Irrelevant sound effects with locally time-reversed speech: Speech reversal and language familiarity

Kazuo Ueda¹, Yoshitaka Nakajima¹, Florian Kattner²,³, and Wolfgang Ellermeier²,³

¹Unit of Perceptual Psychology, Dept. Human Science/Research Center for Applied Perceptual Science, Kyushu University
²Institut für Psychologie, Technische Universität Darmstadt

Key words: memory disruption, speech intelligibility, serial recall

Purpose
To investigate, how altering temporal detail in the interfering auditory stream might affect the magnitude of the irrelevant sound effect (Ellermeier and Zimmer, 2014), we used (a) the original recordings of speech, (b) locally time-reversed speech (Ueda et al., 2017), (c) reversed playback of the entire locally time-reversed sequence, and (d) reversed speech as distractors in the irrelevant speech paradigm.

Further, a variation of the familiarity of the distracting background speech was combined with those systematic manipulations of forward vs. reversed and local vs. global contrasts, to test generality of the results. Specifically, spoken sentences in German and Japanese, which were extracted from the same speech database, and stimuli made from these were employed to test participants in each native language population.

Method
The experiment was performed in two laboratories in parallel, at Technische Universität Darmstadt in Germany, and at Kyushu University in Fukuoka, Japan, with German native listeners \(n = 79\) and with Japanese native listeners \(n = 81\), employing both German and Japanese speech with either sample. About half of each participant group was assigned to native language background conditions, whereas the other half of each group was assigned to non-native language background conditions. The speech stimuli were based on speech samples in the speech database (NTT-AT, Multi-Lingual Speech Database 2002). The auditory stimuli, including the speech stimuli (about 74 dB SPL) and pink noise (about 72 dB SPL), were presented to the participants diotically via a pair of headphones (Beyerdynamic, DT 990) in a soundproof booth. The visual stimuli were random sequence of eight digits without repetition from a set of 1 to 9 and were presented on a computer screen. Digits were presented one by one for 1 s per each, and a 6-s retention interval with a blank screen followed. The participants were instructed to recall serially the digits after the retention intervals, ignoring the background sound.

Results
Figure 1 shows the averaged results. Pink noise (the control condition) was least disruptive to the task, and all the other conditions that employed speech were more disruptive, despite the instruction to ignore the irrelevant sound. For irrelevant speech in the listeners’ native language, locally time-reversed speech with the shortest segment duration (20 ms), which was highly recognizable, was as disruptive as normal speech, whereas locally time-reversed speech with longer segment durations (70 or 120 ms) was less disruptive [significant interaction effects of speech segment duration \(\times\) direction, \(F(3, 2995) = 7.94, p < .0001\), and \(F(3, 3232) = 2.88, p = 0.03\), which correspond to Fig. 1a and c respectively, with repeated-measures ANOVA with arc-sine transformation, were observed]. In contrast, the effect of segment duration vanished completely when locally time-reversed speech was played backward. When the non-native language was presented, the effect of playing directions seemed to be attenuated in locally time-reversed speech.

Discussion
Global integrity of speech affects more on the irrelevant sound effect than local integrity; besides, mechanisms driving the irrelevant sound effect seem to differ between native and non-native languages. Thus, both intelligibility and familiarity of language in background speech did affect the results: Intelligible speech in a familiar language was most disruptive for the participants, whereas unintelligible speech in either familiar or unfamiliar language equally yielded moderate disruptions.

References
Ellermeier, W., and Zimmer, K. (2014). The psychoacoustics of the irrelevant sound effect. Acoustical Science and Technology, 35, 10–16, doi:10.1250/ast.35.10.

Ueda, K., Nakajima, Y., Ellermeier, W., and Kattner, F. (2017). Intelligibility of locally time-reversed speech: A multilingual comparison. Scientific Reports, 7:1782, doi:10.1038/s41598-017-01831-z.

Figure 1. Averaged performance under background sound conditions. (a) German and (b) Japanese background speech was presented to German participants \(n = 38, n = 41\), whereas (c) Japanese and (d) German background speech was presented to Japanese participants \(n = 41, n = 40\). Noise: Pink noise. LTR: Locally time-reversed speech. Error bars represent SEM.