Exploring influences on auditory selective attention by a static and a dynamic binaural reproduction

Josefa Oberem[1], Julia Seibold[2], Iring Koch[2] Janina Fels[1]
[1] Institute of Technical Acoustics, Medical Acoustics Group, RWTH Aachen University, job@akustik.rwth-aachen.de
[2] Institute of Psychology, RWTH Aachen University

Introduction

The topic of the present collaborative project was the exploration of cognitive control mechanisms underlying auditory selective attention. The aim was to examine the influence of variables that increase the complexity of the auditory scene with respect to technical aspects and that influence the efficiency of cognitive processing. As one step towards realistic environments, the dynamic binaural reproduction with consideration of head movements was compared to a static binaural reproduction. To reproduce a realistic auditory scene it has often been proven that a dynamic reproduction is more beneficial compared to a static reproduction. It was reported about a better localization accuracy, significantly less front-back-reversals as well as an impressing plausibility [1, 2]. On account of these findings, it was assumed that a dynamic reproduction would yield to more beneficial results compared to the static reproduction.

Method

Subjects

A number of 24 untrained students (mean age: 26.5 years) participated in the experiment. Subjects were equally divided into male and female listeners and normal-hearing (within 20 dB).

Experimental Task

The developed paradigm, firstly introduced by Koch et al. [3] and extended to a binaural paradigm by Oberem et al. [4, 5], was to analyze the intentional switching in auditory selective attention in realistic environments. It consisted of two simultaneously presented stimuli by two different speakers of opposite sex. One speaker acted as the distractor and the other acted as the target. By a visual cue on a monitor the target-speaker's direction was cued in advance. There were eight different positions equally distributed on the horizontal plane. Speech of target and distractor were never presented from the same direction, but always simultaneously. The used stimuli were spoken digits from 1 to 9, excluding 5 combined with a direction word "oben" or "unten". Recordings were anechoic, loudness adjusted and shortened to 1200 ms. The listener's task was to divide the target's speech into four categories (smaller than 5 (< 5) vs. greater than 5 (> 5), in combination with "oben" or "unten"). The four stimulus categories were mapped to four response buttons on a controller, held in both hands. The procedure of a trial is also depicted in Figure 1.

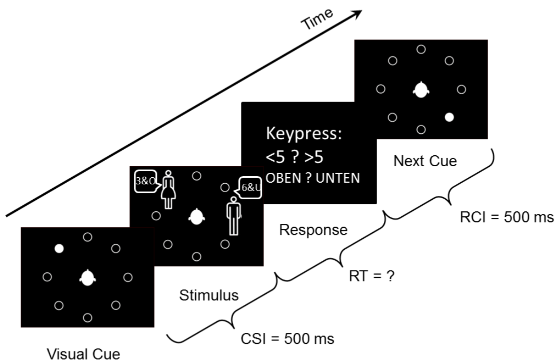

Figure 1: Procedure of a trial with a visual cue indicating the target direction, a cue-stimulus-interval (CSI) of 500 ms, the synchronous presentation of the stimuli, reaction time between onset of stimulus and the response of the subject, and the response-cue-interval (RCI) of 500 ms.

Results

The collected data was submitted to a 3-way-ANOVA with the variables of reproduction method (RepMeth), position (Pos) and attention switch (AttSw). The nonindividual binaural stimuli were either reproduced statically or dynamically. For the dynamic reproduction, a data set of $1° \times 1°$ HRTFs was used. The variable position described the spatial position of the target speaker in space within one trial. The variable had three levels: median plane (front, back), frontal plane (left, right) and diagonal positions (positions shifted by 45° to those of the other levels). Attention Switch described the effect whether the target's spatial position was repeated from one trial to another (e.g. front - front) or switched between trials (e.g. left - back).

Reaction time

The ANOVA yielded no significant main effect of reproduction method [*RepMeth*: $F(1, 23) = 1.14$, $p > .05$] in reaction times. The main effect of position was significant [*Pos*: $F(1.38, 31.7) = 37.95$, $p < .001$], indicating higher reaction times when the target speaker was positioned in median plane than for target's positions in frontal plane (For the target's position, Mauchly's test indicated that the assumption of sphericity is violated $\chi^2(2) = 15.44, p < .05$, therefore the degrees of freedom are corrected using Huynh-Feldt estimates of sphericity ($\epsilon = 0.69$).) (c.f. Fig. 2). Furthermore,

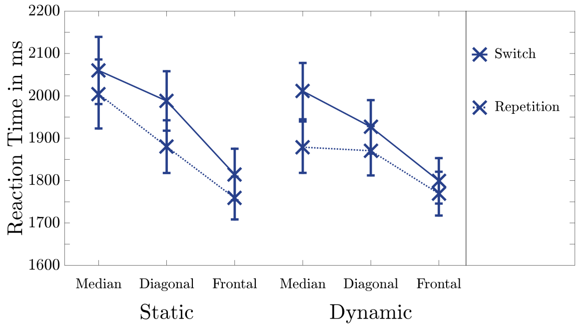

Figure 2: Reaction time (in ms) as a function of reproduction method, position and attention switch (RepMeth×Pos×AttSw). Error bars indicate standard errors.

a significant interaction with the reproduction method was found [*RepMeth*×*Pos*: $F(2,46) = 5.55$, $p < .05$]. The third main effect of attention switch was also significant [*AttSw*: $F(1,23) = 24.68$, $p < .001$], indicating smaller reaction times for repetitions of the target's position than switches. The attention switch effect did not significantly interact with the reproduction method [*RepMeth*×*AttSw*: $F < 1$]. There was also no interaction between the position and the attention switch [*Pos*×*AttSw*: $F(2,46) = 2.04$, $p > .05$] as well as no significant three-way-interaction of all variables [*RepMeth*×*Pos*×*AttSw*: $F(2,46) = 2.49$, $p > .05$].

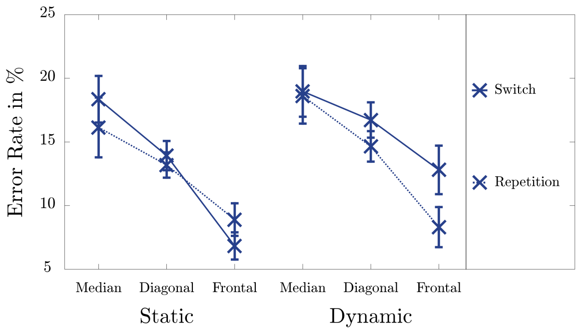

Figure 3: Error rate (in %) as a function of reproduction method, position and attention switch (RepMeth×Pos×AttSw). Error bars indicate standard errors.

Error rate

The ANOVA yielded a significant main effect of reproduction method [*RepMeth*: $F(1,23) = 12.22$, $p < .05$] in error rates, indicating greater error rates when stimuli where reproduced in a dynamic setup. The main effect of position was also significant [*Pos*: $F(1.37, 31.6) = 19.60$, $p < .001$], indicating the same pattern as in reaction times (For the target's position, Mauchly's test indicated that the assumption of sphericity is violated $\chi^2(2) = 15.80, p < .05$, therefore, the degrees of freedom are corrected using Huynh-Feldt estimates of sphericity ($\epsilon = 0.69$).) (c.f. Fig. 3). No significant interaction with the reproduction method was found [*RepMeth*×*Pos*: $F < 1$]. The third main effect of attention switch was also significant [*AttSw*: $F(1,23) = 7.58$, $p < .05$], indicating smaller error rates for repetitions of the target's position than switches. The attention switch effect did not significantly interact with the reproduction method [*RepMeth*×*AttSw*: $F(2,46) = 2.57$, $p > .05$]. There was also no interaction between the position and the attention switch [*Pos*×*AttSw*: $F < 1$]. However, there was a significant three-way-interaction of all variables [*RepMeth*×*Pos*×*AttSw*: $F(2,46) = 5.61$, $p < .05$].

Discussion and Conclusion

The present investigation examined how a static and a dynamic reproduction affected the intentional switching in auditory selective attention. Surprisingly, results are contradictory: On the one hand, in reaction times, there was a significant improvement in median plane for the dynamic reproduction, possibly due to the frequently confirmed effect of diminishing front-back-confusions in dynamic reproductions [1]. On the other hand, in error rates, the dynamic reproduction yields significantly greater error rates than the static reproduction. In a second step, the performed head movements were analyzed and found to be mostly smaller than 1° for both reproduction methods, indicating very few changes of HRTFs in the dynamic reproduction. Therefore, the dynamic reproduction should rather be cosidered as a quasi-static reproduction and should not yield to any significant differences with the static reproduction. In further experiments, these inconsistencies are to be analyzed.

References

[1] P. Minnaar, J. Pedersen: Evaluation of a 3D-audio system with head tracking 120th Audio Engineering Society Convention. Paris, France, 20.-23. May (2006).

[2] D. Begault, E. Wenzel, A. Lee, M. Anderson: Direct Comparison of the Impact of Head Tracking, Reverberation, and Individualized Head-Related Transfer Functions on the Spatial Perception of a Virtual Speech Source. 108th Audio Engineering Society Convention. Paris, France, 19.-22. February (2000).

[3] I. Koch, V. Lawo, J. Fels, M. Vorländer: Switching in the cocktail party: Exploring intentional control of auditory selective attention. Journal of Experimental Psychology / Human Perception and Performance **37** (2011) 1140-1147.

[4] J. Oberem, V. Lawo, I. Koch, J. Fels: Intentional switching in auditory selective attention: Exploring different binaural reproduction methods in an anechoic chamber. Acta Acustica united with Acustica **100** (2014) 1139-1148.

[5] J. Fels, J. Oberem, I. Koch: Examining auditory selective attention in realistic, natural environments with an optimized paradigm. Proceedings of Meetings on Acoustics **28** (2016).