Hearing Research 304 (2013) 20–27

Contents lists available at SciVerse ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research paper

Sound localization in noise and sensitivity to spectral shape☆

Guillaume Andéol[a,*], Ewan A. Macpherson[b], Andrew T. Sabin[c]

[a] Institut de Recherche Biomédicale des Armées, Département Action et Cognition en Situation Opérationnelle, BP 73, 91223 Brétigny sur Orge Cedex, France
[b] School of Communication Sciences and Disorders & National Centre for Audiology, Western University, London, Ontario, Canada N6G1H1
[c] Hearing Aid Laboratory, Department of Communication Sciences and Disorders, Northwestern University, Evanston, IL 60208, United States

ARTICLE INFO

Article history:
Received 22 April 2013
Received in revised form
29 May 2013
Accepted 3 June 2013
Available online 12 June 2013

ABSTRACT

Individual differences exist in sound localization performance even for normal-hearing listeners. Some of these differences might be related to acoustical differences in localization cues carried by the head related transfer functions (HRTF). Recent data suggest that individual differences in sound localization performance could also have a perceptual origin. The localization of an auditory target in the up/down and front/back dimensions requires the analysis of the spectral shape of the stimulus. In the present study, we investigated the role of an acoustic factor, the prominence of the spectral shape ("spectral strength") and the role of a perceptual factor, the listener's sensitivity to spectral shape, in individual differences observed in sound localization performance. Spectral strength was computed as the spectral distance between the magnitude spectrum of the HRTFs and a flat spectrum. Sensitivity to spectral shape was evaluated using spectral-modulation thresholds measured with a broadband (0.2–12.8 kHz) or high-frequency (4–16 kHz) carrier and for different spectral modulation frequencies (below 1 cycle/octave, between 1 and 2 cycles/octave, above 2 cycles/octave). Data obtained from 19 young normal-hearing listeners showed that low thresholds for spectral modulation frequency below 1 cycle/octave with a high-frequency carrier were associated with better sound localization performance. No correlation was found between sound localization performance and the spectral strength of the HRTFs. These results suggest that differences in perceptual ability, rather than acoustical differences, contribute to individual differences in sound localization performance in noise.

1. Introduction

Although audition allows us to accurately localize a sound's origin, the auditory sensory epithelium is not spatially organized. Instead, the auditory system must "rebuild" the auditory space based on acoustic cues, specifically: binaural cues for the left/right dimension and spectral cues for the up/down and front/back dimensions. As indicated by large individual differences in sound localization performance, the quality of these cues and/or the ability to process them could differ among normal hearing listeners. To date, there have been few direct examinations of the factors responsible for such individual differences. Here we tested the role of sensitivity to spectral shape in individual differences in sound localization performance.

Sound localization ability is partially determined by spectral cues, arising from the acoustic filtering of the outer ears, head and upper torso, that shape the spectrum of the incoming sound wave according to the sound source direction (Shaw, 1974, 1997). The function that describes this spectral shaping is called the head related transfer function (HRTF)(Wightman and Kistler, 1989a). The spectral cues in the HRTFs are responsible for front/back as well as up/down localization (Shaw, 1974, 1997). These cues are assumed to be particularly affected by background noise, given the increasing of localization errors in the front/back and up/down dimensions with signal-to-noise ratio degradation (Good and Gilkey, 1996).

Large individual differences are regularly observed in localization in front/back and up/down dimensions (Wenzel et al., 1993; Wightman and Kistler, 1989b; Zahorik et al., 2006). For instance, the proportion of localization trials on which listeners judge that a sound is behind them when it is actually in front of them (or vice versa) can vary by a factor of 20 (from 2% to 40%) among naïve listeners in free field conditions (Wenzel et al., 1993), and the mean localization error in the up/down dimension can range from as little

Abbreviations: DTF, directional transfer function; HRTF, head related transfer function; SLT, sound localization threshold; SMF, spectral modulation frequency; SMT, spectral modulation threshold; SNR, signal-to-noise ratio

* Corresponding author. Tel.: +33 607411905.
E-mail addresses: guillaume.andeol@irba.fr, guillaume.andeol@gmail.com (G. Andéol), ewan.macpherson@nca.uwo.ca (E.A. Macpherson), a-sabin@northwestern.edu (A.T. Sabin).

0378-5955/$ – see front matter © 2013 Authors. Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.heares.2013.06.001

as 5° to as much as 40°, depending on the listener (Wenzel et al., 1993). Moreover, in a noisy environment, individual differences are even larger (Best et al., 2005). One potential explanation for the individual differences in localization performance might be the variations in the features of spectral cues across listeners, due to diversity in outer ear size and/or shape. Some outer ears might thus provide more prominent cues than others.

Consistent with this hypothesis, Butler and Belendiuk (1977) found that a listener with poor localization performance could improve when listening through recordings made using somebody else's ears. Wenzel et al. (1988) argued that sound localization performance can be predicted by the analysis of acoustical properties of the outer ears. They showed that two listeners initially differing in performance, could reach the same performance level if they both individual listened through the same HRTFs (Wenzel et al., 1988). However, these findings were not confirmed in subsequent studies. Using large groups of listeners, Møller et al. (1996) and Middlebrooks (1999b) found that listening through somebody else's ears always resulted in worse performance. Interestingly, Middlebrooks demonstrated that the pattern of localization errors of a listener listening through another listener's HRTFs did not directly depend on the latter's HRTFs, but that it was highly correlated with the magnitudes of the differences between the HRTFs of the latter and the former (Middlebrooks, 1999a,b). Finally, Wightman and Kistler (1999) observed that listeners with similar "spectral detail" (as determined by visual inspection) in their HRTFs strongly differed in their sound localization performance. Nevertheless, to our knowledge, no study has investigated the relationship between a quantification of HRTF spectral detail and sound-localization performance.

Based on findings of Møller et al. (1996), Middlebrooks (1999b) and Wightman and Kistler (1999), we hypothesized another origin for individual differences in sound localization performance than the acoustical characteristics of the outer ear; as suggested by Wightman and Kistler (1999), it is possible that differences in the ability to detect spectral cues (Drennan and Watson, 2001; Eddins and Bero, 2007), also contribute to individual differences in sound localization ability. One approach to testing this possibility would be to examine the correlation between performance in a non-spatial spectral-shape perception task and performance in a spatial hearing task. A similar approach has been used successfully for speech perception studies: for instance, Saoji et al. (2009) found a strong relationship between spectral modulation threshold and vowel/consonant identification performance in cochlear implant listeners.

A correlation between sensitivity to spectral shape and sound localization ability would likely be restricted to those aspects of the spectral shape that convey spatial cues. Because of the limited physical dimensions of the outer ears, spatial cues introduced by outer ear filtering are mainly restricted to the high-frequency part of the spectrum (above 4 kHz). Therefore, assessing sensitivity to spectral shape above 4 kHz could be of particular interest. Likewise, a limited scale of details of the spectral shape seems to be relevant for localization. The results of studies by Macpherson and Middlebrooks (2003) and Qian and Eddins (2008) suggest that spectral details finer than 2 c/o (cycles per octave) do not influence sound localization. Therefore, it appears that spectral localization cues are conveyed by variations in the spectral shape above 4 kHz and at spectral modulation frequencies (SMFs) lower than 2 c/o.

In this study, we explored the extent to which individual variability in sound localization performance was attributable to differences in sensitivity to spectral envelope (the perceptual hypothesis) and/or to differences in HRTF acoustics (the acoustical hypothesis). To maximize individual differences in spatial sensitivity, the spatial task was conducted in noise (Best et al., 2005). Based on previous work, we reasoned that listeners' performance in this spatial task would reflect primarily the detection of spectral cues because these cues are assumed to be more strongly disrupted by noise than are binaural cues (Good and Gilkey, 1996).

To measure sensitivity to spectral shape in a non-spatial context, we used a spectral modulation detection task (Eddins and Bero, 2007). This task allowed us to determine the minimal modulation depth required to discriminate a flat spectrum stimulus from a stimulus with a sinusoidally modulated spectrum. This minimal modulation depth is called the spectral modulation threshold (SMT). We tested spectral modulation detection at different SMFs and audio frequencies because the spectral localization cues vary across these dimensions as do the SMTs. We chose to determine the SMT of stimuli whose carriers were in two different audio frequencies regions: one restricted to spectral region conveying localization cues (4–16 kHz) and one including a larger part of the audible spectrum (0.2–12.8 kHz).

We hypothesized that the relationship between SMT and sound localization performance would be stronger for the high-frequency (4–16 kHz) carrier. We also hypothesized that the correlation would be strongest at the SMFs that are critical for localization. Based on the results of Macpherson and Middlebrooks (2003), correlations should be stronger for SMFs below 2 c/o, and strongest for SMFs around 1 c/o. Significant correlations might also be observed for high-SMF stimuli if the sensitivity to spectral localization cues is related to general ability to detect spectral modulation regardless of the SMF. Finally, to separate the contribution of acoustic factors (spectral details of HRTFs) and perceptual factors (spectral shape sensitivity), we measured the spectral strength of listeners' HRTFs. The spectral strength of individual HRTFs was quantified using the spectral distance, as defined by Middlebrooks (1999a), between the magnitude spectrum of the HRTFs and a flat spectrum.

2. Methods

2.1. Participants

Nineteen participants (nine females; mean age, 30.7 ± 8 years) participated in the study. All had normal hearing (defined as thresholds of 20 dB HL or less at octave frequencies between 0.125 and 8 kHz) and no history of auditory pathology. Otoscopy was also normal. The spectral resolution ability of each participant was checked by a ripple reversal test for a 0.1–5-kHz bandwidth and 30-dB modulation-depth stimulus (Henry et al., 2005). Each participant had a ripple reversal threshold better than 2 c/o. The average ripple reversal threshold was 4.33 c/o and the range was 2.05–7.05 c/o. These results were very close to those obtained by Henry et al. (2005) with a similar population of normal-hearing participants ($n = 12$; mean = 4.84 c/o; range = 2.03–7.55 c/o).

In agreement with the guidelines of the Declaration of Helsinki and of the Huriet law regulating biomedical research in humans in France, participants provided written informed consent before inclusion in the study. All participants were paid (10 €/h) for their services.

2.2. Sound localization task

2.2.1. Task and procedure

The experimental design was similar to that of a previous study (Andeol et al., 2011). The sound localization task was conducted in a semi-anechoic room (Illsonic Sonex Audio). Listeners were seated on an elevated chair whose position was adjusted so that the listener's head was 2.5 m away from each one of eight surrounding loudspeakers (Fig. 1A). The loudspeakers were mounted on the vertices of a cuboid frame (height, 2.76 m; length, 2.94 m; depth, 2.94 m). The loudspeakers' coordinates (azimuth, elevation) were

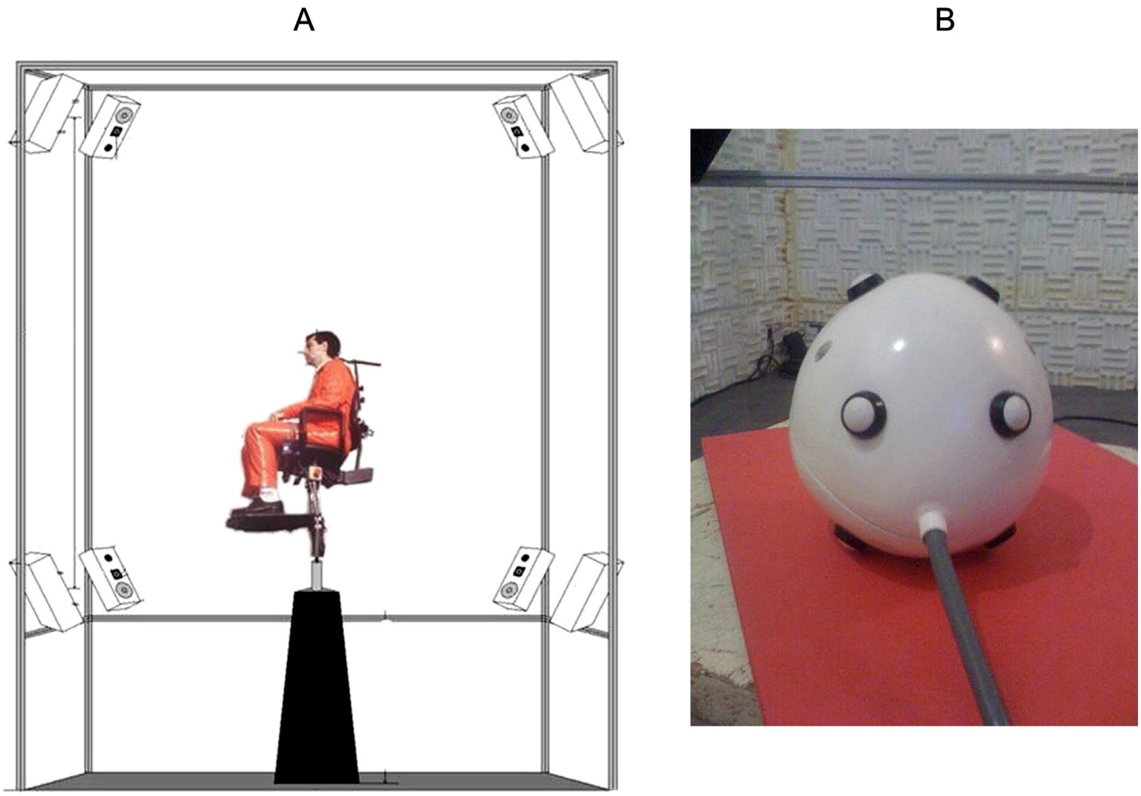

Fig. 1. A. Apparatus used to measure sound localization performance. B. The eight-button response device. Each button is dedicated to one loudspeaker.

as follows: 45°, 35°; 135°, 35°; 225°, 35°; 315°, 35°; 45°, −35°; 135°, −35°; 225°, −35°; and 315°, −35° (see Fig. 1A). During the test sessions, listeners were asked to look straight ahead and to hold their heads and eyes steady. Two video cameras allowed constant monitoring of the head position by both the experimenter and the listener (by way of a video monitor placed in front of the listener). Loudspeakers were visible.

Listeners performed a forced-choice loudspeaker-identification task. On each trial, the signal was emitted from one of the loudspeakers. The listener indicated which loudspeaker had emitted the signal using an eight-button device (Fig. 1B). No feedback was provided. Loudspeaker identification performance was examined for 7 background conditions: a quiet condition and 6 masked conditions with signal-to-noise ratios (SNRs) ranging from −7.5 dB to +5 dB. The loudspeaker emitting the signal was chosen pseudorandomly so that each loudspeaker emitted the target an equal number of times. In the same way, the background condition was chosen pseudorandomly so that each background condition was presented an equal number of times. During a session, each listener completed 5 trials per background condition (7) per loudspeaker (8) resulting in $5 \times 7 \times 8 = 280$ trials per session. Listeners did 4 sessions (1120 trials). Eight training trials were provided at the beginning of each session to allow listeners to check that they were holding the response device securely. The percentage of correct loudspeaker identifications was computed separately for each SNR, session and listener.

2.2.2. Stimulus synthesis

All stimuli were generated digitally at a 48.828 kHz sampling rate using a real-time processor (RX8; Tucker-Davis Technologies) with eight digital-to-analog converters (DACs). The output of each DAC was attenuated (PA5; Tucker-Davis Technologies) and routed to the corresponding loudspeaker via an amplifier (D-75A; Crown). The target was a 200-ms burst of pink noise bandpass filtered between 0.3 and 9 kHz using fourth-order Butterworth filters, and including 36-ms on/off cosine-squared ramps. The masker was a broader (bandpass filtered between 0.125 and 15 kHz using fourth-order Butterworth filters) and longer-duration (500-ms including 36-ms on/off cosine-squared ramps) burst of pink noise. The frequency limits of the target and masker were adjusted in a pilot experiment to ensure both that the loudspeaker that emitted the target was easily identified in quiet, and that the target and masker were easily distinguished when presented together. Each of the eight loudspeakers simultaneously emitted an independent sample of this masker signal to create a diffuse-field sensation (ISO 4869-1). The 200-ms target was temporally centered in the 500-ms masker. The target level was equal to 55 dB SPL. The masker level was set relative to the level of the signal so as to produce six different SNRs ranging from −7.5 to +5 dB in 2.5-dB steps. The target level and the masker level were measured in the center of the loudspeaker array.

2.3. Spectral modulation detection task

2.3.1. Task and procedure

The spectral modulation detection task consisted of distinguishing a target signal with a modulated spectral envelope (ripple), from a flat-envelope “standard” (Eddins and Bero, 2007; Sabin et al., 2012). On each trial, three intervals were presented in random order: two of the three intervals contained the standard; the remaining interval contained the target. Listeners indicated the target interval by clicking with a mouse on a computer screen. After each trial, visual feedback was provided to indicate whether the response was correct or incorrect. The modulation depth (peak to valley difference in dB) was adjusted from trial to trial according to an adaptive procedure in order to estimate the spectral modulation detection threshold (SMT). The adjustment followed a “3 down, 1 up" rule which tracked the 79.4% correct point on the psychometric function (Levitt, 1971). A reversal was defined as a change in the direction of modulation depth from decreasing to increasing (or the opposite). The initial modulation depth was 15 dB. It was adjusted in 3-dB steps until the third reversal, and then in 0.4-dB steps. In each block of 60 trials, the first three reversals were excluded. The modulation depths at the remaining reversals were averaged to obtain the SMT. Blocks containing fewer than 7 reversals were discarded. Listeners completed three blocks (180 trials) for each condition tested. The final SMT was obtained by averaging the SMT across these three blocks. Before starting a series of three blocks, listeners were allowed to familiarize themselves with the stimuli via passive exposure. Following this familiarization phase, listeners had to correctly answer five consecutive trials with a fixed spectral modulation depth of 15 dB before data collection began. SMTs were measured for various stimuli differing in their carrier bandwidth (high-frequency: 4–16 kHz; broadband: 0.2–12.8 kHz), and in their SMF: 0.75, 1.5 and 3 c/o for stimuli with 4–16 kHz bandwidth, and 0.5, 1 and 4 c/o with a 0.2–12.8 kHz bandwidth.

2.3.2. Stimulus synthesis

The procedure for stimulus generation was adapted from a previous study (Eddins and Bero, 2007). The signals were generated digitally with a sampling frequency of 48.828 kHz. A sinusoid on a logarithmic frequency axis with the appropriate SMF and modulation depth (in dB) was used to fill an 8192-point buffer to generate a sinusoidal spectral modulation. This signal was converted from dB to linear magnitude and multiplied by a second buffer filled with random numbers extracted from a Gaussian distribution. This second buffer was then multiplied by the magnitude response of a Butterworth filter (-32 dB/octave) with condition-specific lower and upper cut-off frequencies (0.2–12.8 kHz, or 4–16 kHz). A random phase spectrum was combined with the resulting magnitude. The real part of the inverse FFT of the resulting spectrum was computed. In the time domain, the waveform was shaped by a 150-ms amplitude envelope with 10-ms raised cosine on/off ramps and then scaled to a standard RMS amplitude. To prevent listeners from using local level cues, a roving level of ± 8 dB was applied around a spectrum level of 35 dB SPL, and the modulation phase was chosen randomly from a uniform distribution spanning $0-2\pi$.

2.3.3. Stimulus presentation

All stimuli were presented using custom software written in MATLAB and played using a real-time processor (RX6; Tucker-Davis Technologies), a programmable attenuator (TDT PA5), and a headphone driver (TDT HB7). The sounds were presented through the left earpiece of Sennheiser HD200 headphones. Listeners were tested in a sound-attenuating room.

2.4. Spectral strength

2.4.1. Measurement of transfer functions

HRTFs were measured in a semi-anechoic chamber (Illsonic Sonex Audio). Participants were tested individually. Each participant was seated in an elevated armchair approximately 2.5 m above the ground. The seat height was adjusted in such a way that the participant's head was located at the center of a moveable arc with a radius of 1.4 m. A Fostex 103 Sigma loudspeaker was mounted on the arc, and enclosed in a paralleloid box (15 $\times$ 14 $\times$ 11 cm). The loudspeaker emitted a periodic, pseudo-random signal (a 13th order maximum-length sequence). The signal was generated digitally (48.828 kHz sampling frequency), converted using a TDT RP2.1 processor fitted with a 24-bit sigma-delta digital-to-analog (D/A) converter, and then amplified by a Crown D70A. Each sequence was 168 ms long. The level of the signal measured at the participant's head was equal to 70 dB SPL.

Acoustic signals were recorded simultaneously in the left and right ear canals using miniature microphones (Sennheiser KE211-4), and converted into digital signals (24-bit A/D, 48.828 kHz sampling frequency). Following the "blocked ear meatus" method, each microphone was surrounded in silicon, and the microphone membrane was placed close to the ear-canal entrance. Recordings were performed for 145 different loudspeaker positions corresponding to all combinations of 18 azimuth positions (ranging from 0° to 340° in 20° steps) and eight elevation positions (ranging from $-60°$ to $+80°$ in 20° steps), plus one measure at an elevation of $+90°$ (i.e., vertical), plus eight positions with the same coordinates as the loudspeakers used in the sound localization task, which made a total of 153 ($145 + 1 + 8$) positions. For each loudspeaker position, five consecutive repetitions of the elementary sequence were played and the microphone responses to the last four repetitions were averaged. Circular cross-correlation of the sequence with each averaged response yielded the impulse response of the HRTF (Rife and Vanderkooy, 1989) combined with the transfer functions of the loudspeaker and microphone.

Directional transfer functions (DTFs) were computed according to the method outlined by Middlebrooks (1999a). To eliminate sound propagation times between the loudspeaker and the microphone, and residual sound reflections off of the semi-anechoic chamber walls and seat, temporal windows (boxcar) were applied to the impulse responses, such as that only 128 points (2.6 ms) were retained. The location of the temporal window was defined according to the loudspeaker-microphone distance for the maximal head size in the population of participants.

Each HRTF was obtained by dividing the Fourier transform (FT) of the impulse response measured using the microphones (blocked ear canal) by the FT of the impulse response of the loudspeaker, which was measured using a microphone placed at the position corresponding to the center of the listener's head in the absence of the listener. For each ear, the non-directional component of the HRTFs was computed as the square root of the weighted sum of squared HRTFs that were measured for each sound-source location, with the weights adjusted to take into account the non-uniform distribution of tested sound directions. The magnitude of this non-directional transfer function was inverted, bandpass filtered (50 Hz–14 kHz), and assigned a minimum-phase function. Each of the original HRTFs was then multiplied by the result, yielding an ensemble of normalized HRTFs, which Middlebrooks (1999a) and Middlebrooks and Green (1990) referred to as DTFs. Impulse responses corresponding to these DTFs were computed using an inverse Fourier transform, and were 256-point long each. DTFs contain only directional information and are independent of the characteristics of the microphone and of its insertion into the external ear canal.

2.4.2. Spectral strength

To represent the DTFs and the SMF stimuli on a common logarithmic frequency scale, the procedure described by Middlebrooks (1999a) was employed. Each DTF was filtered by a bank of triangular bandpass filters. Each filter had roll-off slopes of 105 dB/oct (equivalent to a 3-dB bandwidth of 0.057 octaves), and their centre frequencies were spaced by 2% increments in frequency (steps of 0.0286 octaves), yielding 35 bands/octave. The output of each filter was normalized and converted to dB to provide a slightly smoothed estimate of the DTF gain across frequency. The filterbank comprised centre frequencies from 1 kHz to 16 kHz. To compute the overall spectral strength for an individual DTF, the across-frequency variance of the filterbank outputs was computed for centre frequencies between 3.7 kHz and 12.9 kHz (Middlebrooks, 1999a), yielding a value in dB^2. Because only the spectral cues from the ipsilateral ear are thought to be used for location determination of targets with laterality above 30° (Hofman and Van Opstal, 2003; Macpherson

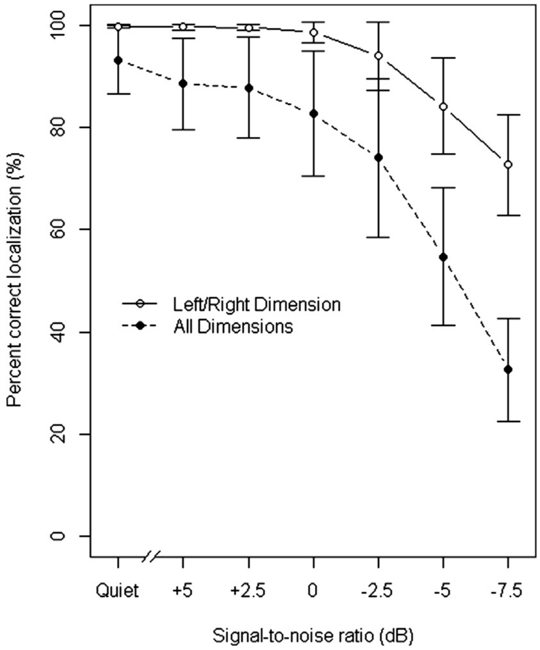

Fig. 2. Mean percent-correct localization performance as a function of signal-to-noise ratio for all dimensions and for the left/right dimension. Bars show 1 standard deviation.

and Sabin, 2007), we computed spectral strength separately for the left ear and for the right ear by averaging the spectral strengths for the four target locations ipsilateral to each ear.

To compute the spectral strength in specific spectral modulation frequency bands, each filter-bank-processed DTF spectrum was first extended by two octaves at the low- and high-frequency ends by duplicating the 1-kHz and 16-kHz values, respectively. The extended spectrum was then convolved with the impulse responses of spectral-modulation-frequency band-pass filters having Gaussian profiles, standard deviations of 0.25 c/o, and center SMFs of 0.5, 0.75, 1, 1.5, 3 and 4 c/o. As for the overall spectral strength, the SMF-specific spectral strength was taken as the variance of the SMF-filtered DTF spectrum between 3.7 and 12.9 kHz. Therefore, the spectral strength was computed for the high frequency region only, in a bandwidth close to the one (4–16 kHz) used in the SMT task.

3. Results

3.1. Sound localization

As expected, sound localization performance decreased with decreasing SNR (repeated measure ANOVA $F(6,108) = 206.16$; $p < 0.001$) (Fig. 2). Sound localization errors occurred mainly in the front/back and/or up/down dimensions; the percent-correct localization score for the left/right dimension (also represented in Fig. 2) was relatively high up to the most adverse SNR. Localization psychometric curves were fitted to the percent correct scores (across all dimensions) across conditions. For each listener, the sound localization threshold (SLT) was determined by the SNR yielding a performance halfway between chance (12.5%) and 100%, i.e. (56.25%). Information concerning the mathematical function used for data fitting can be found in a previous study (Andeol et al., 2011). The average SLT was −4.3(±2.1) dB SNR and the range was −7.8 to −0.5 dB SNR.

3.2. Spectral modulation threshold

Spectral modulation thresholds (SMT) varied across SMFs. They tended to be higher for SMFs below 1 c/o than for higher SMFs (Fig. 3).

This pattern of SMTs differences between SMFs is similar to that observed in a previous study (Eddins and Bero, 2007). The average SMTs measured in the present study (0.5 c/o: 12.4 (±3.4); 1 c/o: 8.3(±2.9); 4 c/o: 8.0(±3.8)) were slightly higher than those measured in the previous study, probably because of the shorter stimulus duration used in the present study.

3.3. Relationship between SLT, SMT and spectral strength

3.3.1. Relationship between SLT and SMT

To determine whether the ability to detect spectral modulation could explain individual differences in sound localization performance, we computed correlations between SMTs and SLT (Table 1

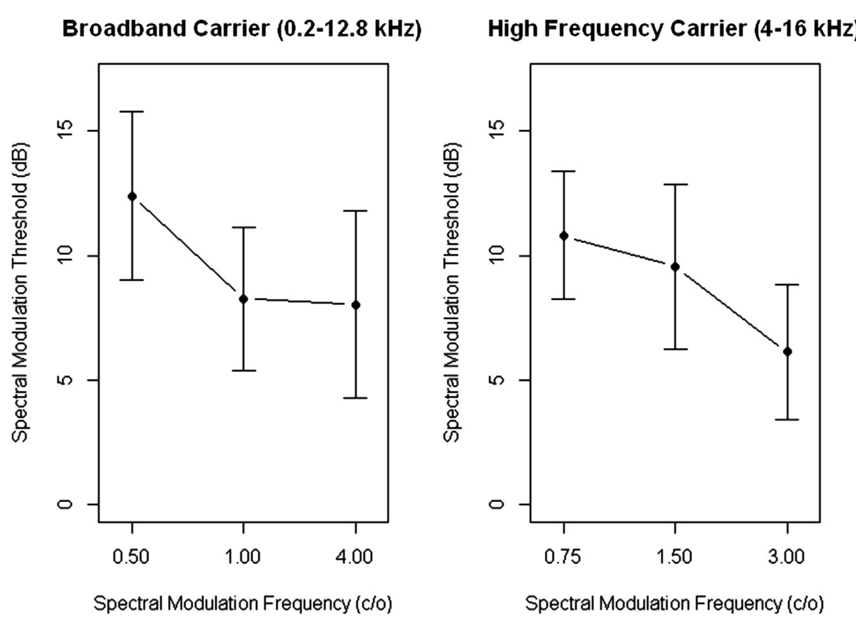

Fig. 3. Mean spectral modulation threshold as a function of spectral modulation frequency. Bars show 1 standard deviation.

Table 1
Correlations between SLTs and SMTs for high-frequency and broadband carriers and different SMFs. The significant correlation (after Bonferroni correction) is in bold.

Correlations between SLT and SMTs			
Carrier	SMF	rho	*p*
High-frequency carrier	0.75 c/o	**0.62**	**0.006**
	1.5 c/o	0.35	0.15
	3 c/o	0.20	0.42
Broadband carrier	0.5 c/o	0.35	0.15
	1 c/o	0.36	0.13
	4 c/o	0.29	0.225

and Fig. 4). The correlation coefficients and corresponding *p* values are listed in Table 1. A significant correlation was found between the SLT and the SMT at 0.75 c/o for the high-frequency carrier (rho = 0.62; $p = 0.006$). This correlation remained significant after Bonferroni correction for multiple significance tests ($n = 6$; $p = 0.05/6 = 0.0083$). No other significant correlation was observed.

The correlation between the SMT at 0.75 c/o for the high-frequency carrier and the SLT might be explained by better ability to detect the signal in the presence of the masker among listeners with better SMT. To address this issue, the correlation between the SMT at 0.75 c/o and the percent-correct localization scores in all dimensions was computed for each SNR (Table 2). This analysis allowed us to check whether the correlation was observed for SNRs at which the signal was easily detectable as shown by a high

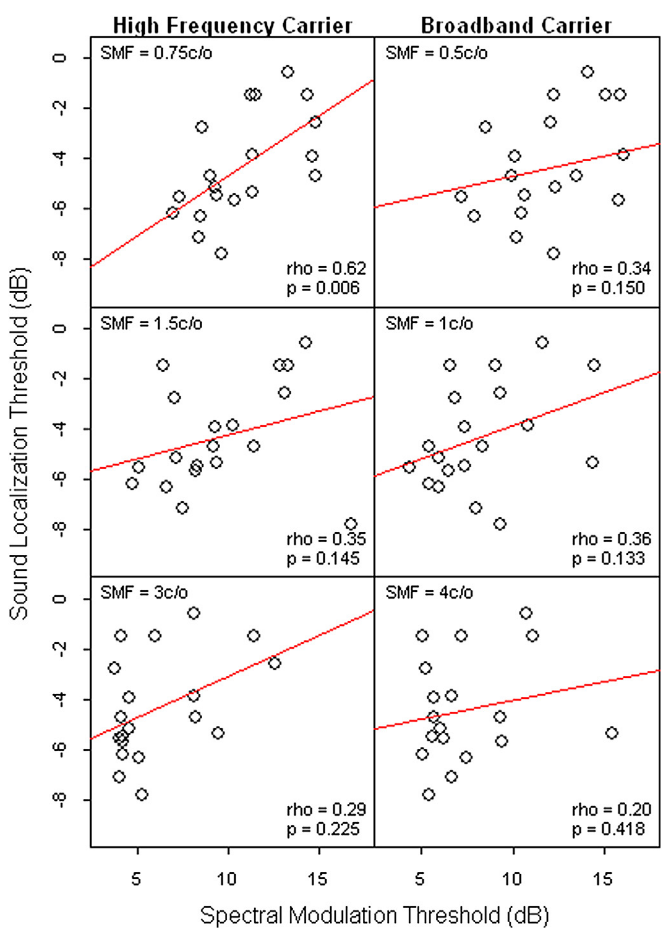

Fig. 4. Spectral modulation thresholds plotted against sound localization thresholds for individual listeners and for each spectral modulation frequency and carrier condition. Linear regressions are shown for each data set. The spearman coefficient of correlation and its degree of significance is indicated in each panel.

percent-correct localization score in left/right dimension (at −2.5 dB SNR, the mean percent-correct localization score in left/right dimension was above 94%(±6.7), Fig. 2). After Bonferroni correction ($n = 7$; $p = 0.05/7 = 0.0071$), the correlations between the SMT at 0.75 c/o and the percent-correct localization in all dimensions were significant for −5 dB (rho = −0.65; $p = 0.0027$) and −2.5 dB (rho = −0.65; $p = 0.0023$) SNRs.

3.3.2. Relationship between spectral strength and SLT

Due to listener and equipment (HRTF measurement device) availability, spectral strength could be computed for only 15 of the 19 listeners. The average left ear spectral strength was equal to 16.9 dB^2(±2.65); the average right ear spectral strength was equal to 17.3 dB^2 (±3.3). No significant correlation across listeners between the spectral strength for the left ear and the SLT for the left-hemisphere targets (rho = 0.11, $p = 0.69$), or between the spectral strength for the right ear and the SLT for the right-hemisphere targets (rho = 0.26, $p = 0.37$) was found.

However, it is important to point out that the overall spectral strength does not precisely represent the spectral strength for the band of SMFs which are believed to be relevant for sound localization, ie SMFs up to 2 c/o (Macpherson and Middlebrooks, 2003; Qian and Eddins, 2008). Fig. 5 shows the overall spectral strength, as well as the spectral strength for different SMFs (0.5; 0.75; 1; 1.5; 3 and 4 c/o), for the left and the right ears. The spectral strength decreased rapidly with the increasing SMF. The average spectral strength was below 2.5 dB^2 for SMFs ≥1 c/o and below 0.2 dB^2 for SMFs ≥ 3 c/o. No significant correlation was found between SLT and the spectral strength computed for the different SMFs. No correlation was computed for SMFs ≥ 3 c/o due to their very low values (Table 3).

3.3.3. Relationship between SMT and spectral strength

It is possible that listeners with low spectral strength may have been naturally trained to improve their sensitivity to spectral shape to detect their own spectral cues. Conversely, listeners with high spectral strength may not have benefited from such training. If so, a positive correlation might exist between spectral strength and SMT. To test this hypothesis, we computed correlations between spectral strength and SMT (see Table 4). No significant correlation was found for any of the SMFs.

4. Discussion

The current study evaluated the role of a perceptual factor, sensitivity to spectral shape, and compared that to an acoustic factor, spectral strength, in individual differences observed in sound localization ability. We found that for localization data obtained in a noisy background which disturbs spectral contrast, the listeners with better sensitivity to spectral shape performed better in the sound localization task. Interestingly, we found that this relationship was observed in noise conditions in which the signal was easily detectable. This result suggests that better sound localization performance in listeners with better sensitivity to spectral shape was not explained by better detection ability in these listeners.

Previous studies have suggested that acoustic factors such as the amount of spectral detail in HRTFs (Butler and Belendiuk, 1977; Wenzel et al., 1988) contribute to individual differences in sound localization performance. Although other studies (Middlebrooks, 1999b; Møller et al., 1996; Wightman and Kistler, 1999) have questioned this conclusion, to date no study has assessed directly the role of these factors, or explored the role of other types of factors (such as perceptual factors) in individual differences. In the current study, we found no significant correlation between a particular acoustic factor (spectral strength) and individual sound localization performance. We cannot, however, exclude the possibility that other types of

Table 2
Correlations between SMT at 0.75 c/o (high frequency carrier) and percent-correct localization scores (all dimensions) for quiet and each SNR conditions. The significant correlations (after Bonferroni correction) are in bold.

Correlations between SMT at 0.75 c/o (high frequency carrier) and percent-correct localization scores for quiet and each SNR conditions														
	Quiet		+5 dB SNR		+2.5 dB SNR		0 dB SNR		−2.5 dB SNR		−5 dB SNR		−7.5 dB SNR	
	rho	*p*	rho	*p*	rho	*p*	rho	*p*	rho	*p*	rho	*p*	rho	*p*
SMT 0.75 c/o high-frequency carrier	−0.38	0.11	−0.49	0.033	−0.55	0.0140	−0.49	0.0339	**−0.65**	**0.0027**	**−0.65**	**0.0023**	−0.47	0.0045

acoustic factors might play a role in individual differences. The correlations observed between sound localization performance and spectral-modulation thresholds suggest that sensitivity to the spectral envelope is one of the factors that contributes to individual variability in sound localization performance. However, such a correlation could also reflect an indirect relationship between these two variables, which can both be influenced by a third variable. A nonspecific ability involved in spectral analysis might indeed influence both spectral modulation threshold (by facilitating detection of spectral modulation) and sound localization performance (by facilitating detection of spectral localization cues). However, the observed correlation pattern suggests a direct relationship: correlations were only observed for the high-frequency carrier and for a SMF equal to 0.75 c/o. This pattern of results is inconsistent with a nonspecific factor related to spectral analysis ability.

The observed correlation is consistent with the location of spectral cues in the high-frequency part of the audio range and in the low spectral modulation frequency range (<2 c/o). High-frequency carrier spectra allow assessment of listeners' ability to detect spectral modulation in the same frequency region as spectral localization cues. Thus, high-frequency carrier stimuli can be considered as more specific to spectral localization cues. Furthermore, the absence of correlation for SMFs higher than 0.75 is consistent with the low spectral strength values we found for SMFs higher than 0.75 c/o. It is also consistent with Macpherson and Middlebrooks' (2003) study, which showed that sound localization in up/down and front/back dimensions was disturbed by adding spectral ripples to a target source spectrum only at SMFs between 0.5 and 2 c/o, and that the addition of ripples at 1 c/o (which is close to 0.75 c/o) was the most disruptive.

The sound localization task used in the current study differs from the tasks usually employed in sound localization studies. While a traditional absolute localization task allows the listener to indicate freely the position of the auditory target (Gilkey et al., 1995; Makous and Middlebrooks, 1990; Wightman and Kistler, 1989a), we chose instead a forced-choice task which limits the listener's answers. Our goal was not to assess precisely the difference between the listener judgment and the correct target position, but rather to assess the potential relationship between a basic auditory task (spectral-modulation detection) and a more complex auditory task (sound localization). The sound localization performance measured depends on many of known and unknown factors. Those factors could be perceptual but also procedural (Djelani et al., 2000; Wightman and Kistler, 2005). Each of those factors has its own variability which reduces the measured correlation. Therefore, reducing procedural variability by simplifying the sound localization task using a forced-choice task may facilitate the measurement of the correlation of interest.

Recent results have shown that perceptual training can lead to an improvement of the spectral modulation threshold for SMFs below 2 c/o (Sabin et al., 2012). If such training also led to an improvement of sound localization performance, then this would provide stronger proof of a relationship between sensitivity to spectral shape and sound localization performance. Moreover, it would be in favor of a causal relationship and not only a correlational relationship between sensitivity to spectral shape and sound localization performance.

The effects on sound localization in quiet of altering the spectral strength of the HRTFs (Brungart and Romigh, 2009; Brungart et al., 2009; Sabin et al., 2005; Wightman and Kistler, 1997; Zhang and Hartmann, 2010) or of specific modulation frequency components of the HRTFs (Qian and Eddins, 2008) are modest at best. The degree of spectral strength modification in those studies was generally much larger than the observed range of individual differences in spectral strength, which may account for the lack of correlation between spectral strength and sound localization performance in the present study. It seems likely that sound localization performance in noise would be more sensitive to the large changes in

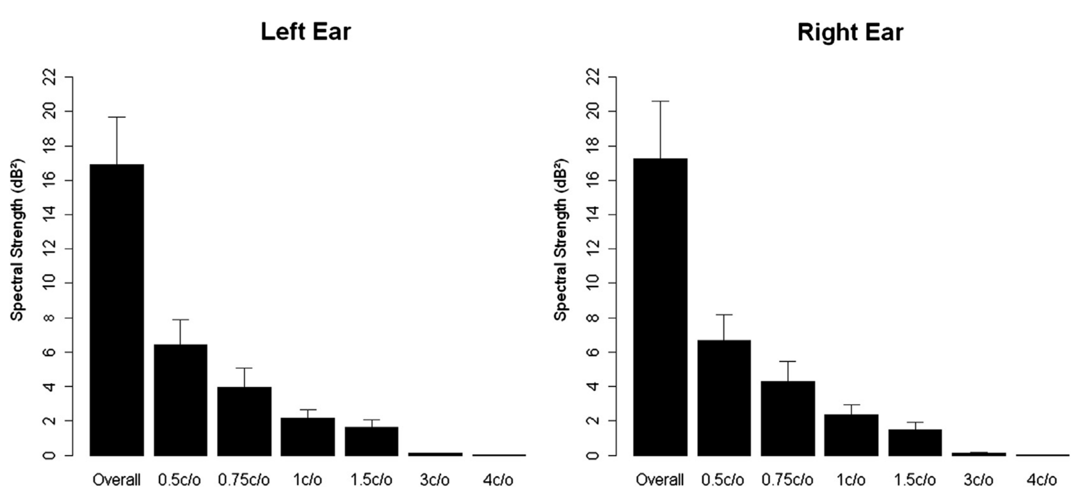

Fig. 5. Mean overall spectral strength and mean spectral strength as a function of spectral modulation frequency. The spectral strengths are computed for the left ear (left) and the right ear (right). Bars show 1 standard deviation.

Table 3
Correlations for each ear between SMF-specific spectral strengths and SLT.

	SMF	Left ear SLT		Right ear SLT	
		rho	*p*	rho	*p*
Spectral strength (left ear/right ear)	0.5 c/o	0.22	0.43	0.12	0.68
	0.75 c/o	0.08	0.80	0.16	0.58
	1 c/o	0.01	0.96	−0.02	0.95
	1.5 c/o	−0.12	0.67	−0.23	0.44

Table 4
Correlations for each ear between SMF-specific spectral strengths and SMTs for high-frequency and broadband carriers.

SMT	Carrier	SMF	Left ear spectral strength		Right ear spectral strength	
			rho	*p*	rho	*p*
	Broadband carrier	0.5 c/o	−0.29	0.29	−0.13	0.64
		1 c/o	−0.03	0.92	−0.16	0.56
		4 c/o	0.22	0.43	0.46	0.09
	High-frequency carrier	0.75 c/o	−0.06	0.76	0.16	0.57
		1.5 c/o	−0.10	0.73	−0.08	0.77
		3 c/o	0.11	0.71	−0.22	0.43

overall or SMF-specific spectral strength employed in those studies than would localization in quiet. Therefore, spectral strength enhancement of SMFs close to 0.75 c/o could improve performance in noise, particularly if tailored to the spectral modulation threshold of each listener.

5. Conclusions

This study was conducted to evaluate the role of sensitivity to spectral shape (a perceptual factor) and the role of HRTF spectral strength (an acoustic factor) in individual differences in sound localization performance. The results indicated that listeners with better sensitivity for spectral modulation frequencies near 0.75 c/o are also better localizers. On the other hand, no relationship was observed between spectral strength per se and sound localization performance. Together those results are in a favor of a perceptual origin rather than an acoustic origin for individual differences in sound localization performance.

Acknowledgments

This work was supported in part by the French Procurement Agency (DGA). Author EM acknowledges funding from the National Science Foundation (USA) and the Natural Sciences and Engineering Research Council of Canada. Author AS acknowledges funding from National Institutes of Health (USA) Grant F31DC9549. We thank Jean Christophe Bouy for help with software development, Lionel Pellieux for HRTF measurements and signal processing manipulations and Brian C.J. Moore, and two anonymous reviewers for many helpful comments.

References

Andeol, G., Guillaume, A., Micheyl, C., Savel, S., Pellieux, L., Moulin, A., 2011. Auditory efferents facilitate sound localization in noise in humans. J. Neurosci. 31, 6759–6763.
Best, V., van Schaik, A., Jin, C., Carlile, S., 2005. Auditory spatial perception with sources overlapping in frequency and time. Acta Acust./Acust. 91, 421–428.
Brungart, D.S., Romigh, G.D., 2009. Spectral HRTF enhancement for improved vertical-polar auditory localization. In: IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 2009. WASPAA '09, pp. 305–308.
Brungart, D.S., Romigh, G.D., Simpson, B.D., 2009. Head related transfer function enhancement for improved vertical polar localization. J. Acoust. Soc. Am. 125, 2691.
Butler, R.A., Belendiuk, K., 1977. Spectral cues utilized in the localization of sound in the median sagittal plane. J. Acoust. Soc. Am. 61, 1264–1269.
Djelani, T., Porschmann, C., Sahrhage, J., Blauert, J., 2000. An interactive virtual-environment generator for psychoacoustic research II: collection of head-related impulse responses and evaluation of auditory localization. Acta Acust./Acust. 86, 1046–1053.
Drennan, W.R., Watson, C.S., 2001. Sources of variation in profile analysis. I. Individual differences and extended training. J. Acoust. Soc. Am. 110, 2491–2497.
Eddins, D.A., Bero, E.M., 2007. Spectral modulation detection as a function of modulation frequency, carrier bandwidth, and carrier frequency region. J. Acoust. Soc. Am. 121, 363–372.
Gilkey, R.H., Good, M.D., Ericson, M.A., Brinkman, J., Stewart, J.M., 1995. A pointing technique for rapidly collecting localization responses in auditory research. Behav. Res. Meth. Instrum. Comput. 27, 1–11.
Good, M.D., Gilkey, R.H., 1996. Sound localization in noise: the effect of signal-to-noise ratio. J. Acoust. Soc. Am. 99, 1108.
Henry, B.A., Turner, C.W., Behrens, A., 2005. Spectral peak resolution and speech recognition in quiet: normal hearing, hearing impaired, and cochlear implant listeners. J. Acoust. Soc. Am. 118, 1111.
Hofman, M., Van Opstal, J., 2003. Binaural weighting of pinna cues in human sound localization. Exp. Brain Res. 148, 458–470.
Levitt, H., 1971. Transformed up-down methods in psychoacoustics. J. Acoust. Soc. Am. 49, 467–477.
Macpherson, E.A., Middlebrooks, J.C., 2003. Vertical-plane sound localization probed with ripple-spectrum noise. J. Acoust. Soc. Am. 114, 430–445.
Macpherson, E.A., Sabin, A.T., 2007. Binaural weighting of monaural spectral cues for sound localization. J. Acoust. Soc. Am. 121, 3677–3688.
Makous, J.C., Middlebrooks, J.C., 1990. Two-dimensional sound localization by human listeners. J. Acoust. Soc. Am. 87, 2188–2200.
Middlebrooks, J.C., 1999a. Individual differences in external-ear transfer functions reduced by scaling in frequency. J. Acoust. Soc. Am. 106, 1480–1492.
Middlebrooks, J.C., 1999b. Virtual localization improved by scaling non-individualized external-ear transfer functions in frequency. J. Acoust. Soc. Am. 106, 1493–1510.
Middlebrooks, J.C., Green, D.M., 1990. Directional dependence of interaural envelope delays. J. Acoust. Soc. Am. 87, 2149–2162.
Møller, H., Sørensen, M.F., Jensen, C.B., HammershØi, D., 1996. Binaural technique: do we need individual recordings? J. Audio Eng. Soc. 44, 451–469.
Qian, J., Eddins, D.A., 2008. The role of spectral modulation cues in virtual sound localization. J. Acoust. Soc. Am. 123, 302–314.
Rife, D.D., Vanderkooy, J., 1989. Transfer-function measurement with maximum-length sequences. J. Audio Eng. Soc. 37, 419–444.
Sabin, A.T., Macpherson, E.A., Middlebrooks, J.C., 2005. Vertical-plane localization of sounds with distorted spectral cues. Assoc. Res. Otolaryngol. Abstr., 257. (Abstract 1389).
Sabin, A.T., Eddins, D.A., Wright, B.A., 2012. Perceptual learning of auditory spectral modulation detection. Exp. Brain Res. 218, 567–577.
Saoji, A.A., Litvak, L., Spahr, A.J., Eddins, D.A., 2009. Spectral modulation detection and vowel and consonant identifications in cochlear implant listeners. J. Acoust. Soc. Am. 126, 955–958.
Shaw, E.A.G., 1974. Transformation of sound pressure level from the free field to the eardrum in the horizontal plane. J. Acoust. Soc. Am. 56, 1848–1861.
Shaw, E.A.G., 1997. Acoustical features of human ear. In: Gilkey, R.H., Anderson, T.R. (Eds.), Binaural and Spatial Hearing in Real and Virtual Environments. Lawrence Erlbaum Associates, Mahwah, NJ, USA, pp. 25–47.
Wenzel, E.M., Wightman, F.L., Kistler, D.J., Foster, S.H., 1988. Acoustic origins of individual differences in sound localization behavior. J. Acoust. Soc. Am. 84, S79.
Wenzel, E.M., Arruda, M., Kistler, D.J., Wightman, F.L., 1993. Localization using non-individualized head-related transfer functions. J. Acoust. Soc. Am. 94, 111–123.
Wightman, F.L., Kistler, D.J., 1989a. Headphone simulation of free field listening. I: stimulus synthesis. J. Acoust. Soc. Am. 85, 858–867.
Wightman, F.L., Kistler, D.J., 1989b. Headphone simulation of free-field listening. II: psychophysical validation. J. Acoust. Soc. Am. 85, 868–878.
Wightman, F.L., Kistler, D.J., 1997. Factors affecting the relative salience of sound localization cues. In: Gilkey, R.H., Anderson, T.R. (Eds.), Binaural and Spatial Hearing in Real and Virtual Environments. Lawrence Erlbaum Associates, Mahwah, NJ, USA, pp. 1–23.
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.
Wightman, F.L., Kistler, D.J., 2005. Measurement and validation of human HRTFs for use in hearing research. Acta Acust./Acust. 91, 429–439.
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, 343–359.
Zhang, P.X., Hartmann, W.M., 2010. On the ability of human listeners to distinguish between front and back. Hear Res. 260, 30–46.