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

INTRODUCTION

- Sound localization in the horizontal plane is largely determined by two physical cues: interaural time difference (ITD) and interaural level difference (ILD).
- For low-frequency (LF) sounds, the interaural time difference (ITD) is the dominant cue, while for high-frequency (HF) sounds, the interaural level difference (ILD) dominates.
- Previous experiments showed that changing the weighting of ITD and ILD while determining the location of a sound source is not always successful (Jeffress & McFadden, 1971), but possible in virtual environment (Ferber, Laback, & Kopčo, 2018).
- Here we examined whether it is possible to change the spectral weighting of high (HF) vs. low (LF) components of sound in real environment.

HYPOTESIS

1. **Change of weighting of HF vs. LF components of sound in real environment will occur as result of training reinforcing either of the two components.**
2. **The reweighting of spectral components will generalize to**
   - **new (untrained) spectral components,**
   - **a change of ITD/ILD weighting in virtual environment.**

METHODS

Design: **Day 1:** Virtual Environment (VE) Pre-training – VE Pretest – Real Env. (RE) Pretest
Day 2-4: RE Training, on **Day 4** followed by RE Posttest & VE Posttest

Participants: 25 subjects with normal hearing, 2 experimental groups:
LF target group (13 S): Trained on low-frequency components
HF target group (12 S): Trained on high-frequency components

Apparatus: VE setup - VR glasses, headphones
RE setup - electromagnetic tracking, loudspeakers, LED projector

Stimuli: **VE**: Narrow-band, 1-octave noises with center frequency of 2.8 kHz; independent combinations of ITD and ILD in range from -70.2° to 70.2°, with an inconsistency between ILD and ITD positions of up to 25.2°

RE: 0.5-octave noise bands centered at HF (11.2 or 5.6 kHz), LF (0.7 or 0.35 kHz) and medium frequency (MF; 2.8 kHz). 3 types of stimuli: 2-channel (1 HF & 1 LF), 4-channel (2 HF & 2 LF), 2-channel stimulus with MF (1 MF & (1 LF or 1 HF)); 11 speakers were spread in the range from -56° to 56° (11.25° spacing) with an inconsistency between positions of HF and LF component up to 22.5°. Visual feedback projected on top of speakers.

Task: - Pretest/Posttest: localize a sound by performing a head-turn towards it.
- Training:

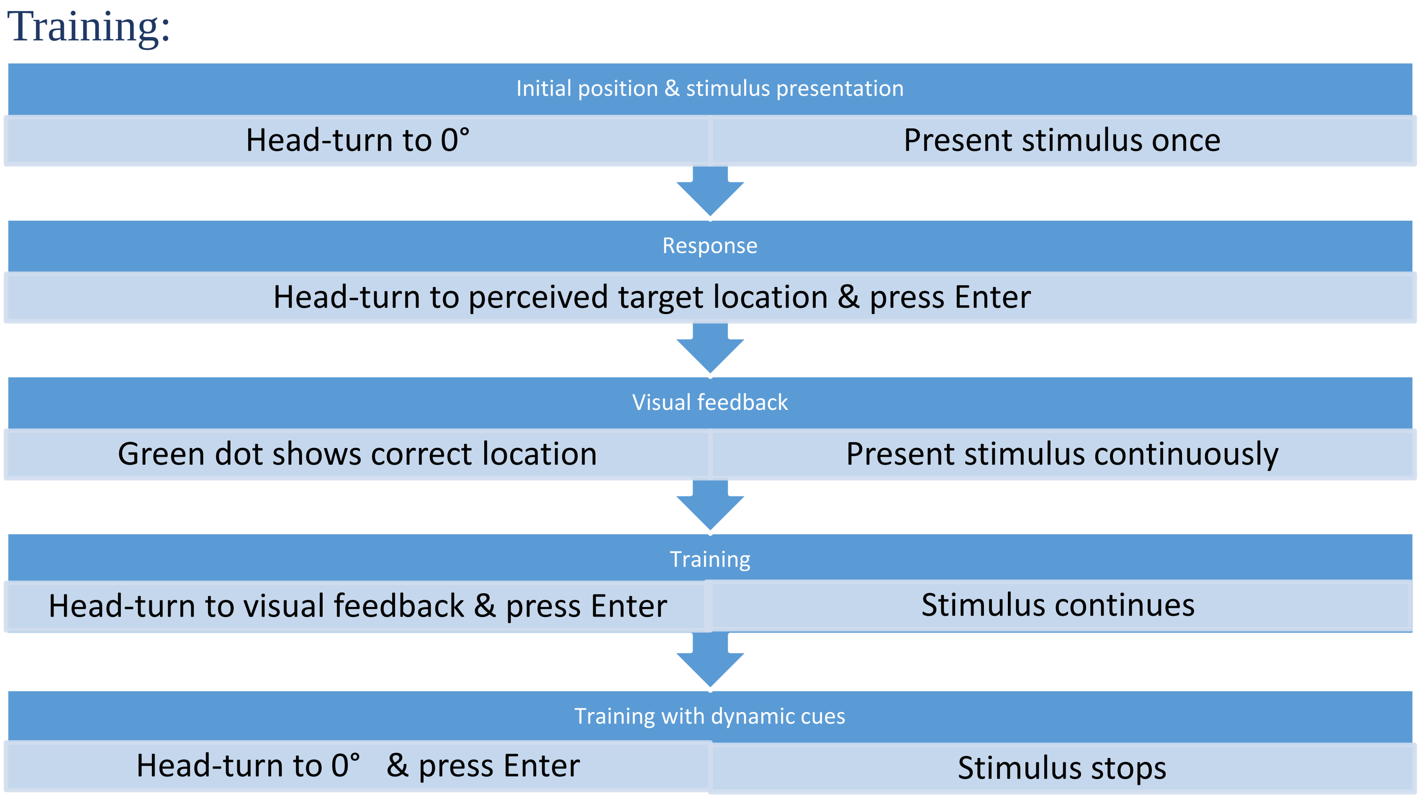

Figure 1. Training procedure.

Analysis using a regression model fitted separately for each target azimuth α.

RE model : $R(\alpha, \Delta_{LF}, \Delta_{HF}) = k_{LF}(\alpha) * \Delta_{LF} + k_{HF}(\alpha) * \Delta_{HF} + Q(\alpha); w_{HF} = \frac{atan\left(\frac{k_{HF}(\alpha)}{k_{LF}(\alpha)}\right)}{90}$

R is a subject's response azimuth in a trial with LF and HF components at positions $\alpha + \Delta_{LF}$ and $\alpha + \Delta_{HF}$, respectively (α is between -56.25° and 56.25° with 11.25° steps). k_{LF}, k_{HF} and Q are approximated parameters of a regression model, where k_{LF} and k_{HF} are regression slopes (determining the weights of the frequency components) and Q is the overall bias for azimuth α.

VE model : $R(\alpha, \Delta_{ITD}, \Delta_{ILD}) = k_{ITD}(\alpha) * \Delta_{ITD} + k_{ILD}(\alpha) * \Delta_{ILD} + Q(\alpha); w_{ILD} = \frac{atan\left(\frac{k_{ILD}(\alpha)}{k_{ITD}(\alpha)}\right)}{90}$

R is a subject's response azimuth in a trial where the ITD and ILD corresponded to positions $\alpha + \Delta_{ITD}$ and $\alpha + \Delta_{ILD}$, respectively (α is between -45° and 45° with 3.6° steps). k_{ITD}, k_{ILD} and Q are approximated parameters of a regression model, where k_{ITD} and k_{ILD} are regression slopes (determining the binaural cue weights) and Q is overall bias for azimuth α.

$\boldsymbol{w_{HF}}$ and $\boldsymbol{w_{ILD}}$ are, respectively, estimated weights of HF vs. LF and ILD vs. ITD components.

RESULTS

REAL ENVIRONMENT

- The HF group's HF weight in the posttest is significantly higher than in the pretest ($F(1,12) = 5.98$;$p = 0.0308$), while no effect of training is observed for the LF group ($F(1,11) = 0.65$;$p=0.4367$).

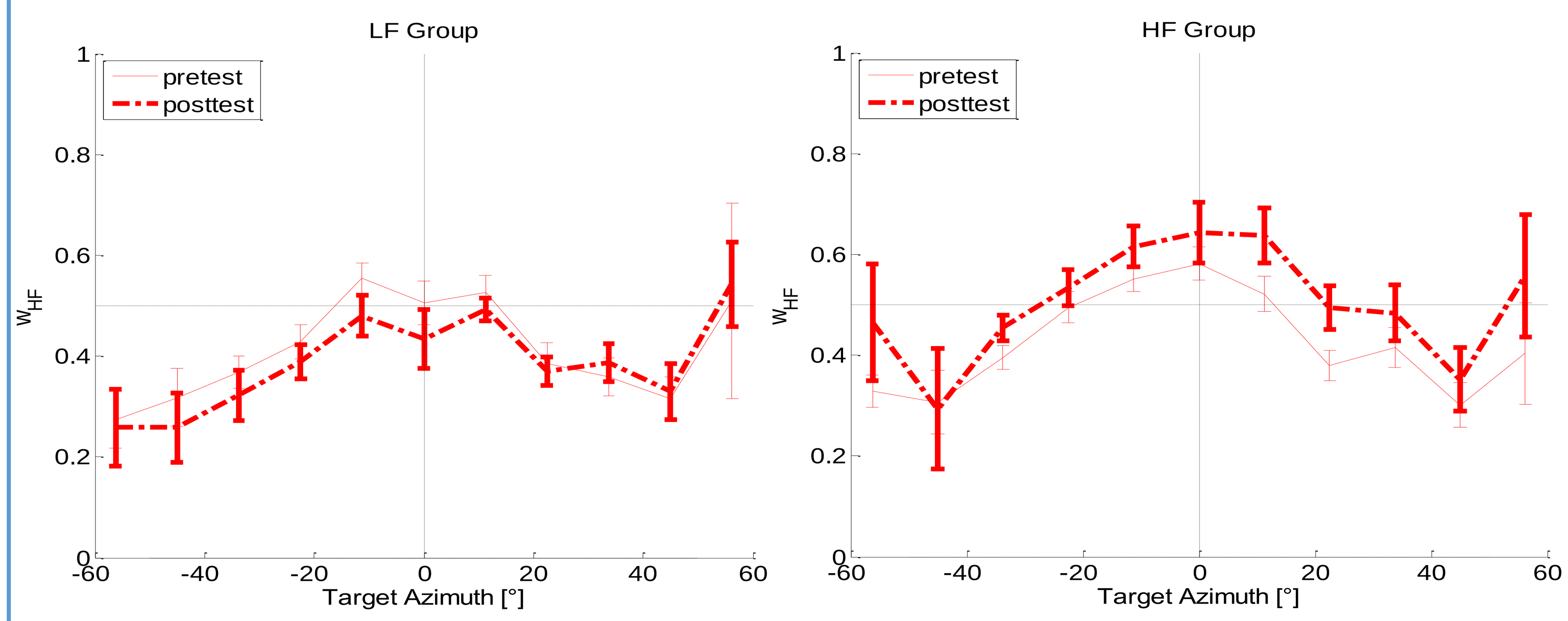

Figure 2. Pretest-Posttest comparison of HF weights for 2-channel sounds in real environment as a function of target azimuth. Both graphs show across-subject means (±SEM).

- HF training generalizes to sounds consisting of trained frequencies and a new 2.8-kHz component, but only for trained low-frequency components (0.35-0.7 kHz).

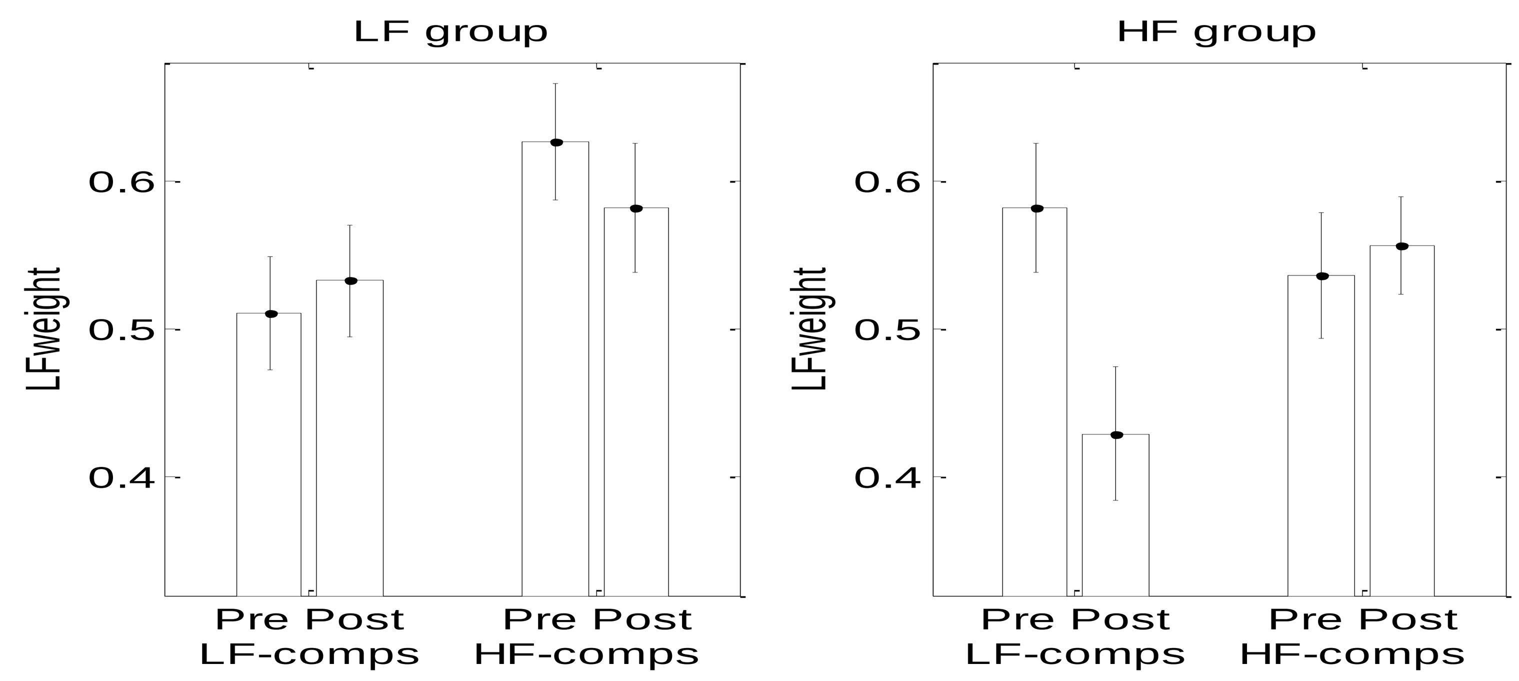

Figure 3. Pretest-Posttest comparison of LF weights in real environment for trials with new component(2.8-kHz).

VIRTUAL ENVIRONMENT

- No significant group-specific change in ILD weights was observed, even though an increase in ILD weights from pre- to posttest is observed for both groups ($F(1,22) = 13.02$; $p = 0.0016$).

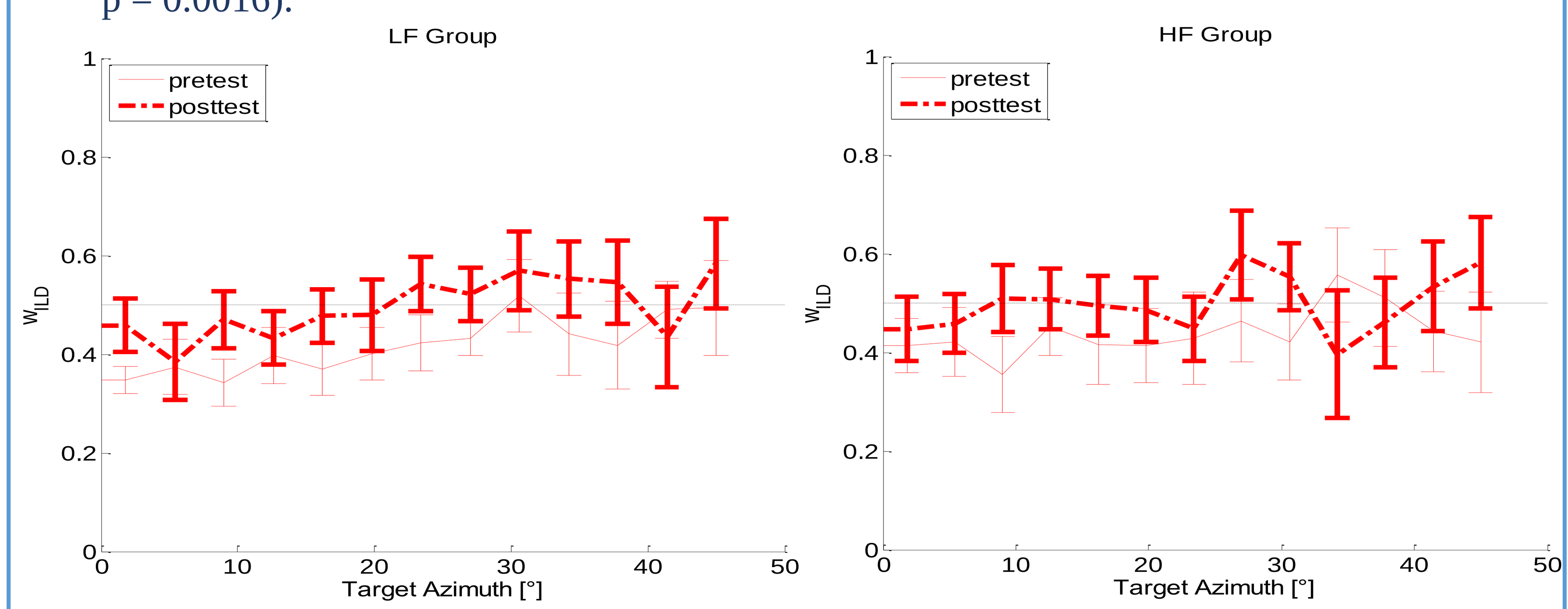

Figure 4. Pretest-Posttest comparison of ILD weights in virtual environment as a function of target azimuth. Data were collapsed around y-axis, assuming left-right symmetry.

CONCLUSIONS

- The results show that it is possible to change the weighting with which individual spectral components contribute to sound localization in a real environment.
- The HF-training was successful but LF-training not, possibly because the LF components are strongly weighted already in the pretest. The HF-training generalized to the untrained frequencies (2-channel stimuli with MF), but only when combined with LF components.
- The HF-training did not generalize to a group-specific change in the ITD/ILD weighting, 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 AND REFERENCES

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

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

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