Pöntynen, Henri; Salminen, Nelli

Resolving front-back ambiguity with head rotation: the role of level dynamics

Published in:
Hearing Research

DOI:
10.1016/j.heares.2019.03.020

Published: 01/06/2019

Document Version
Peer-reviewed accepted author manuscript, also known as Final accepted manuscript or Post-print

Please cite the original version:
Pöntynen, H., & Salminen, N. (2019). Resolving front-back ambiguity with head rotation: the role of level dynamics. *Hearing Research*, *377*, 196-207. https://doi.org/10.1016/j.heares.2019.03.020

Accepted Manuscript

Resolving front-back ambiguity with head rotation: the role of level dynamics

Henri Pöntynen, Nelli Salminen

PII: S0378-5955(18)30555-0

DOI: https://doi.org/10.1016/j.heares.2019.03.020

Reference: HEARES 7728

To appear in: *Hearing Research*

Received Date: 29 November 2018

Revised Date: 25 March 2019

Accepted Date: 27 March 2019

- Head rotations contribute to the perception of sound source location
- Movement resolves front-back ambiguities in free field and virtual auditory space
- Dynamic ITD is a robust front-back localization cue in both environments
- Acoustic-domain phenomena confound ILD dynamics and hinder free field localization
- Simplified ILD dynamics facilitate front-back perception in virtual auditory space

Resolving front-back ambiguity with head rotation: the role of level dynamics

Henri Pöntynen[a,*], Nelli Salminen[a]

[a] *Aalto Acoustics Lab, Department of Signal Processing and Acoustics, School of Electrical Engineering, Aalto University, FI-02150 Espoo, Finland*

Abstract

Making small head movements facilitates spatial hearing by resolving front-back confusions, otherwise common in free field sound source localization. The changes in interaural time difference (ITD) in response to head rotation provide a robust front-back cue, but whether interaural level difference (ILD) can be used as a dynamic cue is not clear. Therefore, the purpose of the present study was to assess the usefulness of dynamic ILD as a localization cue. The results show that human listeners were capable of correctly indicating the front-back dimension of high-frequency sinusoids based on level dynamics in free field conditions, but only if a wide movement range was allowed (±40°). When the free field conditions were replaced by simplistic headphone stimulation, front-back responses were in agreement with the simulated source directions even with relatively small movement ranges (±5°), whenever monaural sound level and ILD changed monotonically in response to head rotation. In conclusion, human listeners can use level dynamics as a front-back localization cue when the dynamics are monotonic. However, in free field conditions and particularly for narrowband target signals, this is often not the case. Therefore, the primary limiting factor in the use of dynamic level cues resides in the acoustic domain behavior of the cue itself, rather than in potential processing limitations or strategies of the human auditory system.

Keywords: active localization, head rotation, front-back ambiguity, binaural cues, directional bands, acoustic bright spot,

1. Introduction

In studies of spatial hearing, listeners are typically instructed to stay still. Under such conditions, sound source localization is based on binaural cues - the interaural time and level differences (ITD and ILD) - and on spectral cues (Middlebrooks and Green, 1991; Blauert, 1997). While the static conditions are important for experimental control, they deprive the listeners of additional sources of spatial information that are available in real life. Motion allows listeners to take advantage of dynamic localization cues, i.e. the changes in acoustic input caused by the movement of the head with respect to the sound source. A role for self-motion in sound source localization was suggested more than eight decades ago (Young, 1931; Wilska, 1938; Wallach, 1939, 1940) and modern psychoacoustic studies have confirmed that head rotations can both enhance (Perrett and Noble, 1997b,a; Wightman and Kistler, 1999; Macpherson, 2011; Brimijoin et al., 2013) and bias (Leung et al., 2008; Cooper et al., 2008; Freeman et al., 2017) sound source localization. The deduction of sound source location from interaural dynamics is an inherently multimodal process, as an unambiguous interpretation of any ITD and ILD sequence always requires the combination of binaural information with another information stream, e.g. self-orientation information during observer movement (Wallach, 1939, 1940) or in the case of moving sources, information about the source trajectory (Wightman and Kistler, 1999).

Amongst the enhancements provided by head rotation, perhaps the most important are the dynamic binaural cues. Normally, when motion is not allowed, binaural cues provide location information only in one dimension (left-right, the interaural axis). As an important example, ITD and ILD do not reveal whether the sound source is in front of or behind the listener (Middlebrooks and Green, 1991; Blauert, 1997). However, when the head is rotated, the direction and rate of change in ITD and ILD depend on the front-back location and elevation of the source, respectively. This pro-

*Corresponding author:
Email address: henri.pontynen@aalto.fi (Henri Pöntynen)

vides the auditory system with an opportunity to resolve front-back confusions by combining the changing self-orientation information furnished by non-auditory sensory modalities with the accompanying changes in spatial cues (Wallach, 1939, 1940; Wightman and Kistler, 1999; Kim et al., 2013; Yost et al., 2015). Dynamic ITD, in particular, appears to be a robust front-back cue. When listeners are presented with stimuli devoid of ILD and spectral cues (e.g. low-pass filtered or low-frequency narrowband sounds), front-back confusions are rare when head rotations are allowed (Perrett and Noble, 1997b,a; Macpherson, 2011).

In contrast to the robust use of dynamic ITD, the usefulness of ILD as a dynamic cue is less clear. When noise stimuli are high-pass filtered, so that phase-locked fine-structure ITD is unavailable, head rotation can be used for solving front-back confusions for sounds presented from the median plane, even when spectral cues are distorted (Perrett and Noble, 1997b; Macpherson, 2013). Because under such conditions dynamic ILD is presumably the only front-back cue available, this finding suggests that at least in the absence of any spectral information and for sources positioned at the midline, dynamic wideband ILD can be useful. However, in another recent study using narrowband stimuli presented from a wide range of horizontal locations, listeners were unable to resolve front-back confusions with head-rotations for high-frequency stimuli in which ILD is the prominent binaural cue (Macpherson, 2011). Yet, under the same conditions low-frequency stimuli were correctly localized based on dynamic ITD. Instead of dynamic ILD, the perceived location of the high-frequency stimuli appeared to be determined by directional band biases: the illusion of sound source elevation induced by narrowband stimuli (Blauert, 1969; Middlebrooks, 1992; Itoh et al., 2007; Thakkar and Goupell, 2014).

Further evidence for the relative weakness of ILD as a dynamic cue comes from studies in which spatial sound presentation is manipulated to introduce cue conflicts. When a conflict is introduced between the spatial information yielded by spectral cues and dynamic ILD (for instance, spectral cues corresponding to a location in the front hemiplane and dynamic cues corresponding to a source in the rear hemiplane), sound sources are localized based on spectral cues and dynamic ILD appears to be ignored (Pöntynen et al., 2016). Also, when a similar manipulation is performed on a wideband sound so that dynamic ITD and ILD are in conflict with spectral cues, location perception may become unstable (Brimijoin and Akeroyd, 2012) or the stimulus may even be perceived as two separate auditory images, with one low-frequency image corresponding to dynamic ITD cues and the other to spectral cues (Pöntynen et al., 2016). This implies that spectral cues do not necessarily dominate dynamic ITD but they might do for dynamic ILD under some conditions. In sum, previous research suggests that dynamic ILD is a relatively weak cue over which spectral cues, and even narrow-band biases, may dominate.

A potential reason for the human auditory system to prefer dynamic ITD over dynamic ILD resides in the differences in the acoustic domain phenomena between low- and high-frequency sound waves. In particular, in the frequency range where fine-structure ITD is a relevant localization cue, the dimensions of the human head are small compared to the wavelength of sound. This dimension mismatch allows sound to diffract around the head, leading to less perturbation of the sound field. Consequently, ITD changes consistently across azimuth and functions as a robust localization cue as long as the wavelength of the incident wave is long enough to result in an unambiguous phase relationship between the ears. Conversely, at high frequencies, where ILD is the dominant binaural cue, the wavelength of sound is small relative to the dimensions of the head. This causes incident waves to reflect and scatter at the air-head boundary rather than to simply diffract around it unaffected. Due to the idiosyncratic and anthropometry-dependent details of the high-frequency interaction between the head and incident sound waves, the narrow-band sound pressure level at each ear - and consequently narrow-band ILD - depends on the angle of incidence in a non-monotonic manner (Blauert, 1997; Middlebrooks et al., 1989; Kuhn, 1987; Macaulay et al., 2010) (see Sec. 2.5 for further discussion on diffraction around a rigid sphere and its relation to ITD and ILD). This non-monotonic relationship between narrow-band ILD and source azimuth confounds static localization, as a single ILD value could arise from multiple azimuthal directions (Macaulay et al., 2010). Likewise, non-monotonicity of ILD is expected to confound dynamic localization because the direction of ILD change is not consistent during large head rotations. These idiosyncrasies may have led the human auditory system to give dominance to spectral cues for elevation and front-back localization of high-frequency sounds. Alternatively, the use of dynamic ILD could be location-dependent so that it influences perceived location whenever it provides consistent information but is ignored when non-monotonic dynamics occur. Currently, these questions about the use of dynamic ILD in sound source localization remain unexplored.

The purpose of the present study was to assess the ability of moving listeners to use dynamic ILD as a localization cue. Due to the limited evidence on dynamic ILD perception, we started by confirming that human listeners are able to solve front-back ambiguities based on dynamic ILD, that is, by rotating the head while localizing a sound void of any other front-back cues. We further assessed the effects of the range of movement and sound source location. Finally, we aimed to understand how partially non-monotonic and location-dependent level dynamics affect the use of dynamic ILD as a localization cue.

2. Methots

2.1. Overview

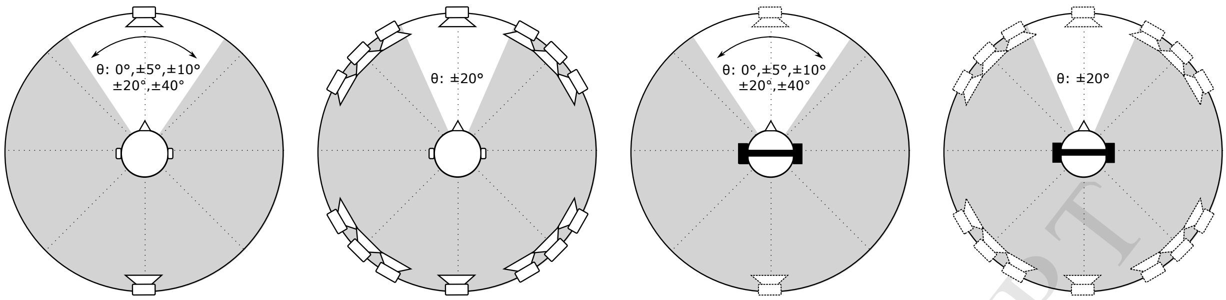
In the present experiments, human listeners participated in a front-back localization task that could be performed accurately only by relying on dynamic binaural cues resulting from head rotation. In order to control the head rotation range available to the listeners, we used a head-orientation-coupled stimulus gating paradigm similar to that used in several earlier studies on dynamic localization (Macpherson and Kerr, 2008; Macpherson, 2011, 2013, 2014) (see Sec. 2.6 for details). Borrowing from the terminology used in these studies, we use the term "movement window" to refer to the head rotation range available to the listeners in each experimental condition. We used sinusoidal stimuli that did not provide informative spectral cues, but instead were expected to induce biased perception of front, rear and elevated locations (Blauert, 1969; Middlebrooks, 1992; Itoh et al., 2007; Thakkar and Goupell, 2014). Therefore, for correct front-back localization, the dynamic cues needed to overcome the possibly contradicting biased perception induced by the sinusoidal narrow-band stimulus. Dynamic ILD perception was assessed with three stimulus frequencies, 2, 4, and 8 kHz. These give rise to ILDs in free field but are above the frequency range of fine-structure ITD perception (Klumpp and Eady, 1956; Zwislocki and Feldman, 1956; Brughera et al., 2013; Verschooten et al., 2019). In order to compare the perception of dynamic ILD to that of dynamic ITD, we included a 500 Hz stimulus as a control case which yielded a robust fine-structure ITD complemented with a relatively weak but monotonic ILD in free field conditions (Hartmann et al., 2016). Despite inducing a small ILD, we expected that free field localization of the low-frequency stimulus would be dominated by the robust ITD dynamics rather than the supplementary ILD dynamics.

Four experiments were conducted (Fig. 1), Experiments I and II in free field and Experiments III and IV in virtual auditory space (VAS) using highly simplified head-related transfer functions (HRTFs) derived from a spherical head model (see Fig. 2). The purpose of Experiment I was, first, to confirm that human listeners can detect dynamic ILD and use it for solving front-back ambiguities and, second, to assess whether performance comparable to dynamic ITD detection could be reached within any movement window. To this end, sounds were presented from the median plane (directly in front or behind the listener) where ILD sensitivity is at its best, and ILD dynamics are monotonic. Further, relatively large movement windows were included to aid dynamic localization. The purpose of Experiment II was to evaluate the dependence of dynamic ILD localization on sound source laterality. We expected differences to occur because of the location-dependence of ILD dynamics and possibly also because of a decline in ILD sensitivity for increasingly lateral sound sources. Therefore, the free field front-back localization task was repeated for sound sources at various lateral angles.

Under free field conditions, the exact ILD dynamics resulting from head rotations are unknown to the experimenter unless a large set of HRTFs is collected for each subject. Here this presents a particular challenge for assessing the role of idiosyncratic level dynamics. The purpose of Experiments III and IV was therefore to assess front-back localization with dynamic ILD under more controlled conditions. These experiments repeated Experiments I and II with simplistic VAS stimulation in which plausible cue dynamics for ITD and ILD were derived from a spherical head model. This stimulation, though unrealistic and distinct from free-field stimulation due to the absence of pinna effects, has the benefit of generating ILD dynamics that include both monotonic and non-monotonic behavior and allowing specific knowledge of the ILDs that the listeners experience. Further, VAS stimulation allowed complete decoupling of ILD and ITD dynamics by holding the confounding binaural cue (ILD for low-frequency control stimuli and ITD for high-frequency stimuli) at a value of zero and updating only the binaural cue of interest according to instantaneous head orientation.

2.2. Subjects

A total of 23 normal hearing volunteers (3 females, 20 males, mean age: 30.3 years, sd: 7.1 years) from the Department of Signal Processing and Acoustics at Aalto University participated in the experiments; the authors did not participate. On average, each subject participated in two (mean: 1.8) experiments. No subject participated in all four experiments. The experiment-wise



(a) Exp. I (b) Exp. II (c) Exp. III (d) Exp. IV

Figure 1: Schematic depictions of the experimental setups. In Experiments I and II, stimuli were presented from loudspeakers and in Experiments III and IV through headphones using spherical HRTFs. Consequently, the loudspeaker symbols denote real sources for Exps. I & II and virtual sources for Exps. III & IV. In Experiments I and III, the sound source was at 0° or 180° and the movement window varied from 0° to ±40°. In Experiments II and IV, stimuli were presented from various lateral angles and the movement window was fixed at ±20°. The size of the loudspeakers is not drawn to scale.

age and gender distributions were as follows: Experiment I: 10 subjects (1 female, 9 male, mean age: 31.0 years, sd: 4.7 years), Experiment II: 11 subjects (1 female, 10 male, mean age: 28.3 years, sd: 8.8 years), Experiment III: 10 subjects (2 female, 8 male, mean age: 32.3 years, sd: 7.2 years), Experiment IV: 11 subjects (all male, mean age: 27.2 years, sd: 6.1 years). The experimental procedures were approved by the Aalto University Research Ethics Committee.

2.3. Apparatus and facilities

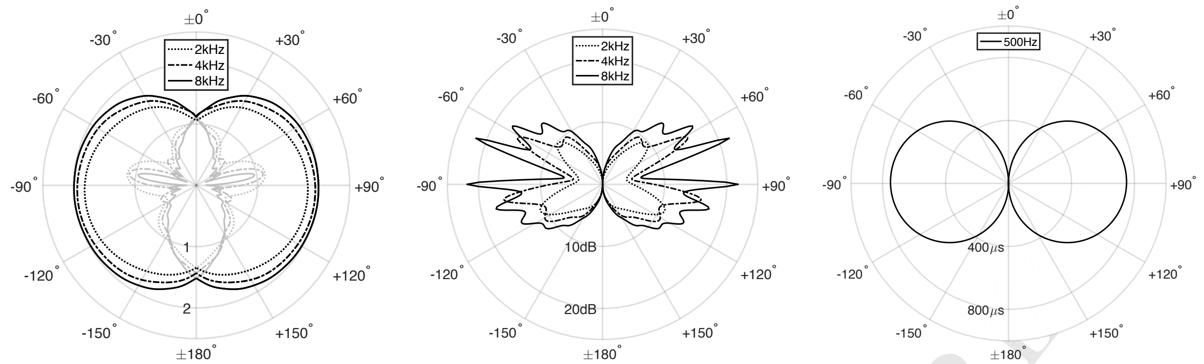
Experiments were conducted in an anechoic chamber fitted with a multichannel loudspeaker system and an infrared camera (OptiTrack Flex 3) motion-tracking system. Subjects wore a headband fitted with reflective markers, whose position and orientation was monitored by the tracking system at a rate of 100 Hz. The loudspeakers were attached to a circular hoop of approximately 2 m radius surrounding the listening position at the center of the chamber. Subjects used a tablet computer to report their responses.

Audio was processed in buffers of 32 samples at a sample rate of 48 kHz and sent to an RME M-32 digital-to-analog converter via an RME MADIface XT audio interface. Depending on the experiment, the analog signals were then sent either to Genelec 8030A active loudspeakers (Experiments I and II) or to Sennheiser HD 600 headphones via a Sound Devices HX-3 headphone amplifier (Experiments III and IV). The average movement-to-stimulus latency varied across experiments but remained below 30 ms in all conditions.

2.4. Stimuli

In order to isolate the influence of ITD and ILD dynamics in active localization, the present experiments made use of sinusoidal stimuli of frequencies 0.5, 2, 4 and 8 kHz. The use of sinusoidal stimuli ensured that informative spectral localization cues were not available and that correct localization in the front-back dimension had to rely on the dynamics in the binaural cue that was dominant at the stimulus frequency. The 0.5 kHz tone acted as a reference stimulus that provided a robust ITD cue and relatively weak level dynamics in free field. Conversely, the high-frequency stimuli were within the frequency range where fine-structure ITD is not available as an effective localization cue (e.g., Klumpp and Eady 1956; Zwislocki and Feldman 1956; Brughera et al. 2013). For the sake of simplicity, we use the term "ITD stimuli" to refer to the 0.5 kHz low-frequency control stimuli and "ILD stimuli" to refer to the high-frequency stimuli at 2, 4 and 8 kHz, despite the fact that under free field conditions, all of these stimuli provide both ITD and ILD cues to some extent. In the VAS experiments (III & IV) however, decoupling of the effects of ILD and ITD dynamics was achieved by tracking only one cue modality for each stimulus. In the case of ILD stimulus trials, the level dynamics were tracked while the ITD value was static at 0 s. Conversely, in ITD stimulus trials only the interaural delay was tracked and the ILD was static at 0 dB.

Stimuli were manipulated according to headtracker data in real time. In each trial, subjects could listen to the stimulus as long as they needed before giving their response. All stimuli were initiated and terminated with 100 ms amplitude ramps and presented on the horizontal plane from azimuth angles 0° and 180° in Experiments I and III, and from angles 0°, ±30°, ±45°, ±60°, ±120°, ±135°, ±150°, and 180° in Experiments II and IV (see Fig. 1). In free field experiments, each stimulus was cal-



(a) Pressure response (b) ILD (c) ITD

Figure 2: Responses of the spherical head model. (a) Pressure responses (relative to free field) of the head model obtained by representing the head as a rigid sphere with ears at ±100°. The ipsilateral responses are plotted in black and the contralateral responses are plotted in grey. (b) ILD responses of the spherical head model. (c) ITD response of the spherical head model at 500Hz. Binaural cue magnitudes are plotted relative to the ipsilateral ear at each azimuth.

ibrated to yield an A-weighted sound pressure level of 60 dBA at the listening position. In VAS experiments, all stimuli yielded 60 dBA from both earphones when the virtual source was positioned at zero azimuth. To reduce the usefulness of possible front-back cues arising from systematic level differences across source positions, ±3 dB randomized level roving was applied between trials (Stevens and Newman, 1936).

2.5. Headphone presentation and the spherical head model

To produce simplistic but plausible, subject-independent stimuli for headphone presentation, ILD dynamics were derived from the acoustic transfer function of a rigid sphere, as described in Duda and Martens (1998). The pressure responses were evaluated for a sphere with a radius of 8.75 cm, speed of sound corresponding to average ambient conditions of 343 m/s and a normalized source distance of 100 sphere radii. The pressure responses are shown in Fig. 2 (a) and expressed relative to conditions where the sphere is absent, i.e. values below 1 denote attenuation and those above 1 denote a pressure increase relative to free field. The sphere responses are characterized by three major phenomena. First, the sphere introduces frequency-dependent acoustic shadowing in the contralateral hemisphere. Second, as the frequency increases, the pressure at the ipsilateral hemisphere increases due to a portion of the wave being reflected back at the air-sphere boundary. The magnitude of this increase approaches a doubling in pressure as the dimensions of the sphere become large compared to the wavelength of the incident wave. Third, the interference of multiple wave components traveling around the sphere results in increasingly dense rippling in the sphere's contralateral pressure response (see Fig. 2a) as the frequency of the incoming wave increases. In particular, at the polar angle opposing the angle of incidence of the incoming wave, multiple components of the diffracted wave are recombined in-phase, leading to the formation of the so-called acoustic "bright spot" that has a significantly higher pressure level than the rest of the contralateral hemisphere (Rabinowitz et al., 1993; Duda and Martens, 1998; Xie, 2013; Macaulay et al., 2010).

The headphone stimuli in Experiments III & IV were based on evaluating the pressure responses at two ears located at ±100° on the surface of the sphere; the azimuth-dependence of the responses (normalized with respect to free field conditions) at the two ears at each stimulus frequency is shown in Fig. 2 (a). The model ILDs obtained as the magnitude ratio of the ipsilateraral and contralateral ear responses are shown in Fig. 2 (b). Qualitatively, this figure shows that while the ILD varies smoothly near the median plane, at lateral azimuths the magnitude ripples and the bright spot at the contralateral ear lead to non-monotonic ILD dynamics that are expected to confound dynamic localization. The ITD associated with the 500 Hz control stimulus was obtained from the phase difference between the complex-valued responses at the two ears as described in Aaronson and Hartmann (2014); the resultant ITD-response is shown in Fig. 2 (c).

The binaural stimuli were implemented by evaluating the scattering equation with the previously declared parameters to a resolution of $1 * 10^{-4}$ at 0.1° azimuthal increments. The resulting gain and delay values were tabulated and used to impose the appropriate binaural differences on the headphone signals according to the instantaneous head orientation information provided by the tracking apparatus. The frequency response of the headphones was accounted for by assigning separate frequency-dependent gain values for each stimulus in the digital domain, so that all stimuli yielded the same sound pressure level from both earphones.

2.6. Limiting the head movement range

The permitted movement ranges (movement windows) were ±5°, ±10°, ±20°, and ±40° in Experiments I and III, and ±20° in Experiments II and IV (see Fig. 1). To control the head rotation range available to the subjects, the target stimuli were gated in a head-orientation-coupled manner (Macpherson, 2013). The gating was implemented by ramping down the stimulus within a 10° angle window beyond the limits of the chosen movement window. In free field experiments this was accompanied with complementary ramping of a semi-diffuse masker consisting of uncorrelated white noise presented from 10 directions within the horizontal plane (0°/180°, ±30° ±60°, ±120°, ±150°). The level of the masker reached 70 dBA if the full 10° transition region was surpassed. Similarly, in headphone experiments both earphone signals were ramped down within a 10° angle window and accompanied with complementarily ramped, uncorrelated white noise maskers presented from both earphones that reached 70 dBA beyond the 10° transition region.

2.7. Experimental scenario and test procedure

During the experiments, the subjects were seated in the middle of a circular array of real or virtual sound sources as shown in Fig. 1. The subjects reported the perceived front-back position of the target stimuli in a single-interval, two-alternative forced-choice task. Individual trials proceeded as follows: to initiate a new trial via the user interface subjects had to be oriented between ±5° azimuth. Once the target stimulus was being presented, subjects could report their answers (”front” or ”back”) only after they had traversed the entire extent of the assigned movement window at least once. Each combination of experimental parameters: stimulus frequency, movement window and front-back location (Experiments I and III) or stimulus frequency and stimulus position (Experiments II and IV) was tested four times in a fully randomized order. As an exception to this, the static condition (0° movement range) in Experiments I and III was performed as a separate stimulation block at the beginning of the experiment. These trials were separated because of the distinct trial structure that involved no head rotation. The order of these trials with respect to the rest of the experiment was not expected to induce any systematic bias in performance, as in the absence of head rotations these stimuli result in chance performance. As in the trials involving head rotation, stimulus duration was unlimited in the static condition trials and subjects could listen to the stimuli as long as was needed to make a front/rear judgment.

Prior to the experimental sessions, all subjects participated in a short familiarization session, where they learned to use the tablet computer response interface and confirmed that they understood the head movement requirements imposed by the experimental design (i.e. the response interface did not allow responses to be reported until the subjects had traversed the entire extent of the set movement window at least once); the number of practice trials subjects undertook during the familiarization sessions was typically less than five. The subjects received no instructions for potential listening strategies nor information regarding the actual (or simulated) sound source location, either during the familiarization sessions or during the experimental sessions.

The single-interval, two-alternative front-back localization task used here (and for instance by Macpherson, 2014) is rather simplistic compared to the commonly used head-pointing (for instance, Perrett and Noble, 1997b; Macpherson, 2011) or verbal responses of perceived azimuth and elevation (for instance, Wallach, 1940; Wightman and Kistler, 1999). The simple front-back localization task provided the significant benefit that naive listeners could perform the task reliably (as shown by the high level of performance reached with the low-frequency control stimulus) after the short familiarization. This made it feasible to conduct the experiments on a relatively large number of untrained participants. Further, the additional information that alternative localization tasks would have provided was not directly relevant to the present aim of understanding how front-back ambiguities are resolved.

2.8. Presentation and analysis of data

For the main analyses of the data, the results were pooled across equivalent angular positions from the front and rear hemiplanes (e.g., 0° & 180° or ±45° & ±135°) and the percentage of correct responses (i.e. responses consistent with the actual source location in free-field experiments and responses consistent with the simulated source location in VAS experiments) for each stimulus was computed, so that each subject contributed a single pseudo-continuous value for each test condition. Since each test condition was tested four times in each spatial location, pooling the results from equivalent angular positions in the front and rear hemiplanes as well as left-right dimensions yielded a total of eight repetitions per stimulus condition for midline sources (front and back results pooled) and 16 repetitions for the lateral source positions in Experiments II and IV (pooled across front-back and left-right dimensions). Furthermore, in order to observe possible directional band biases obscured by the pooling procedure in the main results, the results obtained from the front and back hemiplanes are also presented separately for illustrative purposes (middle and right-hand side panels in result plots). Error bars in figures represent the 95% bias-corrected and accelerated (BC_a) confidence intervals (Efron and Tibshirani, 1994) of the mean scores constructed using a bootstrap based on 1500 resamples from the subject means. This approach was used because the scores were not normally distributed.

For statistical testing, a non-parametric approach was adopted due to the data not fulfilling the assumptions on which parametric statistical tests rely. For evaluating the effect of movement window (in Experiments I and III), sound source azimuth angle (in Experiments II and IV), and stimulus frequency (all experiments), we used the Friedman test, a non-parametric omnibus test suitable for repeated-measures data. In cases where the omnibus test revealed a statistically significant effect ($p<.05$), pairwise comparisons were conducted using the exact Wilcoxon-Pratt signed-rank test (Hothorn et al., 2008; Pratt, 1959) and controlled for multiple comparisons within each set of comparisons using the Bonferroni procedure. Two types of post-hoc comparisons were performed. The first evaluated whether the performance score reached with the high-frequency stimuli was above that reached in the static condition (expected to be at chance, Experiments I and III) or above chance rate (Experiments II and IV). In the case of Experiments II and IV, this was done by subtracting the chance rate of 50% from the observed scores and running the Wilcoxon procedure on the resultant data sets (Hollander et al., 2013). The second post-hoc procedure tested whether performance for each ILD stimulus was lower than that for the ITD stimulus. For the post-hoc comparisons, we report the z-value as a supplementary statistic to the p-values rather than the Wilcoxon statistic, whose interpetation across conditions is more involved than that of the z-value due to the varying number of zeros and tied ranks between comparisons (Pratt, 1959; Rahe, 1974). However, the p-values are derived from the exact distribution of the test statistic rather than the supplementary z-value.

In addition to the front-back localization responses, head-tracker data were also collected. These data showed that the subjects spent more time in active listening for stimulus conditions yielding lower rates of correct responses and conversely performed only the minimum head-rotation required during the conditions yielding near-perfect performance. Because the head-tracking data did not provide additional insight and subject compliance was ensured by the experimental procedure, these data are not described further.

3. Experiment I: front-back localization of free field sources

3.1. Results

The purpose of Experiment I was to confirm that human listeners are able to use dynamic ILD as a localization cue, at least under favorable conditions, and to establish sufficient movement windows for resolving front-back confusions. Fig. 3 displays performance in the front-back localization task as a function of movement window. As expected, in the static condition, performance with all four stimulus frequencies was at chance. Head rotation improved performance even with the smallest movement range of ±5° and further improvements were found when the rotation range became wider. The effect of movement window on localization performance was statistically significant for all stimuli (Friedman tests: 500 Hz: $\chi^2(4) = 35.0, p<.001$, 2 kHz: $\chi^2(4) = 29.2, p<.001$, 4 kHz: $\chi^2(4) = 13.6, p = .009$, 8 kHz: $\chi^2(4) = 15.4, p = .004$).

The movement window sufficient for the dynamic cues to contribute to task performance was statistically evaluated by comparing performance level for each movement condition to that for the static condition. Because no other cues for solving front-back ambiguities were available, a significant result in this test shows that dynamic cues facilitated task performance. The test results are shown in Table 1. For all ILD stimulus frequencies, performance was above chance for the widest movement range. For the 2 kHz and 4 kHz stimuli, above chance performance was observed even for the narrower movement ranges (±5° & ±10°) but for the 8 kHz stimulus this occurred only beyond the ±10° range.

Overall, performance levels reached with the ILD stimuli were below those achieved with the ITD stimulus. This was confirmed by a statistically significant effect of stimulus frequency in the ±5° (Friedman tests: $\chi^2(3) = 9.69, p = .021$), ±10° ($\chi^2(3) = 17.6, p<.001$) and ±20° ($\chi^2(3) = 22.7, p<.001$) movement windows, but not in the 0° ($\chi^2(3) = .932, p = .818$) and ±40°

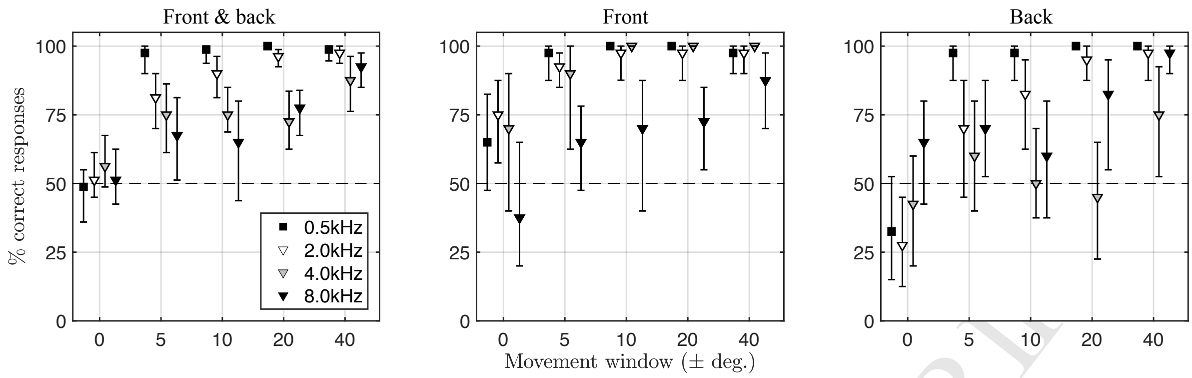


Figure 3: Results from Experiment I (free field sources at 0° and 180°). Percentage of correct responses (averaged over 10 participants) in the front-back localization task is plotted as a function of head-rotation window. Error bars represent 95% BC_a confidence intervals of the means. The first panel presents the pooled data for the main results and the second and third panels show supplementary results from the front and rear hemiplanes respectively.

Table 1: Results of the post-hoc comparisons for Experiment I. † denotes a comparison that was significant before the Bonferroni correction. Comparisons against ITD condition for ±40° window were not conducted due to a non-significant result from the corresponding Friedman test.

Frequency	Movement window	> Chance z	> Chance p*	< ITD z	< ITD p*
2 kHz	±5°	2.68	**.016**	2.56	**.023**
	±10°	2.82	**.004**	2.21	.094†
	±20°	2.87	**.004**	1.73	.375
	±40°	2.82	**.004**	—	—
4 kHz	±5°	2.11	.078†	2.40	**.047**
	±10°	2.51	**.027**	2.69	**.012**
	±20°	1.81	.188†	2.77	**.006**
	±40°	2.77	**.008**	—	—
8 kHz	±5°	1.33	.433	2.45	**.023**
	±10°	1.48	.281	2.67	**.012**
	±20°	2.38	**.047**	2.87	**.003**
	±40°	2.77	**.008**	—	—

($\chi^2(3) = 5.88, p = .118$) movement windows. This shows that performance was independent of stimulus frequency both in the absence of movement and with a wide movement range but frequency-dependent in intermediate movement ranges. Post-hoc comparisons (Table 1) tested whether the scores obtained for the ILD stimuli were significantly below those for the ITD stimuli. These failed to reach significance only for the 2 kHz stimulus within the ±10° and ±20° movement ranges. This signifies that in the intermediate movement windows, only the 2 kHz ILD stimulus led to performance comparable to that for the ITD condition while the higher ILD stimulus frequencies consistently led to performance below that for the ITD condition.

The responses to front and rear hemiplane stimuli are plotted separately in the middle and right panels of Fig. 3 to visualize the potential presence of directional band biases. For the stationary condition, there was a slight bias toward localizing the 2 and 4 kHz stimuli to the front. This bias appeared to be present also in movement conditions: while the 2 and 4 kHz front-hemiplane stimuli led to apparently excellent front-back localization performance even with the limited movement ranges of ±5° and ±10°, performance was much worse when these stimuli were presented from the rear hemiplane. These results are similar to the front bias observed for the stationary condition and therefore are consistent with a frequency-dependent directional bias affecting spatial perception of these stimuli within the smaller movement windows. When larger head rotations were allowed, the effect of these biases was weakened, as displayed by the similar performance for the front and rear hemiplanes with the 2 kHz stimuli within the ±20° and ±40° movement windows.

3.2. Discussion

Experiment I confirmed that ILD dynamics associated with head rotation can be used for resolving front-back confusions. However, the confusion rate decreased at a slower rate as the size of the movement window increased for the ILD than for the ITD stimuli. Particularly for the 4 and 8 kHz stimuli, front-back localization performance with dynamic ILD became comparable to that observed in the ITD condition only for the largest movement window of ±40°. In sum, level dynamics can effectively resolve hemiplane confusions, but only when a relatively wide movement range is allowed.

The narrowband biases found here for the stationary condition and for narrow movement windows are consistent with those reported in previous studies assessing directional band phenomena (Blauert, 1969; Middlebrooks, 1992; Itoh et al., 2007). The 2 and 4 kHz stimuli are often reported to be associated with sound sources in the frontal hemiplane and the 8 kHz center frequency with a sound source above the head. The current experiments only allowed front/back responses and consequently the responses to the 8 kHz stimulus were inconsistent. However, with larger movement windows the level dynamics dominated over the narrowband effects and appeared to be the primary determining factor for localizing the sound source to the front or back.

The present results show that level dynamics can function as a front-back hemiplane cue and overcome directional band biases, if the movement range is sufficiently wide. This differs from a previous study showing no evidence for the use of dynamic ILD in localizing high-frequency narrowband noise (6.0-6.5 kHz) (Macpherson, 2011). It should be noted that in comparison to previous studies the conditions of the present experiment were purposefully designed to be conducive to dynamic localization. Here, the subjects were free to perform horizontal head rotations within the permitted movement range in whatever way they found helpful, stimulus duration was unrestricted, and the sound source location was always at midline, coinciding both with the locations at which human perception of ILD changes is at its most accurate and for which level dynamics resulting from head rotation are relatively monotonic.

4. Experiment II: front-back localization of lateral free field sources

4.1. Results

Experiment II assessed the use of ILD dynamics as a front-back localization cue for lateral sound source locations in the free field. The lateral locations were outside the region of highest auditory spatial acuity and led to a slower rate of change in binaural differences in response to head rotation and presumably to more significant morphological effects (i.e. non-monotonic level dynamics) than for central locations due to e.g. the influence of the acoustic bright spot. The movement range of ±20° was chosen to be sufficient for using dynamic ILD (at least for the median plane locations based on Experiment I) while limiting the overlap of movement windows between source locations.

The results in Fig. 4 showed that task performance did not vary consistently across source locations, save for the fact that slightly higher performance was observed with the 2 kHz stimulus at the median plane than at lateral locations. Accordingly, the effect of sound source laterality was not statistically significant for any of the stimuli (Friedman tests: 500 Hz: $\chi^2(3) = 5.35, p = .148$, 2 kHz: $\chi^2(3) = 3.47, p = .325$, 4 kHz: $\chi^2(3) = 5.97, p = .113$, 8 kHz: $\chi^2(3) = 4.93, p = .177$). Further statistical tests were conducted to evaluate whether task performance was nevertheless above chance. These tests yielded significant results for all ILD stimuli in all sound source locations (Table 2). This shows that despite the relatively low level of performance, dynamic cues contributed to front-back localization.

Table 2: Results of the post-hoc comparisons for Experiment II.

Frequency	Source azimuth	> Chance z	> Chance p*	< ITD z	< ITD p*
2 kHz	0°/180°	2.97	**.001**	2.41	**.047**
	±30°/ ± 150°	2.63	**.009**	2.90	**.003**
	±45°/ ± 135°	2.97	**.001**	2.91	**.003**
	±60°/ ± 120°	2.72	**.006**	2.90	**.003**
4 kHz	0°/180°	2.91	**.003**	2.56	**.018**
	±30°/ ± 150°	2.96	**.001**	2.91	**.003**
	±45°/ ± 135°	2.94	**.001**	2.90	**.003**
	±60°/ ± 120°	2.90	**.003**	2.69	**.009**
8 kHz	0°/180°	2.47	**.023**	2.82	**.006**
	±30°/ ± 150°	2.33	**.029**	2.94	**.001**
	±45°/ ± 135°	2.71	**.012**	2.95	**.001**
	±60°/ ± 120°	2.28	**.038**	2.94	**.001**

As in Experiment I, participants could resolve the front-back location of the low-frequency control stimuli more consistently than they could for the high-frequency stimuli. The dependence of performance on stimulus frequency was statistically significant for all source locations (Friedman tests: 0°/180°: $\chi^2(3) = 15.4, p = .002$, ±30°/ ± 150°: $\chi^2(3) = 22.2, p<.001$, ±45°/ ± 135° : $\chi^2(3) = 22.6, p<.001$, ±60°/ ± 120°: $\chi^2(3) = 18.1, p<.001$). Post-hoc tests were conducted comparing the low-frequency condition with each high-frequency condition. All pair-wise comparisons were significant (Table 2), confirming that the ITD stimulus provided a more robust dynamic front-back cue than any of the ILD stimuli at each source position.

Separation of the results from the two hemiplanes suggested contribution from directional band biases similar to those found for Experiment I, i.e. the 2 and

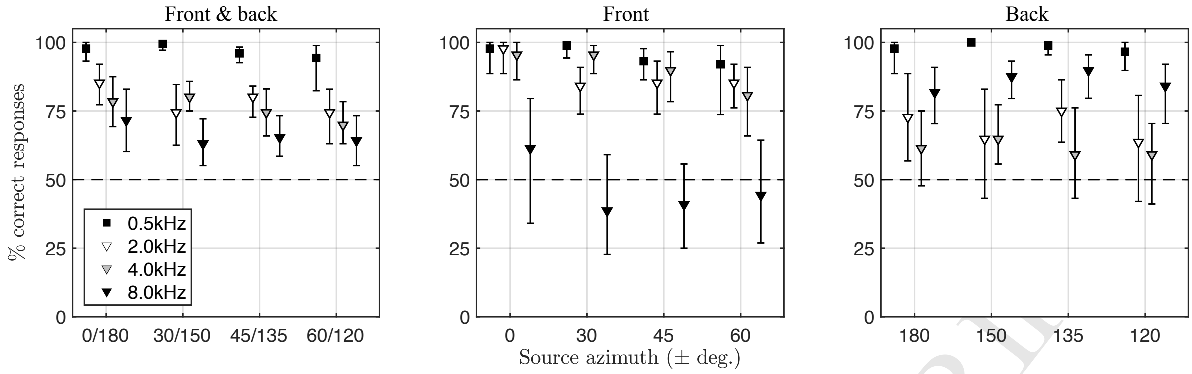


Figure 4: Results from Experiment II (free field sources at lateral positions, movement window fixed at ±20°). Percentage of correct responses (averaged over 11 participants) in the front-back localization task is plotted as a function of free field sound source location. Error bars represent 95% BC_a intervals of the means. The first panel presents the pooled data and the second and third panels show the results from the front and rear hemiplanes, respectively.

4 kHz stimuli received more front than rear responses, suggesting that narrow-band effects were not fully overcome by the dynamics provided by the ±20° movement window.

4.2. *Discussion*

The results of Experiment II show that level dynamics resulting from head rotations contribute to the perception of front-back location also at lateral sound source positions. This was evident in task performance being above chance in the absence of other cues for resolving front-back confusions. Yet, performance was relatively poor for all sound source locations for the ILD stimuli and not comparable to that obtained for the ITD condition. This may have been due to the relatively narrow movement range allowed. However, including a significantly larger movement range, such as the ±40° range leading to high performance in Experiment I, would have led to considerable overlap between the ranges of dynamic cues available for each sound source location. Performance for the 0°/180° source location was slightly lower in Exp. II than in Exp. I for the same (±20°) movement range. This small difference may have been due to the sounds being presented always from the midline in Exp. I, which may have facilitated the maintenance of spatial attention.

Somewhat surprisingly, performance was not dependent on the lateral offset of the sound source. One would expect location-dependent performance because of the different level dynamics arising from head rotations: for central locations these are monotonic but at lateral locations idiosyncratic dynamics become more prevalent due to e.g. the influence of the bright spot (Macaulay et al., 2010). A potential reason why Experiment II did not reveal such effects is the uncontrolled and subject-dependent level dynamics. The experiment was conducted in free field, resulting in natural level dynamics, but as a down side, these dynamics vary across individuals and are unknown to the experimenter unless a dense grid of HRTFs is measured for each subject.

5. Experiment III: front-back localization in virtual auditory space

5.1. *Results*

Experiment III replicated Experiment I with simplistic virtual auditory space stimulation. This allowed the interpretation of the front/back responses in the light of pre-specified azimuth dependence of level dynamics. For the median plane stimuli of Experiment III, level dynamics were monotonic for all stimulus frequencies (Fig. 2). Save for the differences in stimulus presentation method (free field vs. headphones) Experiments I and III were identical.

The results in the left panel of Fig. 5 display a clear improvement in performance with head rotation: the static condition resulted in chance performance, but introducing a movement window of ±5° yielded a large increase in correct response rate (i.e. front-back responses corresponded to the front-back location of the simulated source positions). When the movement range was further extended, front-back confusions no longer occurred. The effect of movement range on front-back localization was statistically significant for all stimuli (Friedman tests: 500 Hz: $\chi^2(4) = 32.4, p<.001$, 2 kHz: $\chi^2(4) = 33.2, p<.001$, 4 kHz: $\chi^2(4) = 28.6, p<.001$, 8 kHz: $\chi^2(4) = 29.5, p<.001$). The contribution of dynamic cues was assessed statistically by comparing each

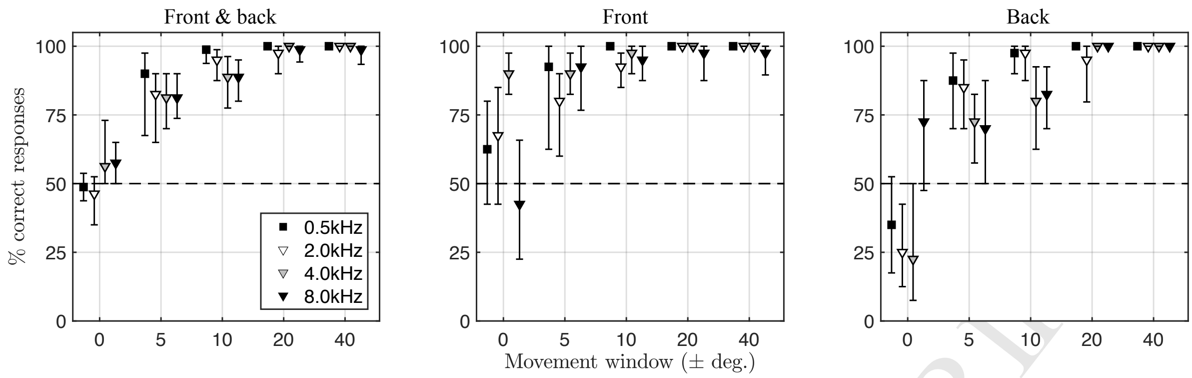


Figure 5: Results from Experiment III (virtual sources at 0° and 180°). The 500 Hz stimuli varied only in ITD and the high-frequency stimuli only in ILD, as the head was turned. Here "correct" responses correspond to front-back responses consistent with the simulated source position. Note that all front and back stimuli in the 0° movement condition were identical in terms of interaural differences. Data are averaged over 10 participants and presented as in Figure. 3.

movement condition to the static condition. This resulted in statistically significant effects in all instances (Table 3), showing that dynamic ILD was used effectively for all stimulus frequencies and for all movement ranges, including the narrowest range tested.

Table 3: Results of the post-hoc comparisons for Experiment III. Comparisons against the ITD condition were conducted only within the ±10° window, as the results from the other windows yielded statistically insignificant results from the Friedman test.

Frequency	Movement window	> Chance z	> Chance p*	< ITD z	< ITD p*
2 kHz	±5°	2.66	**.012**	—	—
	±10°	2.83	**.004**	1.05	.750
	±20°	2.83	**.004**	—	—
	±40°	2.84	**.004**	—	—
4 kHz	±5°	2.46	**.031**	—	—
	±10°	2.36	**.031**	1.99	.188
	±20°	2.82	**.008**	—	—
	±40°	2.82	**.008**	—	—
8 kHz	±5°	2.63	**.016**	—	—
	±10°	2.82	**.004**	2.21	.094†
	±20°	2.83	**.004**	—	—
	±40°	2.83	**.004**	—	—

Statistical testing of differences between ITD and ILD stimuli yielded a significant result for ±10° (Friedman test: $\chi^2(3) = 8.63, p = .035$), but not for the other movement windows (±0°: $\chi^2(3) = 4.37, p = .225$, ±5°: $\chi^2(3) = 5.93, p = .115$, ±20°: $\chi^2(3) = 2.0, p = .572$, ±40°: $\chi^2(3) = 3.0, p = .392$), indicating that task performance was independent of frequency under most movement conditions. Post-hoc tests showed that none of the pair-wise comparisons within the ±10° window yielded significant results (Table 3). In other words, performance based on dynamic ILD was not significantly lower than that reached with the ITD stimuli for any of the movement windows.

Inspection of the results from the two hemiplanes separately revealed that while strong directional band biases were present in the static condition (as suggested for instance by the 2 and 4 kHz stimuli being localized more often to the front than to the back), head rotation allowed dynamic cues to overcome these narrow-band effects.

5.2. *Discussion*

The results from Experiment III show that a front-back location can be assigned to sinusoidal signals presented through headphones if appropriate changes are applied to the earphone signals in response to head rotations. In general, sinusoids presented through headphones are perceived as internalized and the presence of binaural cues is perceived as lateralization along the interaural axis. However, Experiment III shows that when the binaural cues change in response to head rotation, the combination of this dynamic cue with head-position information enabled listeners to consistently report the front-back dimension of the test stimuli according to the simulated directions under most stimulus conditions. Experiment III further displays a notable improvement in front-back localization performance in comparison to the equivalent free-field task of Experiment I (see Figs. 3 & 5). The ILD stimuli led to performance comparable to that for the ITD stimuli for even the smallest movement window of ±5°. This deviates from Experiment I, where task performance with the ILD stimuli was lower for the intermediate movement windows

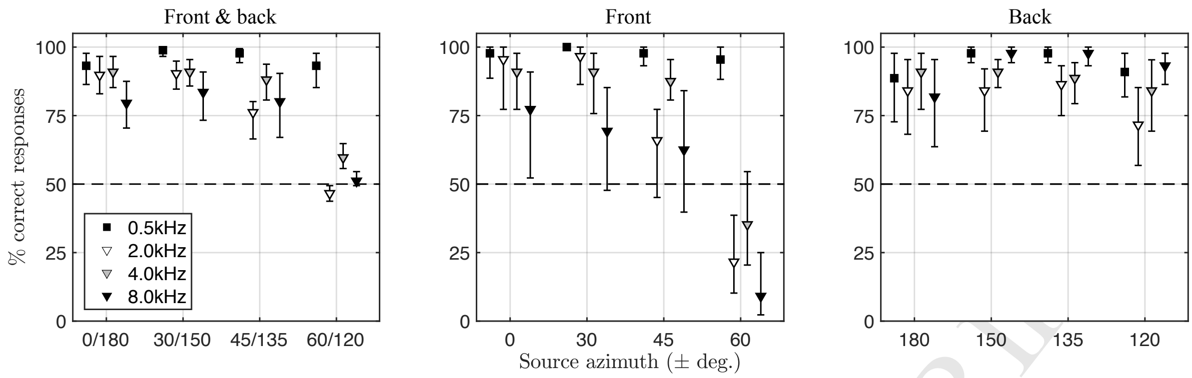


Figure 6: Results from Experiment IV (virtual sources at lateral positions, movement window fixed at ±20°). The 500 Hz stimuli varied only in ITD and the high-frequency stimuli only in ILD, as the head was turned. Data are averaged over 11 participants and presented as in Figure. 4.

and reached the level of ITD stimuli only for the widest range (±40°). The high performance in Experiment III was most likely due to the monotonic level dynamics provided by the VAS stimuli. Referring to the polar plots of Fig. 2, it can be seen that virtual sources located at azimuths 0° and 180° provided smoothly varying dynamics for all ILD stimulus frequencies within all movement windows. The monotonic azimuth dependence near the median plane thus enabled the use of level dynamics as a robust front-back cue.

The combined results from the experiments so far suggest that the auditory system can efficiently integrate both ITD and ILD dynamics with self-motion information to resolve front-back confusions and to overcome biased perception induced by high-frequency narrow-band stimuli. However, this appears possible only when the binaural signals display good continuity throughout the head rotation. This hypothesis was further investigated in Experiment IV, where similar tasks were performed in virtual auditory space with sources at lateral positions.

6. Experiment IV: front-back localization of lateral sources in virtual auditory space

6.1. Results

Experiment IV investigated the effects of monotonic vs. non-monotonic level dynamics on front-back localization. Sources presented in virtual auditory space were placed at lateral positions, where the narrow-band ILD derived from the spherical head model display increasingly non-monotonic azimuth dependence (as shown in Fig. 2b). Compared to the central locations, the ±30°/±150° and ±45°/±135° locations resulted in a slower rate of ILD change and introduced mildly idiosyncratic behavior. For the most lateral locations at ±60°/±120° the level dynamics were qualitatively different: they were non-monotonic for all ILD stimuli (see peaks and notches in Fig. 2b). While being simplistic, this condition provided a test case for whether dynamic ILD modulates location perception even when it provides inconsistent information, i.e. highly non-monotonic ILD changes during head rotation. Apart from the stimulus presentation (headphones vs. free field), Experiment IV was identical to Experiment II.

The results from Experiment IV are shown in Fig. 6. Unlike in Experiment II where the results showed no dependence on source azimuth, here the front-back localization of ILD stimuli showed a clear dependence on the lateral angle of the sound source. As the virtual sources were placed at increasingly lateral locations, performance with ILD stimuli decreased, until settling around chance at the most lateral locations ±60°/±120°. The front-back location of the ITD stimuli in contrast was perceived correctly at all locations. The effect of sound source laterality was statistically significant for all ILD stimuli but not for the ITD stimuli (Friedman tests: 0.5 kHz: $\chi^2(3) = 3.72, p = .293$, 2 kHz: $\chi^2(3) = 26.0, p<.001$, 4 kHz: $\chi^2(3) = 22.0, p<.001$, 8 kHz: $\chi^2(3) = 20.0, p<.001$). Further, performance was significantly above chance for all ILD stimuli at 0°/180°, ±30°/±150°, and ±45°/±135° locations (Table 4). At the most lateral locations (±60°/±120°), performance was at chance except for the 4 kHz stimulus, for which the numerically small difference from chance reached statistical significance.

Though front-back localization of the ILD stimuli was clearly above chance level for most locations, it did not reach the high levels obtained with the ITD stimuli. This was confirmed by a statistically significant effect of stimulus frequency on performance at all azimuths except for the median plane (Friedman tests:

Table 4: Results from the statistical tests for Experiment IV. Comparisons against ITD condition were not performed at 0°/180° because of a non-significant result from the corresponding Friedman test.

Frequency	Source azimuth	> Chance z	> Chance p*	< ITD z	< ITD p*
2 kHz	0°/180°	3.00	**.001**	—	—
	±30°/ ± 150°	2.96	**.001**	2.57	**.023**
	±45°/ ± 135°	2.93	**.003**	2.99	**.001**
	±60°/ ± 120°	-1.92	1.00	2.95	**.001**
4 kHz	0°/180°	2.98	**.001**	—	—
	±30°/ ± 150°	2.95	**.001**	2.39	**.035**
	±45°/ ± 135°	2.95	**.001**	2.57	**.023**
	±60°/ ± 120°	2.85	**.006**	2.90	**.003**
8 kHz	0°/180°	2.93	**.003**	—	—
	±30°/ ± 150°	2.91	**.003**	2.35	**.023**
	±45°/ ± 135°	2.91	**.003**	2.35	**.023**
	±60°/ ± 120°	0.63	1.00	2.99	**.001**

0°/180°: $\chi^2(3) = 6.95, p = .074$, ±30°/ ± 150°: $\chi^2(3) = 10.8, p = .013$, ±45°/ ± 135° : $\chi^2(3) = 16.0, p = .001$, ±60°/ ± 120°: $\chi^2(3) = 26.5, p<.001$). Post-hoc tests (Table 4) compared the scores obtained for the ITD stimuli to those for the ILD stimuli. All pairwise tests were significant, confirming that the task was performed more accurately with ITD than ILD stimuli for all lateral sound source locations.

Plotting the results separately for front and rear source positions revealed a rear bias for all ILD stimulus frequencies at the most lateral source locations (±60°/±120°), for which front-back localization in the pooled responses was at or near chance. Because this bias was independent of stimulus frequency, it is not consistent with directional band biases. Instead, a potential origin may reside in the level dynamics yielded by the spherical head model at this source position (see Discussion below).

6.2. *Discussion*

The results of Experiment IV are perhaps best interpreted in light of the polar plots of the binaural signals yielded by the spherical head model in Fig. 2. The ITD plot in Fig. 2 (c) shows that the ITD is a monotonic function of azimuth in both hemiplanes, thus enabling unambiguous interpretation of the dynamics resulting from head rotations for all sound source locations. Consequently, front-back localization performance was excellent for all sound source locations with the ITD stimuli. In contrast, the magnitude of the ILD changes smoothly with azimuth near the median plane but when the source reaches approximately ±45° in the front hemiplane or ±120° in the rear hemiplane, the ILD responses become idiosyncratic, due to the magnitude ripples in the contralateral side of the sphere's pressure response (see Fig. 2b). Furthermore, the ILD responses have a dip between ±60° and ±90°, where the effect of the bright spot increases the pressure level at the contralateral ear. The pooled responses to the ILD stimuli are broadly in line with the behaviour of the ILD dynamics: performance was accurate for locations yielding good continuity in ILD but dropped to chance for the most lateral locations associated with non-monotonic level dynamics.

Some details in the results of Experiment IV cannot be explained by the simple assumption that non-monotonic ILD dynamics lead to poor performance and possibly to front-back localization based on narrowband biases. In particular, the front-back ambiguity could be resolved at least to some extent, for the 2 kHz ILD stimulus at ±45° even though the corresponding level dynamics were non-monotonic (see Fig. 2b). This stimulus yielded qualitatively similar changes in ILD for both head rotation directions and consequently, dynamic ILD did not offer an unambiguous front-back cue. Further, performance for the 4 kHz stimulus at the most lateral locations was above chance by a small but statistically significant amount. The above-chance performance in these two cases could possibly be explained by considering monaural level dynamics, namely, the input received by the ipsilateral ear alone. The ipsilateral response changes in a monotonic manner between the extreme head orientations and thus could provide an unambiguous front-back cue.

The chance performance for ILD stimuli at ±60° and ±120° is consistent with the highly non-monotonic level dynamics around these locations. However, when inspecting responses to the front and back locations separately, it became evident that the errors were systematic and - unlike in Experiments I, II, and III where frequency-dependent errors reflected directional band biases - independent of stimulus frequency. Instead, there was a systematic rear-bias for all ILD stimuli. This bias suggests that the participants were trying to make use of the level dynamics despite their erratic azimuth dependence. More specifically, the systematic rear bias observed with sources placed at ±60° is likely to arise from the effect of the bright spot. The ±20° rotation range allowed the virtual source to be positioned anywhere between azimuths of 40° and 80° with respect to the head. When the head is rotated towards the end of the movement window so that the sound source is at 80° with respect to the head, the bright spot introduces a strong increase in sound pressure level at the contralateral ear. This increase may have created an impression of a sound source in the rear hemiplane. Because the level variations due to the bright spot are the most prominent dynamic cue present within this movement range, they may have influenced dynamic localization. Similar points have been made in the context of a study assessing the effect of the bright spot on auditory localization with static listeners (Macaulay et al., 2010). An interesting aspect of this systematic rear bias is that the subjects apparently ignored the overall ILD that always favored the ipsilateral ear—and thus a front-hemiplane interpretation of source position—and made their localization judgments based on the dynamics yielded by the bright spot at the contralateral ear. This suggests that strong level dynamics could dominate localization interpretations based on overall ILD.

7. General discussion

The aim of the present study was to establish the usefulness of dynamic ILD as a front-back localization cue and to understand the limitations involved in its use. We found that human listeners are capable of combining level dynamics arising from head rotation with the multimodal information on head position to derive information about the front-back location of a sound source. The usefulness of level dynamics derived from high-frequency sinusoidal signals presented in free field was, however, rather limited. In free field, performance in the front-back task was comparable to that for the more robust dynamic ITD cue only when a relatively large movement range was allowed. Under more challenging conditions, front-back responses reflected a combination of contributions from dynamic ILD and narrowband directional band biases. When a simplistic stimulation derived from a spherical head model was presented through headphones and the level dynamics resulting from head rotation were monotonic, performance with dynamic ILD stimuli approached that observed with dynamic ITD stimuli, even within the smallest tested movement ranges. In general, we observed the results across individual subjects to be qualitatively similar, with individual differences manifesting mainly in the size of the movement window required for ceiling performance and in the effects of directional band biases under static conditions. Overall, the results suggest that the main factor limiting the use of dynamic ILD, or level dynamics in general, as a front-back localization cue resides in the acoustic domain, rather than in potential limitations posed by the human auditory system or the hypothetical dominance of spectral cues or biases over dynamic ILD.

The present study provides an example of how less realistic spatial sound can induce a more robust impression of sound source location. The high-frequency stimuli were poorly localised in free field where spatial cues were natural. However, when simplistic virtual spatial sound was presented in otherwise identical settings, front-back localization judgments were consistently in line with the simulated source positions. Based on the present results, good continuity of the dynamic spatial cues during head rotation appears to be a prerequisite for successful front-back localization, to the extent that realistic variations in cue dynamics can be detrimental. Therefore, the process of combining level dynamics with head orientation information appears not to be calibrated according to the details of the frequency-dependent level dynamics experienced by the listener in everyday life. While such calibration is evident in other aspects of sound source localization (most notably, adaptation to new spectral cues, for instance, Hofman et al., 1998; Van Wanrooij and Van Opstal, 2005; Trapeau et al., 2016) active localization based on level dynamics appears to work under the simplistic assumption that increasing ILD signifies that the head motion results in the sound source being ipsilateral to the ear with the increasing sound pressure level. This assumption is not always valid for narrowband signals in free field and thus localization is poor when only narrow band ILD dynamics are available. However, since the majority of naturally occurring sounds stimulate several auditory bands across multiple octaves, a simple processing scheme based on an assumption that head movement-induced level dynamics are approximately monotonic when computed at the level of auditory objects could prove to be a valid strategy for dynamic localization. Hypothetically, such an approach could be based on averaging the idiosyncratically varying levels across the ensemble of frequency components allocated to any given auditory object by primitive grouping cues such as onset synchrony and coherent envelope fluctuations. Further, an object-based processing scheme would not require the maintenance of multiple frequency-dependent spatial ”templates” based on accumulated experience in dynamic cue processing of narrow-band signals, but rather, would enable level dynamics to facilitate the localization of the vast majority of natural sounds based on the simple assumption that level dynamics are approximately monotonic when computed at the level of perceptually grouped frequency components.

While the headphone stimuli in the present studies violated the expectations of non-monotonic level dynamics during head rotation, they did (at least in the stimulus conditions yielding highest performance in the front-back task) adhere to the principle that object-based level dynamics are approximately monotonic. This suggests that monotonic changes in level dynamics may be a more important factor in active localization than adherence to expected variations within individual auditory bands.

Here, the poor front-back localization in free field is likely to be a consequence of the narrow bandwidth of the sinusoidal stimuli. Changes in the narrowband signal level in each ear can be unstable across azimuth and therefore provide an unreliable localization cue. Wider stimulus bandwidth could allow listeners to overcome these instabilities in dynamic ILD by integrating level information over a wider frequency range across which the overall level dynamics are expected to be more stable across azimuth. Testing this might not be straightforward, as a wide bandwidth would allow spectral localization cues to provide information on sound source location. The use of spectral cues could be prevented by fitting the listeners with short tubes inserted in the ear canals (Perrett and Noble, 1997b) or with moulds that alter the shape of the pinnae (Hofman et al., 1998; Van Wanrooij and Van Opstal, 2005; Trapeau et al., 2016). This however would also alter the frequency-dependent level dynamics. Consequently, assessing the relative contributions of spectral cues and ILD in dynamic localization of wideband sounds is problematic, but may be aided with the application of spherical HRTFs (as was done in e.g. Macpherson, 2013) or other similar abstractions controlled by the experimenter.

Classic studies (e.g. Stevens and Newman, 1936) and also later work (Yost and Zhong, 2014) on the localization of sinusoidal signals show that localization accuracy is frequency dependent. In particular, stimulus frequencies near 2-4 kHz are localized less accurately along the horizontal plane than lower or higher frequency stimuli (e.g. 500 Hz and 8 kHz). The increased localization accuracy for lower frequency sinusoids is attributable to the presence of a robust ITD cue, while improved accuracy for higher frequency sinusoids is due to the larger magnitude of ILD. Our findings are only partially consistent with this general pattern: we found more accurate front-back localization at 500 Hz than for higher stimulus frequencies, but did not find more accurate performance for the 8 kHz than for the 2 and 4 kHz stimuli. This suggests that results on localization accuracy along the left-right dimension obtained under static conditions cannot be directly extrapolated to dynamic front-back localization. The lack of clear frequency dependency in the 2-8 kHz range in our results may be attributable to the different types of binaural information the dynamic front-back task and the classic localization accuracy tasks rely on. Tasks measuring localization accuracy along the horizontal plane require the listener to correctly associate the absolute magnitude of ILD with the correct point along the left-right dimension. In contrast, the dynamic front-back task employed here requires the listener to correctly identify the direction of change in ILD (and to correctly combine this with head-position information). The detection of direction of change, particularly for the relatively large movement windows used here, might be a process less sensitive to the frequency-dependent variations in absolute ILD that cause strong frequency dependence in localization accuracy along the horizontal plane under static conditions.

As in all studies of ILD perception, the present study included not only changes in interaural level but also changes in monaural level. This opens up the possibility that the participants may have based their responses partly on monaural level dynamics. To prevent the use of static monaural level cues, level roving was applied here across trials (as in previous research, for instance, Stevens and Newman, 1936; Francart and Wouters, 2007). However, level roving across trials does not prevent the listeners from using monaural level cues arising from head rotation within each trial. During our dynamic localization task, the listeners could have interpreted an increase in sound level in one ear as signifying that the head rotation resulted in the sound source being nearer to the ear after the rotation, thereby solving the confusion without resorting to the additional information (i.e. decreasing level) provided by the other ear. Some aspects of our findings in fact suggest that, at least under some conditions, listeners do try to make use of monaural level dynamics to infer sound source location. Here, these situations were related to the rather artificial properties of the spherical head model stimulation, although similar situations could arise under natural listening conditions, as suggested by the measurements presented in Macaulay et al. (2010). However, it is not clear whether listeners have access to monaural level information in the presence of binaural cues. There is strong evidence for the inability of listeners to utilize level cues in the presence of roving ITD (Bernstein, 2004). Whether this generalizes to dynamic conditions and localization in the front-back dimension is not clear.

Here, we studied the use of dynamic cues in localization but the fact that the cues were dynamic does not necessarily imply that the auditory system processes these cues as dynamic i.e. as a continuous change in a stimulus parameter. According to the ”snapshot” hypothesis, the perception of auditory motion arises from the detection of a sequence of static locations, rather than from the processing of the continuous change in spatial cues (Grantham, 1986, 1997; Middlebrooks and Green, 1991). Therefore, the present findings on the use of binaural cue dynamics related to head rotation could also arise from the auditory system processing the cues as snapshots. In some previous studies, sound presentation has been limited to occur only during head motion so that the listener hears the localization cues only while they are changing (Macpherson, 2011). Because extracting static location information from a continuously changing cue can be more challenging than from a static one, this could make snapshot processing less effective. Here we did not aim to probe the precise manner in which the change in ILD was used for localization, i.e. whether the ILD information was combined with the multisensory information on head position continuously or as a series of static ILDs. Instead, we aimed for a setup that maximally facilitated the use of binaural dynamics resulting from head rotation. The listeners could sample the stimulus dynamics relatively freely and the auditory input could thereby contain both static and dynamic binaural cues. Consequently, our stimulus and task settings would support the processing of auditory spatial cue dynamics equally well as snapshots and as continuous changes. Therefore, our results cannot be interpreted as in favor of or against the snapshot hypothesis. Independent of whether the stimulus was processed as snapshots or in a continuous manner, the present results nevertheless show that the human auditory system is capable of combining the changes in ILD arising from head rotation with the multisensory information on head position in a manner that facilitates sound source localization.

Acknowledgments

This research was supported by the Academy of Finland (Projects 296751 and 307072). We thank the editor Dr. Brian C. J. Moore and the two reviewers for their thoughtful comments that helped to improve the quality of the manuscript.

References

Aaronson, N. L., Hartmann, W. M., 2014. Testing, correcting, and extending the Woodworth model for interaural time difference. J. Acoust. Soc. Am. 135 (2), 817–823.

Bernstein, L. R., 2004. Sensitivity to interaural intensitive disparities: Listeners use of potential cues. J. Acoust. Soc. Am. 115 (6), 3156–3160.

Blauert, J., 1969. Sound localization in the median plane. Acustica 22, 205–213.

Blauert, J., 1997. Spatial hearing: the psychophysics of human sound localization. MIT press.

Brimijoin, W. O., Akeroyd, M. A., 2012. The role of head movements and signal spectrum in an auditory front/back illusion. i-Perception 3, 179–181.

Brimijoin, W. O., Boyd, A. W., Akeroyd, M. A., dec 2013. The contribution of head movement to the externalization and internalization of sounds. PLoS ONE 8, e83068.

Brughera, A., Dunai, L., Hartmann, W. M., 2013. Human interaural time difference thresholds for sine tones: The high-frequency limit. J. Acoust. Soc. Am. 133, 2839–2855.

Cooper, J., Carlile, S., Alais, D., 2008. Distortions of auditory space during rapid head turns. Exp. Brain. Res. 191, 209–219.

Duda, R. O., Martens, W. L., 1998. Range dependence of the response of a spherical head model. J. Acoust. Soc. Am. 104, 3048–3058.

Efron, B., Tibshirani, R. J., 1994. An introduction to the bootstrap. CRC press.

Francart, T., Wouters, J., 2007. Perception of across-frequency interaural level differences. J. Acoust. Soc. Am. 122, 2826–2831.

Freeman, T. C. A., Culling, J. F., Akeroyd, M. A., Brimijoin, W. O., 2017. Auditory compensation for head rotation is incomplete. J. Exp. Psychol. Hum. Percept. Perform. 43, 371–380.

Grantham, D. W., 1986. Detection and discrimination of simulated motion of auditory targets in the horizontal plane. J. Acoust. Soc. Am. 79 (6), 1939–1949.

Grantham, D. W., 1997. Auditory motion perception: Snapshots revisited. Binaural and spatial hearing in real and virtual environments, 295–313.

Hartmann, W. M., Rakerd, B., Crawford, Z. D., Zhang, P. X., 2016. Transaural experiments and a revised duplex theory for the localization of low-frequency tones. J. Acoust. Soc. Am. 139 (2), 968–985.

Hofman, P. M., Van Riswick, J. G., Van Opstal, A. J., 1998. Relearning sound localization with new ears. Nat. Neurosci. 1, 417–421.

Hollander, M., Wolfe, D. A., Chicken, E., 2013. Nonparametric statistical methods. John Wiley & Sons.

Hothorn, T., Hornik, K., van de Wiel, M., Zeileis, A., 2008. Implementing a class of permutation tests: The coin package. J. Stat. Softw. 28, 1–23.

Itoh, M., Iida, K., Morimoto, M., 2007. Individual differences in directional bands in median plane localization. Appl. Acoust. 68, 909–915.

Kim, J., Barnett-Cowan, M., Macpherson, E. A., 2013. Integration of auditory input with vestibular and neck proprioceptive information in the interpretation of dynamic sound localization cues. In: Proceedings of meetings on acoustics ICA2013. Vol. 19. ASA, p. 050142.

Klumpp, R. G., Eady, H. R., 1956. Some Measurements of Interaural Time Difference Thresholds. J. Acoust. Soc. Am. 28, 859–860.

Kuhn, G. F., 1987. Physical acoustics and measurements pertaining to directional hearing. In: Directional hearing. Springer, pp. 3–25.

Leung, J., Alais, D., Carlile, S., 2008. Compression of auditory space during rapid head turns. Proc. Natl. Acad. Sci. 105, 6492–6497.

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.

Macpherson, E., 2011. Head motion, spectral cues, and wallach's principle of least displacement in sound localization. In: Principles and applications of spatial hearing. World Scientific, pp. 103–120.

Macpherson, E. A., 2013. Cue weighting and vestibular mediation of temporal dynamics in sound localization via head rotation. In: Proceedings of Meetings on Acoustics ICA2013. Vol. 19. ASA, p. 050131.

Macpherson, E. A., 2014. Availability of envelope interaural time-difference cues does not improve front/back localization of narrow-band high-frequency targets via head movement. J. Acoust. Soc. Am. 135 (4), 2282–2282.

Macpherson, E. A., Kerr, D. M., 2008. Minimum head movements required to localize narrowband sounds. In: American Audiology Society 2008 Annual Meeting.

Middlebrooks, J. C., 1992. Narrowband sound localization related to external ear acoustics. J. Acoust. Soc. Am. 92, 2607–2624.

Middlebrooks, J. C., Green, D. M., 1991. Sound localization by human listeners. Annu. Rev. Psychol. 42, 135–159.

Middlebrooks, J. C., Makous, J. C., Green, D. M., 1989. Directional sensitivity of sound-pressure levels in the human ear canal. J. Acoust. Soc. Am. 86, 89–108.

Perrett, S., Noble, W., 1997a. The contribution of head motion cues to localization of low-pass noise. Percept. & Psychophys. 59, 1018–1026.

Perrett, S., Noble, W., 1997b. The effect of head rotations on vertical plane sound localization. J. Acoust. Soc. Am. 102, 2325–2332.

Pöntynen, H., Santala, O., Pulkki, V., 2016. Conflicting dynamic and spectral directional cues form separate auditory images. AES 140th Conv.

Pratt, J. W., 1959. Remarks on zeros and ties in the Wilcoxon signed rank procedures. J. Am. Stat. Assoc. 54 (287), 655–667.

Rabinowitz, W. M., Maxwell, J., Shao, Y. E., Wei, M., 1993. Sound Localization Cues for a Magnified Head: Implications from Sound Diffraction about a Rigid Sphere. Presence 2, 125–129.

Rahe, A. J., 1974. Tables of critical values for the Pratt matched pair signed rank statistic. J. Am. Stat. Assoc. 69 (346), 368–373.

Stevens, S. S., Newman, E. B., 1936. The localization of actual sources of sound. Am. J. Psychol. 48 (2), 297–306.

Thakkar, T., Goupell, M. J., 2014. Internalized elevation perception of simple stimuli in cochlear-implant and normal-hearing listeners. J. Acoust. Soc. Am. 136 (2), 841–852.

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, 879–890.

Van Wanrooij, M. M., Van Opstal, A. J., 2005. Relearning sound localization with a new ear. J. Neurosci. 25 (22), 5413–5424.

Verschooten, E., Shamma, S., Oxenham, A. J., Moore, B. C., Joris, P. X., Heinz, M. G., Plack, C. J., 2019. The upper frequency limit for the use of phase locking to code temporal fine structure in humans: A compilation of viewpoints. Hear. Res.

Wallach, H., 1939. On sound localization. J. Acoust. Soc. Am. 10, 270–274.

Wallach, H., 1940. The role of head movements and vestibular and visual cues in sound localization. J. Exp. Psychol. 27, 339–368.

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.

Wilska, A., 1938. Untersuchungen über das Richtungshören. Acta Societatis Medicorum Fennicae ”Duodecim”.

Xie, B., 2013. Head-related transfer function and virtual auditory display. J. Ross Publishing.

Yost, W. A., Zhong, X., 2014. Sound source localization identification accuracy: Bandwidth dependencies. J. Acoust. Soc. Am. 136 (5), 2737–2746.

Yost, W. A., Zhong, X., Najam, A., 2015. Judging sound rotation when listeners and sounds rotate: Sound source localization is a multisystem process. J. Acoust. Soc. Am. 138 (5), 3293–3310.

Young, P. T., 1931. The role of head movements in auditory localization. J. Exp. Psychol. 14, 95–124.

Zwislocki, J., Feldman, R. S., 1956. Just Noticeable Differences in Dichotic Phase. J. Acoust. Soc. Am. 28, 860–864.