Vision-based Adaptation of the Frequency-dependent Weighting of the Localization Cues

Ondrej Spišák[1], René Šebeňa[1], Peter Lokša[1], Maike Ferber[2], Bernhard Laback[2], Norbert Kopčo[1]

[1]Institute of Computer Science P.J. Šafárik University, Šrobárova 2, 041 80 Košice, Slovakia
[2]Acoustic Research Institute, Austrian Academy of Science, Wohllebengasse 12-14, 1040 Vienna, Austria
ondrej.spisak@student.upjs.sk, rene.sebena@upjs.sk, peter.loksa@student.upjs.sk, mferber@kfs.oeaw.ac.at, Bernhard.Laback@oeaw.ac.at, norbert.kopco@upjs.sk

Abstract

Which cues the auditory system uses to determine the sound source location largely depends on the sound's frequency content. For low-frequency (LF) narrowband sounds, the interaural time difference (ITD) is the dominant cue, while for high-frequency (HF) narrowband sounds, the interaural level difference (ILD) dominates. For mid-frequency narrowband sounds, ITD and ILD both contribute to varying degrees to determining the perceived location. We performed an experiment in which we tested whether it is possible to change spectral weighting of either the HF or the LF components of broadband stimuli by visually guided training in separate subject groups. We also tested whether this reweighting would generalize to a change in the ITD/ILD weighting for mid-frequency sounds. In the subject group trained on HF this training resulted in an increase in the HF weight, but no effect was found in the LF group. However, the change in spectral weighting of the HF group did not generalize to an increase in the relative weighting of the ILD cue for mid-frequency sounds. Thus, the reweighting appears to be only spectral, but not binaural-cue specific.

1 Introduction

Spatial hearing is, from an evolutionary point of view, a very important element by which we can determine the location of a sound source before it is captured visually. This ability has a very important warning function, and it also facilitates everyday life as it allows us to, for example, understand speech in a noisy environments or navigate in space.

Localization of the sound in the horizontal plane is mainly based on two physical parameters, the interaural time difference (ITD) and interaural level difference (ILD), expressing, respectively, the difference in the arrival time and level with which a sound reaches one vs. the other ear. Previous experiments showed that changing the weighting of ITD and ILD to determine the location of a sound source in virtual environment is possible [1] while not always successful [2]. Here, we examined whether it is possible to change the spectral weighing of high vs. low components of sound in real environment, and whether that reweighting would generalize to a change in binaural cue weighing.

2 Methods

The experiment consisted of 3 parts: pretest (performed in virtual reality, VR, and in real environment), training (in real environment), and posttest (identical to pretest) In the VR environment, 1-octave noises with center frequency of 2.8 kHz were used as stimuli. ITD and ILD were independently manipulated to correspond to one of 40 different positions in range from -70.2° to 70.2°, with an inconsistency between ILD and ITD positions of up to 25.2°. The subject's task was to localize a sound by performing a head-turn towards it. No feedback was provided.

In the real environment 11 speakers were spread in the range from -56° to 56° (11.25° spacing). Stimuli consisted of 0.5-octave noise bands centered at high-frequency (HF; 11.2kHz, 5.6kHz), low-frequency (LF; 0.7kHz, 0.35kHz) and medium frequency (MF; 2.8 kHz). 3 types of stimuli were presented: 1) 2-channel stimulus: 1 HF and 1 LF channel from locations separated by 1 or 2 speakers, 2) 4-channel stimulus: 2 HF and 2 LF channels from locations 1-2 speakers apart, and 3) 2-channel stimulus with MF (only used during testing): 1 MF channel and 1 channel at HF or LF, one speaker apart. The subject's task was the same as in VR. After the pretest, participants performed 3 days of training in real environment. The subject's task was the same as in the pretest, with the difference that after confirming the sound position, feedback was provided at the position of either the HF speaker (for HF subject group; 13 subjects) or the low-frequency speakers (for LF group; 12 subjects). Subjects were naïve as to the spectral structure of the stimuli or what the feedback was reinforcing. They were instructed to imagine the sound as coming from the feedback location, and to correct their responses accordingly.

3 Results

The results were analyzed by computing the bias in responses towards LF-components (re. HF-components), and a weight defined as

$$LFweight = (response - HFaz)/(LFaz - HFaz)$$

where HFaz and LFaz represent the location of the two frequency components (in degrees). Figs. 1 and 2 show the LF biases and LFweights for the 2-ch stimuli, separately for the two training groups, for the pretest and posttest, and for the two speaker separations. The HF group's LFweight in the posttest is significantly lower than in the pretest, while no effect of training is observed for the LF group. This effect generalized to the untrained 2-ch stimuli with MF (data not shown).

Figs. 3 and 4 show the results of the VR testing, expressed as the biases towards the ITD location as a function of the separation of the ITD and ILD components, separately for the two training groups, and for the pretest and posttest. No significant difference between the two groups was observed, even though a change in performance from pre- to posttest is observed for both groups. This means that spectral reweighting does not generalize to ITD/ILD reweighting in VR.

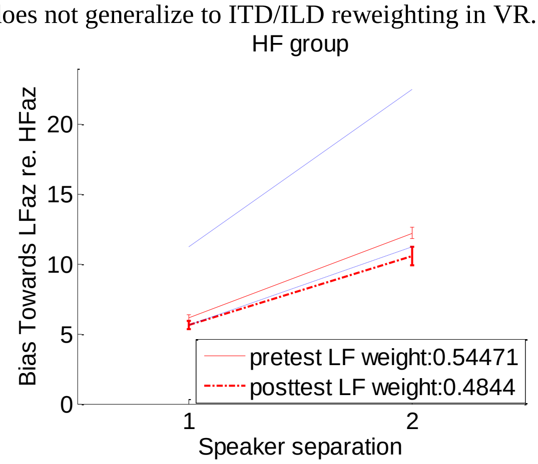

Fig. 1: HF group in real environment.

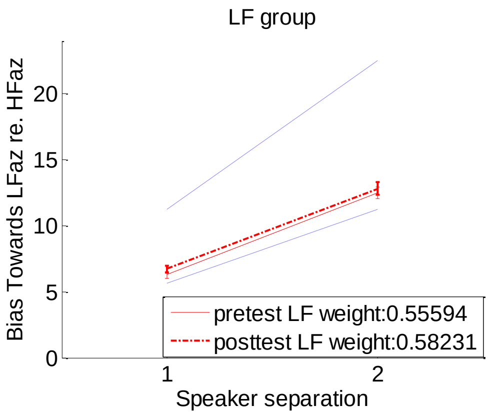

Fig. 2: LF group in real environment.

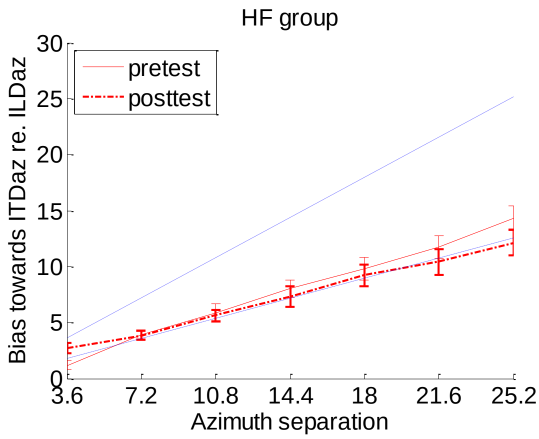

Fig. 3: HF group in virtual environment.

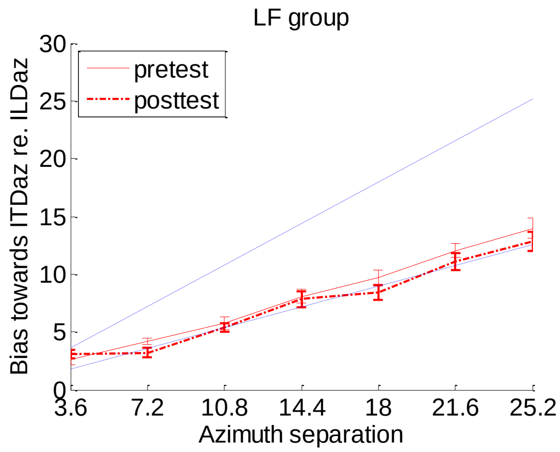

Fig. 4: LF group in virtual environment.

4 Summary

The results show that it is possible to change the weighing with which individual spectral components contribute to sound localization. However, only the HF-training was successful, possibly because the LF components are strongly weighted already in the pretest. Also, the HF training did not generalize to a change in the ITD/ILD weight, suggesting that it is spectrum-specific. However, other differences between the real and VR stimuli, like the presence of reverberation in the real environment, might also have played a role.

Acknowledgment

This work was supported by Danube Partnership project APVV DS-2016-0026, H2020-MSCA-RISE-2015 #691229.

Literature

[1] Ferber M (2018) Plasticity of Spatial Processing in Normal Hearing: Reweighting of Binaural Cues. Unpublished MSc. Thesis. University of Vienna.

[2] Jeffress LA, & McFadden D (1971). Differences of interaural phase and level in detection and lateralization. J Acoust Soc Am, 49(4B), 1169-1179.