On the plasticity of sound localization during the use of hearing protection devices

David J. Audet Jr.

A dissertation

submitted in partial fulfilment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2026

Reading Committee:

Andrew D. Brown, Chair

Yi Shen

Beverly Wright

Program Authorized to Offer Degree:

Speech and Hearing Sciences

University of Washington

Abstract

On the plasticity of sound localization during the use of hearing protection devices

David J. Audet Jr.

Chair of the Supervisory Committee:

Andrew D. Brown

Department of Speech and Hearing Sciences

Hearing is important for communication and situational awareness. Hearing loss due to noise exposure or other insults can degrade these capacities, with corresponding impacts on quality of life and occupational fitness. High-intensity acoustic signals are commonplace in a variety of occupational settings (e.g., military, construction, manufacturing) and are a major cause of permanent hearing loss. Hearing protection devices (HPDs) are therefore essential personal safety equipment for those operating in noisy environments, and are widely available, often in concert with organizational hearing conservation programs that support their use. However, HPDs can exert undesirable perceptual side effects, including acute disruption of spatial hearing. As HPDs are placed in and around the ears, distortions of the natural pinna-related spatial cues result, leading to decreased sound localization abilities, device dissatisfaction and disuse. This dissertation is comprised of three studies that investigate both the impact of hearing protector use on sound localization abilities and the capacity for recovery of those abilities with user training. The first study assessed the horizontal localization abilities of subjects before and after adapting to degraded spatial cues via earplugs. Over the course of eight weeks with intermittent training, subjects demonstrated a vastly improved ability to localize sound sources across both the front and rear hemifields, especially for subjects receiving audiovisual feedback during training. Expanding on these results, the second study aimed to evaluate if similar gains could be realized when sound sources varied across both horizontal and vertical planes simultaneously. Results revealed interactions across spatial domains that partially hindered the recovery of spatial hearing abilities, although measurable gains in performance were nonetheless observed. Finally, utilizing a custom earmuff-style hearing protector, the last study investigated the cues underlying the discrimination of frontal and rear hemifield sound sources and changes that occur following the adaptation to degraded cues. Altogether, these works highlight the plasticity of and barriers to human sound localization with degraded spatial cues. These results directly inform the design of next-generation hearing devices that support both hearing protection and auditory spatial perception.

Table of Contents

Chapter 1: An introduction to sound localization and its disruption via hearing devices

Successful sound source localization is important for safety, social communication, and enjoyment of everyday life. Understanding where different sound sources are in an environment allows for immersion in complex acoustic scenes, keeping the listener aware of potential hazards, engaged with communication partners or focused on sources of entertainment. Three primary acoustic cues support the perception of sound source location. Perception of source azimuth primarily relies on interaural time and level differences (ITD and ILD, respectively) between the two ears, generated as sound travels from the source to the listener's ears (Rayleigh, 1907; Stevens & Newman, 1936; Mills, 1958). Perception of source elevation, and the discrimination of sounds in the front or rear hemifield are primarily achieved via encoding of the spectral shape cue, generated from reflections of sound from the listeners body, head, and pinna (Batteau, 1967; Musicant & Butler, 1984).

The sound source localization ability of humans has been examined using a variety of methods for over 100 years with each cue being probed for its contribution to overall behavioral performance. Largely, these cues have been examined in isolation or are generally categorized into spatial domains, such as horizontal perception, vertical perception, and front-back discrimination. The focus of this dissertation is (1) the use of spectral shape cues toward both vertical perception and front-back discrimination and (2) the adaptation to cues disrupted by an applied source, notably, hearing protection. The remaining sections of this chapter highlight the anatomy and physiology supporting such perception and establish a fundamental expectation for behavioral performance within the vertical and front-back domains.

1.1. A review of sound localization cues and behavior

Accurate perception of a sound source's location in space is driven by the decoding of information provided by the three sound localization cues. Here, the cues will be addressed as they relate to azimuth, elevation, and front-back location perception. During such descriptions, locations are presented within a double-pole coordinate system (Makous & Middlebrooks, 1990; Middlebrooks & Green, 1991) where azimuth is defined as the horizontal angle of deviation from the midsagittal plane and elevation is defined as the angle of deviation from the horizon. In this system, the location directly in-front of a listener is represented as 0°,0° (azimuth and elevation, respectively) with positive azimuth and elevation angles representing locations to the right and upward, respectively.

1.1.1 A note on behavioral measures employed

As will be seen throughout this dissertation, localization behavior has been tested using many different methods, and while no method is universally better than others, it is important to note the differences between methods to understand how best to interpret the results. Two primary categories of methods are employed: source identification (an absolute location judgment) and source discrimination (a relative location judgment). Source identification involves having listeners record the perceived position/direction of a sound in space. This is typically done using some form of a speaker array where sounds are played from numerous locations. If this is done in an open-set procedure with random target locations, it is known as a 'localization' experiment (Stevens & Newman, 1936). The second primary method is source discrimination, exemplified by a minimum audible angle task (MAA), where participants must discriminate between, usually, two sounds along some cue dimension (Mills, 1958). For example, a 2-alternative forced choice (2-AFC) task would ask whether the second sound was to the left or right of the first. A localization task may become a discrimination task by discretizing space into so few locations that it becomes a ‘closed-set’ task. The method used directly impacts the results and suggested interpretation thereof. Localization tasks report absolute error, while discrimination tasks report error threshold or precision (reviewed in Poirier-Quinot et al., 2022).

1.1.2 Localization via binaural cues

The duplex theory of sound localization posits that low and high frequency horizontal perception are guided by the encoding of interaural time difference (ITD) and interaural level difference (ILD) cues (Rayleigh, 1907; Middlebrooks & Greene, 1991). For sounds travelling through a given medium, the use of these cues is reasonably driven by the acoustic properties of the stimulus and its interactions with the head and body. Measured low frequency ITDs are nearly sinusoidal as a function of source azimuth, meaning that as the source moves laterally, changes in the ITD become smaller in a monotonic fashion (Mills, 1958; Kuhn, 1977). Measured ILDs are frequency, azimuth, and distance dependent. Overall, ILDs are typically near 0 dB at the midline, as sound reaches both ears equally. At sources away from the midline, ILDs increase with increasing frequency (Abbagnaro, 1975; Brungart & Rabinowitz, 1999), and concurrently increase as the source moves laterally, albeit with notable non-monotonicities that also vary across frequency (Macaulay et al. 2010). Lastly, while low frequency ILDs are generally small beyond 1 meter, they become increasingly large as sources approach the head (Brungart & Rabinowitz, 1999).

Free-field horizontal localization performance is best around midline (0°) and steadily gets worse toward more lateral positions (4-20° of error) (Stevens & Newman, 1936) and especially in the rear hemifield or at non-zero elevations (Makous & Middlebrooks, 1990). Additionally, it is noted that broadband signals are localized better than tones, since redundant information may be integrated across multiple frequency channels and localization cues. Indeed, in a free-field localization experiment, it is reasonable to assume that listeners are utilizing all cues available, with ITD and ILD perception dominating low and high frequency sound localization, respectively (Macpherson & Middlebrooks, 2002).

1.1.2.1 Horizontal perception during bilaterally symmetrical cue distortion

Prior to discussion of other spatial dimensions, it is interesting to note that horizontal sound perception is generally preserved during the use of bilateral hearing protection devices (HPDs; Brown et al., 2015; Audet et al., 2026; see Chapter 2) or non-occluding earmolds (Hofman et al., 1998; Trapeau et al., 2016), with only minor deficits in perception. Given symmetrical insertion of HPDs, any delay or attenuation provided by the devices is applied equally across ears, preserving ITDs and ILDs, albeit, with signals transmitted with increased latency and at a lower overall intensity, respectively. Earmolds leave the ear canal open, minimally impacting binaural cues of the signal except at the highest frequencies, and generally preserving binaural cue integrity. The small impact on horizontal perception may then be a result of variable device insertion, as perfectly symmetrical insertion is difficult to achieve. Consequently, small shifts in ITD or ILD may be introduced, or interaural correlation may decline, leading to worse horizontal perception (Jeffress et al., 1962). As the remainder of the dissertation concerns monaural cues, binaural cues will be further considered only where relevant to the interpretation of experimental results.

1.1.3 The use of monaural cues for elevation and front-back sound localization

1.1.3.1 Generation of the spectral shape cues

The monaural spectral shape cue is derived primarily from the shape of the pinna, with additional contribution from torso, shoulder, and head reflections. For human listeners, it is generally understood that the pinna's primary purpose is for sound source localization, as its unique shape leads to frequency-dependent reflections (Batteau, 1967) that combine to form a sound's spectral shape. The spectral shape can be employed to disambiguate sources in elevation and front-back hemifields because it changes systematically with source location, via changes that are directly correlated with an individual's pinna shape (Batteau, 1967; Raykar et al., 2005; Stitt & Katz, 2021).

Spectral cues are measured via derivation of a head-related transfer function (HRTF), where the magnitude spectrum of a recording made from within the ear canal of a subject is compared to the spectrum of the stimulus. Inspection of the HRTF reveals the acoustic transformations made by the pinna, head, and body. A conspicuous spectral feature recognized as contributing to the representation of sounds with varying elevation is a high-frequency notch (region of low amplitude; Butler & Belendiuk, 1977). This notch, also observed in non-human animals with pinnae (e.g. cats; Rice et al., 1992), changes in frequency coinciding with changes in elevation and has been a primary target for the study of elevation encoding in the auditory system. Investigation into the origin of the notch in humans, via 3D scanning and acoustic simulation of manipulated pinna landmarks, suggest a primary source as the triangular fossa (Stitt & Katz, 2021).

An additional feature that has been given more attention recently is a low-frequency notch thought to be generated by reflections of sound off the torso (Nakamura & Iida, 2025).

This feature is present only for sounds originating from the front hemifield. In combination with pinna relief (rotation of the pinna away from the side of the head), which is thought to create a 'pinna shadow' that introduces hemifield-dependent features at higher frequencies (Blauert, 1969; Stitt & Katz, 2021), suggests potential cues for front-back discrimination.

1.1.3.2 Sound elevation perception and contributions of spectral shape

At moderate stimulus amplitudes, MAA tasks reveal precision of elevation perception along the midsagittal plane of ~4° (Perrott & Saberi, 1990; Strybel & Fujimoto, 2000). These results are comparable to localization studies assessing elevation perception with broadband stimuli. In such studies, elevation error ranges from ~4° near the horizon to ~10° at extended frontal elevations (Mironovs & Lee, 2018). However, localization perception becomes notoriously poor as sources move overhead (Butler & Humanski, 1992; Morimoto et al., 2003). As source locations move to the rear hemifield, elevation errors get relatively larger while maintaining the same pattern of being better at the horizon (Makous & Middlebrooks, 1990).

Use of the spectral shape cue appears to be significantly impacted by the amplitude of the stimulus. A direct 'level effect' has been demonstrated in the perception of sound elevation, where perception is best between 30-55 dB sensation level (SL; Vliegen & Van Opstal, 2004; Sabin et al., 2005; Zonooz et al., 2019). Positive sensation levels reflect sound amplitudes a certain dB (SL) above threshold. Below 30 dB SL, poor elevation perception has been attributed to interactions between peaks or notches in the sound spectra and a listener's hearing thresholds. Above 55 dB SL, poor perception has been attributed to saturation of cochlear mechanisms, leading to compression of sound spectra at high levels (Macpherson & Middlebrooks, 2000). While not addressed within the context of this document, hearing loss, particularly high-frequency sensorineural hearing loss, is known to degrade elevation perception by preventing access to the prominent spectral features described above (Rakerd et al., 1998).

An implicit challenge in the utilization of spectral shape cues is that the auditory system must make judgments of source locations based on the sound spectra after it has passed through the pinna without direct information about the source spectrum itself (Majdak et al., 2020). The current leading theory for how this takes place, drawn primarily from studies that have evaluated impacts of the spectral shape on elevation perception, is that the auditory system compares the received stimulus to a set of internally stored templates (analogous to HRTFs) that represent all possible locations (Van Opstal et al., 2017; Balachander & Carlile, 2019; Majdak et al., 2020) and 'selects' the direction associated with the template yielding the best match.

Even then, the salient features represented within the stored "HRTFs" remain unclear. While behavioral consequences of manipulating the aforementioned high-frequency notch have been reported (Moore et al., 1989; Rajendran & Gamper, 2019), broadband (or spatially distributed) features within the HRTF may be relevant for perception. Indeed, Zonooz and colleagues (2019) found that even when all features from the 6-9k Hz region (that of the high-frequency notch) were removed, participants were still able to correctly perceive source elevation. They suggest that for their participants to respond correctly to any stimulus that effectively removed the 6-9 kHz notch, consideration of the other frequency regions would be required. Inclusion of wideband spectral features has also been noted to improve modeling techniques aimed at predicting source location from measured HRTFs (Talagala et al., 2014).

1.1.3.3 Contributions of spectral shape to front-back discrimination

Humans' ability to discriminate sources in the front and rear hemifields has been investigated for many decades. In the early 20th century, Rayleigh (1907) commented on the difficulty in front-back discrimination with tonal stimuli. With broadband stimuli, front-back discrimination performance typically ranges from ~2-10% (Makous & Middlebrooks, 1990; Mironovs & Lee, 2018). For long-duration sounds, front-back confusions (FBCs) can be resolved with head movement, where changes in binaural cues over time are most salient (Wightman & Kistler, 1989; Yost, 2023). However, the cues that specifically aid in front-back discrimination of brief sounds remain unclear. The most common result of research into this topic is that there is significant variability of the spectral cues utilized between listeners (Blauert, 1969; Zhang & Hartmann, 2010; Lladó et al., 2022). Additionally, the problem is complicated by the fact that both front-back and elevation perception are supported by the spectral shape cue. Therefore, as discussed in Chapter 4, both perceptual domains are often combined during evaluation of behaviorally relevant spectral features. As such, while many ideas for pertinent front-back spectral regions have been proposed, no clear consensus has emerged. Further investigation of this perceptual capacity is a particular focus of the work presented in this dissertation (see Section 1.5).

1.2. Targeted disruption of spectral cues via experimental stimuli and earmolds

Several prior studies conducted by multiple groups have intentionally disrupted spectral cues in an effort to measure the impact on sound localization abilities that result and, later, to measure the recovery of localization abilities following training (see Section 1.4). One group of studies of this nature uses a form of virtualized stimuli, that is, stimuli presented over headphones, that have been systematically perturbed (Zhang & Hartmann, 2010; Zonooz et al., 2019). Others utilize non-individualized HRTFs, reporting elevation perception that has collapsed to the horizon and FBCs above 20% (Zahorik et al., 2006; Stitt et al., 2019). Still others degrade pinna-derived cues via non-occluding earmolds designed to partially fill the concha. These studies generally report a compression or collapse of elevation responses to the horizon and modest increases of FBCs to ~15% (Hofman et al., 1998; Carlile et al., 2014; Trapeau et al., 2016), defining a performance deficit that is then often modified through the use of auditory training (see Section 1.4).

1.3. Hearing protection devices and the unintended disruption of spectral cues

Other reports of sound localization disruption come from studies utilizing HPDs. HPDs are essential personal protective equipment designed to reduce the amplitude of hazardous noises in the environment to safer levels. Even so, many individuals, who would stand to benefit from use of HPDs, elect not to wear them due to unintended impacts on situational awareness, including sound source localization (reviewed in Audet et al., 2026). Such consequences are a direct result of the application of HPDs to the space immediately within and around the pinnae, thus disrupting open ear spectral shape cues. While introducing perturbations to spectral shape cues that are generally more profound than those introduced in laboratory earmold or virtual audio-based studies, HPDs form an ecological use-case for sound localization research as they represent a real-world application and occurrence of spectral cue distortion. Studies on sound localization performance with HPDs have again reported a loss of systematic elevation perception, and FBC rates of 15-50% depending on the extent of spectral cue disruption (Atherley & Noble, 1970; Russel, 1976; Brown et al., 2015; Audet et al., 2026). Improving our understanding of spectral cues and the distortions caused by such devices stands to inform both hearing conservation practices and the design of HPDs themselves, improving the safety of HPD users. To that end, the impacts of HPDs on sound source localization, and the remediation of those impacts, are major themes explored throughout this dissertation.

1.4. Recovery of disrupted localization through auditory training

While sound source localization abilities are susceptible to disruption via spectral cue distortion, those abilities can be restored via auditory training. The improvement of elevation perception and front-back discrimination have been reported following spectral distortion via non-individualized HRTFs (Zahorik et al., 2006; Poirier-Quinot & Katz, 2021; Meyer & Picinali, 2025), experimentally designed earmolds (Hofman et al., 1998; Carlile et al., 2014; Trapeau et al., 2016), and hearing protection devices (Russel, 1976; Chapters 2, 3, and 4).

The degree of improvement, however, is dependent on multiple factors. While recovery of elevation perception to near-baseline levels has been noted following up to 40 days of exposure to daily life without targeted training (Hofman et al., 1998; Zwiers et al., 2003; Van Wanrooij & Van Opstal, 2005), explicit training rapidly accelerates such adaptation with near-baseline levels of performance observed within 10 days or less (Carlile et al., 2014; Trapeau et al., 2016). Contributing to this acceleration is the presentation of multisensory feedback, often in the form of paired visual stimuli at target locations (Makous & Middlebrooks, 1990; Zonooz & Van Opstal, 2019), additional movement (Casali & Robinette, 2015; Poirier-Quinot & Katz, 2021), or user-adjusted sounds (Trapeau & Schönwiesner, 2018; Stitt et al., 2019).

The magnitude of initial distortion and the duration of training, including the presence or absence of breaks between training blocks (Little et al., 2017) also influences learning gains.

While accommodation to relatively minor spectral distortions can be seen within a single day (Zahorik et al., 2006; Zonooz et al., 2018), studies often report increased benefit when training occurs over multiple consecutive days (Madjak et al., 2010), potentially reflecting contributions of sleep to the consolidation of learning (Carlile, 2014; Irvine, 2018). In other instances, particularly given profound initial disruption of cues, limited gains may occur even with multiple consecutive days of training (treated further in Chapters 2-5).

1.5. Open questions and the foci of this dissertation

While over a century of interdisciplinary research has yielded many insights on the basis of, and constraints upon, sound source localization in basic and applied settings, key questions remain. Over the course of the remaining chapters, three studies completed by the author will be described, demonstrating the author's contribution to the advancement of the sound source localization literature, with an emphasis on factors affecting median plane perception in an applied setting, i.e. during the use of hearing protectors. Each study addresses at least one of the following questions:

1. How is the accommodation of altered spectral cues impacted by the infrequent use of hearing protectors?
2. Can front-back discrimination and elevation perception be concurrently improved through auditory training during the use of hearing protectors?
3. What spectral features are utilized for successful front-back discrimination and how does use of those features change following adaptation to altered spectral cues?

Chapter 2: Audiovisual training rapidly reduces potentially hazardous perceptual errors caused by earplugs

This chapter is a reformatted version of Audet, D. J., Gray, W. O., & Brown, A. D. (2022). Audiovisual training rapidly reduces potentially hazardous perceptual errors caused by earplugs. Hearing research, 414, 108394.

2.1 Introduction

Sensory systems exhibit robust plasticity during ontogeny. In parallel with a developing brain, morphological changes within and between developing sensory structures yield changes in the statistics of sensory inputs, requiring ongoing recalibration of sensory responses against the physical stimuli that drive them (Aslin et al., 1981; Mesulam, 1998; Rauschecker, 1999). Sensory changes throughout adulthood are comparatively slight, and plasticity is generally attenuated. Nonetheless, perturbations of sensory input during adulthood have been shown to induce compensatory changes in multiple sensory modalities, including audition (Merzenich et al., 1984; Knudsen & Brainard, 1991; Roder et al., 1999; Bavelier et al., 2000; Zwiers et al., 2003). The maintenance of such plasticity may serve to accommodate recovery from injury, aging processes, or adaptive coding across dynamic sensory environments (Brenner et al., 2000; Devor & Wall, 1981; Syka, 2002; Dahmen et al., 2010).

In the auditory system, a surprising form of plasticity is observed. The external ears (pinnae) feature ridges and cavities that filter sound in a direction-dependent manner. These pinna-derived spectral shape cues are the primary acoustic cues to sound source location in the vertical dimension and also differentiate forward and rearward sources (Fig. 1A; Batteau, 1967; Musicant & Butler, 1984; Wightman & Kistler, 1999). Pinna morphology and corresponding cues vary across individuals (Wenzel et al., 1993), and the correspondence between spectral shape and source location must be learned through experience (e.g., King et al., 2000). Experimental manipulation of pinna shape distorts spectral shape cues and initially disrupts sound localization; however, after a few weeks of experience with ‘new ears,’ adult human listeners recover near-normal sound localization ability (Hofman et al., 1998, Van Wanrooij & Van Opstal, 2005). Moreover, when the pinnae are restored to their original shape, localization remains accurate. That is, unlike many forms of sensory plasticity, spectral shape cue plasticity does not appear subject to aftereffects, suggesting the emergence of a parallel spectral shape cue representation rather than remapping of a single representation (Hofman et al., 1998).

While altered pinna shape may seem an unnatural distortion, the pinnae naturally increase in size across the lifespan, significantly altering spectral shape cues (Otte et al., 2013), and objects in proximity to the pinnae – including the listener’s own hair – can transiently alter effective spectral shape cues (Treeby et al., 2007). Since its first description (Hofman et al., 1998; cf. Russell et al., 1976), several studies have further examined spectral shape cue plasticity (reviewed in Wright & Zhang, 2006; Carlile, 2014). Most studies have employed chronically worn pinna molds to induce spectral shape distortions, and have evaluated the time course and extent of accommodation, effects of explicit training, and generalization of training effects to novel stimuli (e.g., Van Wanrooij & Van Opstal, 2005; Carlile et al., 2014; Watson et al., 2017). Reports have established that modestly distorted spectral shape cues can be relearned with chronic exposure to the altered cues and that such learning can be accelerated with frequent training sessions (Carlile et al., 2014; Carlile & Blackman, 2014; Zonooz et al., 2018).

Here, we assessed listeners’ ability to accommodate a severe and transient form of spectral shape cue distortion caused by an everyday object: earplugs. While earplugs can effectively reduce the risk of noise induced hearing loss - a problem impacting some 40 million individuals in the US alone (NIDCD, 2019) - prospective users may forgo their use due to undesirable side effects (Berger, 2000; Bolia et al., 2001; Bockstael et al., 2015; Brammer et al., 2014; Nakashima et al., 2018; Reddy et al., 2012). One persistent side effect, a loss of directional hearing ability due to spectral shape cue distortions (Noble & Russell, 1972; Noble, 1981; Noble et al., 1990; Simpson et al., 2005; Zimpfer & Sarafian, 2014; Brown et al., 2015; Smalt et al., 2019), can impact user safety and present unacceptable risk in critical settings (e.g., military settings, Casali et al., 2012; Talcott et al., 2012; Nakashima et al., 2018). Earplug use presents a unique and ecologically important circumstance in which spectral shape cues shift dramatically on a scale of hours or even minutes, such that users do not have the benefit of chronic exposure to altered inputs (cf. Hofman et al., 1998). We demonstrate, using a naturalistic sound localization paradigm, that accommodation of spectral shape cue distortion is possible even under these challenging conditions (cf. Russell, 1976; Casali and Robinette, 2015; Lee and Casali, 2018), and that such accommodation can be significantly accelerated with explicit audiovisual training.

2.2 Materials and methods

2.2.1 Subjects

Twelve adult human subjects aged 23-29 years old (5 male; 7 female) participated in the study. Participation entailed 10 laboratory visits across a period of 10 weeks. All subjects reported normal hearing and passed a hearing test (air conduction thresholds ≤20 dB HL at octave frequencies from 0.25-8 kHz, ≤10 dB interaural asymmetry at all frequencies). Informed consent and all testing procedures followed a protocol approved by the University of Washington Institutional Review Board. Subjects were compensated for their participation.

2.2.2 Apparatus

Sound localization was measured in a hemi-anechoic chamber (ETS-Lindgren HA-160; interior dimensions 9' x 10' x 7'). A circular loudspeaker array (diameter 2.5 m) was positioned inside the chamber with a height-adjustable stool fixed to the floor at the center of the array. Loudspeakers (Orb Audio Mod1) were located at 15° intervals in the azimuthal plane (24 total locations) relative to the center of the array; by convention, the source location directly in front of the subject was designated 0º, source locations to the left were assigned negative azimuth values (e.g., -90º was directly left), and source locations to the right were assigned positive azimuth values (e.g., +90º was directly right). Each loudspeaker was equipped with a tri-color LED mounted centrally that was controlled by an external USB card (National Instruments USB-6501) to provide visual cues during the experiment as described below. Testing was completed in total darkness to limit the use of unintended visual cues. An electromagnetic position-tracking transmitter (Polhemus Fastrak) was suspended from the ceiling, centered above the subject. During all testing, subjects wore an adjustable plastic headband with an affixed lightweight graphite rod. The graphite rod projected forward, terminating at a battery-powered yellow LED that, when worn, was at the center of the subject's neutral forward gaze and rigidly followed the subject's head movement. The position tracking receiver (Polhemus Fastrak) was affixed to the side of the headband to allow for monitoring and recording of head position data.

2.2.3 Stimuli

Stimuli were 100-ms bursts of broadband noise (digitally filtered 0.3-16 kHz, 8th-order Butterworth, new token for each trial), presented at an average level of 70 dBA. Presentation levels were roved ±5 dB across trials to minimize gross level cues, while preserving relative (cross-frequency) level variation that defines the spectral shape cue. (If presentation level is fixed, front-hemifield sources produce slightly but reliably higher levels than rear-hemifield sources - an artificial cue not reflective of real-world listening in which source level varies). Stimuli were designed in MATLAB (MathWorks), synthesized at 44.1 kHz at full 24-bit resolution using a multichannel USB D/A interface (Antelope Orion 32), amplified via 8-channel amplifiers (Crown DriveCore CT875; Harman), and relayed to the target speaker. Stimulus presentation and data acquisition were controlled via MATLAB.

During a majority of testing (excepting 'open ear' pre-training and post-training sessions), subjects wore silicone earplugs (Combat Arms Earplug Gen 4.1; 3M). Devices were worn in 'open mode', selectable with a mechanical switch on the lateral surface of the device, with ear tip size selected to ensure an appropriate fit for each subject. 'Open mode' is explicitly designed to provide less attenuation at typical sound levels (Noise Reduction Rating=8) in order to preserve audibility, e.g., of conversational speech. Presented stimuli (level range 65-75 dBA after roving) were thus comfortably audible in all cases. Prior to earplug placement, subjects were oriented to the device and read a set of manufacturer-provided fitting instructions. Subjects inserted their own earplugs; fit was checked by the experimenter each time.

2.2.4 Procedure

Prior to all testing, the subject was seated at the center of the array and seating height was adjusted to vertically align the ears with the speaker array via a laser-leveling system. Subjects were then instructed to orient toward the loudspeaker at 0º azimuth (‘center’) by aligning the headband-worn yellow LED with a blue LED mounted on the center speaker; for each block of trials, this azimuth and elevation was used as the reference position for the position tracking measurements. In all circumstances, test stimulus presentation was initiated only when the subject was facing center (within ±5° in azimuth and elevation).

Each trial proceeded as follows: While facing center, the subject was presented with a short burst of noise from a randomly selected location. The task was to turn to point the headband-affixed LED at the perceived source location. Once pointing at the perceived location, the subject pressed a button on a handheld keypad to confirm the response. The subject then returned to center to begin the next trial. Subjects were told to remain facing center for the entire duration of each sound and to ensure they had stopped moving before pressing the button to confirm their response. The head was not restrained during testing (‘early’ movement was thus possible), but post-hoc analyses indicated that subjects generally maintained fixation at center (see Discussion). Prior to testing, subjects were read a standard set of instructions (according to the above procedure) and given an opportunity to ask clarifying questions as needed.

Testing was completed over the course of 10 sessions spaced across 10 weeks. Session 1 (pre-training) included (1) a brief block of 24 trials (2 trials at each of 12 locations at 30° increments from 0º) completed with open ears to familiarize the subject with the apparatus and to provide an opportunity for procedural clarification or re-instruction as needed, (2) three blocks of 120 trials (360 total, 15 trials at each of 24 locations at 15º increments from 0º) completed with open ears, and (3) three blocks of 120 trials (360 total, 15 trials at each of the same 24 locations) completed with earplugs. This pre-training testing session lasted approximately 2 hours with breaks provided as needed. (Note: Subject 1904 participated in a pilot session using longer-duration stimuli prior to completing session 1; data not shown). Sessions 2-9 served as training sessions. During this period, subjects were grouped into 'feedback' or 'no-feedback' groups (see below) in a pseudo-random fashion (subjects with odd ID numbers, based on order of enrollment, received feedback, as described below). During each of the eight training sessions, subjects were tested with earplugs only and completed three blocks of 60 trials (180 total, 15 trials at each of 12 locations at 30º increments from 0º; 'trained locations' in Fig. 1B). Training sessions lasted approximately 45 minutes, with no break provided between trial blocks. Continuous training sessions were employed to avoid gaps that can disrupt perceptual learning (Little et al., 2017). Finally, session 10 (post-training) repeated the 'earplug' and 'open ear' tests of session 1 in reverse-order (open ear tested last). Importantly, this session again included all 24 target locations at 15° increments, enabling assessment of spatial generalization of training effects to untrained source locations.

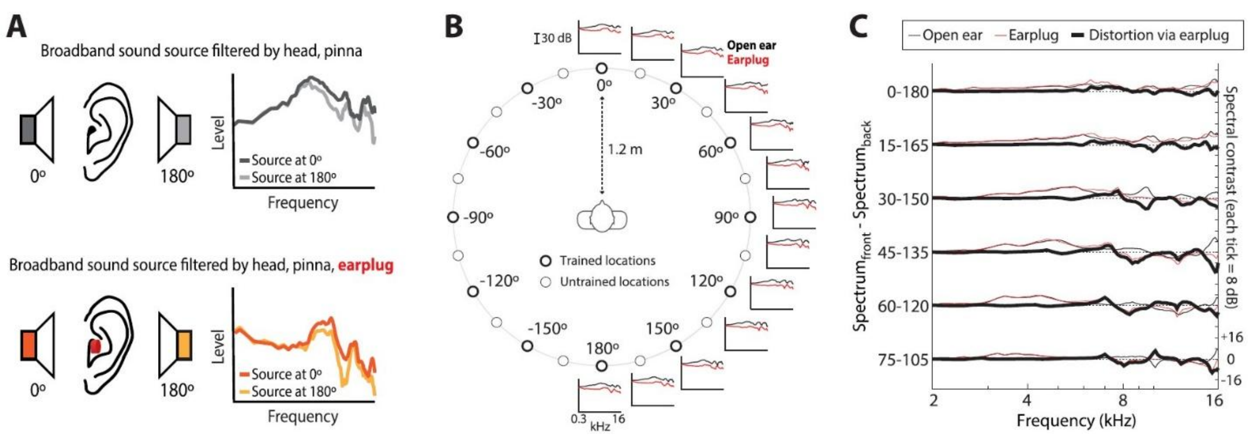

Figure 1. Earplugs distort spatial acoustic information. (A) Sound waves impinging from the front (e.g., 0º azimuth) interact differently with the pinna than sound waves impinging from the back (e.g., 180º). In each case, the unique distributions of sound level across frequency (the 'spectral shape') normally experienced with open ears (upper panel) are altered by earplugs (lower panel). (B) Spectral shapes (insets around perimeter from 0º to 180º) were measured at

15º increments using an acoustic manikin with open ears (black traces) or silicone earplugs designed for military use (red traces). The same loudspeaker array was used for subsequent perceptual testing (a majority of testing used only the loudspeakers at 30º increments; 'Trained locations,' see text). (C) Measured differences in sound spectra for various pairs of front-back locations are plotted across frequency for the open ear condition (thin black lines) and earplug condition (red lines). Thick black lines represent the difference between the open ear and earplug conditions, the presumed primary genesis of the perceptual errors described in the present and previous reports.

2.2.4.1 Feedback procedure

Feedback (presented only to subjects in the 'feedback' group, n=6) consisted of paired auditory and visual stimuli, as follows. A 'correct' localization response was operationally defined as a response within |30°| azimuth of the target.. Responses greater than |30º| from the target (corresponding to >1 loudspeaker from the target during the training portion of the study) were labeled 'incorrect' and are termed large errors (LEs) hereforward. LEs capture responses expected to require reorientation in a real-world localization context, including front-back confusions, consistent with error criteria used in recent related work (see Discussion). Correct responses triggered a green LED to flash at the target location while incorrect (LE) responses triggered a series of events for subjects in the feedback group. First, immediately after the subject's incorrect response was registered and independent of the location of the response, the stimulus was repeated at the target location 8 times, paired with 8 flashes of a red LED at the target location. During this time, the subject was instructed to find and orient to the target location, then maintain fixation on the target speaker until playback ceased. Next, during a 2-s silent period, the subject was instructed to reorient to center (with the aid of a green LED located at 0º). The target stimulus and red LED flash were then repeated 8 additional times at the target location. This additional stage of feedback was intended to re-expose subjects to the spectral information available at the time of the initial error with explicit knowledge of the correct location. The stimulus repetition rate used during both 'sets' of feedback was 2 Hz (100 ms stimulus, 400 ms inter-stimulus interval), for a total feedback sequence duration of ~10 s (including the 2-s pause between 'sets') during which subjects were able to make head movements while experiencing paired auditory and visual stimulation (cf. Carlile et al., 2014). Because feedback sequences added time to each session and were only provided to subjects in the feedback group, an additional pause of 2.6 s was added after every response recorded by the no-feedback group in an effort to balance the total testing time across groups during training sessions. The added time was calculated based on the total time associated with feedback stimulus delivery, assuming a 20% 'incorrect' response rate (consistent with published measurements for previous generations of the Combat Arms earplug, cf. Brown et al., 2015).

2.2.5 Experimental design and analysis

This study employed a longitudinal design, with repeated measures on 12 human subjects across 10 testing sessions. Data, per trial, consisted of a head position track from the starting azimuth (~0º) to the azimuth at the time of button press, taken as the localization response (e.g., Fig. 2A). Examination of head position tracks enabled identification of occasional trials (68 out of 34560 total trials) on which the subject did not register a button press during initial orientation (e.g., did not fully depress the response button) and thus reoriented to center without recording a valid response. These trials were discarded from further analysis. While head elevation was tracked to ensure that the subject was oriented to the horizon (±5º) at the beginning of each trial, stimuli were varied in azimuth only. Each response was thus evaluated in terms of response azimuth versus target azimuth. The LE rate was calculated as the number of LEs (errors >|30º|; see preceding section) divided by the total trial count. Responses below the LE threshold, expected to be of less ecological importance, were quantified in terms of mean angular error - the absolute angular deviation of the response from the target, averaged across target locations. Azimuthal mean angular error is typically minimally affected by bilateral (symmetric) earplug use, whereas the incidence of large errors is markedly increased, an impact of prime concern for end-users (Casali et al., 2012; Brown et al., 2015; Smalt et al., 2019). (For further consideration of localization error criteria, see Discussion Section 4.2.)

Statistical analyses of performance and training effects focused on the incidence of LEs. A sample size of 6 subjects per group (‘feedback’, ‘no-feedback’) was calculated to be sufficient to detect a feedback-induced between-groups difference in LE rate of 10%, assuming a standard deviation of 5% error rate (i.e., Cohen’s d=2), with a power of 0.9. Power was expected to be substantially higher (approaching 1.0) for detection of pre-training increases in LE rate (from open ear to earplug conditions) for which subjects were considered as a single group (n=12) and a large effect size (a ~tenfold increase in LE rate, cf. Brown et al., 2015) was anticipated. In practice, pre-training data were found to be non-normally distributed (Kolmogorov-Smirnov test on both open ear and earplug LE rate $p<0.05$), thus analyses of LEs and summaries thereof, including functions fitted to longitudinal data, were conducted using nonparametric within- and between-subjects comparisons (Wilcoxon signed-rank or Mann-Whitney U test, or exact sign test given asymmetrically distributed open ear LE rates). Tests were conducted using SPSS (Version 26, IBM); standardized test statistics (z values) and p-values are reported for each test along with descriptive values at relevant points in the Results.

2.2.6 Acoustic measurements

In-ear devices have widely variable impacts on the spectral shape information available to wearers, but tend to impact information most at high frequencies, the short wavelengths of which most effectively interact with the acoustically small features of the pinnae (Batteau, 1967; Hofman et al., 1998; Van Wanrooij & Van Opstal, 2005; Zimpfer & Sarafian, 2014; Brown et al., 2015). Fully occluding earplugs can dramatically reshape the spectrum of incoming sound. The precise effect of a given device depends on the geometry and material properties of both the device and the user's pinna (e.g., Brungart et al., 2007), but useful approximations can be obtained using an acoustic manikin with in-ear microphones.

To evaluate the spectral shape distortions produced by the device used in the present study – the Combat Arms Earplug Gen 4.1, an updated version of an earplug long distributed to active-duty U.S. Army personnel (among other usage) – head-related transfer functions (HRTFs) were measured using a manikin designed for hearing protector evaluation (G.R.A.S. 45CB Acoustic Test Fixture, see ANSI (S12.42, 2010). Transfer functions were measured with the same apparatus used for behavioral testing. Excitation signals were exponentially swept sines (0.3-16 kHz); transfer functions were obtained by convolving a time-reversed, amplitude-compensated copy of the excitation signal with the recorded time domain signal and computing the fast Fourier transform of the result (Farina, 2007). To mitigate any potential asymmetries in the chamber sound field, HRTF measurements for a given azimuth angle were taken twice: once with the test fixture facing the 0° speaker (forward) and once with the test fixture facing the 180° speaker (backward). HRTFs measured at corresponding azimuths with respect to the two fixture orientations were then averaged in linear units. Prior to computing fast Fourier transforms, impulse responses were manually trimmed of non-causal artifacts. Measurements completed with earplugs in place (the condition of ultimate interest) were referenced to parallel measurements taken with open ears. Of particular interest in the context of earplug-induced spectral shape cue distortions is the spectral contrast between a 'front' and opposed 'back' location. This contrast defines the spectral information available to differentiate these two locations, the presumed basis of successful localization in the front-back dimension absent visual or auditory motion cues (e.g., Zahorik et al., 2006). Correspondingly, the distortion of this contrast by an earplug was previously found to scale with error rate for a given device (Brown et al., 2015). The Combat Arms Earplug Gen 4.1, the device used in this study, produced large spectral distortions (Fig. 1C) that were similar to those produced by the previous generation (cf. Brown et al., 2015), and are consistent with the relatively high LE rate subsequently observed (see Results).

2.3 Results

2.3.1 Marked initial increase in large errors during earplug use

Human subjects normally localize broadband sounds, such as noise, quickly accurately. Fig. 2A illustrates performance for three representative subjects instructed to localize a sound source (a brief noise burst) via head turn. Stimuli were presented from 24 different source locations in a darkened room (see Materials and Methods). Subjects began each trial facing forward (0º azimuth). Each trace shows, for a given trial, the trajectory of the subject's head position over time in units of degrees relative to the target. Trials with larger positive or negative initial angular distances thus correspond to targets further from 0º (at most 180º, directly behind the subject). While responses to more eccentric targets naturally took longer due to the greater angular distance traversed, subjects generally registered a response in the vicinity of the target within a few seconds (black traces). Occasionally, responses failed to converge near the target, or passed through and diverged from the target. We define such responses (orange traces), simply, as "large errors" (LEs; see Materials and Methods).

Performance with earplugs (Fig. 2B, same subjects) featured a characteristic increase in the occurrence of LEs. LEs frequently entailed mislocalization of a rearward target to the frontal hemifield, manifesting as large initial angular distances that went uncorrected (‘flat’ orange traces near 180º evidencing classical ‘front-back confusions’), or vice-versa, manifesting as small initial angular distances that increased to large magnitudes. LEs are of practical concern in a real-world localization context because they require reorientation (potentially two major reorientations for errant responses in the rear hemifield; cf. Casali et al., 2012; Talcott et al., 2012). LEs can misdirect the visual system and delay (or prevent) an appropriate response to the target, presenting an obvious hazard in critical settings. While LEs occur at a non-zero rate with open ears in some subjects (e.g., S02, Fig. 2A bottom panel, cf. Wightman & Kistler, 1999), earplug use produced a several-fold increase in LEs for all subjects.

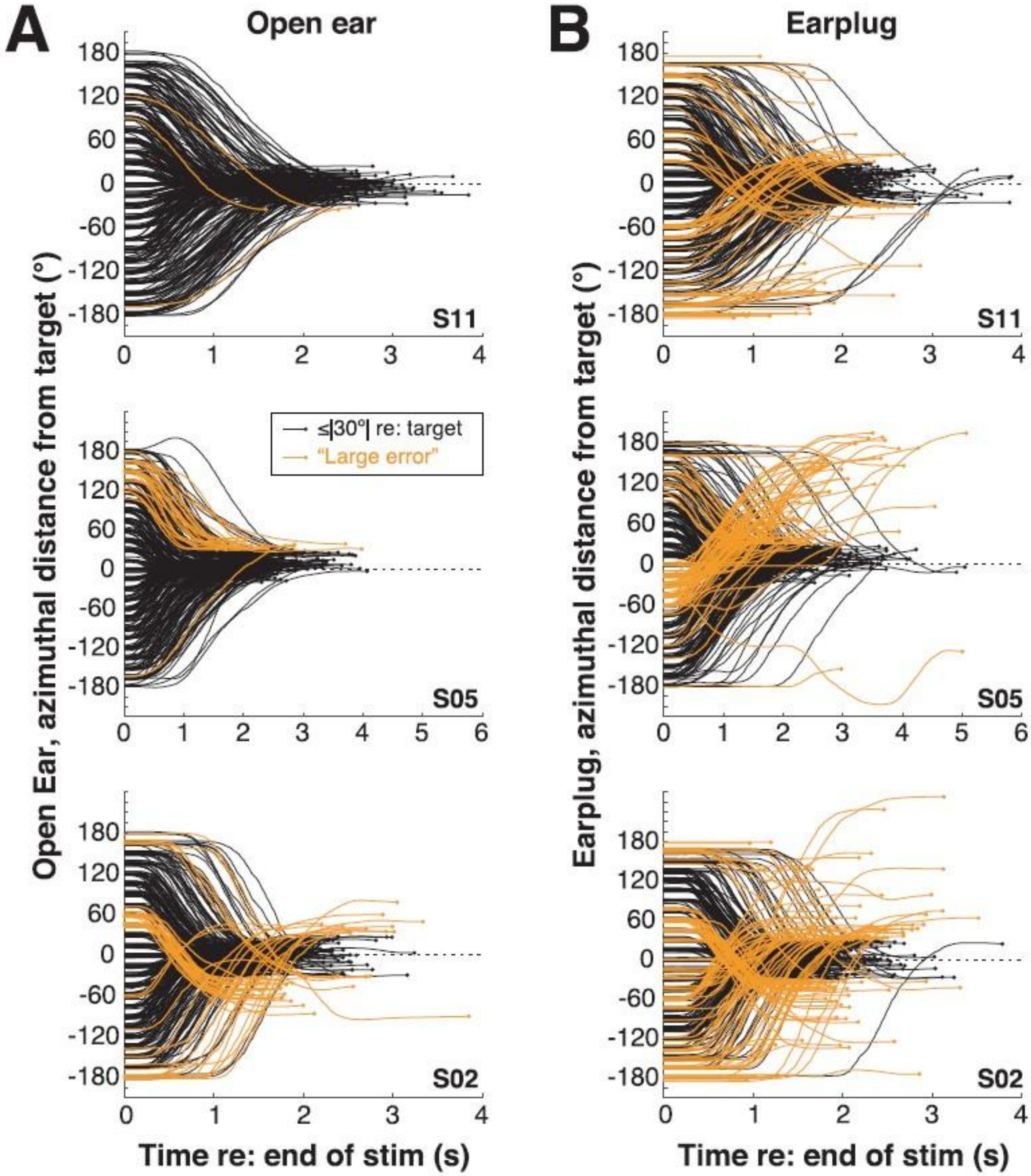

Figure 2. Earplugs cause large errors in orientation to sound sources. (A) Head position as a function of time in three representative subjects (S11, S05, S02) instructed to turn toward a brief noise played from an unknown source location during an initial testing appointment (pre-training). Subjects began each trial facing forward; responses for each trial are plotted in units of degrees of azimuth relative to the source (initial angular distances >|90º| thus indicate sources behind the subject). Subjects completed the task with open ears. Filled circles at the end of each trace indicate the subject's final response. (C) As in A (same subjects), but with the use of silicone earplugs. Whereas responses in the vicinity of the target (black traces) predominated with open ears, the incidence of large errors (>|30º| from target, orange traces; see text) increased markedly during earplug use.

Fig. 3A plots localization data in terms of final response azimuth versus target azimuth pooled across all trials (each point = one trial) for 12 subjects tested during the pre-training session in open ear (left panel) and earplug (right panel) conditions. Vertical spread from y=x indicates response error; dashed lines orthogonal to y=x explicitly demarcate expected locations of front-back confusion responses. To aid in the visualization of observed error magnitudes with respect to the 30º LE criterion, Fig. 3B plots error magnitudes for a subset of response locations (‘trained’ locations, the focus of the following sections) as a swarm chart, in which the horizontal spread of points is scaled according to the density of the underlying distribution. While errors beyond the 30º LE cutoff were relatively uncommon with open ears, errors extending to 180º were quite common with earplugs, with some clustering around putative front-back confusion magnitudes (horizontal black lines as labeled). Interestingly, there was a gap in the error distribution near 90º, the long-term expected value of random performance, suggesting that even very poor performance reflected systematic use of presented cues.

Fig. 3C parses two error measures across subjects: LE rate (left panel) is the proportion of total trials on which subjects recorded an LE, while mean angular error (right panel) is the average final deviation from the target, across all locations, for trials on which LEs did not occur (see Materials and Methods). The observed increase in LE rate during earplug use (open ear median=2.5%; earplug median=26.8%, exact sign test, p=0.0005) is expected, and comparable in magnitude to that observed in previous studies using similar fully occluding silicone or foam earplugs (e.g., Brown et al., 2015; Smalt et al., 2019; see Discussion). In the absence of an LE, the modest increase in mean angular error (open ear median=8.5º; earplug median=11.7º, Wilcoxon signed-rank test, z=2.67, p=0.008) is also expected, but is only a few degrees in magnitude and is of less ecological concern than LEs. Indeed, mean angular error distributions for both open ear and earplug conditions peak in the vicinity of ~10º, reflective of the generally modest impact of bilateral earplugs on localization, excepting their profound impact on the occurrence of LEs, our focus here forward.

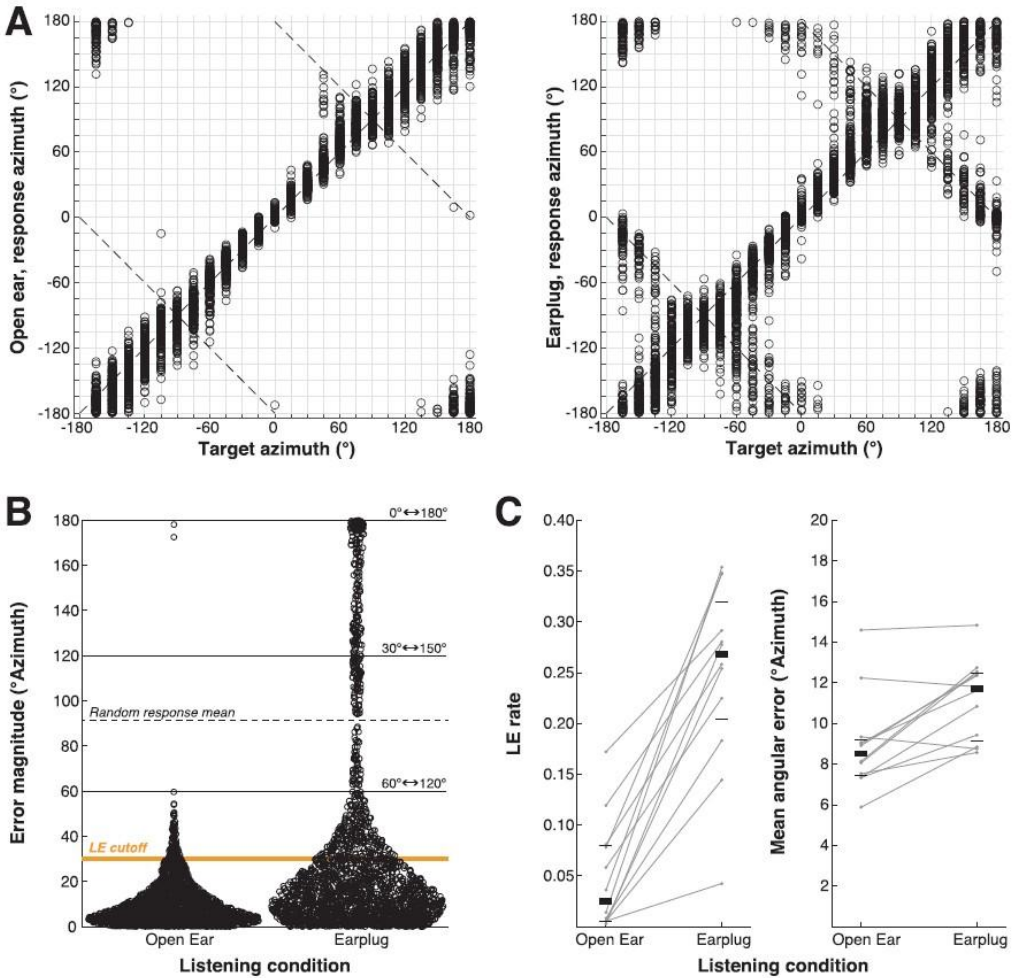

Figure 3. Earplugs markedly increase the incidence of large errors in sound source localization. (A) Response azimuth as a function of target azimuth in open ear (left panel) and earplug (right panel) conditions. Data are pooled across 12 subjects; each point plots the response for a single trial. Dashed lines orthogonal to y=x demarcate putative front-back confusions (see text). (B) Swarm chart of individual trial response error magnitudes across all subjects in open ear and earplug conditions at 'trained locations' used for a majority of subsequent testing. The horizontal spread of points indexes density. Solid black horizontal lines demarcate error magnitudes associated with front-back errors for speaker pairs as labeled. The dotted horizontal line demarcates the error associated with responses orthogonal to the target - the expected mean error obtained via random responses. The orange line demarcates the boundary for operationally defined 'large errors' (LEs; see text). (C) Large error (LE) rate (left panel) and mean angular error for non-LE trials (right panel, see text) averaged across speaker locations for open ear and earplug conditions for each subject (gray points). Hashes give the median and interquartile range across 12 subjects.

2.3.2 Rapid decrease in large errors with audiovisual feedback

Following the initial (pre-training) session, subjects returned once per week for 8 weeks for a training session during which stimuli were presented from a subset of the original 24 source locations ('Trained locations'; see Fig. 1B). All training sessions were completed with earplugs inserted. Fig. 4A plots LEs recorded during the pre-training session and across 8 weekly training sessions for subjects receiving (blue colors) and not receiving (red colors) audiovisual feedback (see Materials and Methods). Cumulative LEs across each session are shown for individual subjects (thinner multicolored lines); the cumulative group median is also shown (thick blue or red lines), and the final per-session LE rate is given at the right ordinate within each panel (bolded points). The two groups recorded very similar performance during the pre-training session (27.0% median LE rate for the feedback group, 28.1% median LE rate for the no-feedback group when computed for only the 12 'trained' locations). While both groups showed improvement over the 8-week training period, the LE rate of the feedback group diverged from (improved more than) that of the no-feedback group during the very first training session (15.0% median LE rate for the feedback group, 21.1% median LE rate for the no-feedback group). By the fourth training session, the median LE rate of the feedback group had improved substantially to 6.7%, less than one-fourth of its pre-training level, while the median LE rate in the no-feedback group had improved modestly to 19.2%. By the final training session, performance had improved further in both groups, but the performance of the feedback group (3.9% median LE rate) remained superior to that of the no-feedback group (12.9%). Capturing individual improvement across the duration of training, Fig. 4B plots pre-training LE rate versus final training LE rate. Individual data for subjects in the feedback (blues) and no-feedback (reds) groups are plotted along with the group medians (bolded red and blue points) and interquartile ranges (horizontal and vertical error bars). Data demonstrate improvement (points below y=x) in all subjects, but notably more improvement in the feedback group (points further from the dashed line), including for feedback vs. no-feedback subjects with similar pre-training LE rates.

Viewed across sessions (Fig. 5A), each subject's accumulation of errors resembled a power function. As the LE count began at zero and the effective slope (LEs per trial) always fell between 0 and 1, data were well-fitted with a single-parameter power function $y=x^b$ (fit R^2 values of 0.93-0.99, median R^2=0.98). Fig. 5B plots the resultant median and interquartile range of fitted power functions for feedback and no-feedback groups. The greater rate of localization improvement in the feedback group - characterized by a decreasing rate of LE accumulation - was captured by the lower value of the fitted slope parameter (b) for the feedback group (blue lines; median b=0.64) versus the no-feedback group (red lines; median b=0.74, Mann-Whitney U Test for independent samples test, z=-2.72, p=0.006). The improvement of both groups relative to pre-training performance was captured by their divergence from the trial-invariant (linear) prediction based on the pre-training earplug LE rate (solid black line with interquartile error lines; open ear linear prediction given by dashed line). Illustrating the decrease in LE rate across trials directly, Fig. 5C plots the gradient of the fitted power functions. While both groups finished training with LE rates lower than their pre-training LE rate (Wilcoxon signed-rank test on final versus pre-training gradient values, feedback group: z=2.20, p=0.028; no-feedback, z=2.20, p=0.028 (identical test statistics a coincidence of the nonparametric analysis), the final LE rate of the feedback group (median fitted LE rate 4.7%) approached the pre-training open-ear LE rate (dashed black lines; Wilcoxon signed-rank test, z=1.78, p=0.075), while that of the no-feedback group (median fitted LE rate 11.4%) remained elevated (z=1.99, p=0.46).

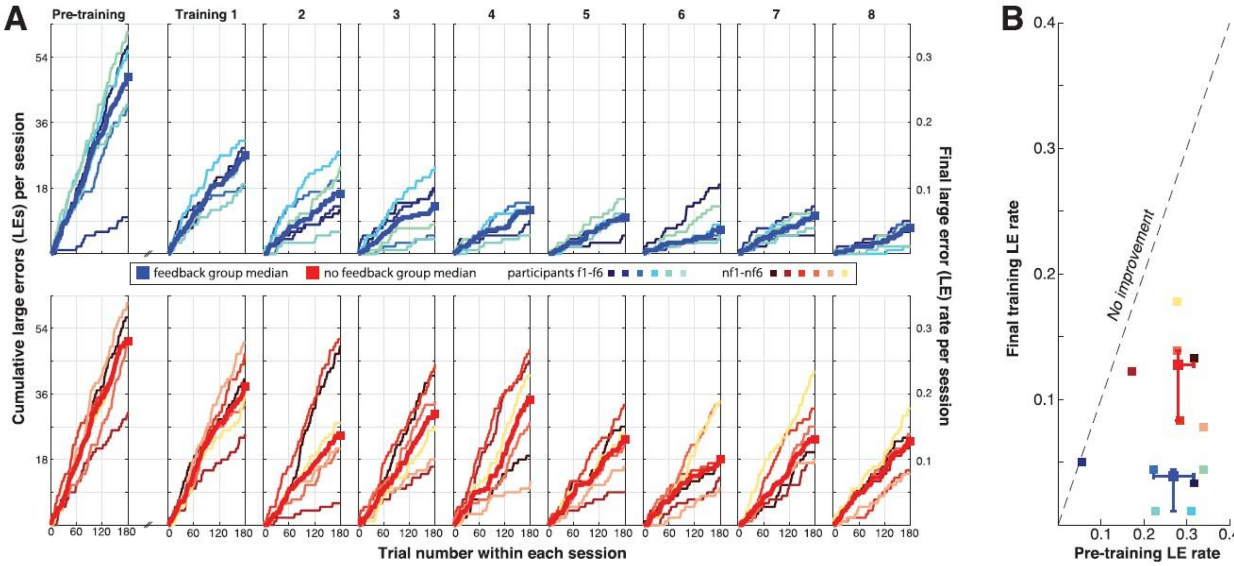

Figure 4. Large sound localization errors dissipate with brief and infrequent training sessions. (A) Each panel plots cumulative large errors (LEs) across 180 trials per testing session (left ordinate). Columns show data for the pre-training 'earplug' session during which no feedback was provided for either group (leftmost column; see also Fig. 3), and 8 weekly training sessions. The upper row plots data for 6 subjects who received explicit audiovisual feedback (f) during training sessions 1-8; the bottom row plots data for 6 subjects who received no explicit feedback (nf). Bold points in each panel show the cumulative median per session; enlarged points at the conclusion of each session show the final median LE count and corresponding final LE rate per session (right ordinate, total LE count divided by 180 total trials). (B) Scatterplot of individual subject LE rate in the final training session versus pre-training earplug condition. Large points give the median and interquartile ranges in each case for feedback (blue) and no-feedback (red) groups.

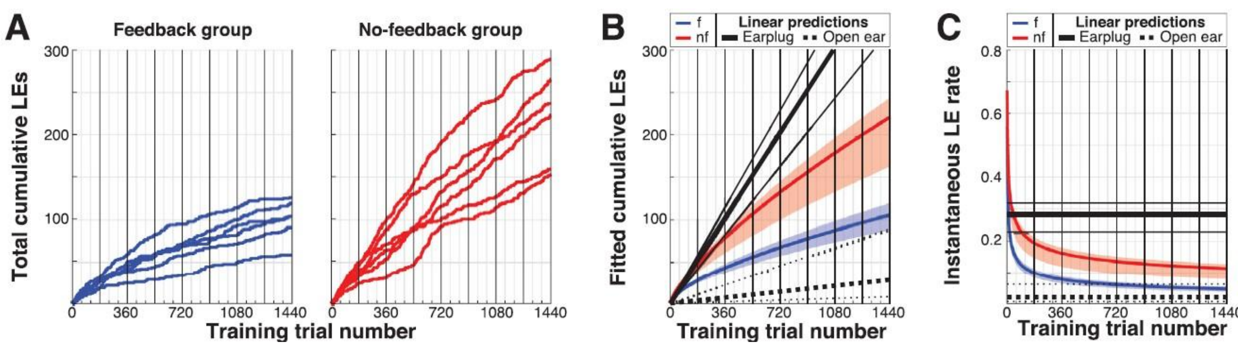

Figure 5. Audiovisual feedback during earplug use augments training benefit. (A) Total large errors (LEs) accumulated per subject across all training sessions for feedback (blue, left) and no-feedback (red, right) groups; bolded vertical lines every 180 trials demarcate sessions. (B) Medians (thick lines) and interquartile range (shaded regions) of total cumulative LEs based on power function fits to individual subject data (R^2=0.93-0.99, median R^2=0.98; see text). Solid and dashed black lines indicate median error counts expected given pre-training earplug (solid) and open ear (dashed) performance; thin lines represent the interquartile range in each case. (C) Instantaneous LE rate based on the gradient of fitted power functions across trials; earplug and open ear predictions (linear slope) are also given (legend as in B).

2.3.3 Generalization of reduced errors and maintenance of open-ear performance

Following the eighth training session, all subjects returned for a final testing session (post-training) during which no feedback was provided and stimuli were presented from all of the original 24 source locations. Testing was completed in earplug and open-ear conditions, enabling (1) assessment of spatial generalization of localization improvement to 'untrained' locations and (2) verification that open-ear performance was unaffected by earplug training. Fig. 6A plots individual post-training versus pre-training LE rates for both earplug (left panel) and open ear (right panel) conditions, with performance at trained locations (filled markers) and untrained locations (open markers) plotted separately in each case. Importantly, greater improvement by the feedback versus no-feedback group continues to be seen (as in Fig. 4B) even for untrained locations. Changes from pre-to-post training are plotted in Fig. 6B, with the size of points weighted by the pre-training LE rate. While not all inclusive, subjects with higher pre-training LE rates (larger points) would have had more potential for improvement, and generally fall lower on the plots. Instances outside of this trend indicate the possibility of limits to the amount of learning to be gained from this training. In the earplug case (upper panel), responses are plotted as normalized change (post-training LE rate divided by pre-training LE rate). The relative decrease in LE rate was similar for trained and untrained locations in both groups (Wilcoxon signed-rank test, feedback group: $z=0.94$, $p=0.345$; no-feedback group: $z=1.36$, $p=0.17$), demonstrating spatial generalization of training benefit (cf. Watson et al., 2017). Finally, the lower panel of Fig. 6B plots the observed change in percent LE rate from pre-training to post-training in the open ear case (normalization is not useful in this case due to infinite/nonsense values produced when dividing by values of/near zero). Changes were generally near zero, demonstrating similar open ear performance before and after earplug training, although somewhat larger shifts occurred in a few individual subjects. Interestingly, the LE rate increased slightly for untrained locations in 5 of 6 'feedback' subjects, albeit not significantly (exact sign test, p=0.063). Nonetheless, minimal change in open ear performance (median shift ≲2% in all cases) is consistent with parallel maintenance of 'accommodated' and 'original' spatial representations as described in prior accounts of spectral shape cue accommodation.

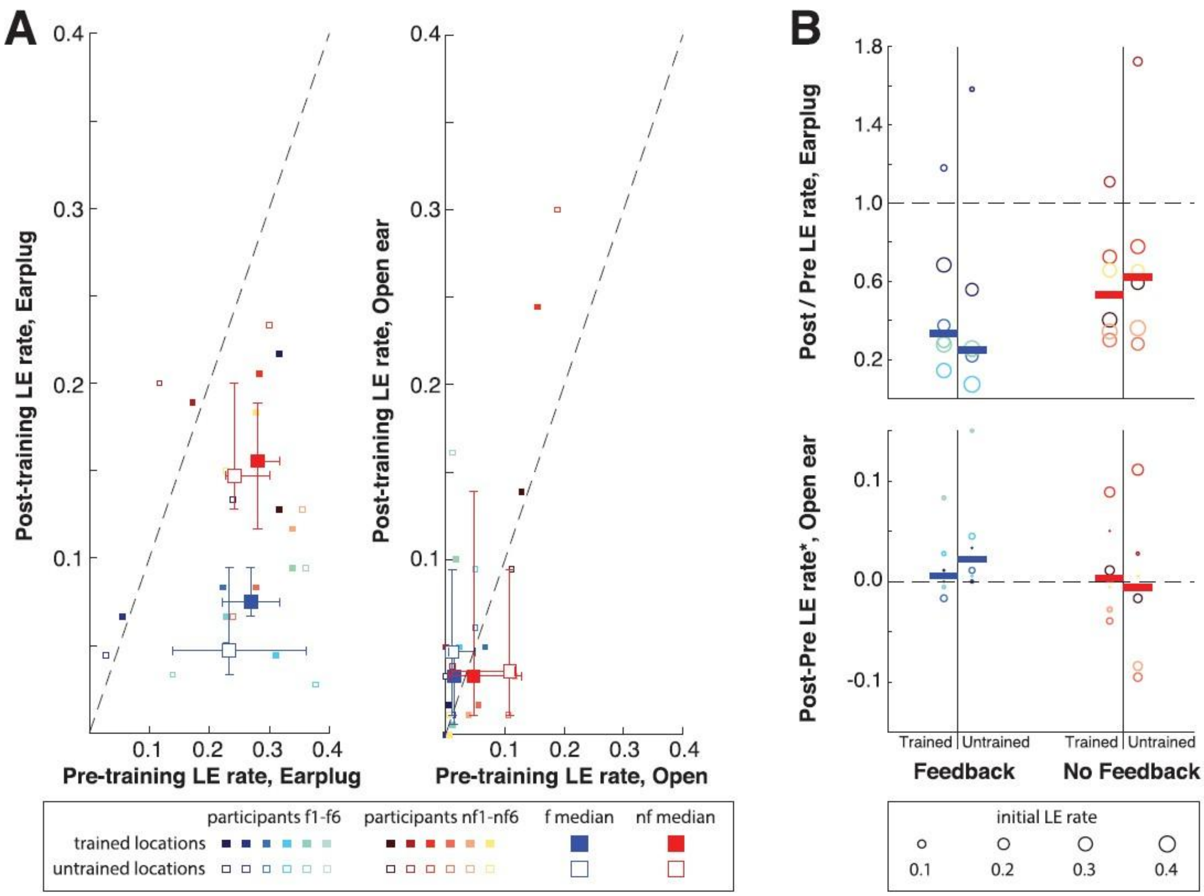

Figure 6. Training benefit generalizes to untrained locations and open-ear performance is maintained. (A) Scatterplots of earplug (left panel) and open ear (right panel) LE rate in post-training versus pre-training sessions for trained locations (filled markers) and untrained locations (open markers). Large points give the median and interquartile range in each case. (C) Change in LE rate from pre- to post-training for trained and untrained locations for feedback and no-feedback groups. Change in earplug performance (upper panel) is expressed as a ratio of post-training/pre-training LE rate, while change in open ear performance (lower panel) is expressed as an absolute difference (see text). In either case, points below the dashed horizontal line represent subjects that recorded improvement (lower LE rate) following training. For display purposes, as small improvements generally reflect subjects with low initial LE rates and thus lower potential to improve, the size of each point is weighted according to the pre-training LE rate (see legend). Thick bars represent group medians.

2.4 Discussion

Peripheral changes in sensory anatomy across the lifespan or from more transient factors can modulate the sensory response elicited by a constant external stimulus (Mistretta, 1984; Otte et al., 2013). Maintenance of accurate perception requires accommodation of such changes. The auditory system possesses a striking capacity to accommodate distortions to sensory anatomy (pinnae and associated spectral shape cues) with chronic exposure to altered inputs, allowing for accurate perception of sound location (see Introduction). The extent of accommodation for more transient changes is less clear. Accommodation of transient changes is a matter of both basic theoretical interest (e.g., whether accommodation is possible without chronic exposure), and also practical importance for users of in-ear devices including earplugs. While earplugs are an essential piece of protective equipment in some settings (military, construction, etc.) their distortion of spectral shape cues is known to degrade situational awareness, compromising safety in those same settings (see Introduction). Here we have demonstrated that (1) auditory spatial plasticity readily extends to the transient and dramatic changes in effective ear shape caused by earplugs, (2) explicit audiovisual feedback during weekly training sessions accelerates accommodation of altered inputs even without exposure to the altered spectral shape cues outside of training sessions, and (3) accommodation effectively generalizes to untrained source locations. These results, building on prior demonstrations of auditory spatial plasticity (e.g., Hofman et al., 1998; Van Wanrooij & Van Opstal, 2005; Carlile & Blackman, 2014; Carlile et al., 2014; Zonooz et al., 2018), point to several fundamental and applied questions for further study, and also suggest that simple training paradigms could be developed to improve the usability of earplugs in critical settings (cf. Russell, 1976; Casali & Robinette, 2015).

2.4.1 Relation to prior studies of spectral shape distortion accommodation

Prior studies of spectral shape cue accommodation have introduced spectral shape distortions of varied severity using low-profile custom earmolds (Hofman et al., 1998; Carlile et al., 2014), virtualized spectral shape cues (Zahorik et al., 2006), and electronic hearing protection and communication systems (Casali & Robinette, 2015; Lee & Casali, 2019). More severe initial spectral shape degradation appears to precipitate worse initial performance (e.g., Van Wanrooij & Van Opstal, 2005; Brown et al., 2015, cf. Russell, 1976) and increase the amount of training required to yield improvement (Van Wanrooij & Van Opstal, 2005; Carlile & Blackman, 2014). In one study imposing modest transient distortion of spectral shape cues via non-individualized virtual head-related transfer functions, two short training sessions with feedback significantly reduced the incidence of front-back confusions (Zahorik et al., 2006). Several studies have imposed substantial long-term (chronic) distortion of spectral shape cues via non-occluding earmolds explicitly designed to alter the effective shape of the pinna. Localization in the vertical plane (Hofman et al., 1998) and front-back dimension (Carlile et al., 2014; Van Wanrooij & Van Opstal, 2005) has been shown to improve over a period of several days to weeks , with significantly reduced front-back confusions after a few days of continuous earmold wear in some subjects (Van Wanrooij & Van Opstal, 2005), particularly with the addition of explicit audiovisual training (e.g., Carlile & Blackman, 2014). An early study of azimuthal discrimination (using 6 discrete locations) across several sessions (Russell, 1976) suggested improvement with training during the transient use of earplugs, but not earmuffs. Earmuffs cover the pinnae entirely and may thus effectively eliminate informative spectral shape cues altogether. Correspondingly, in a more recent study of azimuthal discrimination with electronic hearing protectors, typically designed to preferentially attenuate high-intensity sound, training sessions with feedback spread across 12 days led to significant performance improvement with electronic earplugs, but not with electronic earmuffs (Lee & Casali, 2019; cf. Russell, 1976).

In the present study, spectral shape distortions (see Fig. 1C) were imposed via earplugs with a relatively large form factor (worn in ‘open’ mode; see Methods), and subjects were exposed to distorted cues only during focused but infrequent training sessions. Studies of auditory learning in other contexts have shown that continuous (uninterrupted) training sessions facilitate sustained learning (Little et al., 2017), but that the optimal duration (or number of trials) of training per session is task-specific (Wright & Sabin, 2007; Molloy et al., 2012). Our once-per-week 8-week training schedule approximated the time course of basic training in the armed services (an important use case), while also recapitulating extended periods of non-use. Subjects were provided with approximately 45 minutes (180 trials) of exposure to the distorted spectral cues per week. Despite this limited exposure, we observed relatively rapid improvement in performance during the first training session, particularly in the feedback group, followed by gradual improvement over the remainder of training sessions. As the present study leveraged a naturalistic procedure with which all subjects had extensive practical experience (turning the head toward a sound source) - and as the pre-training session began with several hundred trials of open-ear and earplug listening - such early improvement is presumed to reflect rapid perceptual learning (cf. Hawkey et al., 2004), though contributions of procedural learning cannot be ruled out. The feedback paradigm we employed was designed to provide combined auditory and visual stimulation, which has previously been shown to elicit greater learning than visual feedback alone (cf. Carlile et al., 2014). This result suggests that repeated exposure to auditory spatial cues per se, and not mere knowledge of the locations with which those cues are associated, is an important aspect of effective localization training.

Despite notable improvement with training, accommodation remained incomplete at the end of the training period, with higher LE rates with earplugs than without. It is possible that a more intensive daily training regimen, additional training sessions, or an improved feedback paradigm would have netted more complete accommodation (cf. Wright & Sabin, 2007). Additionally, while efforts were made to achieve reproducible earplug placement from session to session, earplugs were inserted anew each session, introducing the possibility of variation in the resultant distortions of spectral shape. Such variability could have increased session-to-session variability and perhaps also constrained total training benefit. However, elimination of such variability would not be realistic in a real-world use case, and it is notable that the accommodation demonstrated must have occurred despite such variability. Further work is required to specify the practical limits of spectral shape cue accommodation with respect to the severity of initial spectral shape cue distortion and the time course of exposure to distorted cues.

2.4.1.1 Training benefit in the absence of feedback

Notably, while subjects receiving explicit feedback improved more rapidly and more completely than subjects not receiving feedback, improvement relative to pre-training performance was observed in both groups (e.g., Fig. 5). Perceptual improvement in the absence of explicit feedback has been described in both auditory spatial (cf. Russell, 1976, Carlile et al., 2014) and visual spatial (Watanabe et al., 2001; Petrov et al., 2006) domains. Here we explore three potential explanations for the improvement observed in the present study in the absence of explicit feedback.

First, while stimulus presentation only commenced once the subject's head was pointed forward, and subjects were explicitly instructed to fixate at center during stimulus presentation, no head restraint was used. Though stimuli were purposefully brief (100 ms), subtle head movements could have provided for exposure to dynamic spatial acoustic information, which can itself disambiguate forward from rearward sources (Wightman & Kistler, 1999). To evaluate this possibility, and to gain more general insight on potential temporal effects of earplug use, we examined the temporal dynamics of localization responses observed in both groups. Figure 7A plots the average total response time (in seconds) per trial for non-LE trials only (avoiding the influence of long-duration LEs). No systematic effect of earplug use on response time was observed during the pre-training session (upper panel); effects of training were variable and generally minimal across subjects (lower panel). Regarding the specific issue of 'early' head movement during training, Fig. 7B plots exemplar head position traces for all trials at the -30º source location for the first (left column) and final (right column) training session for subjects in 'feedback' (upper rows) and 'no-feedback' (lower rows) groups; total response time is given by the x-coordinate of the final point in each trace. The right inset within each panel plots an expanded view of the first 100 ms of each trace, showing head movement immediately following stimulus offset. While a slight negative-trending slope (movement toward the target) is evident toward the end of this frame on some trials in some subjects, responses are generally 'flat' at the beginning of each window, demonstrating that subjects maintained fixation toward the center speaker as instructed. This pattern recurred at other source locations, suggesting that 'early' head movements were minimal. Summarizing the lack of a relationship between the temporal dynamics of subjects' responses and localization performance, Fig. 7C plots the total LEs observed during the final training session against changes in total response time (upper panel) and cumulative 'early' head movement (i.e., accumulated degrees of movement across the first 100 ms of all non-LE trials; lower panel) for feedback and no-feedback groups. Neither temporal measure was meaningfully associated with performance (note: this held true for correlations computed within speaker locations and across a variety of other summary metrics (data not shown).

A second possibility is that, in making repeated localization judgments, subjects, including no-feedback subjects, benefitted from a kind of implicit statistical learning (cf. Watanabe et al., 2001). While earplugs severely distort spectral shape cues, some spectral features (representing potential cues) are conserved, such as the modestly higher gain at frequencies ~3-7 kHz for forward than rearward source locations (see Fig. 1C; cf. Zahorik et al., 2006). While the initial transformation of spectral shape, including the introduction of unfamiliar spectral notches and peaks, is clearly disruptive to performance, repeated exposure to forward and rearward sources may enable statistical re-learning of altered spectral shape cues that still feature systematic spatial dependencies resembling some aspects of open-ear cues. Future work incorporating individualized HRTF measurements completed with and without earplugs would provide more detailed insight on the statistics of spectral shape cues experienced with respect to observed localization performance.

A third, and rather simpler, possibility is that subjects benefited from exposure to uncontrolled sources of environmental cues. For example, subjects experienced brief but immersive exposure to the environment outside the testing chamber before and after each formal training session. While such exposure was quite limited (<5 minutes per week), subjects inserted and removed their earplugs outside of the testing chamber while verbally interacting with the experimenter. Such exposures, accumulating over the course of several weeks, could have provided a basis for some level of accommodation. A related consideration is that all listeners produced some amount of self-generated sound during task performance (including footfalls during orienting responses to the rear hemifield) that could have provided subtle but consistent exposure to acoustic input from known 'sources' over the course of the training program. Systematic variation of passive environmental exposure could provide an interesting target for future study, bearing on the prospective learning already experienced by earplug users in real-world settings.

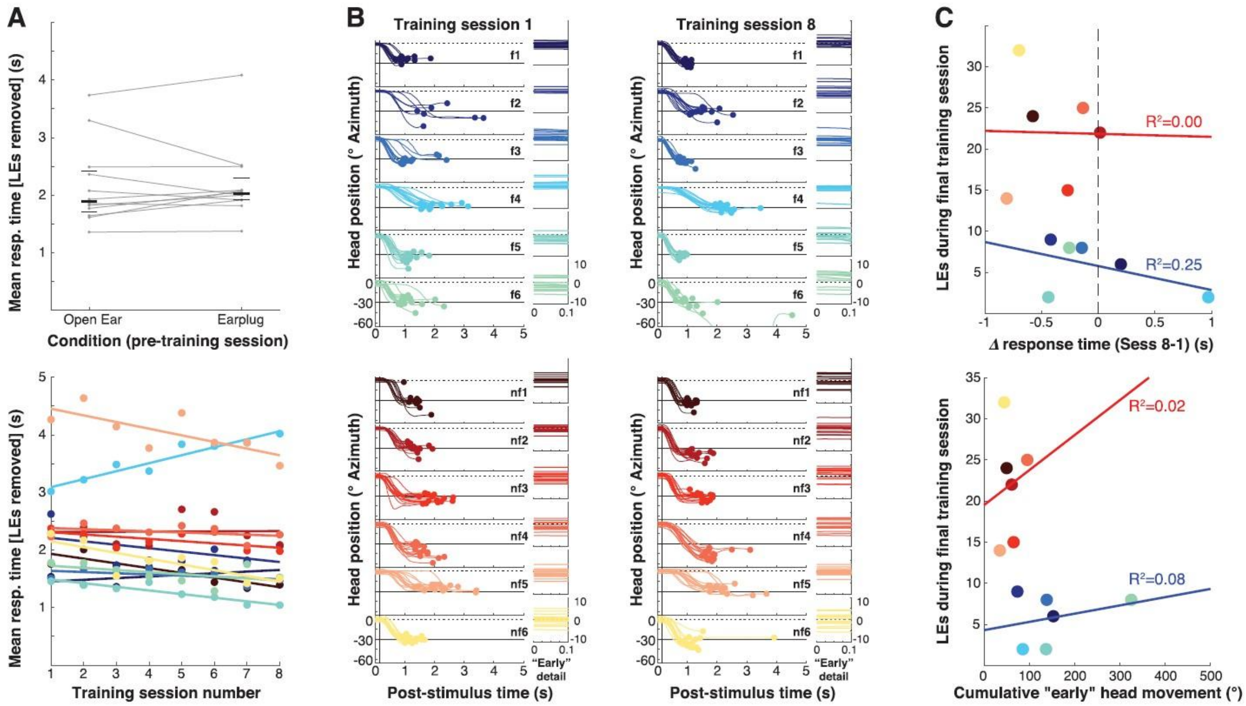

Figure 7. Earplug use and perceptual training benefit show little relation with temporal dynamics of localization responses. (A) (Upper panel) After excluding large-error trials, earplug use did not systematically affect localization response time (averaged across all locations) during the pre-training session. (Lower panel) Average response time across locations for all subjects in each training session (subject color legend given in panel B). Linear fits are plotted in each case, showing little systematic variation for most subjects. (B) Example orienting responses illustrate the temporal dynamics of localization responses, including 'early' movement (insets) evident near the offset of the stimulus in some subjects (see text). Responses (LEs excluded, thus the number of points varies across subjects) are shown for the -30º target location during training sessions 1 (left column) and 8 (right column) for the feedback (upper rows) and no-feedback (lower rows) groups. Negative-trending slopes in right-hand insets indicate 'early' movement occurring near the stimulus offset (see text). (C) (Upper panel) LEs during the final training session versus the change (Δ) in response time over the duration of training. Decreased response time is not associated with lower LE rate in either group (see text). (Lower panel) LEs during the final training session versus the total cumulative 'early' head movement (given in ºAzimuth). Increased early head movement was not associated with lower LE rate in either group (see text).

2.4.1.2 Spatial generalization of training benefit

One critical finding common to both feedback and no-feedback groups was that training at several discrete angles of incidence with earplug-altered spectral shape cues was sufficient to improve localization at untrained source angles. This finding, also reported in an extended-wear earmold study (Watson et al., 2017), suggests that the auditory system effectively interpolates (indeed statistically learns) a functional spectral shape ‘spatial’ representation from a relatively sparse set of spectral shape training data - a certain requirement for an implementable earplug training protocol, since it would be impossible to train all potential source locations. Further reductions in the number of trained locations could increase the efficiency of training, though errors would be expected to increase given an spectral shape representation interpolated from too few trained inputs. A systematic investigation of training ‘density’ versus generalization could prove basically and practically informative.

2.4.2 A note on the definition of localization error boundaries

In addition to variation in the means/extent of spectral shape cue distortion and the testing/training paradigm implemented, prior studies have also varied considerably in the definition of localization errors/error types. Nonetheless, two broad categories can be identified: (1) errors of spread, which are small in magnitude and typical of open-ear and ‘modified’ localization performance alike, and (2) gross errors (often falling into a different hemifield or quadrant than the target), which occur at much higher rates (at least initially) in ‘modified’ than open-ear testing conditions. Errors of spread are generally calculated after removal of gross errors, and expressed as an averaged degree value - although caution should be taken in the interpretation of such values obtained using discretized response paradigms (e.g., discrimination of labeled speaker locations), which naturally quantize ‘spread’ values. The quantification of gross errors has been substantially more variable, with error boundaries starting anywhere between |15º| and |45º|. Zahorik et al. (2006) proposed criteria for front-back error labeling that depended on the target, response, and subject-specific mean angular error. Carlile et at. (2014) defined front-back confusions as hemispheric errors that were >|25°| from the interaural axis. Casali and colleagues defined a “ballpark” boundary for errors >|22.5º| (not in the ballpark) in one study (Casali et al., 2012) and errors >|15º| in another (Lee and Casali, 2018). Brown et al. (2015) evaluated the incidence of errors >|45º|, noting that a majority of such errors appeared to reflect front-back confusions. Most recently, Smalt and colleagues (2019) proposed a >|30°| categorical threshold, with the provision that such errors occurred in a quadrant different from the source (defined by intersection of front/back and left/right hemifields; termed “quadrant errors”), though the ecological significance of quadrant boundaries above and beyond the >|30°| threshold is unclear. In view of prior literature, the training paradigm employed in the present study, and indeed the ecology of spatial orientation, we enforced a simple >|30º| ‘large error’ criterion. This threshold captures orientation errors that place the target, at best, well outside of central vision (cf. Strasburger et al., 2011), as well as true front-back confusions that place the target outside the visual field altogether - in either case potentially delaying (or preventing) appropriate responses to critical signals.

2.4.3 Future directions

The accommodation of degraded spectral shape cues demonstrated here underscores the rapidity and robustness of this unusual form of sensory plasticity. Additional studies should work to better specify limitations on the completeness of spectral shape cue accommodation (i.e., convergence toward open ear performance) and its generalization to novel conditions, toward improved basic understanding of the underlying adaptive processes. Such efforts will also support the development of practical training applications for end-users of earplugs and other hearing devices (cf. Lee and Casali, 2018).

2.5 Acknowledgements

This work was funded by the University of Washington Royalty Research Fund [ADB, WOG] and the National Institute on Deafness and Other Communication Disorders (Grant No. T32-DC005361 [DJA]). The authors wish to thank Drs. Nathaniel Greene, Erick Gallun, and Daniel Tollin for useful discussions.

Chapter 3: Impacts of audiovisual training on two-dimensional sound localization with earplugs

This chapter is a slightly modified version of: Audet, D. J., Hunsaker, A., Curry, L., & Brown, A. D. (submitted). Impacts of audiovisual training on two-dimensional sound localization with earplugs

3.1 Introduction

Spatial hearing relies on both binaural and monaural cues. The binaural cues, interaural time difference (ITD) and interaural level difference (ILD), primarily aid in horizontal localization (Rayleigh, 1907). The monaural spectral shape cue, derived from interaction of impinging high-frequency sound with the complex shape of the pinna and adjacent anatomy, primarily aids in front-back and vertical localization (Batteau, 1967; Musicant & Butler, 1984; Wightman & Kistler, 1999). With all cues working in tandem under optimal conditions, humans can accurately and precisely localize sound in both azimuth and elevation (e.g., Makous & Middlebrooks, 1990; Carlile et al., 1997; reviewed in Middlebrooks & Green, 1991). However, in real-world settings, many factors can intervene to degrade spatial hearing cues and resultant performance.

Hearing devices that modify the signal reaching the ear(s) are a common source of spatial cue degradation (reviewed in Zheng et al., 2022; Smalt et al., 2020). Even symmetrically placed/programmed bilateral hearing devices modestly impact binaural cues and sound source localization in azimuth (e.g., Van Den Bogaert et al., 2006; Brown et al., 2016). Impacts on monaural cues are often more severe; in- and on-ear devices alter the effective shape of the pinna, and provide nonuniform gain or attenuation across frequency, dramatically altering the spectral features of the transmitted signal and associated spectral shape cues. These disruptions result in a decreased ability to discern front from back and up from down, even after accounting for source audibility (Reddy et al., 2012; Bolia et al., 2000; Denk et al., 2019). On average, without visual cues, earplug users confuse frontward and rearward targets ~20% of the time, although the incidence of error varies across devices (e.g., Vause & Grantham, 1999; Brown et al., 2015; Chapter 2). For example, error rate increases toward chance (i.e. 50% correct) during the use of earmuffs, which effectively eliminate pinna cues altogether (Casali & Robinette, 2015). In the vertical domain, disruption of the spectral shape cue generally leads to a collapse of responses toward the horizon, with subjects demonstrating a lack of sensitivity to changes in source elevation (Hofman et al., 1998; Zahorik et al., 2006; Friedrich & Schönwiesner, 2025). Such disruptions to source localization are major perceptual errors that can misdirect the visual system (Montagne & Zhou, 2018) and delay responses to critical signals, decreasing user safety and device satisfaction, ultimately leading to device disuse (Byrne & Noble, 1998; Wong et al., 2003; Simpson et al., 2005 Casali et al., 2009; Montagne & Zhou, 2018).

Localization errors can sometimes be prevented or corrected. For longer duration sounds, listeners can rotate their heads, effectively allowing the auditory system to collect more samples of the incoming sound, with dynamic binaural cues serving to disambiguate source hemifield (Wightman & Kistler, 1999; Archer-Boyd et al., 2015). However, short-duration and unexpected sounds (e.g. single words, impulsive signals, etc.) do not permit useful head movements or prior preparation. Sound source localization errors in this context are more successfully corrected via training. Training and subsequent improvement in front-back perception has been demonstrated in many contexts, such as following unilateral earplugging (Van Wanrooij & Van Opstal, 2005; Kacelnik et al., 2006; Kumpik et al., 2010; Irving & Moore, 2011), with virtualized stimuli filtered through individualized head-related transfer functions (HRTFs; Madjak et al., 2010), with virtualized stimuli filtered through unfamiliar HRTFs (Zahorik et al., 2006), and with bilateral use of passive hearing protectors (e.g. Russell, 1976; Casali & Robinette, 2015; Chapter 2; Golob et al., 2024). However, not all studies assess or report improvement in both front-back and elevation equally. Some that report improvement in both front-back and elevation only present sound sources in the frontal hemifield (Van Wanrooij & Van Opstal, 2005; Trapeau et al., 2020; Valzolgher et al., 2020; Friedrich & Schönwiesner, 2025), while others present sources in 360° space with relatively minor disruption of the spectral shape cues (Zahorik et al., 2006; Madjak et al., 2010; Poirier-Quinot & Katz, 2021). Questions remain regarding the simultaneous improvement in front-back discrimination and vertical perception following significant disruption of spectral shape cues.

The present study aimed to evaluate localization performance and improvement in azimuth and elevation following significant spectral shape cue disruption via binaurally inserted earplugs. Localization improvement is assessed throughout a 10-day audiovisual training paradigm. Results are evaluated in terms of changes in front-back confusions and elevation perception, generalization to untrained source locations, and associations with individual audiometric factors.

3.2 Materials and methods

3.2.1 Subjects and study enrollment

Nine adult human subjects aged 18-27 years old (5 male; 4 female) participated in the study. All subjects reported normal hearing and passed a hearing test (air conduction thresholds ≤30 dB HL at octave frequencies from 0.25-8 kHz, ≤10 dB interaural asymmetry at all frequencies). Extended high frequency hearing thresholds (9-16 kHz) were also measured but were not used as exclusionary criteria. All subjects had previously participated in a study evaluating impacts of hearing protectors on auditory perception, including sound localization (Audet et al., 2026). Subjects invited to participate in the present study all demonstrated a large increase in localization errors during earplug use (specifically, a minimum 20% increase in errors |>30º| relative to open ear listening while using the tested earplug). The rationale for this inclusionary criterion was to enroll subjects with known potential to improve performance in one or more spatial dimensions, i.e., to avoid attempting to train subjects already near a performance ceiling with little room to improve. The open ear and earplug measurement sessions that established candidacy for the study comprised the baseline "pretest" sessions (Sessions 1 and 2) described below. Including these sessions, study participation entailed 14 laboratory visits across a period of 6 weeks. Subjects were remunerated for their participation. Written informed consent and all testing procedures followed a protocol approved by the University of Washington Institutional Review Board (STUDY00004023, last approved on April 20, 2023).

3.2.2 Apparatus

Sound localization was measured in a hemi-anechoic chamber (ETS-Lindgren HA-160; interior dimensions 9' x 10' x 7'). A circular loudspeaker array (diameter 2.5 m) was positioned inside the chamber with a height-adjustable stool fixed to the floor at the center of the array. Tested loudspeakers (Gallo Acoustics Micro) were located at 45° intervals in the azimuthal plane and 15° intervals in the vertical plane relative to the center of the array; by convention, the source location directly in front of the subject was designated 0º ('center'), source locations to the left were assigned negative azimuth values (e.g., -90º was directly left), and source locations to the right were assigned positive azimuth values (e.g., +90º was directly right). Each loudspeaker was equipped with a tri-color LED mounted centrally that was controlled by an external USB card (National Instruments USB-6501) to provide visual cues during the experiment as described below. Testing was completed in total darkness to limit the use of unintended visual cues. An inertial measurement unit (WitMotion WT901BLECL) securely fastened to an adjustable headband was used to measure head motion relative to the standard starting position on each trial (see 'Behavioral Procedure' below). The headband also supported a graphite rod that projected forward, terminating at a battery-powered yellow LED; when worn, the LED was at the center of the subject's neutral forward gaze and rigidly followed the subject's head movement.

3.2.3 Stimuli

Stimuli were 100-ms bursts of broadband noise (digitally filtered 0.3-15 kHz, 8th-order Butterworth, new token for each trial), presented at an average level of 70 dBA. Presentation levels were roved ±5 dB across trials to minimize gross level cues; if presentation level is fixed, front-hemifield sources produce slightly but reliably higher levels than rear-hemifield sources, potentially providing an artificial cue not reflective of real-world listening in which source level varies. Stimuli were designed in MATLAB (MathWorks, Natick), synthesized at 44.1 kHz at full 24-bit resolution using a multichannel USB D/A interface (Antelope Galaxy 64), amplified via 8-channel amplifiers (Crown DriveCore CT875; Harman), and relayed to the target loudspeaker. Trial-to-trial stimulus presentation and data acquisition were also controlled via MATLAB.

During a majority of testing (excepting 'open ear' pretest and posttest sessions), subjects wore silicone earplugs (Combat Arms Earplug Gen 4.1; 3M). Devices were worn in 'open mode', selectable with a mechanical switch on the lateral surface of the device, with ear tip size selected to ensure an appropriate fit for each subject (Figure 1A). 'Open mode' is explicitly designed to provide low attenuation at typical sound levels (Noise Reduction Rating=8) to preserve audibility, e.g., of conversational speech. The stimuli presented (level range 65-75 dBA after roving) were thus comfortably audible in all cases. Prior to hearing protection device (HPD) placement, subjects were oriented to the device and read a set of manufacturer-provided fitting instructions. Subjects inserted their own earplugs; fit was checked and corrected as needed by the experimenter.

3.2.4 Acoustic measurements

Given a broadband stimulus of moderate intensity, normal-hearing listeners with open ears are assumed to have access to all spatial cues (both binaural and monaural) potentially present in the stimulus. Reduced access to a given spectral region – for example, due to hearing loss (Wang et al., 2021), occlusion of the ear canals by earplugs, or both factors – could reduce access to one or more spatial cues. In order to evaluate the level of transmitted signals for later evaluation of audibility at the single-subject level, a binaural manikin (45CB, G.R.A.S.) was used to measure open-ear and earplug-attenuated experimental stimuli. The manikin was positioned at the center of the loudspeaker array used for human subject testing. Noise burst stimuli identical to those presented to subjects were played from the same loudspeaker locations used for training (16 replicates per location to facilitate de-noising through signal averaging). Resultant in-ear signals were measured via the manikin's high-sensitivity ear simulators (G.R.A.S. 43BB) with open ears and with earplugs inserted, digitized (Antelope Galaxy64), and saved to disk for offline processing. Recorded noise bursts were evaluated in the frequency domain via fast Fourier transform; averaged magnitude spectra (expressed in dB SPL) were then compared to subject audiograms.

3.2.5 Behavioral procedures

This study utilized 2 different sets of speaker locations for testing and training of localization abilities. Localization testing (pretest and posttest) sessions were completed using a 24-speaker array with speakers positioned at eight unique azimuths (every 45° from 0-315°) and three unique elevations (-30°, 0°, +30°). Localization training sessions utilized a 12-speaker mid-sagittal array including six source locations at 0° azimuth and six locations at 180° azimuth with elevations spanning 15° increments from -30° to +45° in each case. Figure 1B provides a visual representation of the speaker arrays used for testing and training.

3.2.5.1 Sound source localization procedure

Prior to all testing, the subject was seated at the center of the array. Seat height was adjusted vertically to align the ears with the speaker array via a laser-leveling system. Subjects were then instructed to orient toward the loudspeaker at 0º azimuth (‘center’) by aligning the headband-worn yellow LED with a blue LED mounted on the center speaker; for each block of trials, this azimuth and elevation was used as the reference position for the position tracking measurements.

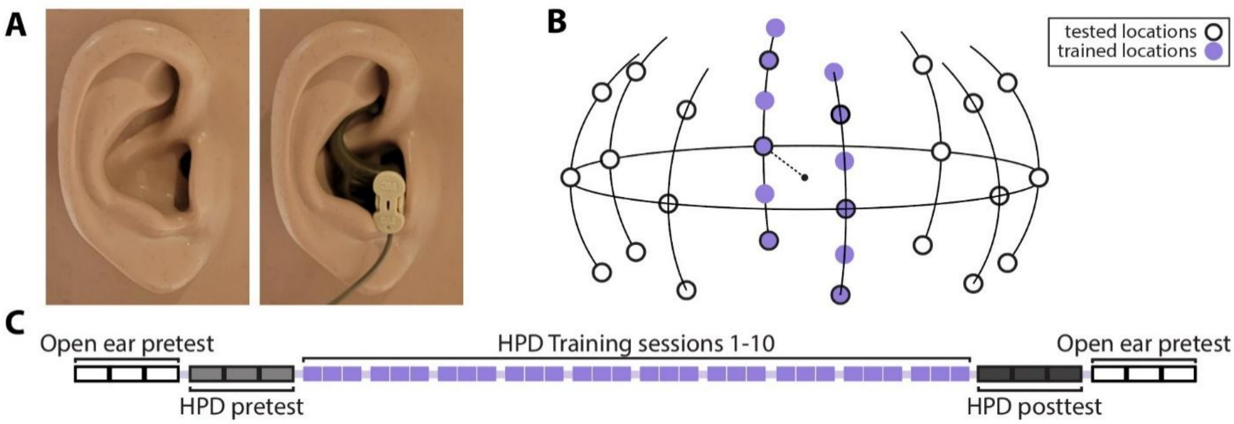

Figure 1. Key methodological components. (A) Pre- and posttest sessions were completed with open ears and with bilaterally inserted earplugs (Combat Arms Earplug Gen 4.1; HPD). This device includes a retention hook placed in the cymba concha and a lateral toggle switch. The toggle switch was pressed down for all subjects, setting the device in a low-attenuation mode (see text). Subjects always wore earplugs during training sessions. (B) Two overlapping sets of speaker locations were employed. Pretest and posttest speakers (black edged circles) spanned 0 to 315° in azimuth at 45° increments and -30 to 30° in elevation at 30° increments. Training speakers (filled circles) were located on the mid-sagittal plane only, at two azimuths (0 and 180°) and six elevations (spanning -30 to 45° at 15° increments). (C) Schematic of the study timeline. Subjects first completed open ear and HPD pretest, followed by 10 consecutive sessions of training, ending participation with HPD and open ear posttest.

Each trial proceeded as follows: While facing center, the subject was presented with a short burst of noise from a randomly selected location. The task was to turn and point the headband-affixed LED at the perceived source location. Once pointing at the perceived location, the subject pressed a button on a handheld keypad to confirm the response. The subject then returned to 'center' to begin the next trial. Subjects were told to remain facing center for the entire duration of each sound and to ensure they had stopped moving before pressing the button to confirm their response. The lights in the testing chamber remained off throughout all measurements. The head was not restrained during testing; while this introduces the possibility of 'early' movement, a prior study utilizing similar procedures revealed no significant head movement prior to cessation of the stimuli (Chapter 2). Prior to testing, subjects were read a standard set of instructions (according to the above procedure) and given an opportunity to ask clarifying questions as needed.

Testing was completed over the course of 14 sessions (Figure 1C) spaced across 6 weeks. Session 1 (open-ear pretesting) included (1) a brief block of 24 trials (1 trial at each of 24 locations) completed with open ears to familiarize the subject with the apparatus and to provide an opportunity for procedural clarification or re-instruction as needed, (2) three blocks of 120 trials (360 total, 15 trials at each of 24 locations) completed with open ears. Session 2 (HPD pretesting) consisted of three blocks of 120 trials (360 total, 15 trials at each of the same 24 locations) completed with the HPD in place. Sessions 1 and 2 lasted approximately two hours each and generally occurred over the span of two weeks (one session per week). Sessions 3-12 served as training sessions. During each of these ten sessions, subjects wore the HPDs throughout the session and completed three consecutive blocks of 60 trials (180 total, 15 trials at each of 12 locations). Training sessions lasted approximately 45 minutes, with no break provided between trial blocks. Continuous training sessions were employed to avoid gaps that can disrupt perceptual learning (Little et al., 2017). Sessions 3-12 occurred over the span of two weeks (10 consecutive days, excluding weekends). Finally, Session 13 and 14 (posttesting) repeated the 'open ear' and 'HPD' tests of session 1 and 2, respectively (open ear tested last). These sessions included all 24 target locations evaluated in pretesting sessions, enabling assessment of spatial generalization of training effects to untrained source locations. Notably, subjects only wore the HPDs in the lab environment and were not permitted to wear them outside of the lab.

3.2.5.2 Feedback procedure

During training sessions, subjects were provided audiovisual feedback as follows. A 'correct' localization response was operationally defined as a response within |7.5°| of the target and triggered a green LED to flash at the target location. Responses greater than |7.5º| from the target were labeled ‘incorrect’ and triggered a series of events. First, immediately after the subject’s incorrect response was registered, the stimulus was repeated at the target location eight times, paired with eight flashes of a red LED at the target location. During this time, the subject was instructed to find and orient to the target location, then maintain fixation on the target speaker until playback ceased. Next, during a two-second silent period, the subject was instructed to reorient to center (with the aid of a green LED located at 0º). Finally, the target stimulus and red LED flash were repeated eight additional times at the target location. This additional stage of feedback was intended to re-expose subjects to the spectral information available at the time of the initial error with explicit knowledge of the correct location. The stimulus repetition rate used during both ‘sets’ of feedback was 2 Hz (100 ms stimulus, 400 ms inter-stimulus interval), for a total feedback sequence duration of ~10 s (including the two-second pause between ‘sets’) during which subjects were able to make head movements while experiencing paired auditory and visual stimulation (cf. Carlile et al., 2014; Chapter 2).

3.2.6 Behavioral design and analysis

This study employed a longitudinal design, with repeated measures on nine human subjects across 14 testing sessions. Data, per trial, consisted of a head position track from the starting location (~0º azimuth and elevation) to the location at the time of button press, taken as the localization response. Examination of head position tracks enabled identification of occasional trials (57 out of 29,160 total trials) on which the subject did not register a button press during initial orientation (e.g., did not fully depress the response button) and thus reoriented to center without recording a valid response. These trials were discarded from further analysis.

Overall performance (pre- and posttest) was quantified via the spherical correlation coefficient (SCC), describing the similarity between the patterns of target and response locations (Fisher et al., 1993; Carlile et al., 1997; Watson et al., 2017; Poirier-Quinot et al., 2022). Azimuth- and elevation-specific performance was described via two different metrics. In azimuth, 'front-back confusions' (FBCs) were defined as errors $>|30^{\circ}|$ that fell in the opposite hemifield as the target (cf. Carlile & Blackman, 2014). FBC rates were calculated as the number of FBCs divided by the total number of trials within a given session. Target locations at $\pm 90^{\circ}$ were excluded when calculating FBC rate for pretest and posttest sessions as responses in the opposite hemifield are impossible at these locations. In elevation, multiple lines of evidence demonstrate that localization responses are compressed toward the horizon (0° elevation) following acute spectral shape cue disruption (Hofman et al., 1998; Zahorik et al., 2006; Zimpfer & Sarafian, 2014; Butler & Humanski, 1992; Bolia et al., 2001); such compression can be usefully quantified by the gain (slope of a linear fit) of response versus target elevation (e.g., Hofman et al., 1998). It should be noted that SCC and elevation gain approach 1 as performance nears 'perfect,' but this result can theoretically also be obtained for responses that are systematically shifted by an equal amount across all target locations.

In quantifying performance across subjects, pretest data were found to be non-normally distributed (Kolmogorov-Smirnov test on both open ear and earplug FBC rate $p<0.05$), thus all following analyses were conducted using nonparametric within- and between-subjects comparisons; medians and interquartile ranges are given in place of means and standard deviations throughout. Wilcoxon signed-rank tests including standardized test statistics (z-values) and p-values are reported with descriptive values at appropriate points in the Results.

The Holm-Bonferroni method was employed to manage family-wise error rate and address the challenge of multiple comparisons (Holm, 1979). All analyses were conducted using MATLAB.

3.3 Results

3.3.1 Earplugs change the directional dependence of spectral shape

The acoustic impact of earplug use was measured via manikin-based HRTF recordings made both with and without the HPD in place. The vertical transfer function, in both conditions, obtained for the right ear of the manikin for all trained locations is shown in Figure 2. Measurements are displayed with magnitudes relative to that obtained, for each condition, at 1 kHz for the center speaker (0°, 0°) location. In general, open ear data are similar to previous reports with a peak from 2-4 kHz, reflecting a non-directional ear canal resonance, and a notch from 6-9 kHz that increases and decreases in frequency with changes in source elevation, described as the 'elevation notch' (Rice et al., 1992; Brungart & Rabinowitz, 1999; Van Wanrooij & Van Opstal, 2005). In contrast, the HPD data are notably distorted across the frequency range. These data demonstrate a peak from 4-5 kHz that is independent of source elevation, and a notch from 6-9 kHz that is weakly dependent on elevation. The differences between open ear and HPD HRTFs represent the disparity between the expected open ear spectral information and the HPD-filtered spectral information. The data demonstrate that, while some spatially dynamic features may be conveyed, the HPD reshaped the transmitted signal, with notable attenuation from 5-10 kHz – the region thought to provide the primary cues for open-ear performance.

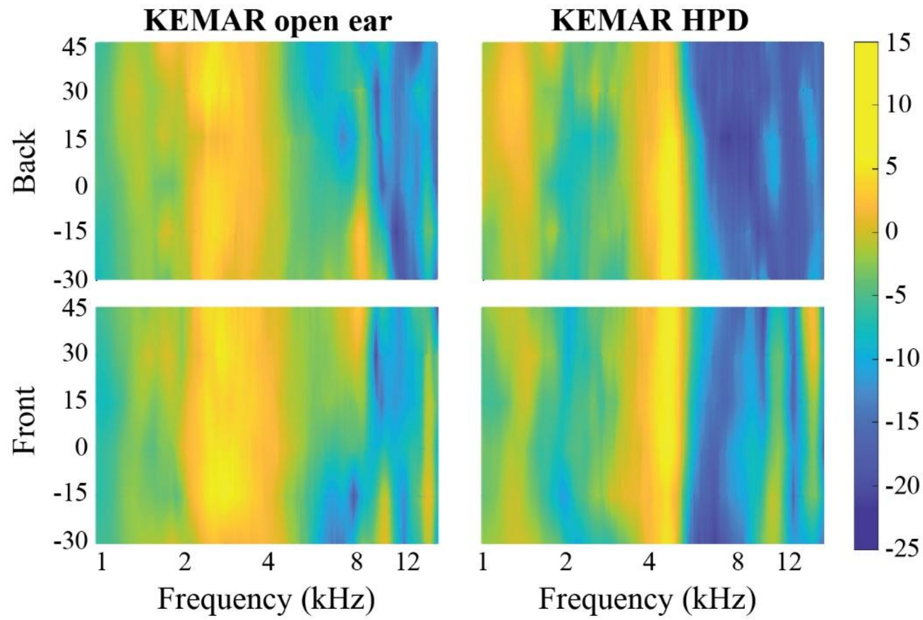

Figure 2. Manikin-based HRTFs demonstrating the transition of spectral features across vertical positions along the midsagittal plane. Specific locations recorded include -30° to 45° elevation at both 0° and 180° azimuth at 15° intervals. Data shown are from the recorded right ear of a KEMAR manikin both without (left) and with (right) the HPD in place. All magnitudes (color) are plotted relative to the magnitude recorded from the 'center' speaker at 1 kHz.

3.3.2 Acutely degraded sound localization improves in azimuth and elevation with daily audiovisual training

Source localization ability was assessed in 'pretest' sessions across 24 target locations with open ears (Session 1) and with the HPD inserted (Session 2). Changes in localization performance were quantified using three error metrics: SCC, FBC rate, and elevation gain (Figure 3A, B, and C, respectively). With open ears, subjects demonstrated a median SCC of 0.80, decreasing to 0.26 upon insertion of the HPD (Wilcoxon signed-rank test, $z=2.67$, $p=0.0039$). Median FBC rate more than doubled from 8.99% to 20.28% across all subjects ($z=-2.38$, $p=0.016$) and median elevation gain decreased from 0.59 to 0.17 ($z=2.67$, $p=0.0039$). While significant degradation of sound localization during earplug use is anticipated at the population level based on a volume of prior literature (see Introduction), this result was inevitable in the present sample due to the enrollment procedure used; as described in Methods and Materials, subjects were invited for participation in the present study based on participation in a previous study only if they had demonstrated a large decrement in performance with the tested earplug (>20% increase in errors |>30º| re: open ear). This substantial decrement in performance for each subject could be distributed across one or more error dimensions, defining the bases for and magnitude of potential improvements during localization training.

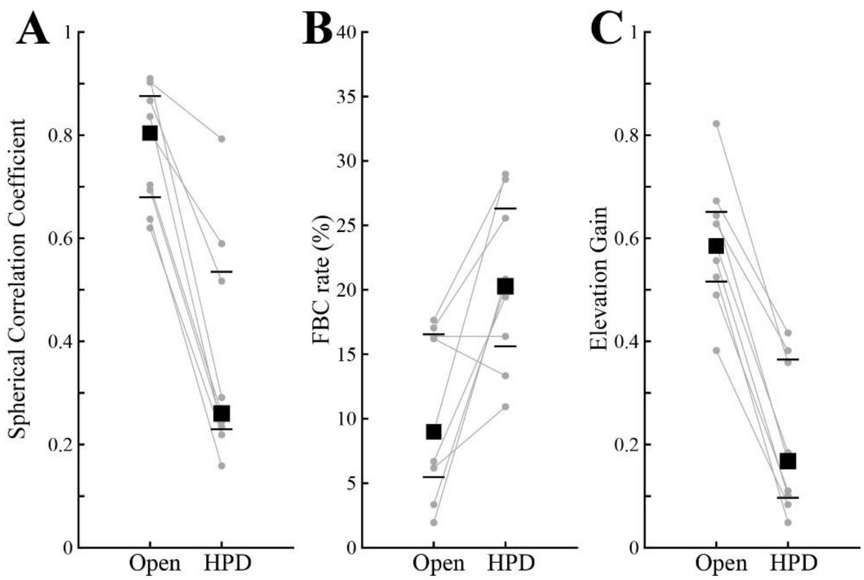

Figure 3. Sound localization metrics for two-dimensional localization performance with open ears and immediately after insertion of earplugs in enrolled subjects of the present study (see Materials and Methods). (A) The spherical correlation coefficient (SCC; e.g., Carlile et al., 1997) reflects the normalized dispersion of localization response locations re: target locations and can range from -1 (all responses reflected) to 1 (responses identical to target or shifted by a constant value); front-back confusions were removed prior to calculation of SCC. (B) The front-back confusion (FBC) rate is calculated across locations other than ±90º azimuth locations and indicates the proportion of trials on which subjects responded in the opposite hemifield from the target (see text). (C) Elevation gain, indicating the slope of a linear fit between target and response elevations, can also range from -1 to 1. In (A-C), thin gray points and thin traces reflect individual subjects (n=9), bold dashes indicate medians, and thin dashes indicate interquartile ranges.

During training, subjects returned to the lab once a day for 10 days (excluding weekends) to complete ~45-minutes of localization training per day across 12 target locations with the HPD inserted. Subjects were provided feedback during training (see Feedback procedure). Subjects had no exposure to sound through the tested HPDs outside of the laboratory environment.

Because all training locations were on the mid-sagittal plane (see Fig. 1), performance over the course of training was quantified in terms of FBC rate and elevation gain only, and improvement can be understood in terms of changes in these measures (SCC does not provide meaningful additional insight since horizontal location was not varied beyond two diametrically opposed locations).

Figure 4 shows performance on the final day of training (day 10) versus the first day (day 1) in terms of FBC rate and elevation gain. On these axes, improvement is represented as points that fall opposite the gray shaded regions. Shaded symbols represent individual subjects, and bold squares represent the median ± interquartile range.

For the subset of locations included in the training paradigm, intersubject variability in day 1 FBC rates was large (range approximately 5% to 45%). However, by day 10, FBC rates approached or reached zero (minimum possible value) for all subjects (all final FBC rates <10%). The lack of final FBC rate dependence on initial FBC rate is demonstrated by a line of best fit (thin solid line) yielding a slope near zero (slope=-0.046, R^2=0.047, p=0.58).

While variance in day 1 elevation gain was also large (range approximately 0.1-0.8), in contrast to FBC rate, variance remained large at day 10 (range 0.3-0.9). A line of best fit to elevation gain data yielded a slope near 1 (slope=1.06, R^2=0.65, p=0.0087), demonstrating a dependence of final elevation gain on initial elevation gain, and that the improvement in elevation gain was similar across subjects.

To further parse differences between azimuth (front-back) and elevation adaptation, data were plotted across all training days separately for forward (Fig 4B, top row) and rearward (bottom row) target locations. Some published evidence suggests that performance for rear hemifield locations (outside the visual field), while initially worse, improves similarly to that for front hemifield locations (Carlile & Blackman, 2014). In terms of FBCs, this difference was apparent on day 1 and 2 but quickly faded as both types of hemifield confusions were reduced to median values of 5% or less after day 4. This result is consistent with previous evidence for rapid learning in the front-back dimension (cf. Zahorik et al., 2006; Chapter 2; Golob et al., 2024; see 'Discussion'). Elevation gain also demonstrated a front-rear difference, with generally reduced gain in the rear hemifield. However, unlike FBCs, this difference persisted across the duration of training, with median elevation gain remaining lower in the rear hemifield through day 10. Further, elevation gain did not reach its theoretical ceiling value of one nor an apparent plateau (cf. Zahorik et al., 2006); improvement was approximately linear, suggesting that further adaptation could have occurred given additional training.

Improvement from day 1 to day 10 was additionally assessed at all trained speaker locations (Fig. 4C). Improvement in median FBC rate, collapsed across all subjects, was similar across all locations, with marginally worse day 10 performance demonstrated for the 30° and 45° speakers in the rear hemifield. Because elevation gain is a cross-location metric that cannot be calculated for individual speaker locations, location-specific elevation performance for purposes of this sub-analysis was assessed as the median absolute elevation error magnitude, collapsed across all subjects. In the frontal hemifield, while error was reduced by training at all locations, error was lowest for the center loudspeaker (0°, 0°) throughout testing. In the rear hemifield the two most extreme elevations (-30° and 45°) initially yielded the largest errors but also netted the largest improvement; all rear hemifield loudspeakers yielded on the order of 10° elevation error by day 10.

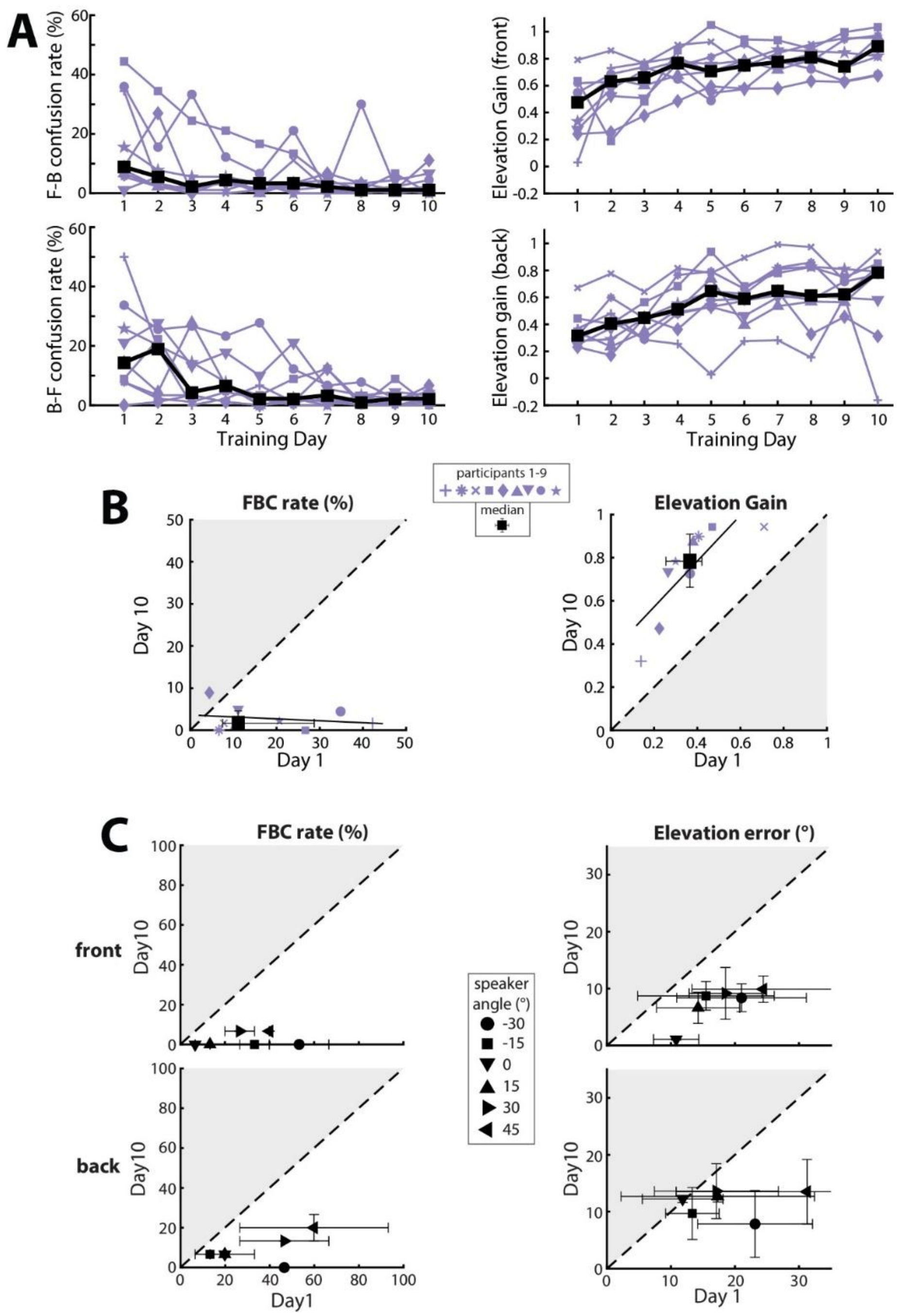

Figure 4. Changes in sound localization metrics over the course of daily training. (A) Performance is plotted across all training days in terms of FBC rate and elevation gain across forward and rearward targets (left/right, top/bottom, respectively). Individual subjects (thin purple lines) as well as the group median (thick black line) are shown. (B) FBC rate (left) and elevation gain (right) for training day 10 versus day 1. Points falling outside the shaded region indicate improvement from day 1 to day 10. A linear fit to individual subject data (thin line) evaluates the relationship between initial and final performance. (C) Median FBC rate (left) and absolute elevation error magnitude (right) at each speaker location for day 1 versus day 10. Points falling outside the shaded region indicate improvement.

3.3.3 Posttest performance and generalization

Approximately one week after the final day of training, subjects completed post-testing sessions with stimuli presented from the original 24 source locations; no feedback was provided. Posttesting sessions were completed both with (Session 13) and without (Session 14) the trained HPD inserted. The primary purpose of posttest measurements was to evaluate post-training HPD performance across a more ecological distribution of target locations and, by extension, to evaluate transfer of learning to untrained source locations, i.e. spatial generalization.

Figure 5A plots HPD posttest against pretest SCC; again, the SCC value is a measure of response dispersion across all 24 locations. SCC values significantly improved after training (earplug median pretest SCC = 0.26; posttest = 0.49; $z=-2.67$, $p=0.0077$), demonstrating increased coherence between target and response locations following training. A line of best fit yielded a slope less than 1 (slope=0.70, $R^2=0.63$, $p=0.0105$), indicating that SCC was generally increased more for subjects with worse pretest SCC. Performance changes for trained and untrained locations could not be usefully compared in terms of SCC, since trained locations were all located on the mid-sagittal plane.

Figure 5B plots HPD pretest versus posttest FBC rates for trained (black markers) and untrained (gray markers) source locations. While a majority of points fall below y=x, indicating improvement at posttest, untrained locations (which included 12 sources at ±45º and ±135º azimuth) elicited lower initial FBC rates and also smaller improvement than trained locations (which included 6 sources at 0º and 180º only). Nonetheless, FBCs were reduced at both trained (pretest median=36.67%, posttest median=7.78%, $z=2.67$, $p=0.0076$) and untrained (pretest median=13.33%, posttest median=7.81%, $z=2.24$, $p=0.025$) target locations (cf. Chapter 2). Notably, while some subjects still recorded near-zero FBC rates for trained locations at posttest (compare with Fig. 4A), FBC rates increased somewhat relative to training day 10 for other subjects. Trained locations in the posttest array included only 6 of the 12 mid-sagittal locations comprising the training array (see Fig. 1B), complicating direct comparison. For trained locations, the extent of FBC change varied across subjects; a linear fit (thin black solid line) was not significant (slope=0.24, R^2=0.10, p=0.42). However, a linear fit for untrained locations was significant (thin dotted gray line, slope=0.81, R^2=0.59, p=0.016) suggesting that front-back discrimination may be more readily learned at locations off of the midsagittal plane.

Figure 5C plots HPD pretest versus posttest elevation gain in the same format as Fig. 5B. In this case, points above y=x indicate improvement from pretest to posttest. All points fall above y=x; elevation gain was significantly increased for both trained (pretest median=0.070, posttest median=0.58, z=-2.67, p=0.0077) and untrained (pretest median=0.19, posttest median=0.37, z=-2.67, p=0.0077) locations. Like FBC rate, the extent of elevation gain change also varied across subjects; linear fits demonstrated slopes that were positive for both trained locations (thin black solid line, slope=1.58, R^2=0.51, p=0.031, not significant after alpha correction) and untrained locations (thin grey dotted line, slope=1.09, R^2=0.72, p=0.0036). These slopes suggest greater relative improvement for subjects with better pretest performance (although only significant for untrained locations).

Notably, significant changes in FBC rate and elevation gain reported here index performance gains at azimuths that were never presented in the course of training, suggesting that generally informative cues for front-back discrimination and elevation perception may be accessed based on presentation of stimuli in the mid-sagittal plane only. At the same time, performance gains at untrained locations were less than those at trained locations for both FBC rate (change in FBCs for trained locations=28.89%, change for untrained=5.53%) and elevation gain (change in elevation gain for trained locations=0.51, change for untrained=0.18), suggesting that a spatially distributed training array could have netted greater improvement across space. The implications of this result for spatial cue learning in applied contexts are treated further in the Discussion.

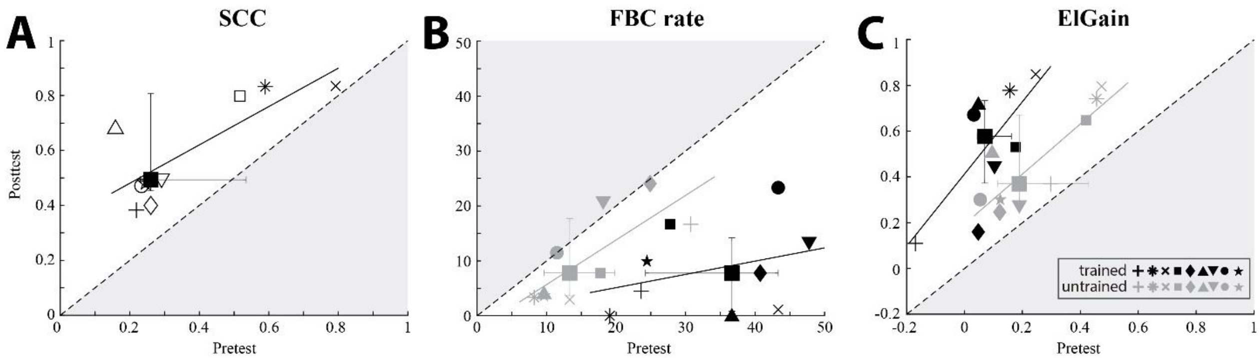

Figure 5. Posttest versus pretest HPD localization performance across 24 target locations including trained and untrained sources. (A) SCC (as in Fig. 3A) summarizes the dispersion of responses relative to targets at posttest versus pretest. Trained and untrained locations are not parsed (see text). A linear fit to individual subject data (thin line) evaluates the relationship between pretest and posttest performance. (B) FBC rate for trained locations (black symbols) and untrained locations (gray symbols) at pretest versus posttest. symbols that fall outside the shaded regions (below y=x) indicate improvement. Linear fits to individual subject data evaluate the relationship between pretest and posttest performance at trained (thin black solid line) versus untrained (thin gray dotted line) locations. (C) Elevation gain for trained locations and untrained locations. symbols that fall outside the shaded regions (above y=x) indicate improvement. Legend otherwise as in (B).

Following measurement of posttest performance with the HPD, subjects completed a final session with the open ear. When assessing all 24 locations tested at pretest and posttest, posttest open ear performance slightly improved relative to pretest. While SCC did not significantly change (pre=0.80; post=0.81; z=-1.24; p=0.21), both FBC rate (pre=8.99; post=4.78; z=2.43; p=0.015) and elevation gain (pre=0.59; post=0.72; z=2.67; p=0.0077) improved relative to pretest performance. This result unexpected on the basis of past reports (e.g., Hofman et al., 1998; Carlile et al., 2014; Chapter 2) and is addressed further in the Discussion. FBC and elevation gain performance was also assessed for trained and untrained locations individually. Figure 6 plots open ear pretest vs posttest for the three major metrics, as in Figure 5, separating performance for the trained and untrained source locations where applicable. Highlighted in this figure is the apparent difference in FBC rate improvement for trained and untrained locations. While results appear quantitatively similar to that of the HPD condition, differences in the case of open ear performance were not statistically significant.

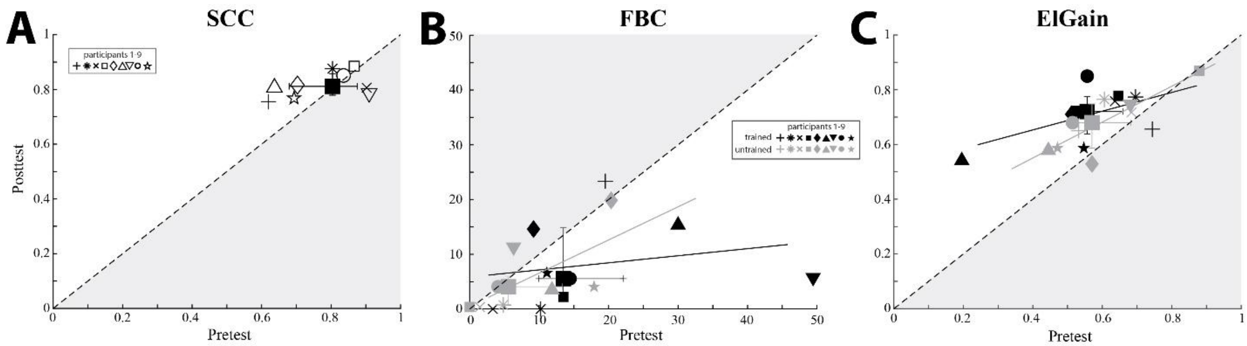

Figure 6. Posttest versus pretest localization performance for the open ear condition. Otherwise, layout and legend as in Figure 5.

3.3.4 Assessing impacts of target audibility on localization errors

Criteria for inclusion in the present study included audiometric thresholds within "normal hearing" limits at frequencies 0.25-8 kHz; subjects also completed extended high frequency threshold testing from 9-16 kHz. Although informative spatial cues, particularly elevation cues, have been suggested to occur in this spectral region (e.g., see Baumgartner et al., 2014), extended high frequency audiometric thresholds are often not reported for subjects completing localization tasks. While the present study used a low-attenuation earplug that preserved the audibility of target stimuli (i.e., subjects had no trouble hearing the broadband noise bursts presented on each trial), we reasoned that reduced audibility at extended high frequencies and increased attenuation by the earplug at extended high frequencies might combine to limit access to informative cues.

To illustrate this point, Figure 7A plots the range of subjects' hearing thresholds in dB SPL for the right ear (red shading gives range across 9 subjects; dashed line gives median) and left ear (blue). Amplitude spectra of in-ear signals for a source directly in front (0º azimuth, 0º elevation) in open ear (dashed line) and earplug (solid line) conditions are plotted on the same axes. The range of observed thresholds increases at higher frequencies, while the in-ear level of target stimuli, particularly in the earplug condition, decreases. The earplug-affected "sensation level" (SL) of spectral features falls below 30 dB SL above 8 kHz and reduces to 0 dB SL starting around 12 kHz, suggesting inaudible spectral regions were present for some subjects. While audibility for individual subjects would ultimately depend on the unique interactions of ear shape (e.g. Wenzel et al., 1993) and the insertion loss provided by the earplug as-worn, the plotted manikin-based values provide a useful approximation (see Discussion).

To assess the potential relationship between stimulus audibility and acute impacts on localization performance, the change in subject localization performance from open ear to earplug pretest was plotted for FBC rate (Figure 7B) and elevation gain (Figure 7C) as a function of the average extended high frequency audiometric threshold (in dB SPL, averaged across right and left ears per subject). No relationship was evident in the case of FBCs, consistent with indications (e.g., Blauert, 1969; Zahorik et al., 2006) that salient front-back cues may arise at 'mid' frequencies (i.e., frequencies below the extended high frequency range). In the case of elevation gain, a weak trend is observed (slope=0.008, R^2=0.5, p=0.033), suggesting that worse extended high-frequency thresholds lead to larger decrements in elevation gain, and in turn, that the disruption of vertical localization ability observed after earplug insertion among subjects of the present study may be related to the attenuation provided by the earplug in addition to other factors, ostensibly including disruption of audible spectral shape cues.

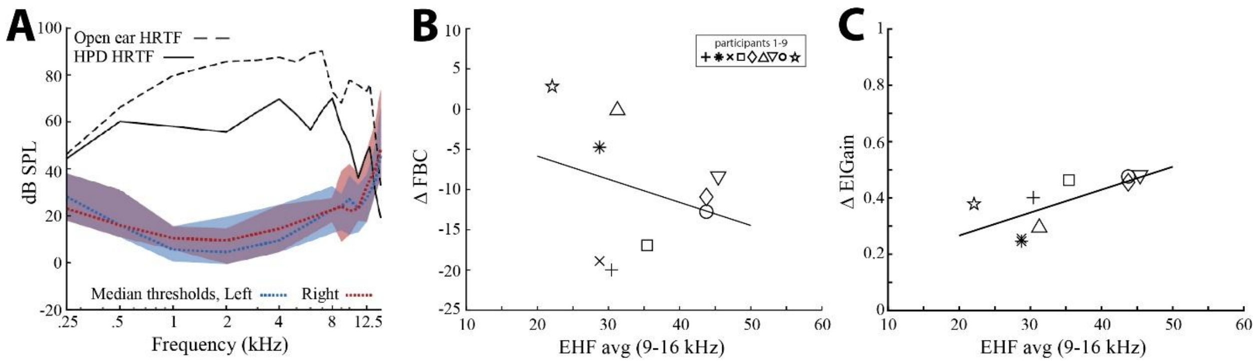

Figure 7. Evaluation of the relationship between high-frequency stimulus audibility and localization errors. (A) Subject thresholds from 0.25-16 kHz for right (red) and left (blue) ears are plotted as a function of frequency. Shaded regions give the entire range of observed thresholds across 9 subjects; dotted lines give the median threshold. Amplitude spectra of in-ear recordings of noise burst stimuli presented from the 'center' loudspeaker (0° azimuth, 0° elevation) are plotted on the same axes. Recordings were obtained using a binaural manikin in open ear (dashed line) and earplug (solid line) conditions as tested behaviorally. Spectral segments intersecting or falling below subject thresholds are presumed to be inaudible by one or more subjects. (B) Changes in FBC rate between open ear and earplug pretest are plotted against average extended high frequency audiometric thresholds (dB SPL, averaged across right and left ears) for each subject. Negative values indicate worse performance in the earplug condition. A line of best fit is also shown. (C) As in (B), but for elevation gain, such that positive values indicate worse performance in the earplug condition.

3.4 Discussion

The present study evaluated changes in sound localization during the use of bilaterally worn earplugs over the course of a brief (10-day) training paradigm. Earplugs are known to acutely disrupt spatial hearing performance, a key constraint on their usability in high-risk settings (e.g., Noble & Russell, 1972; Simpson et al., 2005; Talcott et al., 2012). While sufficient attenuation can of course limit access to stimulus features including spatial cues, the earplugs employed in the present study were low-attenuation passive devices (Combat Arms Gen. 4.1, used in 'open mode') that, similar to active electronic earplugs (e.g., Casali & Robinette, 2015) and hearing aids (Van den Bogaert et al., 2006), transmit highly audible signals but nonetheless significantly degrade spatial hearing performance, ostensibly due to disruption of monaural/pinna cues in particular (Russell, 1976; Zimpfer & Sarafian, 2014; Brown et al., 2015; Chapter 2). In this sense, low-attenuation passive earplugs may be viewed as an applied analog of the silicone earmolds used to overtly disrupt pinna-derived spatial cues in multiple studies spanning the past 25 years (e.g., Hofman et al., 1998; Van Wanrooij & Van Opstal, 2005; Carlile et al., 2014; Watson et al., 2017). In the following sections we consider insights gained from the present measurements primarily in the context of this substantial literature. We also consider, at points, how insights on spatial cue learning may also be applied to impact outcomes in users of earplugs and other real-world hearing devices.

3.4.1 Acutely induced increases in localization errors and training-induced improvement across error type

Previous studies of sound localization adaptation have employed stimuli varying in both azimuth and elevation within the frontal hemifield/visual field only (e.g., Hofman et al., 1998; Van Wanrooij & Van Opstal, 2005), in azimuth only but spanning front and rear hemifields (e.g., Casali & Robinette, 2015; Chapter 2), or in both azimuth and elevation spanning both front and rear hemifields (e.g., Zahorik et al., 2006; Carlile & Blackman, 2014; Watson et al., 2017). The present study employed stimuli varied in both azimuth and elevation across front and rear hemifields, with two different arrays of target locations utilized at different points in testing (Figure 1B). Whether induced via experimental earmolds (e.g., Hofman et al., 1998; Carlile et al., 2014), in/on-ear devices (e.g., Casali & Robinette, 2015; Chapter 2; present study), or insufficiently individualized virtual stimuli (Zahorik et al., 2006; cf. Wenzel et al., 1993), all studies have demonstrated that spectral cue disruption leads to acute increases in sound localization errors (an inevitable result of the present study due to the enrollment procedure used). The kinds of errors that can be or have been then quantified and tracked as a function of training/exposure are dependent on the spatial configuration of tested locations.

When sources are present in both front and rear hemifields, subjects initially demonstrate acute increases in front-back confusions. In previous studies (e.g., Zahorik et al., 2006; Carlile & Blackman, 2014) and the present, the incidence of front-back confusions (rear hemifield responses for front hemifield targets) is initially lower than the incidence of back-front confusions (front hemifield responses for rear hemifield targets). The reason for this disparity is not fully clear (nor is the mechanistic basis of front-back discrimination; see following section), but it is consistent with the general observation that localization ability is poorer in the rear hemifield than in the frontal hemifield (e.g., Makous & Middlebrooks, 1990; Carlile et al., 1997). However, despite this initial disparity, FBC rate rapidly declines over the course of exposure to altered cues (e.g., Zahorik et al., 2006; Golob et al., 2024; present study) such that, with training, FBC rates approach open-ear levels for sources in both front and rear hemifields (see Fig. 4C).

When sources are present at multiple elevations, subjects initially demonstrate acute degradation of elevation perception, which may be quantified in terms of error magnitude (increased re: open ear, expressing the increased divergence of response versus target elevations) or response gain (decreased re: open ear, expressing the reduced dependence of response elevation on target elevation). In the present study, we used elevation gain (e.g. Hofman et al., 1998) to summarize the variation of response elevation with target elevation for each subject. Like FBCs, values of elevation gain were also initially lower in the rear hemifield than in the front hemifield, again consistent with generally poorer localization performance for rear hemifield targets (Makous & Middlebrooks, 1990; Carlile et al., 1997). However, this difference generally persisted across the period of training, and subjects did not reach a performance asymptote in either case (see following section).

It follows from the foregoing that training studies limiting the occurrence (or measurement) of FBCs or errors in elevation via exclusion of stimulus variation along those dimensions may overestimate functional training-induced gains in localization performance. Nonetheless, it is notable that FBCs – the largest auditory spatial orientation errors possible, and a striking form of perceptual error in general – decrease similarly whether source elevation is varied (e.g., Zahorik et al., 2006; Carlile & Blackman, 2014; present study) or is held constant at the horizon (Chapter 2; Golob et al., 2024). In service of eliciting significant gains in localization in applied settings, e.g., reducing or eliminating front-back confusions during the use of earplugs, training on a sparse set of speakers near the horizon may be sufficient to net measurable benefit.

3.4.2 Impacts of training time course and interactions with error type

Adaptation to altered spectral cues has been evaluated across a variety of exposure time courses, from two 30-minute sessions (Zahorik et al., 2006) to one week of daily sessions (Golob et al., 2024) to two weeks of focused training sessions additional to continuous daily wear (Carlile et al., 2014) to several weeks of extended daily wear with intermittent laboratory visits (Hofman et al., 1998; Carlile & Blackman, 2014) to two months of once-weekly training sessions (Chapter 2). In light of such wide temporal variation, the reproducibility of the fundamental result – that localization performance significantly improves – is quite remarkable. Nonetheless, as suggested in the previous section, improvements in front-back discrimination and elevation perception appear subject to different time courses. In terms of putative FBCs, we have observed similar improvement across a series of audiovisual training sessions spread across two months (Chapter 2) or two weeks (present study), with performance approaching an apparent asymptote for most subjects after a few sessions. Consistent with this finding, prior studies have reported significant reductions in FBCs over the span of one week (Golob et al., 2024) and even two sessions (Zahorik et al., 2006).

In contrast, improvement in elevation perception has generally been reported across a scale of weeks, with graded improvement continuing for several weeks in some subjects (Hofman et al., 1998; Carlile & Blackman, 2014). Significant improvement was evident across a period of 10 days in the present study (cf. Carlile et al., 2014), but subjects did not reach a performance asymptote, suggesting that improvement could have continued given additional training sessions. In the same study that reported significant FBC reduction with two 30-minute training sessions (Zahorik et al., 2006), no improvement in elevation performance was observed. The authors speculated that differences in the form and/or extent of training-induced changes in front-back discrimination versus elevation perception might suggest separate mechanisms subserving the two capacities, even though both apparently depend on 'spectral' cues. Data from the present study are consistent with this speculation. Elevation may be encoded in high-frequency spectral notches and/or gradients that vary systematically with source elevation, initially by brainstem neurons sensitive to those features (see Baumgartner et al., 2014; Hebrank & Wright, 1974). Potential cue(s) and mechanism(s) subserving front-back discrimination are less clear, but relatively higher gain at mid-frequencies (e.g., Zahorik et al., 2006) and/or multimodal tuning for audiovisual convergence (e.g., Montagne & Zhou, 2018) for sources in the front hemifield may contribute. Deeper understanding of the cues and mechanisms underlying perception along these two dimensions is necessary to understand the bases and time courses of training-induced modification.

3.4.3 Generalization of training to untrained source locations

Insofar as the goal of sound localization training is to improve real-world localization in which target sounds could come from any location, spatial generalization of training to "untrained' source locations is an essential outcome. Additionally or alternatively, the success or failure of generalization may be evaluated with respect to acoustic information conveyed to yield fundamental insight on the salient cues/basis of performance. Past studies (Mendonça et al., 2013; Watson et al., 2017; Chapter 2) have generally demonstrated significant generalization of sound localization training to untrained source locations (but see Zonooz & Van Opstal, 2019). Generalization is likely successful because of the continuous nature of spectral shape change with continuous change in source location (Middlebrooks, 1997; see Chapter 2), enabling the auditory system to effectively interpolate untrained locations from adjacent trained locations. Consistent with this account, a previous study by our group (Chapter 2) demonstrated robust generalization of localization training across sources separated by 15º in azimuth for subjects that received explicit audiovisual feedback during training, and even those that did not receive explicit feedback.

Toward more efficient training paradigms, the question arises: What is the minimal set of trained locations that can produce useful generalization? To probe this question, the present study utilized a spatially sparse set of training locations designed to target the most problematic spatial perceptual impacts of spectral shape cue disruption (and earplug use), increased front-back and elevation errors. Although azimuthal accuracy, excluding FBCs, is slightly degraded following spectral shape cue disruption (e.g., Carlile & Blackman, 2014, Brown et al., 2015), the inclusion of training locations at multiple azimuths significantly increases training time (assuming a given number of training stimuli per location). If FBCs could be resolved at multiple azimuths via relearning of a generic cue (e.g., higher relative gain at mid-frequencies for forward sources; Blauert, 1969; Zahorik et al., 2006), significant gains in ecological performance could be realized with a minimal set of front and rear hemifield training locations. Similarly, if elevation perception could be augmented via re-learning of a generic elevation cue, elevation might be recovered at multiple azimuths via training at a small set of azimuths.

In practice, significant improvements in front-back discrimination (Fig. 5B) and elevation gain (Fig. 5C) were observed at untrained locations, suggesting that a spatially sparse training array was sufficient to induce a degree of improvement across space. However, variability across subjects was notable, and in the case of FBCs, three (of nine) subjects showed little or no improvement at untrained locations. Moreover, for both FBCs and elevation gain, improvement was less at untrained than at trained locations. Thus, results also suggest that untrained locations conveyed spatial cues that were sufficiently different from those at trained locations to limit generalization. A training array including locations distributed in both azimuth and elevation is likely required to facilitate robust generalization across two-dimensional space; the optimal distribution/density of such an array would be an interesting topic for future study.

3.4.4 Examination of individual variability and future directions

While most subjects are able to adapt to altered spatial cues to some extent, the form and degree of adaptation varies substantially across subjects (Hofman et al., 1998; Zahorik et al., 2006; Carlile et al., 2014; Chapter 2). One source of variability is the form and degree of the cue alterations themselves, which depend on the interactions of the cue-modifying device with the subject's unique ear shape (e.g., Van Wanrooij & Van Opstal, 2005). Another source of variability might be the audibility of salient cues. Changes in localization response gain and variance have been noted for sources below approximately 20 dB sensation level (SL) (Macpherson & Sabin, 2013; Vliegen & Van Opstal, 2004; Sabin et al., 2005); as stimuli approach threshold (0 dB SL), spectral regions become inaudible, and cues conveyed within those regions become inaccessible. In the case of elevation perception, in particular, salient cues may occur at frequencies above the conventional audiometric range (Best et al., 2005; Baumgartner et al., 2016). In the present study, extended high frequency thresholds were measured for all subjects and compared to earplug-induced decrements in localization performance; a marginal trend was observed in the case of elevation perception, with greater decrements in performance for subjects with poorer extended high frequency thresholds. Associations between localization errors and functional audibility in the context of earplug / in-ear device use would be interesting to explore further in a future study with a higher number of subjects and larger range of hearing thresholds. Such an analysis would be particularly informative if combined with individualized measures of spectral shape cues before and after device insertion (cf. Van Wanrooij & Van Opstal, 2005) to provide estimates of audible spectral shape cue distortions.

3.5 Acknowledgements

This work was funded by the University of Washington, the Department of Defense (Joint Warfighter Medical Research Program) [Contract W81XWH-20-C-0077], and the National Institute on Deafness and Other Communication Disorders (Grant No. T32-DC005361). The authors would like to thank Drs. Nathaniel Greene, Ted Argo, and Bev Wright for useful discussions.

Chapter 4: Spectral cues supporting front-back discrimination: individualized measurement and accommodation following overt spectral manipulation

4.1 Introduction

Directional hearing is important for communication, situational awareness, and safety. Humans frequently find themselves in situations that require reflexive sound source localization, such as when someone calls their name in a crowded room, or a car horn blares from an approaching vehicle (Byrne & Noble, 1998). Normal sound localization relies on three main cues: interaural level difference (ILD), interaural time difference (ITD), and spectral shape cues. ILDs and ITDs are binaural cues that arise from the separation of the ears on opposite sides of the head and aid in horizontal source localization (azimuth; Rayleigh, 1907). Spectral shape cues are monaural and arise from direction-dependent interactions of impinging sound waves with the complex shape of the ear, allowing for perception of source position along the vertical plane (e.g. up-down and front-back) (Batteau, 1967; Blauert, 1969; Hofman et al., 1998).

As summarized in the introduction to Chapter 3, bilaterally worn hearing protection devices (HPDs) introduce minor distortion of binaural cues while significantly impacting the use of spectral cues for the purposes of elevation and front-back perception (Reddy et al., 2012). Additionally, the degree of sound localization disruption is strongly correlated with the degree of spectral cue degradation caused by the HPDs (Brown et al., 2015; Audet et al., 2026). Performance is significantly worse with the use of earmuff-style devices compared to all other HPD styles, either when considering raw percent correct or accumulation of specific error types, including front-back confusions (Zimpfer & Sarafian, 2014). When considering the behavioral consequence of localization errors, front-back confusions (FBCs) stand out as major perceptual errors that can delay responses to critical signals and disorient users in hazardous environments (Montagne & Zhou, 2018).

However, HPDs are a valuable tool in the prevention of noise induced hearing loss (NIHL), accounting for 16% of all disabling hearing loss cases worldwide (Nelson et al., 2005). HPDs help prevent hearing loss by way of broadband attenuation that ranges from ~10 to 40 dB depending on frequency and HPD style (Park & Casali, 1991; Buck & Zimpfer-Jost, 2004). However, unintended side effects of HPD use, such as degraded front-back discrimination ability, introduce barriers that lead to poor device retention (Simpson et al., 2005; Wong et al., 2003; Yahya et al., 2016).

The prevention and possible correction of such localization errors via training is summarized in Chapter 3. Interestingly, while previous evidence suggests that information conveyed by sound spectra contribute to successful front-back discrimination, the specific spectral features related to such perception remain unclear. This contrasts with elevation perception, which is typically associated with a 'spectral notch' around 6-9 kHz that systematically increases in frequency as source elevation increases (e.g., Rice et al., 1992; Brungart & Rabinowitz, 1999; Van Wanrooij & Van Opstal, 2005). Initial efforts to describe the spectral feature(s) aiding discrimination of front versus back led to the notion of 'directional bands'; specific frequency regions where a narrowband noise can induce either a forward or rearward response (Blauert, 1969). Within the 'directional bands' framework, increased sound energy from 0.25-0.5 kHz, from 2-6 kHz, and at 16 kHz results in forward perception, while increased energy from 1-2 kHz and from 8-12 kHz results in rearward perception. This framework was further supported by Hebrank & Wright (1974b; broadband importance), Langendijk & Bronkhorst (2002; high frequency importance), and Zahorik et al. (2006; 3-7 kHz importance). However, including the original study of Blauert (1969), these studies did not employ trial-to-trial roving in their stimuli presentation, possibly resulting in a gross level cue that would not be available in real-world listening conditions with variable source levels. When roving is included in stimuli presentation, key spectral regions are more variable across individuals (Lladó et al., 2022) but are likely to fall within the broadband range of 2.5-10 kHz (Yost, 2023).

In total, reports of spectral cue importance for front-back perception suggest potential contributions from multiple frequency regions. Recent work by Lladó and colleagues (2025) investigated this variance, retrospectively assessing the ability of five different models of sagittal plane localization to predict behavioral responses of human listeners. Results from this work revealed that no single model effectively predicted front-back perception across all listeners. The three models with the highest predictive accuracy across multiple listeners included one based on the 6-9 kHz notch region (Zonooz et al., 2019), one that emphasized 6 kHz specifically (Ahrens & Brimijoin, 2019), and a novel one based on spectral variance. 'Spectral variance,' as defined by Lladó et al. (2025), refers to frequency regions with spectral magnitudes that are most variable as source location changes. While the variability in model selection across listeners could be informed by higher level central auditory processing (Lladó et al., 2022), peripheral factors, such as the significant variability in pinna shape across every individual (Chaudhuri et al., 2017), may also contribute to the variability of spectral importance across listeners (Lladó et al., 2025).

In the absence of a yet-to-be-defined spectral feature set that is common across listeners, the primary working hypothesis tested in this chapter is that the frequency information supporting front-back discrimination is strongly related to each individual's unique anatomy and resultant acoustic properties. Under this hypothesis, we expect the spectral regions supporting front-back discrimination to be variable across individuals, since detailed ear anatomy and resultant acoustic properties are also variable across individuals (summarized in preceding chapters). However, because listeners' utilization of spectral features is known to be malleable with training, we further expect that listeners can be trained to accommodate an introduced front-back discrimination cue. For purposes of this study, this cue will be generated by a newly designed custom earmuff, further building upon the applied theme of hearing protection as an ecologically important source of spatial cue disruption explored throughout this dissertation. We expect that after successful accommodation to the custom earmuff, the spectral regions supporting front-back discrimination will deviate from those of the open ear to demonstrate increased weighting of frequencies related to the specific spectral feature created by the earmuff.

Alternatively, the spectral regions supporting front-back perception with open ears could be relatively more reproducible across listeners. Additionally, even with modified auditory periphery, the key regions and/or the task of front-back discrimination could be mediated by more central auditory processes, limiting the impact of training on measured spectral weights. For example, while listeners may still improve in their front-back discrimination abilities because of training, the improvement could reflect improvement in the utilization of information within static spectral regions.

4.2 Materials and methods

4.2.1 Subjects

Sixteen adult human subjects aged 18-45 years old (5 male; 11 female) completed the study. Participation entailed between 2 and 12 laboratory visits across a period of one month. An additional five subjects enrolled in the study but did not complete the full test battery and/or did not comply with training requirements (see section 4.2.5); their data are not further considered. All subjects reported normal hearing and passed a hearing test composed of air conduction thresholds ≤25 dB HL at octave frequencies from 0.25-8 kHz, ≤10 dB interaural asymmetry at all frequencies. In addition, extended high frequencies (9-16 kHz) were measured with ≤25 dB HL thresholds needed for inclusion in the study. Informed consent and all testing procedures followed a protocol approved by the University of Washington Institutional Review Board. Subjects were compensated for their participation.

4.2.2 Apparatus

Sound localization was measured in a hemi-anechoic chamber (ETS-Lindgren HA-160; interior dimensions 9' x 10' x 7'). A circular loudspeaker array (diameter 2.5 m) was positioned inside the chamber with a height-adjustable stool fixed to the floor at the center of the array. A chin-rest was placed at head level such that listeners' heads remained stationary during all testing. Six loudspeakers (Gallo Acoustics MicroSE) were positioned in two clusters of three speakers. The two clusters were centered directly in front and behind the listener. Each cluster included three speakers: a central speaker, one speaker 5° to the left, and another 5° to the right in the azimuthal plane relative to the center of the array. By convention, the source location directly in front of the subject was designated 0º, source locations to the left were assigned negative azimuth values (e.g., -5º was slightly left), and source locations to the right were assigned positive azimuth values (e.g., +5º was slightly right). Following this scheme, the speaker directly behind the listener was designated 180°. The 0° and 180° loudspeakers were equipped with a tri-color LED mounted centrally that was controlled by an external USB card

(National Instruments USB-6501) to provide visual cues during the experiment as described below. Testing was completed in total darkness to limit the use of unintended visual cues.

4.2.3 Stimuli

'Base stimuli' were broadband (digitally filtered 0.2-21 kHz) noise tokens. The tokens were generated in an iterative process, using a custom MATLAB script, in which an initial random-noise token was filtered by a 10th-order Butterworth filter and divided by its own temporal envelope. This process of filtering and dividing by its own envelope was repeated 16 times. The goal of this process was to reduce temporal fluctuations in the noise tokens, allowing for greater homogeneity of digitally amplitude-scaled stimulus tokens. Base stimuli were used to generate different stimuli for each experimental procedure, as described below (Figure 1).

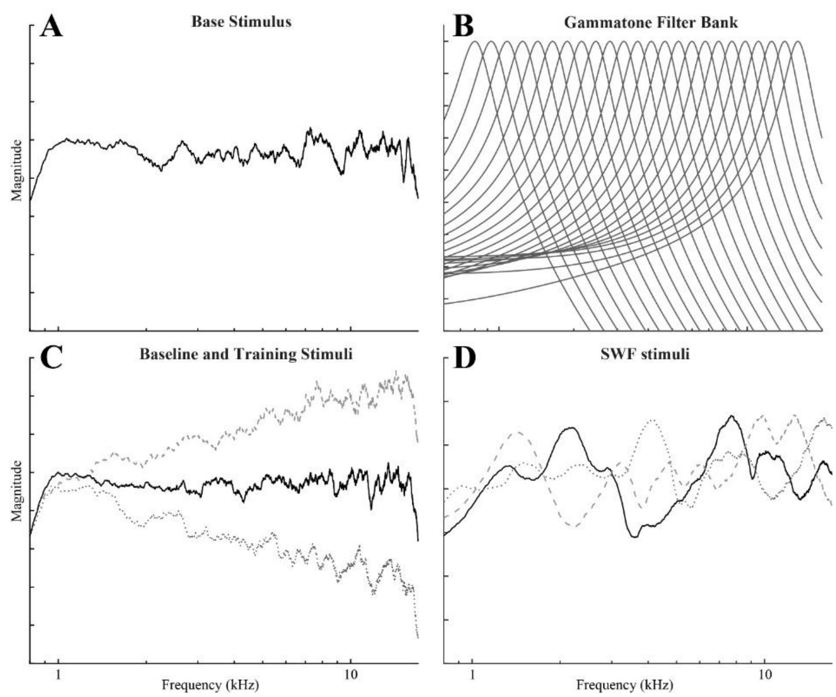

Figure 1. Example depictions of stimuli generation. (A) Base stimuli, generated as low-noise noise with a spectrally flat profile are passed through a gammatone filter bank (B). Energy in each filter band is manipulated to generate either (C) baseline and training stimuli with -3, 0, or 3 dB of spectral tilt from 0.8-16 kHz or (D) SWF stimuli. SWF stimuli have energy in each band pseudo-randomly adjusted by either -80, -60, -40, -20, or 0 dB before being recombined into a single stimulus (three examples shown).

For baseline and training procedures, base stimuli were modified to include a -3, +3, or 0 dB/octave spectral tilt. This was done to control for cross-trial comparisons of pitch/timbre within a 2AFC procedure (see Section 4.2.5). The spectral tilt was implemented by passing base stimulus tokens through a 24-channel gammatone filter bank (see Figure 1; Glasberg and Moore, 1990; Slaney, 1993) with frequency range 0.8-16 kHz, compensating for group delay, and extracting filter coefficients for each frequency band. Spectral tilt vectors were designed to be -3, +3, or 0 dB/octave at octave frequencies from 0.5-16 kHz. These vectors were then interpolated to the center frequencies of the gammatone filters, exponentiated, and applied to the contents of each filter band. Each final token was then constructed as a linear sum of each adjusted output. The number of tokens with each spectral tilt was balanced within each task. As such, for baseline testing, a set of 100 ‘baseline stimuli’ were made (33, 33, and 34 tokens with -3, +3, or 0 dB spectral tilt, respectively); for training, a smaller set of 45 ‘training stimuli’ were made (15 tokens of each spectral tilt).

Where indicated, ‘SWF stimuli’ consisted of base stimulus tokens with -80, -60, -40, -20, or 0 dB of magnitude adjustment pseudo-randomly assigned to each of 24 frequency bands (see Figure 1). The magnitude adjustments were applied in a similar procedure to that of the baseline stimuli described above, with the following differences. Toward pseudo-random application of magnitude adjustments for each noise token, a 24x10 matrix (24 frequency bands; 10 noise tokens) was generated. For a given frequency band (row), the five proposed magnitude adjustments were randomly arranged with predetermined counts (-80:3, -60:2, -40:2, -20:2, 0:1). For example, a single row might have read: [-80, -40, -40, -80, 0, -60, -80, -20, -20, -60]. Each column, therefore, represented a random assortment of magnitude adjustments that were applied to the 24-channel ‘base stimulus’ token. The matrix was created such that while each noise (column) imposed random magnitude adjustments, the average of the 10 noises reflected equal assignment of each manipulation across each band. Therefore, over the course of testing, the long-term average was a flat spectrum, broadband stimulus. 10 sets of these stimulus matrices were generated, resulting in 100 total SWF stimuli. In addition to the manipulated stimuli, 20 'baseline stimuli' were added to the set of 100, acting as trials to evaluate the stability of subject response accuracy and bias throughout testing.

All stimuli were 100 ms in duration with 12.5-ms on/off ramps, presented at 65 dBA with ±10 dB roving. Roving was applied to minimize the use of across-trial level cues. Stimuli were designed in MATLAB, synthesized at 44.1 kHz at full 24-bit resolution using a 64-channel USB D/A interface (Antelope Galaxy 64), amplified (Crown DriveCore CT875; Harman), and relayed to the target speaker. All procedures including stimulus presentation and data acquisition described to this point were controlled with custom software developed in MATLAB.

Of note, the degree of amplitude adjustment for SWF stimuli was motivated by previous studies assessing elevation perception (Zonooz et al., 2019) and front-back perception (Yost, 2023; Lladó et al., 2022) during spectral randomization. As Zonooz et al. (2019) suggested, HRTF manipulations must balance being too small to influence perception and so large/unnatural that subjects experience a loss of externalization or even perceive a second source. The adjustments used here were inspired by that of Lladó et al. (2022), who employed magnitude adjustments of -90, -40, or 0 dB randomly assigned to individual frequency bands such that -90 dB was five times more likely to occur than other adjustments, resulting in ~7-25% front-back confusion rates across all subjects.

4.2.4 Custom earmuff

The motivation behind designing a custom earmuff was two-fold. Firstly, in assessing the malleability of spectral cue processing for sound localization, it was desirable to provide all subjects with known and similarly distorted spectral information. Therefore, between-subject comparisons of SWFs after accommodation were not only more straightforward but could be more readily related to the known altered spectral information provided by the custom earmuff. Secondly, there is a large gap in current hearing protection design outcomes; although effective at reducing dangerous sound intensities to safe levels, HPDs reduce the user's situational awareness via degraded spectral cue information. Disruptions are notably worse for circumaural earmuffs, which are thought to eliminate pinna cues altogether (see Introduction). Developing a passive earmuff that allows access to informative spectral cues toward successful front-back discrimination would remove a significant barrier to HPD use and retention in situations for which earmuffs are preferred (e.g. ergonomically) or required (e.g. part of double hearing protection to protect against the highest-intensity noise environments). As an aside, in early efforts to measure spectral weighting of cues for front-back discrimination, we evaluated the use of virtualized sources presented over earphones, which allowed for full control of spectral features available at the level of the ear canal. Virtualized playback failed to produce compelling externalization despite several adjustments to playback methods, and this approach lacked the applied value of testing with a real head-worn device.

In developing the new earmuff, the intention was to design a device that introduced a novel and overt spectral feature that was highly directional such that it could reasonably be used to disambiguate forward and rearward sound sources. To create such a feature, and to evaluate whether observed perceptual gains were leveraging the created feature, it would need to be in a frequency region without a preexisting directional feature, perceptible, and hemifield-dependent. When designing the earmuff for this study, the intention was to avoid spectral regions previously implicated in front-back perception (i.e. 'directional bands' described in Section 4.1), introducing a feature in a novel frequency region. Specifically, our goal was to generate a mid-frequency spectral notch unique to rear hemifield sources.

In order for the mid-frequency spectral notch created by the custom earmuff to guide front-back perception, it must be perceptible. Existing spectral notch perception literature, although sparse, suggests that detection thresholds of ¼ octave bandwidth notches range from 3-6 dB when centered at 1 kHz compared to 10-20 dB when centered at 8 kHz (Moore et al., 1989). Furthermore, these thresholds are reduced (improve) as notch bandwidth increases. Therefore, we aimed to create an earmuff that would introduce a suprathreshold notch with a depth of at least 10 dB within the target frequency range.

The selected design was a cylindrical earmuff shell with symmetrical cavities facing the forward and rearward directions with a Helmholtz resonator placed in a single (rearward) cavity (see Figure 2). The two symmetric cavities opened into a common central chamber that was then routed into a single tube. This tube was then coupled to a Comply ® Foam ear tip which was inserted into the ear canal. This design was found to yield high spectral symmetry between front and back sound sources disrupted primarily and intentionally by the addition of a Helmholtz resonator in the rear cavity, creating a spectral notch at a designed frequency. The use of a tube and insert ear tip bypasses the pinna, with the intent to prevent individualized/pinna-specific interactions that may cause additional uncontrolled spectral differences.

All earmuff parts were designed in Autodesk Inventor Professional 2024 and fabricated via 3D printing (Lulzbot® Mini) using 2.85mm PolyLite™ polylactic acid (filament). The earmuff was printed in 3 individual parts, attached via snap-fittings embedded in the design. The assembled parts were then coupled to commercially available ear cushions and a recycled headband removed from a pair of commercially available earmuffs.

The resonant frequency of the Helmholtz resonators used in this study were calculated following this formula:

$$f = \frac{c}{2\pi} * \sqrt{\frac{A}{V * 0.85L}}$$

Where f is the tuned frequency of the resonator, c is the speed of sound in air (343m/s), A is the cross-sectional area of the resonator neck, V is the volume of the resonator cavity, and L is the length of the neck (Kinsler et al., 2000). Following this formula, resonators were iteratively 3D printed and tested (via manikin-based recordings) until the desired frequency was attained.

Figure 2. 3D rendering of the designed earmuff (top) with symmetrical cavities. A Helmholtz resonator tuned to 2 kHz can be seen in the rear cavity. Tubing is used to couple the earmuff to a Comply ® Foam ear tip (bottom right). The earmuff is attached to a headband and worn by a KEMAR manikin for acoustic testing (bottom left).

Toward demonstration of the impact of the earmuff on specific spectral regions, Figure 3 demonstrates acoustic measurement of the front-back spectral differences, as described above, for both the open ear and the newly-designed earmuff. All measurements were completed with a Knowles Electronic Manikin for Acoustic Research (KEMAR) manikin placed in the center of our speaker array, facing forward, with ear canals aligned to the loudspeaker array using a laser-leveling system. Notably, with the earmuffs in place, a 15 dB notch is present, centered around 2 kHz with a ~1 kHz bandwidth, attributed to the Helmholtz resonator tuning. However, other spectral differences are also present, such as a 10 dB difference at 7 kHz, -5 dB differences at 5.5 and 8-9 kHz, and a narrowband 12 dB difference at 10 kHz. While these features were expected to be less salient than the introduced notch and potentially imperceptible (Moore et al., 1989), subjects had access to full-bandwidth signals while wearing the earmuffs.

These data also permit the measurement of the broadband attenuation provided by the earmuff. Traditionally, the impact of an HPD on transmitted sound is indicated by the device's noise reduction rating (NRR), describing broadband attenuation measured under ideal conditions (Protection of Environment, 1979). In practice, labeled NRR is calculated through cross-frequency averaging of real ear attenuation at threshold (REAT) measures obtained from human listeners wearing the device. Here, an acoustic approximation of NRR was calculated by taking the average difference, from 1-8 kHz, between recordings of a 100 dB SPL pink-noise stimulus presented to an acoustic manikin both with and without the earmuffs in place. This approximation yielded a value of 15, reflecting mild attenuation relative to traditional earmuffs (NRRs generally between 25-30).

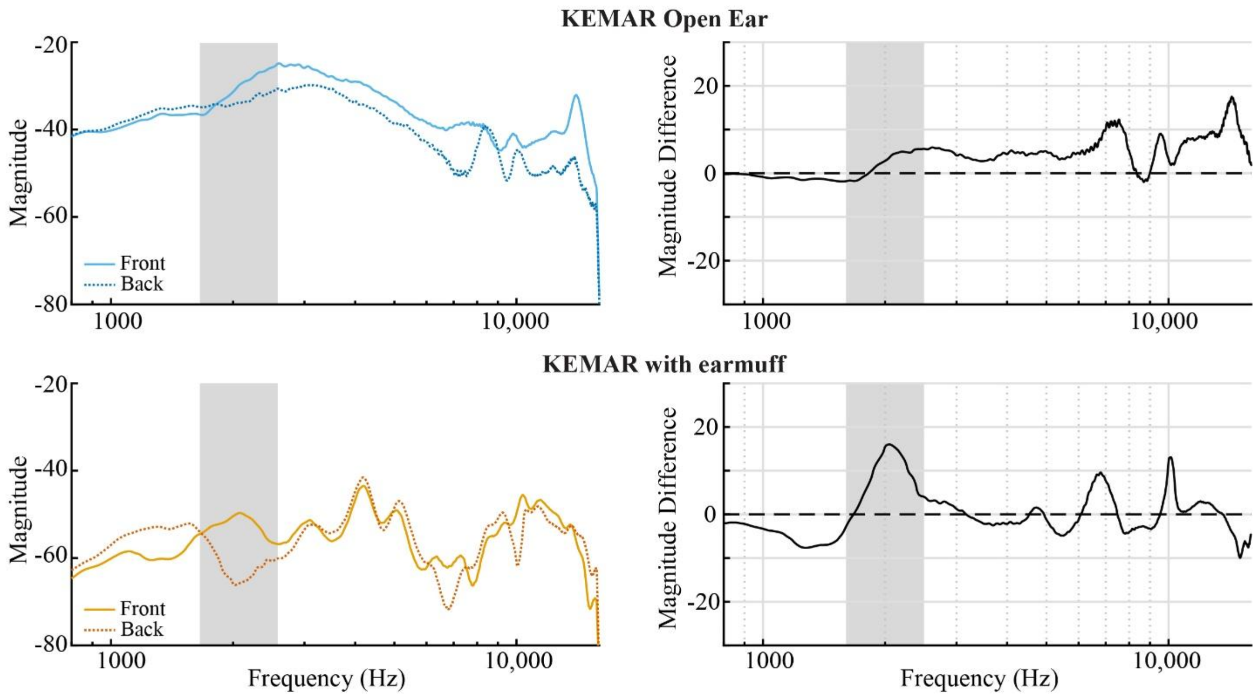

Figure 3. Acoustic measurement of KEMAR's right ear HRTF without (top) and with (bottom) the earmuff in place. Measured front (0˚ azimuth; solid line) and back (180˚ azimuth; dotted line) recordings (left) are subtracted, revealing a difference curve (right) used to describe possible acoustic information present for the discrimination of forward and rearward sound source locations. The shaded region indicates the frequency range of the Helmholtz resonator-derived notch.

4.2.5 Procedure

Figure 4 depicts the timeline/structure of all measures completed by subjects. Each type of measure is described further, below. Prior to data collection, subjects were assigned to one of two groups based on scheduling availability. Group TR completed all measures, including pretesting, training, and posttesting. Group NT (controls) completed only the pretesting and posttesting appointments with a break between the two that aligned with the training time of Group TR.

4.2.5.1 Centering

Prior to all testing, the subject was seated at the center of the array and seating height was adjusted to vertically align the ears with the speaker array via a laser-leveling system. Subjects then placed their chin in the chin-rest, facing the center speaker (0˚), to complete a behavioral centering procedure. This involved listening to 10-seconds of a white-noise stimulus from the 0˚ speaker that was immediately followed by the same stimulus from the 180˚ speaker. The subject was instructed to listen to both sounds and ensure that both the front and back stimulus sounded as if they were centered in space. Specifically, subjects were told to note if either sound was to the left or right of center. If so, the subject or earmuffs could be adjusted to ensure proper centering. This procedure was repeated until subjects were both objectively (with laser leveling) and subjectively (per reported percept) centered. Centering occurred at the beginning of each appointment, after every break, and before each run of the training procedure (described below). Notably, this procedure had to be completed more often during earmuff testing, as many variables complicated centering (see Section 4.2.7 for additional details).

4.2.5.2 Baseline testing (pretest and posttest)

During the first and last testing appointment, baseline measurement of front-back discrimination performance was completed both with subject's open ears and with the earmuffs donned. In both cases, centering was completed prior to data collection. On each trial, a single baseline stimulus was pseudo-randomly presented from one of the six active loudspeakers. Using a handheld keypad, subjects indicated the perceived stimulus hemifield by pressing one of two buttons labeled "F" (front) and "B" (back) with glow-in-the-dark adhesive stickers. For each condition, one block of 200 trials (each of the 100 baseline stimuli played once from each location) was completed, providing a 0.5% resolution in the determination of baseline front-back discrimination percent correct.

It should be noted that while the specific stimuli presented on each trial was random, overall stimulus presentation was balanced such that 100 stimuli each were presented from the front and back hemifields, and stimuli were presented from each of the three speakers in each hemifield equally. Additionally, prior to stimulus presentation, speaker compensation filters were used to minimize speaker-specific spectral coloration. Stimuli were then normalized and scaled to a per-trial-randomized level within 55-75 dB (roving). These controls guided the subjects to utilize cues embedded within the spectral shape of incoming sounds instead of other acoustic cues unlikely to be available in real-world listening contexts.

Finally, during the last appointment (posttesting), four total front-back discrimination testing runs were completed. In addition to completing testing with their open ears and the trained earmuff, the earmuff was modified such that the primary Helmholtz resonator was positioned in the forward-facing cavity, effectively reversing the trained 2 kHz spectral cue. Finally, a different Helmholtz resonator was placed in the rear-facing cavity that was tuned to 4.5 kHz. Comparison of the three earmuff-based posttesting runs was intended to provide insight on learning specificity and/or generalization to untrained spectral shapes.

4.2.5.3 Spectral weight function (SWF) measurement

Spectral weighting data for each subject was collected in the same testing environment used for other components of the experiment. Primary differences included the stimuli used (see Section 4.2.3) and the duration of testing. The spectral weighting task was completed with open ears at both pretest and posttest to evaluate train-induced changes in spectral weighting. Spectral weighting measurements were also completed with the earmuff donned at only the posttest appointment. On each trial, a SWF stimulus was randomly selected from the full set of 120 stimuli (100 manipulated stimuli plus 20 baseline stimuli trials). 960 total trials were presented over 4 blocks. Within each block (240 trials), each of the 120 SWF stimuli were played once from the forward and rearward hemifields, with speaker presentation balanced as in baseline testing. In total, each of the 120 SWF were played four times each from both front and back locations.

Although the stimuli differed markedly between this and the baseline testing appointment, the task was similar. Subjects heard a single manipulated stimulus through one of the six active speaker locations. Subjects then selected one of the two available buttons on their handheld keypad to indicate the perceived hemifield of the stimulus.

The SWF for each condition was then determined via a multivariate regularized logistic regression (using 'lassoglm' within the MATLAB Statistics Toolbox; see Analysis for additional details). Frequency band weights were extracted as the feature coefficients of the final derived model. The SWF thus indicates the frequency bands that best predicted the behavioral performance of the subject and their relative importance to other bands.

4.2.5.4 Earmuff training

During the accommodation phase, subjects completed 10 days of targeted auditory training, with the earmuff in place, following protocols similar to those described in Chapters 2 and 3. Specifically, during a single session, subjects completed three rounds of training. The first and last rounds included 90 trials (2 hemifields, 45 stimulus tokens) of 'active' audiovisual training, where subjects heard a single training stimulus from a random target speaker, responded using the keypad (identical procedure as in baseline testing), and were then presented with audiovisual feedback. If correct, subjects were presented with a single green LED at the target location and a 1-second chime presented from the target speaker. If incorrect, subjects were presented with a red flashing LED at the target location simultaneous with repetition of the target stimuli 8 times with a 500 ms interstimulus interval. The second round of training, deviating from the procedure used in Chapters 2 and 3, and inspired by Wright et al. (2010), was called 'passive training.' During this training round, subjects sat in the testing chamber with their head in the chinrest while listening to sequential presentations of training stimuli from each of the six locations in a random order, in tandem with presentation of a blue LED on the central speaker of the target hemifield. Stimuli were presented 180 times (~4 minutes). Subjects were not prompted to respond during this time but were encouraged to actively listen to presented sound sources. A single session of training thus included active, passive, and active training rounds, in that order. Subjects completed one session of training each day for 10 days. Ideally, these sessions occurred over the course of two weeks (daily training), however, due to unavoidable scheduling conflicts, some subjects extended appointments into a third week.

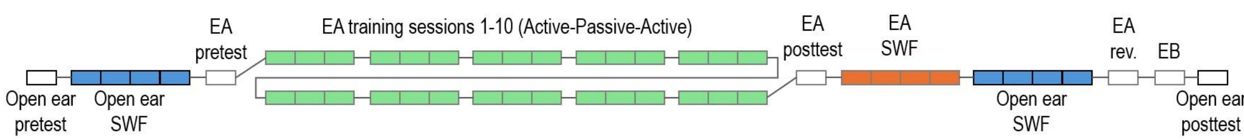

Figure 4. Diagramed study timeline. Individual boxes represent blocks of testing. Each test does not necessarily occur on a separate day (see Section 4.2.5). EA: Earmuff. EA rev: Primary earmuff with 2 kHz Helmholtz resonator in the frontal cavity. EB: 4.5 kHz Helmholtz resonator in the rear cavity

4.2.6 Analysis

Front-back discrimination performance, for the open ear pre- and posttest, earmuff pre- and posttest, and for additional earmuff testing completed at posttest, was assessed in terms of front-back confusion rate overall and for each location/hemifield independently. Performance near 50% indicates random guessing given the two-hemifield paradigm, while performance >50% indicates worse-than-chance identification of the target hemifield. Changes in performance from pre- to posttest, and comparison between Group TR and NT were assessed with non-parametric Wilcoxon signed rank tests.

Although all subjects were both objectively and subjectively aligned in the center of the speaker array (see Section 4.2.5.1), it became evident that various factors could misalign subjects, leading to asymmetric performance across target speakers. Subject performance, relative to subject centering, was assessed via comparison of performance outcomes across three pairs of speakers, (A[-5°:175°], B[0°:180°], C[5°:-175°]). With pair B representing idealized centering, balanced presentation of stimuli from the other speaker locations was intended to prevent salient left/right cues. However, if ideal centering was not achieved, presentation of sound from the left or right was no longer balanced, which resulted in decreased FBC rate for the uncentered speaker pairs, ostensibly via access to nonzero binaural differences that introduced the possibility of performance based on source lateralization rather than front-back discrimination. Therefore, the FBC rate for each speaker pair within each block was fit with a linear regression. All blocks with a regression slope>|20| were flagged. The prevalence of asymmetrical listening was 17% (51 out of 292 total test blocks). Such runs were removed from all analyses.

During SWF testing, subjects heard an equal number of stimuli from either the front or back hemifield, in randomized order. In the context of observer weighting, the fundamental question to be answered is not whether the subject was "correct" or "incorrect," which becomes tenuous to specify once the HRTF is manipulated. Rather, the question is whether stimulus adjustment (in this case manipulation of a selected frequency band) increased or decreased the probability that the subject responded 'front' vs. 'back.' For example, if more attenuation within a given band was associated with 'front' responses, more attenuation in that band for an otherwise ‘back’ stimulus could lead to reproducible ‘front’ perception (cf. Blauert, 1969). Following this framework, a multivariate regularized logistic regression algorithm was used to determine the relative importance of stimulus manipulations in their ability to predict forward or rearward perception.

Input into the regression algorithm was an 800x24 matrix of stimulus characteristics and an 800x1 vector of subject responses. The matrix of stimulus characteristics was derived as a combination of measured subject hearing thresholds and the pseudorandomized set of modified stimuli (see Section 4.2.3) where each of 24 frequency bands, determined via a gammatone filter bank from 0.8-16 kHz, was adjusted by -80, -40, -20, or 0 dB. However, as seen in Figure 1B, the frequency bands overlap, impacting the realized attenuation in each frequency band. For example, a band nominally attenuated by 80 dB that is flanked by bands with 0 dB of attenuation, will result in significantly less than 80 dB of realized attenuation. Recognizing this interaction, all final stimuli were processed via fast Fourier transform, revealing the resulting spectral level within each target frequency band. Within each stimulus, each band was then assigned a ‘realized manipulation’ value that was the difference between the spectral level within the band and the average level of the stimulus, effectively standardizing all frequency band manipulations across all stimuli. Finally, the stimulus manipulations were combined with trial-specific roving (±10 dB) and individual subject hearing thresholds which were measured during the enrollment appointment to define band-specific amplitudes in terms of sensation level. With the input of the logistic regression algorithm being this set of stimulus characteristics, the resultant model weights reflect how, for a particular listener, the spectral level at 24 unique frequency bands between 0.8 and 16 kHz, relative to the listener’s hearing thresholds, influenced their perception of sound sources in the front and back hemifields.

Practically, this relationship is modeled using a logistic regression algorithm (MATLAB's 'lassoglm' function) with elastic net regularization and 10-fold cross validation, allowing for estimation of the contribution of multiple frequency bands while controlling for multicollinearity. Elastic net regression (Zou & Hastie, 2005) employs a combination of both ridge and LASSO regularization, penalizing learned model coefficients that get too large and pushing very small coefficients to 0. These penalties help prevent overfitting of any learned model and help simplify, or prune, the number of features utilized. During model training, 10-fold cross validation is employed to tune the hyperparameters lambda (λ; penalty magnitudes) and alpha (α; balance between ridge and LASSO regularization) (Friedman et al., 2010). This method separates the data (800 total trials of manipulated stimuli) into 10 subsets (80 trials each). Across 10 cycles, a single subset is set aside to be used as a validation-set for assessing the model that is created on the remaining data. Model predictions on the validation set are used to quantify model accuracy and deviance. The estimated regression coefficients (weights) are stored for later comparison. The λ and α (regression hyperparameters) that result in the lowest deviance on the validation data are selected. A final model is then retrained on all datapoints (nothing left out) with the selected hyperparameters. Frequency band 'weights' are extracted from the resulting model coefficients. This procedure was repeated 50 times, bootstrapping each dataset to assess model stability via selection frequency for each feature coefficient (weight). Interpretation of model weights is initially in the form of log-odds changes in the prediction of forward responses given a single standard deviation increase of filter magnitude.

Assessment of SWFs addresses two primary questions: (1) are SWFs after accommodation different from those at pretest and (2) do the posttest SWFs reflect the use of earmuff-related acoustic features? Considering the working hypothesis that spectral regions supporting front-back perception are related to the acoustics of the ear, it is necessary to directly relate post-accommodation SWFs to the learned acoustics of the earmuffs. To accomplish this, the difference is taken between forward and rearward sound spectra, measured via acoustic manikin recordings with the earmuff in place (see Figure 3). This difference curve is then binned according to the gammatone filter bands used for both stimuli and SWF generation. Difference magnitudes within each bin are averaged, creating a 24-point vector that can be directly compared to SWF profiles.

Finally, condition-level SWF comparisons evaluate the convergence of SWFs between subjects. This analysis is especially useful following the expectation that open ear SWFs are highly variable across subjects (Lladó et al., 2022). Following this notion, if spectral weighting is driven by the peripheral acoustics of the ear, then training all subjects on the same set of spectral cues provided by the earmuffs should lead to improved group-level convergence across earmuff SWFs. Within-subject changes in SWFs were assessed via Kendall's Tau correlation. The use of Kendall's Tau correlation allows for nonparametric comparison of SWFs while ignoring the absolute magnitude of assigned weights. Additionally, Kendall's Tau is more robust to ties, which are likely given elastic net regularization's tendency to force small and uninformative weights to 0. Group-level assessment of training outcomes (pre- and posttest) indicate the effect of training and modification of SWFs therein. Where necessary, to correct for multiple comparisons and control the family-wise error rate, the Holm-Bonferroni correction is applied, which sequentially adjusts p-values based on their rank (Holm, 1979). Significance is then determined relative to an alpha (α) level of 0.05.

4.3 Results

4.3.1 Initially degraded performance improves with training

Figure 5 plots front-back discrimination performance for the open ear and earmuff conditions at pretest. Results demonstrate excellent open ear performance for all subjects (median=1.25%) that is significantly degraded with initial wear of the earmuff (median=49.5%; z=-3.49; p<0.001).

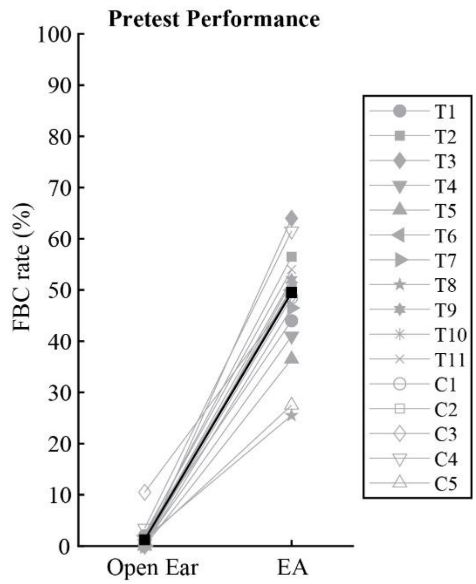

Figure 5. Front-back discrimination performance with the open ear is degraded with acute wear of the earmuff. FBC rate for all subjects individually (gray lines and symbols) as well as median performance (black lines and symbols) are shown.

As noted in Section 4.2.5, five subjects were randomly assigned to a separate ‘control’ group. These subjects did not complete training and are therefore removed from training-related analyses. Figure 6A plots FBC rate across all 10-days of training with the earmuff. Notably, each point represents the average of two ‘active training’ runs that occurred on each day. While the general trend appears to be gradual improvement over the 10 days, much variability exists. Explicit improvement is demonstrated by comparing the first (median=41.11%) and last (median=20.56%) training days (z=1.94; p=0.026; Figure 6B). While 10 out of 11 subjects recorded some improvement by training day 10, the magnitude of improvement varied substantially, with no apparent relationship between starting performance and benefit magnitude (R^2=0.03; p=0.66).

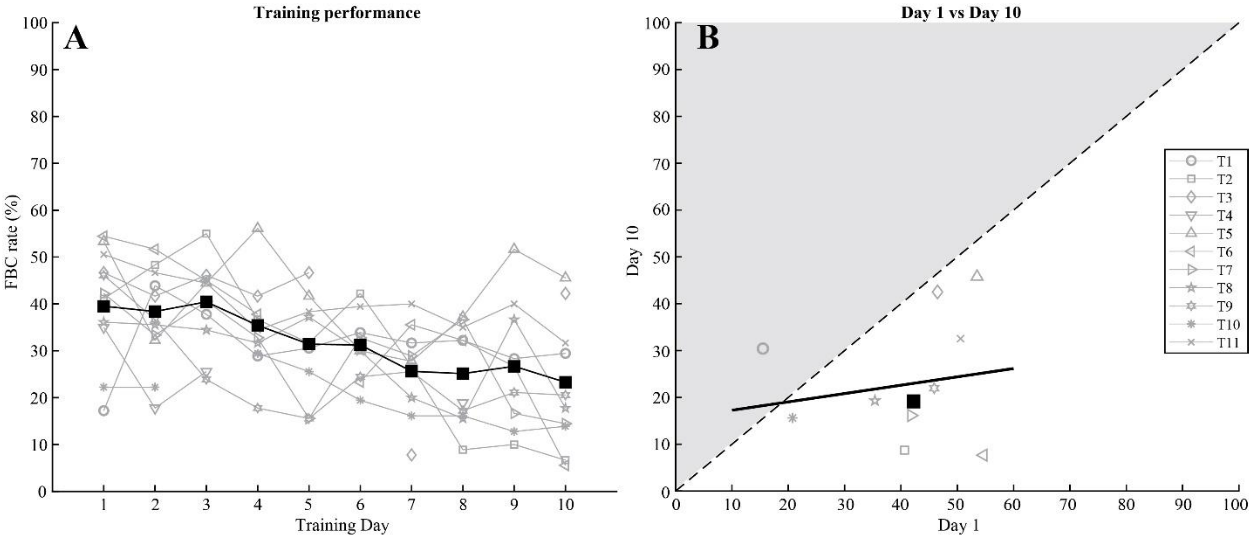

Figure 6. Improvement during training is noted for most subjects. (A) FBC rate for subjects in Group TR across 10 days of training. Individual subject performance (gray lines and symbols) as well as median performance (black lines and symbols) are shown. (B) FBC rate demonstrated on Day 1 against performance on Day 10. Individual subject performance (gray symbols) and median performance (black symbol) are demonstrated. Points falling below the black dotted line indicate improvement. A linear regression (black solid line) demonstrates no significant relationship between Day 1 and Day 10 performance.

At posttest, front-back perception with the trained earmuff (EA), see Figure 7, was significantly improved for Group TR (pretest median=50.0%; posttest median= 21.0%; z=2.13; p=0.017). Linear regression of pretest FBC rate on posttest FBC rate revealed that subjects who performed worse at pretest improved more than those who performed better at pretest (R^2=0.47; p=0.043). Performance improvement was not observed for subjects in Group NT (Figure 7, right panel), suggesting that performance improvement in Group TR was specifically due to participation in the training paradigm. While the median front-back error rate was more than halved by participation in training, it is noted that Group TR did not reach baseline open ear performance with the earmuff, with front-back error rates remaining elevated at a median of approximately 20%. This result is further considered in the context of previous spectral shape perturbation and hearing protector training studies in the Discussion.

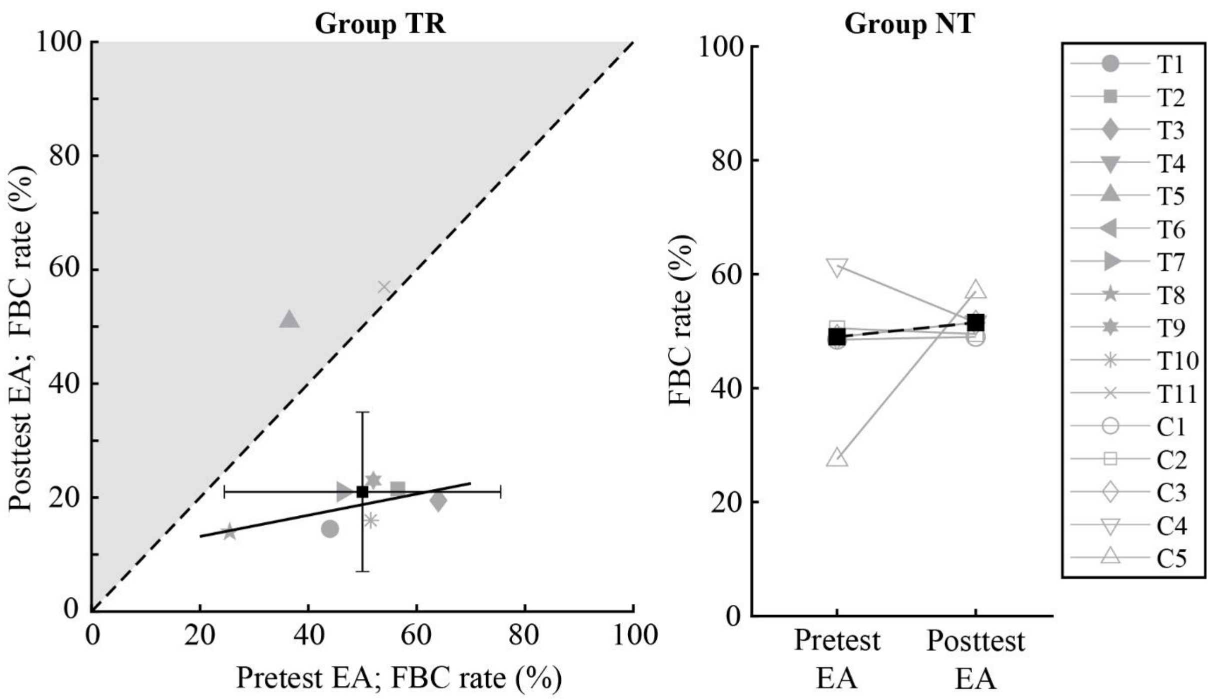

Figure 7. Poor pretest performance with the earmuff was improved following training for Group TR (left) and remains poor for Group NT (right). Individual subject FBC rate (gray lines and symbols) as well as median rate (black lines and symbols) are shown. Linear regression (black solid line) demonstrates a significant relationship between pretest and posttest FBC rate for Group TR, suggesting that subjects performing worse at pretest improved more than those performing better at pretest.

4.3.2 The specificity of learning

In addition to measurement of postest performance with the trained earmuff (EA), earmuff performance was measured in two alternative earmuff conditions. The first, condition EArev, challenged the specificity of the learned spectral feature, assessing if the intended 2 kHz spectral notch was indeed linked to rearward perception. In this condition, the 2 kHz Helmholtz resonator, previously in the rear earmuff cavity for training, was moved to the front-facing cavity. If learning was specific to a 2 kHz notch representing sources from the rear hemifield, the expected result is that subjects who demonstrated learning should be systematically incorrect, mistakingly categorizing sound spectra with a 2 kHz notch as being from the rear. Results from this task, plotted in Figure 8A, demonstrate a significant increase in FBC rate compared to the trained earmuff (EArev median=43%; z=-2.03; p=0.021), but do not reflect systematically inverted (i.e. worse-than-chance) performance compared to condition EA. This result suggests that subjects may have been utilizing information across frequency, i.e. features additional to the salient low-frequency cue provided by the earmuffs.

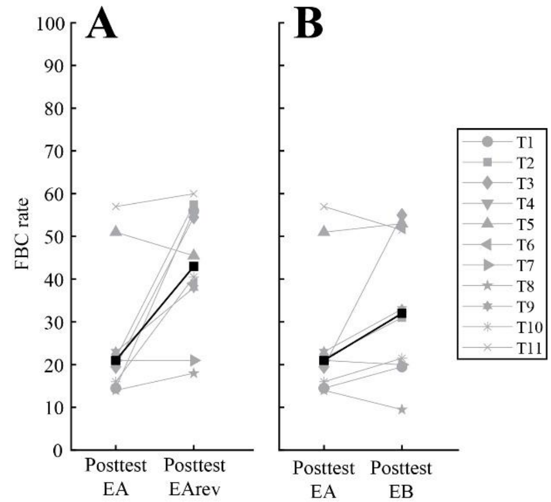

Figure 8. Posttesting with earmuffs different from that used during training suggests the use of spectral features apart from the 2 kHz notch introduced by the trained earmuff. Posttest FBC rate with the trained earmuff (EA) was compared to a condition that positions the 2 kHz Helmholtz resonator in the forward-facing cavity (EArev; panel A) and a condition that replaces the 2 kHz resonator with one tuned to 4.5 kHz (EB; panel B). Individual subject performance (gray lines and symbols) as well as median performance (black lines and symbols) are shown.

The second alternative earmuff condition replaced the trained 2 kHz Helmholtz resonator with one tuned to 4.5 kHz (EB), seeking to assess generalization of learning. Generalization, in a sound localization context, has been demonstrated to occurr from trained to untrained sound source locations (Chapter 2), but not between multiple sets of learned spectral cues (Friedrich & Schönwiesner, 2025). As designed, the dual-cavity design of the earmuff effectively removed open ear front-back spectral differences while introducing a prominent feature at the Helmholtz resonator-tuned frequency (see Figure 3). As such, by replacing only the Helmholtz resonator within the rear cavity of the earmuff, the spectral cues presented to subjects in the EB condition were quantitaviely similar to those learned during training, including the imposition of a notch for rearward sources. However, the frequency of the EB notch (4.5 kHz) was approximately one octave higher than the frequency of the trained notch (2 kHz). As shown in Figure 8B, subjects demonstrated a slight, though non-significant, increase in FBC rate compared to their posttest performance with the trained earmuff (EB median=32%; p=0.069). Furthermore, EB FBC rate remained reduced compared to EA pretest (z=2.14; p=0.016). Taken together, results from EArev and EB testing suggest that other broadband spectral features, common between EA, EArev, and EB earmuffs were utilized to aid front-back discrimination across condition.

4.3.3 Maintenance of training benefit over time

A subset of five subjects from Group TR were asked to return to complete a retest of their earmuff front-back discrimination abilities following the final posttest, assessing the stability of learning over time. Figure 9 plots the difference between FBC rate at posttest (EA) and the return test (EA2) as a function of how many days lapsed between their final posttest and this return test. For some subjects, both initial and follow-up posttesting demonstrate similar FBC rates (points near 0), reflecting maintenance of learning, while for others, performance at the return posttest reflects increased FBC rate compared to initial posttest (negative differences). Despite a negative-going trend, no significant trend was found across the subset of subjects for FBC rate change relative to time since posttest (R^2=0.3; p=0.2).

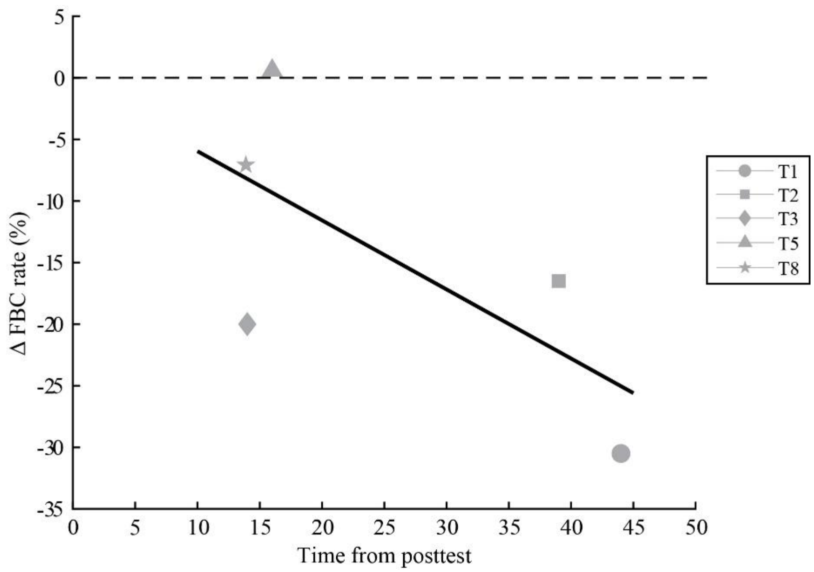

Figure 9. The difference between final posttest (EA) and the return posttest (EA2) for each subject (gray symbols) is plotted against days lapsed between EA and EA2. A line of best fit (black line) shows a non-significant trend for less performance at EA2 as more time passes.

4.3.4 Evaluation of spectral cues for front-back discrimination

As described above, all subjects completed spectral weighting function (SWF) testing prior to and following training (or in the case of Group NT subjects, following a break in study participation). This testing was conducted by presenting subjects with 800 trials of broadband noise bursts with pseudo-randomly perturbed spectra. Parameters of the known spectra were used as input, along with the binary (front or back) perceptual responses from the subjects, into a logistic regression algorithm, revealing model weights for each of 24 discrete frequency bands ranging from 0.8-16 kHz. The model weights reflect the contribution of each frequency band to the overall perception of front or back. To demonstrate the derivation of the model weights, Figure 10D plots the SWF from one example subject along with response data from three specific frequency bands (Figure 10A-C). Specifically, as all spectral perturbations are known, the percentage of frontal perceptions per perturbation magnitude can be calculated. This type of calculation is plotted across binned perturbations for three unique frequency bands in Figure 10A-C. Note, for example, that the percentage of frontal perception in Figure 10A does not change across perturbation bin. Therefore, this bin would receive a weight near 0 within the SWF, as seen in Figure 10D. In contrast, Figure 10B demonstrates a frequency band that shows an increased percentage of frontal percepts as the perturbations become more positive, and therefore, a positive weight in the SWF. Another way to describe this would be that as sound energy in that frequency bin increases, the subject is more likely to perceive the whole stimulus as originating from the front hemifield. The opposite is true for the band shown in Figure 10C, which demonstrates fewer forward percepts as energy in the bin increases. This then results in a negative weight in the SWF. Additionally note that because the slope is steeper in Figure 10C than that's shown in Figure 10B, the relative magnitude of the weight is also greater.

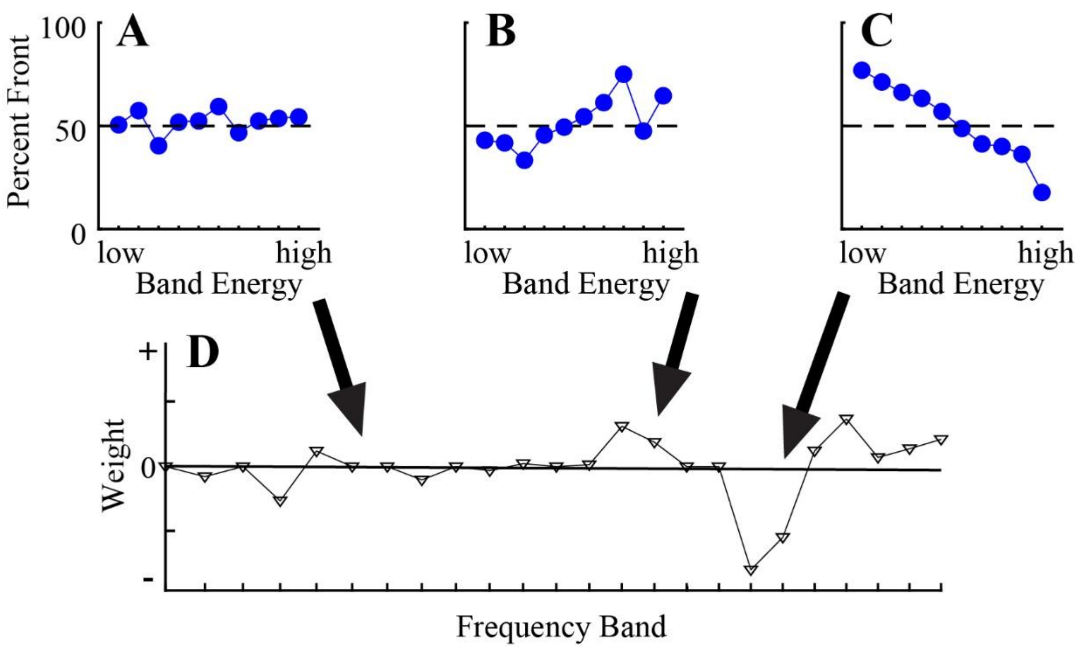

Figure 10. An example of the derivation of a SWF for a single subject. (A-C) percent front perception plotted against perturbation magnitude for three unique frequency regions. The trends seen within these plots directly relate to the representative SWF model weights for each frequency region. (D) SWF model output depicting weights for each of 24 discrete frequency bands. Weights represent the contribution to perception 'front' for a broadband stimulus.

Figure 11A plots the results of the spectral weighting analysis for all subjects, measured during the study pretest. Each subject's model weights (gray lines and symbols) have been normalized to aid in across-subject comparison. Also shown is the median weight within each frequency band (black line and symbols). The contribution of each subject's data to the median is weighted relative to the prediction accuracy for their respective model output. Prediction accuracy for all subjects is plotted in Figure 11B, demonstrating a median accuracy of 64.06%. Notably, data from Group TR and NT are plotted together, as all subjects experienced identical testing conditions at this point in data collection. While large intersubject variability is evident, the pattern of the median SWF suggests that less energy around 1.2 kHz and from 7-12 kHz, and more energy around 5 kHz and 16 kHz all lead to increased frontal perception. The implications of such pattern with respect to previous reports of 'directional bands' at similar frequencies is discussed in Section 4.4.1.3. SWFs from three example subjects are shown as bar plots in Figure 11C. These examples demonstrate the variability in spectral weighting across subjects, with some suggesting that performance is guided by information in very few frequency bands, while others suggest a more broadband use of information across the frequency range during open ear listening.

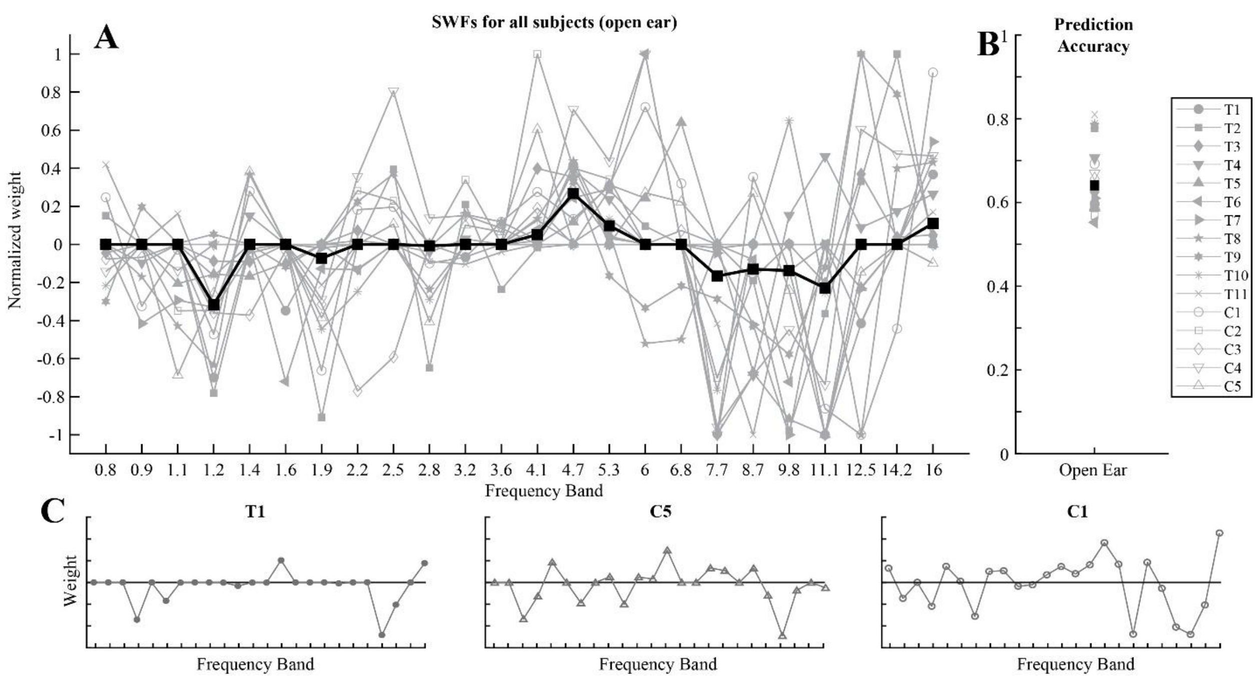

Figure 11. (A) Open ear SWF model output for all subjects, depicting weights for each of 24 discrete frequency bands. Individual subject data (gray lines and symbols) for Group TR (filled

symbols) and Group NT (open symbols) are combined for calculation of a weighted median (black line and symbols), with each subject's data weighted by the prediction accuracy (B) of the final SWF model. (C) Models from three example subjects demonstrate variability in model output across subjects

At the posttest appointment, i.e. following training on a novel set of spectral cues, each subject's open ear SWF was measured again. The purpose of this repetition is to assess the extent to which training modified spectral weights across frequency. To this end, functions are treated separately for subjects who received training (Group TR) and those who did not (Group NT). Figure 12A-B plots the posttest open ear SWFs for subjects in each group (gray lines and symbols) as well as a within-group weighted median (black lines and symbols). Also plotted are the prediction accuracies for all subjects (median=65.6%; Figure 12C). One-sided Wilcoxon signed rank tests across Kendall's Tau correlations revealed that the posttest and pretest SWFs were correlated, at the group-level, for subjects in both Group TR (median Kendall's Tau=0.49; z=2.89; p=0.002) and Group NT (median Kendall's Tau=0.41; z=1.89; p=0.030). Inspection of the pretest and posttest SWFs for individual subjects in Group TR (Figure 12D) reveal no clear increase in the weighting of information near the 2 kHz trained notch frequency. This result, consistent with posttest FBC data shown in Figure 8, suggests that learning was not solely or even primarily based on information available at the resonator frequency, but rather dependent on the use of information across frequency. This is further suggested by persistently higher weight magnitudes for high-frequency spectral bins, where spectral contrast information provided by the earmuff was reduced, but nonzero (see Figure 3). To further investigate the relationship between spectral information in high frequency regions (<4 kHz) and subject spectral weighting, the front-back acoustic contrast measured with the acoustic manikin was assessed for its correlation to the change in SWFs from pretest to posttest. This analysis revealed no significant correlations suggesting that any changes to the open ear SWFs from pretest to posttest were not directly related to the acoustics of the earmuff (described further in Section 4.4 Discussion).

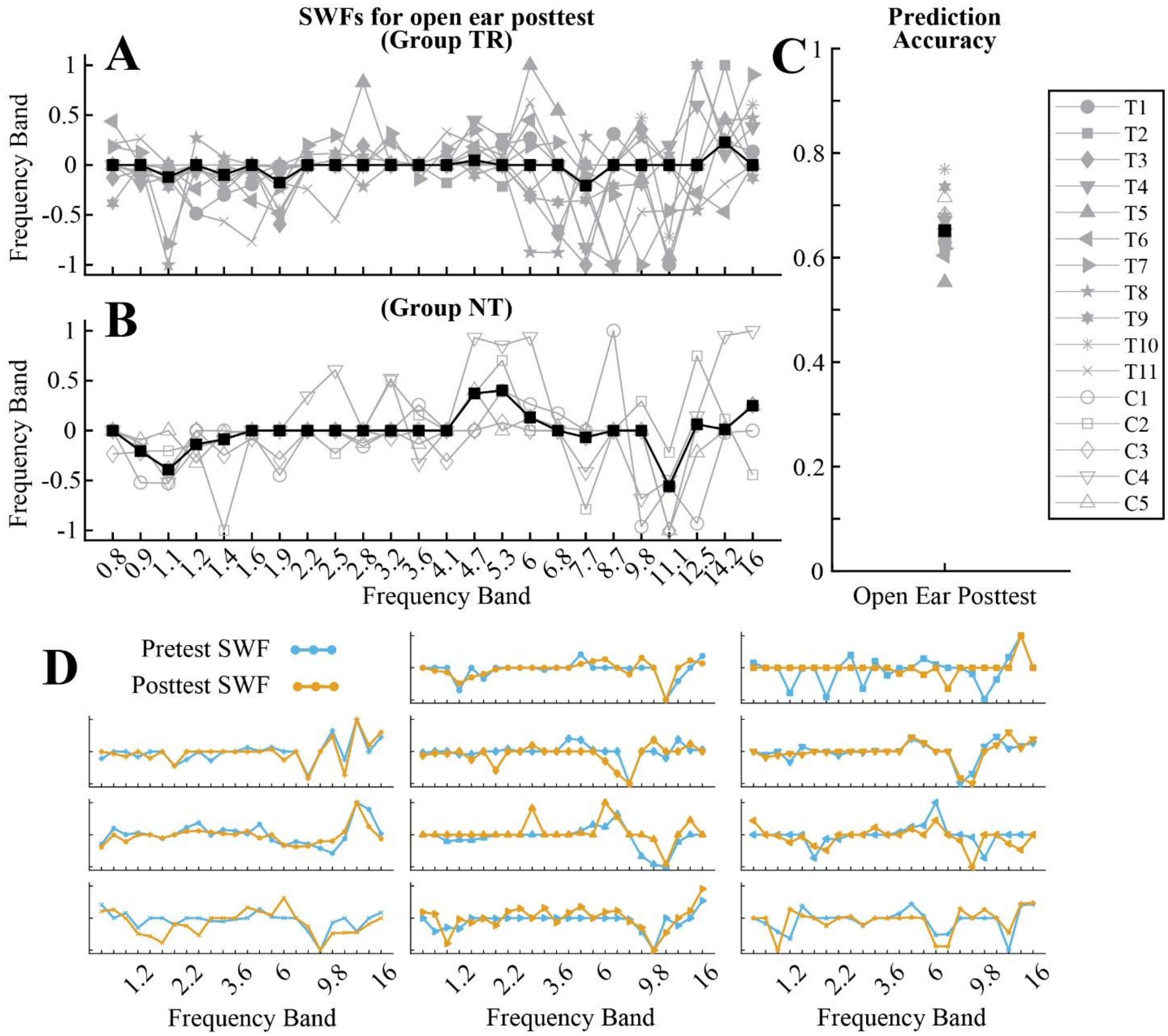

Figure 12. Posttest open ear SWF model outputs for all subjects in Group TR (A) and Group NT (B). (C) Prediction accuracy is plotted for all subjects in Group TR (filled symbols) and Group NT (open symbols). Plot details as in Figure 11. (D) Pretest and posttest SWFs for individual subjects in Group TR.

SWF testing was also completed at the posttest appointment explicitly while subjects were wearing the earmuffs. Recall from description of the SWF stimuli (Section 4.2.3), that SWF measurement also included occasional presentation of stimuli with no perturbations, meant to the stability of performance when stimuli are embedded within a more challenging context (randomly interspersed within trials of perturbed noises). As such, one might expect low FBC rate on these trials during open ear SWF testing, as subjects are well accustomed to listening with their open ears. Indeed, analysis of the unperturbed trials during both pre- and posttest open ear SWF testing reveal FBC rates <2%. During earmuff SWF testing, Group TR subjects demonstrated a median of 25.66% FBC rate across unperturbed trials (similar to their posttest EA performance). This suggests that the magnitude of improvement achieved by Group TR was reliable and robust to testing within a complex listening context.

Analyses of SWFs to this point (the comparison of pre- and posttest open ear SWFs) operate under the assumption that spectral weighting is condition-agnostic. That is, perception, regardless of hearing condition, is guided by the same set of spectral weights. However, it is also possible that the manner in which spectral information is weighted may not be static, e.g. may be subject to top-down influence that allows for the switching between different sets of learned spectral weights. Measurement and comparison of earmuff SWFs to pre- and posttest open ear SWFs was intended to provide insight on this possibility.

The measured earmuff SWFs for Group TR again demonstrated significant correlations to both pretest open ear (median Kendall's Tau=0.25; z=1.83; p=0.033) and posttest open ear (median Kendall's Tau=0.28; z=2.75; p=0.003) SWFs. However, Wilcoxon signed rank tests on pairwise Kendall's Tau correlations across subjects revealed increased variability across earmuff SWFs (z=0.61; p=0.27) compared to open ear SWFs (z=6.33; p<0.001). Earmuff SWFs for Group NT also demonstrated similarity to pretest open ear SWFs (median Kendall's Tau=0.26; z=1.89; p=0.03), although the similarity was marginal when compared with posttest open ear SWFs (p=0.053).

Finally, the earmuff SWFs of Group TR were compared (via Kendall's Tau correlation) to a manikin-based recording of the front-back difference presented by the earmuff (see Figure 3 and Section 4.2.6), directly assessing the similarity of earmuff acoustic features and the adapted spectral weighting. A one-sided Wilcoxon signed rank test revealed that the correlation of earmuff SWFs with the earmuff acoustic contrast was not significantly different from zero ($p=0.69$).

In summary, it appears that the adaptation demonstrated by subjects in Group TR over the course of 10-days of training was not driven primarily by use of the prominent 2 kHz spectral notch relayed by the earmuff, but rather by information distributed across higher frequency, for which the largest perceptual weights were generally observed.

4.4 Discussion

4.4.1 Comparison of results to other studies

4.4.1.1 Baseline performance

Previous studies of open ear front-back discrimination abilities, both within sound source localization tasks (Madjak et al., 2010; Brown et al., 2015; Trapeau et al., 2016; Chapter 2) and discrimination tasks (Blauert, 1969; Zhang & Hartmann, 2010) have reported low rates of FBCs. The performance demonstrated within this study is aligned with those observations. When spectral cues are disrupted, performance is degraded to a degree that covaries with the magnitude of spectral disruption (Russel, 1976; Carlile, 2014; Brown et al., 2015; Audet et al., 2026). While the earmuffs employed in this work were unique to this study, a method for evaluation of performance across hearing protectors was described explicitly by Audet and colleagues (2026). In this effort, behavioral performance across multiple error metrics was measured for six different hearing protection devices. A model was then built to predict behavioral performance that utilized electromechanical measurements from all six HPDs. A major component of their model referenced the degradation of open ear spectral shapes and its relationship to increased FBCs. In the present study, the earmuffs were designed specifically to both minimize any natural open ear front-back spectral differences and introduce a novel spectral feature at a previously unutilized frequency (see Figure 3). The high baseline earmuff FBC rate of 49.5% is consistent with the high (by design) spectral distortion introduced by the earmuff.

4.4.1.2 Training performance and improvement following spectral disruption

Front-back discrimination appears to be a highly adaptable behavior. Significant improvement is demonstrated with the use of virtualized audio (Madjak et al., 2010), non-occluding earmolds (Carlile et al., 2014), and hearing protection (Chapter 2 and 3). In the current study, training benefit was reduced compared to previous studies. Possible explanations for this result are considered here.

Firstly, as acknowledged above, the earmuffs developed for this study degraded spectral cues profoundly, with the intent to provide listeners with a single prominent and relatively narrow frequency region from which to extract information. Additionally, the selected region (~2 kHz) is notably uninformative in typical open ear recordings of spectral shape, an intentional selection intended to evaluate the emergence of increased spectral weighting within a novel region, if such changes in weighting were to occur. Multiple observations from the present study, including reduced learning relative to prior studies that have presented listeners with some variety of broadband distorted pinna cue(s), as well as the failure of notch hemifield inversion

(‘EArev’ condition) to invert FBC rate, suggest that front-back discrimination involves integration of features across multiple frequency regions.

While significant improvement was noted by the end of the 10-day training period, it is possible that additional training could have augmented training benefit, leading to greater adaptation to the novel introduced feature. The 10-day training paradigm employed here was anticipated to be sufficient to elicit substantial improvement in FBC rate based on prior studies which have all reported significant reduction of errors within 10 days or less across a majority of subjects (Zahorik et al., 2006; Carlile et al., 2014; Golob et al., 2024; Chapter 2). However, all of the foregoing studies utilized earmolds or earplug-style HPDs to induce spectral shape perturbations. These methods still allowed impinging sounds to interact with the pinnae, which, as noted above, defines a source of residual cues that we specifically sought to eliminate.

Apart from difficulty associated with the earmuff-related spectral cues, the task itself may have affected training outcomes. The training procedure employed in this study included many measures intended to control for the potential use of extraneous cues beyond the sound spectrum. These controls included (1) roving of overall stimulus level per trial to prevent a cross-trial level effect, (2) randomized presentation of stimuli with three different spectral tilts to prevent cross-hemifield recognition of a timbre/pitch percept, and (3) randomized presentation of stimuli from three different azimuths within each hemifield to reduce likelihood of a left-right percept that can accompany head movement. In an attempt to isolate the available perceptual cues to the earmuff-related spectral notch, task difficulty could have been consequently increased, leading to decreased task engagement, reducing training benefit.

4.4.1.3 Spectral weighting toward the identification of salient spectral features

A primary goal of this work was to elucidate important spectral features for the discrimination of sounds from the front and rear hemifields. As described in Section 4.1, front-back discrimination-relevant features have been investigated before. Most notable are the ‘directional bands’ described by Blauert (1969), that detail specific frequency regions that elicit either forward, upward, or rearward perception when stimulated with narrowband noise. When comparing the ‘directional bands’ that reflect forward perception with the SWF results outlined in Figure 10, there is a striking amount of overlap: less energy around 1 kHz and 8 kHz, and more energy around 4 kHz and 16 kHz all lead to increased forward perception with open ears. One critique of the results from Blauert was the lack of roving utilized in their stimuli presentation, potentially leading to an additional cross-trial level-cue. However, the current work demonstrated similar results while controlling for that and other potentially confounding cues.

Other studies exploring spectral weighting for sound localization typically combine front-back and elevation perception into the single domain of ‘sagittal plane localization.’ The assumption made in such studies is that rearward perception is perception of sounds >90° along the ‘elevation’ dimension. While results from these studies emphasize the importance of the 6-9 kHz ‘notch’ region (Ahrens & Brimijoin, 2019; Zonooz et al., 2018; Lladó et al., 2025), largely due to the elevation-dependence of notch frequency (Brungart & Rabinowitz, 1999; Van Wanrooij & Van Opstal, 2005), that same region has been implicated in distinguishing source hemifield (Blauert, 1969; Figure 10). It would be valuable to further investigate how such a region can be utilized for decisions along both a continuous (elevation) and binary (front-back) set of locations simultaneously.

4.4.2 Effects of the earmuff on measured spectral weights

The primary goal of measuring the spectral weighting functions (SWFs) of the listeners both before and after training was to investigate the influence of learned spectral shape features on spectral weighting. The hypothesis followed that if specific spectral features defined at the level of the ear were the primary factor driving decoding of source hemifield, then the SWFs from pretest to posttest would differ, as a new set of features would have to be learned to successfully determine front and back with the earmuff. In contrast, if the incoming spectra were compared to less flexible representations of spectral shapes (as described in Chapter 1), then the SWFs would be relatively more static, reflecting the auditory system's attempt to leverage nuanced front-back information within persistent frequency regions.

The results described above for the trained subjects (Group TR), demonstrated correlation between all SWF conditions, and, importantly, a lack of increased weighting of the earmuff-specific 2 kHz region. However, while trained subjects generally did not produce increased weights in the region surrounding the introduced notch, comparison of pre- and posttest SWFs did suggest that SWFs were not static (i.e., while the correlation of pre- and posttest was greater than zero, it was also well less than 1). Changes were observed in median SWFs from pretest to posttest with the open ear. However, those changes were not correlated with the acoustic information provided by the earmuffs, even when assessing only information above 4 kHz. In theory, in order to maintain improved performance with the earmuff concurrent with near-perfect performance with open ears, as still observed at posttest, subjects would produce a SWF that accommodates information available both via the earmuffs and the open ear. This idea, called 'many-to-one' mapping, describes a situation where the spectral shape profile for a single location is adapted to accommodate multiple spectral profiles (Trapeau et al., 2016; Friedrich &

Schönwiesner, 2025). Support for such a schema comes from studies referencing a decision-model based on the correlation between incoming sound spectra and stored spectral shapes (Middlebrooks, 1992; Langendijk & Bronkhorst, 2002; Baumgartner et al., 2014). Interpreted in this way, the positive but incomplete training outcomes for our subjects suggest that the acoustic features introduced by the earmuff were partially but insufficiently adapted within static internal spectral profiles.

4.5 Future directions

The training outcomes demonstrated here, while incomplete, highlight both the potential of and a limit to the plasticity of spectral cue adaptation observed here and in prior studies. Future work should aim to further assess this limit, investigating the effects of training time on similarly degraded spectral cues. It is possible that additional training e.g., over a period of a month or more, could have produced a larger training benefit with the earmuff. It is correspondingly possible that more substantial and more consistent changes in subjects' SWFs may have been observed with additional training. A longer training time-course was not anticipated to be necessary based on past work discussed herein, but future work could evaluate more extensive training schedules. Additionally, other spectral perturbations, including the introduction of prominent features at higher frequencies overlapping with natural pinna cues, which may be intrinsically easier to relearn, could be explored. Further, given sufficient training, SWFs with the earmuff in place could be more reliably measured, providing additional insight on whether altered spectral cues are adapted within a single representation of spectral shapes ('many-to-one' mapping) or learned as an independent set of spatial representations.

The present results suggest that, in addition to improved spatial hearing through training previously demonstrated with earplugs, some training benefit might also be observed with appropriately designed earmuffs. In a hearing conservation and occupational health context, earmuffs are the preferred choice within some organizations, as they provide certain practical (e.g. ergonomic and/or hygienic) advantages over earplugs. Moreover, high-intensity noise environments, prompting the use of double hearing protection, require the use of earmuffs. In acute laboratory studies, earmuffs tend to produce substantially larger spatial hearing deficits than earplugs (Russell, 1976; Audet, 2026). Development of earmuffs that lead to improved spatial hearing relative to currently available devices would reduce the burden of use currently placed on the user, potentially leading to improved device retention and increased prevention of noise-induced hearing loss.

4.6 Acknowledgements

This work was supported in part by the University of Washington and by the US Army Medical Research and Development Command (Contract HT9425-24-9-0022). The authors would like to thank Drs. Yi Shen and Bev Wright for useful discussions regarding experimental design, and Drs. David Dall’Osto and Erik Peterson for useful discussions regarding earmuff design.

Chapter 5: Conclusion

Successful sound source localization is important for safety, social communication, and enjoyment of everyday life. This dissertation has explored circumstances under which sound localization can be acutely degraded, factors associated with such degradation and means by which localization abilities may be partially restored.

In particular, when wearing hearing protection devices, sound localization cues derived from the auditory periphery are disrupted, leading to significant localization errors. Generally, these errors present as a collapsing of elevation perception to the horizon and significantly increased front-back confusions. However, accommodation of the altered sound localization cues can be achieved with enough time and/or explicit auditory training. In the preceding chapters, factors related to the improvement of sound localization abilities following training were explored, including the degree of spectral cue distortion caused by the hearing devices, training duration, and feedback type. At the end of Chapter 1, three questions were proposed. These questions, and insights gained through the work presented, are revisited in the following sections.

5.1 How is the accommodation of altered spectral cues impacted by the infrequent use of hearing protectors?

Chapter 2 detailed a study exploring front-back perception along the horizontal plane both before and after auditory training during the use of commercial, military-grade earplugs. Importantly, the earplugs were only worn during the research appointments (~45 minutes per week) and training included robust audiovisual feedback via paired LEDs and target sound repetition. Results of this work demonstrated significant improvements over the course of eight weekly (once per week) training sessions. These results were in-line with previous work demonstrating similar improvements via daily training sessions using non-occluding earmolds (Carlile et al., 2014), but the gains observed are especially noteworthy given the significant deviations of the spectral shape cues from the open ear to the earplug. Additionally, the improvement during training generalized to untrained locations. This generalization suggested that listeners learned a continuous representation of spectral changes across azimuth, allowing for improvement at locations intermediate to the trained locations.

Results described in Chapter 2, in the context of other work, highlight the plasticity of spectral cues for front-back discrimination. Synthesizing work across studies, improvement in front-back discrimination has now been demonstrated rapidly over a single day (Zahorik et al., 2006; Madjak et al., 2010), over several consecutive days (Golob et al., 2024; Chapter 3), and over short weekly sessions (Chapter 2; Stitt et al., 2019). These results collectively indicated the capacity for improvement of front-back discrimination given access to distorted pinna-derived spectral cues. However, in assessing the ecological relevance of results thus far, given that performance was only assessed in azimuth, questions remained regarding the resolution of errors along multiple spatial dimensions at once.

5.2 Can front-back discrimination and elevation perception be concurrently improved through auditory training with hearing protectors?

A limitation of the results described in Chapter 2 is that subjects were only presented with source locations along the horizontal plane. A similar limitation is present in studies measuring improvement in elevation perception after spectral shape distortion while utilizing sound source locations only in the front hemifield (Hofman et al., 1998; Carlile et al., 2014; Trapeau et al., 2016; Friedrich & Schönwiesner, 2025). Notably, while both types of studies yield significant improvement in localization perception within their target domain, adaptation across both front-back and elevation perception simultaneously has seldom been reported, or even explored (Zahorik et al., 2006). Chapter 3 described a study aiming to explore the potential interaction between these domains. While utilizing similar procedures and earplugs as those in Chapter 2, the speaker locations were substantially different. During pre- and posttesting, noise bursts were presented from a 24-speaker array sampling 360° space at eight unique azimuths and three elevations. Then, during 10 daily training sessions, participants were presented stimuli from 12 locations spanning six elevations either directly in front or directly behind them. Changes in both front-back discrimination and elevation perception were measured. Improvement was demonstrated in terms of both front-back discrimination and elevation accuracy, although front-back discrimination improved both more rapidly and more completely at all trained locations along the midsagittal plane. Further, the generalization of performance gains to locations at other azimuths was limited. Taken together, results suggested an interaction between the spectral cues for front-back discrimination and elevation perception that may have limited simultaneous adaptation over the two-week time course studied. Nonetheless, the asymptotic improvement of front-back discrimination toward near-normal levels across a two-week time course, coupled with the results of Chapter 2, motivated the design of a final experiment in which the basis of and limitations on successful front-back discrimination were studied more fully.

5.3 What spectral features are utilized for successful front-back discrimination and how does use of those features change following adaptation to altered spectral cues?

In an attempt to elucidate the specific spectral features involved in front-back discrimination, Chapter 4 described a study aimed at revealing key spectral regions utilized by listeners specifically completing a front-back discrimination task while wearing custom-designed earmuffs that profoundly disrupted natural spectral shape cues. Measurement of those regions was accomplished via a spectral weighting procedure that involved collecting binary (front or back) responses to stimuli with pseudorandom perturbations. Analysis of the open ear data, via logistic regression, revealed multiple frequency regions that influenced response direction. Those regions resembled Blauert's 'directional bands' (1969), which indicated that forward perception was guided by decreased energy around 1 kHz, increased energy around 4 kHz, and decreased energy around 10 kHz.

Following measurement of the spectral weighting function with open-ears, subjects were fit with a pair of earmuffs designed to eliminate prominent front-back spectral differences apart from a single introduced spectral notch at 2 kHz. Subjects were then trained, over the course of 10 sessions (2-3 weeks), leading to improvement in front-back perception while wearing the earmuffs. However, the rate of FBCs only improved to roughly 20%. This is notably reduced compared to open ear baseline performance of ~1% or the improvement of earplug-mediated performance, on the order of 5-10%, reported in prior chapters. This result suggests a limit to the plasticity of front-back perception that was not anticipated a priori, given the consistent success of past studies in efficiently training front-back perception with distorted cues. The most obvious explanation is that the manipulation imposed here eliminated, rather than merely distorted, natural pinna cues. It was also the case that the spectral feature relayed by the earmuff was both low-frequency and narrowband, perhaps too dissimilar from the primarily high-frequency features provided by natural pinnae, and potentially not suitable to stimulate the auditory neurons thought to be responsible for encoding spectral notches (Voigt & Young, 1980; Nelken & Young, 1994; Davis, 2002).

5.4 Future Directions

The work described throughout this dissertation inspires multiple avenues of future work. In an applied context, the work improves our understanding of pathways to restore sound source localization during the use of hearing protection devices – perhaps the most ubiquitous and consequential instance of disrupted median plane localization. Results demonstrate that meaningful improvement in localization can be realized in a matter of hours of auditory training, whether spread across daily sessions or infrequent weekly sessions. Audiovisual feedback, which could be included as part of a training program, both accelerates and augments training benefit.

In lieu of developing a training paradigm to improve the use of current hearing protection devices that disrupt sound localization cues, novel hearing devices could be developed that minimize spectral cue distortion. While this also requires a more complete understanding of the spectral cues for elevation and front-back perception, as well as how those cues manifest for the individual, results from the final experiment reported here suggest that introducing narrowband cues in restricted spectral regions, even if acoustically large and unambiguous, may be less effective than seeking to restore more broadband (pinna-like) cues that the system is naturally adapted to leverage. However, development of hearing devices that introduce novel sets of salient and easy-to-learn broadband cues could provide another path to restore performance.

Finally, some of the same limitations observed during the use of hearing protectors are observed during the use of other hearing devices, such as hearing aids for the treatment of hearing loss. Considering that some of the most common hearing losses involve the loss of sensitivity to high-frequency sound important for sound localization, future studies could assess the impact of hearing loss on both spectral cue adaptation and spectral cue weighting.

Additionally, hearing aid technology could be leveraged to enhance informative spectral features potentially promoting restoration of sound localization abilities.

Bibliography

Abbagnaro, L. A., Bauer, B. B., & Torick, E. L. (1975). Measurements of diffraction and interaural delay of a progressive sound wave caused by the human head. II. J. Acoust. Soc. Am., 58(3), 693-700.

Abel, S.M., Hay, V.H. (1996). Sound localization the interaction of aging, hearing loss and hearing protection. Scandinavian Audiology 25:3-12. doi:10.3109/01050399609047549

Ahrens A & Brimijoin O. (2019). Perceptual weighting of elevation localization cues across frequency. In 177th Meeting of the Acoustical Society of America.

Anderson, S., Kraus, N. (2013). Auditory training: evidence for neural plasticity in older adults. Perspectives on hearing and hearing disorders. Research and research diagnostics 17:37–57. https://doi.org/10.1044/hhd17.1.37

ANSI (2010). Methods For The Measurement Of Insertion Loss Of Hearing Protection Devices In Continuous Or Impulsive Noise Using Microphone-In-Real-Ear Or Acoustic Test Fixture Procedures (S12.42). Retrieved from https://webstore.ansi.org/Standards/ASA/ANSIASAS12422010

Aslin, R.N., Pisoni, D.B., Hennessy, B.L., & Perey, A.J. (1981). Discrimination of voice onset time by human infants: new findings and implications for the effects of early experience. Child development 52:1135–1145.

Atherley, G. R. C., & Noble, W. G. (1970). Effect of ear-defenders (ear-muffs) on the localization of sound. Occupational and Environmental Medicine, 27(3), 260-265.

Audet, D. J., Hunsaker, A. A., Butler, M., Sammeth, C., Podolski, A., Argo, T. F., ... & Brown, A. D. (2026). Two-dimensional sound localization during hearing protector use: Human performance and acoustic prediction. J. Acoust. Soc. Am., 159(4), 3130-3147.

Balachandar, K., & Carlile, S. (2019). The monaural spectral cues identified by a reverse correlation analysis of free-field auditory localization data. J. Acoust. Soc. Am., 146(1), 29–40. https://doi.org/10.1121/1.5113577

Batteau, D. W. (1967). The role of the pinna in human localization. Proceedings of the Royal Society of London. Series B. Biological Sciences, 168(1011), 158–180. https://doi.org/10.1098/rspb.1967.0058

Baumgartner, R., Majdak, P., & Laback, B. (2014). Modeling sound-source localization in sagittal planes for human listeners. J. Acoust. Soc. Am., 136(2), 791-802. doi: 10.1121/1.4887447

Baumgartner, R., Majdak, P., & Laback, B. (2016). Modeling the effects of sensorineural hearing loss on sound localization in the median plane. Trends in Hearing, 20, 2331216516662003. doi: 10.1177/2331216516662003

Bavelier, D., Tomann, A., Hutton, C., Mitchell, T., Corina, D., Liu, G., & Neville, H. (2000). Visual attention to the periphery is enhanced in congenitally deaf individuals. J. Neurosci. 20:RC93. doi:10.1523/jneurosci.20-17-j0001.2000

Berger, E.H. (2000). Hearing protection devices. In The Noise Manual. (pp. 379-454). Fairfax: American Industrial Hygiene Association.

Best, V., Carlile, S., Jin, C., & Van Schaik, A. (2005). The role of high frequencies in speech localization. J. Acoust. Soc. Am., 118(1), 353-363. doi: 10.1121/1.1926107

Blauert, J. (1969). Sound Localization in the Median Plane. 22.

Bockstael, A., Keppler, H., Botteldooren, D. (2015). Musician earplugs: Appreciation and protection. Noise and Health, 17:198–208. doi:10.4103/1463-1741.160688

Bolia, R.S., Dangelo, W.R., Mishler, P.J., & Morris, L.J. (2001). Effects of hearing protectors on auditory localization in azimuth and elevation. Human Factors: The Journal of the Human Factors and Ergonomics Society 43:122-128. doi:10.1518/001872001775992499

Bolia, R.S., & McKinley, R.L. (2000). The Effects of Hearing Protectors on Auditory Localization: Evidence From Audio-Visual Target Acquisition. International Journal of Occupational Safety and Ergonomics, 6(3), 309-319. doi: 10.1080/10803548.2000.11076457

Brammer, A.J., Yu, G., Bernstein, E.R., Cherniack, M.G., & Peterson, D.R. (2014). Understanding speech when wearing communication headsets and hearing protectors with subband processing. J. Acoust. Soc. Am. 136:671–681. https://doi.org/10.1121/1.4883385

Brenner, N., Bialek, W., & de Ruyter van Steveninck, R. (2000). Adaptive rescaling maximizes information transmission. Neuron. 26:695-702. doi: 10.1016/s0896-6273(00)81205-2.

Brown, A. D., Beemer, B. T., Greene, N. T., Argo, T., Meegan, G. D., & Tollin, D. J. (2015). Effects of Active and Passive Hearing Protection Devices on Sound Source Localization, Speech Recognition, and Tone Detection. PLOS ONE, 10(8), e0136568. https://doi.org/10.1371/journal.pone.0136568

Brown, A.D., Rodriguez, F.A., Portnuff, C.D., Goupell, M.J., & Tollin, D.J. (2016). Time-varying distortions of binaural information by bilateral hearing aids: effects of nonlinear frequency compression. Trends in hearing, 20, 2331216516668303. doi: 10.1177/2331216516668303

Brungart, D.S., Hobbs, B.W., Hamil, J.T. (2007). A comparison of acoustic and psychoacoustic measurements of pass-through hearing protection devices. 2007 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics. doi:10.1109/aspaa.2007.4393042

Brungart, D. S., & Rabinowitz, W. M. (1999). Auditory localization of nearby sources. Head-related transfer functions. J. Acoust. Soc. Am., 106(3), 1465–1479.

Buck, K., & Zimpfer-Jost, V. (2004). Active hearing protection systems and their performance. NATO Lecture Series HFM-111 on Personal Active Noise Reduction.

Butler, R. A., & Belendiuk, K. (1977). Spectral cues utilized in the localization of sound in the median sagittal plane. J. Acoust. Soc. Am., 61(5), 1264–1269. https://doi.org/10.1121/1.381427

Butler, R. A., & Humanski, R. A. (1992). Localization of sound in the vertical plane with and without high-frequency spectral cues. Perception & Psychophysics, 51(2), 182–186. https://doi.org/10.3758/BF03212242

Byrne, D., & Noble, W. (1998). Optimizing sound localization with hearing aids. Trends in Amplification, 3(2), 51-73. doi: 10.1177/108471389800300202

Carlile, S. (2014). The plastic ear and perceptual relearning in auditory spatial perception. Frontiers in Neuroscience, 8. https://doi.org/10.3389/fnins.2014.00237

Carlile, S., Balachandar, K., Kelly, H. (2014). Accommodating to new ears: the effects of sensory and sensory-motor feedback. J. Acoust. Soc. Am. 135, 2002–2011. doi: 10.1121/1.4 86 8369

Carlile, S., Blackman, T. (2014). Relearning auditory spectral cues for locations inside and outside the visual field. J. Assoc. Res. Otolaryngol. 15:249-263. doi:10.1007/s10162-013-0429-5

Carlile, S., Delaney, S., & Corderoy, A. (1999). The localisation of spectrally restricted sounds by human listeners. Hearing research, 128(1-2), 175-189. doi: 10.1016/S0378-5955(98)00205-6

Carlile, S., Leong, P., & Hyams, S. (1997). The nature and distribution of errors in sound localization by human listeners. Hearing research, 114(1-2), 179-196. doi: 10.1016/S0378-5955(97)00161-5

Casali, J. G., Ahroon, W. A., & Lancaster, J. A. (2009). A field investigation of hearing protection and hearing enhancement in one device: For soldiers whose ears and lives depend upon it. Noise and Health, 11(42), 69-90. doi: 10.4103/1463-1741.48564

Casali, J. G., & Robinette, M. B. (2015). Effects of user training with electronically-modulated sound transmission hearing protectors and the open ear on horizontal localization ability. International Journal of Audiology, 54(sup1), S37–S45. https://doi.org/10.3109/14992027.2014.973538

Casali, J.G., Talcott, K.A., Keady, J.P., & Killion, M.C. (2012). Warfighter auditory situation awareness: Locating the shooter with and without hearing protection. Naval Engineers Journal 124:149-159.

Chaudhuri S., Dey D., Bandyopadhyay S. & Das A. (2017). A study on variability of anthropometric data of various pinna patterns of human ear for individualization of head related transfer function: A technical review. Computer, Communication and Electrical Technology, 313–317.

Dahmen, J.C., Keating, P., Nodal, F.R., Schulz, A.L., & 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

Davis, K. A. (2002). Evidence of a Functionally Segregated Pathway From Dorsal Cochlear Nucleus to Inferior Colliculus. Journal of Neurophysiology, 87(4), 1824–1835. https://doi.org/10.1152/jn.00769.2001

Denk, F., Ewert, S.D., & Kollmeier, B. (2019). On the limitations of sound localization with hearing devices. J. Acoust. Soc. Am., 146(3), 1732-1744. doi: 10.1121/1.5126521

Devor, M., Wall, P. (1981). Plasticity in the spinal cord sensory map following peripheral nerve injury in rats. J. Neurosci. 1:679-684. doi:10.1523/jneurosci.01-07-00679.1981

Farina, A. (2007, May 5-8). Advancements in impulse response measurements by sine sweeps [Paper presentation]. Audio Engineering Society 122nd Convention, Vienna, Austria.

Fisher, N.I., Lewis, T., & Embleton, B.J. (1993). Statistical analysis of spherical data. Cambridge University Press.

Friedman JH., Hastie T. & Tibshirani R. (2010). Regularization paths for generalized linear models via coordinate descent. Journal of Statistical Software, 33, 1–22.

Friedrich, P., & Schönwiesner, M. (2025). Adaptation rate and persistence across multiple sets of spectral cues for sound localization. J. Acoust. Soc. Am., 157(3), 1543–1553. https://doi.org/10.1121/10.0036056

Glasberg BR. & Moore BCJ. (1990). Derivation of auditory filter shapes from notched-noise data. Hearing Research, 47(1), 103–138. https://doi.org/10.1016/0378-5955(90)90170-T

Golob, E. J., Nelson, J. T., Walasek, J. A., Piedra, E. T., Mock, J. R., Adams, M. J. K., ... & Spear, C. S. A. (2024). Training to improve spatial hearing and situation awareness when wearing hearing protection. Military medicine, 189(1-2), e306-e312.

Hawkey, D.J., Amitay, S., Moore, D.R. (2004). Early and rapid perceptual learning. Nature Neuroscience 7:1055-1056. doi:10.1038/nn1315

Hebrank, J., & Wright, D. (1974). Spectral cues used in the localization of sound sources on the median plane. J. Acoust. Soc. Am., 56(6), 1829-1834. doi: 10.1121/1.1903520

Hofman, P. M., Van Riswick, J. G., & Van Opstal, A. J. (1998). Relearning sound localization with new ears. Nature neuroscience, 1(5), 417-421.

Holm S. (1979). A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics, 65–70.

Irvine, D. R. F. (2018). Auditory perceptual learning and changes in the conceptualization of auditory cortex. Hearing Research, 366, 3–16. https://doi.org/10.1016/j.heares.2018.03.011

Jeffress, L. A., Blodgett, H. C., & Deatherage, B. H. (1962). Effect of Interaural Correlation on the Precision of Centering a Noise. J. Acoust. Soc. Am., 34(8), 1122–1123. https://doi.org/10.1121/1.1918257

King, A.J., Parsons, C.H., & Moore, D.R. (2000) Plasticity in the neural coding of auditory space in the mammalian brain. Proceedings of the National Academy of Sciences 97:11821-11828. https://doi.org/10.1073/pnas.97.22.11821

Knudsen, E., Brainard, M. (1991). Visual instruction of the neural map of auditory space in the developing optic tectum. Science 253:85-87. doi:10.1126/science.2063209

Kuhn, G. F. (1977). Model for the interaural time differences in the azimuthal plane. J. Acoust. Soc. Am., 62(1), 157-167.

Kumpik, D.P., Campbell, C., Schnupp, J.W., & King, A.J. (2019). Re-weighting of sound localization cues by audiovisual training. Frontiers in neuroscience, 13, 1164. doi: 10.3389/fnins.2019.01164

Kumpik, D.P., & King, A.J. (2019). A review of the effects of unilateral hearing loss on spatial hearing. Hearing research, 372, 17-28. doi: 10.1016/j.heares.2018.08.003

Langendijk EHA. & Bronkhorst AW. (2002). Contribution of spectral cues to human sound localization. J. Acoust. Soc. Am., 112(4), 1583–1596. https://doi.org/10.1121/1.1501901

Lee, K., Casali, J.G. (2019). Learning to localize a broadband tonal complex signal with advanced hearing protectors and TCAPS: The effectiveness of training on open-ear vs. device-occluded performance. International Journal of Audiology 58:Sup1. doi:10.1080/14992027.2018.1532118

Little, D.F., Zhang, Y.X., & Wright, B.A. (2017). Disruption of perceptual learning by a brief practice break. Current Biology 27:3699-3705. https://doi.org/10.1016/j.cub.2017.10.032

Lladó, P., Hyvärinen, P., Kim, T., & Pulkki, V. (2022). Using reverse correlation to derive spectral weights for median plane localisation. International Congress on Acoustics.

Lladó P., Majdak P., Barumerli R. & Baumgartner R. (2025). Spectral weighting of monaural cues for auditory localization in sagittal planes. Trends in Hearing, 29, 23312165251317027.

Macaulay, E. J., Hartmann, W. M., & Rakerd, B. (2010). The acoustical bright spot and mislocalization of tones by human listeners. J. Acoust. Soc. Am., 127(3), 1440–1449. https://doi.org/10.1121/1.3294654

Macpherson, E. A., & Middlebrooks, J. C. (2000). Localization of brief sounds: Effects of level and background noise. J. Acoust. Soc. Am., 108(4), 1834–1849. https://doi.org/10.1121/1.1310196

Macpherson, E. A., & Middlebrooks, J. C. (2002). Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited. J. Acoust. Soc. Am., 111(5), 2219–2236. https://doi.org/10.1121/1.1471898

Macpherson, E.A., & Sabin, A.T. (2013). Vertical-plane sound localization with distorted spectral cues. Hearing research, 306, 76-92. doi: 10.1016/j.heares.2013.09.007

Majdak, P., Baumgartner, R., & Jenny, C. (2020). Formation of Three-Dimensional Auditory Space. In J. Blauert & J. Braasch (Eds.), The Technology of Binaural Understanding (pp. 115–149). Springer International Publishing. https://doi.org/10.1007/978-3-030-00386-9_5

Majdak, P., Goupell, M. J., & Laback, B. (2010). 3-D localization of virtual sound sources: Effects of visual environment, pointing method, and training. Attention, Perception, & Psychophysics, 72(2), 454–469. https://doi.org/10.3758/APP.72.2.454

Makous, J. C., & Middlebrooks, J. C. (1990). Two-dimensional sound localization by human listeners. J. Acoust. Soc. Am., 87(5), 2188–2200. https://doi.org/10.1121/1.399186

May, B. J. (2000). Role of the dorsal cochlear nucleus in the sound localization behavior of cats. Hearing Research, 148(1–2), 74–87. https://doi.org/10.1016/S0378-5955(00)00142-8

Mendonça, C., Campos, G., Dias, P., & Santos, J.A. (2013). Learning auditory space: Generalization and long-term effects. PloS one, 8(10), e77900. doi: 10.1371/journal.pone.0077900

Merzenich, M.M., Nelson, R.J., Stryker, M.P., Cynader, M.S., Schoppmann, A., & Zook, J.M. (1984). Somatosensory cortical map changes following digit amputation in adult monkeys. Journal of Comparative Neurology 224:591-605. doi:10.1002/cne.902240408

Mesulam, M. (1998). From sensation to cognition. Brain 121:1013-1052. doi:10.1093/brain/121.6.1013

Meyer, J., & Picinali, L. (2025). On the generalization of accommodation to head-related transfer functions. J. Acoust. Soc. Am., 157(1), 420-432.

Middlebrooks, J. C. (1992). Narrow-band sound localization related to external ear acoustics. J. Acoust. Soc. Am., 92(5), 2607-2624.

Middlebrooks, J.C., & Green, D.M. (1991). Sound localization by human listeners. Annual review of psychology, 42(1), 135-159. doi: 10.1146/annurev.ps.42.020191.001031

Mills, A. W. (1958). On the Minimum Audible Angle. J. Acoust. Soc. Am., 30(4), 237–246. https://doi.org/10.1121/1.1909553

Mironovs, M., & Lee, H. (2018, May). On the accuracy and consistency of sound localization at various azimuth and elevation angles. In 144th Audio Engineering Society International Pro Audio Convention (pp. Convention-Paper).

Mistretta, C.M. (1984). Aging effects on anatomy and neurophysiology of taste and smell. Gerodontology 3:131-136. doi:10.1111/j.1741-2358.1984.tb00364.x

Molloy, K., Moore, D.R., Sohoglu, E., & Amitay, S. (2012) Less is more: latent learning is maximized by shorter training sessions in auditory perceptual learning. PLOS One 7. doi:10.1371/journal.pone.0036929

Montagne, C., & Zhou, Y. (2018). Audiovisual interactions in front and rear space. Frontiers in Psychology, 9, 713. doi: 10.3389/fpsyg.2018.00713

Moore, B. C. J., Oldfield, S. R., & Dooley, G. J. (1989). Detection and discrimination of spectral peaks and notches at 1 and 8 kHz. J. Acoust. Soc. Am., 85(2), 820–836. https://doi.org/10.1121/1.397554

Mueller, M.F., Kegel, A., Schimmel, S. M., Dillier, N., & Hofbauer, M. (2012). Localization of virtual sound sources with bilateral hearing aids in realistic acoustical scenes. J. Acoust. Soc. Am., 131(6), 4732-4742. doi: 10.1121/1.4705292

Musicant, A. D., & Butler, R. A. (1984). The influence of pinnae-based spectral cues on sound localization. J. Acoust. Soc. Am., *75*(4), 1195-1200.

Nakamura, F., & Iida, K. (2025). Cue for rear sound image localization in head-related transfer function below 4 kHz. Applied Acoustics, 229, 110370.

Nakashima, A., Abel, S.M., & Smith, I. (2018). Communication in military environments: influence of noise, hearing protection and language proficiency. Applied Acoustics 131:38-44. https://doi.org/10.1016/j.apacoust.2017.10.010

Nelken, I., & Young, E. D. (1994). Two separate inhibitory mechanisms shape the responses of dorsal cochlear nucleus type IV units to narrowband and wideband stimuli. Journal of Neurophysiology, 71(6), 2446–2462. https://doi.org/10.1152/jn.1994.71.6.2446

Nelson DI., Nelson RY., Concha-Barrientos M. & Fingerhut M. (2005). The global burden of occupational noise-induced hearing loss. American Journal of Industrial Medicine, 48(6), 446–458. https://doi.org/10.1002/ajim.20223

NIDCD (2019). Noise-induced hearing loss. National Institutes of Health 14-4233. https://www.nidcd.nih.gov/health/noise-induced-hearing-loss

Noble, W.G., Russell, G. (1972). Theoretical and practical implications of the effects of hearing protection devices on localization ability. Acta Otolaryngologica 74:29-36. doi: 10.3109/00016487209128419

Noble, W.G. (1981). Earmuffs, exploratory head movements, and horizontal and vertical sound localization. The Journal of Auditory Research 21:1-12.

Noble, W., Murray, N., & Waugh, R. (1990). The effect of various hearing protectors on sound localization in the horizontal and vertical planes. American Industrial Hygiene Association Journal 51:370-377. doi:10.1080/15298669091369808

Otte, R.J., Agterberg, M.J., Van Wanrooij, M.M., Snik, A.F., & Van Opstal, A.J. (2013). Age-related hearing loss and ear morphology affect vertical but not horizontal sound-localization performance. J. Assoc. Res. Otolaryngol.: JARO 14:261–273. https://doi.org/10.1007/s10162-012-0367-7

Park MY. & Casali JG. (1991). A Controlled Investigation of In-Field Attenuation Performance of Selected Insert, Earmuff, and Canal Cap Hearing Protectors. Human Factors, 33(6), 693–714. https://doi.org/10.1177/001872089103300606

Petrov, A.A., Dosher B.A., Lu, Z. (2006). Perceptual learning without feedback in non-stationary contexts: Data and model. Vision Research 46:3177-3197.

Perrott, D. R., & Saberi, K. (1990). Minimum audible angle thresholds for sources varying in both elevation and azimuth. J. Acoust. Soc. Am., 87(4), 1728–1731. https://doi.org/10.1121/1.399421

Poirier-Quinot, D., & Katz, B. F. G. (2021). On the improvement of accommodation to non-individual HRTFs via VR active learning and inclusion of a 3D room response. Acta Acustica, 5, 25. https://doi.org/10.1051/aacus/2021019

Poirier-Quinot, D., S. Lawless, M., Stitt, P., & F.G. Katz, B. (2022). HRTF Performance Evaluation: Methodology and Metrics for Localisation Accuracy and Learning Assessment. In B. F.G. Katz & P. Majdak (Eds.), Advances in Fundamental and Applied Research on Spatial Audio. IntechOpen. https://doi.org/10.5772/intechopen.104931

Protection of Environment, 40 C.F.R § 211.207 (Subpart B) (1979). https://www.ecfr.gov/current/title-10/chapter-III/part-745

Rajendran, V. G., & Gamper, H. (2019). Spectral manipulation improves elevation perception with non-individualized head-related transfer functions. J. Acoust. Soc. Am., 145(3), EL222–EL228. https://doi.org/10.1121/1.5093641

Rakerd, B., Vander Velde, T. J., & Hartmann, W. M. (1998). Sound localization in the median sagittal plane by listeners with presbyacusis. Journal of the American Academy of Audiology, 9(6), 466–479.

Raykar, V. C., Duraiswami, R., & Yegnanarayana, B. (2005). Extracting the frequencies of the pinna spectral notches in measured head related impulse responses. J. Acoust. Soc. Am., 118(1), 364–374. https://doi.org/10.1121/1.1923368

Rauschecker, J.P. (1999). Auditory cortical plasticity: A comparison with other sensory systems. Trends in Neurosciences 22:74-80. doi:10.1016/s0166-2236(98)01303-4

Rayleigh, Lord. (1907). On our perception of sound direction. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 13(74), 214–232. https://doi.org/10.1080/14786440709463595

Reddy, R., Thorne, P., Welch, D., Ameratunga, S. (2012). Hearing protection use in manufacturing workers: A qualitative study. Noise and Health 14:202. doi:10.4103/1463-1741.99896

Rice, J. J., May, B. J., Spirou, G. A., & Young, E. D. (1992). Pinna-based spectral cues for sound localization in cat. Hearing Research, 58(2), 132–152. https://doi.org/10.1016/0378-5955(92)90123-5

Röder, B., Teder-S ä lej ä rvi, W., Sterr, A., Rösler, F., Hillyard, S.A., & Neville, H.J. (1999). Improved auditory spatial tuning in blind humans. Nature 400:162-166. https://doi.org/10.1038/22106

Russell, G. (1976). Practice Effects for Auditory Localization: A Test of a Differentiation Theory of Perceptual Learning and Development. Perceptual and Motor Skills, 43(2), 647–653. https://doi.org/10.2466/pms.1976.43.2.647

Sabin, A. T., Macpherson, E. A., & Middlebrooks, J. C. (2005). Human sound localization at near-threshold levels. Hearing Research, 199(1–2), 124–134. https://doi.org/10.1016/j.heares.2004.08.001

Simpson, B.D., Bolia, R. S., McKinley, R. L., & Brungart, D.S. (2005). The impact of hearing protection on sound localization and orienting behavior. Human Factors: The Journal of the Human Factors and Ergonomics Society 47:188-198. doi:10.1518/0018720053653866

Slaney M. (1993). An efficient implementation of the Patterson-Holdsworth auditory filter bank. Apple Computer, Perception Group, Tech. Rep, 35(8). https://engineering.purdue.edu/~malcolm/apple/tr35/PattersonsEar.pdf

Smalt, C. J., Calamia, P. T., Dumas, A. P., Perricone, J. P., Patel, T., Bobrow, J., Quatieri, T. F. (2019). The effect of hearing-protection devices on auditory situational awareness and listening effort. Ear & Hearing, 41:82-94. doi:10.1097/aud.0000000000000733

Stevens, S. S., & Newman, E. B. (1936). The Localization of Actual Sources of Sound. The American Journal of Psychology, 48(2), 297. https://doi.org/10.2307/1415748

Stitt, P., & Katz, B. F. G. (2021). Sensitivity analysis of pinna morphology on head-related transfer functions simulated via a parametric pinna model. J. Acoust. Soc. Am., 149(4), 2559–2572. https://doi.org/10.1121/10.0004128

Stitt, P., Picinali, L., & Katz, B. F. G. (2019). Auditory Accommodation to Poorly Matched Non-Individual Spectral Localization Cues Through Active Learning. Scientific Reports, 9(1), 1063. https://doi.org/10.1038/s41598-018-37873-0

Strasburger, H., Rentschler, I., & Juttner, M. (2011). Peripheral vision and pattern recognition: A review. Journal of Vision, 11:13-13. doi:10.1167/11.5.13

Strybel, T. Z., & Fujimoto, K. (2000). Minimum audible angles in the horizontal and vertical planes: Effects of stimulus onset asynchrony and burst duration. J. Acoust. Soc. Am., 108(6), 3092–3095. https://doi.org/10.1121/1.1323720

Syka, J. (2002). Plastic changes in the central auditory system after hearing loss, restoration of function, and during learning. Physiological Reviews 82:601-636. doi:10.1152/physrev.00002.2002

Talagala, D. S., Zhang, W., Abhayapala, T. D., & Kamineni, A. (2014). Binaural sound source localization using the frequency diversity of the head-related transfer function. J. Acoust. Soc. Am., 135(3), 1207–1217. https://doi.org/10.1121/1.4864304

Talcott, K.A., Casali, J.G., Keady, J.P., & Killion, M.C. (2012). Azimuthal auditory localization of gunshots in a realistic field environment: effects of open-ear versus hearing protection-enhancement devices (HPEDs), military vehicle noise, and hearing impairment. International Journal of Audiology 51(Sup1):20–30. https://doi.org/10.3109/14992027.2011.631591

Trapeau, R., Aubrais, V., & Schönwiesner, M. (2016). Fast and persistent adaptation to new spectral cues for sound localization suggests a many-to-one mapping mechanism. J. Acoust. Soc. Am., 140(2), 879–890. https://doi.org/10.1121/1.4960568

Trapeau, R., & Schönwiesner, M. (2018). The Encoding of Sound Source Elevation in the Human Auditory Cortex. Journal of Neuroscience, 38(13), 3252–3264.

Treeby, B.E., Pan, J., & Paurobally, R.M. (2007). The effect of hair on auditory localization cues. J. Acoust. Soc. Am. 122:3586-3597. doi:10.1121/1.2793607

Van den Bogaert, T., Klasen, T.J., Moonen, M., Van Deun, L., & Wouters, J. (2006). Horizontal localization with bilateral hearing aids: Without is better than with. J. Acoust. Soc. Am., 119(1), 515-526. doi: 10.1121/1.2139653

Van Opstal, A. J., Vliegen, J., & Van Esch, T. (2017). Reconstructing spectral cues for sound localization from responses to rippled noise stimuli. PLOS ONE, 12(3), e0174185. https://doi.org/10.1371/journal.pone.0174185

Van Wanrooij, M. M., & Van Opstal, A. J. (2005). Relearning Sound Localization with a New Ear. J. Neurosci., 25(22), 5413–5424. https://doi.org/10.1523/JNEUROSCI.0850-05.2005

Vause, N.L., & Grantham, D.W. (1999). Effects of earplugs and protective headgear on auditory localization ability in the horizontal plane. Human factors, 41(2), 282-294. doi: 10.1518/001872099779591213

Vliegen, J., & Van Opstal, A. J. (2004). The influence of duration and level on human sound localization. J. Acoust. Soc. Am., 115(4), 1705–1713. https://doi.org/10.1121/1.1687423

Voigt, H. F., & Young, E. D. (1980). Evidence of inhibitory interactions between neurons in dorsal cochlear nucleus. Journal of Neurophysiology, 44(1), 76–96. https://doi.org/10.1152/jn.1980.44.1.76

Wang, M., Ai, Y., Han, Y., Fan, Z., Shi, P., & Wang, H. (2021). Extended high-frequency audiometry in healthy adults with different age groups. Journal of Otolaryngology-Head & Neck Surgery, 50, 1-6. doi: 10.1186/s40463-021-00534-w

Watanabe, T., Nanez, J.E., Sasaki, Y. (2001). Perceptual learning without perception. Nature 413:844-848.

Watson, C.J.G., Carlile, S., Kelly, H., & Balachander, K. (2017). The generalization of auditory accommodation to altered spectral cues. Scientific Reports 7. https://doi.org/10.1038/s41598-017-11981-9

Wenzel, E.M., Arruda, M., Kistler, D.J., & Wightman, F.L. (1993). Localization using nonindividualized head-related transfer functions. J. Acoust. Soc. Am. 94:111-123. doi:10.1121/1.407089

Wightman, F. L., & Kistler, D. J. (1989). Headphone simulation of free-field listening. II: Psychophysical validation. J. Acoust. Soc. Am., 85(2), 868–878. https://doi.org/10.1121/1.397558

Wightman, F.L., Kistler, D.J. (1999). Resolution of front-back ambiguity in spatial hearing by listener and source movement. J. Acoust. Soc. Am. 105:2841-2853. doi: 10.1121/1.426899.

Wong, L.L., Hickson, L., & McPherson, B. (2003). Hearing aid satisfaction: what does research from the past 20 years say?. Trends in amplification, 7(4), 117-161. doi: 10.1177/108471380300700402

Wright, B.A., Sabin, A.T. (2007). Perceptual learning: how much daily training is enough? Experimental Brain Research 180:727–736. https://doi.org/10.1007/s00221-007-0898-z

Wright, B.A., Zhang, Y. (2006). A review of learning with normal and altered sound-localization cues in human adults. International Journal of Audiology 45(Sup1):92-98. doi:10.1080/14992020600783004

Yost, W. A. (2023). Randomizing spectral cues used to resolve front-back reversals in sound-source localization. J. Acoust. Soc. Am., 154(2), 661–670. https://doi.org/10.1121/10.0020563

Zahorik, P., Bangayan, P., Sundareswaran, V., Wang, K., & Tam, C. (2006). Perceptual recalibration in human sound localization: Learning to remediate front-back reversals. J. Acoust. Soc. Am., 120(1), 343–359. https://doi.org/10.1121/1.2208429

Zhang, P. X., & Hartmann, W. M. (2010). On the ability of human listeners to distinguish between front and back. Hearing Research, 260(1–2), 30–46. https://doi.org/10.1016/j.heares.2009.11.001

Zheng, Y., Swanson, J., Koehnke, J., & Guan, J. (2022). Sound localization of listeners with normal hearing, impaired hearing, hearing aids, bone-anchored hearing instruments, and cochlear implants: a review. American Journal of Audiology, 31(3), 819-834. doi: 10.1044/2022_AJA-22-00006

Zimpfer, V., Sarafian, D. (2014). Impact of hearing protection devices on sound localization performance. Frontiers in Neuroscience 8. doi:10.3389/fnins.2014.00135

Zonooz, B., Arani, E., Körding, K. P., Aalbers, P. A. T. R., Celikel, T., & Van Opstal, A. J. (2019). Spectral Weighting Underlies Perceived Sound Elevation. Scientific Reports, 9(1), 1642. https://doi.org/10.1038/s41598-018-37537-z

Zonooz, B., Arani, E., Van Opstal, A.J. (2018). Learning to localize weakly-informative sound spectra with and without feedback. Scientific Reports 8. doi:10.1038/s41598-018-36422-z

Zonooz, B., & Van Opstal, A. J. (2019). Differential Adaptation in Azimuth and Elevation to Acute Monaural Spatial Hearing after Training with Visual Feedback. Eneuro, 6(6), ENEURO.0219-19.2019. https://doi.org/10.1523/ENEURO.0219-19.2019

Zou H. & Hastie T. (2005). Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society Series B: Statistical Methodology, 67(2), 301–320.

Zwiers, M. P., Van Opstal, A. J., & Paige, G. D. (2003). Plasticity in human sound localization induced by compressed spatial vision. Nature Neuroscience, 6(2), 175–181. https://doi.org/10.1038/nn999