Hearing Research xxx (2017) 1–8

Contents lists available at ScienceDirect

Hearing Research

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

Intentional switching in auditory selective attention: Exploring attention shifts with different reverberation times

Josefa Oberem [a, *], Julia Seibold [b], Iring Koch [b], Janina Fels [a]

[a] Institute of Technical Acoustics, Medical Acoustics Group, RWTH Aachen University, Kopernikusstraße 5, 52074 Aachen, Germany
[b] Institute of Psychology, RWTH Aachen University, Jägerstraße 17, 52066 Aachen, Germany

A R T I C L E I N F O

Article history:
Received 6 July 2017
Received in revised form
11 December 2017
Accepted 18 December 2017
Available online xxx

Keywords:
Auditory selective attention
Task switching
Reverberation
Binaural hearing

A B S T R A C T

Using a well-established binaural-listening paradigm the ability to intentionally switch auditory selective attention was examined under anechoic, low reverberation (0.8 s) and high reverberation (1.75 s) conditions. Twenty-three young, normal-hearing subjects were tested in a within-subject design to analyze influences of the reverberation times. Spoken word pairs by two speakers were presented simultaneously to subjects from two of eight azimuth positions. The stimuli consisted of a single number word, (i.e., 1 to 9), followed by either the direction “UP” or “DOWN” in German. Guided by a visual cue prior to auditory stimulus onset indicating the position of the target speaker, subjects were asked to identify whether the target number was numerically smaller or greater than five and to categorize the direction of the second word. Switch costs, (i.e. reaction time differences between a position switch of the target relative to a position repetition), were larger under the high reverberation condition. Furthermore, the error rates were highly dependent on reverberant energy and reverberation interacted with the congruence effect, (i.e. stimuli spoken by target and distractor may evoke the same answer (congruent) or different answers (incongruent)), indicating larger congruence effects under higher reverberation times.

1. Introduction

Studies on auditory selective attention were firstly introduced by Cherry (1953) and since then analyzed with several dichotic [Broadbent (1958); Pashler (1999); Ihlefeld and Shinn-Cunningham (2008); Bronkhorst (2015); Koch et al. (2011)] and binaural [Best et al. (2007, 2010); Kidd et al. (2005b); Allen et al. (2009); Oberem et al. (2014)] paradigms. In real-life scenes reverberant energy distorts the signal [Náběleк and Robinson (1982); Darwin and Hukin (2000a); Lavandier and Culling (2008)] and therefore it is of interest how auditory selective attention is affected by reverberant energy.

Using an attention task where subjects were asked to repeat four consecutive digits spoken by the target speaker always positioned in front in the presence of two other distracting speakers located to the sides, Ruggles and Shinn-Cunningham (2011) varied the amount of reverberant energy from “anechoic ($RT_{60} = 0$ s), intermediate reverberation ($RT_{60} = 0.4$ s) to high reverberation ($RT_{60} = 3$ s)”. They reported a great impact on performance when adding reverberation, especially differences in performance between anechoic (60–80% correct) and intermediate reverberation (40–50% correct) were noteworthy. On account of these results they conclude that reverberant energy interferes with spatial selective attention.

Similar reverberation times were analyzed by Culling et al. (2003) who measured Speech Reception Thresholds (SRTs) under anechoic ($RT_{60} = 0$ s) and reverberant ($RT_{60} = 0.4$ s) conditions. Target and distractor were collocated in front of the subject or spatially separated ($-60^\circ/+60^\circ$). SRTs were found to be significantly lower under anechoic conditions, which was reconfirmed by Lavandier and Culling (2007). The reverberant energy also interacted with the location of target and distractor, indicating no improvement in SRT for spatially separated speakers in the reverberant condition.

Contradictory to that were findings by Kidd et al. (2005b). They reported that the effect of reverberation was greater when target and masker were spatially separated rather than collocated at the

* Corresponding author.
E-mail addresses: job@akustik.rwth-aachen.de (J. Oberem), julia.seibold@rwth-aachen.de (J. Seibold), koch@psych.rwth-aachen.de (I. Koch), jfe@akustik.rwth-aachen.de (J. Fels).

https://doi.org/10.1016/j.heares.2017.12.013
0378-5955/

same position. Instead of using simulated reverberation times, Kidd and colleagues changed the reverberation of the laboratory by mounting foam and plexiglas to the walls. Further findings were that the amount of masking increased as reverberation times increased and that these acoustic differences also significantly affected the performance in the speech identification task.

Related to the cited investigations [Ruggles and Shinn-Cunningham (2011); Culling et al. (2003); Kidd et al. (2005b)], Darwin and Hukin (2000b) explored the effect of reverberation ($RT_{60} = 0.4$ s) on the ability of listeners to maintain their attention to one speaker across time. Using a paradigm with minimal intelligibility requirements it was found that the influences of reverberant energy on inter-aural time differences (ITD) were significant. The use of ITD differences was impaired by reverberation and therefore maintaining attention to the target was more complicated. However, natural prosody and vocal-tract size differences between talkers, being two further cues for selective attention, were not affected by reverberation.

In the present study, a paradigm focusing on the intentional switching of auditory attention rather than maintaining the listener's attention to a single source was used. Different to paradigms of cited studies the present paradigm offers the possibility to analyze reaction times of the participants and their error rates. Firstly introduced by Koch et al. (2011), the paradigm has been used and tested [Lawo et al. (2014); Lawo and Koch (2015)] with dichotic reproduction and is by now well-established [Bronkhorst (2015)]. Koch et al. (2011) explicitly examined the endogenous, voluntary attention switches, therefore cued attention switches referred to the target's gender and its location (e.g. the target's gender switched between trials; in the preceding trial the target was a male speaker on the left side and in the following trial the target was a female speaker on the right side). The main finding was that a cued switch of the relevant target resulted in a worse performance than in cued repetitions of the relevant target's speaker gender and location [Lawo et al. (2014); Lawo and Koch (2015)]. Furthermore, the role of attentional control in processing of task-irrelevant information in auditory attention switching has been explored by the authors. The participants' task was to always categorize the relevant number word presented by the target speaker as smaller than or greater than five and press the corresponding response button. The two presented stimuli of one trial could be congruent (both number words smaller than five or both greater than five) and therefore call for the same response, or they could be incongruent (one digit was smaller and one was greater than five) and therefore call for different responses. The “congruency effect” [Kiesel et al. (2010)], showing that participants respond faster in congruent trials than in incongruent trials, was confirmed [Koch et al. (2011)], suggesting some processing of irrelevant information (i.e. of distractor's speech).

The paradigm was also extended into a binaural version to reproduce more realistic scenarios [Oberem et al. (2014, 2017a,b); Fels et al. (2016)]. For this purpose, a scene with more combinations of the speaker's locations than in the dichotic set-up was provided in an anechoic chamber with different binaural reproduction methods. Besides a set-up with real loudspeakers, head-related-transfer-functions were used to present binaural stimuli via headphones. In the investigation by Oberem et al. (2014) it was found that a required switch of the attention focus yielded longer reaction times and increased error rates than a repetition of the target's location, which were also dependent from the target's location itself.

In the present investigation participants had to categorize one of two binaurally presented couple of a number word and a direction word according to numerical size and direction. The target and distracting speaker were always positioned in one out of eight different positions around the listener but never collocated [Fels et al. (2016)]. As outcome measures reaction times as well as error rates were observed.

Inspired by the findings of Ruggles and Shinn-Cunningham (2011), reverberation times in three levels from anechoic ($RT_{60} = 0$ s), low reverberation ($RT_{60} = 0.8$ s, comparable to an acoustically untreated classroom instead of $RT_{60} = 0.4$ s, comparable to a damped recording room) to high reverberation ($RT_{60} = 1.75$ s, comparable to a auditorium instead of $RT_{60} = 3$ s, comparable to a medium-sized church) were simulated. The underlying room model was also designed with comparable diameters, however, walls were not set to be parallel and the listener was not positioned in the center of the room to prevent unwanted acoustical effects due to nodal points or echos [Hartmann (1983); Rakerd and Hartmann (1985); Giguère and Abel (1993)] (c.f. Section 2.4).

It was postulated that reverberant energy would increase reaction times and error rates in the present investigation, based on the cited findings [Ruggles and Shinn-Cunningham (2011); Culling et al. (2003); Kidd et al. (2005a); Darwin and Hukin (2000b)]. Furthermore, Ruggles and Shinn-Cunningham (2011) showed how maintaining auditory selective attention on a single sound source in presence of interfering sources is degraded by reverberant energy. These findings led to the hypothesis of increased reaction times and error rates for repetition trials (i.e. where a listener is asked to focus on the same direction in two consecutive trials (c.f. Section 2.6)), under increased reverberation in the present investigation. Since it was known, how reverberation degrades ITD timing information, which results in a blurred localization information [Ruggles and Shinn-Cunningham (2011)], it was predicted in the present investigation that localizing a new sound source and focusing attention on that source would also degrade with increasing reverberation times. This is the case for switch trials where the listener has to switch his/her attention to a new spatial position between trials (c.f. Section 2.6).

Spatial separation turned out to be beneficial in findings by Kidd et al. (2005a) under increasing reverberation times, however, Culling et al. (2003) reported an opposite effect. Therefore, in this investigation special attention is focused on the spatial location of target and distractor as well as their angular separation.

2. Methods

2.1. Participants

A number of 23 paid (8 Euro) students aged between 19 and 34 years (mean age: 23.8 ± 3.4 years) participated in the experiment. Subjects were equally divided into male (12) and female (11) listeners. Listeners were screened to ensure that they had normal hearing (within 20 dBHL) for frequencies between 250 Hz and 10 kHz via pure-tone ascending standard audiometry. All listeners could be considered as non-expert listeners since they had never participated in a listening test on auditory selective attention.

2.2. Stimulus material

Speech material was recorded under anechoic conditions with two male and two female professional, native German speakers. The used hardware, a large diaphragm condenser microphone TLM170 by Neumann and Zoom H6 Handy Recorder (both: cardioid directivity pattern), allowed recordings with a frequency range from 70 Hz to 20 kHz. The stimuli consisted of a single spoken digits (1–9, excluding 5) which was followed by one of two German disyllabic direction words (“UP”, in German “OBEN” and “DOWN”, in German “UNTEN”)(e.g. the combined stimulus could be “Four