Hearing Research 365 (2018) 28–35

Contents lists available at ScienceDirect

Hearing Research

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

Research Paper

Effect of sound level on virtual and free-field localization of brief sounds in the anterior median plane

Frederic Marmel [a, b, *], Miriam I. Marrufo-Pérez [a, b], Jan Heeren [d], Stephan Ewert [d], Enrique A. Lopez-Poveda [a, b, c]

[a] Instituto de Neurociencias de Castilla y León, Universidad de Salamanca, Calle Pintor Fernando Gallego 1, 37007 Salamanca, Spain
[b] Instituto de Investigación Biomédica de Salamanca, Paseo de San Vicente 58-182, 37007 Salamanca, Spain
[c] Facultad de Medicina, Departamento de Cirugía, Universidad de Salamanca, Avenida Alfonso X "El Sabio" s/n, 37007 Salamanca, Spain
[d] Medizinische Physik and Cluster of Excellence Hearing4All, Universität Oldenburg, 26111 Oldenburg, Germany

ARTICLE INFO

Article history:
Received 18 December 2017
Received in revised form
31 May 2018
Accepted 8 June 2018

Keywords:
Sound level effect
Head-related transfer function
Vertical localization
High-frequency spectral cues
Pinna cues
Spectral discrimination

ABSTRACT

The detection of high-frequency spectral notches has been shown to be worse at 70–80 dB sound pressure level (SPL) than at higher levels up to 100 dB SPL. The performance improvement at levels higher than 70–80 dB SPL has been related to an 'ideal observer' comparison of population auditory nerve spike trains to stimuli with and without high-frequency spectral notches. Insofar as vertical localization partly relies on information provided by pinna-based high-frequency spectral notches, we hypothesized that localization would be worse at 70–80 dB SPL than at higher levels. Results from a first experiment using a virtual localization set-up and non-individualized head-related transfer functions (HRTFs) were consistent with this hypothesis, but a second experiment using a free-field set-up showed that vertical localization deteriorates monotonically with increasing level up to 100 dB SPL. These results suggest that listeners use different cues when localizing sound sources in virtual and free-field conditions. In addition, they confirm that the worsening in vertical localization with increasing level continues beyond 70–80 dB SPL, the highest levels tested by previous studies. Further, they suggest that vertical localization, unlike high-frequency spectral notch detection, does not rely on an 'ideal observer' analysis of auditory nerve spike trains.

1. Introduction

The ability to localize broadband sounds in the median plane has been shown to deteriorate with increasing sound pressure level (SPL) above about 50 dB SPL (Hartmann and Rakerd, 1993; Macpherson and Middlebrooks, 2000; Rakerd et al., 1998; Vliegen and Van Opstal, 2004). This level effect is more apparent for brief sounds (a few tens of milliseconds or shorter) than for longer sounds (80 ms or more; Macpherson and Middlebrooks, 2000; Vliegen and Van Opstal, 2004; but see Rakerd et al., 1998 for a level effect with 1-s white noises). Localization in the vertical dimension is thought to rely on high-frequency spectral peaks and notches generated by the filtering action of the pinna (Asano et al., 1990; Iida et al., 2007; Lopez-Poveda and Meddis, 1996; Macpherson and Sabin, 2013; Middlebrooks and Green, 1991). Presumably, auditory peripheral compression, the broadening of cochlear filters, and/or the saturation of auditory nerve fibers' discharge rates make it harder for the peripheral auditory system to resolve pinna cues at high levels (e.g., Rice et al., 1995; Lopez-Poveda, 1996; Macpherson and Middlebrooks, 2000; Macpherson and Sabin, 2013; Reiss et al., 2011; Sachs and Young, 1979). These notions, however, are controversial.

Alves-Pinto and Lopez-Poveda (2005, 2008) studied the level dependence of the discriminability of broadband noises with and without a rectangular high-frequency notch centered at 8 kHz. They observed that for some participants the ability to detect the high-frequency spectral notch varied non-monotonically with level: just-detectable notch depths were generally smaller for levels below 70 dB SPL and above 80 dB SPL than for levels of 70–80 dB SPL. Insofar as localization in the median plane relies to

* Corresponding author. Present address: Manchester Centre for Audiology and Deafness (ManCAD), Division of Human Communication, Development & Hearing, School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester Academic Health Science Centre, Oxford Road, M13 9PL, UK.

E-mail address: frederic.marmel@gmail.com (F. Marmel).

https://doi.org/10.1016/j.heares.2018.06.004
0378-5955/

some extent on high-frequency spectral notches, the work of Alves-Pinto and Lopez-Poveda suggests that vertical localization of brief broadband sounds may not deteriorate monotonically with increasing level. Instead, the level effect reported previously for levels up to 70–80 dB SPL (Hartmann and Rakerd, 1993; Macpherson and Middlebrooks, 2000; Vliegen and Van Opstal, 2004) could give way to an improvement with increasing level for levels higher than about 80 dB SPL. The present study investigated this possibility. A first experiment (Section 2) collected elevation judgments in the anterior median plane using virtual sound acoustics for short (10 ms) broadband noises at sound levels ranging from 30 to 95 dB SPL. Virtual localization was at chance at 60–80 dB SPL and better than chance at 40–50 and 90 dB SPL, which is consistent with the results of Alves-Pinto and Lopez-Poveda (2005). However, a second experiment (Section 3) using loudspeakers placed in the anterior median plane (free-field sound acoustics) showed that localization deteriorated monotonically with increasing level from 50 to 100 dB SPL for most participants, hence providing evidence against the hypothesis of an improvement for levels higher than about 80 dB SPL. The results are discussed (Section 4) in terms of the difference between the two tasks and the different mechanisms they may involve.

2. Experiment 1. Virtual sound localization with non-individualized head-related transfer functions

2.1. Methods

2.1.1. Participants

Eleven participants (10 females) were tested, but one was excluded because of an inability to perform the task. The remaining 10 participants were between 22 and 34 years old, with a mean age of 26 years. All of them had audiometric thresholds less than or equal to 20 dB HL in both ears at octave frequencies spanning 250–8000 Hz (American National Standards Institute, 2004). Participants were volunteers and were not paid for their service. The human experimentation ethical review board of the University of Salamanca approved all procedures.

2.1.2. Stimuli

On each trial, the stimulus was a pair of identical 10-ms Gaussian noise bursts separated by a 500-ms silent interval. For each pair, a fresh Gaussian noise was generated in the time domain by drawing pseudorandom sample values from the standard normal distribution (441 samples, corresponding to 10-ms noises generated at a 44100-Hz sampling rate). The sample amplitudes were adjusted to get a root mean square (RMS) level equal to the desired level, and 2-ms onset and offset raised-cosine ramps were applied. The resulting noise was duplicated and the two noise segments concatenated with a 500-ms silence between them. Silent 500-ms intervals were added before and after the pair of noises. Out-of-the-head vertical location was simulated by convolving the noises with head-related impulse responses (HRIRs) for a KEMAR (Knowles Electronics model DB-4004 with the “small” pinna model DB-061), as measured by Gardner and Martin (1995). The “full” HRIRs (containing the response of the measurement system and the KEMAR ear canal resonance) were used rather than the equalized “diffuse” versions, because localization seemed easier to the first two authors this way. Fourteen virtual vertical sources were simulated in the anterior median plane, with elevation angles spanning -40° to 90° in 10°-steps. Fig. 1 shows the frequency spectra of the HRIR for each elevation angle, hereinafter referred to as head-related transfer functions (HRTFs). The noises had levels of 30, 40, 50, 60, 70, 80, 90, and 95 dB SPL.

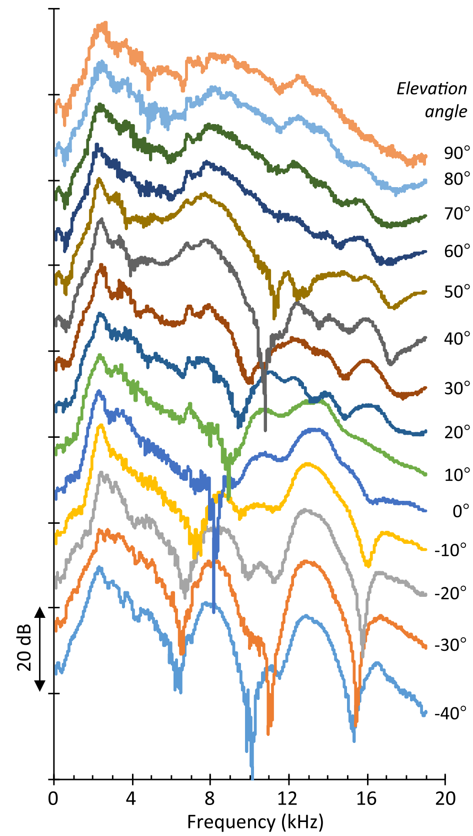

Fig. 1. Frequency spectra of the KEMAR head-related impulse responses used in Experiment 1. Each trace is for a different elevation angle (in degrees), as indicated by the numbers to the right of the trace. To minimize clutter, an arbitrary vertical shift of 10 dB has been introduced across traces.

2.1.3. Apparatus

All stimuli were generated digitally using custom Matlab software (The Mathworks, Natick, Massachusetts, USA). Stimuli were digital-to-analog converted using an RME Fireface 400 sound card at a sampling rate of 44100 Hz and a resolution of 24 bits, and presented bilaterally via ER2 insert earphones (Etymotic Research, Inc., Elk Grove Village, IL, USA). The ER2 insert earphones are designed to give a flat frequency response at the eardrum from 250 to 8000 Hz. Participants sat in a double-wall sound booth during testing. Stimulus level (in dB SPL) was specified relative to the acoustic sound level of a 1-kHz digital sinusoidal wave with maximal digital amplitude played via the right ER2 earphone. This calibration value was measured by fitting the right-ear ER2 in a Zwislocki coupler (Knowles DB-100, Knowles Electronics Inc., Itasca, IL, USA) connected to a sound level meter (B&K 2238, Brüel & Kjær, Naerum, Denmark).

2.1.4. Procedure

The experimental procedure was controlled via custom Matlab software. Each participant completed four sessions of eight blocks (corresponding to the eight target levels). Each session started with the 60-dB-SPL block followed by the 50-, 40-, 30-, 70-, 80-, 90-, and 95-dB-SPL blocks, in that order. Within each block, 14 stimuli corresponding to the 14 input angles were presented randomly. The computer screen displayed a picture of a head in a circle with 27 buttons representing potential virtual sources between -40° and 90°, in 5° steps. Upon each stimulus presentation, participants had to judge the elevation of the sound by clicking on one of the buttons. The participant's response (no time limit) triggered the