Evaluation of experiments on auditory selective attention in an anechoic environment and a reverberant room with nonindividual binaural reproduction

Josefa Oberem[1], Vera Lawo[2], Iring Koch[2], Janina Fels[1,3]

[1] *Institute of Technical Acoustics, Medical Acoustics Group, RWTH Aachen University, job@akustik.rwth-aachen.de*

[2] *Institute of Psychology, RWTH Aachen University*

[3] *Research Centre Jülich, Institute of Neuroscience and Medicine*

Introduction

When auditory selective attention started to be in the interest of acoustical and psychological research dichotic-listening-paradigms were frequently used, but still today many psychological experiments on the cocktail-party effect are performed using a dichotic reproduction [1]. For more realistic scenes and a better spatial impression paradigms should be binaural. In previous work a dichotic-listening-paradigm to study intentional switches in auditory selective attention was extended to a binaural listening test setup [2]. As a next step towards realistic cocktail-party-scenes room acoustic needs to be included in the auralization. In this study the effect on auditory selective attention by the reverberation time of two different rooms (anechoic room vs. "acoustically good" seminar room) is examined.

Focusing one's attention on a target speaker and following his/her speech requires and is dependent from for example the localization performance as well as a the speech intelligibility. Localization in reverberant environments and anechoic rooms was studied and compared by Giguère and Abel [3] resulting in a worse localization performance for greater reverberation times. However, Rychtáriková et al. [4] observed a better localization performance in reverberant rooms when nonindividual HRTFs are used for a binaural synthesis than in anechoic rooms. Furthermore, speech intelligibility is negatively affected by the reverberation time [5]. On account of these findings, it is not clear whether the intentional switching in auditory selective attention is negatively affected by a reverberant environment.

Method

Experimental Task

The developed paradigm, firstly introduced by Koch et al. [1], to analyze the intentional switching in auditory selective attention consists of two simultaneously presented stimuli by two different speakers of opposite sex. One speaker acts as the distractor and the other acts as the target. By a visual cue on a monitor the target-speaker's direction is cued in advance. There are eight different positions equally distributed on the horizontal plane. Speech of target and distractor are never presented from the same direction, but always simultaneously. Hence, there are 56 combinations of the target's and the distractor's position. The used stimuli are spoken digits from 1 to 9, excluding 5. Recordings are anechoic, loudness adjusted and shorten to 730 ms. The listener's task is to categorize the target's speech into two categories (smaller than 5 (< 5) vs. greater than 5 (> 5)). The two stimulus categories are mapped to two response buttons, held in hand, to be pressed by the left and right thumb. The procedure of a trial is also depicted in Figure 1.

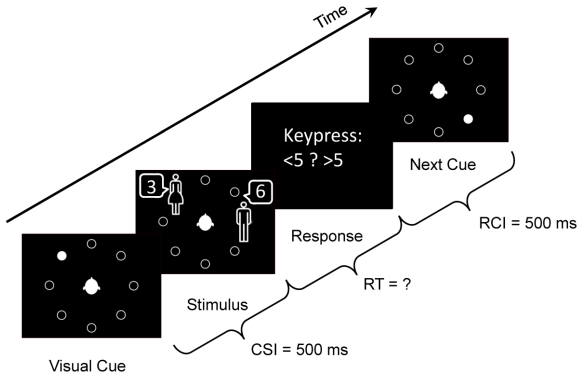

Figure 1: Procedure of a trial with a visual cue indicating the target direction, a cue-stimulus-interval (CSI) of 500 ms, the synchronous presentation of the stimuli, reaction time between onset of stimulus and the response of the subject, and the response-cue-interval (RCI) of 500 ms.

Room Conditions

To analyze the influence of reverberant environments the experiment is carried out in "two different rooms". For both conditions a binaural synthesis is used which is presented via headphones (*Sennheiser HD 600*). Therefore, the experiment actually takes place in one quite test room, but the reproduced stimuli are virtually affected by different reverberation times.

For the room condition "Anechoic" free-field HRTFs measured with an artificial head (ITA) are used to synthesize the binaural stimuli. This condition is to be compared to the condition of "Roomacoustics", where the named HRTFs are also convolved with BRIR, which are simulated in RAVEN. The simulated room with a volume of 137 m^3 is designed after DIN 18041 [6] and appropriate for speech. Therefore, the frequency dependent reverberation time varies between $T = 0.4\,\mathrm{s}$ and $T = 0.8\,\mathrm{s}$. An example would be an acoustically good seminar room.

Subjects

A number of 48 students aged between 18 and 35 with tested normal-hearing participated in the experiment and were paid (8 Euros). The experiment lasted for an hour

including the audiogramm and a training preceeding the real experiment.

Results

The collected data is submitted to a 3-way-ANOVA with the variables of room configuration (R), transition (T) and congruency (C).
The two room conditions are described above.
Transition describes the effect whether the targets spatial position is repeated from one trial to another (e.g. front - front) or switched between trials (e.g. left - back).
The variable congruency refers to the stimuli of target and distractor. The variable has two different levels. On the one hand, the two stimuli can be congruent, which is the case when both digits are smaller than 5 or both greater than 5 (e.g. 2 and 4, 6 and 9) and on the other hand, the two stimuli can be incongruent, which is the case when one digit is smaller and one greater than 5 (e.g. 1 and 7, 8 and 3).

Reaction time

The ANOVA yields no significant main effect of room configuration (R) $F < 1$ in reaction times. The main effect of transition (T) is significant $F(1, 46) = 61.10$, $MSE = 31,933$, $p < 0.001$, $\eta_p^2 = 0.571$, indicating a higher reaction time for switches than for repetitions (c.f. Fig. 2). The third main effect of congruency is also significant $F(1, 46) = 20.80$, $MSE = 98,555$, $p < 0.001$, $\eta_p^2 = 0.311$, indicating smaller reaction times for congruent trials than for incongruent trials. All interactions of the given variables are not significant.

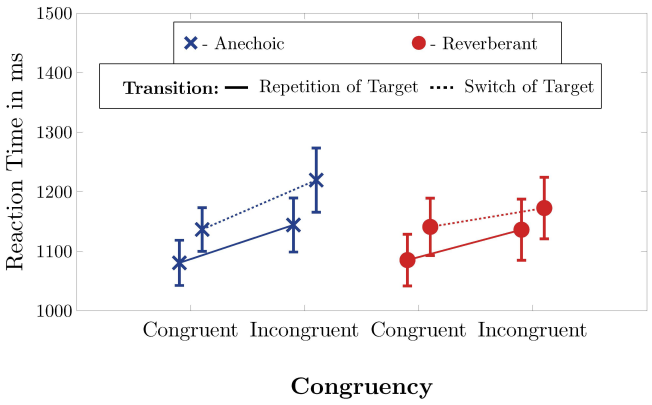

Figure 2: Reaction time (in ms) as a function of room configuration, transition and congruency (RxTxC). Error bars indicate standard errors.

Error rate

The ANOVA yields no significant main effect of room configuration (R) $F(1, 46) = 2.83$, $MSE = 0,000$, $p > 0.05$, $\eta_p^2 = 0.058$ in error rates. The main effect of transition (T) is also not significant $F < 1$ (c.f. Fig. 3). However, the third main effect of congruency is significant $F(1, 46) = 309.31$, $MSE = 0,000$, $p < 0.001$, $\eta_p^2 = 0.871$, indicating smaller reaction times for congruent trials than for incongruent trials. All interactions of the given variables are not significant.

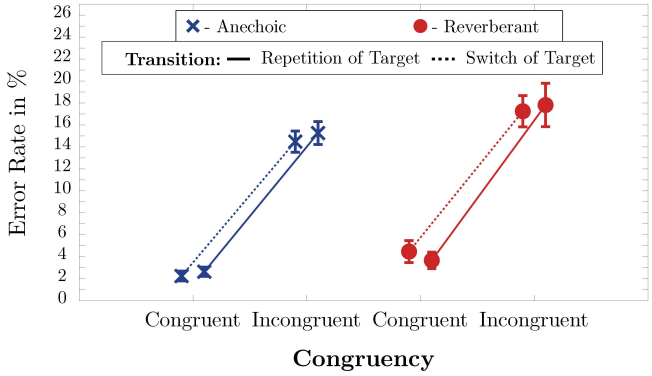

Figure 3: Error rate (in %) as a function of room configuration, transition and congruency (RxTxC). Error bars indicate standard errors.

Summary and Outlook

No significant difference between room conditions could be found. Furthermore, the effects of transition and congruency are identical within room conditions. The chosen room configuration is acoustically very good. It can be assumed that reflections of walls, ceiling and floor do not have a negative impact on auditory selective attention. In rooms with large reverberation times, which are not adequate for speech [6] significantly worse reaction times and error rates are expected and are going to be examined in a further study.

References

[1] 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.

[2] 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 (under review) (2014).

[3] C. Giguère, S. M. Abel: Sound localization: Effects of reverberation time, speaker array, stimulus frequency, and stimulus rise/decay. JASA, (1993) 94(2):769–776.

[4] M. Rychtáriková, T. van den Bogaert, G. Vermeir, J. Wouters: Perceptual validation of virtual room acoustics: Sound localization and speech understanding. Applied Acoustics, (2011) 72(4):196–204.

[5] T. Houtgast, H. J. M. Steeneken, R. Plomp: Predicting Speech Intelligibility in Rooms from the Modulation Transfer Function. I. General Room, Acoustics. Acta Acustica united with Acustica, (1980) 46(1):60–72.

[6] DIN 18041, Ausgabe: 2004-06 Hörsamkeit in kleinen bis mittelgroßen Räumen (17.140.01)