S., and Fitch, R. H. (1993). Neurobiological basis of speech: a case for the preeminence of temporal processing. *Ann. N. Y. Acad. Sci.* 682, 27–47. doi: 10.1111/j.1749-6632.1993.tb22957.x

Tanaka, K., Ohta, S., Kinno, R., and Sakai, K. L. (2017). Activation changes of the left inferior frontal gyrus for the factors of construction and scrambling in a sentence. *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.* 93, 511–522. doi: 10.2183/pjab.93.031

Tecchio, F., Benassi, F., Zappasodi, F., Gialloreti, L. E., Palermo, M., Seri, S., et al. (2003). Auditory sensory processing in autism: a magnetoencephalographic study. *Biol. Psychiatry* 54, 647–654. doi: 10.1016/s0006-3223(03)00295-6

Van der Molen, M. J., Van der Molen, M. W., Ridderinkhof, K. R., Hamel, B. C., Curfs, L. M., and Ramakers, G. J. (2012a). Auditory and visual cortical activity during selective attention in fragile X syndrome: a cascade of processing deficiencies. *Clin. Neurophysiol.* 123, 720–729. doi: 10.1016/j.clinph.2011.08.023

Van der Molen, M. J., Van der Molen, M. W., Ridderinkhof, K. R., Hamel, B. C., Curfs, L. M., and Ramakers, G. J. (2012b). Auditory change detection in fragile X syndrome males: a brain potential study. *Clin. Neurophysiol.* 123, 1309–1318. doi: 10.1016/j.clinph.2011.11.039

Verleger, R., Jaskowski, P., and Wauschkuhn, B. (1994). Suspense and surprise: on the relationship between expectancies and P3. *Psychophysiology* 31, 359–369. doi: 10.1111/j.1469-8986.1994.tb02444.x

Vlaskamp, C., Oranje, B., Madsen, G. F., Mollegaard Jepsen, J. R., Durston, S., Cantio, C., et al. (2017). Auditory processing in autism spectrum disorder: mismatch negativity deficits. *Autism Res.* 10, 1857–1865. doi: 10.1002/aur.1821

Wass, S. V., Daubney, K., Golan, J., Logan, F., and Kushnerenko, E. (2019). Elevated physiological arousal is associated with larger but more variable neural responses to small acoustic change in children during a passive auditory attention task. *Dev. Cogn. Neurosci.* 37:100612. doi: 10.1016/j.dcn.2018.12.010

Whitehouse, A. J., and Bishop, D. V. (2008). Do children with autism 'switch off' to speech sounds? An investigation using event-related potentials. *Dev. Sci.* 11, 516–524. doi: 10.1111/j.1467-7687.2008.00697.x

Wunderlich, J. L., Cone-Wesson, B. K., and Shepherd, R. (2006). Maturation of the cortical auditory evoked potential in infants and young children. *Hear. Res.* 212, 185–202. doi: 10.1016/j.heares.2005.11.010

Yamaguchi, S., and Knight, R. T. (1991). P300 generation by novel somatosensory stimuli. *Electroencephalogr. Clin. Neurophysiol.* 78, 50–55. doi: 10.1016/0013-4694(91)90018-y

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

BRIEF RESEARCH REPORT
published: 01 November 2021
doi: 10.3389/fnins.2021.721922

Temporary Unilateral Hearing Loss Impairs Spatial Auditory Information Processing in Neurons in the Central Auditory System

*Jennifer L. Thornton[1], Kelsey L. Anbuhl[2] and Daniel J. Tollin[1]**

[1] Department of Physiology and Biophysics, University of Colorado School of Medicine, Aurora, CO, United States, [2] Center for Neural Science, New York University, New York, NY, United States

Temporary conductive hearing loss (CHL) can lead to hearing impairments that persist beyond resolution of the CHL. In particular, unilateral CHL leads to deficits in auditory skills that rely on binaural input (e.g., spatial hearing). Here, we asked whether single neurons in the auditory midbrain, which integrate acoustic inputs from the two ears, are altered by a temporary CHL. We introduced 6 weeks of unilateral CHL to young adult chinchillas via foam earplug. Following CHL removal and restoration of peripheral input, single-unit recordings from inferior colliculus (ICC) neurons revealed the CHL decreased the efficacy of inhibitory input to the ICC contralateral to the earplug and increased inhibitory input ipsilateral to the earplug, effectively creating a higher proportion of monaural responsive neurons than binaural. Moreover, this resulted in a ∼10 dB shift in the coding of a binaural sound location cue (interaural-level difference, ILD) in ICC neurons relative to controls. The direction of the shift was consistent with a compensation of the altered ILDs due to the CHL. ICC neuron responses carried ∼37% less information about ILDs after CHL than control neurons. Cochlear peripheral-evoked responses confirmed that the CHL did not induce damage to the auditory periphery. We find that a temporary CHL altered auditory midbrain neurons by shifting binaural responses to ILD acoustic cues, suggesting a compensatory form of plasticity occurring by at least the level of the auditory midbrain, the ICC.

Keywords: sound localization, inferior colliculus, conductive hearing loss, interaural level difference, mutual information

OPEN ACCESS

Edited by:
Victoria M. Bajo Lorenzana, University of Oxford, United Kingdom

Reviewed by:
Fernando R. Nodal, University of Oxford, United Kingdom
Joel I. Berger, The University of Iowa, United States

****Correspondence:***
Daniel J. Tollin
Daniel.tollin@cuanschutz.edu

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience

Received: *07 June 2021*
Accepted: *29 September 2021*
Published: *01 November 2021*

Citation:
Thornton JL, Anbuhl KL and Tollin DJ (2021) Temporary Unilateral Hearing Loss Impairs Spatial Auditory Information Processing in Neurons in the Central Auditory System. Front. Neurosci. 15:721922. doi: 10.3389/fnins.2021.721922

INTRODUCTION

Conductive hearing loss (CHL) during development can change auditory system structure and function (see reviews by Moore and King, 2004; Tollin, 2010; Whitton and Polley, 2011). Early life exposure to CHL, particularly unilateral, can lead to impairments in binaural hearing, including sound localization accuracy and acuity as well as spatial unmasking, even after resolution of the CHL and hearing sensitivity in both ears returns to normal (Moore et al., 1991; Ludwig et al., 2019). Binaural hearing deficits can recover, but can take months or years. During recovery, a child may present as audiologically normal, yet have lingering difficulty with speech perception in noisy, reverberant environments. As language is often learned in such environments, these impairments may contribute to deficits in language acquisition which is often observed in children with a history of chronic CHL (for review, see Whitton and Polley, 2011). CHL induced in adulthood can also be

detrimental to binaural hearing. For example, adults who experience prolonged periods of mild-moderate HL display worse performance on a binaural hearing task even long after corrective surgery (Hall and Derlacki, 1986; Ferguson et al., 1998). Decades of studies of the neural, anatomical and behavioral consequences of experimentally induced CHL in animal models have revealed effects related to the timing of onset, the duration, and the severity of the deprivation (Clopton and Silverman, 1977; Silverman and Clopton, 1977; Popescu and Polley, 2010; Keating and King, 2013; Polley et al., 2013; Keating et al., 2015; see Kumpik and King, 2019 for review). Regarding neural processing, these studies have generally examined how CHL alters basic neural response properties. However, it remains unclear how these alterations contribute to the persistent binaural hearing deficits. Here, we utilized the novel framework of information theory (Dayan and Abbott, 2001) to quantify how much information auditory midbrain neurons carry about a binaural cue to sound location (i.e., interaural level differences, ILD). We then examine how a temporary CHL induced in young adult animals alters neural information processing of binaural information, providing insight to underlying mechanisms of binaural function.

MATERIALS AND METHODS

All surgical and experimental procedures complied with the guidelines of the University of Colorado Anschutz Medical Campus Animal Care and Use Committees and the National Institutes of Health. Eleven young adult (∼P70; all female) chinchillas were used for the deprivation experiments while 19 normal-hearing animals (∼P70; 10 female, 9 male) were used for control data. This age was chosen as the head, pinnae, ear canal, and associated acoustical cues to sound location reach adult-like ranges at this time (Jones et al., 2011b). Similar to that described in Lupo et al. (2011), a small foam earplug (AO Safety, Indianapolis, IN, United States) was custom fit to the external ear canal of the animal and was then inserted into the left ear canal for 6 weeks. Foam earplugs (Lupo et al., 2011) produce similar spectrotemporal alterations of sound input to the cochlea as fluid in the middle ear (Thornton et al., 2012, 2013). Earplugs were checked daily to ensure stable placement within the ear canal.

Cochlear Microphonic Recordings

Animals (CHL: n = 11; control: n = 19) were anesthetized and prepared for electrophysiology as described in Jones et al. (2011a). Animals were anesthetized with an intramuscular (i.m.) injection of ketamine hydrochloride (KetaVed, 30 mg/kg i.m.) and xylazine hydrochloride (TranquiVed, 5 mg/kg im); supplementary injections were administered to maintain an adequate level of anesthesia. Next, a hole (2–3 mm diameter) was made in each bulla through which a silver ball electrode was placed on the round windows and fixed in place with dental acrylic, resealing the bullae. The cochlear microphonic (CM) was differentially amplified, filtered, and visually verified by oscilloscope. To quantify the magnitude of the CHL due to the earplug, free-field CM measurements were taken for both the left (earplugged in CHL group) and right (no earplug) ears for three different conditions: with the earplug in place (left ear), after the earplug was removed, and the right ear (within-animal control). Stimuli consisted of 10–ms sinusoids (2.5–ms rise/fall, 5–ms plateau) with octave steps from 0.25 to 20 kHz. Each stimulus was presented at least 25 times with a 40–ms interstimulus period.

Electrophysiological Recordings

Single unit, extracellular responses were recorded from neurons in the ICC as described in Jones et al. (2015). Briefly, extracellular electrical activity was measured with Parylene-coated tungsten microelectrodes (1–2 MΩ; Microprobe, Clarksburg, MD). Electrical activity was amplified (ISO-80, WPI, Sarasota, FL; Stanford Research Systems SRS 560, Sunnyvale, CA) and filtered (300–3,000 Hz). The colliculus was located stereotaxically and confirmation of ICC was determined by systematic changes in the characteristic frequencies (CFs) of neurons as the electrode was advanced. Candidate extracellular responses were isolated with a BAK amplitude-time window discriminator (model DDIS-1, Mount Airy, MD), and spike times were stored at a precision of 1μs via a TDT RV8. Neurons were studied if their single-unit spike waveforms exhibited good signal-to-noise ratio along with amplitude and temporal action potential morphology that was consistent from spike to spike. All recordings in the group of animals with CHL were performed the same day as the earplug removal.

Frequency-intensity response areas were measured with tone pips to estimate the CF and threshold. Next, neurons were classified in terms of their binaural input patterns in response to high level (20–30 dB re: threshold) CF tones. Briefly, neurons are given a classification that corresponds with the amount of excitation (or lack thereof, or "O") that they receive from each ear. For example, a neuron that receives excitatory inputs from both the contralateral and ipsilateral ears is classified as EE (one excitatory input, "E," from the contralateral ear and one excitatory input, "E," from the ipsilateral ear). A neuron with excitation from the contralateral and inhibition from the ipsilateral ear is designated EI. And some neurons only responded to stimulation of the contralateral ear and are designated EO.

Finally, in a smaller subset of neurons, ILD sensitivity was examined using 50 repetitions of 50-ms duration CF tones by holding the signal level to the contralateral ear (∼20 dB re: each neuron threshold) constant and varying the level at the ipsilateral ear from ≥ 25 dB below to 25 dB above ipsilateral threshold (5-dB steps). Following Brown and Tollin (2016), the rate vs. ILD for each neuron was fitted with a 4-parameter sigmoid, where rate (ILD) = $y0 + \alpha/(1\text{-}exp(\text{-}(ILD\text{-}ILD0)/\beta))$. Before fitting, the data were normalized to the maximum rate (**Figure 1**). The fits described the data for all neurons ($R > 0.9$); therefore, the fit parameters were used for all subsequent analysis. The parameters examined were half-max ILD (i.e., ILD at 50% of the maximal rate), rate-ILD slope (i.e., spikes/s/dB, not normalized; computed at half-max ILD), ILD dynamic range defined between 90 and 10% of maximum rate, and spike rate (or spike count) modulation defined as the difference between the maximum and minimum rate (or count) from the fitted rate-ILD function divided by the maximum rate (or count).

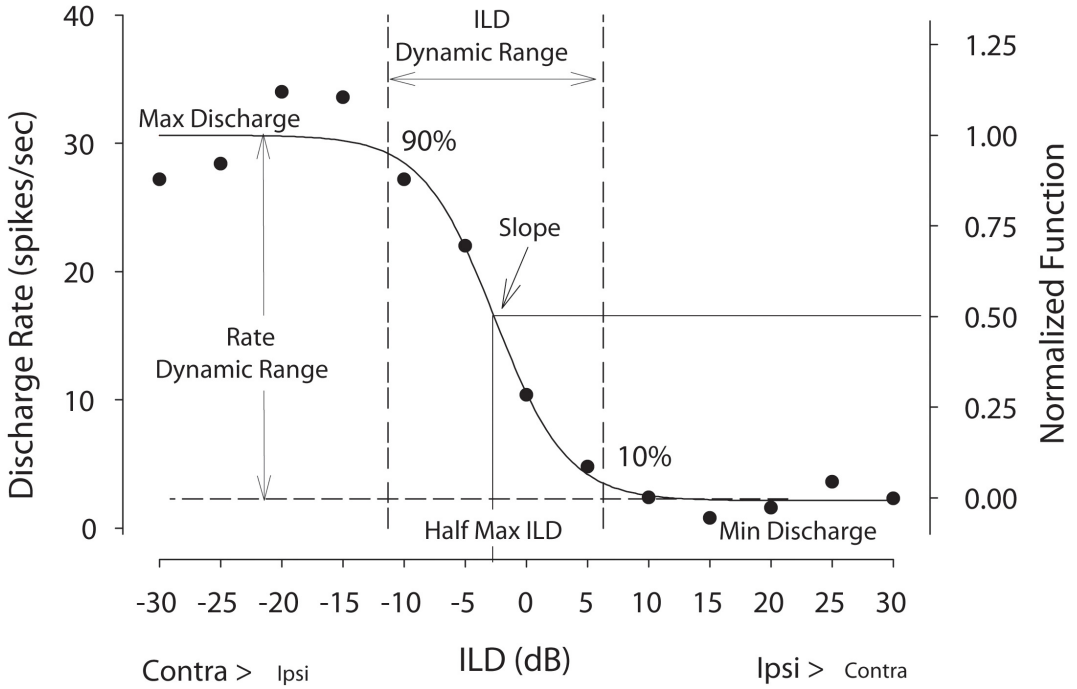

FIGURE 1 | Example discharge rate vs. ILD neural tuning function. Raw spike rate data were fitted with a 4-parameter logistic function from which key tuning parameters were computed including half-maximal ILD—the ILD producing 50% of maximal firing—and ILD dynamic range—the range of ILDs producing 10–90% of maximal firing. Rate dynamic range is defined as the difference between the maximum and minimum rates from the fitted curves divided by the maximum rate. For population level analyses, firing rates were normalized. This figure was adapted from Jones et al. (2015).

Neural Information Analysis—Mutual Information Computation

The mutual information (MI) is a measure of the strength of the association between two random variables, such as a spike count, *r*, and a given stimulus, *S* (Dayan and Abbott, 2001). MI is given by

$$MI(r, S) = \sum_i \sum_j p\left(S_j\right) p\left\langle r_i \mid S_j\right\rangle log_2 \left[\frac{p\left\langle r_i \mid S_j\right\rangle}{p\left(r_i\right)}\right]$$

where *p(Sj)* is the probability that the stimulus (S) had a particular value [S-values (i.e., ILDs) were presented with equal probability], *p(ri)* is the probability that the count was r_i at any value of *S*, and *p(ri| Sj)* is the probability that the count was r_i when the stimulus was S_j. Intuitively, MI will be high when the count variability is larger when computed across different stimuli than the variability computed within single presentations of a particular stimulus. MI summarizes all information contained in a neuronal response into a single, meaningful number as measured in bits. The benefit of using MI as a measure of neural sensitivity is that it allows one to make very few assumptions about the form of the neural response properties. The MI represents the upper bound on the information that even the "best "decoder" could represent (Benichoux et al., 2017). Thus, if CHL changes the information carrying capacity of ICC neurons then the MI will capture and quantify it.

RESULTS

Conductive Hearing Loss Due to Earplug Does Not Alter Auditory Periphery

To quantify the CHL caused by the earplug, as well as assay the function of the peripheral auditory system, sound-evoked CM responses were measured while the earplug was still in place and immediately after earplug removal (see Lupo et al., 2011; Thornton et al., 2012, 2013 for detailed methods). CM data from the right (unplugged) ear was used as a within-animal control, as unilateral CHL does not alter the CM in the non-occluded contralateral ear (Larsen and Liberman, 2010). The CHL was ∼10–15 dB for frequencies < 4 kHz increasing to ∼30 dB for frequencies > 4 kHz, consistent with Lupo et al. (2011). The magnitude of the CHL induced by earplugs is comparable to a CHL due to middle-ear effusion (Thornton et al., 2012, 2013). With earplugs, the mean CM thresholds across all frequencies and animals (n = 11) was 46.2 ± 7.1 dB. After removal of the earplug, CM thresholds were reduced to 30.7 ± 8.3 dB. Thus, the plug produced an across-frequency attenuation of 15.5 dB. A two-way repeated-measures ANOVA revealed that there was a significant decrease in attenuation after the earplug was removed [$F_{(1,\ 10)} = 103.3$, $p < 0.0001$]. There was no significant difference between the CM thresholds in the control ear and thresholds in the experimental ear after the earplug was removed [$F_{(1,\ 15)} = 3.71$, $p = 0.073$]. The return of CM thresholds to normal levels after earplug removal indicates that the hearing

loss induced by the plug was reversible, from 0.25 to 20 kHz. These data indicate that the earplug-induced CHL produced a reversible hearing loss without damaging the auditory periphery.

Unilateral Conductive Hearing Loss Does Not Alter Frequency Selectivity or Thresholds of Monaural and Binaural Sensitive Neurons

Monaural and binaural response properties were studied for a total of 323 IC neurons in normal control animals and 134 neurons from CHL animals after earplug removal. Neurons from the ICC contralateral to the previous CHL were analyzed.

The CFs for neurons with EE responses averaged 1.63 ± 2.0 kHz (median: 0.840, range: 0.140–8.95 kHz, n = 140) in controls and 1.76 ± 2.85 kHz (0.963, 0.320–13.77 kHz, n = 23) for neurons in the ICC contralateral to the CHL; CFs for EE neurons were not significantly different between control and experimental group [$t_{(159)}$ = –0.27, p = 0.79]. Thresholds for EE neurons averaged 14.44 dB SPL (median: 10, range: –15–50 dB, n = 140) for controls and 18.0 ± 12.5 dB SPL (15, 0–40 dB, n = 23) for neurons contralateral to the CHL; thresholds were not significantly different ([$t_{(159)}$ = –1.04, p = 0.3]. The CFs for neurons with EO responses averaged 5.1 ± 4.34 kHz (median: 4.2, range: 0.08–17.3 kHz, n = 95) in controls and 5.68 ± 3.42 kHz (5.15, 0.162–15.7 kHz, n = 82) for neurons contralateral to the CHL; CFs for EO neurons were not significantly different [$t_{(175)}$ = –0.97, p = 0.33]. Thresholds for EO neurons averaged 20.91 ± 14.18 dB SPL (median: 15, range: –10–55 dB SPL, n = 95) in controls and 23.84 ± 13.8 dB SPL (20, 0–60 dB SPL, n = 82) for neurons contralateral to the CHL; thresholds for EO neurons were not significantly different [$t_{(175)}$ = –1.38, p = 0.17]. Finally, CFs for neurons with EI responses averaged 7.25 ± 4.07 kHz (median: 6.28, range: 0.795–24.25 kHz, n = 90) in controls and 6.0 ± 3.56 kHz (5.38, 0.807–14.37 kHz, n = 26) for neurons contralateral to the CHL; CF for EI neurons were not significantly different [$t_{(114)}$ = 1.41, p = 0.16]. Thresholds for neurons with EI responses averaged 24.22 ± 12.36 dB SPL (median: 25, range: −10–55 dB SPL, n = 90) in controls and 29.62 ± 12.64 dB SPL (30, 10–50, n = 26) for neurons contralateral to the CHL; thresholds for EI neurons were not significantly different [$t_{(114)}$ = −1.95, p = 0.054].

Distribution of Monaural and Binaural Neural Responses Is Altered by Unilateral Conductive Hearing Loss

While there were no significant differences with the range of CFs or thresholds of ICC neurons between groups, there were substantial changes in the distributions of monaural and binaural responses. In controls, 42.7% of the neurons were EE, 29.4% EO and 27.9% EI while in animals with unilateral CHL, the proportions were 17.6% for EE, 62.6% for EO and 19.8% for EI. A chi-square test indicated that there was a significant difference in the proportions of monaural and binaural categories due to the CHL [$\chi^2(2, N = 457) = 72.13$, $p < 0.0001$]. This suggests that the unilateral CHL shifted ICC neurons (contralateral to the previous CHL) from primarily binaurally responsive units (i.e., EE, EI) to primarily monaural responsive units (i.e., EO).

Neural Coding of Interaural-Level-Difference Cues Is Altered by Unilateral Conductive Hearing Loss

Figure 1 illustrates the parameters of rate-vs.-ILD functions considered in this section. Half-max ILD values were compared between normal-hearing control animals and animals that received a unilateral earplug as adults. A shift in half-max ILD indicates a shift in the entire rate-ILD curve, signifying that that specific neuron encodes a different range of ILDs. For control animals, the mean half-max ILD was 1.9 ± 8.3 dB (n = 31 units), with a median value of 0.94 dB and a range of from −20.6 to 17.1 dB (**Figure 2C**, left and right panels, gray bars). For CHL animals, neurons in the ICC contralateral to the CHL ear (**Figure 2C**, left panel) exhibited a shift in the mean half-max ILD value to −10.26 ± 12.2 dB (n = 19 neurons) with a median value of −14.1 dB. The overall range of ILDs encoded by these neurons was shifted toward negative ILDs, ranging from −28.4 to 15.0 dB (**Figure 2C**, left panel, gray hatched bars). Relative to controls, the CHL produced an effective shift in ILD coding of 12.2 dB for neurons contralateral to the CHL. An unpaired t-test indicated a significant difference in half-max ILDs between normal and earplugged neurons [$t_{(48)}$ = 3.85, p = 0.0004].

CHL also altered the ILD dynamic range of the rate-ILD curves. The mean ILD dynamic range for control neurons was 26.1 ± 10.1 dB (median: 25.7 dB). Following CHL, neurons in the ICC contralateral to the CHL displayed a mean dynamic range of 19.4 ± 9.8 dB (median: 17.6 dB), which was significantly lower than control neurons [$t_{(48)}$ = 2.24, p = 0.031].

The slopes of the ILD functions in the ICC contralateral to the hearing loss were not significantly different between normal animals that animals that had a unilateral earplug-induced CHL [$t_{(48)} = -1.75, p > 0.05$]. Across ICC neurons in normal animals, the average rate-ILD slope was 2.34 ± 1.97 sp/s/dB with a median value of 2.06 sp/s/dB while in neurons contralateral to the CHL, the average rate-ILD slope was 2.9 ± 1.8 sp/s/dB with a median value of 2.4 sp/s/dB.

Finally, CHL altered the modulation of discharge rate due to ILD. In control neurons, the firing rate was modulated by 82 ± 16% (difference between unnormalized maximum and minimum rates derived from the fitted function divided by the maximum) with presentation of a range of ILD values (± 25 dB to the ipsilateral ear). Neurons contralateral to the CHL were modulated by 58 ± 18%, significantly less than controls [$t_{(48)}$ = 4.9, $p < 0.0001$]. The overall reduction in rate-modulation to ILD in neurons from CHL animals results in shallower rate-ILD functions. This is due to an elevation of the discharge rate at the minima of the ILD function, with no change to the maximum discharge rate. This suggests that CHL induced an overall reduction in inhibition in ICC neurons.

Finally, linear regression of the ILD tuning curve properties considered above against the CF of the neuron were conducted. The earplugs produced larger CHL at frequencies ∼ > 4 kHz

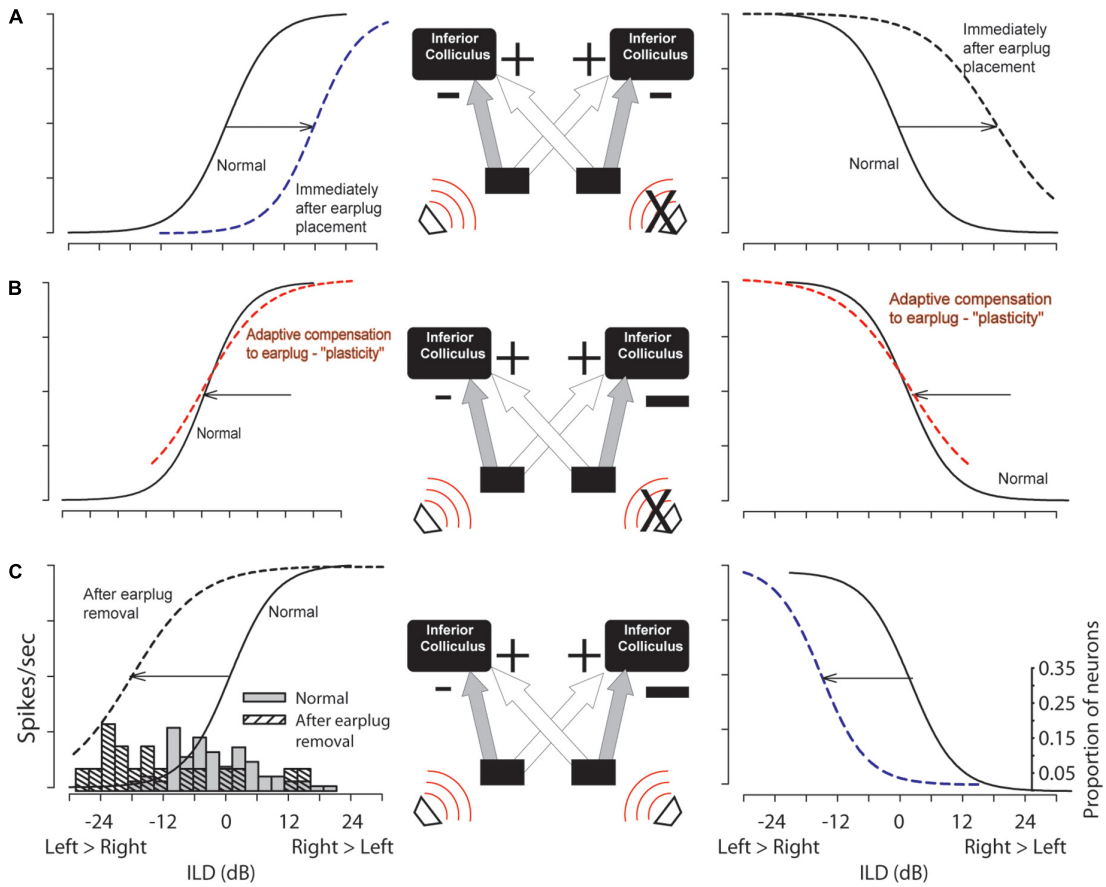

FIGURE 2 | Hypothesized changes due to CHL (right ear, "X") in circuit function and the sensitivity to ILDs in the left (contralateral to earplug) and right (ipsilateral to earplug) ICC are shown in **(A,B)** with empirical data shown in **(C)**. The simplified circuit shows ipsilateral inhibitory ("–") and contralateral excitatory ("+") inputs to the ICC (strengths indicated by sizes of the symbols). **(A)** Relative to ILD coding in normal hearing controls (solid lines) ILD coding shifts to the right due to earplug insertion (dashed). **(B)** Hypothetical shifts in ILD coding if circuit plasticity has compensated for unilateral earplug over time. **(C)** Upon earplug removal, ILD coding shifts left due to the CHL-induced altered circuitry. Histograms: empirical half-max ILDs form ICC neurons in normal hearing controls (n = 31) and experimental group after earplug removal (n = 19).

than those below and it is possible that binaural neurons with higher CF might have been more impacted by the earplug than neurons with lower CF. Of all of the regression analyses, only one test was significant. **Figure 3** shows that the half-max ILD was shifted significantly more for neurons of higher CF than lower after earplug removal ($r = 0.53$, $p = 0.02$); there was no correlation ($r = 0.14$, $p = 0.47$) between half-max ILD and CF in the control animal neurons, consistent with what we have reported before in a much more extensive dataset (Jones et al., 2015).

Inferior Colliculus Responses Carry Less Information Regarding Interaural-Level Difference Cues Following Conductive Hearing Loss

Figure 4 shows mutual information computed for 31 neurons from control animals (black bars, symbols) and for 19 neurons measured from CHL animals in the ICC contralateral to the CHL (white bars and symbols). Neurons from control animals exhibited a mean MI of 0.7 ± 0.29 (median: 0.64) bits, whereas neurons from experimental animals (mean: 0.44 ± 0.15; median: 0.41 bits) were significantly lower [$t_{(48)} = -3.4$, $p = 0.0015$]. The responses of ICC neurons thus carried ∼37% less mutual information regarding ILD cues following a unilateral CHL as compared to normal-hearing controls. For neurons contralateral to the CHL, reduction in the information-carrying capacity was consistent with the significant reduction in the amount by which ILD modulated the spike count relative to the max count revealed in the earlier section. This is consistent with an effective reduction in inhibition due to the CHL. While there was a significant correlation between the half-max ILD and CF of neurons after earplug removal, there was no significant correlations between either MI or spike count modulation and neuron CF.

DISCUSSION

Altered inputs to the auditory system can result in anatomical, physiological, and behavioral changes that persist long

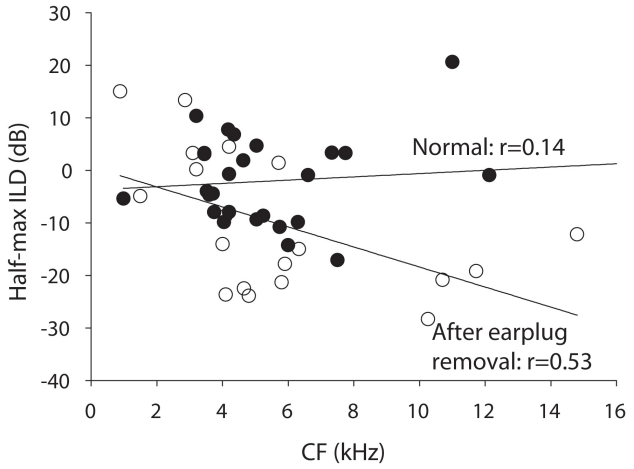

FIGURE 3 | Half-maximal ILD as a function of CF for the normal control animals (black circles) and the experimental animals after 6 weeks of unilateral CHL (open circles).

beyond resolution of the hearing impairment (reviewed by Moore and King, 2004; Tollin, 2010; Whitton and Polley, 2011; Kumpik and King, 2019). The majority of evidence for CHL-induced plasticity in the auditory system comes from developmental studies in humans and animals. However, studies in adult humans and animals have also suggested that CHL can induce plasticity and that subjects can adapt to altered auditory inputs particularly via behavioral training paradigms. The data presented here suggests a compensatory mechanism for plasticity by at least the level of the inferior colliculus (ICC) as well as reduced information processing of ICC neurons. **Figure 2** illustrates our general hypothesis regarding compensatory changes in the ascending circuitry to the IC in response to a unilateral CHL. In normal-hearing circuitry (**Figure 2A**, solid lines), spike rate is modulated by ILD sigmoidally with maximum responses for ILDs favoring the excitatory contralateral ear and reduced responses for ILDs favoring the inhibitory ipsilateral ear. Immediately after introduction of a CHL (in this case, earplug insertion), the rate-ILD curves would shift toward the right simply because the acoustical input from the contralateral ear has been attenuated (i.e., less effective excitatory input). This is represented by the dashed lines in **Figure 2A**.

Figure 2B illustrates the hypothesized circuit changes that would occur if mechanisms were to compensate for the altered sound localization cues due to CHL (see Lupo et al., 2011; Thornton et al., 2012, 2013). Compensatory mechanisms could potentially work to shift the rate-ILD curves back toward normal (dashed line overlapping normal curves). To achieve this kind of adaptive compensation in the ICC contralateral to the CHL, the strength (or gain) of inhibitory input from the ipsilateral normal-hearing ear (left side in example) is hypothesized to be reduced and/or the strength (or gain) of the excitation from the contralateral CHL-ear increased; the size of the "+" and "−" symbols have been adjusted in **Figure 2B** to illustrate this change. After removing the CHL, the effective changes to the ILD-coding pathways to the ICC can be revealed. If the circuit had been altered as in **Figure 2C**, then after CHL removal the rate-ILD curves are hypothesized to shift toward the left (**Figure 2C**, black dashed line, left column), demonstrating a reduced ipsilateral inhibitory (and/or increased contralateral excitatory) response when compared to normal. Our data is consistent with this hypothesis. **Figure 3** shows that after earplug removal the shift was larger for neurons with higher CF than lower. These results are consistent with the hypothesis that the earplugs induced larger CHL for higher frequencies than lower and that the compensation for the CHL by the circuit was larger at higher CFs.

A similar compensatory response is expected in ICC neurons that are ipsilateral to the CHL (**Figure 2**, right panels), although insufficient data were collected in these studies to test this hypothesis. Immediately after introduction of a CHL, there will be an effective reduction in the strength of inhibition to the ICC ipsilateral to the CHL simply due to the attenuation of sound (**Figure 2A**, dashed line, right column). If adaptive compensation occurs, the strength of excitatory contralateral inputs will be reduced in order to match the reduced inhibitory inputs and/or an increase in the strength of the ipsilateral inhibitory input to match the normal contralateral excitation. These changes would effectively shift the rate-ILD curves back to normal with the CHL in place (**Figure 2B**, dashed line, right column). Immediately after CHL removal, an overall large inhibitory response would remain, causing the rate-ILD curves to shift to the left of normal (**Figure 2C**, dashed line, right column). Future studies will be required to test these predictions.

The present results are inconsistent with some previous physiological results in the ICC of animals with unilateral CHL (summarized by Moore and King, 2004; Tollin, 2010; Whitton and Polley, 2011). Prior studies in rats demonstrated that CHL persistently reduced the effectiveness of inputs to the two ICCs from the ear with the CHL, a finding that produces data consistent with illustrations in **Figure 2A** as opposed to **Figure 2C** (Clopton and Silverman, 1977; Silverman and Clopton, 1977; Popescu and Polley, 2010). One possible reason for this may be that the experiments in the current study were performed in the chinchilla which is a precocious species (Jones et al., 2011a,b) that also has good low-frequency hearing. Additionally, the results of the prior studies could potentially be due to an altered periphery due to the CHL, where there might be a residual hearing loss due to peripheral damage due to the CHL even after CHL removal, which would also yield results as in **Figure 2A** (dashed lines) even without central auditory-system plasticity. However, the most plausible explanation for the difference is that the prior studies were conducted in developing animals that were earplugged at or around hearing onset, while the present studies were done in audiologically mature young adult animals. More studies are needed to reveal the sources of the differences in the results.

While compensatory plasticity may or may not occur as illustrated in **Figure 2**, there is no doubt that CHL exerts

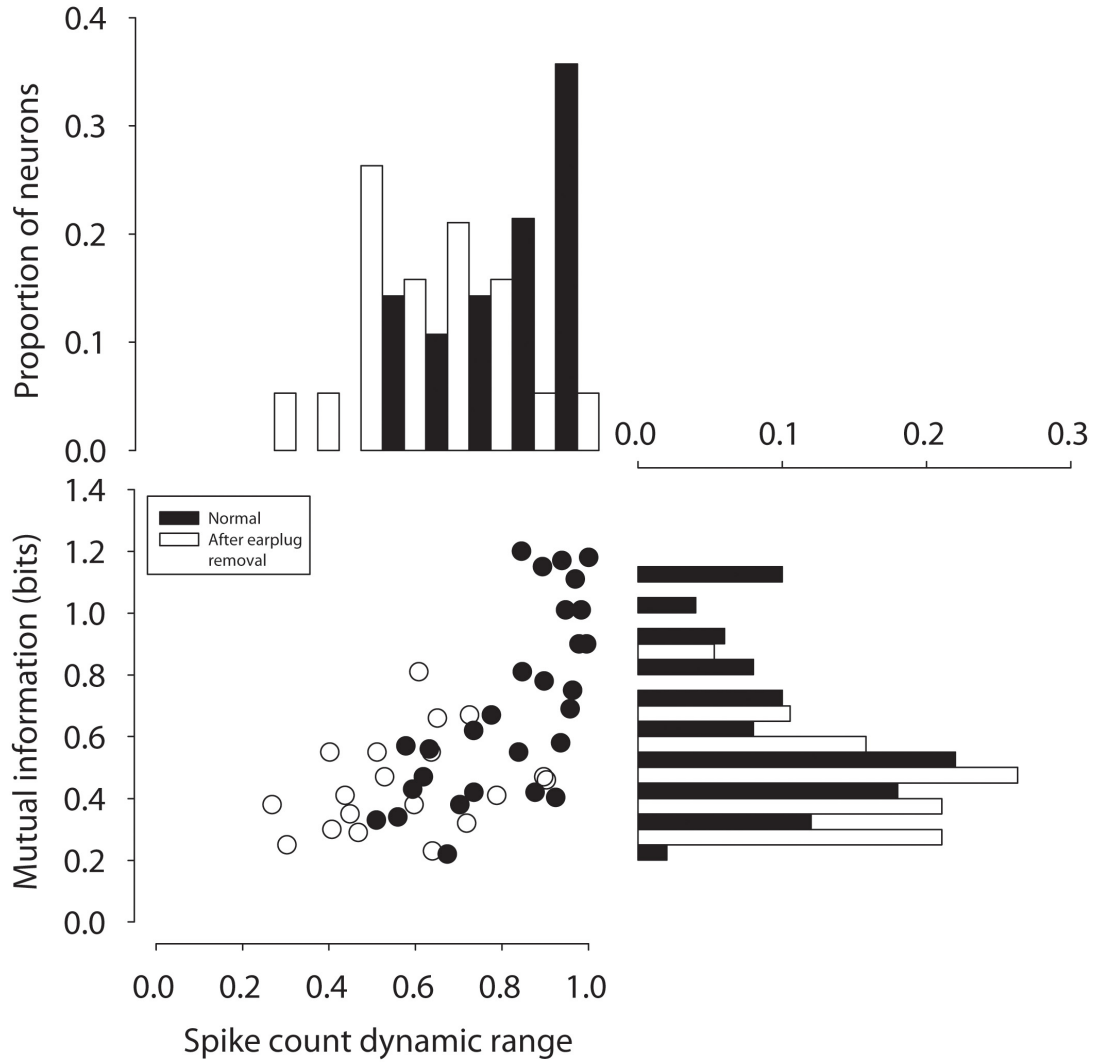

FIGURE 4 | Mutual information between spike count and ILD in ICC neurons in normal-hearing (black circles) adults and after 6 weeks of unilateral CHL (open circles) is plotted as a function of the spike count modulation dynamic range by ILD (% re: max count).

an effect on the neural coding of spatial information in ICC neurons that persists after CHL resolution, as demonstrated by the 37% reduction in the capacity of neurons to carry information about ILDs (**Figure 4**). Reduced MI may suggest alterations in the responsiveness (spike rates), reliability (spike rate variability), as well as the general sensitivity of ICC neurons and/or their inputs to the cues to location, including ILD. The results suggest that at least for ICC neurons contralateral to the CHL, a reduction in the capacity of ILD to modulate spiking was correlated with a reduction in information-carrying capacity of these neurons. The impaired neural information processing demonstrated here may provide a basis for the persistent behavioral deficits in binaural and spatial hearing tasks that have been observed clinically after chronic CHL both during development and in adulthood. Since we have found persistent reductions in the capability of critical neural circuits in the ascending auditory pathway to encode spatial attributes of sound, it may logically follow that there will be a similar reduction in the perceptual capabilities as well. One caveat of this study is that the effects of the unilateral CHL were studied within 24–36 h after earplug removal. It is possible that the effects we observed were transient, and the binaural deficits may resolve over time. Toward this end, ongoing studies are examining the behavioral consequences of reduced information processing due to CHL induced during both development and into adulthood as well as longitudinal studies post CHL resolution.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The animal study was reviewed and approved by the University of Colorado Anschutz Medical Campus Animal Care and Use Committees.

AUTHOR CONTRIBUTIONS

JT and DT designed the research and secured the funding. JT, KA, and DT collected, analyzed the data, and wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the NIH NIDCD grants R01-DC011555 (DT) and F31-DC011198 (JT).

ACKNOWLEDGMENTS

The content of this manuscript has been presented in part at the 4th International Symposium on Auditory and Audiological Research, August 28–30, 2013 Hotel Nyborg Strand, Denmark. We thank members of the Tollin lab for support and assistance in data collection and analyses.

REFERENCES

Benichoux, V., Brown, A. D., Anbuhl, K. L., and Tollin, D. J. (2017). Representation of multidimensional stimuli: quantifying the most informative stimulus dimension from neural responses. *J. Neurosci.* 37, 7332–7346. doi: 10.1523/jneurosci.0318-17.2017

Brown, A. D., and Tollin, D. J. (2016). Slow temporal integration enables robust neural coding and perception of a cue to sound source location. *J. Neurosci.* 36, 9908–9921. doi: 10.1523/jneurosci.1421-16.2016

Clopton, B. M., and Silverman, M. S. (1977). Plasticity of binaural interaction. II. Critical period and changes in midline response. *J. Neurophysiol.* 40, 1275–1280. doi: 10.1152/jn.1977.40.6.1275

Dayan, P., and Abbott, L. F. (2001). *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. Cambridge, MA: MIT Press.

Ferguson, M. O., Cook, R. D., Hall, J. W., Grose, J. H., and Pillsbury, H. C. (1998). Chronic conductive hearing loss in adults: effects on the auditory brainstem response and masking-level difference. *Arch. Otolaryngol. Head Neck Surg.* 124, 678–685. doi: 10.1001/archotol.124.6.678

Hall, J. W., and Derlacki, E. L. (1986). Effect of conductive hearing loss and middle ear surgery on binaural hearing. *Ann. Otol. Rhinol. Laryngol.* 95, 525–530. doi: 10.1177/000348948609500516

Jones, H. G., Brown, A. D., Koka, K., Thornton, J. L., and Tollin, D. J. (2015). Sound frequency-invariant neural coding of a frequency-dependent cue to sound source location. *J. Neurophysiol.* 114, 531–539. doi: 10.1152/jn.00062.2015

Jones, H. G., Koka, K., and Tollin, D. J. (2011b). Postnatal development of cochlear microphonic and compound action potentials in a precocious species, chinchilla lanigera. *J. Acoust. Soc. Am.* 130, 38–43.

Jones, H. G., Koka, K., Thornton, J., and Tollin, D. J. (2011a). Concurrent development of the head and pinnae and the acoustical cues to sound location in a precocious species, the chinchilla (*Chinchilla laniger*). *J. Assoc. Res. Otolaryngol.* 12, 127–140. doi: 10.1007/s10162-010-0242-3

Keating, P., and King, A. J. (2013). Developmental plasticity of spatial hearing following asymmetric hearing loss: context-dependent cue integration and its clinical implications. *Front. Syst. Neurosci.* 7:123. doi: 10.3389/fnsys.2013.00123

Keating, P., Dahmen, J. C., and King, A. J. (2015). Complementary adaptive processes contribute to the developmental plasticity of spatial hearing. *Nat. Neurosci.* 18, 185–187. doi: 10.1038/nn.3914

Kumpik, D. P., and King, A. J. (2019). A review of the effects of unilateral hearing loss on spatial hearing. *Hear. Res.* 372, 17–28. doi: 10.1016/j.heares.2018.08.003

Larsen, E., and Liberman, M. C. (2010). Contralateral cochlear effects of ipsilateral damage: no evidence for interaural coupling. *Hear. Res.* 260, 70–80. doi: 10.1016/j.heares.2009.11.011

Ludwig, A. A., Zeug, M., Schonwiesner, M., Fuchs, M., and Meuret, S. (2019). Auditory localization accuracy and auditory spatial discrimination in children with auditory processing disorders. *Hear. Res.* 377, 282–291.

Lupo, J. E., Koka, K., Thornton, J. L., and Tollin, D. J. (2011). The effects of experimentally induced conductive hearing loss on the spectral and temporal aspects of sound transmission through the ear. *Hear. Res.* 272, 30–41.

Moore, D. R., and King, A. J. (2004). "Plasticity of binaural systems," in *Development of the Auditory System, Springer Handbook of Auditory Research*, eds T. N. Parks, E. W. Rubel, R. R. Fay, and A. N. Popper (New York: Springer-Verlag), 96–172. doi: 10.1007/978-1-4757-4219-0_4

Moore, D. R., Hutchings, M. E., and Meyer, S. E. (1991). Binaural masking level differences in children with a history of otitis media. *Audiology* 30, 91–101. doi: 10.3109/00206099109072874

Polley, D. B., Thompson, J. H., and Guo, W. (2013). Brief hearing loss disrupts binaural integration during two early critical periods of auditory cortex development. *Nat.Commun.* 4:2547. doi: 10.1038/ncomms3547

Popescu, M. V., and Polley, D. B. (2010). Monaural deprivation disrupts development of binaural selectivity in auditory midbrain and cortex. *Neuron* 65, 718–731. doi: 10.1016/j.neuron.2010.02.019

Silverman, M. S., and Clopton, B. M. (1977). Plasticity of binaural interaction. I. Effect of early auditory deprivation. *J. Neurophysiol.* 40, 1266–1274. doi: 10.1152/jn.1977.40.6.1266

Thornton, J. L., Chevallier, K. M., Koka, K., and Tollin, D. J. (2013). Conductive hearing loss induced by experimental middle ear effusion in a chinchilla model reveals impaired TM-coupled ossicular chain movement. *J. Assoc. Res. Otolaryngol.* 14, 451–464. doi: 10.1007/s10162-013-0388-x

Thornton, J. L., Chevallier, K. M., Koka, K., Lupo, J. E., and Tollin, D. J. (2012). The conductive hearing loss due to an experimentally-induced middle ear effusion alters the interaural level and time difference cues to sound location. *J. Assoc. Res. Otolaryngol.* 13, 641–654. doi: 10.1007/s10162-012-0335-2

Tollin, D. J. (2010). "The development of sound localization mechanisms," in *Oxford Handbook of Developmental Behavioral Neuroscience*, eds M. S. Blumberg, J. H. Freeman, and S. R. Robinson (New York: Oxford University Press), 262–282.

Whitton, J. P., and Polley, D. B. (2011). Evaluating the perceptual and pathophysiological consequences of auditory deprivation in early postnatal life: a comparison of basic and clinical studies. *J. Assoc. Res. Otolaryngol.* 12, 535–546. doi: 10.1007/s10162-011-0271-6

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 11 November 2021
doi: 10.3389/fnins.2021.772943

Myelination Deficits in the Auditory Brainstem of a Mouse Model of Fragile X Syndrome

Alexandra Lucas[1], Shani Poleg[1], Achim Klug[1†] and Elizabeth A. McCullagh[2†]*

[1] Department of Physiology and Biophysics, University of Colorado Anschutz Medical Campus, Aurora, CO, United States,
[2] Department of Integrative Biology, Oklahoma State University, Stillwater, OK, United States

Auditory symptoms are one of the most frequent sensory issues described in people with Fragile X Syndrome (FXS), the most common genetic form of intellectual disability. However, the mechanisms that lead to these symptoms are under explored. In this study, we examined whether there are defects in myelination in the auditory brainstem circuitry. Specifically, we studied myelinated fibers that terminate in the Calyx of Held, which encode temporally precise sound arrival time, and are some of the most heavily myelinated axons in the brain. We measured anatomical myelination characteristics using coherent anti-stokes Raman spectroscopy (CARS) and electron microscopy (EM) in a FXS mouse model in the medial nucleus of the trapezoid body (MNTB) where the Calyx of Held synapses. We measured number of mature oligodendrocytes (OL) and oligodendrocyte precursor cells (OPCs) to determine if changes in myelination were due to changes in the number of myelinating or immature glial cells. The two microscopy techniques (EM and CARS) showed a decrease in fiber diameter in FXS mice. Additionally, EM results indicated reductions in myelin thickness and axon diameter, and an increase in g-ratio, a measure of structural and functional myelination. Lastly, we showed an increase in both OL and OPCs in MNTB sections of FXS mice suggesting that the myelination phenotype is not due to an overall decrease in number of myelinating OLs. This is the first study to show that a myelination defects in the auditory brainstem that may underly auditory phenotypes in FXS.

Keywords: myelination, auditory brainstem, Fragile X Syndrome, coherent anti-stokes Raman scattering, medial nucleus of the trapezoid body

OPEN ACCESS

Edited by:
Marc Schönwiesner, Leipzig University, Germany

Reviewed by:
Christian Keine, University of Oldenburg, Germany
Adrian Rodriguez-Contreras, City College of New York (CUNY), United States

***Correspondence:**
Elizabeth A. McCullagh
elizabeth.mccullagh@okstate.edu

†These authors share senior authorship

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience

Received: *09 September 2021*
Accepted: *21 October 2021*
Published: *11 November 2021*

Citation:
Lucas A, Poleg S, Klug A and McCullagh EA (2021) Myelination Deficits in the Auditory Brainstem of a Mouse Model of Fragile X Syndrome. Front. Neurosci. 15:772943. doi: 10.3389/fnins.2021.772943

INTRODUCTION

Fragile X Syndrome (FXS) is the most common monogenic form of autism spectrum disorder (ASD) and occurs in 1:4,000–1:8,000 people in the United States (Hagerman and Hagerman, 2008). FXS has been commonly used as a model for studying ASD, especially because there are several commercially available animal models such as mouse and rat (The Dutch-Belgian Fragile X Consorthium et al., 1994; Till et al., 2015; Tian et al., 2017). There have been many proposed mechanisms for how loss of Fragile X Mental Retardation Protein (FMRP), the protein encoded by the gene *Fmr1*, leads to FXS and ASD phenotypes (Bear et al., 2004; Hagerman et al., 2009; Osterweil et al., 2013; Gantois et al., 2017; Rajaratnam et al., 2017, among others). However, drug therapies in animal models do not always "rescue" human phenotypes when extended to the clinic

(Dahlhaus, 2018). Recent work has shown that myelination may underly some the phenotypes common in ASD and FXS (Pacey et al., 2013; Phan et al., 2020). Auditory phenotypes, such as auditory hypersensitivity, are often conserved between mouse and human FXS, and it is possible that myelination deficits are also conserved (McCullagh et al., 2020b; Phan et al., 2020). There is reduced or delayed myelination in FXS in many brain areas (Pacey et al., 2013; Phan et al., 2020) and one of the targets of FMRP is myelin basic protein (Jeon et al., 2017), suggesting that deficits seen in FXS may be caused at least in part by alterations to myelination. Additionally, research has shown FMRP expression in mature oligodendrocytes (OLs), potentially explaining how loss of FMRP impacts myelination (Wang et al., 2004; Giampetruzzi et al., 2013). The mechanisms by which binaural hearing is impaired in people with FXS and ASD are unknown, however, there is evidence the auditory brainstem is involved (Kulesza and Mangunay, 2008; Kulesza et al., 2011; Wang et al., 2014; Rotschafer et al., 2015; Garcia-Pino et al., 2017; McCullagh et al., 2017, 2020a; Zorio et al., 2017; Curry et al., 2018; El-Hassar et al., 2019; Lu, 2019).

Previous work has shown that FXS impacts the auditory brainstem and sound localization processing in a myriad of ways, from protein dysfunction to behavioral changes [reviewed in McCullagh et al. (2020b)]. FMRP expression is high in the brainstem, suggesting an important role in brainstem function (Wang et al., 2014; Zorio et al., 2017). Alterations to proteins which directly interact with FMRP can critically disrupt the maintenance of neuronal activity patterns in the brainstem (Brown et al., 2010; Strumbos et al., 2010; El-Hassar et al., 2019). In addition, previous anatomical and physiological work has shown that excitation/inhibition balance within the auditory brainstem circuit is altered in FXS (Rotschafer et al., 2015; Garcia-Pino et al., 2017; McCullagh et al., 2017; Rotschafer and Cramer, 2017; Curry et al., 2018; Lu, 2019). This circuit relies on precise timing of excitatory and inhibitory inputs for computation of sound location information (reviewed in Grothe et al., 2010). Behavioral studies have shown subtle, but persistent, auditory specific behavioral phenotypes in FXS mice that likely originate in altered auditory brainstem processing (McCullagh et al., 2020a) however, the exact mechanism that underlies these alterations is unknown.

The ability to localize sound is dependent on a very precise comparison of sound input from the two ears, either through interaural timing differences (ITD) or interaural sound intensity/level differences (IID). Processing and interpretation of ITD and IID relies on extremely fast and temporally precise synaptic transmission, leading to some of the most heavily myelinated axons in the brain (Ford et al., 2015). In the auditory brainstem, the medial nucleus of the trapezoid body (MNTB) afferent axons stand out due to their large diameters, which are about ∼3 μm in healthy hearing animals (Ford et al., 2015). The large diameter is due to a high demand on speed and temporal precision by this synapse (Kim J.H. et al., 2013). Any alteration in this pathway that changes the precision and timing of this circuit will lead to substantial difficulties in the ability to separate competing auditory streams and localize sound, as occurs in ASD. Interestingly, increased changes to latency for both behavioral and physiological measures seem to be one of the more repeatable phenotypes in FXS mice (Kim H. et al., 2013; McCullagh et al., 2020a, but see Rotschafer et al., 2015; El-Hassar et al., 2019). One way in which changes to myelination in the auditory brainstem would likely manifest is through changes in the speed of electrical propagation, and therefore latency of auditory brainstem timing.

Several studies have shown that myelination deficiencies contribute to ASD phenotypes in several brain areas. This suggests that an area of the brain most critically dependent on myelination for properly encoding sound location information, such as the auditory brainstem, should be especially impacted (Pacey et al., 2013; Phan et al., 2020). We use a combination of microscopy techniques [coherent anti-stokes Raman spectroscopy (CARS) and electron microscopy (EM)] to examine fine myelin microstructure in Fmr1 KO mice compared to wildtype controls (B6). Additionally, we measured the number of mature and precursor oligodendrocytes (OLs and OPCs) as a potential mechanism through which myelin microstructure is altered in FXS mice. We hypothesize that there will be reduced myelin (thickness and diameter) and concomitant changes in number of OLs and OPCs in FXS mice. The goal of this study is to provide a possible mechanism through which auditory phenotypes might arise in FXS.

MATERIALS AND METHODS

All experiments complied with all applicable laws, National Institutes of Health guidelines, and were approved by the University of Colorado Anschutz Institutional Animal Care and Use Committee.

Animals

All experiments were conducted in either C57BL/6J (stock #000664, B6) background or hemizygous male and homozygous *Fmr1* knock-out strain (B6.129P2-*Fmr1*^{TM1Cgr}/J stock #003025, Fmr1 KO) obtained from The Jackson Laboratory (Bar Harbor, ME, United States) (The Dutch-Belgian Fragile X Consorthium et al., 1994). Animals from both sexes were used in the experiments for both B6 and Fmr1 KO mice. There were no significant differences in any of the measures based on sex, so data for both sexes are combined for analysis (p-values shown are the main effect of sex; CARS $p = 0.7454$, OL count $p = 0.9529$, OPC count $p = 0.2103$, EM data was one animal of each sex so differences do not apply here). Exact number of animals used are listed in the figure legends for each experiment but ranged from 4–7 for each genotype and were between P72-P167 (postnatal day).

Tissue Preparation

Mice were overdosed with pentobarbital (120 mg/kg body weight) and transcardially perfused with phosphate-buffered saline (PBS; 137 mM NaCl, 2.7 mM KCl, 1.76 mM KH_2PO_4, 10 mM Na_2HPO_4 Sigma-Aldrich, St. Louis, MO, United States) followed by 4% paraformaldehyde (PFA). After perfusion, the animals were decapitated, and the brains removed from the skull. Brains were kept overnight in 4% PFA before

transferring to PBS. Brainstems were embedded in 4% agarose (in PBS) and sliced coronally using a Vibratome (Leica VT1000s, Nussloch, Germany) at either 200 μm thickness for myelination analysis with CARS or 70 μm thickness for labeling oligodendrocytes.

For EM imaging, mice were perfused with PBS followed by a solution containing 2.5% glutaraldehyde, 2% paraformaldehyde, and 0.1 M cacodylate buffer at pH 7.4. The brains were stored in the same solution for 24 h and sectioned at 500 μm using the same protocol as above. The sections were immersed in glutaraldehyde solution for a minimum of 24 h at 4°C. The tissue was then processed by rinsing in 100 mM cacodylate buffer followed by immersion in 1% osmium and 1.5% potassium ferrocyanide for 15 min. Next, the tissue was rinsed five times in the cacodylate buffer and immersed again in 1% osmium for 1.5 h. After rinsing five times for 2 min each in cacodylate buffer and two times briefly in water, e*n bloc* staining with 2% uranyl acetate was done for at least 1 h at 4°C, followed by three rinses in water. The tissue was transferred to graded ethanols (50, 70, 90, and 100%) for 15 min each and then finally to propylene oxide at room temperature, after which it was embedded in LX112 and cured for 48 h at 60°C in an oven. Ultra-thin parasagittal sections (55 nm) were cut on a Reichert Ultracut S (Leica Microsystems, Wetzlar, Germany) from a small trapezoid positioned over the tissue and were picked up on Formvar-coated slot grids (EMS, Hatfield, PA, United States).

Immunohistochemistry

For staining oligodendrocytes, six to eight free-floating sections from each brain were submerged in L.A.B. solution (Liberate Antibody Binding Solution, Polysciences, Inc., Warrington, PA, United States, Cat No. 24310) for ten min to help expose epitopes related to labeling mature and precursor oligodendrocytes. Next, the slices were washed two times in PBS and blocked in a solution containing 0.3% Triton-X (blocking solution), 5% normal goat serum (NGS) and PBS for 1 h on a laboratory shaker. After blocking, slices were stained with primary antibodies rabbit anti-Aspartoacylase (GeneTex, Irvine, CA, United States; 1:1,000) and mouse Sox-10 (A-2) (Santa Cruz, Dallas, TX, United States; sc-365692, 1:500) diluted in blocking solution with 1% NGS and incubated overnight (**Table 1**). Slices were then washed three times (10 min each wash) in PBS and incubated in secondary antibodies (**Table 2**) diluted in blocking solution with 1% NGS for 2 h. After three additional washes in PBS (5 min each wash), Nissl (Neurotrace 425/435 Blue-Fluorescent Nissl Stain, Invitrogen, Carlsbad, CA, United States, 1:100) in antibody media [AB media: 0.1 M phosphate buffer (PB: 50 mM KH_2PO_4, 150 mM Na_2HPO_4), 150 mM NaCL, 3 mM Triton-X, 1% bovine serum albumin (BSA)] was applied for thirty min. Stained slices were then briefly washed in PBS and slide-mounted with Fluoromount-G (Southern Biotech, Cat.-No.: 0100-01, Birmingham, AL, United States). Slides were stored at 4°C. Slices used for myelination analysis with CARS (four to five slices per brain) were stained with Nissl as described above, then stored free-floating at 4°C until imaged. All antibody labeling was performed at room temperature.

Antibody Characterization

The primary antibody for mature oligodendrocytes (Aspartoacylase (ASPA), 1:1,000, GeneTex, Irvine, CA, United States; GTX113389; RRID:AB_10727411, **Table 1**) is a rabbit polyclonal antibody specific to mature oligodendrocytes (OLs). The ASPA antibody is specific to a recombinant protein encompassing a sequence within the center region of the human aspartoacylase. Aspartoacylase is responsible for the conversion of *N*-acetyl_L-aspartic acid (NAA) to aspartate and acetate. Hydrolysis by aspartoacylase is thought to help maintain white matter (Bitto et al., 2007). The ASPA antibody has been shown to label mature oligodendrocytes in mouse cerebral cortex and has been validated by protein overexpression and western blot analysis (Larson et al., 2018; Orthmann-Murphy et al., 2020). Additional western blot analyses with various whole cell extracts also showed that ASPA antibody detects aspartoacylase protein. The primary antibody Sox-10 [Sox-10 (A-2), 1:500, Santa Cruz Biotechnology, Heidelberg, Germany; sc-365692; RRID:AB_10844002] was used to label the entire oligodendrocyte lineage [oligodendrocyte precursor cells (OPCs) and OLs]. Sox-10 is a mouse monoclonal antibody specific to an epitope mapping between amino acids 2–29 at the N-terminus of the human Sox-10 gene. Sox-10 has been shown to specifically

TABLE 1 | Primary antibodies used in immunofluorescence.

Antibody	Immunogen	Manufacturer, species, mono or polyclonal, cat. or lot no., RRID	Concentration
Aspartoacylase (ASPA)	Recombinant protein encompassing a sequence within the center region of human Aspartoacylase	GeneTex (Irvine, CA, United States), rabbit polyclonal, GTX113389, RRID:AB_10727411	1:1,000
Sox-10 (A-2)	An epitope between amino acids 2-29 at the N-terminus of Sox-10 protein.	Santa Cruz Biotechnology, (Dallas, TX, United States), mouse monoclonal, sc-365692; RRID:AB_10844002	1:500

TABLE 2 | Secondary antibodies used in immunofluorescence.

Catalog No., RRID	Manufacturer	Reactivity	Conjugate	Concentration
A-21428, AB_2535849	Invitrogen	Goat anti-Rabbit IgG (H + L)	Alexa Fluor 555	1:500
A-21235, AB_2535804	Invitrogen	Goat anti-Mouse IgG (H + L)	Alexa Fluor 647	1:500

label the oligodendrocyte lineage in the CNS and in mouse brains (Zuo et al., 2018; Barak et al., 2019; Imamura et al., 2020). Both primary antibodies were visualized using two fluorescent-conjugated secondary antibodies which are listed in **Table 2**.

Imaging

Brainstem slides for immunofluorescence were imaged using an Olympus FV1000 confocal microscope (Olympus, Tokyo, Japan) with lasers for 488, 543, and 635 nm imaging. Once MNTB was identified by the distinctive trapezoidal morphology and cell size, z-stacks were taken using a 20x objective (UPLSAPO20X, NA 0.75) so that the entire nucleus could be visualized and quantified. The nucleus was separated into medial, central, and lateral regions following the protocol in Weatherstone et al., 2016. Briefly, the MNTB was digitally extracted using FIJI software, and the tonotopic axis was estimated by drawing the longest possible dorsomedial-to-ventrolateral line through the nucleus (Weatherstone et al., 2016). This line was measured using FIJI and divided into thirds wherein perpendicular lines were drawn. These lines delineated the lateral, central and medial regions of MNTB. For CARS microscopy, brainstem sections were imaged using an Olympus FV-1000 (Olympus Corp., Tokyo, Japan) fitted with non-descanned detectors in both the forward and epi CARS directions. Sections were placed in a culture dish with coverslip (for inverted microscopy) and custom weight to keep tissue near coverslip. Z-stacks were taken with a 60X, 1.2NA infrared corrected water objective which served for collection of the CARS signal in the epi direction for medial and lateral MNTB. Signal in the forward direction was collected through an Olympus 0.55NA condenser. The APE picoEmerald laser set up contained an NKT fiber laser which provided the 80 MHz clock, and an OPO (Optical Parametric Oscillator) laser with a tunable range of 770–990 nm. Water absorption peak was at 1,388 cm^{-1} which implies a 901.1 nm pump/probe beam with the fixed Stokes beam at 1,031 nm and a CARS signal at 801.6 nm. To excite the CH_2 stretch at 2,845 cm^{-1} the APE picoEmerald lasers required a pump/probe beam set to 797.2 nm which resulted in a CARS signal at 649.8 nm, thus laser settings were always set to 797.2 nm prior to imaging. Pulse duration for the two lasers was approximately 2 ps and the polarization was linear and horizontal. The synchronization was primarily based on the OPO cavity length with the pulse pumped by the frequency doubled output from the NKT laser at 515.5 nm. The tunable pump/probe beam was used for the TPEF (two-photon excitation fluorescence), which was separated by a dichroic from the CARS signal. The NDD (non-descanned detector) epi-direction unit contained two PMTs of the same type. The CARS and fluorescence do not share PMTs. These settings were used for optimal lipid signal to selectively image myelination in this area.

As an additional control, one brain from a wild type mouse and one brain from a knockout were processed and imaged with electron microscopy. The main purpose of this control was to verify that CARS microscopy has sufficient resolution to discriminate between the small axon diameter differences that are studied here. Sections for EM were first imaged on a compound microscope, and MNTB was identified by the characteristic cell size and shape, as well as the location of the cells in the parasagittal slices. Once MNTB location was determined, slices were imaged on a FEI Tecnai G2 transmission electron microscope (Hillsboro, OR, United States) with an AMT digital camera (Woburn, MA, United States).

Cell Counting

Quantification of OL's and entire OL lineage [oligodendrocyte precursor cells (OPCs) and OLs] was performed using the optical fractionator approach and FIJI software (Schindelin et al., 2012). In a pilot study to determine the appropriate stereological parameters, five brains were sliced at 50 μm using a freezing microtome (SM2010R Leica, Nussloch, Germany) and z-stacks from five sections per brain (every other section) were taken with a 20x objective the same as above. Using FIJI each stack was brightened, the background subtracted, the image scaled and the MNTB was digitally extracted. Separate grid plugins were used to create the dissector and count frames for each image stack. To account for tissue shrinkage, the section thickness was measured at multiple points within MNTB using the Olympus FV1000 microscope and FV10-ASW microscopy software. Measurements were taken by recording the first z-level where tissue features came into focus, and then recording the last z-level where tissue features were in focus. The difference between the recorded levels for each measured area of MNTB was averaged and represented the section thickness for stereological calculations. Stacks were set to contain 20 slices regardless of the section thickness and the step size was kept constant within, but not between brains. Cells were counted by moving through the stack and using a cell counter plugin to mark the top of the cell when it first came into focus within the probe unless it contacted the exclusion lines of the count frame (Schmitz and Hof, 2005).

The pilot data indicated that selecting four 70 μm sections (every other section starting from a randomly selected slice) to give a section subfraction (SSF) of 0.280 was appropriate for calculating the total number of cells in a single MNTB. In addition, counts would use a probe area of (1,250 μm × 1,250 μm) and dissector (5,000 μm × 5,000 μm) giving an area subfraction (ASF) of 0.25. It should be noted that two brains had fewer than four slices (3.5 and three slices, respectively) which were suitable for quantification. OL analyses also included three brains from the pilot study that had adequate immunohistochemical staining for quantification. In these cases, the SSF was adjusted, or a fifth slice was counted to achieve an SSF that was closest to 0.280 (0.20–0.25). The height subfraction (HSF) was calculated separately for each brain using the mean dissector height (OL mean = 19.5–34.4 μm; OPC mean = 21.06–34.4 μm) and the mean section thickness (OL mean = 26.75–45.75 μm; OPC mean = 27.75–45.75 μm). Using these parameters, OL's and OPC's were counted and recorded in excel. OPC's were identified as cells that were labeled with Sox 10, but not co-labeled with ASPA (**Figure 1**). Calculation of the total cell count for each complete MNTB nucleus was done by multiplying the counted cells (ΣQ^-) with the reciprocal sampling fractions such

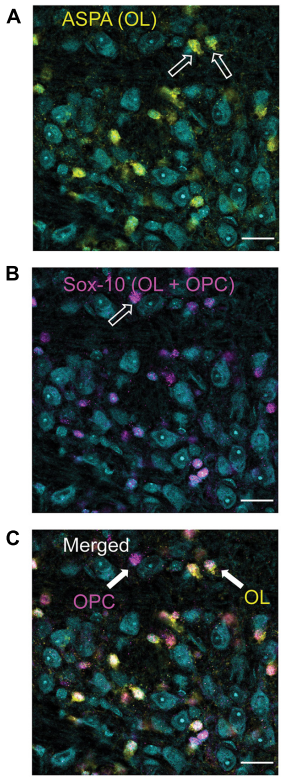

FIGURE 1 | Differentiation of oligodendrocyte (OLs, ASPA) and oligodendrocyte precursor cells (OPCs, Sox-10 no ASPA). ASPA, a stain specific for mature OLs is shown in yellow [panel **(A)**]. These yellow cells were used to calculate the number of mature OLs in each image. Sox-10 is a marker specific for the entire OL lineage (OPCs and OLs) and is labeled in magenta **(B)**. To determine the number of OPCs in each image, cells that were labeled with Sox-10 but not ASPA (as indicated by the arrow labeled OPC) were counted **(C)**. MNTB cell bodies are indicated in cyan with a Nissl stain.

that:

$$N = \Sigma Q^{-} \times 1/\mathrm{HSF} \times 1/\mathrm{ASF} \times 1/\mathrm{SSF}$$

Error in the stereological approach was calculated using the Gundersen-Jensen coefficient of error (CE) estimator and smoothness class of $m = 1$ (Gundersen et al., 1999). CE measures error as a function of the biological variance (noise) and the sampling variance from systemic uniformly random sampling. The coefficient of variance (CV) within and between brains was high ($CV > 1$) due to natural variations in the population of glial cells throughout MNTB. CE for OL and OPC estimates was greater than 10% (Mean CE = 0.51), but the sampling parameters chosen showed CE was negligible in terms of overall variation ($CE^2/CV^2 < 0.003$) (Gundersen et al., 1999). Thus, the stereological parameters were determined to be sufficient for quantification estimates.

Myelination Analysis

Analysis of myelination morphology was done using FIJI software. Myelin diameter was measured from CARS images by using the line tool to measure the fiber diameter (**Figure 2A**) for 20 axons per image. The resolution was not sufficient in CARS images to quantify thickness or g-ratio of axons. Tissue shrinkage was not corrected for in this analysis. In addition, diameters of

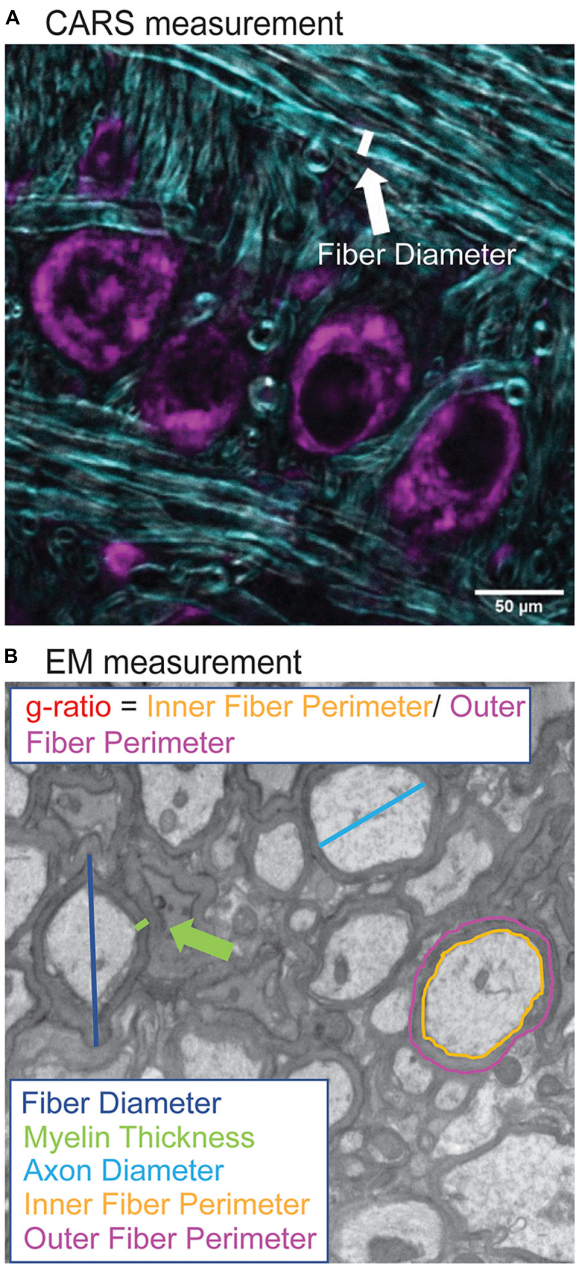

FIGURE 2 | Quantification of myelination morphology in CARS and EM images. CARS images were quantified using the FIJI line segment tool. A line, as indicated by white line **(A)**, was drawn and the measure function used to measure the fiber diameter for 20 axons per image (purple indicates cell bodies stained with Nissl, CARS signal in cyan). For EM images, increased resolution allowed for more detailed analysis including fiber diameter (dark blue), axon diameter (cyan), myelin thickness (green), inner (orange) and outer (purple) fiber diameter **(B)**. Inner and outer fiber perimeter were then used to calculate g-ratio, which is the inner fiber diameter/outer fiber perimeter.

240 axons were also measured in the same way in EM images, however, myelin thickness was measured directly, and g-ratios were calculated from fiber and axon diameters (**Figure 2B**; Ford et al., 2015). G-ratios were calculated from EM images using inner and outer fiber perimeter ratio (g-ratio = inner fiber perimeter /outer fiber perimeter). EM data is representative of one animal from each genotype and was performed primarily to confirm CARS diameter changes.

Statistical Analyses

Figures were generated in R (R Core Team, 2013) using ggplot2 (Wickham, 2016). Boxplots dislay the median and 25th–75th percentiles (or 1st and 3rd quartiles, respectively) the whiskers represent ± 1.5 times the interquartile range. Data that falls outside the range are plotted as individual points. Data for all experiments were analyzed using linear mixed-effects models [lme4; (Bates et al., 2014)]. To account for repeated measures and variability with animals, Animal was considered a random effect with fixed effects of genotype and location (medial, center, lateral, and total when appropriate). Additional analyses used fixed effects of age or sex to determine if these were contributing factors in the differences shown. It was expected that there may be differences between different regions of the MNTB (medial, center, lateral, total) therefore *a priori*, it was determined that estimated marginal means [emmeans; (Lenth, 2019)] would be used for pairwise comparisons between region and genotype. To control for multiple comparisons, emmeans implements a Tukey method for these contrasts. Asterisks are used to indicate statistical significance between the two genotypes, as follows: $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$. Figures were prepared for publication using Photoshop and Illustrator (Adobe, San Jose, CA, United States).

RESULTS

We used several (EM and CARS) microscopy techniques to examine the myelin microstructure of putative globular bushy cell axons in the auditory brainstem that innervate the MNTB and form the Calyx of Held in Fmr1 KO mice compared to wildtype controls (B6). We also counted the number of both mature and precursor oligodendrocytes in the same region to show if structural changes are due to changes in number or type of oligodendrocytes which myelinate these axons.

Myelin Diameter as Measured by Coherent Anti-stokes Raman Spectroscopy

We first examined the diameter of axons in the medial and lateral MNTB using CARS microscopy (**Figure 3**). CARS microscopy for myelination imaging was used over electron microscopy (EM) due to its speed and compatibility with immunofluorescence (Wang et al., 2005), in this case we used a Nissl stain for simultaneous cell body staining. We measured axon fiber diameters in both medial and lateral MNTB because previous work has shown the potential for differences in myelin based on tonotopic location (medial or lateral) (Ford et al., 2015; Stange-Marten et al., 2017). Representative images for both location (medial or lateral columns) for B6 and Fmr1 KO (rows) are shown in **Figures 3A–D**. We found that there was a significant decrease in fiber diameter in Fmr1 KO compared to B6 for both medial and lateral MNTB (**Figure 3E**). Not surprisingly when data were summed between medial and lateral locations, there was still a significant decrease in fiber diameter in Fmr1 KO mice.

Coherent Anti-stokes Raman Spectroscopy Myelin Characteristics Confirmed by Electron Microscopy

To confirm that the CARS technology was sensitive enough to determine differences between fiber diameters in Fmr1 KO and B6 mice, we also performed EM controls on one animal of each genotype (across multiple sections with multiple measurements, **Figure 4**). The advantage of EM, while costly and time consuming, is that it allows for higher resolution and thus more detailed description of changes to myelination than CARS microscopy techniques. Since alterations in fiber diameter seen in CARS data was not dependent on medial or lateral localization, and it is difficult to determine location with parasagittal EM sections, these data are not separated by location. Representative EM images for B6 and Fmr1 KO are shown in **Figures 4A,B**, respectively. Consistent with CARS analysis, EM data showed smaller fiber diameters in Fmr1 KO animals compared to B6 (**Figure 4C**). Additionally, axon diameter and myelin thickness are also decreased in Fmr1 KO compared to B6 (**Figures 4D,E**). Lastly, g-ratio, an often-used index of optimal functional and structural myelination, was increased in Fmr1 KO animals compared to B6 (**Figure 4F**; Chomiak and Hu, 2009). These results are consistent with the relatively large axon diameters in the Fmr1 KO mice compared to fiber diameters (a measure for g-ratio) despite thinner myelin and smaller axons than B6 mice.

Number of Mature Oligodendrocytes

Oligodendrocytes (OLs) are the glial cells responsible for myelinating axons in the central nervous system. FMRP has been shown to be expressed in OLs, potentially providing a mechanism through which loss of FMRP can lead to changes in myelination (Wang et al., 2004; Giampetruzzi et al., 2013). For OL counts, the MNTB was divided into three sections, medial, central, and lateral. **Figures 5A,B** show representative images of the entire MNTB stained with ASPA (a marker for mature OLs) and sox-10 (a marker for the entire OL lineage) for B6 and Fmr1 KO mice, respectively. As shown by arrowheads, OLs are represented by staining with ASPA (which also coincides with sox-10 staining). Cells that are only stained for sox-10 (magenta) and not ASPA (yellow) are quantified as OPCs (see below and **Figures 1**, **5**). Like CARS data, there was no difference in OL number based on specific tonotopy of the MNTB (medial, center, or lateral) however, when data was combined for all regions, there was a significant increase in OL number in Fmr1 KO mice compared to B6 (**Figure 5C** top panel). Using the sampling Q count technique, to estimate number of cells within the entire MNTB, similar results were found to raw counts.

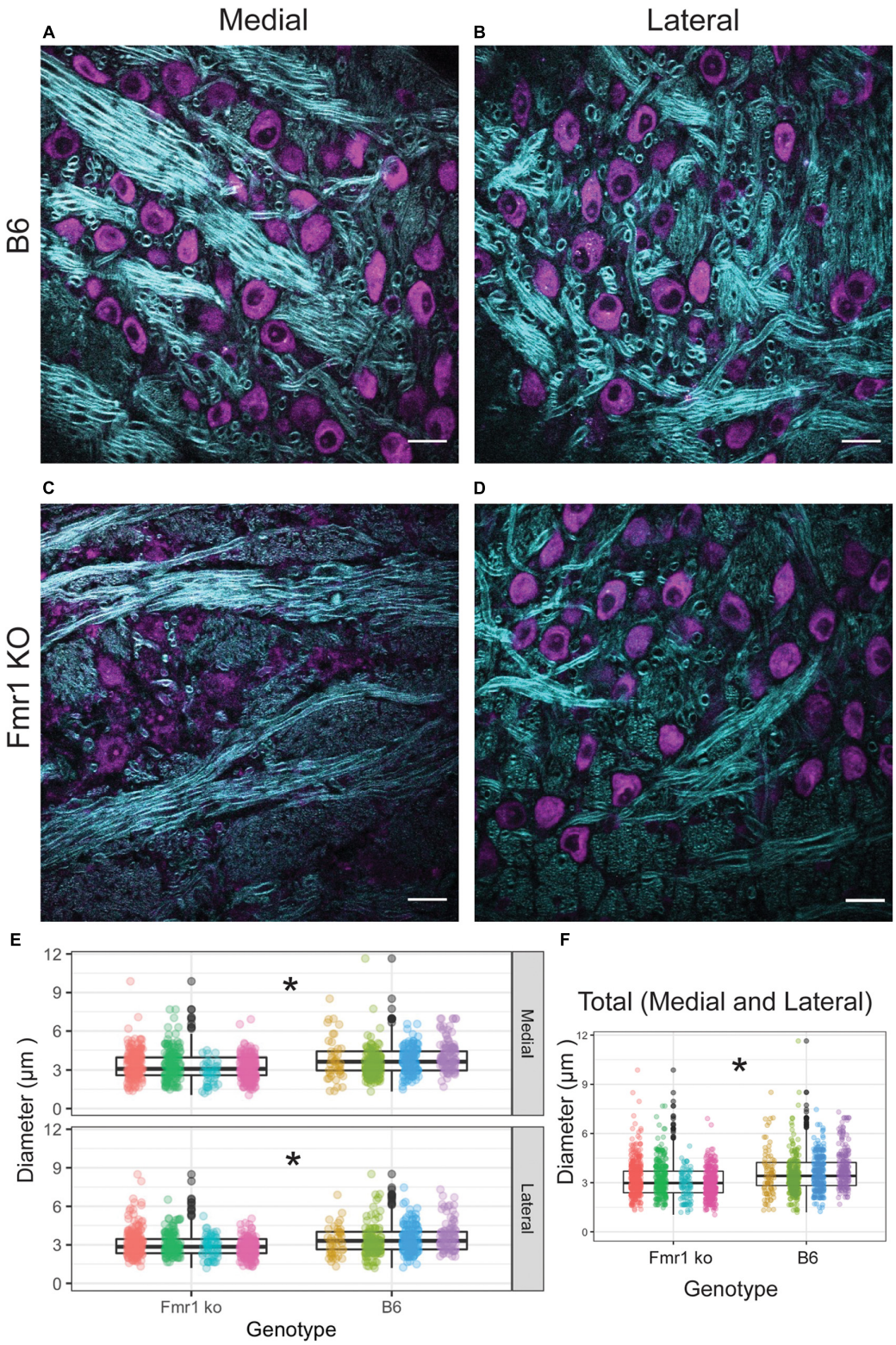

FIGURE 3 | Fiber diameter size is reduced in Fmr1 KO mice compared to B6 using CARS. Representative images for medial and lateral (columns) MNTB in B6 and Fmr1 KO mice (rows) (purple indicates cell bodies stained with Nissl, CARS signal in cyan) **(A–D)**. Scale bar is 20 μm. Box plots show data from each individual animal (color spread) and population data represented by the box for Fmr1 KO (left) and B6 (right) for either medial (top) or lateral (bottom) MNTB **(E)**. N for each genotype was four animals with 20 measurements per section and 4–6 sections per animal. Panel **(F)** shows the data independent of medial or lateral location (total) for each animal across genotype. *$p < 0.05$.

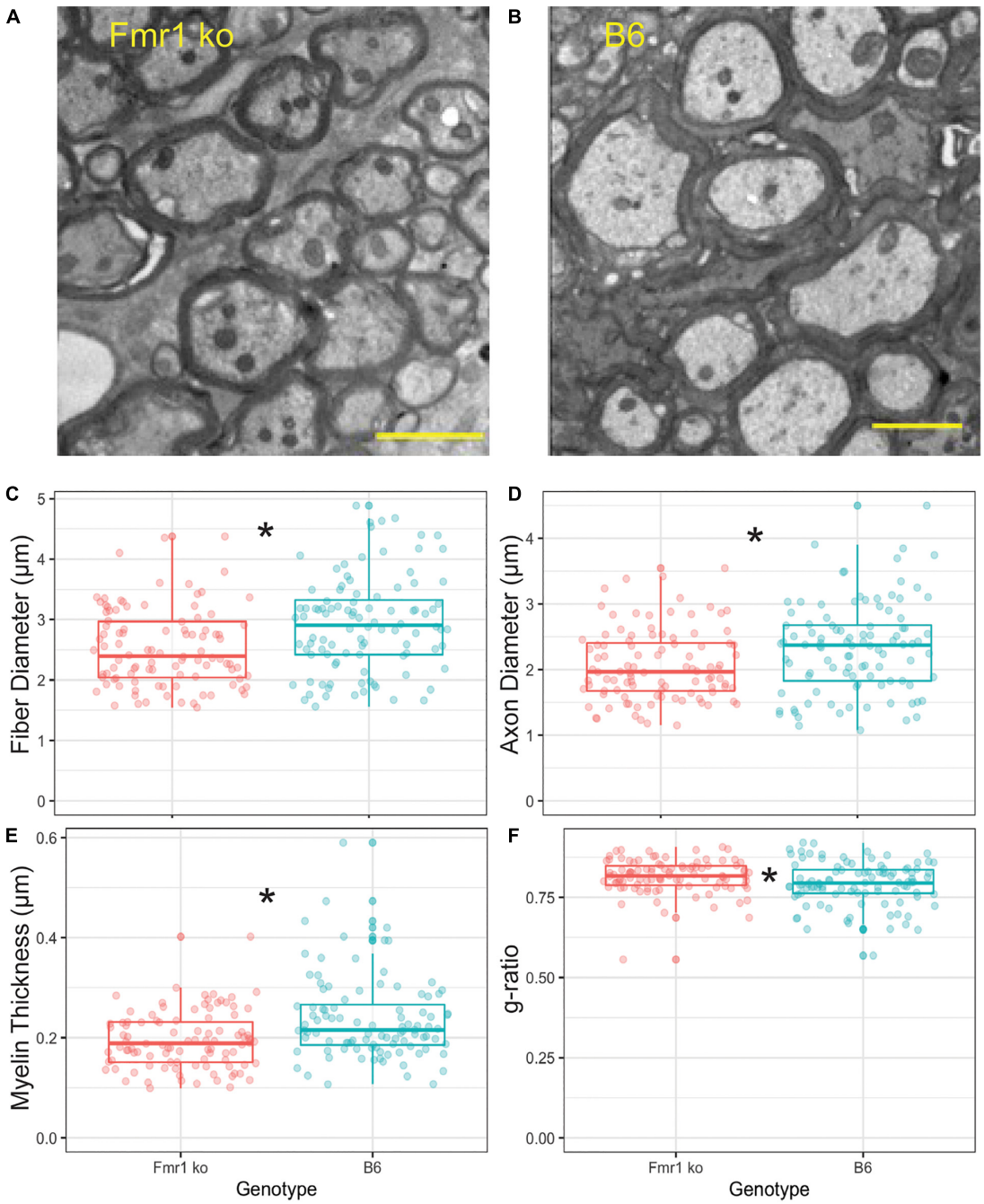

FIGURE 4 | Myelin characteristics are also altered in Fmr1 KO compared to B6 mice measured by EM. Representative images for EM in Fmr1 KO **(A)** and B6 **(B)** mice. Scale bar is 2 µm. Box plots show data from each individual measurement for Fmr1 KO (left) and B6 (right) mice. Measured parameters include fiber diameter **(C)**, axon diameter **(D)**, myelin thickness **(E)**, and g-ratio **(F)**. N is one animal per genotype with multiple sections and multiple measurements/section represented by individual puncta. $^*p < 0.05$.

Specifically, there was no difference in medial, center, or lateral estimates, but total combined regions showed a highly significant ($p < 0.0001$) increase in OLs in Fmr1 KO animals compared to B6. In addition to changes in number of mature OLs, another possible mechanism underlying alterations to myelination could be incomplete or impaired maturation of OLs.

Number of Precursor Oligodendrocytes

Other brain areas (cerebellum, neocortex, and others) show changes in myelination in FXS related to impaired maturation or function of oligodendrocyte precursor cells (Wang et al., 2005; Pacey et al., 2013; Jeon et al., 2017). Like OL count measures, the MNTB was divided into medial, central, and lateral subsections

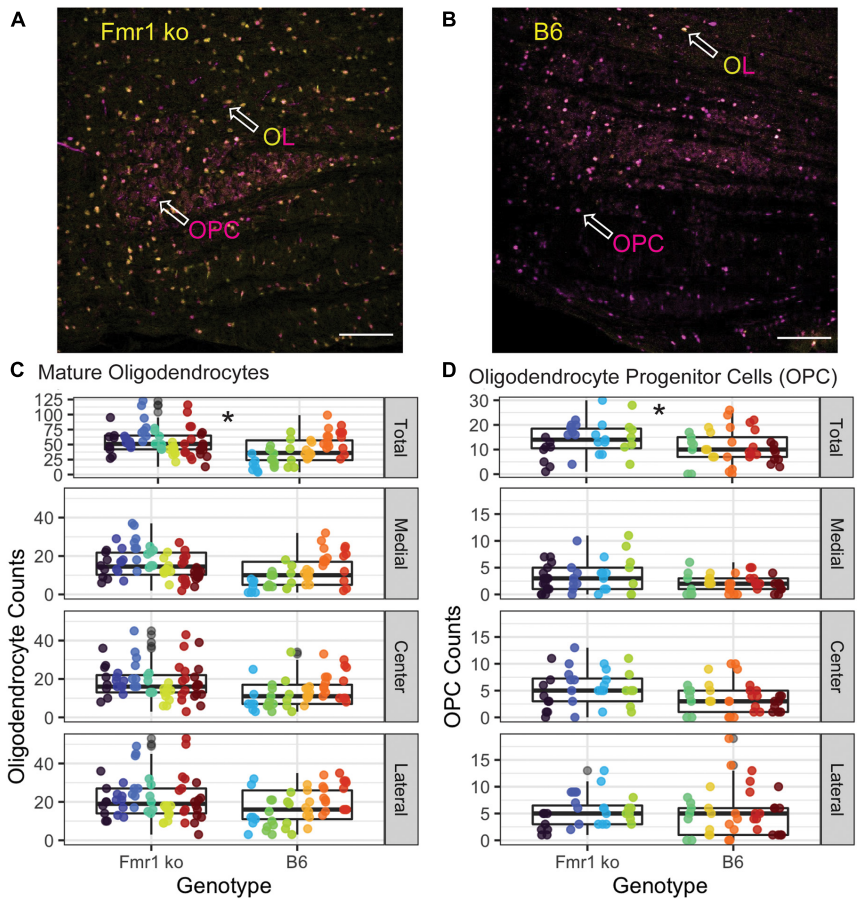

FIGURE 5 | Oligodendrocyte and OPC number are increased in the MNTB independent of tonotopy in Fmr1 KO mice compared to B6. Representative images for OLs (yellow/magenta ASPA/sox-10 marker, arrowhead) and OPCs (magenta sox-10 marker, arrowhead without ASPA (yellow) staining) in Fmr1 KO **(A)** and B6 **(B)** mice. Scale bar is 100 μm. Box plots show data from each individual measurement for Fmr1 KO (left) and B6 (right) for each animal (as indicated by different colors within the boxplots) for the total, medial, center, and lateral MNTB for OLs **(C)**. N is 7 Fmr1 KO animals and 6 B6 animals. Box plots show data from each individual measurement for Fmr1 KO (left) and B6 (right) for each animal (as indicated by different colors within the boxplots) for the total, medial, center, and lateral MNTB **(D)** for OPCs. N is 4 Fmr1 KO animals and 5 B6 animals. *$p < 0.05$.

(**Figure 5D**), with representative images of the entire MNTB shown in **Figure 5** for B6 (**Figure 5B**) and Fmr1 KO (**Figure 5A**). As described above, OPCs were quantified as cells that were positive for the sox-10 (magenta) marker without ASPA (yellow) staining (**Figures 5A,B** arrowheads, OPC label). Like OL counts, there was no significant difference between Fmr1 KO and B6 mice based on tonotopic location (medial, center, or lateral, **Figure 5D**), however, when data was summed for a total MNTB count, there was a significant increase in OPC number in Fmr1 KO mice compared to B6 (**Figure 5D**, top panel). Consistent with data from OLs, whole nucleus Q counts were consistent with raw counts such that there was no difference between genotypes based on tonotopy, but total estimates were significantly higher in Fmr1 KO animals compared to B6 ($p < 0.0014$). Together, these OPC and OL data show that changes in myelination observed through EM and CARS are not due to inherent decreases in numbers of mature or immature oligodendrocytes, however, there are changes in overall number of these cells compared to wildtype animals.

DISCUSSION

Mechanisms underlying auditory symptoms among patients with FXS are still poorly understood. The auditory brainstem is the first location along the ascending auditory pathway where sound information is processed, and it is likely key to identifying auditory difficulties since any changes at this level alter midbrain and cortical processing as well. Here we show decreased myelination in adult FXS mice (fiber diameter, axon diameter, thickness, and increased g-ratio) compared to wildtype in the region of the MNTB, which contains one of the largest and most heavily myelinated axons in the central nervous system. One possible mechanism for impaired myelination is changes in oligodendrocyte number and type. We show an increase in mature and precursor oligodendrocytes, suggesting perhaps an overcompensation for earlier reductions in OL numbers that others have shown (Pacey et al., 2013). Results from these measures suggest impaired myelination in the MNTB of Fmr1 KO mice, which could be a

potential mechanism for auditory brainstem physiological and behavioral phenotypes.

While this is the first study to examine myelination changes in the auditory brainstem in FXS, previous work has shown reduced myelin sheath growth, decreases in total myelin (Doll et al., 2020), fewer myelinated and thinner axons, reductions in myelin basic protein, and development impacts of loss of FMRP (Pacey et al., 2013). FMRP is present in both precursor and mature oligodendrocytes (Wang et al., 2004; Giampetruzzi et al., 2013) and located subcellularly within myelin sheaths (Doll et al., 2020). However, recent work suggests that FMRP is not acting directly to regulate myelin basic protein (MBP) or other myelin protein transcripts, suggesting a diversity of roles for FMRP in myelin and oligodendrocyte dynamics (Giampetruzzi et al., 2013; Doll et al., 2020). Interestingly, inconsistent with the literature, and unexpectedly based on reductions in myelin characteristics measured here, we saw an increase in number of both OLs and OPCs. These differences can potentially be explained by an overcompensation for OL number reductions during development in this brain area, or an increased number of OLs that are not necessarily myelinating efficiently (fewer myelin sheaths per OL, non-myelinating OLs, etc.). In addition, it is possible that there are changes in proliferation ability of OPCs and OLs that were not measured here with only single time point measurements (Bu et al., 2004). Lastly, we did not measure a direct effect of FMRP on myelination, and amount of myelination in the MNTB also depends on neuronal activity (Sinclair, 2017), we cannot rule out the possibility of loss of FMRP having secondary effects besides an impact on OLs. What consequences these reductions in myelination have on conduction speed and physiological properties of the Calyx of Held in synaptic transmission are potential areas for future research.

G-ratio is a common measure for not just structural but potential physiological properties of myelin conduction along an axon. The g-ratios we measured for Fmr1 KO and B6 animals are similar to previous work (particularly for wildtype B6 animals) in the MNTB (Ford et al., 2015; Stange-Marten et al., 2017). Conduction velocity is an important factor in reliable processing of sound location information, resulting in the Calyx of Held which is highly specialized for speed and temporal fidelity. Increased g-ratio suggests slower conduction velocities and thinner myelin in Fmr1 KO mice (Rushton, 1951; Berman et al., 2019). However, note that there may be other factors which contribute to conduction velocity, for example Ranvier node and internode spacing, length, and diameter (Ford et al., 2015). In addition, while we did not see differences in myelin properties based on location or tonotopic area, it is possible that node dynamics vary more with tonotopic area while myelin thickness and diameter are less impacted based on tonotopy. However, note that such differences may also be explained by the relative importance of ILD versus ITD processing in different animal models. The previous work showing tonotopic differences was performed in gerbils, a rodent model suitable to study both low and high-frequency hearing. In that animal model, one important role for MNTB neurons is to provide inhibitory input to ITD processing (Ford et al., 2015). By contrast, mice are primarily high-frequency specialists. Consistent with their hearing spectrum and the fact that mice primarily use ILD cues, previous work has shown that there are no differences in myelination properties based on tonotopy (Stange-Marten et al., 2017), similar to what we see in this study.

In the current study we did not directly label Calyx of Held axons. Rather, we made measurements within the MNTB, measuring axons that were projecting in the coronal plane. Although axons projecting to MNTB are significantly larger than any other passing fibers and can be discerned by eye relatively easily and reliably, we cannot rule out the possibility that some measured axons may be "passing through." If our analysis did include some passing fibers, these would likely decrease the observed differences between wild type and mutant, such that the results presented in this study may be a lower limit of the true differences between wild type and mutant mouse model. This would also limit the interpretation of conclusions about tonotopic location of the measurements and whether indeed there are differences based on frequency-coding. Indeed, we have no way of discerning if the medial MNTB measurements included lateral axons passing through to the lateral MNTB. Note that CARS nor EM techniques inherently discriminate between different types of axons. However, based on the size of the axons that we measured in both EM and CARS, fibers are consistent with the expected size of Calyx axons in the same area (Ford et al., 2015; Sinclair, 2017; Stange-Marten et al., 2017), therefore we are confident that at least the large majority of the axons are Calyceal projections.

This is the first study to show myelination changes in the auditory brainstem sound localization pathway in FXS mice. Thinner and smaller diameter Calyx axons with increased g-ratio may underly sound localization and auditory hypersensitivity issues seen behaviorally in mice and described by patient's with FXS. Interestingly, we saw an increase in both mature and immature OLs suggesting that myelin deficits are not due simply to fewer myelinating OLs. Future studies elucidating when during development myelin deficits begin and how OLs develop from precursors into mature OLs in this brain area would be crucial to understanding the complete picture of auditory brainstem phenotypes in FXS.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The animal study was reviewed and approved by University of Colorado Anschutz Institutional Animal Care and Use Committee.

AUTHOR CONTRIBUTIONS

EM, SP, AL, and AK developed the ideas and methods. EM, SP, and AL collected the data for the manuscript. EM and AL created the statistical analyses and figures for the manuscript. All authors helped wrote and revised the manuscript.

FUNDING

Supported by NIH R01 DC 17924, R01 DC 18401 (Klug), and NIH 1R15HD105231-01, T32DC012280, and FRAXA (McCullagh). The CARS imaging was performed in the Advanced Light Microscopy Core part of the NeuroTechnology Center at the University of Colorado Anschutz Medical Campus supported in part by NIH P30 NS048154 and NIH P30 DK116073.

ACKNOWLEDGMENTS

We would like to thank Jennifer Bourne for EM imaging. We would like to thank Clara Bacmeister and Ethan Hughes (University of Colorado, Anschutz) for providing antibodies, staining protocols, and overall support of this project. We would also like to thank Marc Ford for his help with troubleshooting and advice on this project.

REFERENCES

Barak, B., Zhang, Z., Liu, Y., Nir, A., Trangle, S. S., Ennis, M., et al. (2019). Neuronal deletion of Gtf2i, associated with Williams syndrome, causes behavioral and myelin alterations rescuable by a remyelinating drug. *Nat. Neurosci.* 22, 700–708. doi: 10.1038/s41593-019-0380-9

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2014). Fitting linear mixed-effects models using lme4. *J. Stat. Softw. arXiv* 1406:v067.i01. doi: 10.18637/jss.v067.i01

Bear, M. F., Huber, K. M., and Warren, S. T. (2004). The mGluR theory of fragile X mental retardation. *Trends Neurosci.* 27, 370–377. doi: 10.1016/j.tins.2004.04.009

Berman, S., Filo, S., and Mezer, A. A. (2019). Modeling conduction delays in the corpus callosum using MRI-measured g-ratio. *Neuroimage* 195, 128–139. doi: 10.1016/j.neuroimage.2019.03.025

Bitto, E., Bingman, C. A., Wesenberg, G. E., McCoy, J. G., and Phillips, G. N. (2007). Structure of aspartoacylase, the brain enzyme impaired in Canavan disease. *Proc. Natl. Acad. Sci.* 104, 456–461. doi: 10.1073/pnas.0607817104

Brown, M. R., Kronengold, J., Gazula, V.-R., Chen, Y., Strumbos, J. G., Sigworth, F. J., et al. (2010). Fragile X mental retardation protein controls gating of the sodium-activated potassium channel Slack. *Nat. Neurosci.* 13, 819–821. doi: 10.1038/nn.2563

Bu, J., Banki, A., Wu, Q., and Nishiyama, A. (2004). Increased NG2+ glial cell proliferation and oligodendrocyte generation in the hypomyelinating mutant shiverer. *Glia* 48, 51–63. doi: 10.1002/glia.20055

Chomiak, T., and Hu, B. (2009). What is the optimal value of the g-ratio for myelinated fibers in the rat CNS? A theoretical approach. *PLoS One* 4:e7754. doi: 10.1371/journal.pone.0007754

Curry, R. J., Peng, K., and Lu, Y. (2018). Neurotransmitter- and release-mode-specific modulation of inhibitory transmission by Group I metabotropic glutamate receptors in central auditory neurons of the mouse. *J. Neurosci.* 38, 8187–8199. doi: 10.1523/JNEUROSCI.0603-18.2018

Dahlhaus, R. (2018). Of men and mice: modeling the fragile X syndrome. *Front. Mol. Neurosci.* 11:41. doi: 10.3389/fnmol.2018.00041

Doll, C. A., Yergert, K. M., and Appel, B. H. (2020). The RNA binding protein fragile X mental retardation protein promotes myelin sheath growth. *Glia* 68, 495–508. doi: 10.1002/glia.23731

El-Hassar, L., Song, L., Tan, W. J. T., Large, C. H., Alvaro, G., Santos-Sacchi, J., et al. (2019). Modulators of Kv3 potassium channels rescue the auditory function of fragile X mice. *J. Neurosci.* 39, 4797–4813. doi: 10.1523/JNEUROSCI.0839-18.2019

Ford, M. C., Alexandrova, O., Cossell, L., Stange-Marten, A., Sinclair, J., Kopp-Scheinpflug, C., et al. (2015). Tuning of Ranvier node and internode properties in myelinated axons to adjust action potential timing. *Nat. Commun.* 6:8073. doi: 10.1038/ncomms9073

Gantois, I., Khoutorsky, A., Popic, J., Aguilar-Valles, A., Freemantle, E., Cao, R., et al. (2017). Metformin ameliorates core deficits in a mouse model of fragile X syndrome. *Nat. Med.* 23, 674–677. doi: 10.1038/nm.4335

Garcia-Pino, E., Gessele, N., and Koch, U. (2017). Enhanced excitatory connectivity and disturbed sound processing in the auditory brainstem of fragile X mice. *J. Neurosci.* 37, 7403–7419. doi: 10.1523/JNEUROSCI.2310-16.2017

Giampetruzzi, A., Carson, J. H., and Barbarese, E. (2013). FMRP and myelin protein expression in oligodendrocytes. *Mol. Cell. Neurosci.* 56, 333–341. doi: 10.1016/j.mcn.2013.07.009

Grothe, B., Pecka, M., and McAlpine, D. (2010). Mechanisms of sound localization in mammals. *Physiol. Rev.* 90, 983–1012. doi: 10.1152/physrev.00026.2009

Gundersen, H. J., Jensen, E. B., Kiêu, K., and Nielsen, J. (1999). The efficiency of systematic sampling in stereology–reconsidered. *J. Microsc.* 193, 199–211. doi: 10.1046/j.1365-2818.1999.00457.x

Hagerman, P. J., and Hagerman, P. J. (2008). The fragile X prevalence paradox. *J. Med. Genet.* 45, 498–499. doi: 10.1136/jmg.2008.059055

Hagerman, R. J., Berry-Kravis, E., Kaufmann, W. E., Ono, M. Y., Tartaglia, N., Lachiewicz, A., et al. (2009). Advances in the treatment of fragile X syndrome. *Pediatrics* 123, 378–390. doi: 10.1542/peds.2008-0317

Imamura, O., Arai, M., Dateki, M., Oishi, K., and Takishima, K. (2020). Donepezil-induced oligodendrocyte differentiation is mediated through estrogen receptors. *J. Neurochem.* 155, 494–507. doi: 10.1111/jnc.14927

Jeon, S. J., Ryu, J. H., and Bahn, G. H. (2017). Altered translational control of fragile X mental retardation protein on myelin proteins in neuropsychiatric disorders. *Biomol. Ther. (Seoul)* 25, 231–238. doi: 10.4062/biomolther.2016.042

Kim, H., Gibboni, R., Kirkhart, C., and Bao, S. (2013). Impaired critical period plasticity in primary auditory cortex of fragile X model mice. *J. Neurosci.* 33, 15686–15692. doi: 10.1523/JNEUROSCI.3246-12.2013

Kim, J. H., Renden, R., and von Gersdorff, H. (2013). Dysmyelination of auditory afferent axons increases the jitter of action potential timing during high-frequency firing. *J. Neurosci.* 33, 9402–9407. doi: 10.1523/JNEUROSCI.3389-12.2013

Kulesza, R. J. Jr., Lukose, R., and Stevens, L. V. (2011). Malformation of the human superior olive in autistic spectrum disorders. *Brain Res.* 1367, 360–371. doi: 10.1016/j.brainres.2010.10.015

Kulesza, R. J., and Mangunay, K. (2008). Morphological features of the medial superior olive in autism. *Brain Res.* 1200, 132–137. doi: 10.1016/j.brainres.2008.01.009

Larson, V. A., Mironova, Y., Vanderpool, K. G., Waisman, A., Rash, J. E., Agarwal, A., et al. (2018). Oligodendrocytes control potassium accumulation in white matter and seizure susceptibility. *eLife* 7:e34829. doi: 10.7554/eLife.34829

Lenth, R. (2019). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. Available Online at: CRAN.R-project.org/package=emmeans.

Lu, Y. (2019). Subtle differences in synaptic transmission in medial nucleus of trapezoid body neurons between wild-type and Fmr1 knockout mice. *Brain Res.* 1717, 95–103. doi: 10.1016/j.brainres.2019.04.006

McCullagh, E. A., Poleg, S., Greene, N. T., Huntsman, M. M., Tollin, D. J., and Klug, A. (2020a). Characterization of auditory and binaural spatial hearing in a fragile X syndrome mouse model. *eNeuro* 7:ENEURO.300–ENEURO.319. doi: 10.1523/ENEURO.0300-19.2019

McCullagh, E. A., Rotschafer, S. E., Auerbach, B. D., Klug, A., Kaczmarek, L. K., Cramer, K. S., et al. (2020b). Mechanisms underlying auditory processing deficits in Fragile X syndrome. *FASEB J.* 34, 3501–3518. doi: 10.1096/fj.201902435R

McCullagh, E. A., Salcedo, E., Huntsman, M. M., and Klug, A. (2017). Tonotopic alterations in inhibitory input to the medial nucleus of the trapezoid body in a mouse model of Fragile X syndrome. *J. Compar. Neurol.* 262:375. doi: 10.1002/cne.24290

Orthmann-Murphy, J., Call, C. L., Molina-Castro, G. C., Hsieh, Y. C., Rasband, M. N., Calabresi, P. A., et al. (2020). Remyelination alters the pattern of myelin in the cerebral cortex. *eLife* 9:e56621. doi: 10.7554/eLife.56621

Osterweil, E. K., Chuang, S.-C., Chubykin, A. A., Sidorov, M., Bianchi, R., Wong, R. K. S., et al. (2013). Lovastatin corrects excess protein synthesis and prevents epileptogenesis in a mouse model of fragile X syndrome. *Neuron* 77, 243–250. doi: 10.1016/j.neuron.2012.01.034

Pacey, L. K. K., Xuan, I. C. Y., Guan, S., Sussman, D., Henkelman, R. M., Chen, Y., et al. (2013). Delayed myelination in a mouse model of fragile X syndrome. *Hum. Mol. Genet.* 22, 3920–3930. doi: 10.1093/hmg/ddt246

Phan, B. N., Bohlen, J. F., Davis, B. A., Ye, Z., Chen, H.-Y., Mayfield, B., et al. (2020). A myelin-related transcriptomic profile is shared by Pitt–Hopkins syndrome models and human autism spectrum disorder. *Nat. Neurosci.* 23, 375–385. doi: 10.1038/s41593-019-0578-x

R Core Team (2013). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rajaratnam, A., Shergill, J., Salcedo-Arellano, M., Saldarriaga, W., Duan, X., and Hagerman, R. (2017). Fragile X syndrome and fragile X-associated disorders. *F1000Research* 6:2112. doi: 10.12688/f1000research.11885.1

Rotschafer, S. E., and Cramer, K. S. (2017). Developmental emergence of phenotypes in the auditory brainstem nuclei of fmr1knockout mice. *eNeuro* 4, 1–21. doi: 10.1523/ENEURO.0264-17.2017

Rotschafer, S. E., Marshak, S., and Cramer, K. S. (2015). Deletion of Fmr1 alters function and synaptic inputs in the auditory brainstem. *PLoS One* 10:117266. doi: 10.1371/journal.pone.0117266

Rushton, W. A. H. (1951). A theory of the effects of fibre size in medullated nerve. *J. Physiol.* 115, 101–122.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682. doi: 10.1038/nmeth.2019

Schmitz, C., and Hof, P. R. (2005). Design-based stereology in neuroscience. *Neuroscience* 130, 813–831. doi: 10.1016/j.neuroscience.2004.08.050

Sinclair, J. L. (2017). Sound-evoked activity influences myelination of brainstem axons in the trapezoid body. *J. Neurosci.* 37, 8239–8255. doi: 10.1523/JNEUROSCI.3728-16.2017

Stange-Marten, A., Nabel, A. L., Sinclair, J. L., Fischl, M., Alexandrova, O., Wohlfrom, H., et al. (2017). Input timing for spatial processing is precisely tuned via constant synaptic delays and myelination patterns in the auditory brainstem. *Proc. Natl. Acad. Sci. U S A* 114, E4851–E4858. doi: 10.1073/pnas.1702290114

Strumbos, J. G., Brown, M. R., Kronengold, J., Polley, D. B., and Kaczmarek, L. K. (2010). Fragile X mental retardation protein is required for rapid experience-dependent regulation of the potassium channel Kv3.1b. *Journal of Neuroscience* 30, 10263–10271. doi: 10.1523/JNEUROSCI.1125-10.2010

Tian, Y., Yang, C., Shang, S., Cai, Y., Deng, X., Zhang, J., et al. (2017). Loss of FMRP impaired hippocampal long-term plasticity and spatial learning in rats. *Front. Mol. Neurosci.* 10:269. doi: 10.3389/fnmol.2017.00269

Till, S. M., Asiminas, A., Jackson, A. D., Katsanevaki, D., Barnes, S. A., Osterweil, E. K., et al. (2015). Conserved hippocampal cellular pathophysiology but distinct behavioural deficits in a new rat model of FXS. *Hum. Mol. Genet.* 24, 5977–5984. doi: 10.1093/hmg/ddv299

Wang, H., Fu, Y., Zickmund, P., Shi, R., and Cheng, J.-X. (2005). Coherent Anti-Stokes Raman Scattering Imaging of Axonal Myelin in Live Spinal Tissues. *Biophys. J.* 89, 581–591. doi: 10.1529/biophysj.105.061911

Wang, H., Ku, L., Osterhout, D. J., Li, W., Ahmadian, A., Liang, Z., et al. (2004). Developmentally-programmed FMRP expression in oligodendrocytes: a potential role of FMRP in regulating translation in oligodendroglia progenitors. *Hum. Mol. Genet.* 13, 79–89. doi: 10.1093/hmg/ddh009

Wang, Y., Sakano, H., Beebe, K., Brown, M. R., de Laat, R., Bothwell, M., et al. (2014). Intense and specialized dendritic localization of the fragile X mental retardation protein in binaural brainstem neurons: A comparative study in the alligator, chicken, gerbil, and human: FMRP localization in NL/MSO dendrites. *J. Compar. Neurol.* 522, 2107–2128. doi: 10.1002/cne.23520

Weatherstone, J. H., Kopp-Scheinpflug, C., Pilati, N., Wang, Y., Forsythe, I. D., Rubel, E. W., et al. (2016). Maintenance of neuronal size gradient in MNTB requires sound-evoked activity. *J. Neurophysiol.* 117, 756–766. doi: 10.1152/jn.00528.2016

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.

X Consorthium, Bakker, C. E., Verheij, C., Willemsen, R., van der Helm, R., Oerlemans, F., et al. (1994). Fmr1 knockout mice: A model to study fragile X mental retardation. *Cell* 78, 23–33. doi: 10.1016/0092-8674(94)90569-X

Zorio, D. A. R., Jackson, C. M., Liu, Y., Rubel, E. W., and Wang, Y. (2017). Cellular distribution of the fragile X mental retardation protein in the mouse brain: Fmrp distribution in the mouse brain. *J. Compar. Neurol.* 525, 818–849. doi: 10.1002/cne.24100

Zuo, H., Wood, W. M., Sherafat, A., Hill, R. A., Lu, Q. R., and Nishiyama, A. (2018). Age-dependent decline in fate switch from NG2 cells to astrocytes after Olig2 deletion. *J. Neurosci.* 38, 2359–2371. doi: 10.1523/JNEUROSCI.0712-17.2018

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

REVIEW
published: 10 February 2022
doi: 10.3389/fnins.2022.799787

Hearing in Complex Environments: Auditory Gain Control, Attention, and Hearing Loss

Benjamin D. Auerbach[1,2]* *and Howard J. Gritton*[2,3,4]

[1] Department of Molecular and Integrative Physiology, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL, United States, [2] Neuroscience Program, University of Illinois at Urbana-Champaign, Urbana, IL, United States, [3] Department of Comparative Biosciences, University of Illinois at Urbana-Champaign, Urbana, IL, United States, [4] Department of Bioengineering, University of Illinois at Urbana-Champaign, Urbana, IL, United States

Listening in noisy or complex sound environments is difficult for individuals with normal hearing and can be a debilitating impairment for those with hearing loss. Extracting meaningful information from a complex acoustic environment requires the ability to accurately encode specific sound features under highly variable listening conditions and segregate distinct sound streams from multiple overlapping sources. The auditory system employs a variety of mechanisms to achieve this auditory scene analysis. First, neurons across levels of the auditory system exhibit compensatory adaptations to their gain and dynamic range in response to prevailing sound stimulus statistics in the environment. These adaptations allow for robust representations of sound features that are to a large degree invariant to the level of background noise. Second, listeners can selectively attend to a desired sound target in an environment with multiple sound sources. This selective auditory attention is another form of sensory gain control, enhancing the representation of an attended sound source while suppressing responses to unattended sounds. This review will examine both "bottom-up" gain alterations in response to changes in environmental sound statistics as well as "top-down" mechanisms that allow for selective extraction of specific sound features in a complex auditory scene. Finally, we will discuss how hearing loss interacts with these gain control mechanisms, and the adaptive and/or maladaptive perceptual consequences of this plasticity.

Keywords: adaptation, gain control, attention, auditory scene analysis, cocktail party problem, hearing loss

OPEN ACCESS

Edited by:
Dan Tollin,
University of Colorado, United States

Reviewed by:
Matthew James McGinley,
Baylor College of Medicine, United States
Frederick Jerome Gallun,
Oregon Health & Science University, United States

***Correspondence:**
Benjamin D. Auerbach
bda5@illinois.edu

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience

Received: *22 October 2021*
Accepted: *18 January 2022*
Published: *10 February 2022*

Citation:
Auerbach BD and Gritton HJ (2022) Hearing in Complex Environments: Auditory Gain Control, Attention, and Hearing Loss. Front. Neurosci. 16:799787. doi: 10.3389/fnins.2022.799787

INTRODUCTION

Auditory scene analysis— the ability to segregate specific sound features from multiple overlapping sources— is essential for extracting meaningful information from a complex sound environment (Bregman, 1990). The classic example of this problem is the cocktail party effect, where a listener can selectively focus on one specific speaker while filtering out a range of other stimuli (Cherry, 1953; Bregman, 2008). While the cocktail party problem represents a particularly challenging situation for the auditory system, as both the target and background sounds are comprised of similar acoustic features, most behaviorally-relevant sounds (such as a person talking) occur against a background

of everyday noise (e.g., traffic noise, a loud TV, etc.). Thus, adapting to noisy environments is a fundamental feature of the auditory system important for a range of listening conditions (Willmore et al., 2014; King and Walker, 2020). Understanding how the auditory system adapts to complex sound environments has important clinical implication as well, as individuals with age-related hearing loss or other hearing impairments often have great difficulties listening in noise, even when cochlear amplification is accounted for (Johannesen et al., 2016). How the auditory system solves the problem of auditory scene analysis remains incompletely understood.

Like humans, many animals— such as birds (Hulse et al., 1997), frogs (Endepols et al., 2003), and other mammals (Ma et al., 2010; Chapuis and Chadderton, 2018; Noda and Takahashi, 2019)— are capable of listening to a single sound source in a mixture of sources. Here we will discuss recent evidence from animal and human literature regarding the neurophysiological mechanisms for auditory scene analysis and hearing in complex environments. In particular, we will focus on gain control mechanisms— adjustments to the slope and dynamic range of neural input–output (I/O) relationships— that allow neurons to actively regulate their response sensitivity to the current environmental or behavioral demands (Robinson and McAlpine, 2009; Ferguson and Cardin, 2020). First, we will discuss how the auditory system adapts its response properties to changes in the overall distribution of incoming stimulus features. This bottom-up adaptation to stimulus statistics allows for extraction and invariant representation of key auditory features used to segregate sound sources in complex and continually changing acoustic environments. Next, we will discuss top-down contextual and attentional gain control mechanisms that can highlight behaviorally relevant sound information while selectively filtering distracting sources, even with overlapping acoustic features. Finally, we will examine how the central auditory system adapts to cochlear hearing loss and how this compensatory plasticity can have both adaptive and maladaptive consequences for sound perception and listening in complex auditory environments.

BOTTOM-UP ADAPTATION TO SOUND STATISTICS

Most natural sounds, including human speech, are characterized by dynamic changes in acoustic energy across spectral and temporal domains (Davenport, 1952; Singh and Theunissen, 2003; Santoro et al., 2014). In order to efficiently analyze an auditory scene and accurately represent the vast range of sounds encountered in the world, auditory neurons must be able to continually adapt their response properties to the prevailing acoustic environment. There is ample evidence that neural representations of sound are sensitive to statistical regularities in the acoustic environment (Winkler et al., 2009). For instance, many neurons across the auditory neuraxis exhibit stimulus-specific adaptation (SSA), in that they become less responsive to frequently occurring or repetitive stimuli but retain their sensitivity to rare stimuli, allowing for an intrinsic capacity to selectively encode unpredictable or novel sounds (Ulanovsky et al., 2003; Nelken, 2014). In addition to adapting to their own stimulus history, auditory neurons can also modify their response properties to match the statistics of the entire distribution of sounds encountered in the environment. Auditory neurons adapt their dynamic range and gain in response to a variety of stimulus statistics (**Figure 1**), including: mean sound level (Dean et al., 2005; Wen et al., 2009; Barbour, 2011), sound level variance or contrast (Nagel and Doupe, 2006; Rabinowitz et al., 2011; Willmore et al., 2014), interaural sound cues (Dahmen et al., 2010; Stange et al., 2013), and spectral-temporal correlations (Kvale and Schreiner, 2004; Natan et al., 2016; Homma et al., 2020). In this manner, neuronal responses are continuously rescaled to match dynamically changing sound conditions while maintaining overall firing rates across stimuli with different statistics. This adaptation to sound statistics enables auditory neurons to efficiently encode a wide-range of stimulus features under highly variable conditions and may be an effective mechanism for generating relatively invariant sound representations that are robust to the presence of background noise. Below we will discuss evidence for different forms of stimulus statistic adaptation as well as our current understanding of the neurophysiological mechanism and perceptual consequences of these adaptations.

Dynamic Range Adaptation

Natural acoustic scenes are characterized by stimuli that can vary over a wide range of sound levels, roughly 10–12 orders of magnitude (Baccus, 2006; Robinson and McAlpine, 2009). The auditory system maintains a remarkable sensitivity to small differences in sound level over this enormous range of intensities despite the relatively restricted dynamic range of individual auditory neurons, typically 30–50 dB SPL (**Figure 1A**) (Sachs and Abbas, 1974; Viemeister, 1988). This so-called dynamic range problem is compounded by noisy environments that act to increase the steady-state firing rate of auditory neurons, thereby limiting their dynamic range even further (**Figure 1A**) (Costalupes et al., 1984; Young and Barta, 1986). One potential solution to the dynamic range problem is to have distinct subsets of auditory neurons with different thresholds and dynamic ranges, such that combining or stitching together individual response functions would allow for representation of intensities across the full range of hearing at the population level (Barbour, 2011). Dynamic range stitching is observed to some degree at the level of the auditory nerve (AN), where fibers can be classified into at least three distinct subsets based on their response threshold and spontaneous firing rates (SR) (Evans, 1972; Sachs and Abbas, 1974; Liberman, 1978). Low-SR fibers have high thresholds and large dynamic ranges, medium-SR fiber have intermediate thresholds and dynamic range, while high-SR fibers have low thresholds and narrow dynamic range. Thus, while individual AN fibers have a restricted dynamic range, their sensitivity is distributed across a range of intensities. Moreover, the high threshold and larger dynamic range of low-SR fibers make them better suited for encoding intensity at higher sound levels and more resistant to background noise, suggesting they may be important for hearing in a noisy

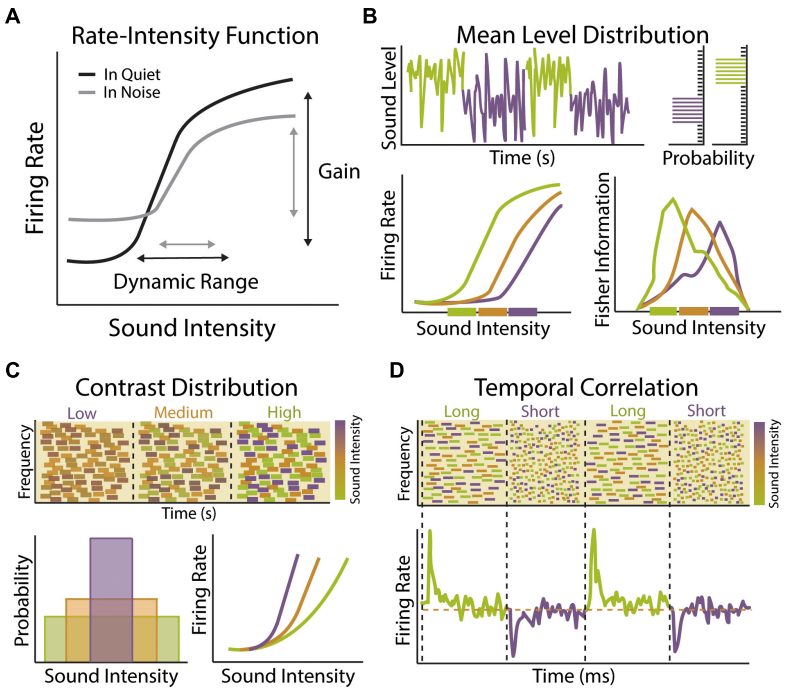

FIGURE 1 | Stimulus statistic adaptation in the auditory system. **(A)** Example rate-intensity function showing input–output relationship of sound intensity and auditory neuron firing rate in quiet (black) or in the presence of background noise (gray). Auditory neurons encode sound level information through changes in mean firing rate. The neuron's response gain is defined as the rate at which neuronal firing increases as a function of increasing sound level input. The dynamic range of a neuron is defined as the range of stimulus values encoded by a neuron though changes in its firing rate. A majority of auditory neurons exhibit low thresholds with firing rates that saturate to low or moderate sound levels, limiting their dynamic range. Both the gain and dynamic range of auditory neurons become compressed in the presence of background noise, which can be compensated for in part by adaptation to sound level statistics. **(B)** Example of dynamic range adaptation in a central auditory neuron or auditory nerve (AN) fiber. Top: Switching stimulus with high probability density region (HPR) at low (green) or high (purple) sound intensity levels. Within each environment, the range of intensities is drawn from a statistically defined high probability region confined to a narrow range of intensity while the remaining stimuli are drawn from a broader range of intensities outside the HPR (top right). Bottom: Auditory neurons adapt their threshold and dynamic range so that they are most sensitive to sound intensities within in the HPR, indicated by thick colored bars on *x*-axis (bottom left). Fisher information measure of coding accuracy for rate-intensity functions indicates that dynamic range adaptation acts to improve accuracy of sound intensity coding for sound levels most likely to be encountered in current environment (bottom right). Schematized data adapted from Dean et al. (2005). **(C)** Example of contrast gain control in a central auditory neuron. Top: Spectrogram of dynamic random chord (DRC) stimuli with low, medium or high contrast in sound intensity levels. *Bottom:* Theses spectrotemporally complex stimuli have same mean sound level but different sound level variance (bottom left). Neuronal gain shifts with changes in contrast (bottom right). Response gain is steep in low contrast environments (purple), allowing neurons to be sensitive to small changes in sound intensity, but becomes progressively shallower in medium (orange) and low (green) contrast environments, allowing neurons to maintain sensitivity to a larger range of intensities. Schematized data adapted from Rabinowitz et al. (2011). **(D)** Example of gain adaptation to changes in temporal correlation of sounds in central auditory neuron. Top: A series of DRCs with dynamically changing temporal correlation structure. Bottom: Auditory neurons respond to transition in temporal correlation structure by brief increases or decreases in firing rate, followed by steady state firing rates (orange dashed line). These transient changes in firing rate are indicative of adaptation to statistical change in stimulus, indicating that auditory neurons can adapt to the temporal dynamic range of the inputs to preserve encoding efficiency under varying statistical constraints without changing overall activity levels. Schematized data adapted from Natan et al. (2016).

environment (Costalupes et al., 1984). Dynamic range stitching is also observed in the central auditory system, with a subset of central auditory neurons exhibiting non-monotonic response functions that respond best to a particular sound level rather than exhibiting a constant increase in firing rate with increasing sound intensity (Suga and Manabe, 1982; Phillips et al., 1994; Sadagopan and Wang, 2008). However, the vast majority of auditory neurons have thresholds and dynamic ranges that are still heavily skewed toward lower sound intensities and it is therefore unlikely that dynamic range stitching can fully account for the maintenance of consistent sound level sensitivity across the entire range of hearing (Sachs and Abbas, 1974; Watkins and Barbour, 2011). It is becoming increasingly clear that a major mechanism for solving the dynamic range problem is that individual auditory neurons can dynamically adapt their threshold and dynamic range to compensate for changes in mean stimulus level, thereby maintaining maximum sensitivity over the most commonly encountered sound levels in the prevailing acoustic environment (**Figure 1B**) (Dean et al., 2005).

Dynamic range adaptation to mean sound level is observed across multiple levels of the auditory system, most notably at the AN (Wen et al., 2009), inferior colliculus (IC) (Dean et al., 2005, 2008), and auditory cortex (ACx) (Watkins and Barbour, 2008). When sounds are drawn from a distribution with a high probability of loud sounds, such that the mean sound level is high, neurons shift their dynamic range upward, increasing their

sensitivity to louder sounds (**Figure 1B**). Sound-level adaptation is thus compensatory, so that neuronal responses are relatively invariant to changes in background level. The degree of dynamic range adaptation observed in the IC is higher than at the level of the AN, suggesting that adaptation in the auditory midbrain is only partially accounted for by changes occurring in the periphery. This suggest that additional adaptation is occurring within the IC and potentially in the auditory brainstem as well. Indeed, neurons in the ventral cochlear nucleus (VCN) of the brainstem have been shown to adapt their threshold and dynamic range to the presence of background noise (May and Sachs, 1992). However, these studies employed a stationary background noise and experiments using the same dynamic switching stimulus used to measure mean level adaptation in the AN and IC have not been performed at the level of the brainstem. Similar magnitudes of adaptation are seen in subcortical and cortical neurons, suggesting that many features of mean level adaptation in cortex are inherited from lower levels of the auditory system. However, there are unique features of adaptation in the ACx, such as a subset of cortical neurons with non-monotonic rate-level functions that do not undergo adaptive recoding and are thus maladapted to high level sounds but preserve coding accuracy for quiet sounds even in the context of high sound level environments (Watkins and Barbour, 2008).

Contrast Gain Control

The auditory system is not only exposed to a wide range of sound levels, but any acoustic scene may be comprised of a relatively large or restricted subset of intensities across this spectrum. Thus, in addition to adapting to mean sound level, auditory neurons must be able to modulate their response properties to changes in the variance or contrast of sound levels present in the environment (**Figure 1C**). Contrast-invariant tuning is one of the most well-characterized examples of gain modulation observed across sensory systems (Finn et al., 2007; Olsen and Wilson, 2008; Rabinowitz et al., 2011) and contrast gain control in the auditory system is thought to be an important physiological mechanisms for encoding sound stimuli in noisy background conditions (Willmore et al., 2014). When an auditory neuron is exposed to a wide range of sound intensities, such that the contrast of the input is high, the gain of that neuron is low (**Figure 1C**). In this manner, the neuron has a broader dynamic range that is relatively insensitive to changes in sound level. When the contrast of the input is low, the gain of the neuron increases, making it more sensitive to small changes in intensity. Thus, like mean level adaptations, contrast gain control is compensatory, allowing neurons to adjust their gain in a manner that allows for representations that are relatively invariant to the level of background noise. Such adaptation allows for sounds that are structurally similar but with different contrast levels to be represented in a similar manner. It should also be noted that auditory neurons adjust their response properties to higher-order stimulus statistics like skewness or kurtosis as well (Kvale and Schreiner, 2004). Studies in ferrets have found that contrast gain control is more complete in the ACx compared to subcortical stations (Rabinowitz et al., 2013). However, more recent work in mice has found similar levels of contrast gain adaptation in the ACx, auditory thalamus (medial geniculate body; MGB), and IC (Lohse et al., 2020). Notably, the authors did find that adaptation time constants become longer at ascending levels of the auditory system, resulting in progressively more stable representations. Thus, there may be progressive changes to contrast gain control along the ascending auditory pathway, similar to that observed with mean level adaptation.

The combined effect of dynamic range adaptation and contrast gain control is to minimize the influence of background noise on auditory feature encoding. Indeed, by the level of the ACx, adaptation to mean level and contrast enables speech sounds to be represented in a way that is robust to the presence of background noise (Rabinowitz et al., 2013; Mesgarani et al., 2014b). However, it is important to note that adaptation to other acoustic features beyond sound intensity is likely important for speech perception and auditory scene analysis as well. Spectral features are a fundamental component of communication signals in mammalian vocalizations (Suga et al., 1983; Kadia and Wang, 2003). Human speech is comprised of several harmonic features and the use of these features can be helpful for identifying a speaker in a complex environment (Ehret and Riecke, 2002). Speech also varies in its temporal profile, including elements of fast temporal modulation and slower changes associated with periodicity of the speech signal, and the temporal structure of human vocalizations plays a crucial role in speech comprehension (Rosen, 1992; Shannon et al., 1995; Hickok and Poeppel, 2007). Frequency-specific adaptations have been observed in the human ACx that depend on the spectral range of acoustic stimuli, suggesting that there are neural adjustments to spectral stimulus statistics of sound stimuli (Herrmann et al., 2014). ACx neurons also display gain adaptations to changes in the temporal properties of sound input, allowing them to maintain their dynamic range across a range of temporal correlations (Natan et al., 2016) and use non-linear sensitivity to temporal and spectral content for adaptation (**Figure 1D**) (Angeloni and Geffen, 2018). Thus, in addition to sound level adaptations, neuronal adjustments to spectral and temporal sounds properties are also likely important for neural representation of speech in different background conditions (Ding and Simon, 2013; Mesgarani et al., 2014b; Khalighinejad et al., 2019).

Adaptation in Sound Localization

In a natural environment, multiple sound sources often originate from different locations and being able to identify the spatial location of distinct sound sources is a key component to auditory scene analysis. Binaural cues, such as interaural time (ITD) and level (ILD) differences, are essential for localizing sounds in space. The medial superior olive (MSO) and lateral superior olive (LSO) of the auditory brainstem are the initial sites of ITD and ILD processing in the mammalian auditory system, respectively. These brainstem nuclei contain coincidence detecting neurons that encode ITD and ILD differences by comparing the timing of converging inputs from the ipsilateral and contralateral ear with submillisecond precision (Park et al., 2004; Grothe et al., 2010). Because of the degree of precision required for these computations, and the fact that accurate representation of

absolute stimulus values may be more important for sound-source localization than for other acoustic features like sound level, traditional models of sound localization have proposed that ITDs and ILDs are encoded via a fixed labeled-line mechanism resulting in a hard-wired place code or map of auditory space (Jeffress, 1948; Grothe and Koch, 2011). However, it has now been shown that sound localization cues and spatial perception are also subject to short-term adaption based on prior stimulus history (Phillips and Hall, 2005; Vigneault-MacLean et al., 2007; Dahmen et al., 2010; Stange et al., 2013). Indeed, coincidence-detector neurons in both the MSO and LSO exhibit dynamic gain adaptations driven by feedback loops in early parts of the binaural pathway that modulate their sensitivity to ITD and ILD differences in response to prior activity levels (Finlayson and Adam, 1997; Magnusson et al., 2008; Park et al., 2008; Stange et al., 2013; Lingner et al., 2018). Adaption to sound source localization is observed in downstream auditory structures as well. By presenting sound sequences in which ILDs rapidly fluctuate according to a Gaussian distribution, it was shown that IC neurons also adapt the dynamic range and gain of their ILD-rate functions to match the mean and variance of the stimulus distribution (Dahmen et al., 2010). Thus, adaptive gain control mechanisms can also modulate the population code for auditory space in a stimulus history dependent manner.

Perceptual Consequences of Stimulus Statistic Adaptation

The above studies demonstrate that the auditory system uses multiple adaptive coding strategies to most efficiently represent and extract features from the sound environment. However, elucidating the perceptual consequences of these adaptations is crucial for determining if and how they facilitate our ability to analyze an auditory scene. Several recent studies have found that perceptual adaptations to stimulus statistics in humans parallel neurophysiological adaptations in animal models using near identical paradigms (Dahmen et al., 2010; Stange et al., 2013; Lohse et al., 2020). For instance, there is a close correspondence between changes to the perceived laterality of a stimulus in humans and adaptations to ILD-rate functions in the IC of ferrets when both are presented with noise sequences with rapidly fluctuating ILDs (Dahmen et al., 2010). Likewise, acuity in an intensity discrimination task is rapidly adjusted with changes to sound contrast in humans and the strength of this perceptual contrast adaptation could be predicted from physiological contrast adaptation observed in mice (Lohse et al., 2020). Chronic *in vivo* recordings from the ACx of mice trained to detect a target sound in background noise shortly after a change in the background contrast have provided some of the first evidence that cortical gain modulation and sound detection behavior are modulated by contrast in a parallel manner in the same subjects (Angeloni et al., 2021). This study found that ACx activity is necessary for detection of targets in background noise and that inter-subject variability in the magnitude of contrast gain control observed in the ACx predicted behavioral performance. These findings provide evidence that adaptive coding in the ACx has direct implications on perceptual behavior. In contrast, single unit recordings from the IC of macaques performing a masked tone detection task found that, despite observing dynamic range adaptation in the IC of these animals, behavioral detection thresholds were not affected by this neuronal adaptation (Rocchi and Ramachandran, 2018). Likewise, MEG and EEG studies have found evidence for dynamic range adaptation in the ACx of humans but parallel behavioral studies found that perceptual sensitivity to sound level was actually affected in an opposite manner than predicted by dynamic range adaptation, with increased sensitivity to sound intensities in the low probability region of the intensity distribution rather than high (Simpson et al., 2014; Herrmann et al., 2020). Thus, while there is growing evidence that adaption to stimulus statistics does influence perception, more work is needed to determine how different forms of adaptation across levels of the auditory system contribute to sound perception and auditory scene analysis.

Many studies have now shown that auditory neurons adapt their response properties to a range of stimulus statistics and tremendous progress has been made in the neurophysiological characterization of these bottom-up adaptations. However, there are many open questions that remain to be addressed. For instance, while the above studies indicate that adaptive coding is gradually built along the auditory pathway, the relative contributions of different auditory structures remain incompletely understood. More studies utilizing simultaneous recordings from multiple auditory regions are needed to determine how different forms of adaptation emerge along the ascending auditory pathway. Indeed, a recent study using this approach has uncovered a previously underappreciated role for subcortical processing in contrast gain control (Lohse et al., 2020). Second, the underlying cellular and circuit mechanisms driving adaption to sound statistics need to be fully elucidated, as will be discussed in subsequent sections. This knowledge is essential for understanding the biophysical constraints on theses adaptive processes as well as for generating novel strategies for manipulating these processes to better investigate their contribution to auditory scene analysis. Finally, more studies performing neurophysiological recordings from actively behaving animals are necessary to directly assess the impact of bottom-up adaptions on perception (Angeloni et al., 2021). One difficulty in assessing the perceptual consequences of bottom-up adaptations is that, in most cases, measuring behavioral sensitivity to changes in stimulus statistics requires subjects to be engaged in a perceptual decision-making task. Task engagement itself will invoke a multitude of adaptive changes in the auditory system, as will be discussed in the next section.

TOP-DOWN CONTRIBUTIONS TO AUDITORY SCENE ANALYSIS

Bottom-up adaptations to the prevailing sound statistics enable the auditory system to more efficiently encode target sounds in complex or noisy environments, particularly when the statistics of foreground and background sounds are distinct (**Figure 1**). However, background sounds that share acoustic features or statistical properties that significantly overlap with the signals

of interest, such as is the case for the cocktail party problem, pose unique challenges for auditory scene analysis and additional mechanisms must exist to selectively extract specific sound sources from structurally similar background noise (King and Walker, 2020). Attention is a cognitive process by which organisms filter the most relevant behavioral information from their environment to enhance perception of one particular stimulus over another. Selective attention has been proposed to contribute to auditory scene analysis by acting as a form of sensory gain control, enhancing the representation of an attended sound source while suppressing responses to unattended sounds (Fritz et al., 2007; Kerlin et al., 2010; Zion Golumbic et al., 2013). This process can occur when the stimulus itself directs attention through enhanced salience, referred to as bottom-up or "pop-out" attention (Kayser et al., 2005), or can be endogenously generated through top-down or "task-modulated" processes. The focus of this section will be to discuss these top-down mechanisms and how selective attention contributes to gain modulation, feature selection, and stream separation in the auditory system, which work in concert to improve auditory scene analysis. First, we will discuss the growing body of evidence from animal studies showing that sensory encoding is fundamentally modulated by behavioral state. Then, we will discuss evidence that task-engagement, a proxy for attention in animal models, is associated with receptive field changes that act to maximize encoding of task-relevant information. Finally, we will discuss evidence from human studies showing that selective attention does indeed influence perception and listening performance in complex auditory environments.

Behavioral State and Attentional Modulation of Sensory Processing

Behavioral states have strong influences on neuronal responses associated with sensory processing (Bennett et al., 2014; Lee et al., 2014). Early studies that monitored pupil dilation as a proxy for arousal found that cortical neurons in sensory regions are strongly modulated by dilation onset and that neuronal firing rates correlate with the level of dilation even in the absence of sensory input (Iriki et al., 1996). These results suggest that internal state changes may influence how we process incoming sensory information. Indeed, performance on a tone-in-noise detection task has been shown to be highly state-dependent, with peak behavioral performance being associated with intermediate levels of arousal (McGinley et al., 2015a). These intermediate arousal states coincided with periods of stable hyperpolarization in auditory cortical neurons. Subsequent studies using optogenetic manipulations of inhibitory interneuron populations found that such hyperpolarized states increase encoding specificity by augmenting the threshold for responsivity of excitatory neurons while simultaneously narrowing the frequency tuning properties in principle cell populations (Hamilton et al., 2013; Aizenberg et al., 2015; Phillips and Hasenstaub, 2016). Selective attention also profoundly impacts behavioral sensitivity to sensory stimuli and is in fact operationally defined as an improvement in psychophysical performance for attended versus unattended stimuli (e.g., Carrasco, 2011). Neurophysiological studies have revealed that attentional effects on sensory processing are due at least in part to gain modulation that increases stimulus-evoked response size (Mitchell et al., 2007; Maunsell, 2015). Indeed, tone-evoked responses in the ACx are greater in animals performing a tone-detection task compared to passive listening conditions, indicative of attentional gain modulation (Francis et al., 2018). Taken together, these findings offer insight into how behavioral state changes influence sensory processing irrespective of the sound statistics being conveyed to the sensory system.

While behavioral state and attentional gain increases enhance the magnitude of sensory-evoked responses, it is important to note that a non-selective increase in neuronal activity is not necessarily beneficial to stimulus detection. Rather, attention appears to enhance feature encoding by modulating not only the magnitude of the sensory stimulus but also the spontaneous activity or "noise" of neural responses (Harris and Thiele, 2011). Background noise can include non-stimulus specific activity represented by highly correlated neurons that act to reduce the amount of information that can be encoded for a particular stimulus or through competing distractors in the stimulus field (Zohary et al., 1994; Fries et al., 2001; Moreno-Bote et al., 2014). Selective attention not only increases stimulus-evoked responses, but also reduces the effect of intrinsic background noise, thereby enhancing signal-to-noise ratios for sensory representations and decreasing trial-to-trial variability (Mitchell et al., 2009; Downer et al., 2017; Francis et al., 2018). Selective attention simultaneously reduces variability and noise correlations across populations of cortical neurons in large part by reducing low frequency firing rate correlations to produce a sparse and temporally reliable code (Mitchell et al., 2009; Francis et al., 2018). This appears to be the case for arousal-dependent changes in sound processing as well, as changes in pupil diameter produce bi-modal effects on spontaneous and sensory-evoked activity that improve signal-to-noise ratios of sound-evoked responses (McGinley et al., 2015b). Such findings argue that reduction in spontaneous neural activity is as critical to feature discrimination as gain modulated increases in firing rate. Indeed, attentional control associated with the act of behavioral engagement appears to enhance feature encoding by altering the spontaneous activity of cortical circuits prior to sensory processing. For example, the process of self-directed trial initiation decreases the rate of spontaneous activity in the ACx of rats performing a tone-detection task and optogenetic disruption of cortical activity before tone presentation acts to impair performance (Carcea et al., 2017). Thus, attention has the effect of both increasing responsiveness to a target stimulus while simultaneously reducing the influence of background activity and distracting inputs, which together act to stabilize sensory representations and promote feature tracking in complex environments.

Task-Dependent Modulation of Sound Feature Encoding

The above studies suggest that auditory responses rely not only on the external sounds reaching the ear, but also on the behavioral

context and internal state of the subject. While it is clear that attentional modulation of auditory neuron response properties can act to improve signal-to-noise ratios and the reliability of sensory encoding, does selective attention allow subjects to focus on specific sound features in a complex auditory environment? Ideally, a subject in a complex auditory scene could utilize the spectrotemporal content of relevant features to separate attended streams from background unattended streams to better isolate the target. Task engagement has indeed been shown to result in rapid adaptions to auditory neuron response properties in a manner that optimizes encoding of task-specific features. Combined measures of temporal and frequency sensitivity to sound stimuli can be measured by calculating the spectrotemporal receptive fields (STRFs) of cortical neurons. In a series of experiments where ferrets were trained to discriminate a tonal target in the presence of background noise stimuli that were comprised of TORCs (temporally orthogonal ripple combinations), it was demonstrated that the STRFs of ACx neurons dynamically adapt to the stimulus features, enhancing responses to the target frequency while reducing responses to the non-target spectral and temporal features (**Figures 2Ai,ii**) (Fritz et al., 2003, 2005a). These changes in receptive properties were rapidly and specifically modulated by task-engagement, with the STRFs returning to their original fields shortly after the behavioral task was over. Moreover, STRF changes were highly dependent upon the nature of the task and revealed task-specific signatures based on whether the animal was taxed with spectral or temporal feature discrimination (Fritz et al., 2007). Task reward structure also modulates attention-driven receptive field plasticity, with positive or negative reinforcement for the same target tone resulting in rapid and selective changes in cortical STRFs at the target frequency in equal magnitude but opposite direction (David et al., 2012). Thus, attention reshapes cortical tuning properties in manner that enhances the contrast between task relevant stimulus classes. Encoding of sound spatial location is also highly sensitivity to task demands. In a task where cats were trained to identify changes in sound source origin across trials, spatial selective tuning emerged within seconds of task onset and was mediated via suppression of tuning responsivity in least preferred spatial locations (**Figures 2Bi,ii**) (Lee and Middlebrooks, 2011). Thus, while spatial tuning is typically broad in auditory cortical neurons, selective attention can rapidly sharpen spatial tuning properties to improve localization of attended sound sources. These findings suggest that auditory selective attention can mediate short-term cortical plasticity to modulate spectrotemporal and spatial sound encoding, thereby improving perceptual performance in a task-specific manner.

While the above studies indicate that top-down attentional signals dynamically reshape receptive fields in the primary ACx in a task-specific manner, an important question that remains is the anatomical locus of these top-down signals. As sound information ascends through the auditory system, neurons preferentially encode more abstract sound entities or categorizations rather than detailed spectrotemporal features (Chechik and Nelken, 2012; Mesgarani et al., 2014a). This abstraction along cortical hierarchies is likely important for building invariant representations of foreground target sounds that are robust to different background sound conditions. Indeed, neurons in non-primary ACx exhibit greater invariance in encoding acoustically distorted communicative signals compared to neurons in primary ACx (Carruthers et al., 2013, 2015). Similarly, dual recordings from primary and secondary ACx in ferrets trained to detect streams of repeated noise samples embedded in a stream of random background samples found that stream-specific gain enhancement was stronger in secondary cortical areas compared to primary ACx (Saderi et al., 2020). Importantly, categorical sound representations in higher-order cortical regions are often behaviorally-gated, adaptively assuming different states or filter properties depending upon the demands of the ongoing task. In ferrets engaged in a tone-discrimination task, belt regions of ACx that typically reflect stimulus properties similar to primary ACx (**Figure 2Ci**) take on less faithful representation of the stimulus and more abstract properties that reflect components of the motivational properties of the behavioral task (**Figure 2Cii**) (Atiani et al., 2014). This effect of task engagement is seen to even greater degree in frontal cortical regions, where neurons that rarely responded to sound stimuli during passive listening selectively responded to target sounds during behavior (**Figure 2Ciii**) (Fritz et al., 2010). These task-specific responses in frontal cortical regions could in principal provide top-down signals that mediate receptive field changes in primary ACx based on task category expectations (Fritz et al., 2007). Consistent with this notion, changes in frontal cortex representations coincide with augmented inter-areal coherence between frontal cortex and regions of primary ACx that were most responsive to target sounds (Fritz et al., 2010). Likewise, pairing electrical stimulation of orbitofrontal cortex with sound stimuli in passively listening animals induced rapid changes in the frequency receptive fields of primary ACx neurons in a manner similar to the effects of task-engagement (Winkowski et al., 2013). These findings suggest that functional interactions between frontal and primary sensory areas can shape the flow of relevant auditory information during active listening and is thus likely to play an important role in auditory scene analysis.

Top-Down Modulation of Stimulus Statistic Adaption

Behavioral task engagement is not only associated with attention driven changes to sound feature encoding but has also been shown to directly influence the degree of bottom-up adaptation to statistical changes in sound. For instance, a recent study revealed that the magnitude of dynamic range adaptation in the IC of macaques was enhanced in animals actively engaged in a tone-in-noise detection task compared to when they were passively listening to the same stimuli (Rocchi and Ramachandran, 2020). Recordings of IC neurons in guinea pigs repeatedly exposed to a switching stimulus that alternates between loud and quiet environment found that auditory midbrain neurons adapt more rapidly with repeated exposure to a loud environment, a phenomenon termed meta-adaptation (Robinson et al., 2016). This meta-adaptation suggests that auditory scene analysis is not only influenced by the statistical properties of sound input but our prior knowledge of the sound environment.

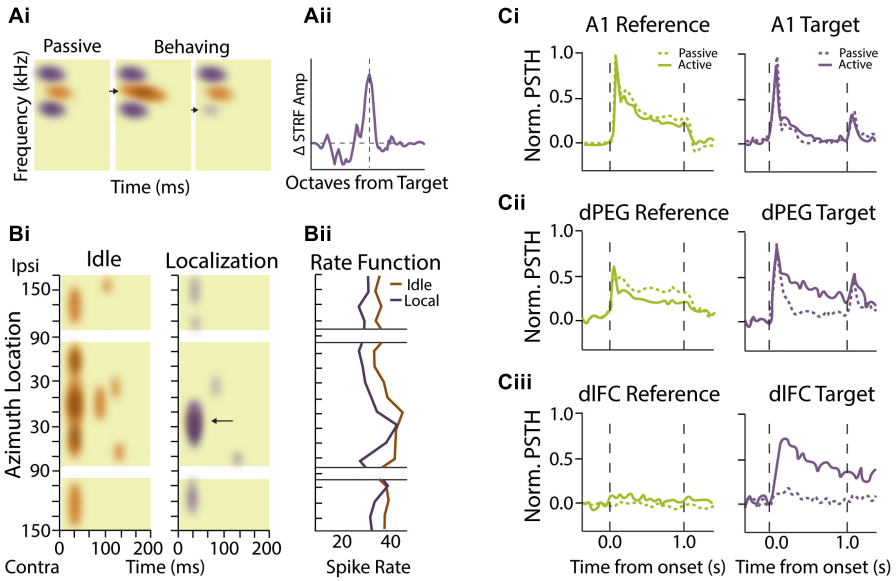

FIGURE 2 | Attentional and task-dependent modulation of sound feature encoding. **(Ai,ii)** Spectrotemporal receptive fields (STRFs) in auditory cortical neurons change based on task engagement and target frequency. **(Ai)** Example STRF showing enhanced sensitivity (orange) and sideband inhibition (purple) during passive presentation of broadband temporally orthogonal ripple combinations (TORC) stimuli (left). Performance of a tone detection task with peak STRFs near the target frequency (arrow) enhances the excitatory region in the STRF during behavior (middle). When tonal targets were presented with frequencies that coincided with inhibitory STRF, (arrow) the STRF showed local decreases or elimination of inhibitory sidebands (right). **(Aii)** Summarized data showing that STRF plasticity adaptation effects are most substantial when near the target frequency with facilitation occurring over ∼1 octave from the target stimulus. Schematized data adapted from Fritz et al. (2003). **(Bi,ii)** Spatial sensitivity modulated by task performance. **(Bi)** Heat maps demonstrating primary auditory cortex (A1) neural activity as a function of time (horizontal axis) and stimulus location (vertical axis) from a single behavioral session. This neuron shows burst activity at sound onset and is strongly responsive to probe trials originating from all locations during idle conditions (non-task performing condition). During the sound localization task where the cat is rewarded for discriminating changes in elevation, neural responses become more selective for probe trial origin, responding best to stimuli located between contralateral 10° and 50°. Arrow indicates increased specificity for this unit at the spatial localization. Colors indicate changes in mean intensity firing rates for the two conditions. **(Bii)** Rate functions in response to sound onset are shown to the right for the passive and sound location task conditions as a function of stimulus location. Schematized data adapted from Lee and Middlebrooks (2011). **(Ci,ii,iii)** Effects of task performance on auditory responsivity in auditory and frontal cortices. **(Ci)** Average behavior-dependent change in reference (green) and target (purple) responses in A1. Reference targets included TORC or narrowband white noise stimuli while targets consisted of pure tones. Dashed lines represent pre-task passive responses while solid lines represent task-engaged response. The average reference and target response as measured by normalized peri-stimulus time histograms (PSTH) amplitude were not significantly different between passive and behavior conditions. **(Cii)** Target and reference comparison for dorsal posterior ectosylvian gyrus (dPEG) of the ferret which is a belt region receiving A1 input. dPEG shows an average target response augmentation during task performing conditions. **(Ciii)** Target and reference PSTH comparison for dorsal lateral frontal cortex (dlFC), an executive region important for cue-directed behavior. dlFC neurons show almost no responsivity during passive conditions for either target or reference stimuli; however, they are strongly regulated by the target exclusively during behavior. Schematized data adapted from Atiani et al. (2014).

Interestingly, cortical inactivation via cryoloop cooling disrupted meta-adaptation in the IC, indicating the top-down nature of this phenomenon. Thus, adaptation to mean sound level is accelerated and more efficient when animals have been previously exposed to an environment or are engaged in an actively listening paradigm. Together, these attentional effects on spectral-temporal receptive properties, spatial tuning, and sound level adaptations are likely to aid in our ability to identify target sound sources in complex listening conditions.

Attentional Contributions to Auditory Scene Analysis in Humans

Animal studies have clearly demonstrated that attentional state and experience can influence auditory response properties and adaption to sound stimulus statistics. Parallel human studies have provided evidence that this attention-driven plasticity is indeed important for auditory scene analysis. In human ACx, selective attention enhances psychophysical performance through increases in neural gain (Kauramäki et al., 2007; Kerlin et al., 2010; Zion Golumbic et al., 2013). Increases in multiplicative gain in ascending pathways as well as enhancement of feature selectivity in secondary auditory cortices associated with "what" and "where" processing pathways appears to occur during auditory scene analysis (Ahveninen et al., 2006). The adoption of stimulus parameters from electrophysiological studies in rodents and ferrets have further revealed complex associations in auditory regions in human studies. For instance, rapid changes in the spectrotemporal response of recorded neurons in human ACx can occur in seconds, mimicking the effects seen in ferrets. Such changes also corresponded with improved perceptual performance in extraction of speech from a degraded stimulus (Holdgraf et al., 2016). In binaural task designs, selective auditory attention enhances the neural representation of relevant sound streams while reducing the neural representation of irrelevant sound streams in distracting

environments (Bidet-Caulet et al., 2007). For example, in a study with competing sound streams that were modulated at different frequencies, the auditory steady state response was modulated at the sound stream frequency in the direction of the attended stream within the precentral sulcus (Bharadwaj et al., 2014), a region important for visual spatial attention (Wardak et al., 2006). Imaging studies have further identified a dichotomy of processing where primary regions show enhanced sensitivity to temporal coherence and associate auditory regions show strong activation by sounds with reduced temporal correlations (Zatorre and Belin, 2001; Schönwiesner et al., 2005; Overath et al., 2008).

The level of noise invariance is also highly regulated by the directionality of the attended source in humans, suggesting that hierarchical cortical processing allows for spectrotemporal feature extraction that is strongly spatially modulated (Mesgarani and Chang, 2012; Schneider and Woolley, 2013). High-density EEG has revealed spatial speech stream segregation occurs during selective attention for an attended talker. Importantly, differences in alpha power (8–12 Hz) across hemispheres at parietal sites indicated the direction of auditory attention (Kerlin et al., 2010). Interestingly, analysis of high gamma (75–150 Hz) LFPs in the posterior temporal lobe reveals that reconstruction of the speech spectrograms from neural activity reflect the attended speaker alone despite being presented in the presence of a competing speaker. Importantly, on counterbalanced trials, the reconstructed spectrograms in the same region reflected the change to the new attended speaker, suggesting cortical representation of speech gives rise to the perceptual aspects relevant for the listener's intended goal (Mesgarani and Chang, 2012). In subsequent studies, Deng et al., found that directed attention cues occurring before the auditory discrimination task promoted supramodal alpha activity ipsilateral to the area of directed attention. Further, this relative ratio of ipsilateral/contralateral alpha activity shifted smoothly across hemispheres as the target source location was moved from the ipsilateral to the contralateral location (Deng et al., 2020). Such findings suggest that an ability to attend to localized sound statistics reflecting a relevant target are an important feature of auditory scene analysis, although not all listeners can do this with the same level of precision. For instance, a recent study in individuals with normal levels of hearing and speech understanding, found that reduced performance for non-speech auditory selective attention accounted for the greatest variation in individual task performance in a cocktail-party listening task (Oberfeld and Klöckner-Nowotny, 2016). This raises the intriguing possibility that listener performance in complex environments is largely a function of attentional capacity. These findings offer insight into the complex interactions between sound feature statistic adaptation and the role of cognitive capacity in attentional gain control during auditory scene analysis.

This section has highlighted our current understanding of behavioral state influences on auditory processing and the evidence for top-down regulation of feature encoding and adaptation in the auditory system. However, several questions remain to be answered. For instance, despite strong evidence for attentional gain control and task-dependent receptive field plasticity in the auditory system, it is still unclear if gain modulation can sufficiently account for the receptive field changes seen with task engagement (Otazu et al., 2009; Lopez Espejo et al., 2019). In addition, while frontal executive control regions are implicated in top-down modulation of auditory feature encoding through studies of coherence, the specific cortical regions involved— and how they are recruited to impart influence in primary sensory regions— remains unclear. Most prominently, little is known about the local and long-range circuit mechanisms and neurotransmitter systems that allow for such dynamic attentional adaptations. In the subsequent section we will discuss potential candidates, including distinct neuronal subtypes which confer specialization to cortical and subcortical circuits during sensory processing.

MECHANISMS FOR BOTTOM-UP AND TOP-DOWN ADAPTATIONS

The combination of bottom-up adaptations to the sound statistical environment and top-down modulation of receptive field properties greatly effects sound feature encoding in the auditory system and is likely to impact our ability to listen to target sound sources in a noisy environment. There is also evidence that these bottom-up and top-down adaptations directly interact, as task engagement can modulate how auditory neurons adapt to changes in incoming stimulus statistics. An important question, therefore, is if these bottom-up and top-down adaptations converge on common neuronal mechanisms. Does attention co-opt the circuits that mediate bottom-up gain adaptations, or do top-down and bottom-up gain control rely on independent mechanisms that can interact in complex ways? There are number of cellular and circuit mechanisms that may be used to implement gain adaptations to changes in incoming stimulus statistics and/or selective attention (**Figure 3**). This section will discuss our current understanding of how these changes are implemented within the auditory system and potential interactions between them.

Synaptic Mechanisms Contributing to Sound Stimulus Adaptations

Efficient information processing in neural circuits is dependent on tightly regulated interactions between excitatory and inhibitory neurons, which may or may not necessarily be balanced depending on conditions or behavioral state (Haider et al., 2006; Shew et al., 2011; Yizhar et al., 2011). Under conditions of tight excitatory-inhibitory balance (E/I balance), as sometimes seen in the ACx (Wehr and Zador, 2003; Zhang et al., 2003), synaptically driven fluctuations in membrane potential (V_m) can multiplicatively regulate tuned neural responses (Chance et al., 2002; Shu et al., 2003). Thus, noisy background synaptic input can exert gain control on neuronal output under different environmental conditions, providing a potential cellular mechanism for gain adaptations to stimulus statistics (Finn et al., 2007). However, *in vivo* whole cell recordings from ACx neurons during high and low contrast stimulation found that membrane conductance was not significantly modulated by stimulus

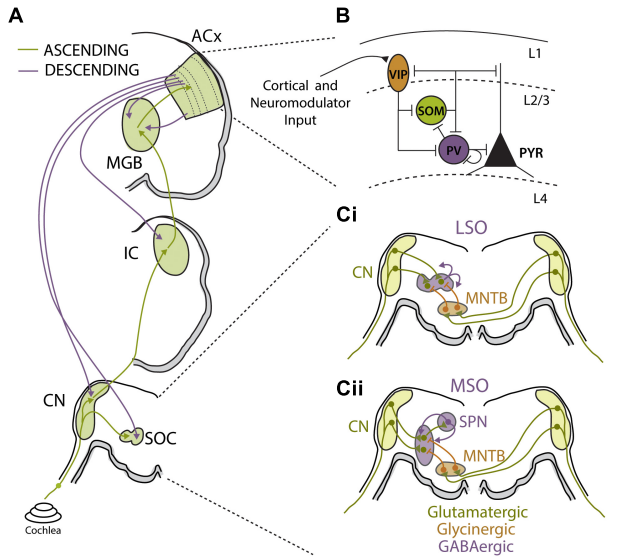

FIGURE 3 | Auditory circuit mechanisms for bottom-up and top-down adaptations. **(A)** Schematics of major auditory ascending (green) and descending (purple) pathways and associated auditory processing nuclei from the cochlea to auditory cortex (ACx). Major sound processing nuclei are highlighted in green, including the cochlear nucleus (CN), superior olive complex (SOC), inferior colliculus (IC), medial geniculate body (MGB), and ACx. Ascending pathways primarily terminate in layer 4 of ACx while corticofugal projections originate in layer 5b and layer 6 and can terminate at every level of the ascending pathway. **(B)** Summary figure showing known inhibitory relationships between parvalbumin (PV), somatostatin (SST), and vasoactive intestinal peptide (VIP)-positive neurons and their combined influences on excitatory cell populations. In this largely accepted model, locomotion, or top-down input, preferentially activates VIP cells, reducing SST cell output and releasing PV and excitatory cells from inhibition. PV activity can reduce background noise that improves signal-to-noise encoding of sensory stimuli by excitatory cell populations. Cortical layers are shown and separated by dashed lines. Figure was adapted from Pakan et al. (2016). **(Ci,ii)** Schematics of sound localization circuits of auditory brainstem used for processing interaural level **(Ci)** and timing **(Cii)** differences. **(Ci)** Principal neurons of lateral superior olive (LSO) receive excitatory glutamatergic inputs (green) from ipsilateral CN and glycinergic inhibition (orange) from ipsilateral medial nucleus for the trapezoid body (MNTB). MNTB receives excitatory input from contralateral CN. The interaction of ipsilateral excitation and contralaterally driven inhibition drives LSO firing in manner that can be used to calculate interaural level differences (ILD). LSO principal neurons release GABA (purple) in activity-dependent manner to asymmetrically modulate function of both glutamatergic and glycinergic inputs via activation of pre-synaptic $GABA_B$ receptors. **(Cii)** Principal neurons of the medial superior olive (MSO) receive excitatory glutamatergic inputs (green) from ipsilateral CN and contralaterally driven glycinergic inhibition (orange) from ipsilateral MNTB. MSO neurons also receive ipsilaterally driven inhibition from lateral nucleus of the trapezoid body (not shown). MSO neurons send excitatory projections to superior olivary nuclei (SPN), which in turn send feedback GABAergic projections (purple) to the MSO.

contrast and contrast-dependent V_m fluctuations could not account for contrast gain control in the ACx (Cooke et al., 2020). Short-term synaptic plasticity is another potential mechanism for gain modulation (Abbott et al., 1997), as short-term synaptic depression at thalamocortical synapses is believed to contribute to contrast gain control in the visual system (Carandini et al., 2002; Banitt et al., 2007). Synaptic depression is observed at synapses across levels of the auditory system (Nelson et al., 2009; Yang and Xu-Friedman, 2009; Blundon et al., 2011), including hair cell-ribbon synapses (Goutman, 2017). While this form of short-term plasticity has been proposed to play a role in SSA (Motanis et al., 2018), its role in stimulus statistic adaptation remains to be determined. Interestingly, STRF models that incorporate both synaptic depression and gain control show that there are additive effects of these properties on the robustness of cortical neuron STRFs to background noise, suggesting that these are complementary processes that may not reflect the same underlying mechanism (Mesgarani et al., 2014b; Pennington and David, 2020).

Modulation of presynaptic synaptic transmission has been shown to play an integral role for adaptive gain control in auditory brainstem circuits that use precise comparison of excitatory and inhibitory synaptic inputs to compute ITDs and ILDs for sound localization (**Figures 3Ci,ii**) (Finlayson and Caspary, 1989; Park et al., 1996). Interestingly, both excitatory and inhibitory pre-synaptic terminals in the MSO and LSO are dynamically adjusted by GABA via activation of pre-synaptic $GABA_B$ receptors that modulate neurotransmitter release (Magnusson et al., 2008; Grothe and Koch, 2011; Stange et al., 2013). In the LSO, GABA is released from principle cells in an activity-dependent manner and bind to pre-synaptic $GABA_B$ receptors to mediated gain adaptation on the time scale of seconds (**Figure 3Ci**) (Magnusson et al., 2008). Retrograde activation of presynaptic $GABA_B$ receptors has asymmetric effects on excitatory and inhibitory synaptic terminals in the LSO, suppressing glutamatergic transmission more strongly than glycinergic transmission. The net effect is to decrease excitability of LSO neurons, resulting in a shift in the dynamic range of ILD functions and narrowing the binaural receptive field of LSO neurons so that ipsilateral stimuli are preferentially encoded and perceived as more intense (Magnusson et al., 2008). MSO neurons have also been shown to modulate their sensitivity to ITD through a $GABA_B$ feedback mechanism from the superior periolivary nucleus (SPN), which also receives collateral inputs from the MSO (**Figure 3Cii**). This di-synaptic feedback loop activates pre-synaptic $GABA_B$ receptors, causing a slow-acting and long-lasting decrease in MSO neuronal activity in a manner proportional to their prior activity levels. This activity-dependent rate adaptation does not directly alter preferred ITDs in MSO neurons, but results in a form output normalization gain modulation that produces asymmetry in hemispheric population code for sound space (Stange et al., 2013). In this manner, strongly lateralized sound sources induce unequal adaptation preferentially in the contralateral hemisphere, thereby shifting perceived location of a subsequently presented sound source. Parallel psychophysical experiments found that the same paradigm used to evoke $GABA_B$ receptor-mediated adaptation in gerbils caused predictable shifts in sound localization percepts in humans (Phillips and Hall, 2005; Stange et al., 2013). Thus, dynamic adjustments to the balance between excitation and inhibition in MSO and LSO neurons via regulation of presynaptic transmitter release are used to modulate sound localization cues and spatial perception.

Cortical Circuit Mechanisms Contributing to Bottom-Up and Top-Down Adaptations

Local inhibitory interneuron networks are prominent regulators of neuronal gain, particularly within cortical circuits (Katzner et al., 2011; Ferguson and Cardin, 2020), and E/I balance is thought to be essential for proper regulation of sensory encoding (Wehr and Zador, 2003; Isaacson and Scanziani, 2011). Indeed, there is evidence that alterations to E/I balance underlie rapid receptive field changes seen in cortical neurons following experimental conditioning or learning (Carcea and Froemke, 2013; Froemke et al., 2013). E/I balance is also highly state-dependent (Haider et al., 2013; Zhou et al., 2014) and top-down modulation of E/I balance may contribute to the observed behavioral state and attentional gain modulation of sensory processing discussed in the previous section (Harris and Thiele, 2011). However, the diversity of inhibitory interneurons subtypes, in terms of physiological properties and computational functions, has complicated our understanding of how local inhibitory networks contribute to sensory gain adaptations. Most recent work has focused on three major GABAergic cell types (**Figure 3B**): Fast-spiking parvalbumin positive (PV) interneurons that target perisomatic regions of excitatory neurons; low-threshold spiking somatostatin positive (SST) interneurons that target dendrites; and sparse, dendritic targeting vasoactive intestinal peptide (VIP) interneurons that often target other inhibitory interneurons to form disinhibitory circuits (Wood et al., 2017).

PV interneurons act as key mediators of response gain in cortical principal cells (Cardin et al., 2009; Sohal et al., 2009; Atallah et al., 2012; Wilson et al., 2012) and have been broadly implicated in feedforward circuits contributing to frequency tuning, adaptation, and gap encoding in ACx (Moore and Wehr, 2013; Li et al., 2014, 2015; Aizenberg et al., 2015; Natan et al., 2015; Keller et al., 2018). For these reasons, recent work has attempted to elucidate the role of PV neurons in contrast gain control in the ACx (Cooke et al., 2020). Optogenetic manipulation of PV neurons did indeed modulate the overall gain of ACx principal neurons. However, PV-mediated inhibition was minimally involved in gain adaptations to changes in sound variance and PV interneuron activity itself was not modulated by stimulus contrast. Thus, PV neuron activity modulates the gain of auditory cortical responses but not in a contrast-specific manner. Consistent with these findings, both background noise and PV neuron activation alter response gain of ACx principle neurons but the effects of these manipulations are additive, suggesting they involve independent mechanisms (Christensen et al., 2019). It is important to note, however, that similar levels of contrast gain adaptation are observed in the cortex and subcortical structures in mice (Lohse et al., 2020). The fact that local manipulation of PV neurons in the ACx does not influence contrast gain modulation does not preclude a role for inhibitory circuits in subcortical auditory structures. It will be important to use similar approaches to examine the role of inhibition in subcortical auditory areas as well as the role of other inhibitory interneuron cell types in the cortex.

On the other hand, recent work has demonstrated that PV neuron activity in the ACx is strongly modulated by behavioral state, suggesting that this class of interneurons may be involved in top-down attentional gain modulation. *In vivo* whole-cell recordings in awake mice found that spontaneous and sensory-evoked responses of both excitatory and PV neurons in the ACx are scaled down when animals transitioned from quiescence to active behavior (Zhou et al., 2014). This behavioral-state dependent gain modulation preserved tuning properties of ACx principle neurons but increased signal-to-noise ratios by relatively suppressing spontaneous activity more than evoked activity. PV neurons are also strongly regulated by motor cortical projections that act to suppress ACx activity associated with internally generated acoustic stimuli during locomotion (Nelson et al., 2013; Schneider et al., 2014). While it remains to be determined how attention engages PV neurons in the auditory system, PV neuron activity in the prefrontal cortex is increased with goal-driven attentional processing and PV neuron activity levels correlated with behavioral performance on the 5-choice serial reaction time task, a common rodent attentional task (Kim et al., 2016). Importantly, PV neurons play in integral role in generation of gamma oscillations in the cortex (Cardin et al., 2009; Sohal et al., 2009) and attentional processing is characterized by increases in gamma activity in sensory regions (Fries et al., 2001; Gregoriou et al., 2015; Ni et al., 2016). Gamma activity has been suggested to modulate the gain of incoming sensory input (Tiesinga et al., 2004, 2008; Börgers et al., 2005; Ni et al., 2016), providing a link between PV neuron function and attentional gain control (Tiesinga et al., 2004, 2008; Börgers et al., 2005; Ni et al., 2016). Thus, while PV neurons in the cortex may not be necessary for bottom-up contrast gain control, they are likely to be important mediators of top-down attentional gain modulation.

SST interneurons have been implicated in a variety of forms of auditory cortical adaptations, including SSA (Kato et al., 2015; Natan et al., 2015, 2017) and forward suppression (Phillips et al., 2017). While a role for SST neurons in contrast gain control or other forms of stimulus statistic adaptation remain to be determined, cortical SST neurons do exhibit properties that make them well-suited for these types of computations compared to PV neurons. For instance, while PV neurons are co-tuned for frequency with neighboring excitatory neurons in the ACx (Moore and Wehr, 2013), SST neurons are involved in a form of network-level lateral inhibition in the cortex (Kato et al., 2017). This lateral inhibitory network could provide a substrate for divisive normalization, a canonical computational strategy used throughout sensory systems to implement gain modulation for invariant sensory representation (Olsen and Wilson, 2008; Carandini and Heeger, 2012). SST neurons are also recruited slightly later than PV or excitatory cells, and SST neurons are more tightly tuned with higher intensity thresholds, suggesting they may contribute to feedback modulation of cortical circuits in response to stimulus history (Li et al., 2014, 2015; Kato et al., 2017).

A third important cortical inhibitory cell-type is the VIP expressing interneuron. VIP neurons represent only 1–2% of cortical neurons but can have broad impact on cortical circuit

function, as they target other cortical interneurons in superficial layers (**Figure 3B**) (Pfeffer et al., 2013). VIP neurons also receive strong neuromodulator input and are highly innervated by intracortical projections from outside of primary sensory areas (Zhang et al., 2014). These properties make VIP neurons well-suited to implement top-down modulation of cortical response gain via disinhibition. Consistent with this notion, VIP neuron activity in visual cortex is upregulated during locomotion and optogenetic activation of VIP neurons increases response gain of visual cortical excitatory neurons, mimicking the effect of locomotion (Fu et al., 2014). Optogenetic stimulation of VIP interneurons during a visual contrast detection task improves performance, while activating either SST or PV interneurons reduces the ability of the mouse to detect lower contrasts (Cone et al., 2019). Similar effects are seen for frequency tuning in the ACx, where VIP activation transiently suppresses SST and, to a lesser extent, PV neuron activity, leading to disinhibition in a subset of tone-responsive neurons and an increase in the gain of the corresponding tuning curves (Pi et al., 2013). Moreover, this study demonstrated that VIP neurons were strongly recruited in response to reinforcement signals during a tone discrimination task. Thus, VIP neurons are particularly well-suited to mediate top-down gain modulations via a disinhibitory cortical microcircuit that is engaged under specific behavioral conditions, and may therefore play an important role in attentional modulation of auditory processing.

Cortico-Fugal Circuits Contributing to Bottom-Up and Top-Down Adaptations

Descending projections from the ACx are far more numerous than ascending projections, and these massive yet poorly understood corticofugal projections target virtually every level of the auditory system, including the MGB, IC, cochlear nucleus (CN), superior olivary complex (SOC) and even the cochlea (**Figure 3A**) (Winer et al., 2002; Xiao and Suga, 2002; Meltzer and Ryugo, 2006; Llano and Sherman, 2009; Jäger and Kössl, 2016). While we are only beginning to understand how these descending projections influence sound perception, there is strong evidence for top-down regulation of subcortical sound processing via corticofugal projections. As with local cortical inhibitory neurons, corticofugal neurons are a heterogeneous set of cells with diverse properties and projection targets. Early studies revealed that stimulating cortico-thalamic (CT) projecting fibers egocentrically enhances tuning to match the origin of the descending cortical region (Yan and Suga, 1996; Zhang et al., 1997). More recent work has identified the complexity of this pathway in serving to balance the competing demands of increasing neuronal sensitivity for rapid signal detection or dampening excitability to enhance fine-tuned feature discrimination (Happel et al., 2014; Guo et al., 2017; Homma et al., 2017). While activation of CT neurons has been shown to decrease cortical response gain via direct activation of local inhibitory interneurons and/or projections to the TRN (Olsen et al., 2012; Bortone et al., 2014), recent evidence indicates that CT neurons can drive both increases and decreases to cortical gain depending on the timing of their activation relative to ascending input (Crandall et al., 2015; Guo et al., 2017). Cortico-collicular (CC) projecting neurons make-up a distinct class of corticofugal neurons with different projection targets and response properties than CT neurons (Winer et al., 2002; Williamson and Polley, 2019). CC projections to the IC enhance SSA and sharpen frequency tuning, biasing the receptive fields of subcortical auditory neurons toward frequently occurring or highly salient stimuli (Zhang et al., 2005; Bajo et al., 2010; Blackwell et al., 2016).

How do descending auditory projections contribute to stimulus statistic adaptation? While cortical silencing has significant effects on neuronal excitability in the MGB and IC, this manipulation does not affect contrast gain control (Lohse et al., 2020) or mean level adaptation (Robinson et al., 2016) in these structures, indicating that subcortical gain and dynamic range adaptations occur independently of top-down cortical feedback. However, cortical inactivation did interfere with meta-adaptation in the IC (Robinson et al., 2016) and several studies have shown that auditory attentional tasks modulate efferent projections back to the cochlea (Giard et al., 1994; Marian et al., 2018). Thus, it has been proposed that corticofugal projections play an important role in providing contextual information to upstream auditory areas that is necessary for interpreting ambiguous signals, such as those encountered in complex or noisy acoustic environments (Asilador and Llano, 2021). Indeed, electrical stimulation of ACx in humans was shown to modulate subcortical auditory pathways and enhance speech recognition under challenging conditions (de Boer and Thornton, 2008; Srinivasan et al., 2012; Shastri et al., 2014). Together, these studies suggest that descending corticofugal projections likely play an important role in top-down modulation of auditory gain in response to changes in behavioral state or context but may not be necessary for adaption to sound statistics observed in subcortical auditory areas.

Within this section we have provided a summary of our current understanding of cellular and circuit mechanisms that contribute to bottom-up and top-down adaptation throughout the auditory system. We have focused on both local synaptic and circuit interactions between excitatory and inhibitory neurons as well long-range connections between auditory regions that play essentials roles in adaptive sound processing. The role of specific interneuron subclasses and their specialized contributions to subcortical and cortical microcircuits remains an active area of interest in sensory processing. In particular, the contribution of specific interneuron classes to bottom-up stimulus statistic adaptation and top-down task-dependent receptive-field plasticity remains to be fully elucidated. While recent work has indicated that PV neurons are unlikely to mediate contrast gain control in the ACx, the role of other interneuron subtypes is less well understood. Recent computational studies have suggested that top-down inhibitory neurons that disinhibit bottom-up cortical circuits, similar to the VIP neuron circuit motif described above, can explain the attentional effects of auditory tuning properties (Chou and Sen, 2021). However, this model remains to be tested experimentally. Furthermore, much less is known about the role of interneurons subcortically as it relates to whether the behavioral state modulations that

are actively present in cortex contribute to adaptation and meta-adaptation observed in ascending subcortical structures. Finally, a major challenge for the field going forward is how to best isolate the contributions of these cell classes due to their interconnected nature and the presence of disinhibitory effects that are difficult to disentangle while using traditional methods of manipulation.

HEARING LOSS AND HEARING IN NOISY ENVIRONMENTS

One of most prominent and disabling disruptions associated with hearing loss is the inability to hear in noisy environments. Difficulties hearing in noise could be due to a general reduction in audibility and degraded encoding of incoming sound input as a consequence of hearing loss. However, studies have shown that speech perception and hearing in noise difficulties are present even when cochlear amplification is accounted for Peters et al. (1998); Johannesen et al. (2016). Moreover, hearing in noise difficulties often occur even in the absence of overt audiometric threshold shifts (Kraus et al., 2000; Zeng et al., 2005). Hearing loss fundamentally alters the pattern and level of incoming sound, and thus will greatly affect the stimulus statistics the auditory system is exposed to. Indeed, hearing loss is often associated with central auditory gain enhancement in attempts to preserve sound detection levels (Auerbach et al., 2014; Chambers et al., 2016; Salvi et al., 2016). Thus, it is possible that compensatory gain changes with pathological changes to auditory input and adaptations to more physiological changes to sound statistics engage overlapping mechanisms and may interfere with each other. Hearing loss may also interfere with attentional mechanisms important for sound perception. Selective attention performs best when auditory streams can be segregated based on select features. Degradation of spectrotemporal structure impairs adaptation accuracy and reduces the efficiency of anticipated noise. Inaccuracies build at multiple levels, delaying and reducing the efficiency with which attention groups relevant objects. In complex scenes, where background noise statistics and the spectrotemporal features of the target can rapidly fluctuate, hearing impaired listeners have more difficulty forming perceptual objects from their environments (Shinn-Cunningham and Best, 2008). Under these conditions, the benefits of knowing what features to direct attention to are degraded and reduce the capacity for cognitive control to benefit the listener. Thus, hearing loss not only affects bottom-up gain adaptations, but these changes are compounded by reducing the capacity for top-down attentional mechanisms to help contribute to auditory scene analysis. In the preceding sections, we have reviewed the primary bottom-up and top-down adaptations that contribute to auditory scene analysis, as well as the potential mechanisms underlying their generation. In this section we will discuss how our understanding of these adaptive coding strategies in normal hearing can help us understand the often devastating effects of hearing loss on listening in complex acoustic environments.

Hidden Hearing Loss

Sensorineural hearing loss is often associated with overt damage to sensory hair cells, resulting in elevated sound detection thresholds (**Figures 4Ai,ii,iii**) (Schmiedt, 1984; Cunningham and Tucci, 2017). Permanent threshold shifts in the clinical pure tone audiogram have thus traditionally been a key criterion for diagnosing hearing loss (Simel et al., 2016). However, many individuals with clinically normal audiometric thresholds nonetheless report significant auditory perceptual disruptions, including temporal processing deficits, impaired speech perception, and most prominently, difficulties hearing in noisy environments (Starr et al., 1996; Zeng et al., 2005; Cunningham and Tucci, 2017; Ralli et al., 2019). Recent evidence from animal models has suggested that cochlear neuronal degeneration can occur even without overt hair cell damage or permanent threshold shifts (Kujawa and Liberman, 2009). This so called "hidden hearing loss" (HHL), due to the fact that this dysfunction is not revealed by standard audiometric tests, is estimated to occur in ∼12–15% of individuals (Tremblay et al., 2015; Spankovich et al., 2018). HHL is likely a key contributor to difficulties hearing in noise in the absence of clinically diagnosed hearing loss (Plack et al., 2014; Ralli et al., 2019) and has also been suggested to contribute to auditory perceptual disorders like tinnitus and hyperacusis that are often associated with hearing impairment (Schaette and McAlpine, 2011; Hickox and Liberman, 2014). Below we discuss the consequences of cochlear degeneration on peripheral and central auditory function, with a particular focus on how central adaptations to this form of hearing loss may interfere with our ability to compensate for noisy environments.

Cochlear Synaptopathy

While there are three times as many outer hair cells (OHC) than inner hair cells (IHC), virtually all (∼95%) afferent signals from the cochlea are relayed to the central auditory system via IHCs synapsing on type 1 spiral ganglion neurons (SGNs) (**Figure 4Ai**) (Spoendlin, 1972). Accumulating evidence suggests that the synapses between IHCs and type I SGNs, whose axons comprise the AN tract, appear to be most vulnerable to noise- or age-related hearing loss. Indeed, animal studies have found that there is a marked reduction of IHC-Type 1 SGN synaptic contacts following exposure to ototoxic drugs, environmental noise, or aging, and this synaptopathy often proceeds overt hair cell damage (**Figure 4Aii**) (Liberman and Kujawa, 2017; Wu et al., 2019; Kohrman et al., 2020). Remarkably, it has been shown that animals with damage restricted to the IHC-type I SGN complex maintain normal hearing thresholds in quiet despite severely reduced afferent drive to the central auditory system (**Figure 4B**) (Lobarinas et al., 2013; Chambers et al., 2016). However, these same animals perform much poorer than control animals when challenged with a more difficult task, such as tone detection in background noise, gap-in-noise detection, or a remote masking paradigm (**Figure 4C**) (Salvi et al., 2016; Lobarinas et al., 2020; Resnik and Polley, 2021). These results indicate that cochlear degeneration could contribute to real-world listening difficulties even in the absence of threshold shifts in the clinical audiogram. It

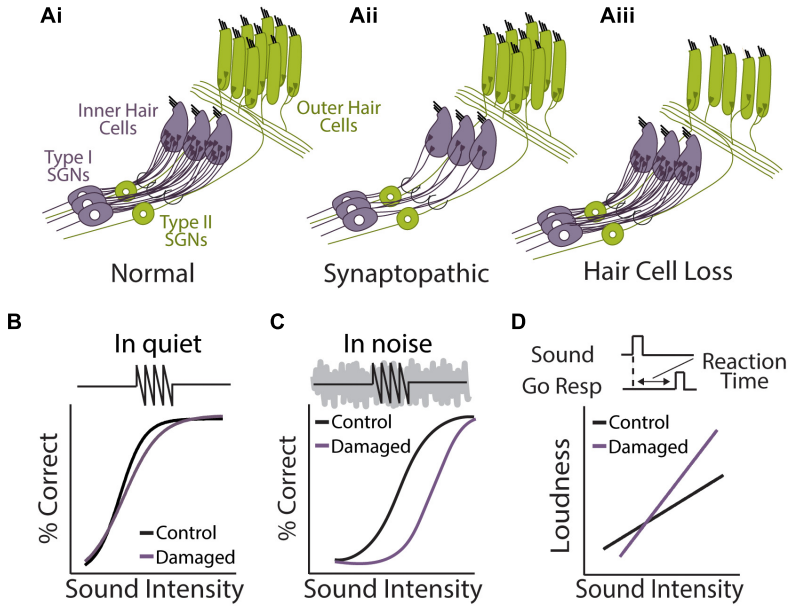

FIGURE 4 | Perceptual consequences of sensorineural hearing loss. **(Ai,ii,iii)** Diagram of cochlear hair cells and spiral ganglion neuron connectivity under normal **(Ai)** or pathological conditions of synaptopathy **(Aii)**, or sensory hair cell damage **(Aiii)**. **(Ai)** Purple: over 95% of afferent input to central auditory system comes from type 1 spiral ganglion neurons (SGN) that form synaptic contacts inner hair cells (IHCs), the main conventional sensory receptors of the cochlea. IHCs are innervated by multiple (10–20) type I neurons but each type I neuron only contacts a single IHC. Green: unmyelinated type II SGNs form synaptic contacts with multiple outer hair cells (OHCs) but each OHC only receives one contact from one Type-II neuron. Type-II SGNs represent only 5% of afferent input and are not involved with transmission of acoustic information to brain. Rather, the major role of OHCs is to amplify the cochlear mechanical response to low-level input, providing increased sensitivity to low intensity sounds. **(Aii)** In the synaptopathic ear, many of the synaptic contacts between type I SGNs and IHCs have degenerated, leaving fewer afferent nerve fibers to relay sound information from the ear to the brain, which may underly hidden hearing loss and impaired speech-in-noise perception. **(Aiii)** Many forms of acquired sensorineural hearing loss are associated with damage to OHCs and disruption to mechanical cochlear gain control mechanisms, leading to permanent threshold shifts, loudness recruitment, and broader frequency tuning. **(B)** Tone detection behavior in animals with selective damage the type I SGN-IHC complex (purple) is remarkably normal under quiet conditions, even with moderate to severe cochlear deafferentation. **(C)** Tone-in-noise detection is more severely impaired in animals with selective damage the type I SGN-IHC complex (purple) even though thresholds in quiet are maintained. Schematized data in panels **(B,C)** adapted from Resnik and Polley (2021). **(D)** Auditory reaction time (RT) measures of loudness growth in animal models (Radziwon and Salvi, 2020) have demonstrated that some forms of hearing loss can result in abnormal increases to the slope of RT-intensity functions, consistent with loudness recruitment and/or hyperacusis.

should be noted that AN fiber loss following kainic acid treatment in budgerigars, a common avian model of hearing (Dooling and Popper, 2000), does not result in deficits to tone-in-noise detection (Henry and Abrams, 2021), suggesting there may be species-specific effects of cochlear degeneration and/or central adaptation to hearing loss.

The difference in detection between quiet and noisy conditions following cochlear degeneration may be due in part to peripheral mechanisms. Spared AN fibers maintain normal thresholds and tuning following IHC or SGN degeneration and detecting tones in quiet may only require a small fraction of surviving peripheral afferents (Wang et al., 1997; Kujawa and Liberman, 2009; Salvi et al., 2016). Interestingly, the AN fibers most susceptible to noise-induced synaptopathy are low and medium SR fibers, which have higher thresholds and are thought to be useful for hearing in noisy environments (Wang et al., 1997; Furman et al., 2013). However, substantial recovery of sound detection thresholds is seen even with ototoxic treatments that cause near complete loss (∼95%) of IHC-SGN synapses (Chambers et al., 2016) or selective lesion of IHCs, which are similarly contacted by all subsets of AN fibers (Lobarinas et al., 2013; Salvi et al., 2016). In fact, IHC lesions were actually shown to result in an increased proportion of low SR nerves relative to medium and high, opposite to what is observed with noise-induced synaptopathy (Salvi et al., 2016). Thus, while peripheral mechanisms certainly contribute to perceptual alterations associated with cochlear degeneration, there is growing awareness that adaptations in the central auditory system are essential for fully understanding the perceptual consequences of cochlear hearing impairment.

Central Gain Enhancement Following Hearing Loss

Loss of afferent drive to the central auditory system— be it due to ototoxic drugs, sensorineural hearing loss, or acoustic deprivation— have been shown to result in a compensatory increase in neuronal gain in the central auditory system, a phenomenon termed central gain enhancement (**Figure 5A**) (Gerken et al., 1984; Syka, 2002; Auerbach et al., 2014; Chambers et al., 2016; Salvi et al., 2016). Gain increases due to sensorineural hearing loss have been observed at every level of the auditory system. AN synapses onto CN bushy cells have been shown to

homeostatically adapt their pre-synaptic strength in response to changes in acoustic input (Zhuang et al., 2020), indicating that dynamic gain adaptations are in place at the earliest points of the central auditory system. Indeed, neuronal gain increases in response to hearing loss have been observed in several auditory brainstem nuclei, albeit in a cell-type specific and time-restricted manner (Boettcher and Salvi, 1993; Brozoski et al., 2002; Cai et al., 2009). However, in studies where concurrent recordings from multiple levels of the auditory system were performed, it has been consistently found that gain changes with drug or noise-induced hearing loss are more rapid and more complete in the ACx compared to subcortical structures like in the IC or CN (**Figure 5A**) (Popelar et al., 1987; Syka et al., 1994; Auerbach et al., 2014; Chambers et al., 2016; Salvi et al., 2016). Thus, like more rapid adaptations to stimulus statistics, sustained gain increases following pathological changes to sound input progressively develop through the ascending auditory system, with the most complete recovery being observed at the level of the ACx (**Figure 5A**). Central gain enhancement following cochlear degeneration has also been shown to result in enhanced sound-evoked activity in corticofugal projections, suggesting that in addition to changes along the ascending auditory pathway, descending projections are also altered in response to hearing loss (Asokan et al., 2018).

What are the perceptual consequences of central gain enhancement? There is growing evidence that central gain enhancement is associated with restoration of hearing thresholds in quiet (**Figure 4B**). Parallel behavioral and neurophysiological studies in mice given round window application of the Na^+/K^+ ATPase pump inhibitor ouabain, which selectively destroys type-I SGNs, or chinchillas treated with the anti-cancer agent carboplatin, which selectively destroys IHCs in these animals, have shown that recovery in tone detection thresholds in quiet (**Figure 4B**) corresponds with recovery of intensity-response functions in ACx (**Figure 5A**) (Chambers et al., 2016; Salvi et al., 2016). However, this compensatory plasticity is unable to restore all aspects of auditory processing that are disrupted by sensorineural hearing loss and, in fact, may actively contribute a range of auditory perceptual deficits associated with hearing impairment as well. For instance, while central gain enhancement restores mean firing rates and intensity coding in the ACx, this adaptation cannot compensate for degradation of temporal processing with hearing loss, which depends on specialized subcortical circuits optimized for fast time scales (Chambers et al., 2016). Interestingly, gain increases in the ACx following hearing loss can often overshoot baseline levels, resulting in sound-evoked hyperactivity (**Figure 5A**). This excessive increase in central gain may contribute to the development of hyperacusis, a sound intolerance disorder often associated with hearing loss (Zeng, 2013; Auerbach et al., 2014; Pienkowski et al., 2014). Indeed, cortical gain increases are associated with maladaptive changes to loudness perception following ototoxic (Auerbach et al., 2018) or noise-induced hearing loss (**Figure 4D**) (Radziwon et al., 2019). These results highlight the perceptual trade-offs that inevitably arise when sensory systems must adapt their neural representations to changes in the environment.

Recent evidence suggests the intriguing notion that central gain changes that support restoration of sound processing in quiet backgrounds may actively interfere with auditory circuit mechanisms that normally support adaptation to background noise. While thresholds in quiet are remarkably normal in animals with severe cochlear degeneration, performance in hearing damaged animals was much worse when challenged with a tone-in-noise detection task (**Figure 4C**) (Lobarinas et al., 2013; Salvi et al., 2016; Resnik and Polley, 2021). Examination of mean-level adaptation in the IC of mice given a noise exposure that produces HHL and central gain increases found significant impairment in adaptive coding for loud environments (Bakay et al., 2018). While both dynamic range and gain adaptations were still observed in the IC of noise-exposed animals, threshold shifts were significantly reduced compared to controls and intensity-response functions were less informative about the sound level distribution the animals were exposed to, particularly for loud environments (**Figure 5B**). This impairment in adaptive coding could contribute to difficulties hearing in noisy environment. Indeed, a follow-up study found that noise-induced HHL in gerbils, which have a hearing range more similar to humans than mice, reduced the ability of IC neurons to discriminate between speech tokens presented in background noise at high sound intensities (75 dB SPL), although discriminability at moderate sound intensities (60 dB SPL) was surprisingly improved (Monaghan et al., 2020). A phenomenological model of cochlear synaptopathy that selectively impairs high threshold, low-SR AN fibers and result in enhanced central gain could reproduce this pattern of improved discrimination at moderate levels but decreased performance at high levels, providing a link between peripheral pathology and central plasticity.

The mechanisms of central gain enhancement remain to be completely elucidated; however, several lines of evidence suggest that a combination of increased excitatory neuronal function and, in particular, decreased inhibitory function contribute to this experience-dependent plasticity (Yang et al., 2011; Auerbach et al., 2014; Resnik and Polley, 2017; Balaram et al., 2019). Recent studies using optogenetic manipulation (Resnik and Polley, 2017) or chronic two-photon calcium imaging of genetically labeled PV inhibitory neuron populations in the ACx (Resnik and Polley, 2021) have demonstrated that central gain changes and perceptual restoration of detection thresholds are correlated with decreased PV-mediated inhibition in the ACx. Intriguingly, these recent studies have indicated that alterations to cortical E/I balance that help restore hearing thresholds in quiet may be actively interfering with hearing in noise. Chronic two photon imaging of putative excitatory and PV inhibitory neuronal populations in the ACx found that cochlear degeneration was associated with distinct forms of plasticity in cortical excitatory and inhibitory neurons, with near complete recovery in sound-evoked responses for cortical excitatory neurons but a persistent decrease in PV neuron activity (**Figure 5Ci**) (Resnik and Polley, 2021). The combined effect of these changes was an increase in cortical gain that corresponded with recovery of tone detection in quiet, but also an imbalance in spontaneous activity rates between excitatory and PV inhibitory neurons that led to random surges of correlated activity that impaired tone-detection

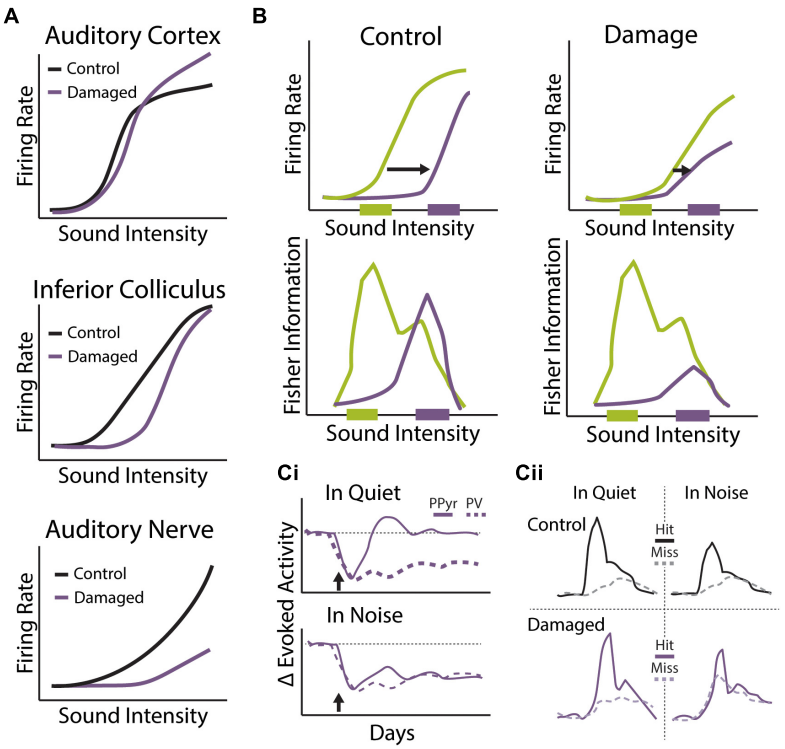

FIGURE 5 | Central gain enhancement following sensorineural hearing loss. **(A)** Schematics of rate-intensity functions from multiple levels of the auditory system under control conditions (black) or following cochlear damage via noise or ototoxic drug exposure that results in the central gain enhancement (purple). While output from the AN is severely degraded in terms of evoked-response threshold and suprathreshold intensity coding, rate-intensity functions gradually recover at ascending levels of the auditory system so that thresholds and suprathreshold responses are nearly normal at the level of the ACx and, in some cases, exhibit rebound hyperactivity. **(B)** Mean sound level adaptation to loud sound environments is altered with noise-induced hidden hearing loss. Rate-intensity functions from the IC of control (left) and noise exposed (right) mice when exposed to dynamic sound stimulus that switches between distributions with high probability of low sound levels (green) and high probability of high sound levels (purple). Noise exposed animals exhibit less dynamic range adaptation (top) and response functions carry less information about loud sound environments (bottom) compared to control animals. Schematized data adapted from Bakay et al. (2018). **(Ci,ii)** Cochlear degeneration triggers compensatory changes to cortical excitatory/inhibitory balance that differentially effects tone detection in quiet and noise. **(Ci)** Distinct changes to tone-evoked calcium transients in putative excitatory pyramidal neurons (PPyr, solid lines) and genetically-labeled PV inhibitory neurons (dashed lines) in the ACx following ouabain induced cochlear degeneration (arrow). Following transient loss of evoked activity, PPyr neurons exhibit near complete recovery of evoked-response size in quiet but not in background noise. Sustained decreases to tone-evoked activity in PV neurons are observed following ouabain treatment in both quiet and noise conditions. **(Cii)** Combined behavioral and imaging sessions showing differences in tone-evoked responses in cortical PPyr neurons on hit (solid lines) vs. miss (dashed lines) trials in quiet or background noise from animals before (control, black) and after ouabain-induced cochlear degeneration (damaged, purple). Decreased tone-in-noise detection in ouabain-treated animals is not only associated with diminished tone-evoked responses on hit trials but also increased activity on miss trials. These results suggest that altered E/I balance in the ACx following cochlear degeneration may lead to impaired adaption to background noise and decreased signal-to-noise ratios for detection of foreground stimuli. Schematized data adapted from Resnik and Polley (2021).

in background noise. Interestingly, combined behavioral and imaging sessions demonstrated that impairments to tone-in-noise detection following cochlear degeneration was not only the result of diminished evoked activity to target sounds but increased sensitivity to background noise, so that signal-to-noise ratios were decreased for foreground sounds (**Figure 5Cii**). Indeed, perceptual misses in noise were better predicted by levels of neural synchronization during the pre-stimulus period than the size of stimulus-evoked responses (Resnik and Polley, 2021). Thus, diminished PV neuron-mediated inhibition in ACx following hearing loss may be responsible for both adaptive recovery of sound detection in quiet as well as impaired adaption to background noise that disrupts perception in more challenging conditions. This degraded ability to adapt to noisy conditions could reflect impairments in bottom-up contrast gain adaptation, altered top-down modulation of cortical inhibitory circuits that act to reduce spontaneous activity during behavioral engagement, or a combination of the two.

Hearing Loss and Top-Down Cognition

In addition to having sizeable impact on bottom-up sound processing and adaptation, hearing loss is also likely to affect top-down regulation of sound feature encoding. There is a well-characterized relationship between hearing loss and cognitive decline, although the directionality and mechanisms are strongly debated (Lin et al., 2013). It is not clear if cognitive decline or age-related hearing loss precede one another or if any such effects would even be generalizable more broadly across individuals. We will not review this debate here, except to acknowledge that cognitive decline impacts cortical circuits essential for

attentional sound processing. As discussed previously, top-down circuits are critical for segregating attended streams in complex environments and reductions in cognitive capacity can alter performance in auditory scene analysis. Cognitive decline is a near universal phenomenon associated with normal aging with decline levels highly correlated with age (Park et al., 2003). Older adults are more influenced by the presence of sensory perceptual conflicts during tasks of focused attention and this coincides with reduced measures of conflict in fronto-parietal ERP markers associated with greater attentional control (Passow et al., 2014). One potential source for this reduction in performance is diminished contextual adaptation of sound level statistics within the listening environment. Herrmann et al. (2018) found that older listeners exposed to a sound distribution with two-levels showed similar neural response magnitudes but reduced capacity for sensory adaptation relative to young listeners. This finding suggests reduced capacity for adaptation to the statistical properties of the context and impaired ability to filter unattended auditory streams. While much more work is needed to elucidate the relationship between hearing loss, cognitive decline, and auditory scene analysis, current evidence suggests that hearing impairments that arise with age are likely the combined effect of disruptions to bottom-up sound processing and top-down auditory attentional regulation.

CONCLUSION

We live in a world full of sounds. The auditory system employs a variety of adaptive coding strategies (**Figures 1**, **2**) to navigate this cacophonous environment, including: compensatory dynamic range and gain adaptations to incoming stimulus statistics in order to build level and contrast invariant tuning of sound features under different background conditions (Rabinowitz et al., 2013); adaptive spatial tuning for localizing and focusing on specific sound sources to aid in the segregation of auditory streams in the presence of complex sound environment (Reed et al., 2020); and top-down attentional mechanisms that modulate auditory response and receptive field properties to selectively amplify behaviorally relevant sound features (Fritz et al., 2005b). These adaptations are observed throughout the ascending and descending auditory hierarchy to various degrees and can be both rapid, as seen in task-engaged subjects in perceptual decision-making paradigms, as well as sustained, as seen with long-term changes to auditory input associated with hearing loss.

A number of synaptic and circuit mechanisms are used to implement adaptive coding strategies in the auditory system (**Figure 3**), including: use-dependent changes in synaptic transmission; regulation of E/I balance to modulate response gain and minimize the influence of background noise; and top-down disinhibitory circuit motifs that can selectively modify sound encoding in response to changes in behavioral state. Interestingly, it appears that bottom-up and top-down gain changes are mediated by distinct mechanisms, suggesting the individual contributions of these different forms of adaptation are at a minimum additive or perhaps even work synergistically to enhance performance in challenging auditory scenes. Future work must investigate this possibility further by comparing neurophysiological adaptations to sound statistics in passively listening versus task-engaged animals in combination with *in vivo* manipulation of putative generators of bottom-up and top-down adaptations. It is also possible that different forms of auditory plasticity can interfere with each other, as may be the case with sensorineural hearing loss.

Listening in noisy environments poses additional challenges for those with hearing loss (**Figure 4**). This difficulty is due in part to degraded encoding of incoming stimuli, leading to impoverished representation of spectrotemporal sound features and disrupted ability to segregate sound sources based on select features. However, recent evidence suggests that compensatory plasticity mechanisms that help restore rapid signal detection following loss of afferent drive may actively interfere with the auditory system's ability to adapt to more challenging listening conditions as well. For instance, increased central auditory excitability following hearing loss allows for the amplification of diminished sound-driven input from the periphery, but it may also make the auditory system more sensitive to the influence of background sounds and impair adaptation to noisy environments (**Figure 5**). A future challenge will be to identify whether central gain enhancement seen with hearing loss reflects bottom-up gain adaptations in response to changes in sound level statistics, reduced top-down modulation of cortical inhibitory circuits that coincide with disruptions in attentional mechanisms, or some interaction between these components.

In summary, the studies reviewed here indicate that the auditory system is highly adaptive, modulating its response properties to best fit the current environmental and/or behavioral goals. These adaptations appear to be crucial for optimal representation of sounds under diverse conditions and for listening in complex auditory environments. Further understanding of the mechanisms mediating bottom-up and top-down adaptations to sound processing, as well as the interaction between them, is crucial for harnessing the auditory system's vast potential to compensate for difficult listening conditions, particularly following sensorineural hearing loss.

AUTHOR CONTRIBUTIONS

BA and HG wrote the manuscript together and approved of the submitted version.

REFERENCES

Abbott, L. F., Varela, J. A., Sen, K., and Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science* 275, 220–224. doi: 10.1126/science.275.5297.221

Ahveninen, J., Jääskeläinen, I. P., Raij, T., Bonmassar, G., Devore, S., Hämäläinen, M., et al. (2006). Task-modulated "what" and "where" pathways in human auditory cortex. *Proc. Natl. Acad. Sci. U.S.A.* 103, 14608–14613. doi: 10.1073/pnas.0510480103

Aizenberg, M., Mwilambwe-Tshilobo, L., Briguglio, J. J., Natan, R. G., and Geffen, M. N. (2015). Bidirectional regulation of innate and learned behaviors that rely on frequency discrimination by cortical inhibitory neurons. *PLoS Biol.* 13:e1002308. doi: 10.1371/journal.pbio.1002308

Angeloni, C. F., Młynarski, W., Piasini, E., Williams, A. M., Wood, K. C., Garami, L., et al. (2021). Cortical efficient coding dynamics shape behavioral performance. *bioRxiv [Preprint]* doi: 10.1101/2021.08.11.455845

Angeloni, C., and Geffen, M. N. (2018). Contextual modulation of sound processing in the auditory cortex. *Curr. Opin. Neurobiol.* 49, 8–15. doi: 10.1016/j.conb.2017.10.012

Asilador, A., and Llano, D. A. (2021). Top-down inference in the auditory system: potential roles for corticofugal projections. *Front. Neural Circuits* 14:89. doi: 10.3389/fncir.2020.615259

Asokan, M. M., Williamson, R. S., Hancock, K. E., and Polley, D. B. (2018). Sensory overamplification in layer 5 auditory corticofugal projection neurons following cochlear nerve synaptic damage. *Nat. Commun.* 9:2468. doi: 10.1038/s41467-018-04852-y

Atallah, B. V., Bruns, W., Carandini, M., and Scanziani, M. (2012). Parvalbumin-expressing interneurons linearly transform cortical responses to visual stimuli. *Neuron* 73, 159–170. doi: 10.1016/j.neuron.2011.12.013

Atiani, S., David, S. V., Elgueda, D., Locastro, M., Radtke-Schuller, S., Shamma, S. A., et al. (2014). Emergent selectivity for task-relevant stimuli in higher-order auditory cortex. *Neuron* 82, 486–499. doi: 10.1016/j.neuron.2014.02.029

Auerbach, B. D., Radziwon, K., and Salvi, R. (2018). Testing the central gain model: loudness growth correlates with central auditory gain enhancement in a rodent model of hyperacusis. *Neuroscience* 407, 93–107. doi: 10.1016/j.neuroscience.2018.09.036

Auerbach, B. D., Rodrigues, P. V., and Salvi, R. J. (2014). Central gain control in tinnitus and hyperacusis. *Front. Neurol.* 5:206. doi: 10.3389/fneur.2014.00206

Baccus, S. A. (2006). From a whisper to a roar: adaptation to the mean and variance of naturalistic sounds. *Neuron* 51, 682–684. doi: 10.1016/j.neuron.2006.09.007

Bajo, V. M., Nodal, F. R., Moore, D. R., and King, A. J. (2010). The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nat. Neurosci.* 13, 253–260. doi: 10.1038/nn.2466

Bakay, W. M. H., Anderson, L. A., Garcia-Lazaro, J. A., McAlpine, D., and Schaette, R. (2018). Hidden hearing loss selectively impairs neural adaptation to loud sound environments. *Nat. Commun.* 9:4298. doi: 10.1038/s41467-018-06777-y

Balaram, P., Hackett, T. A., and Polley, D. B. (2019). Synergistic transcriptional changes in AMPA and GABAA receptor genes support compensatory plasticity following unilateral hearing loss. *Neuroscience* 407, 108–119. doi: 10.1016/j.neuroscience.2018.08.023

Banitt, Y., Martin, K. A. C., and Segev, I. (2007). A biologically realistic model of contrast invariant orientation tuning by thalamocortical synaptic depression. *J. Neurosci.* 27, 10230–10239. doi: 10.1523/JNEUROSCI.1640-07.2007

Barbour, D. L. (2011). Intensity-invariant coding in the auditory system. *Neurosci. Biobehav. Rev.* 35, 2064–2072. doi: 10.1016/j.neubiorev.2011.04.009

Bennett, C., Arroyo, S., and Hestrin, S. (2014). Controlling brain states. *Neuron* 83, 260–261. doi: 10.1016/j.neuron.2014.07.007

Bharadwaj, H. M., Lee, A. K. C., and Shinn-Cunningham, B. G. (2014). Measuring auditory selective attention using frequency tagging. *Front. Integr. Neurosci.* 8:6. doi: 10.3389/fnint.2014.00006

Bidet-Caulet, A., Fischer, C., Besle, J., Aguera, P.-E., Giard, M.-H., and Bertrand, O. (2007). Effects of selective attention on the electrophysiological representation of concurrent sounds in the human auditory cortex. *J. Neurosci.* 27, 9252–9261. doi: 10.1523/JNEUROSCI.1402-07.2007

Blackwell, J. M., Taillefumier, T. O., Natan, R. G., Carruthers, I. M., Magnasco, M. O., and Geffen, M. N. (2016). Stable encoding of sounds over a broad range of statistical parameters in the auditory cortex. *Eur. J. Neurosci.* 43, 751–764. doi: 10.1111/ejn.13144

Blundon, J. A., Bayazitov, I. T., and Zakharenko, S. S. (2011). Presynaptic gating of postsynaptically expressed plasticity at mature thalamocortical synapses. *J. Neurosci.* 31, 16012–16025. doi: 10.1523/JNEUROSCI.3281-11.2011

Boettcher, F. A., and Salvi, R. J. (1993). Functional changes in the ventral cochlear nucleus following acute acoustic overstimulation. *J. Acoust. Soc. Am.* 94, 2123–2134.

Börgers, C., Epstein, S., and Kopell, N. J. (2005). Background gamma rhythmicity and attention in cortical local circuits: a computational study. *Proc. Natl. Acad. Sci. U.S.A.* 102, 7002–7007. doi: 10.1073/pnas.0502366102

Bortone, D. S., Olsen, S. R., and Scanziani, M. (2014). Translaminar inhibitory cells recruited by layer 6 corticothalamic neurons suppress visual cortex. *Neuron* 82, 474–485. doi: 10.1016/j.neuron.2014.02.021

Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: A Bradford Book.

Bregman, A. S. (2008). "3.49 - Auditory scene analysis," in *The Senses: A Comprehensive Reference*, eds R. H. Masland, T. D. Albright, T. D. Albright, R. H. Masland, P. Dallos, D. Oertel, et al. (New York, NY: Academic Press), 861–870. doi: 10.1016/B978-012370880-9.00009-8

Brozoski, T. J., Bauer, C. A., and Caspary, D. M. (2002). Elevated fusiform cell activity in the dorsal cochlear nucleus of chinchillas with psychophysical evidence of tinnitus. *J. Neurosci.* 22, 2383–2390.

Cai, S., Ma, W.-L. D., and Young, E. D. (2009). Encoding intensity in ventral cochlear nucleus following acoustic trauma: implications for loudness recruitment. *J. Assoc. Res. Otolaryngol.* 10, 5–22. doi: 10.1007/s10162-008-0142-y

Carandini, M., and Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nat. Rev. Neurosci.* 13, 51–62. doi: 10.1038/nrn3136

Carandini, M., Heeger, D. J., and Senn, W. (2002). A synaptic explanation of suppression in visual cortex. *J. Neurosci.* 22, 10053–10065. doi: 10.1523/JNEUROSCI.22-22-10053.2002

Carcea, I., and Froemke, R. C. (2013). "Chapter 3 - Cortical Plasticity, excitatory–inhibitory balance, and sensory perception," in *Progress in Brain Research*, eds M. N. Michael, M. Merzenich, and M. V. V. Thomas (Amsterdam: Elsevier), 65–90. doi: 10.1016/B978-0-444-63327-9.00003-5

Carcea, I., Insanally, M. N., and Froemke, R. C. (2017). Dynamics of auditory cortical activity during behavioural engagement and auditory perception. *Nat. Commun.* 8:14412. doi: 10.1038/ncomms14412

Cardin, J. A., Carlen, M., Meletis, K., Knoblich, U., Zhang, F., Deisseroth, K., et al. (2009). Driving fast-spiking cells induces gamma rhythm and controls sensory responses. *Nature* 459, 663–667. doi: 10.1038/nature08002

Carrasco, M. (2011). Visual attention: the past 25 years. *Vision Res.* 51, 1484–1525. doi: 10.1016/j.visres.2011.04.012

Carruthers, I. M., Laplagne, D. A., Jaegle, A., Briguglio, J. J., Mwilambwe-Tshilobo, L., Natan, R. G., et al. (2015). Emergence of invariant representation of vocalizations in the auditory cortex. *J. Neurophysiol.* 114, 2726–2740. doi: 10.1152/jn.00095.2015

Carruthers, I. M., Natan, R. G., and Geffen, M. N. (2013). Encoding of ultrasonic vocalizations in the auditory cortex. *J. Neurophysiol.* 109, 1912–1927. doi: 10.1152/jn.00483.2012

Chambers, A. R., Resnik, J., Yuan, Y., Whitton, J. P., Edge, A. S., Liberman, M. C., et al. (2016). Central gain restores auditory processing following near-complete cochlear denervation. *Neuron* 89, 867–879. doi: 10.1016/j.neuron.2015.12.041

Chance, F. S., Abbott, L. F., and Reyes, A. D. (2002). Gain modulation from background synaptic input. *Neuron* 35, 773–782. doi: 10.1016/S0896-6273(02)00820-6

Chapuis, G. A., and Chadderton, P. T. (2018). Using temporal expectation to assess auditory streaming in mice. *Front. Behav. Neurosci.* 12:205. doi: 10.3389/fnbeh.2018.00205

Chechik, G., and Nelken, I. (2012). Auditory abstraction from spectro-temporal features to coding auditory entities. *Proc. Natl. Acad. Sci. U.S.A.* 109, 18968–18973. doi: 10.1073/pnas.1111242109

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *J. Acoust. Soc. Am.* 25, 975–979. doi: 10.1121/1.1907229

Chou, K. F., and Sen, K. (2021). AIM: a network model of attention in auditory cortex. *PLoS Comput. Biol.* 17:e1009356. doi: 10.1371/journal.pcbi.1009356

Christensen, R. K., Lindén, H., Nakamura, M., and Barkat, T. R. (2019). White noise background improves tone discrimination by suppressing cortical tuning curves. *Cell Rep.* 29, 2041–2053.e4. doi: 10.1016/j.celrep.2019.10.049

Cone, J. J., Scantlen, M. D., Histed, M. H., and Maunsell, J. H. R. (2019). Different inhibitory interneuron cell classes make distinct contributions to visual contrast

perception. *eNeuro* 6:ENEURO.0337-18.2019. doi: 10.1523/ENEURO.0337-18.2019

Cooke, J. E., Kahn, M. C., Mann, E. O., King, A. J., Schnupp, J. W. H., and Willmore, B. D. B. (2020). Contrast gain control occurs independently of both parvalbumin-positive interneuron activity and shunting inhibition in auditory cortex. *J. Neurophysiol.* 123, 1536–1551. doi: 10.1152/jn.00587.2019

Costalupes, J. A., Young, E. D., and Gibson, D. J. (1984). Effects of continuous noise backgrounds on rate response of auditory nerve fibers in cat. *J. Neurophysiol.* 51, 1326–1344. doi: 10.1152/jn.1984.51.6.1326

Crandall, S. R., Cruikshank, S. J., and Connors, B. W. (2015). A corticothalamic switch: controlling the thalamus with dynamic synapses. *Neuron* 86, 768–782. doi: 10.1016/j.neuron.2015.03.040

Cunningham, L. L., and Tucci, D. L. (2017). Hearing loss in adults. *N. Engl. J. Med.* 377, 2465–2473. doi: 10.1056/NEJMra1616601

Dahmen, J. C., Keating, P., Nodal, F. R., Schulz, A. L., and King, A. J. (2010). Adaptation to stimulus statistics in the perception and neural representation of auditory space. *Neuron* 66, 937–948. doi: 10.1016/j.neuron.2010.05.018

Davenport, W. B. (1952). An experimental study of speech-wave probability distributions. *J. Acoust. Soc. Am.* 24, 390–399. doi: 10.1121/1.1906909

David, S. V., Fritz, J. B., and Shamma, S. A. (2012). Task reward structure shapes rapid receptive field plasticity in auditory cortex. *Proc. Natl. Acad. Sci. U.S.A.* 109, 2144–2149. doi: 10.1073/pnas.1117717109

de Boer, J., and Thornton, A. R. D. (2008). Neural correlates of perceptual learning in the auditory brainstem: efferent activity predicts and reflects improvement at a speech-in-noise discrimination task. *J. Neurosci.* 28, 4929–4937. doi: 10.1523/JNEUROSCI.0902-08.2008

Dean, I., Harper, N. S., and McAlpine, D. (2005). Neural population coding of sound level adapts to stimulus statistics. *Nat. Neurosci.* 8, 1684–1689. doi: 10.1038/nn1541

Dean, I., Robinson, B. L., Harper, N. S., and McAlpine, D. (2008). Rapid neural adaptation to sound level statistics. *J. Neurosci.* 28, 6430–6438. doi: 10.1523/jneurosci.0470-08.2008

Deng, Y., Choi, I., and Shinn-Cunningham, B. (2020). Topographic specificity of alpha power during auditory spatial attention. *Neuroimage* 207:116360. doi: 10.1016/j.neuroimage.2019.116360

Ding, N., and Simon, J. Z. (2013). Adaptive temporal encoding leads to a background-insensitive cortical representation of speech. *J. Neurosci.* 33, 5728–5735. doi: 10.1523/JNEUROSCI.5297-12.2013

Dooling, R. J., and Popper, A. N. (2000). *Comparative Hearing: Birds and Reptiles*. New York, NY: Springer, doi: 10.1007/978-1-4612-1182-2_1

Downer, J. D., Rapone, B., Verhein, J., O'Connor, K. N., and Sutter, M. L. (2017). Feature-selective attention adaptively shifts noise correlations in primary auditory cortex. *J. Neurosci.* 37, 5378–5392. doi: 10.1523/JNEUROSCI.3169-16.2017

Ehret, G., and Riecke, S. (2002). Mice and humans perceive multiharmonic communication sounds in the same way. *Proc. Natl. Acad. Sci. U.S.A.* 99, 479–482. doi: 10.1073/pnas.012361999

Endepols, H., Feng, A. S., Gerhardt, H. C., Schul, J., and Walkowiak, W. (2003). Roles of the auditory midbrain and thalamus in selective phonotaxis in female gray treefrogs (Hyla versicolor). *Behav. Brain Res.* 145, 63–77. doi: 10.1016/S0166-4328(03)00098-6

Evans, E. F. (1972). The frequency response and other properties of single fibres in the guinea-pig cochlear nerve. *J. Physiol.* 226, 263–287. doi: 10.1113/jphysiol.1972.sp009984

Ferguson, K. A., and Cardin, J. A. (2020). Mechanisms underlying gain modulation in the cortex. *Nat. Rev. Neurosci.* 21, 80–92. doi: 10.1038/s41583-019-0253-y

Finlayson, P. G., and Adam, T. J. (1997). Excitatory and inhibitory response adaptation in the superior olive complex affects binaural acoustic processing. *Hear. Res.* 103, 1–18. doi: 10.1016/S0378-5955(96)00158-X

Finlayson, P. G., and Caspary, D. M. (1989). Synaptic potentials of chinchilla lateral superior olivary neurons. *Hear. Res.* 38, 221–228. doi: 10.1016/0378-5955(89)90067-1

Finn, I. M., Priebe, N. J., and Ferster, D. (2007). The emergence of contrast-invariant orientation tuning in simple cells of cat visual cortex. *Neuron* 54, 137–152. doi: 10.1016/j.neuron.2007.02.029

Francis, N. A., Winkowski, D. E., Sheikhattar, A., Armengol, K., Babadi, B., and Kanold, P. O. (2018). Small networks encode decision-making in primary auditory cortex. *Neuron* 97, 885–897.e6. doi: 10.1016/j.neuron.2018.01.019

Fries, P., Reynolds, J. H., Rorie, A. E., and Desimone, R. (2001). Modulation of oscillatory neuronal synchronization by selective visual attention. *Science* 291, 1560–1563. doi: 10.1126/science.1055465

Fritz, J. B., David, S. V., Radtke-Schuller, S., Yin, P., and Shamma, S. A. (2010). Adaptive, behaviorally gated, persistent encoding of task-relevant auditory information in ferret frontal cortex. *Nat. Neurosci.* 13, 1011–1019. doi: 10.1038/nn.2598

Fritz, J. B., Elhilali, M., and Shamma, S. A. (2005a). Differential dynamic plasticity of A1 receptive fields during multiple spectral tasks. *J. Neurosci.* 25, 7623–7635. doi: 10.1523/JNEUROSCI.1318-05.2005

Fritz, J. B., Elhilali, M., David, S. V., and Shamma, S. A. (2007). Does attention play a role in dynamic receptive field adaptation to changing acoustic salience in A1? *Hear. Res.* 229, 186–203. doi: 10.1016/j.heares.2007.01.009

Fritz, J., Elhilali, M., and Shamma, S. (2005b). Active listening: task-dependent plasticity of spectrotemporal receptive fields in primary auditory cortex. *Hear. Res.* 206, 159–176. doi: 10.1016/j.heares.2005.01.015

Fritz, J., Shamma, S., Elhilali, M., and Klein, D. (2003). Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nat. Neurosci.* 6, 1216–1223. doi: 10.1038/nn1141

Froemke, R. C., Carcea, I., Barker, A. J., Yuan, K., Seybold, B. A., Martins, A. R. O., et al. (2013). Long-term modification of cortical synapses improves sensory perception. *Nat. Neurosci.* 16, 79–88. doi: 10.1038/nn.3274

Fu, Y., Tucciarone, J. M., Espinosa, J. S., Sheng, N., Darcy, D. P., Nicoll, R. A., et al. (2014). A cortical circuit for gain control by behavioral state. *Cell* 156, 1139–1152. doi: 10.1016/j.cell.2014.01.050

Furman, A. C., Kujawa, S. G., and Liberman, M. C. (2013). Noise-induced cochlear neuropathy is selective for fibers with low spontaneous rates. *J. Neurophysiol.* 110, 577–586. doi: 10.1152/jn.00164.2013

Gerken, G. M., Saunders, S. S., and Paul, R. E. (1984). Hypersensitivity to electrical-stimulation of auditory nuclei follows hearing-loss in cats. *Hear. Res.* 13, 249–259. doi: 10.1016/0378-5955(84)90078-9

Giard, M.-H., Collet, L., Bouchet, P., and Pernier, J. (1994). Auditory selective attention in the human cochlea. *Brain Res.* 633, 353–356. doi: 10.1016/0006-8993(94)91561-X

Goutman, J. D. (2017). Mechanisms of synaptic depression at the hair cell ribbon synapse that support auditory nerve function. *Proc. Natl. Acad. Sci. U.S.A.* 114, 9719–9724. doi: 10.1073/pnas.1706160114

Gregoriou, G. G., Paneri, S., and Sapountzis, P. (2015). Oscillatory synchrony as a mechanism of attentional processing. *Brain Res.* 1626, 165–182. doi: 10.1016/j.brainres.2015.02.004

Grothe, B., and Koch, U. (2011). Dynamics of binaural processing in the mammalian sound localization pathway – The role of GABAB receptors. *Hear. Res.* 279, 43–50. doi: 10.1016/j.heares.2011.03.013

Grothe, B., Pecka, M., and McAlpine, D. (2010). Mechanisms of sound localization in mammals. *Physiol. Rev.* 90, 983–1012. doi: 10.1152/physrev.00026.2009

Guo, W., Clause, A. R., Barth-Maron, A., and Polley, D. B. (2017). A corticothalamic circuit for dynamic switching between feature detection and discrimination. *Neuron* 95, 180–194.e5. doi: 10.1016/j.neuron.2017.05.019

Haider, B., Duque, A., Hasenstaub, A. R., and McCormick, D. A. (2006). Neocortical network activity *in vivo* is generated through a dynamic balance of excitation and inhibition. *J. Neurosci.* 26, 4535–4545. doi: 10.1523/JNEUROSCI.5297-05.2006

Haider, B., Häusser, M., and Carandini, M. (2013). Inhibition dominates sensory responses in the awake cortex. *Nature* 493, 97–100. doi: 10.1038/nature11665

Hamilton, L. S., Sohl-Dickstein, J., Huth, A. G., Carels, V. M., Deisseroth, K., and Bao, S. (2013). Optogenetic activation of an inhibitory network enhances feedforward functional connectivity in auditory cortex. *Neuron* 80, 1066–1076. doi: 10.1016/j.neuron.2013.08.017

Happel, M. F. K., Deliano, M., Handschuh, J., and Ohl, F. W. (2014). Dopamine-modulated recurrent corticoefferent feedback in primary sensory cortex promotes detection of behaviorally relevant stimuli. *J. Neurosci.* 34, 1234–1247. doi: 10.1523/JNEUROSCI.1990-13.2014

Harris, K. D., and Thiele, A. (2011). Cortical state and attention. *Nat. Rev. Neurosci.* 12, 509–523. doi: 10.1038/nrn3084

Henry, K. S., and Abrams, K. S. (2021). Normal tone-in-noise sensitivity in trained budgerigars despite substantial auditory-nerve injury: no evidence of hidden hearing loss. *J. Neurosci.* 41, 118–129.

Herrmann, B., Augereau, T., and Johnsrude, I. S. (2020). Neural responses and perceptual sensitivity to sound depend on sound-level statistics. *Sci. Rep.* 10:9571. doi: 10.1038/s41598-020-66715-1

Herrmann, B., Maess, B., and Johnsrude, I. S. (2018). Aging affects adaptation to sound-level statistics in human auditory cortex. *J. Neurosci.* 38, 1989–1999. doi: 10.1523/JNEUROSCI.1489-17.2018

Herrmann, B., Schlichting, N., and Obleser, J. (2014). Dynamic range adaptation to spectral stimulus statistics in human auditory cortex. *J. Neurosci.* 34, 327–331. doi: 10.1523/JNEUROSCI.3974-13.2014

Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402. doi: 10.1038/nrn2113

Hickox, A. E., and Liberman, M. C. (2014). Is noise-induced cochlear neuropathy key to the generation of hyperacusis or tinnitus? *J. Neurophysiol.* 111, 552–564. doi: 10.1152/jn.00184.2013

Holdgraf, C. R., de Heer, W., Pasley, B., Rieger, J., Crone, N., Lin, J. J., et al. (2016). Rapid tuning shifts in human auditory cortex enhance speech intelligibility. *Nat. Commun.* 7:13654. doi: 10.1038/ncomms13654

Homma, N. Y., Happel, M. F. K., Nodal, F. R., Ohl, F. W., King, A. J., and Bajo, V. M. (2017). A role for auditory corticothalamic feedback in the perception of complex sounds. *J. Neurosci.* 37, 6149–6161. doi: 10.1523/JNEUROSCI.0397-17.2017

Homma, N. Y., Hullett, P. W., Atencio, C. A., and Schreiner, C. E. (2020). Auditory cortical plasticity dependent on environmental noise statistics. *Cell Rep.* 30, 4445–4458.e5. doi: 10.1016/j.celrep.2020.03.014

Hulse, S. H., MacDougall-Shackleton, S. A., and Wisniewski, A. B. (1997). Auditory scene analysis by songbirds: stream segregation of birdsong by European starlings (Sturnus vulgaris). *J. Comp. Psychol.* 111, 3–13. doi: 10.1037/0735-7036.111.1.3

Iriki, A., Tanaka, M., and Iwamura, Y. (1996). Attention-induced neuronal activity in the monkey somatosensory cortex revealed by pupillometrics. *Neurosci. Res.* 25, 173–181. doi: 10.1016/0168-0102(96)01043-7

Isaacson, J. S., and Scanziani, M. (2011). How inhibition shapes cortical activity. *Neuron* 72, 231–243. doi: 10.1016/j.neuron.2011.09.027

Jäger, K., and Kössl, M. (2016). Corticofugal modulation of DPOAEs in gerbils. *Hear. Res.* 332, 61–72. doi: 10.1016/j.heares.2015.11.008

Jeffress, L. A. (1948). A place theory of sound localization. *J. Comp. Physiol. Psychol.* 41, 35–39. doi: 10.1037/h0061495

Johannesen, P. T., Pérez-González, P., Kalluri, S., Blanco, J. L., and Lopez-Poveda, E. A. (2016). The influence of cochlear mechanical dysfunction, temporal processing deficits, and age on the intelligibility of audible speech in noise for hearing-impaired listeners. *Trends Hear.* 20:2331216516641055. doi: 10.1177/2331216516641055

Kadia, S. C., and Wang, X. (2003). Spectral integration in A1 of awake primates: neurons with single- and multipeaked tuning characteristics. *J. Neurophysiol.* 89, 1603–1622. doi: 10.1152/jn.00271.2001

Kato, H. K., Asinof, S. K., and Isaacson, J. S. (2017). Network-level control of frequency tuning in auditory cortex. *Neuron* 95, 412–423.e4. doi: 10.1016/j.neuron.2017.06.019

Kato, H. K., Gillet, S. N., and Isaacson, J. S. (2015). Flexible sensory representations in auditory cortex driven by behavioral relevance. *Neuron* 88, 1027–1039. doi: 10.1016/j.neuron.2015.10.024

Katzner, S., Busse, L., and Carandini, M. (2011). GABAA inhibition controls response gain in visual cortex. *J. Neurosci.* 31, 5931–5941. doi: 10.1523/jneurosci.5753-10.2011

Kauramäki, J., Jääskeläinen, I. P., and Sams, M. (2007). Selective attention increases both gain and feature selectivity of the human auditory cortex. *PLoS One* 2:e909. doi: 10.1371/journal.pone.0000909

Kayser, C., Petkov, C. I., Lippert, M., and Logothetis, N. K. (2005). Mechanisms for allocating auditory attention: an auditory saliency map. *Curr. Biol.* 15, 1943–1947. doi: 10.1016/j.cub.2005.09.040

Keller, C. H., Kaylegian, K., and Wehr, M. (2018). Gap encoding by parvalbumin-expressing interneurons in auditory cortex. *J. Neurophysiol.* 120, 105–114. doi: 10.1152/jn.00911.2017

Kerlin, J. R., Shahin, A. J., and Miller, L. M. (2010). Attentional gain control of ongoing cortical speech representations in a "Cocktail Party.". *J. Neurosci.* 30, 620–628. doi: 10.1523/JNEUROSCI.3631-09.2010

Khalighinejad, B., Herrero, J. L., Mehta, A. D., and Mesgarani, N. (2019). Adaptation of the human auditory cortex to changing background noise. *Nat. Commun.* 10:2509. doi: 10.1038/s41467-019-10611-4

Kim, H., Ährlund-Richter, S., Wang, X., Deisseroth, K., and Carlén, M. (2016). Prefrontal parvalbumin neurons in control of attention. *Cell* 164, 208–218. doi: 10.1016/j.cell.2015.11.038

King, A. J., and Walker, K. M. (2020). Listening in complex acoustic scenes. *Curr. Opin. Physiol.* 18, 63–72. doi: 10.1016/j.cophys.2020.09.001

Kohrman, D. C., Wan, G., Cassinotti, L., and Corfas, G. (2020). Hidden hearing loss: a disorder with multiple etiologies and mechanisms. *Cold Spring Harb. Perspect. Med.* 10:a035493. doi: 10.1101/cshperspect.a035493

Kraus, N., Bradlow, A. R., Cheatham, M. A., Cunningham, J., King, C. D., Koch, D. B., et al. (2000). Consequences of neural asynchrony: a case of auditory neuropathy. *J. Assoc. Res. Otolaryngol.* 1, 33–45. doi: 10.1007/s101620010004

Kujawa, S. G., and Liberman, M. C. (2009). Adding insult to injury: cochlear nerve degeneration after "Temporary" noise-induced hearing loss. *J. Neurosci.* 29, 14077–14085.

Kvale, M. N., and Schreiner, C. E. (2004). Short-term adaptation of auditory receptive fields to dynamic stimuli. *J. Neurophysiol.* 91, 604–612. doi: 10.1152/jn.00484.2003

Lee, A. M., Hoy, J. L., Bonci, A., Wilbrecht, L., Stryker, M. P., and Niell, C. M. (2014). Identification of a brainstem circuit regulating visual cortical state in parallel with locomotion. *Neuron* 83, 455–466. doi: 10.1016/j.neuron.2014.06.031

Lee, C.-C., and Middlebrooks, J. C. (2011). Auditory cortex spatial sensitivity sharpens during task performance. *Nat. Neurosci.* 14, 108–114. doi: 10.1038/nn.2713

Li, L., Ji, X., Liang, F., Li, Y., Xiao, Z., Tao, H. W., et al. (2014). A feedforward inhibitory circuit mediates lateral refinement of sensory representation in upper layer 2/3 of mouse primary auditory cortex. *J. Neurosci.* 34, 13670–13683. doi: 10.1523/JNEUROSCI.1516-14.2014

Li, L.-Y., Xiong, X. R., Ibrahim, L. A., Yuan, W., Tao, H. W., and Zhang, L. I. (2015). Differential receptive field properties of parvalbumin and somatostatin inhibitory neurons in mouse auditory cortex. *Cereb. Cortex* 25, 1782–1791. doi: 10.1093/cercor/bht417

Liberman, M. C. (1978). Auditory-nerve response from cats raised in a low-noise chamber. *J. Acoust. Soc. Am.* 63, 442–455. doi: 10.1121/1.381736

Liberman, M. C., and Kujawa, S. G. (2017). Cochlear synaptopathy in acquired sensorineural hearing loss: Manifestations and mechanisms. *Hear. Res.* 349, 138–147. doi: 10.1016/j.heares.2017.01.003

Lin, F. R., Yaffe, K., Xia, J., Xue, Q.-L., Harris, T. B., Purchase-Helzner, E., et al. (2013). Hearing loss and cognitive decline in older adults. *JAMA Int. Med.* 173, 293–299. doi: 10.1001/jamainternmed.2013.1868

Lingner, A., Pecka, M., Leibold, C., and Grothe, B. (2018). A novel concept for dynamic adjustment of auditory space. *Sci. Rep.* 8:8335. doi: 10.1038/s41598-018-26690-0

Llano, D. A., and Sherman, S. M. (2009). Differences in intrinsic properties and local network connectivity of identified layer 5 and layer 6 adult mouse auditory corticothalamic neurons support a dual corticothalamic projection hypothesis. *Cereb. Cortex* 19, 2810–2826. doi: 10.1093/cercor/bhp050

Lobarinas, E., Salvi, R., and Ding, D. (2013). Insensitivity of the audiogram to carboplatin induced inner hair cell loss in chinchillas. *Hear. Res.* 302, 113–120. doi: 10.1016/j.heares.2013.03.012

Lobarinas, E., Salvi, R., and Ding, D. (2020). Gap detection deficits in chinchillas with selective carboplatin-induced inner hair cell loss. *J. Assoc. Res. Otolaryngol.* 21, 475–483. doi: 10.1007/s10162-020-00744-5

Lohse, M., Bajo, V. M., King, A. J., and Willmore, B. D. B. (2020). Neural circuits underlying auditory contrast gain control and their perceptual implications. *Nat. Commun.* 11:324. doi: 10.1038/s41467-019-14163-5

Lopez Espejo, M., Schwartz, Z. P., and David, S. V. (2019). Spectral tuning of adaptation supports coding of sensory context in auditory cortex. *PLoS Comput. Biol.* 15:e1007430. doi: 10.1371/journal.pcbi.1007430

Ma, L., Micheyl, C., Yin, P., Oxenham, A. J., and Shamma, S. A. (2010). Behavioral measures of auditory streaming in ferrets (Mustela putorius). *J. Comp. Psychol.* 124, 317–330. doi: 10.1037/a0018273

Magnusson, A. K., Park, T. J., Pecka, M., Grothe, B., and Koch, U. (2008). Retrograde GABA signaling adjusts sound localization by balancing excitation and inhibition in the brainstem. *Neuron* 59, 125–137. doi: 10.1016/j.neuron.2008.05.011

Marian, V., Lam, T. Q., Hayakawa, S., and Dhar, S. (2018). Top-down cognitive and linguistic influences on the suppression of spontaneous otoacoustic emissions. *Front. Neurosci.* 12:378. doi: 10.3389/fnins.2018.00378

Maunsell, J. H. R. (2015). Neuronal mechanisms of visual attention. *Annu. Rev. Vis. Sci.* 1, 373–391. doi: 10.1146/annurev-vision-082114-035431

May, B. J., and Sachs, M. B. (1992). Dynamic range of neural rate responses in the ventral cochlear nucleus of awake cats. *J. Neurophysiol.* 68, 1589–1602. doi: 10.1152/jn.1992.68.5.1589

McGinley, M. J., David, S. V., and McCormick, D. A. (2015a). Cortical membrane potential signature of optimal states for sensory signal detection. *Neuron* 87, 179–192. doi: 10.1016/j.neuron.2015.05.038

McGinley, M. J., Vinck, M., Reimer, J., Batista-Brito, R., Zagha, E., Cadwell, C. R., et al. (2015b). Waking state: rapid variations modulate neural and behavioral responses. *Neuron* 87, 1143–1161. doi: 10.1016/j.neuron.2015.09.012

Meltzer, N. E., and Ryugo, D. K. (2006). Projections from auditory cortex to cochlear nucleus: a comparative analysis of rat and mouse. *Anat. Rec. A. Discov. Mol. Cell Evol. Biol.* 288, 397–408. doi: 10.1002/ar.a.20300

Mesgarani, N., and Chang, E. F. (2012). Selective cortical representation of attended speaker in multi-talker speech perception. *Nature* 485, 233–236. doi: 10.1038/nature11020

Mesgarani, N., David, S. V., Fritz, J. B., and Shamma, S. A. (2014b). Mechanisms of noise robust representation of speech in primary auditory cortex. *Proc. Natl. Acad. Sci. U.S.A.* 111, 6792–6797.

Mesgarani, N., Cheung, C., Johnson, K., and Chang, E. F. (2014a). Phonetic feature encoding in human superior temporal gyrus. *Science* 343, 1006–1010. doi: 10.1126/science.1245994

Mitchell, J. F., Sundberg, K. A., and Reynolds, J. H. (2007). Differential attention-dependent response modulation across cell classes in macaque visual area V4. *Neuron* 55, 131–141. doi: 10.1016/j.neuron.2007.06.018

Mitchell, J. F., Sundberg, K. A., and Reynolds, J. H. (2009). Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. *Neuron* 63, 879–888. doi: 10.1016/j.neuron.2009.09.013

Monaghan, J. J. M., Garcia-Lazaro, J. A., McAlpine, D., and Schaette, R. (2020). Hidden hearing loss impacts the neural representation of speech in background noise. *Curr. Biol.* 30, 4710–4721.e4. doi: 10.1016/j.cub.2020.09.046

Moore, A. K., and Wehr, M. (2013). Parvalbumin-expressing inhibitory interneurons in auditory cortex are well-tuned for frequency. *J. Neurosci.* 33, 13713–13723. doi: 10.1523/JNEUROSCI.0663-13.2013

Moreno-Bote, R., Beck, J., Kanitscheider, I., Pitkow, X., Latham, P., and Pouget, A. (2014). Information-limiting correlations. *Nat. Neurosci.* 17, 1410–1417. doi: 10.1038/nn.3807

Motanis, H., Seay, M. J., and Buonomano, D. V. (2018). Short-term synaptic plasticity as a mechanism for sensory timing. *Trends Neurosci.* 41, 701–711. doi: 10.1016/j.tins.2018.08.001

Nagel, K. I., and Doupe, A. J. (2006). Temporal processing and adaptation in the songbird auditory forebrain. *Neuron* 51, 845–859. doi: 10.1016/j.neuron.2006.08.030

Natan, R. G., Briguglio, J. J., Mwilambwe-Tshilobo, L., Jones, S. I., Aizenberg, M., Goldberg, E. M., et al. (2015). Complementary control of sensory adaptation by two types of cortical interneurons. *Elife* 4:e09868. doi: 10.7554/eLife.09868

Natan, R. G., Carruthers, I. M., Mwilambwe-Tshilobo, L., and Geffen, M. N. (2016). Gain control in the auditory cortex evoked by changing temporal correlation of sounds. *Cereb. Cortex* 27, 2385–2402. doi: 10.1093/cercor/bhw083

Natan, R. G., Rao, W., and Geffen, M. N. (2017). Cortical interneurons differentially shape frequency tuning following adaptation. *Cell Rep.* 21, 878–890. doi: 10.1016/j.celrep.2017.10.012

Nelken, I. (2014). Stimulus-specific adaptation and deviance detection in the auditory system: experiments and models. *Biol. Cybern* 108, 655–663. doi: 10.1007/s00422-014-0585-7

Nelson, A., Schneider, D. M., Takatoh, J., Sakurai, K., Wang, F., and Mooney, R. (2013). A circuit for motor cortical modulation of auditory cortical activity. *J. Neurosci.* 33, 14342–14353. doi: 10.1523/JNEUROSCI.2275-13.2013

Nelson, P. C., Smith, Z. M., and Young, E. D. (2009). Wide-dynamic-range forward suppression in marmoset inferior colliculus neurons is generated centrally and accounts for perceptual masking. *J. Neurosci.* 29, 2553–2562. doi: 10.1523/JNEUROSCI.5359-08.2009

Ni, J., Wunderle, T., Lewis, C. M., Desimone, R., Diester, I., and Fries, P. (2016). Gamma-rhythmic gain modulation. *Neuron* 92, 240–251. doi: 10.1016/j.neuron.2016.09.003

Noda, T., and Takahashi, H. (2019). Behavioral evaluation of auditory stream segregation in rats. *Neurosci. Res.* 141, 52–62. doi: 10.1016/j.neures.2018.03.007

Oberfeld, D., and Klöckner-Nowotny, F. (2016). Individual differences in selective attention predict speech identification at a cocktail party. *Elife* 5:e16747. doi: 10.7554/eLife.16747

Olsen, S. R., and Wilson, R. I. (2008). Lateral presynaptic inhibition mediates gain control in an olfactory circuit. *Nature* 452, 956–960. doi: 10.1038/nature06864

Olsen, S. R., Bortone, D. S., Adesnik, H., and Scanziani, M. (2012). Gain control by layer six in cortical circuits of vision. *Nature* 483, 47–52. doi: 10.1038/nature10835

Otazu, G. H., Tai, L.-H., Yang, Y., and Zador, A. M. (2009). Engaging in an auditory task suppresses responses in auditory cortex. *Nat. Neurosci.* 12, 646–654. doi: 10.1038/nn.2306

Overath, T., Kumar, S., von Kriegstein, K., and Griffiths, T. D. (2008). Encoding of spectral correlation over time in auditory cortex. *J. Neurosci.* 28, 13268–13273. doi: 10.1523/JNEUROSCI.4596-08.2008

Pakan, J. M., Lowe, S. C., Dylda, E., Keemink, S. W., Currie, S. P., Coutts, C. A., et al. (2016). Behavioral-state modulation of inhibition is context-dependent and cell type specific in mouse visual cortex. *eLife* 5:e14985. doi: 10.7554/eLife.14985

Park, H. L., O'Connell, J. E., and Thomson, R. G. (2003). A systematic review of cognitive decline in the general elderly population. *Int. J. Geriatr. Psychiatry.* 18, 1121–1134. doi: 10.1002/gps.1023

Park, T. J., Brand, A., Koch, U., Ikebuchi, M., and Grothe, B. (2008). Dynamic changes in level influence spatial coding in the lateral superior olive. *Hear. Res.* 238, 58–67. doi: 10.1016/j.heares.2007.10.009

Park, T. J., Grothe, B., Pollak, G. D., Schuller, G., and Koch, U. (1996). Neural delays shape selectivity to interaural intensity differences in the lateral superior olive. *J. Neurosci.* 16, 6554–6566.

Park, T. J., Klug, A., Holinstat, M., and Grothe, B. (2004). Interaural level difference processing in the lateral superior olive and the inferior colliculus. *J. Neurophysiol.* 92, 289–301. doi: 10.1152/jn.00961.2003

Passow, S., Westerhausen, R., Hugdahl, K., Wartenburger, I., Heekeren, H. R., Lindenberger, U., et al. (2014). Electrophysiological correlates of adult age differences in attentional control of auditory processing. *Cereb. Cortex* 24, 249–260. doi: 10.1093/cercor/bhs306

Pennington, J. R., and David, S. V. (2020). Complementary effects of adaptation and gain control on sound encoding in primary auditory cortex. *eNeuro* 7:ENEURO.205–ENEURO.220. doi: 10.1523/ENEURO.0205-20.2020

Peters, R. W., Moore, B. C., and Baer, T. (1998). Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people. *J. Acoust. Soc. Am.* 103, 577–587. doi: 10.1121/1.421128

Pfeffer, C. K., Xue, M., He, M., Huang, Z. J., and Scanziani, M. (2013). Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. *Nat. Neurosci.* 16, 1068–1076. doi: 10.1038/nn.3446

Phillips, D. P., and Hall, S. E. (2005). Psychophysical evidence for adaptation of central auditory processors for interaural differences in time and level. *Hear. Res.* 202, 188–199. doi: 10.1016/j.heares.2004.11.001

Phillips, D. P., Semple, M. N., Calford, M. B., and Kitzes, L. M. (1994). Level-dependent representation of stimulus frequency in cat primary auditory cortex. *Exp. Brain Res.* 102, 210–226. doi: 10.1007/BF00227510

Phillips, E. A. K., Schreiner, C. E., and Hasenstaub, A. R. (2017). Cortical interneurons differentially regulate the effects of acoustic context. *Cell Rep.* 20, 771–778. doi: 10.1016/j.celrep.2017.07.001

Phillips, E. A., and Hasenstaub, A. R. (2016). Asymmetric effects of activating and inactivating cortical interneurons. *Elife* 5:e18383. doi: 10.7554/eLife.18383

Pi, H.-J., Hangya, B., Kvitsiani, D., Sanders, J. I., Huang, Z. J., and Kepecs, A. (2013). Cortical interneurons that specialize in disinhibitory control. *Nature* 503, 521–524. doi: 10.1038/nature12676

Pienkowski, M., Tyler, R. S., Roncancio, E. R., Jun, H. J., Brozoski, T., Dauman, N., et al. (2014). A review of hyperacusis and future directions: part II. Measurement, mechanisms, and treatment. *Am. J. Audiol.* 23, 420–436. doi: 10.1044/2014_AJA-13-0037

Plack, C. J., Barker, D., and Prendergast, G. (2014). Perceptual consequences of "Hidden" hearing loss. *Trends Hear.* 18:2331216514550621. doi: 10.1177/2331216514550621

Popelar, J., Syka, J., and Berndt, H. (1987). Effect of noise on auditory evoked-responses in awake guinea-pigs. *Hear. Res.* 26, 239–247. doi: 10.1016/0378-5955(87)90060-8

Rabinowitz, N. C., Willmore, B. D. B., King, A. J., and Schnupp, J. W. H. (2013). Constructing noise-invariant representations of sound in the auditory pathway. *PLoS Biol.* 11:e1001710. doi: 10.1371/journal.pbio.1001710

Rabinowitz, N. C., Willmore, B. D. B., Schnupp, J. W. H., and King, A. J. (2011). Contrast gain control in auditory cortex. *Neuron* 70, 1178–1191. doi: 10.1016/j.neuron.2011.04.030

Radziwon, K., and Salvi, R. (2020). Using auditory reaction time to measure loudness growth in rats. *Hear. Res.* 395:108026. doi: 10.1016/j.heares.2020.108026

Radziwon, K., Auerbach, B. D., Ding, D., Liu, X., Chen, G. D., and Salvi, R. (2019). Noise-induced loudness recruitment and hyperacusis: insufficient central gain in auditory cortex and amygdala. *Neuroscience* 422, 212–227. doi: 10.1016/j.neuroscience.2019.09.010

Ralli, M., Greco, A., De Vincentiis, M., Sheppard, A., Cappelli, G., Neri, I., et al. (2019). Tone-in-noise detection deficits in elderly patients with clinically normal hearing. *Am. J. Otolaryngol.* 40, 1–9. doi: 10.1016/j.amjoto.2018.09.012

Reed, D. K., Chait, M., Tóth, B., Winkler, I., and Shinn-Cunningham, B. (2020). Spatial cues can support auditory figure-ground segregation. *J. Acoust. Soc. Am.* 147:3814. doi: 10.1121/10.0001387

Resnik, J., and Polley, D. B. (2017). Fast-spiking GABA circuit dynamics in the auditory cortex predict recovery of sensory processing following peripheral nerve damage. *Elife* 6:e21452. doi: 10.7554/eLife.21452

Resnik, J., and Polley, D. B. (2021). Cochlear neural degeneration disrupts hearing in background noise by increasing auditory cortex internal noise. *Neuron* 109, 984–996.e4. doi: 10.1016/j.neuron.2021.01.015

Robinson, B. L., and McAlpine, D. (2009). Gain control mechanisms in the auditory pathway. *Curr. Opin. Neurobiol.* 19, 402–407. doi: 10.1016/j.conb.2009.07.006

Robinson, B. L., Harper, N. S., and McAlpine, D. (2016). Meta-adaptation in the auditory midbrain under cortical influence. *Nat. Commun.* 7:13442. doi: 10.1038/ncomms13442

Rocchi, F., and Ramachandran, R. (2018). Neuronal adaptation to sound statistics in the inferior colliculus of behaving macaques does not reduce the effectiveness of the masking noise. *J. Neurophysiol.* 120, 2819–2833. doi: 10.1152/jn.00875.2017

Rocchi, F., and Ramachandran, R. (2020). Foreground stimuli and task engagement enhance neuronal adaptation to background noise in the inferior colliculus of macaques. *J. Neurophysiol.* 124, 1315–1326. doi: 10.1152/jn.00153.2020

Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 336, 367–373. doi: 10.1098/rstb.1992.0070

Sachs, M. B., and Abbas, P. J. (1974). Rate versus level functions for auditory-nerve fibers in cats: tone-burst stimuli. *J. Acoust. Soc. Am.* 56, 1835–1847. doi: 10.1121/1.1903521

Sadagopan, S., and Wang, X. (2008). Level invariant representation of sounds by populations of neurons in primary auditory cortex. *J. Neurosci.* 28, 3415–3426. doi: 10.1523/JNEUROSCI.2743-07.2008

Saderi, D., Buran, B. N., and David, S. V. (2020). Streaming of repeated noise in primary and secondary fields of auditory cortex. *J. Neurosci.* 40, 3783–3798. doi: 10.1523/JNEUROSCI.2105-19.2020

Salvi, R., Sun, W., Ding, D., Chen, G. D., Lobarinas, E., Wang, J., et al. (2016). Inner hair cell loss disrupts hearing and cochlear function leading to sensory deprivation and enhanced central auditory gain. *Front. Neurosci.* 10:621. doi: 10.3389/fnins.2016.00621

Santoro, R., Moerel, M., Martino, F. D., Goebel, R., Ugurbil, K., Yacoub, E., et al. (2014). Encoding of natural sounds at multiple spectral and temporal resolutions in the human auditory cortex. *PLoS Comput. Biol.* 10:e1003412. doi: 10.1371/journal.pcbi.1003412

Schaette, R., and McAlpine, D. (2011). Tinnitus with a normal audiogram: physiological evidence for hidden hearing loss and computational model. *J. Neurosci.* 31, 13452–13457. doi: 10.1523/jneurosci.2156-11.2011

Schmiedt, R. A. (1984). Acoustic injury and the physiology of hearing. *J. Acoust. Soc. Am.* 76, 1293–1317. doi: 10.1121/1.391446

Schneider, D. M., and Woolley, S. M. N. (2013). Sparse and background-invariant coding of vocalizations in auditory scenes. *Neuron* 79, 141–152. doi: 10.1016/j.neuron.2013.04.038

Schneider, D. M., Nelson, A., and Mooney, R. (2014). A synaptic and circuit basis for corollary discharge in the auditory cortex. *Nature* 513, 189–194. doi: 10.1038/nature13724

Schönwiesner, M., Rübsamen, R., and von Cramon, D. Y. (2005). Hemispheric asymmetry for spectral and temporal processing in the human antero-lateral auditory belt cortex. *Eur. J. Neurosci.* 22, 1521–1528. doi: 10.1111/j.1460-9568.2005.04315.x

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

Shastri, U., Mythri, H. M., and Kumar, U. A. (2014). Descending auditory pathway and identification of phonetic contrast by native listeners. *J. Acoust. Soc. Am.* 135, 896–905. doi: 10.1121/1.4861350

Shew, W. L., Yang, H., Yu, S., Roy, R., and Plenz, D. (2011). Information capacity and transmission are maximized in balanced cortical networks with neuronal avalanches. *J. Neurosci.* 31, 55–63. doi: 10.1523/JNEUROSCI.4637-10.2011

Shinn-Cunningham, B. G., and Best, V. (2008). Selective attention in normal and impaired hearing. *Trends Amplif.* 12, 283–299. doi: 10.1177/1084713808325306

Shu, Y., Hasenstaub, A., Badoual, M., Bal, T., and McCormick, D. A. (2003). Barrages of synaptic activity control the gain and sensitivity of cortical neurons. *J. Neurosci.* 23, 10388–10401.

Simel, D. L., Bagai, A., Thavendiranathan, P., and Detsky, A. S. (2016). "Hearing impairment," in *The Rational Clinical Examination: Evidence-Based Clinical Diagnosis*, eds D. L. Simel and D. Rennie (New York, NY: McGraw-Hill Education).

Simpson, A. J. R., Harper, N. S., Reiss, J. D., and McAlpine, D. (2014). Selective adaptation to "Oddball" sounds by the human auditory system. *J. Neurosci.* 34, 1963–1969. doi: 10.1523/JNEUROSCI.4274-13.2013

Singh, N. C., and Theunissen, F. E. (2003). Modulation spectra of natural sounds and ethological theories of auditory processing. *J. Acoust. Soc. Am.* 114, 3394–3411. doi: 10.1121/1.1624067

Sohal, V. S., Zhang, F., Yizhar, O., and Deisseroth, K. (2009). Parvalbumin neurons and gamma rhythms enhance cortical circuit performance. *Nature* 459, 698–702. doi: 10.1038/nature07991

Spankovich, C., Gonzalez, V. B., Su, D., and Bishop, C. E. (2018). Self reported hearing difficulty, tinnitus, and normal audiometric thresholds, the National Health and Nutrition Examination Survey 1999-2002. *Hear. Res.* 358, 30–36. doi: 10.1016/j.heares.2017.12.001

Spoendlin, H. (1972). Innervation densities of the cochlea. *Acta Otolaryngol.* 73, 235–248. doi: 10.3109/00016487209138937

Srinivasan, S., Keil, A., Stratis, K., Woodruff Carr, K. L., and Smith, D. W. (2012). Effects of cross-modal selective attention on the sensory periphery: cochlear sensitivity is altered by selective attention. *Neuroscience* 223, 325–332. doi: 10.1016/j.neuroscience.2012.07.062

Stange, A., Myoga, M. H., Lingner, A., Ford, M. C., Alexandrova, O., Felmy, F., et al. (2013). Adaptation in sound localization: from GABAB receptor–mediated synaptic modulation to perception. *Nat. Neurosci.* 16, 1840–1847. doi: 10.1038/nn.3548

Starr, A., Picton, T. W., Sininger, Y., Hood, L. J., and Berlin, C. I (1996). Auditory neuropathy. *Brain* 119(Pt. 3), 741–753.

Suga, N., and Manabe, T. (1982). Neural basis of amplitude-spectrum representation in auditory cortex of the mustached bat. *J. Neurophysiol.* 47, 225–255. doi: 10.1152/jn.1982.47.2.225

Suga, N., O'Neill, W. E., Kujirai, K., and Manabe, T. (1983). Specificity of combination-sensitive neurons for processing of complex biosonar signals in auditory cortex of the mustached bat. *J. Neurophysiol.* 49, 1573–1626. doi: 10.1152/jn.1983.49.6.1573

Syka, J. (2002). Plastic changes in the central auditory system after hearing loss, restoration of function, and during learning. *Physiol. Rev.* 82, 601–636. doi: 10.1152/physrev.00002.2002

Syka, J., Rybalko, N., and Popelar, J. (1994). Enhancement of the auditory-cortex evoked-responses is awake guinea-pigs after noise exposure. *Hear. Res.* 78, 158–168. doi: 10.1016/0378-5955(94)90021-3

Tiesinga, P. H., Fellous, J.-M., Salinas, E., José, J. V., and Sejnowski, T. J. (2004). Inhibitory synchrony as a mechanism for attentional gain modulation. *J. Physiol. Paris* 98, 296–314. doi: 10.1016/j.jphysparis.2005.09.002

Tiesinga, P., Fellous, J.-M., and Sejnowski, T. J. (2008). Regulation of spike timing in visual cortical circuits. *Nat. Rev. Neurosci.* 9, 97–107. doi: 10.1038/nrn2315

Tremblay, K. L., Pinto, A., Fischer, M. E., Klein, B. E. K., Klein, R., Levy, S., et al. (2015). Self-reported hearing difficulties among adults with normal audiograms: the beaver dam offspring study. *Ear. Hear.* 36, e290–e299. doi: 10.1097/AUD.0000000000000195

Ulanovsky, N., Las, L., and Nelken, I. (2003). Processing of low-probability sounds by cortical neurons. *Nat. Neurosci.* 6, 391–398. doi: 10.1038/nn1032

Viemeister, N. F. (1988). Intensity coding and the dynamic range problem. *Hear. Res.* 34, 267–274. doi: 10.1016/0378-5955(88)90007-x

Vigneault-MacLean, B. K., Hall, S. E., and Phillips, D. P. (2007). The effects of lateralized adaptors on lateral position judgements of tones within and across frequency channels. *Hear. Res.* 224, 93–100. doi: 10.1016/j.heares.2006.12.001

Wang, J., Powers, N. L., Hofstetter, P., Trautwein, P., Ding, D., and Salvi, R. (1997). Effects of selective inner hair cell loss on auditory nerve fiber threshold, tuning and spontaneous and driven discharge rate. *Hear. Res.* 107, 67–82.

Wardak, C., Ibos, G., Duhamel, J.-R., and Olivier, E. (2006). Contribution of the monkey frontal eye field to covert visual attention. *J. Neurosci.* 26, 4228–4235. doi: 10.1523/JNEUROSCI.3336-05.2006

Watkins, P. V., and Barbour, D. L. (2008). Specialized neuronal adaptation for preserving input sensitivity. *Nat. Neurosci.* 11, 1259–1261. doi: 10.1038/nn.2201

Watkins, P. V., and Barbour, D. L. (2011). Level-tuned neurons in primary auditory cortex adapt differently to loud versus soft sounds. *Cereb. Cortex* 21, 178–190. doi: 10.1093/cercor/bhq079

Wehr, M., and Zador, A. M. (2003). Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. *Nature* 426, 442–446. doi: 10.1038/nature02116

Wen, B., Wang, G. I., Dean, I., and Delgutte, B. (2009). Dynamic range adaptation to sound level statistics in the auditory nerve. *J. Neurosci.* 29, 13797–13808. doi: 10.1523/JNEUROSCI.5610-08.2009

Williamson, R. S., and Polley, D. B. (2019). Parallel pathways for sound processing and functional connectivity among layer 5 and 6 auditory corticofugal neurons. *Elife* 8:e42974. doi: 10.7554/eLife.42974

Willmore, B. D., Cooke, J. E., and King, A. J. (2014). Hearing in noisy environments: noise invariance and contrast gain control. *J. Physiol.* 592, 3371–3381. doi: 10.1113/jphysiol.2014.274886

Wilson, N. R., Runyan, C. A., Wang, F. L., and Sur, M. (2012). Division and subtraction by distinct cortical inhibitory networks in vivo. *Nature* 488, 343–348. doi: 10.1038/nature11347

Winer, J. A., Chernock, M. L., Larue, D. T., and Cheung, S. W. (2002). Descending projections to the inferior colliculus from the posterior thalamus and the auditory cortex in rat, cat, and monkey. *Hear. Res.* 168, 181–195. doi: 10.1016/s0378-5955(02)00489-6

Winkler, I., Denham, S. L., and Nelken, I. (2009). Modeling the auditory scene: predictive regularity representations and perceptual objects. *Trends Cogn. Sci.* 13, 532–540. doi: 10.1016/j.tics.2009.09.003

Winkowski, D. E., Bandyopadhyay, S., Shamma, S. A., and Kanold, P. O. (2013). Frontal cortex activation causes rapid plasticity of auditory cortical processing. *J. Neurosci.* 33, 18134–18148. doi: 10.1523/JNEUROSCI.0180-13.2013

Wood, K. C., Blackwell, J. M., and Geffen, M. N. (2017). Cortical inhibitory interneurons control sensory processing. *Curr. Opin. Neurobiol.* 46, 200–207. doi: 10.1016/j.conb.2017.08.018

Wu, P., Liberman, L., Bennett, K., de Gruttola, V., O'Malley, J., and Liberman, M. (2019). Primary neural degeneration in the human cochlea: evidence for hidden hearing loss in the aging ear. *Neuroscience* 407, 8–20. doi: 10.1016/j.neuroscience.2018.07.053

Xiao, Z., and Suga, N. (2002). Modulation of cochlear hair cells by the auditory cortex in the mustached bat. *Nat. Neurosci.* 5, 57–63. doi: 10.1038/nn786

Yan, J., and Suga, N. (1996). Corticofugal modulation of time-domain processing of biosonar information in bats. *Science* 273, 1100–1103. doi: 10.1126/science.273.5278.1100

Yang, H., and Xu-Friedman, M. A. (2009). Impact of synaptic depression on spike timing at the endbulb of Held. *J. Neurophysiol.* 102, 1699–1710. doi: 10.1152/jn.00072.2009

Yang, S., Weiner, B. D., Zhang, L. S., Cho, S.-J., and Bao, S. (2011). Homeostatic plasticity drives tinnitus perception in an animal model. *Proc. Natl. Acad. Sci. U.S.A.* 108, 14974–14979. doi: 10.1073/pnas.1107998108

Yizhar, O., Fenno, L. E., Prigge, M., Schneider, F., Davidson, T. J., O'Shea, D. J., et al. (2011). Neocortical excitation/inhibition balance in information processing and social dysfunction. *Nature* 477, 171–178. doi: 10.1038/nature10360

Young, E. D., and Barta, P. E. (1986). Rate responses of auditory nerve fibers to tones in noise near masked threshold. *J. Acoust. Soc. Am.* 79, 426–442. doi: 10.1121/1.393530

Zatorre, R. J., and Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cereb. Cortex* 11, 946–953. doi: 10.1093/cercor/11.10.946

Zeng, F. G. (2013). An active loudness model suggesting tinnitus as increased central noise and hyperacusis as increased nonlinear gain. *Hear. Res.* 295, 172–179. doi: 10.1016/j.heares.2012.05.009

Zeng, F. G., Kong, Y. Y., Michalewski, H. J., and Starr, A. (2005). Perceptual consequences of disrupted auditory nerve activity. *J. Neurophysiol.* 93, 3050–3063. doi: 10.1152/jn.00985.2004

Zhang, L. I., Tan, A. Y. Y., Schreiner, C. E., and Merzenich, M. M. (2003). Topography and synaptic shaping of direction selectivity in primary auditory cortex. *Nature* 424, 201–205. doi: 10.1038/nature01796

Zhang, S., Xu, M., Kamigaki, T., Do, J. P. H., Chang, W.-C., Jenvay, S., et al. (2014). Long-range and local circuits for top-down modulation of visual cortex processing. *Science* 345, 660–665. doi: 10.1126/science.1254126

Zhang, Y., Hakes, J. J., Bonfield, S. P., and Yan, J. (2005). Corticofugal feedback for auditory midbrain plasticity elicited by tones and electrical stimulation of basal forebrain in mice. *Eur. J. Neurosci.* 22, 871–879. doi: 10.1111/j.1460-9568.2005.04276.x

Zhang, Y., Suga, N., and Yan, J. (1997). Corticofugal modulation of frequency processing in bat auditory system. *Nature* 387, 900–903. doi: 10.1038/43180

Zhou, M., Liang, F., Xiong, X. R., Li, L., Li, H., Xiao, Z., et al. (2014). Scaling down of balanced excitation and inhibition by active behavioral states in auditory cortex. *Nat. Neurosci.* 17, 841–850. doi: 10.1038/nn.3701

Zhuang, X., Wong, N. F., Sun, W., and Xu-Friedman, M. A. (2020). Mechanisms and functional consequences of presynaptic homeostatic plasticity at auditory nerve synapses. *J. Neurosci.* 40, 6896–6909. doi: 10.1523/JNEUROSCI.1175-19.2020

Zion Golumbic, E. M., Ding, N., Bickel, S., Lakatos, P., Schevon, C. A., McKhann, G. M., et al. (2013). Mechanisms underlying selective neuronal tracking of attended speech at a "Cocktail Party.". *Neuron* 77, 980–991. doi: 10.1016/j.neuron.2012.12.037

Zohary, E., Shadlen, M. N., and Newsome, W. T. (1994). Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* 370, 140–143. doi: 10.1038/370140a0

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 14 March 2022
doi: 10.3389/fnins.2022.840983

Predicting the Influence of Axon Myelination on Sound Localization Precision Using a Spiking Neural Network Model of Auditory Brainstem

*Ben-Zheng Li[1,2,3,4], Sio Hang Pun[3], Mang I. Vai[3,4], Tim C. Lei[1,2] and Achim Klug[1]**

[1] Department of Physiology and Biophysics, University of Colorado Anschutz Medical Campus, Aurora, CO, United States, [2] Department of Electrical Engineering, University of Colorado, Denver, Denver, CO, United States, [3] State Key Laboratory of Analog and Mixed Signal Very-Large-Scale Integration (VLSI), University of Macau, Taipa, Macau SAR, China, [4] Department of Electrical and Computer Engineering, Faculty of Science and Technology, University of Macau, Taipa, Macau SAR, China

Spatial hearing allows animals to rapidly detect and localize auditory events in the surrounding environment. The auditory brainstem plays a central role in processing and extracting binaural spatial cues through microsecond-precise binaural integration, especially for detecting interaural time differences (ITDs) of low-frequency sounds at the medial superior olive (MSO). A series of mechanisms exist in the underlying neural circuits for preserving accurate action potential timing across multiple fibers, synapses and nuclei along this pathway. One of these is the myelination of afferent fibers that ensures reliable and temporally precise action potential propagation in the axon. There are several reports of fine-tuned myelination patterns in the MSO circuit, but how specifically myelination influences the precision of sound localization remains incompletely understood. Here we present a spiking neural network (SNN) model of the Mongolian gerbil auditory brainstem with myelinated axons to investigate whether different axon myelination thicknesses alter the sound localization process. Our model demonstrates that axon myelin thickness along the contralateral pathways can substantially modulate ITD detection. Furthermore, optimal ITD sensitivity is reached when the MSO receives contralateral inhibition via thicker myelinated axons compared to contralateral excitation, a result that is consistent with previously reported experimental observations. Our results suggest specific roles of axon myelination for extracting temporal dynamics in ITD decoding, especially in the pathway of the contralateral inhibition.

Keywords: sound localization, auditory brainstem, medial superior olive, myelin alteration, interaural time difference, spiking neural network, computational model

OPEN ACCESS

Edited by:
Dong Song,
University of Southern California, United States

Reviewed by:
Petr Marsalek,
Charles University, Czechia
Joshua H. Goldwyn,
Swarthmore College, United States

****Correspondence:***
Achim Klug
achim.klug@cuanschutz.edu

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience

Received: *21 December 2021*
Accepted: *18 February 2022*
Published: *14 March 2022*

Citation:
Li B-Z, Pun SH, Vai MI, Lei TC and Klug A (2022) Predicting the Influence of Axon Myelination on Sound Localization Precision Using a Spiking Neural Network Model of Auditory Brainstem. Front. Neurosci. 16:840983. doi: 10.3389/fnins.2022.840983

INTRODUCTION

In the mammalian brain, the precise temporal information encoded in action potentials and trains of action potentials is one major mechanism contributing to accurate neural integration and information processing. Such temporal precision is especially crucial for the localization process of sound sources in the auditory brainstem, especially for low-frequency sound sources. This

process works through the microsecond-precise integration of signals between the two ears such that whenever the temporal precision is even slightly altered, sound localization accuracy suffers significantly (reviewed in Grothe, 2003; Tollin and Yin, 2009; Grothe et al., 2010). Myelination of afferent fibers is a key mechanism in mammalian brains to ensure reliable, energy-efficient and temporally precise action potential propagation, and therefore, myelination is tightly controlled and actively managed in the brain (reviewed in Debanne, 2004). It is therefore not unexpected that such mechanisms have been described in the sound localization pathway as well (Ford et al., 2015; Seidl and Rubel, 2016; Stange-Marten et al., 2017).

In the auditory brainstem, sound localization along the azimuth is accomplished by the two principal localization nuclei, the lateral and the medial superior olive (LSO and MSO, respectively). High-frequency sounds are localized in the LSO by calculating the interaural level difference (ILD; Boudreau and Tsuchitani, 1968), while low-frequency sounds are localized in the MSO by calculating the interaural time difference (ITD; Goldberg and Brown, 1969). Mammals, including human listeners, are typically capable of resolving two sound sources that are just a few degrees separated from each other, by resolving ITDs as small as several microseconds (Grothe et al., 2010). How exactly the sound localization circuit can accomplish this extraordinary computational result within a set of fixed exterior constraints has been the subject of a number of studies. The speed of sound in air, as well as an animal's head size, dictate particular ITDs. For different species with different head sizes, the relationship between ITD and corresponding spatial angle varies, but for most species and most spatial angles, ITDs are one to two orders of magnitude shorter than the duration of an action potential. To operate at such temporal precision, several neural mechanisms along the afferent pathways have been described, including kinetically fast ion channels, large and electrically compact synapses, tuning of cochlear, synaptic, post-synaptic, and transmission delays (reviewed in Trussell, 1999; Grothe, 2003; Tollin and Yin, 2009; Grothe et al., 2010). Some of these mechanisms remain plastic during an individual's lifetime, allowing for a recalibration of the network, for example in response to hearing loss. Other mechanisms, such as axon myelination, are plastic during a period in which the animal's head is still growing and thus ITDs are changing, but do not re-calibrate after adulthood (Sinclair et al., 2017). Once calibrated, the various inputs to MSO neurons show different axon myelination patterns, which precisely preserve but also actively control the timing of action potentials propagated in these afferents, thereby controlling the sound localization process at MSO neurons.

In this study, we investigate contralateral excitatory and inhibitory inputs to the MSO, which show different axon myelination patterns (Morest, 1968; Grothe et al., 2010). The contralateral inhibitory pathway consists of axons with thicker layers of myelin, resulting in higher conduction velocities, compared to those of the ipsilateral excitatory pathway. Experimental results suggest that axonal myelination may be specifically adapted for tuning the input timing to the MSO, thereby actively contributing to spatial hearing perception (Schwartz, 1992; Ford et al., 2015; Seidl and Rubel, 2016; Stange-Marten et al., 2017). However, axon myelination as a factor in circuit modeling is underexplored and is simply included as a constant in most models—probably due to our still incomplete understanding of the structure-function relationships (Nave, 2010). In this study, the role of myelin morphology as a contributing factor to the MSO sound localization circuit is specifically explored.

Prior modeling efforts of ITD coding at MSO primarily focused on the effects of post-synaptic integration of MSO neurons (Brand et al., 2002; Zhou et al., 2005; Leibold, 2010; Brughera et al., 2013; Myoga et al., 2014). Other studies modeled the axonal propagation time as a constant delay, not including axonal morphology (Wang et al., 2014; Encke and Hemmert, 2018). On the other hand, some studies included action potential timing difference by varying axonal propagation delays (Glackin et al., 2010; Pan et al., 2021). These models, however, were based on the Jeffress model of a delay line structure (Jeffress, 1948), which is anatomically inconsistent with neural inhibition observed in mammalian ITD circuits (Brand et al., 2002; Grothe and Pecka, 2014; Franken et al., 2015).

This study employed a spiking neural network (SNN) model to investigate how axonal structure and synaptic adaptation between the excitatory and inhibitory inputs to MSO can affect ITD decoding. Specifically, the myelination thickness and the synaptic conductance were modeled in detail and compared with pure tone and natural sound stimulation regarding ITD coding accuracy and minimum temporal discrimination. Based on this SNN model and our decoding analysis, we found that the axon myelination patterns of both contralateral excitatory and inhibitory pathways can significantly modulate ITD decoding. The variation of myelin thickness, which results in conduction velocity variations along the excitatory pathways, can significantly shift the ITD tuning curve. On the other hand, axonal myelination and synaptic strength variations on the inhibitory pathway can significantly influence ITD sensitivity and precision.

MATERIALS AND METHODS

Neuron and Synapse Model

The spiking neurons were modeled under a conductance-based leaky integrate-and-fire scheme. The membrane potential (v_m) of a spiking neuron was described by the following first-order differential equation:

$$C_m \frac{dv_m(t)}{dt} = g_l\left(E_l - v_m\left(t\right)\right) + I_{syn}(t),$$

where I_{syn} is the total synaptic current comprising an excitatory and an inhibitory component:

$$I_{syn}(t) = g_e(t)\left(E_e - v_m\left(t\right)\right) + g_i(t)(E_i - v_m\left(t\right)),$$

$$\tau_e \frac{dg_e\left(t\right)}{dt} = -g_e(t),$$

$$\tau_i \frac{dg_i(t)}{dt} = -g_i(t).$$

Descriptions of the parameters and the values used in the simulation are listed in **Table 1**. The excitatory and inhibitory synaptic conductances were modeled as first-order time decaying parameters with lifetimes of τ_e and τ_i. When an action potential arrived at the presynaptic membrane, the conductance increased by Δg_{exci} or Δg_{inhi}, and subsequently decayed as described by the time constants. The spiking neuron will elicit an action potential when the membrane potential $v_m(t)$ reaches the firing threshold V_{th}, and the membrane potential $v_m(t)$ is then reset to the resting potential V_{reset} after a short refractory period τ_{ref}.

Network Architecture

The architecture of the proposed SNN sound localization model (**Figure 1A**) consists of a left and a right MSOs and related afferent nuclei, including cochlear nucleus and trapezoid body. The cochlea first encodes acoustic stimuli, which are then sent as action potentials into the model through the auditory nerve (AN) and received by the Cochlear Nuclei (CN). At the CN, the Spherical Bushy Cells (SBCs) innervate the two MSOs bilaterally, and Globular Bushy Cells (GBCs) innervate the contralateral Medial Nucleus of the Trapezoid Body (MNTB) as well as the ipsilateral Lateral Nucleus of the Trapezoid Body (LNTB). Under this architecture, MSO cells receive bilateral excitation from the SBCs, ipsilateral inhibition from the LNTB and contralateral inhibition from the MNTB.

Every neural population (AN, GBC, MSO, etc.) in each hemisphere contained 1,000 spiking neurons that stochastically connected to neurons in other populations based on a connection probability (p_{connect}). This stochastic connectivity introduced heterogeneity of input sources across all modeled neurons and rich diversity of neural responses, which was comparable to biological neural circuits. The synaptic connections were also associated with an axonal transmission delay (t_{trans}) and a synaptic delay (t_{syn}). These delays were normally distributed with a standard deviation of 0.05 ms. The transmission delays of the connections that traveled across the midline (SBC to contralateral MSO and GBC to contralateral MNTB) were derived from the corresponding conduction velocities and myelin thicknesses. The SNN model was implemented using the Brian2 simulator (Stimberg et al., 2019) and was simulated on a supercomputing cluster (RMACC Summit, University of Colorado Boulder). The simulation was repeated twenty times with different random seeds to compensate for random effects induced during stimuli encoding, network building, and decoding. Averaged metrics from all random permutations are reported as results. The specific parameters of the connections were adopted from previous studies (Brand et al., 2002; Spirou et al., 2005; Couchman et al., 2010; Roberts et al., 2013, 2014; Encke and Hemmert, 2018) and are listed in **Table 2**.

Stimulus Encoding

For sound stimulation to the cochlea, two spike generators were used at the left and right AN to encode the sound signals into cochlear action potential responses. For pure tone stimulation, sinusoidal sound waves of 300 Hz with a 50 dB sound pressure level (SPL) were used, except where otherwise indicated. The envelope duration of the sinusoidal wave was 100 ms with 20 ms ramp-up and ramp-down periods and sampled at 100 kHz. For natural sound stimulation, sound samples with dominant frequencies ranging from 318 to 546 Hz were created based on 60 bird song clips of the long-eared owl collected from the Xeno-canto project.[1] The bird song clips were adjusted to 50 dB SPL with 20 ms ramp-up and ramp-down, and were up-sampled from 44.1 to 100 kHz. Much of the previously published physiological data which informed our model were recorded in Mongolian gerbils. Thus, vocalizations of an owl species—a predator of this species—seemed appropriate.

Acoustic stimuli were sent to two ears with ITDs ranging from −1 to +1 ms with step sizes in log-scale. Although the simulated ITD range was beyond the biologically relevant ITD range of gerbils, typically ± 130 μs, this exceeded range was chosen to increase the comparability of our simulated results to existing physiological recordings using these broader ITD ranges (Pecka et al., 2008; Franken et al., 2015). ITD is defined as the onset time difference of the same sound between the left and the right ear, with positive ITDs defined as sound leading at the right ear. For each ITD, the stimulation was repeated ten times for pure tones and eight times for natural sounds. A peripheral hearing model (Zilany et al., 2014; Rudnicki et al., 2015) was used to generate action potentials from the sound waves for the AN spike generators. The simulated AN was configured as a composition of 60% high spontaneous firing rate fibers, 20% medium spontaneous firing rate fiber, and 20% low spontaneous firing rate fibers.

[1] www.xeno-canto.org

TABLE 1 | List of parameters for neurons and synapses.

Neuron parameter	Value	Description	Synapse parameter	Value	Description
C_m	70 pF	Membrane capacitance	E_e	0 mV	Excitatory reversal potential
V_{th}	−50 mV	Threshold potential	E_i	−70 mV	Inhibitory reversal potential
V_{reset}	−55.8 mV	Reset potential	Δg_{exci}	15 nS	Excitatory postsynaptic conductance increment
τ_{ref}	5 ms	Refractory period	Δg_{inhi}	75 nS	Inhibitory postsynaptic conductance increment
E_l	−55.9 mV	Leaky reversal potential	τ_e	0.23 ms	Excitatory time constant
g_l	13 nS	Leaky conductance	τ_i	2 ms	Inhibitory time constant

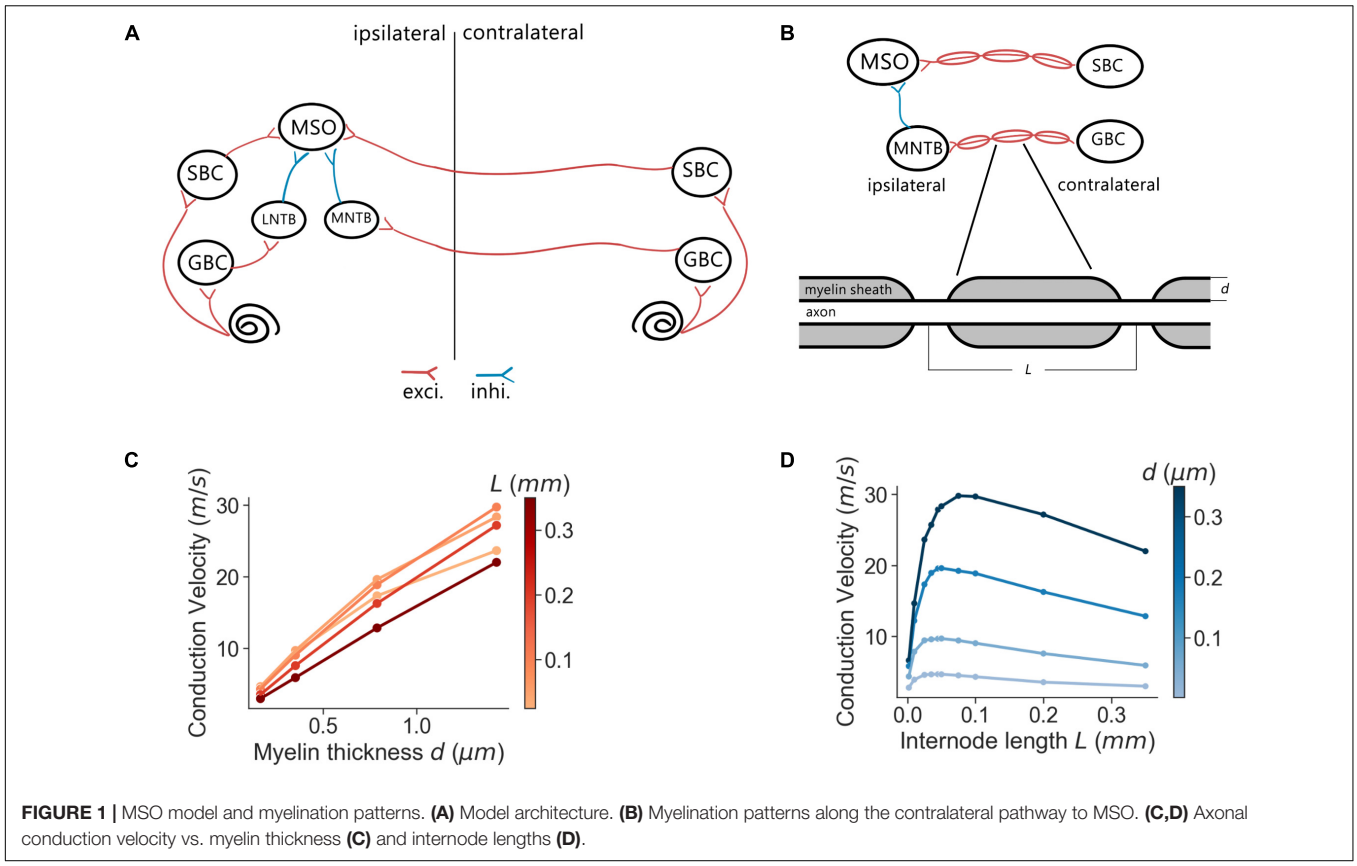

FIGURE 1 | MSO model and myelination patterns. **(A)** Model architecture. **(B)** Myelination patterns along the contralateral pathway to MSO. **(C,D)** Axonal conduction velocity vs. myelin thickness **(C)** and internode lengths **(D)**.

TABLE 2 | List of parameters for neurons and synapses.

Population	Type	Connection	$p_{connect}$	$\bar{t}_{trans}$ (ms)	$\bar{t}_{syn}$ (ms)
AN	Spike generator	AN → ipsi. GBC	4%	0.2	0.6
		AN → ipsi. SBC	4%	0.2	0.6
SBC	Excitatory	SBC → ipsi. MSO	0.6%	0.6	0.6
		SBC → cont. MSO	0.6%	*various*[b]	0.6
GBC	Excitatory	GBC → ipsi. LNTB	0.3%	0.3	0.3
		GBC → cont. MNTB[a]	one-to-one	*various*[c]	0.2
MNTB	Inhibitory	MNTB → ipsi. MSO	0.3%	0.1	0.3
LNTB	Inhibitory	LNTB → ipsi. MSO	0.3%	0.1	0.3
MSO	Excitatory				

[a] *The Calyx of Held with* Δg_{inhi} *= 250 Ns.*
[b] *Depends on the SBC myelin thickness* d_{SBC}.
[c] *Depends on the GBC myelin thickness* d_{GBC}.

Axon Myelination and Conduction Velocity

The conduction velocity of myelinated axons was computed using the multi-compartment axon model of Halter and Clark (1991) following the methods and parameters proposed by Ford et al. (2015) for simulating GBC and SBC fibers. Briefly, axons were compartmentalized into nodes and internodes represented by Hodgkin-Huxley type differential equations, which describe the axial current flow in accordance with the kinetics of the inactivating sodium, low-threshold potassium, and leak channels. In the simulation, action potentials were elicited at the first node of the axon by brief current stimulation, and the traveling time of the action potential across twenty internodes was used to calculate the conduction velocity.

The simulations were implemented with the myelinated axon model (Arancibia-Cárcamo et al., 2017)[2] in MATLAB. In the simulations, the axon internodal length was 0.187 mm (the average of SBC and GBC fibers measured in Ford et al., 2015) except where specified otherwise, and the node diameter was defined as 60% of the axon diameter (**Figure 1B**). Meanwhile,

[2]https://github.com/AttwellLab/MyelinatedAxonModel/

the myelin thickness of the axon varied from 0.2 to 0.6 μm. The anatomical arrangement of the model was based on adult Mongolian gerbil with stable myelination (Seidl and Rubel, 2016; Sinclair et al., 2017). The axon length from the SBC in CN to MSO on the ipsilateral side and from MNTB to MSO were assumed to have a length of 4.5 mm estimated from a gerbil brain atlas (Radtke-Schuller et al., 2016). As previously reported (Ford et al., 2015), the internodal length decreased at a distance of more than 0.5 mm from the branching area in MSO and 0.7 mm from the heminode near MNTB. At that point, the conduction velocity became more uniform along the rest of the axon, and the transmission delay was directly computed from the conduction velocity and the corresponding axon length with the uniform internodal length, i.e., 4 mm for the SBC-MSO projections and 3.8 mm for the GBC-MNTB projections.

Data Analysis

ITD responses to various sound stimuli of each MSO neuron were quantified as the firing rates in response to these stimuli (**Figure 2A**). The overall ITD tuning curves of the MSO were computed as the mean of single-neuron ITD responses in the MSO population trimmed by firing rates between 20 and 80% to omit non-responding neurons (**Figure 2B**). For a more accurate peak firing rate analysis, the ITD tuning curves were interpolated with a precision of 0.1 μs and subsequently smoothed by a Savitzky–Golay filter with an 80 μs smoothing window. After the ITD curves were smoothed, the peak firing rate positions were then used to define the best ITD. The peak amplitude and the Full-Width-at-Half-Maximum (FWHM) are illustrated in **Figure 2C**.

For the ITD decoding analysis, the spike counts of MSO neurons during the course of the stimuli were used as decoding features. ITDs were predicted using a Support Vector Machine (SVM) classifier with a linear kernel trained with the leave-one-out cross-validation approach. Three hundred MSO neurons were then randomly selected from both left and right MSO and sent to the SVM classifier to predict ITD in response to each stimulation. The decoding accuracy and the mean squared error (MSE) were determined from the predicted ITDs with 17 ITD sub-classes in the range from −1 to 1 ms. The classification accuracy between 10 and −10 μs ITDs was also computed in the same manner to estimate the precision of the ITD detection.

The sensitivity of ITDs was accessed using the Just-Noticeable-Difference (JND) that quantifies the smallest perceptible change. It was computed by comparing ITD responses symmetrical to the zero time, e.g., −50 and 50 μs ITDs. For each pair of symmetrical ITD responses, the difference of firing rates between left and right MSO were compared using the one-tailed Mann-Whitney *U*-test. The smallest symmetrical ITD that reached the minimum significant level indicates the JND.

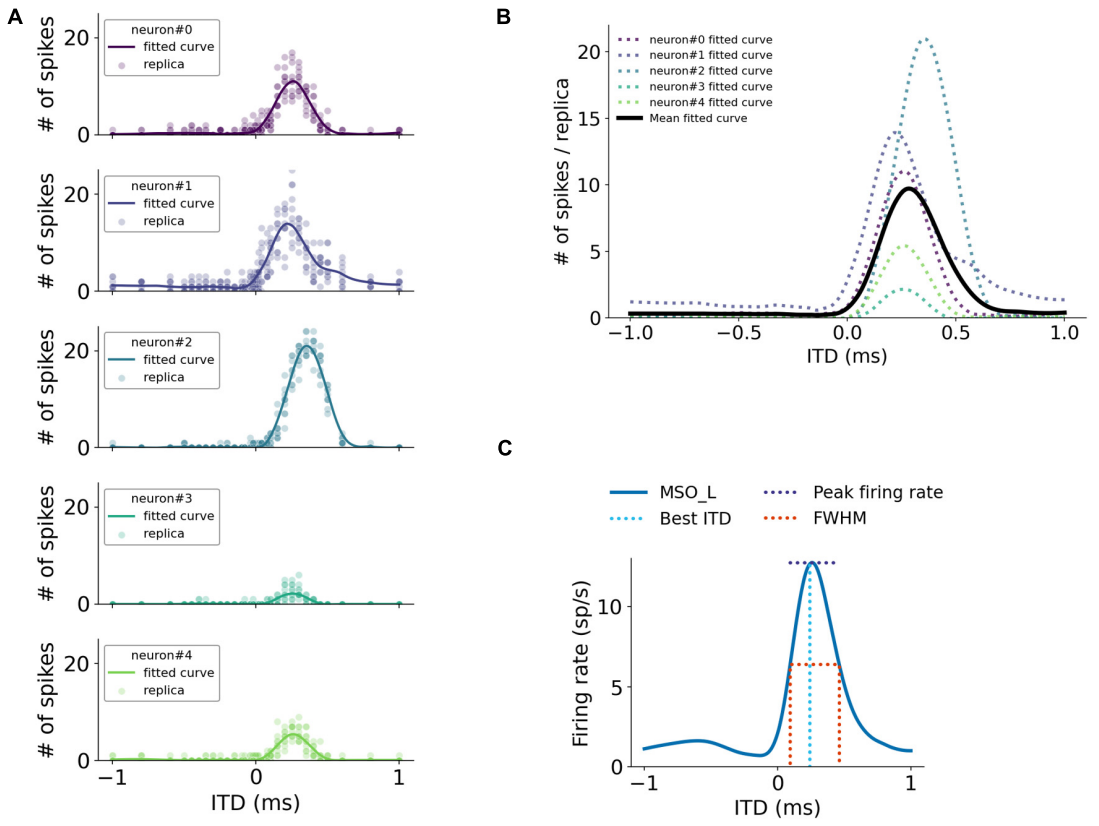

FIGURE 2 | Quantification of simulated MSO responses to ITDs. **(A)** ITD functions of single MSO neurons calculated by curve fitting. **(B)** Overall MSO response to ITDs (black curve) estimated by taking average of single-neuron ITD functions listed in **(A)**. **(C)** Annotation of metrics used for quantifying MSO responses.

Data Accessibility

The implementation source code and natural sound clips are available on the GitHub repository.[3]

RESULTS

Conduction Velocity Varies With Axon Myelination Patterns

We first assessed the correlation between axon myelination patterns and conduction velocity. Our overall results indicate that both the myelin thickness and the internodal length affect conduction velocity. While the level of myelin is directly proportional to the conduction velocity (**Figure 1C**), the internodal length has a non-linear relationship with the conduction velocity (**Figure 1D**). Theoretically, increasing the myelin thickness should increase the axial current flow, which in turn increases the propagation speed of action potentials. On the other hand, an increased internodal length results in a greater myelin coverage of the axon, which also increases the conduction velocity. In addition, fewer nodes result in less frequent regeneration of the action potential, additionally speeding up conduction. When the internodal length becomes even longer, the conduction velocity decreases since the transfer efficiency of the depolarization between nodes is lower (Brill et al., 1977; Ford et al., 2015). For the following results, we used the myelin thickness as the major variable for the comparison between different myelinated fibers, while the internodal length was set as a constant for comparison and simplification purposes.

Spherical Bushy Cell Axon Myelin Thickness and Interaural Time Difference Tuning

The firing rates of MSO neurons were studied during pure tone sound wave stimulation with varying ITDs, and were additionally recorded against changing myelin thicknesses of the SBC axon (d_{SBC}) along the contralateral excitatory pathway (**Figure 3A**). The ITD tuning curve shifted toward the center (0-ITD) when SBC myelination increased (**Figures 3B–D**). The observed shift was qualitatively the same for different sound frequencies (**Figures 3B–D**) but note that the absolute firing rate decreased with increasing frequency. This decrease is consistent with the reported frequency-dependent thresholds in MSO neurons (Remme et al., 2014; Mikiel-Hunter et al., 2016) and is most likely due to lower phase-locking at higher frequencies and periods of inhibition overlapping more with periods of excitation at higher frequencies. The corresponding best ITD, the FWHM, and the peak firing rate of the ITD tuning curves for the left MSO were quantified at 300 Hz and are shown in **Figure 4**.

The best ITD decreased linearly proportional to the myelin thickness of the SBC axon (**Figures 4A,B**). A positive best ITD value indicates that the left MSO fired more rapidly when the sound stimulus first arrived at the right ear. This result is in line with experimental observations and has been described by the opponent-channel coding model (Magezi and Krumbholz, 2010; Encke and Hemmert, 2018), where the ITD is encoded by contralateral MSO, for example, the right-leading ITD evokes the left MSO. In addition, as the myelin thickness increased ($d_{SBC} > 0.35$ μm), the best ITD value assumed values smaller than 0 ms. This shift caused the opponent-channel coding to fail and ITD responses as well as the coding scheme to reverse, predicting a sound location on the ipsilateral side of the brainstem. This result has not been experimentally reported. The SBC myelin thickness also altered the peak width (**Figures 4D,E**) and peak firing rate (**Figures 4G,H**) of the ITD tuning curve.

Globular Bushy Cell Axon Myelin Thickness and Interaural Time Difference Tuning

An increase in myelination (d_{GBC}) of the GBC axon along the contralateral inhibitory pathway (**Figure 3E**) showed significant changes to both the shapes and the scales of the ITD tuning curves with pure tone sound wave stimulation (**Figures 3F–H**). The best ITD value varied non-monotonically with GBC myelin thickness as shown in **Figures 4A,C**, where it first shifted toward the center (0 ITD), and later away from the center with a swing of more than 100 μs ITD (**Figures 4A,C**). The best ITD value reached its minimum when the GBC myelin thickness was around 0.25 μm and plateaued with a higher best ITD when the thickness was thicker than 0.3 μm. The peak width decreased when the GBC myelin thickness was increased, and subsequently became steady with thicker GBC myelination (**Figures 4D,F**). Note that a wider FWHM indicates broader tuned ITDs for the MSO, and a narrower FWHM suggests the tuning curve has a higher sensitivity and precision. Therefore, the results from this simulation indicate that GBC axons with relatively thicker myelination may yield more precise ITDs. Finally, the peak firing rate dropped substantially when the GBC myelin thickness increased from 0.2 to 0.3 μm, resulting in roughly stabilized peak firing rates (**Figures 4G,I**).

Both Spherical Bushy Cell and Globular Bushy Cell Myelination Influence Interaural Time Difference Tuning

The interaction between myelin thickness of two contralateral inputs (one excitatory and one inhibitory) toward shaping ITD are shown in **Figure 4**. The increase of the SBC myelin thickness (d_{SBC}) had a stronger effect on the best ITD, where the best ITD was increasing with thinner SBC myelin thickness (**Figures 4A,B**). Similarly, GBC myelin thickness (d_{GBC}) also shifted the best ITD within a 100 μs range, especially for d_{GBC} from 0.2 to 0.3 μm (**Figures 4A,C**). The FWHM and peak firing rate were affected by both SBC and GBC myelin thickness. Thinner SBC myelin and thicker GBC myelin tended to produce narrower FWHM (**Figures 4D–F**) and lower peak firing rates (**Figures 4G–I**).

[3]https://github.com/libenzheng/MedialSuperiorOliveModel

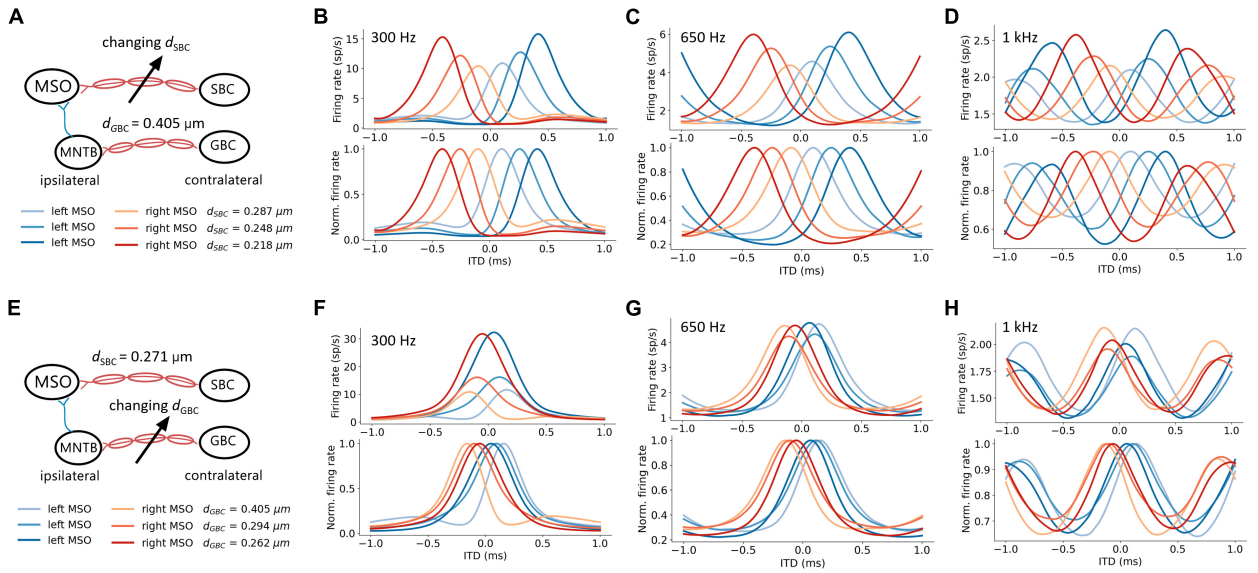

FIGURE 3 | Axon myelination along contralateral pathways influence ITD tuning. **(A–D)** Effects of changing SBC axon myelin thickness. **(E–H)** Effects of changing GBC axon myelin thickness. **(A,E)** Schematic of tuned parameter and legends of ITD tuning curves. **(B–D,F–H)** ITD tuning curve (top), and normalized ITD tuning curve (bottom). **(B,F)** 300 Hz 50 dB SPL pure tone stimulation. **(C,G)** 650 Hz 50 dB SPL pure tone stimulation. **(D,H)** 1 kHz 50 dB SPL pure tone stimulation.

Myelination Affecting Interaural Time Difference Encoding Accuracy

An ITD encoding analysis was conducted using population responses of MSO neurons to repeated pure tone stimuli with different ITDs. The encoding accuracy can be interpreted as the amount of ITD information extracted by the MSO (**Figures 5A–C**). The results indicate that the encoding accuracy reached its optimum when either SBC myelin was much thicker than GBC myelin, or when GBC myelin became much thicker than SBC myelin. The ability to accurately encode ITDs can also be represented as the mean squared error (MSE) between true ITDs and predicted ITDs (**Figures 5D–F**). Similar to the conclusion on ITD decoding accuracy, the contrast between SBC and GBC myelin thickness could result in smaller MSEs.

Myelin Thickness and the Precision of Interaural Time Difference Decoding

Natural sound clips presented with small ITDs were utilized to stimulate the circuit and calculate the precision of ITD decoding. The just noticeable difference (JND) of ITD was calculated by comparing MSO responses to symmetrical ITDs, and the calculated JND can be regarded as the sensitivity to ITD stimuli. The best sensitivity was obtained when SBC myelin was thinner than 0.27 μm, or GBC myelin was thicker than 0.38 μm and at the same time SBC myelin was thinner than 0.3 μm (**Figures 6A–C**). Apart from this, the sensitivity became far worse when the best ITD approached zero. The worst sensitivity was obtained when the best ITD was negative. This result can be explained since the opponent-channel coding scheme was used in the JND calculation together with a one-tailed test. Besides calculating JND, the decoding accuracy in the range between 10 and −10 μs ITDs was computed to quantify the precision of the circuit to capture very small ITDs. The most precise accuracy was also acquired when the GBC myelin was much thicker than the SBC myelin (**Figures 6D–F**).

The Effects of Contralateral Inhibition on Interaural Time Difference Decoding

The SNN circuit was simulated with different synaptic strengths of the contralateral inhibitory inputs to estimate the effects of the contralateral inhibition on the ITD computation (**Figure 7A**). The shape of the ITD tuning curve changed as the inhibitory synaptic conductance (Δg_{inhi}) from MNTB to MSO increased (**Figures 7B,C**). The best ITD shifted closer to zero-ITD when the contralateral inhibition increased and reached a plateau with a Δg_{inhi} above 25∼50 nS (**Figure 7D**). Both peak width and peak firing rate dropped with an increasing Δg_{inhi} (**Figures 7E,F**). Combining the results with those obtained from varying inhibition strength and myelin thickness, the maximum decoding accuracy, the minimum MSE and the minimum JND could be obtained when the SNN circuit has an optimal Δg_{inhi} of about 50 nS and the GBC myelin was 0.2 μm thicker than the SBC myelin (**Figures 7G–I**).

DISCUSSION AND CONCLUSION

Myelination, Conduction Velocity, and Spike Timing

In this study, the mammalian MSO circuit was modeled and tested with varying axon myelination properties to determine which role myelination plays in conduction velocity and transmission delay of action potentials. Physiological

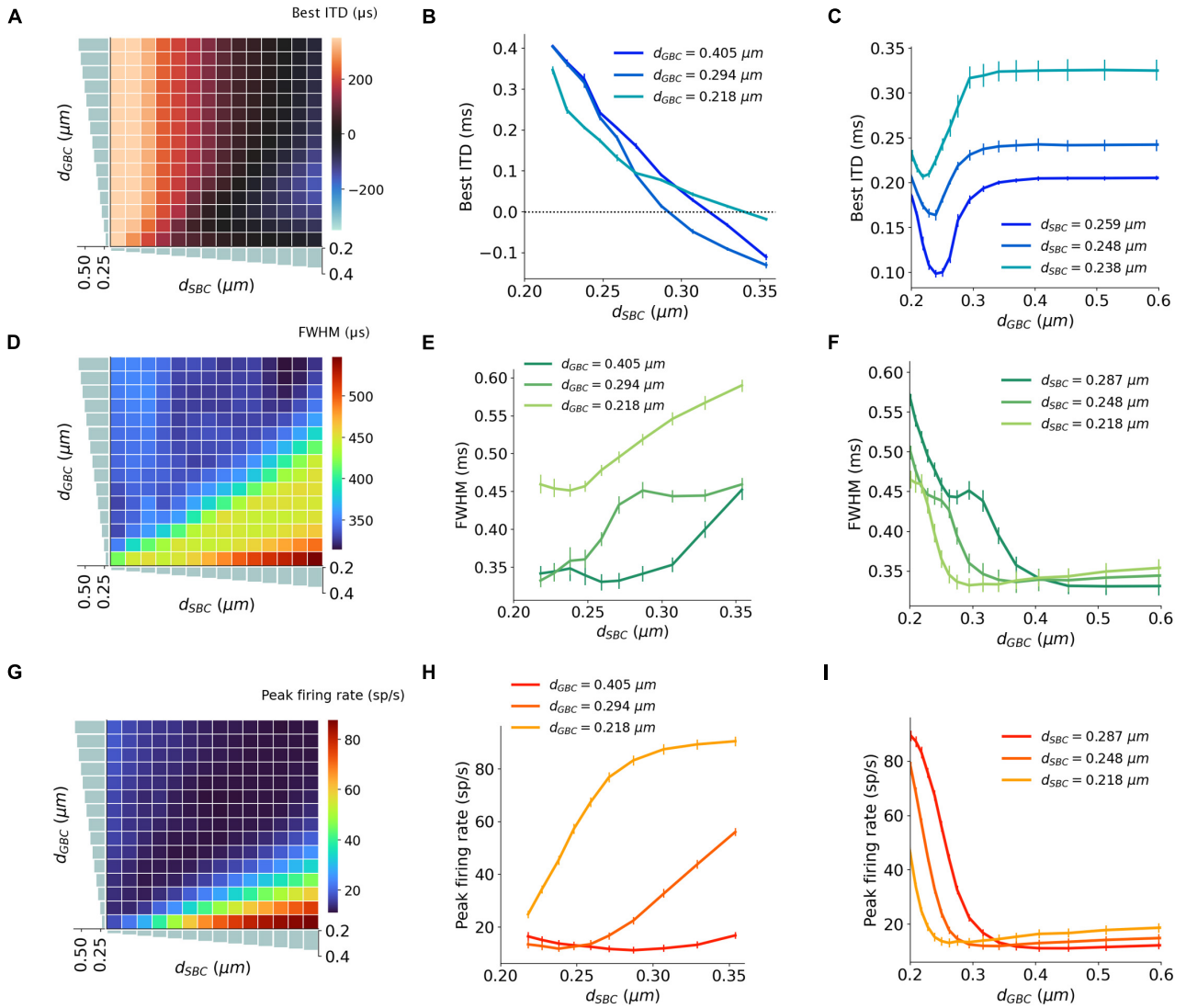

FIGURE 4 | Interaction between SBC and GBC axon myelination on ITD tuning, tested at 300 Hz. **(A–C)** Best ITD as a function of different SBC and GBC myelin thicknesses. **(D–F)** FWHM as a function of different SBC and GBC myelin thicknesses. **(G–I)** Peak firing rate as a function of different SBC and GBC myelin thicknesses. Error bars represent standard errors.

parameters included in our model were obtained from a number of published studies from Mongolian gerbils (Meriones unguiculatus), such that the resulting model should be a close representation of the gerbil.

We found that changes in axon myelin thickness affected the average propagation delays of action potentials between nuclei. A thicker myelin layer on GBC axons resulted in faster action potential propagation for the contralateral inhibition, and at the same time, a thinner myelin layer on the SBC axons resulted in longer delays for the contralateral excitation to MSO.

The model simulation demonstrated that the myelination of the contralateral input axons not only shifted the best ITDs but also influenced the peak firing rate and peak width of the ITD tuning curves, as shown in **Figures 3**, **4**. This change in the shape of the ITD tuning curve was induced by the relative input timing of the excitation and inhibition, which controls the duration of the net post-synaptic potential and the time window of the binaural coincidence detection. For instance, when the time window became shorter, MSO neurons elicited fewer spikes during the sound stimulation period within a narrower ITD range, subsequently, its ITD tuning curve was altered toward a lower peak firing rate and a shorter peak width.

Our decoding results (**Figures 5**, **6**) suggest that a larger GBC myelin layer combined with a smaller SBC myelin layer produced an optimal decoding accuracy and resulted in optimal ITD precision. In this scenario, action potentials associated with contralateral inhibition propagated faster than those associated with the excitation. This phenomenon had also been reported in previous experimental studies (Brand et al., 2002; Roberts et al., 2013). Specifically, it was postulated that MSO receives

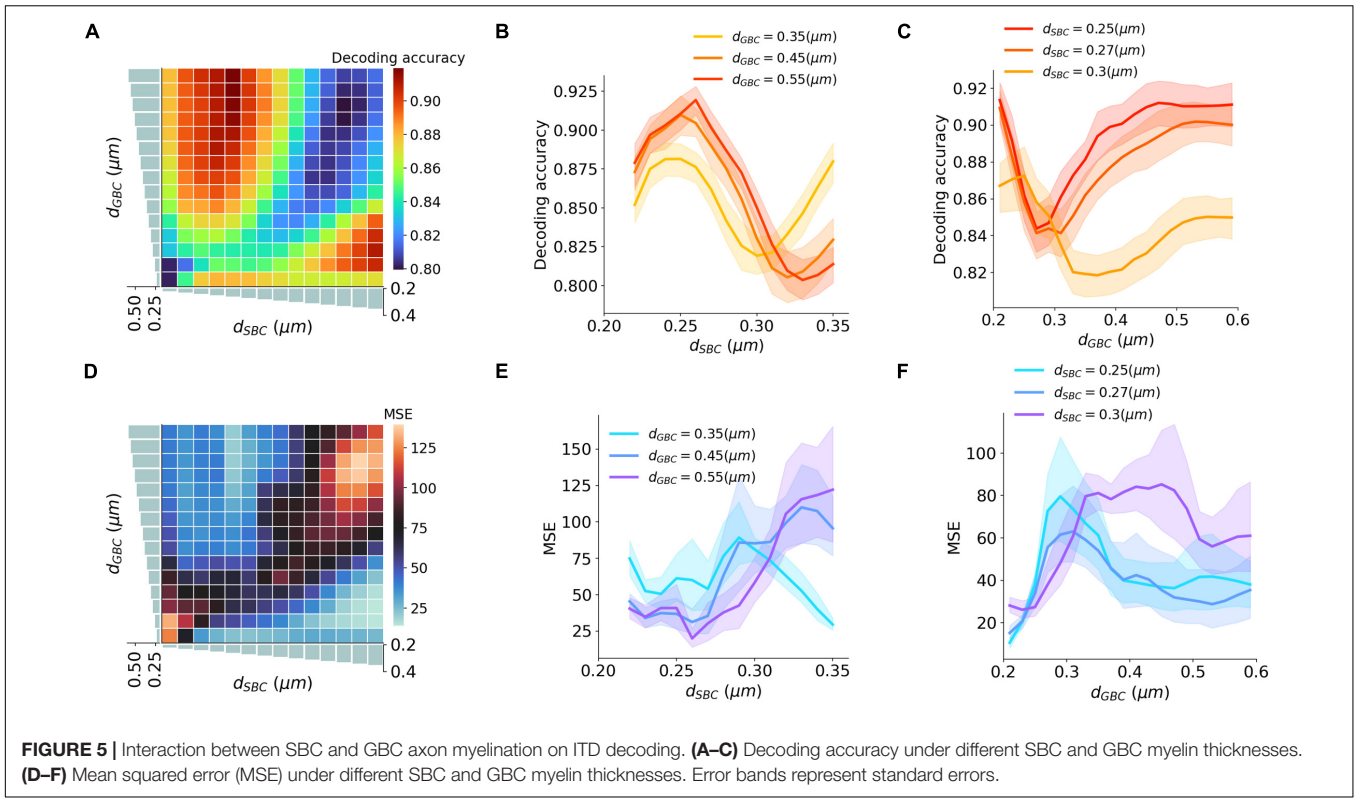

FIGURE 5 | Interaction between SBC and GBC axon myelination on ITD decoding. **(A–C)** Decoding accuracy under different SBC and GBC myelin thicknesses. **(D–F)** Mean squared error (MSE) under different SBC and GBC myelin thicknesses. Error bands represent standard errors.

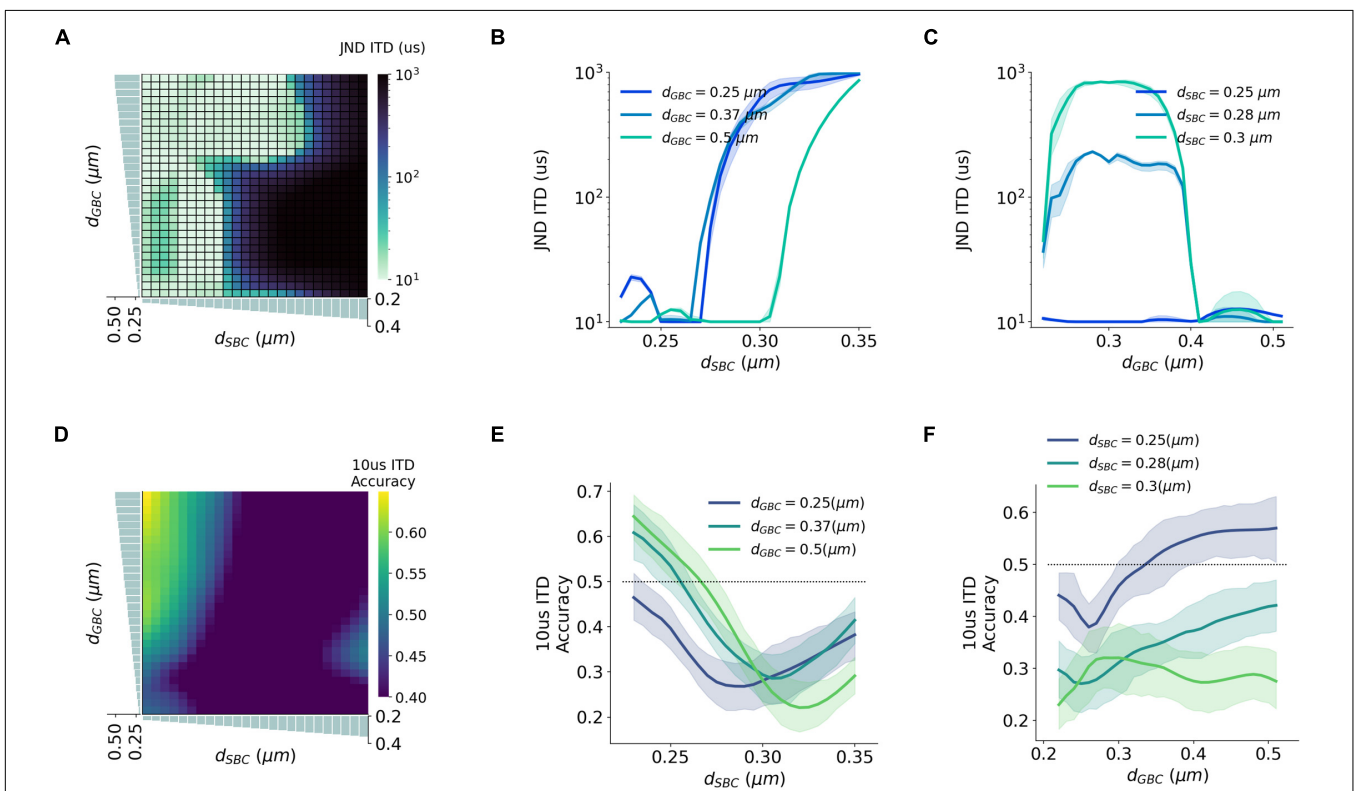

FIGURE 6 | Interaction between SBC and GBC axon myelination on ITD sensitivity. **(A–C)** Just noticeable difference (JND) of ITD for different SBC and GBC myelin thicknesses. **(D–F)** Decoding accuracy of 10-us ITD for different SBC and GBC myelin thicknesses. Error bands represent standard errors.

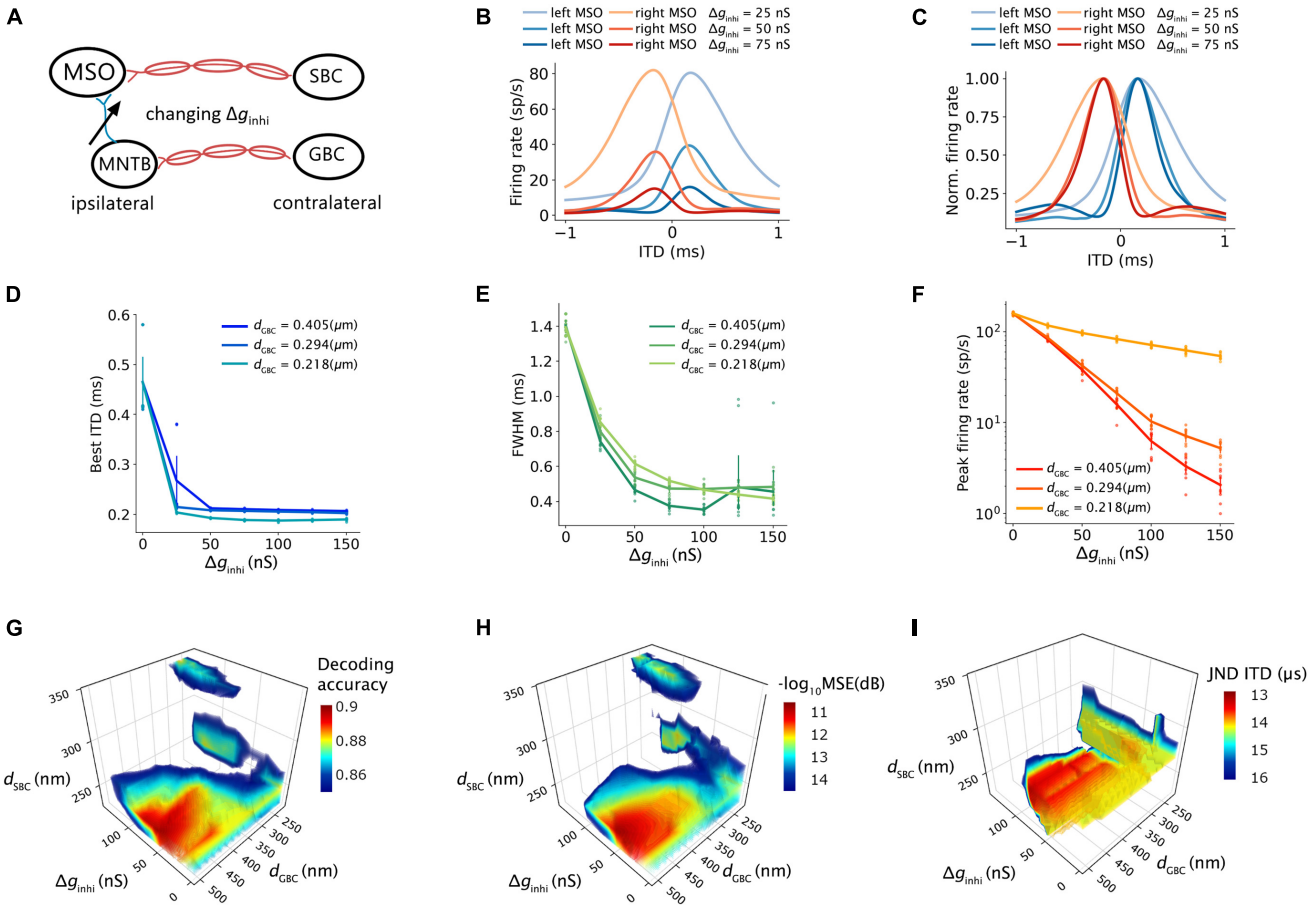

FIGURE 7 | Effects of contralateral inhibition on ITD tuning and decoding. **(A)** Sketch describing tuned parameter. **(B,C)** ITD tuning curve **(B)** and normalized ITD tuning curve **(C)** as a function of different contralateral inhibitory synaptic strength. **(D)** Best ITD vs. contralateral inhibitory synaptic strength with an SBC myelin thickness of 0.248 um. **(E)** Full width at half maximum (FWHM) vs. contralateral inhibitory synaptic strength with the SBC myelin thickness of 0.248 um. **(F)** Peak firing rate vs. contralateral inhibitory synaptic strength with the SBC myelin thickness of 0.248 um. **(G–I)** Decoding accuracy **(G)**, mean square error in decibel **(H)**, and just noticeable difference of ITD **(I)** change with contralateral inhibitory synaptic strength and contralateral myelin thickness; values above 75% quantile were colored.

contralateral inhibition earlier than contralateral excitation, despite an additional synapse in the inhibitory pathway and a longer distance to cover. These findings may indicate that this tuning of myelination of axons on the contralateral pathways could be the consequence of structural adaptation of the sound localization pathway toward more accurate ITD detection.

It has been shown that myelination is fully established in the gerbils auditory brainstem significantly after hearing onset (Seidl and Rubel, 2016; Sinclair et al., 2017). Moreover, myelination patterns differ between axons responding to low vs. high frequency sound (Ford et al., 2015) and are altered when the animal's sound experience is experimentally altered (Sinclair et al., 2017), suggesting that sound activity is involved in the establishment, and perhaps the maintenance of myelination. On the other hand, alterations in myelination as they occur in some conditions such as Fragile X (Lucas et al., 2021) or multiple sclerosis (Levine et al., 1993) or in animals with myelination deficits (Kim E. K. et al., 2013; Kim J. H. et al., 2013) result in temporally less well-timed activity in the sound localization pathway and/or impaired sound localization abilities, highlighting the functional consequences of these alterations. Taken together, these findings highlight the need for precisely controlled myelination patterns and suggest a possible mechanism to exert this control. The results of the present study are consistent with this body of work, highlighting how the computation of sound location in MSO changes with different myelination patterns.

Negative Best Interaural Time Difference

The optimal decoding accuracy and small MSE can mathematically be also achieved with a completely opposite myelination pattern consisting of thicker SBC myelin and thinner GBC myelin (**Figure 5**). Under this opposite scheme, the peak firing rate became much higher, and the best ITDs were negatively shifted away from the zero-ITD (**Figure 4**). From a purely mathematical perspective, the higher peak firing rate achieved in this scheme increases the signal dynamic range, thus improving the signal-to-noise ratio for the encoded ITDs. The

sign of the ITD can be thought of being irrelevant for the total amount of information, since negative or the positive best ITDs encode the same ITD value. However, from an experimental standpoint, such a non-opponent-channel coding scheme has not been observed.

Possible Roles of Inhibition in Medial Superior Olive

Several competing models have been suggested for the role of inhibition in the MSO in ITD tuning (Brand et al., 2002; Pecka et al., 2008; Couchman et al., 2010; Grothe et al., 2010; Roberts et al., 2013; van der Heijden et al., 2013; Myoga et al., 2014; Franken et al., 2015). Some studies support the hypothesis that this inhibition modulates the peak timing of the excitation and tunes the coincidence detection. Notably, the pharmacological blockage of inhibition shifts the best ITD toward the zero-ITD (Brand et al., 2002; Pecka et al., 2008). Moreover, conductance clamp recordings (Myoga et al., 2014) demonstrated that precisely timed inhibition could tune the best ITD by modulating the net excitatory post-synaptic potential (EPSP), and the leading contralateral inhibition biased the coincidence detection timing about 50∼150 μs. The experimental result is comparable to our computational results (**Figures 4A,C**) in which the increased GBC myelin thickness modestly shifted the best ITD by only 100 μs but resulted in much improved sensitivity and decoding accuracy.

On the other hand, other studies have challenged this inhibition-tuning model. Although well-timed leading contralateral inhibition was observed in an *in vitro* study (Roberts et al., 2013), the ITD function and EPSP did not significantly differ with an inhibitory conductance of a 300-μs leading contralateral inhibition. In addition, although shift of best ITD toward zero was shown for pharmacological blockage, Roberts et al. (2013) described the role of inhibition as transient and less significant over time, inferring that the removal of inhibition should not systemically shift the best ITD (Franken et al., 2015). In these studies, the occurrence of inhibition decreased overall firing rates across ITDs and narrowed the ITD functions without shifting the best ITD. This trend is consistent with our results to some degree (**Figure 7**), as the best ITD was not shifted with varying inhibitory synaptic conductances unless the inhibitory conductance was lower than 50 nS (**Figure 7D**). The increased inhibition also reduced the peak firing rate (**Figure 7E**) and peak width (**Figure 7F**) in a way similar to the increased myelin thickness and conduction velocity on the contralateral inhibitory pathway (**Figures 3E–H**). On the other hand, we note that the ITDs produced by varying myelination in the afferent excitatory pathway exceed the biologically relevant range of at least most mammalian species (**Figures 2B,C**), while the smaller range of ITDs produced by varying myelination in the inhibitory pathway matches that range closer. It is, therefore, possible that this study underestimates that role.

The relative timing of the binaural excitation was concluded to be the dominant factor for ITD tuning due to its apparent capability to regulate the best ITD compared to the inhibition (Roberts et al., 2013; van der Heijden et al., 2013; Seidl and Rubel, 2016). However, through the decoding analysis, the linkage between the best ITD and the estimated ITD sensitivity was unexpectedly shown to be more indirect but in a profound way in which the best ITD shifts could not simply be used to predict the precision of the ITD detection. Even though the presence of the leading contralateral inhibition reduced the peak firing rate and narrowed the peak width of the ITD tuning curves, the decoding results revealed that the timing and synaptic strength of the contralateral inhibition largely attributed to the pinpoint precision and sensitivity of the ITD computation (**Figures 7G–I**). Therefore, our findings imply that the complexity of ITD tuning depends on the temporal interaction between the excitation and inhibition.

Limitations

Our computational model was designed to probe the influence of axon myelination. We simplified the model and omitted several possible mechanisms of ITD encoding. First, a low-threshold potassium current shown to interact with the synaptic inhibition in MSO and sharpen the temporal sensitivity of the binaural integration (Khurana et al., 2011; Roberts et al., 2013; Myoga et al., 2014) was omitted. Second, post-inhibitory facilitation that can raise the firing rate under certain conditions (Beiderbeck et al., 2018; Ma et al., 2021) and had been observed in the MSO of juvenile mice (Dodla et al., 2006) was also not considered. This phenomenon could possibly compensate for a decreased firing rate induced by the leading contralateral inhibition in MSO. Third, a basic leaky integrate-and-fire model was utilized to improve computational efficiency and avoid some detailed physiological parameters, which have not been well characterized in gerbils. Nonetheless, the current model could be modified with the inclusion of diverse physiological parameters (Remme et al., 2014) to simulate rich membrane dynamics and temporal responses. Furthermore, besides spike conduction latency, spike conduction jitter in the auditory brainstem could affect precise temporal integration in sound localization (Reed et al., 2002; Marsalek and Kofranek, 2005; Kim E. K. et al., 2013; Kim J. H. et al., 2013). In our model, the conduction jitter was simplified as a constant 0.05 ms standard deviation that did not vary with myelin thickness.

Conclusion

By using an SNN model of the auditory brainstem, we found that axon myelination regulated ITD computation. The myelination of contralateral excitatory pathways shifted the best ITD. Moreover, the myelination and synaptic strength of contralateral inhibition influenced the peak firing rate and width of the ITD tuning curve, and subsequently modulated the ITD precision and sensitivity.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

B-ZL designed the study, conducted the study, analyzed the data, and wrote the manuscript. SP and MV designed the study. AK and TL designed the study and wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the NIH R01 DC 17924 and R01 DC 18401. This work was also funded by the National Key R&D Program of China (No. 2020YFB1313502), the Shenzhen-Hong Kong-Macau S&T Program (Category C) of SZSTI (No. SGDX20201103094002009), the University of Macau (File nos. MYRG2018-00146-AMSV, MYRG2019-00056-AMSV), the Science and Technology Development Fund, Macau SAR [File nos. 088/2016/A2, 0144/2019/A3, 0022/2020/AFJ, SKL-AMSV (FDCT-funded), SKL-AMSV-ADDITIONAL FUND, SKL-AMSV (UM)-2020-2022]. Additionally, this work utilized the Summit supercomputer, which was supported by the National Science Foundation (awards ACI-1532235 and ACI-1532236) to the University of Colorado Boulder and Colorado State University.

REFERENCES

Arancibia-Cárcamo, I. L., Ford, M. C., Cossell, L., Ishida, K., Tohyama, K., and Attwell, D. (2017). Node of Ranvier length as a potential regulator of myelinated axon conduction speed. *eLife* 6:e23329. doi: 10.7554/elife.23329

Beiderbeck, B., Myoga, M. H., Müller, N. I. C., Callan, A. R., Friauf, E., Grothe, B., et al. (2018). Precisely timed inhibition facilitates action potential firing for spatial coding in the auditory brainstem. *Nat. Commun.* 9:1771. doi: 10.1038/s41467-018-04210-y

Boudreau, J. C., and Tsuchitani, C. (1968). Binaural interaction in the cat superior olive S segment. *J. Neurophysiol.* 31, 442–454. doi: 10.1152/jn.1968.31.3.442

Brand, A., Behrend, O., Marquardt, T., McAlpine, D., and Grothe, B. (2002). Precise inhibition is essential for microsecond interaural time difference coding. *Nature* 417, 543–547. doi: 10.1038/417543a

Brill, M. H., Waxman, S. G., Moore, J. W., and Joyner, R. W. (1977). Conduction velocity and spike configuration in myelinated fibres: computed dependence on internode distance. *J. Neurol. Neurosurg. Psychiatry* 40, 769–774. doi: 10.1136/jnnp.40.8.769

Brughera, A., Dunai, L., and Hartmann, W. M. (2013). Human interaural time difference thresholds for sine tones: the high-frequency limit. *J. Acoust. Soc. Am.* 133, 2839–2855. doi: 10.1121/1.4795778

Couchman, K., Grothe, B., and Felmy, F. (2010). Medial superior olivary neurons receive surprisingly few excitatory and inhibitory Inputs with balanced strength and short-term dynamics. *J. Neurosci.* 30, 17111–17121. doi: 10.1523/jneurosci.1760-10.2010

Debanne, D. (2004). Information processing in the axon. *Nat. Rev. Neurosci.* 5, 304–316. doi: 10.1038/nrn1397

Dodla, R., Svirskis, G., and Rinzel, J. (2006). Well-timed, brief inhibition can promote spiking: postinhibitory facilitation. *J. Neurophysiol.* 95, 2664–2677. doi: 10.1152/jn.00752.2005

Encke, J., and Hemmert, W. (2018). Extraction of inter-aural time differences using a spiking neuron network model of the medial superior olive. *Front. Neurosci.* 12:140. doi: 10.3389/fnins.2018.00140

Ford, M. C., Alexandrova, O., Cossell, L., Stange-Marten, A., Sinclair, J., Kopp-Scheinpflug, C., et al. (2015). Tuning of Ranvier node and internode properties in myelinated axons to adjust action potential timing. *Nat. Commun.* 6:8073. doi: 10.1038/ncomms9073

Franken, T. P., Roberts, M. T., Wei, L., Golding, N. L., and Joris, P. X. (2015). *In vivo* coincidence detection in mammalian sound localization generates phase delays. *Nat. Neurosci.* 18, 444–452. doi: 10.1038/nn.3948

Glackin, B., Wall, J. A., McGinnity, T. M., Maguire, L. P., and McDaid, L. J. (2010). A spiking neural network model of the medial superior olive using spike timing dependent plasticity for sound localization. *Front. Comput. Neurosci.* 4:18. doi: 10.3389/fncom.2010.00018

Goldberg, J. M., and Brown, P. B. (1969). Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. *J. Neurophysiol.* 32, 613–636. doi: 10.1152/jn.1969.32.4.613

Grothe, B. (2003). New roles for synaptic inhibition in sound localization. *Nat. Rev. Neurosci.* 4, 540–550. doi: 10.1038/nrn1136

Grothe, B., and Pecka, M. (2014). The natural history of sound localization in mammals – a story of neuronal inhibition. *Front. Neural Circuits* 8:116. doi: 10.3389/fncir.2014.00116

Grothe, B., Pecka, M., and McAlpine, D. (2010). Mechanisms of sound localization in mammals. *Physiol. Rev.* 90, 983–1012. doi: 10.1152/physrev.00026.2009

Halter, J. A., and Clark, J. W. (1991). A distributed-parameter model of the myelinated nerve fiber. *J. Theor. Biol.* 148, 345–382. doi: 10.1016/s0022-5193(05)80242-5

Jeffress, L. A. (1948). A place theory of sound localization. *J. Comp. Physiol. Psychol.* 41, 35–39. doi: 10.1037/h0061495

Khurana, S., Remme, M. W. H., Rinzel, J., and Golding, N. L. (2011). Dynamic interaction of Ih and IK-LVA during trains of synaptic potentials in principal neurons of the medial superior olive. *J. Neurosci.* 31, 8936–8947. doi: 10.1523/jneurosci.1079-11.2011

Kim, J. H., Renden, R., and von Gersdorff, H. (2013). Dysmyelination of auditory afferent axons increases the jitter of action potential timing during high-frequency firing. *J. Neurosci.* 33, 9402–9407. doi: 10.1523/jneurosci.3389-12.2013

Kim, E. K., Turkington, K., Kushmerick, C., and Kim, J. H. (2013). Central dysmyelination reduces the temporal fidelity of synaptic transmission and the reliability of postsynaptic firing during high-frequency stimulation. *J. Neurophysiol.* 110, 1621–1630. doi: 10.1152/in.00117.2013

Leibold, C. (2010). Influence of inhibitory synaptic kinetics on the interaural time difference sensitivity in a linear model of binaural coincidence detection. *J. Acoust. Soc. Am.* 127, 931–942. doi: 10.1121/1.3282997

Levine, R. A., Gardner, J. C., Fullerton, B. C., Stufflebeam, S. M., Carlisle, E. W., Furst, M., et al. (1993). Effects of multiple sclerosis brainstem lesions on sound lateralization and brainstem auditory evoked potentials. *Hear. Res.* 68, 73–88. doi: 10.1016/0378-5955(93)90066-A

Lucas, A., Poleg, S., Klug, A., and McCullagh, E. A. (2021). Myelination deficits in the auditory brainstem of a mouse model of fragile X syndrome. *Front. Neurosci.* 15:772943. doi: 10.3389/fnins.2021.772943

Ma, H., Jia, B., Li, Y., and Gu, H. (2021). Excitability and threshold mechanism for enhanced neuronal response induced by inhibition preceding excitation. *Neural Plast.* 2021:6692411. doi: 10.1155/2021/6692411

Magezi, D. A., and Krumbholz, K. (2010). Evidence for opponent-channel coding of interaural time differences in human auditory cortex. *J. Neurophysiol.* 104, 1997–2007. doi: 10.1152/jn.00424.2009

Marsalek, P., and Kofranek, J. (2005). Spike encoding mechanisms in the sound localization pathway. *Biosystems* 79, 191–198. doi: 10.1016/j.biosystems.2004.09.022

Mikiel-Hunter, J., Kotak, V., and Rinzel, J. (2016). High-frequency resonance in the gerbil medial superior olive. *PLoS Comput. Biol.* 12:e1005166. doi: 10.1371/journal.pcbi.1005166

Morest, D. K. (1968). The growth of synaptic endings in the mammalian brain: a study of the calyces of the trapezoid body. *Z. Anat. Entwicklungsgesch.* 127, 201–220. doi: 10.1007/bf00526129

Myoga, M. H., Lehnert, S., Leibold, C., Felmy, F., and Grothe, B. (2014). Glycinergic inhibition tunes coincidence detection in the auditory brainstem. *Nat. Commun.* 5:3790. doi: 10.1038/ncomms4790

Nave, K.-A. (2010). Myelination and support of axonal integrity by glia. *Nature* 468, 244–252. doi: 10.1038/nature09614

Pan, Z., Zhang, M., Wu, J., Wang, J., and Li, H. (2021). Multi-tone phase coding of interaural time difference for sound source localization with spiking neural networks. *IEEE ACM Trans. Audio Speech Lang. Process.* 29, 2656–2670. doi: 10.1109/taslp.2021.3100684

Pecka, M., Brand, A., Behrend, O., and Grothe, B. (2008). Interaural time difference processing in the mammalian medial superior olive: the role of glycinergic inhibition. *J. Neurosci.* 28, 6914–6925. doi: 10.1523/jneurosci.1660-08.2008

Radtke-Schuller, S., Schuller, G., Angenstein, F., Grosser, O. S., Goldschmidt, J., and Budinger, E. (2016). Brain atlas of the Mongolian gerbil *(Meriones unguiculatus)* in CT/MRI-aided stereotaxic coordinates. *Brain Struct. Funct.* 221, 1–272. doi: 10.1007/s00429-016-1259-0

Reed, M., Blum, J., and Mitchell, C. (2002). Precision of neural timing: effects of convergence and time-windowing. *J. Comput. Neurosci.* 13, 35–47. doi: 10.1023/a:1019692310817

Remme, M. W. H., Donate, R., Mikiel-Hunter, J., Ballestero, J. A., Foster, S., Rinzel, J., et al. (2014). Subthreshold resonance properties contribute to the efficient coding of auditory spatial cues. *Proc. Natl. Acad. Sci. U.S.A.* 111, E2339–E2348. doi: 10.1073/pnas.1316216111

Roberts, M. T., Seeman, S. C., and Golding, N. L. (2013). A mechanistic understanding of the role of feedforward inhibition in the mammalian sound localization circuitry. *Neuron* 78, 923–935. doi: 10.1016/j.neuron.2013.04.022

Roberts, M. T., Seeman, S. C., and Golding, N. L. (2014). The relative contributions of MNTB and LNTB neurons to inhibition in the medial superior olive assessed through single and paired recordings. *Front. Neural Circuits* 8:49. doi: 10.3389/fncir.2014.00049

Rudnicki, M., Schoppe, O., Isik, M., Völk, F., and Hemmert, W. (2015). Modeling auditory coding: from sound to spikes. *Cell Tissue Res.* 361, 159–175. doi: 10.1007/s00441-015-2202-z

Schwartz, I. R. (1992). "The Superior Olivary Complex and Lateral Lemniscal Nuclei," in *The Mammalian Auditory Pathway: Neuroanatomy*, eds D. B. Webster, A. N. Popper, R. R. Fay (New York, NY: Springer), 117–167. doi: 10.1007/978-1-4612-4416-5_4

Seidl, A. H., and Rubel, E. W. (2016). Systematic and differential myelination of axon collaterals in the mammalian auditory brainstem. *Glia* 64, 487–494. doi: 10.1002/glia.22941

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

Spirou, G. A., Rager, J., and Manis, P. B. (2005). Convergence of auditory-nerve fiber projections onto globular bushy cells. *Neuroscience* 136, 843–863. doi: 10.1016/j.neuroscience.2005.08.068

Stange-Marten, A., Nabel, A. L., Sinclair, J. L., Fischl, M., Alexandrova, O., Wohlfrom, H., et al. (2017). Input timing for spatial processing is precisely tuned via constant synaptic delays and myelination patterns in the auditory brainstem. *Proc. Natl. Acad. Sci. U.S.A.* 114, E4851–E4858. doi: 10.1073/pnas.1702290114

Stimberg, M., Brette, R., and Goodman, D. F. (2019). Brian 2, an intuitive and efficient neural simulator. *eLife* 8:e47314. doi: 10.7554/elife.47314

Tollin, D. J., and Yin, T. C. T. (2009). "Sound localization: neural mechanisms," in *Encyclopedia of Neuroscience*, eds M. D. Binder, N. Hirokawa, U. Windhorst (Madison, WI: University of Wisconsin), 137–144.

Trussell, L. O. (1999). Synaptic mechanisms for coding timing in auditory neurons. *Annu. Rev. Physiol.* 61, 477–496. doi: 10.1146/annurev.physiol.61.1.477

van der Heijden, M., Lorteije, J. A. M., Plauška, A., Roberts, M. T., Golding, N. L., Borst, J., et al. (2013). Directional hearing by linear summation of binaural inputs at the medial superior olive. *Neuron* 78, 936–948.

Wang, L., Devore, S., Delgutte, B., and Colburn, H. S. (2014). Dual sensitivity of inferior colliculus neurons to ITD in the envelopes of high-frequency sounds: experimental and modeling study. *J. Neurophysiol.* 111, 164–181. doi: 10.1152/jn.00450.2013

Zhou, Y. Carney, L. H, and Colburn H. S (2005). A model for interaural time difference sensitivity in the medial superior olive: interaction of excitatory and inhibitory synaptic inputs, channel dynamics, and cellular morphology. *J. Neurosci.* 25, 3046–3058. doi: 10.1523/jneurosci.3064-04.2005

Zilany, M. S. A., Bruce, I. C., and Carney, L. H. (2014). Updated parameters and expanded simulation options for a model of the auditory periphery. *J. Acoust. Soc. Am.* 135, 283–286. doi: 10.1121/1.4837815

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership