Acoustics'08 Paris
June 29-July 4, 2008
www.acoustics08-paris.org

An algorithm modelling the Irrelevant Sound Effect (ISE)

S. J. Schlittmeier[a], T. Weissgerber[b], S. Kerber[b], H. Fastl[b] and J. Hellbrueck[a]

[a]Work, Environmental and Health Psychology, Catholic University of Eichstaett-Ingolstadt, Ostenstr. 26-28, 85072 Eichstaett, Germany
[b]AG Technische Akustik, MMK, TU München, Arcisstr. 21, 80333 München, Germany
sabine.schlittmeier@ku-eichstaett.de

Verbal short-term memory capacity is reduced significantly during certain background sounds. Remembering a series of digits is significantly impaired by speech or music with prominent staccato passages whereas, for example, music with prominent legato passages does not disturb performance in comparison to silence. This so-called Irrelevant Sound Effect (ISE) occurs although the background sounds are irrelevant with respect to the digit sequences to be remembered. Until now, a multitude of cognitive psychological experiments explored the ISE and collected behavioral performance data during different sound conditions.
The talk presents an algorithm which models performance data in ISE experiments, i.e. the detrimental impact of background sounds on memory performance. The database of this algorithm is about 50 background sounds and corresponding performance data, which have been collected in cognitive psychological experiments at the KU Eichstätt-Ingolstadt. The algorithm is based on the instrumental measuring of the hearing sensation fluctuation strength and is able to emulate the performance results in about 90 % of cases within the interquartile ranges. The algorithm will be discussed within the scope of cognitive short-term memory models, which claim to explain the ISE and with respect to practical implications.

1 Introduction

Cognitive performance very rarely takes places during silence. However background sound can reduce cognitive performance, even if it is irrelevant to the task and is intended to be ignored. This phenomenon has been verified in a multitude of behavioral experiments for verbal short-term memory performance and is called the Irrelevant Sound Effect (ISE) (cf. e.g. [1]). Verbal short-term memory provides the retention of verbal material over a short-period of time, e.g. for further processing, reasoning or for long-term storage. Thus, it is central to human information processing.

The capacity of verbal short-term memory is significantly reduced when background sounds are present, which are characterized by distinct temporal-spectral variations with successive varying perceptual tokens (changing-state sounds). Thus, an ISE is evoked by background speech or music with prominent staccato passages, whereas music with prominent legato passages or continuous noise disturbs performance significantly less – or even not at all in comparison to silence (e.g. [2]). The relevance of the changing-state characteristic for ISE evocation is expressed in terms of the cognitive psychological changing-state hypothesis proposed by Jones and co-authors (e.g. [3]) and has been widely accepted.

Although a multitude of cognitive psychological experiments have explored the ISE and collected behavioral performance data during different sound conditions (see [1, 2] for overviews), no external psycho-acoustical criterion existed to predict the ISE's occurrence. Thus, behavioral experiments are still indispensable to identify the effect of a given background sound on short-term memory performance. This poses problems in an applied context. In office environments, for example, the potential beneficial – or less advantageous – effects of noise abatement on cognitive performance can not be calculated before their realization (cf. e.g. [4]).

Consequently, a quantitative parametric approach which allows the irrelevant sound's inherent changing-state to be measured – and with this its disturbance impact on short-term memory performance (i.e. the ISE) – is greatly needed. In the following study we present an algorithm which models performance data in ISE experiments on the basis of instrumental measurement of the hearing sensation fluctuation strength (cf. also [5]).

2 The behavioral experiments

The database used for algorithm modeling of the ISE comprises of about 50 background sounds and corresponding performance data, which have been collected in cognitive psychological experiments at the KU Eichstätt-Ingolstadt. Background sound conditions included speech, music, pink noise, animal sounds, sequences of tones, office and traffic noise.

In all of these experiments the verbal serial recall task was applied; this is the standard method for testing verbal short-term memory and the ISE. Here, the experiment involved the digits from 1 to 9 being presented successively in a randomized order during the different background sound conditions. The participants were asked to recall the numbers after a short retention interval of 10 s in their exact presentation order. Each digit not recalled in the correct serial position was counted as an error. Items were presented either visually or auditorily. In the latter case, perfect intelligibility of the auditory items was ensured by speech identification tests.

A trail in this experiment encompassed item presentation, retention interval and recall. Depending on the experiment, 15-20 trials were performed during each sound condition. The error rates over all trials within one sound condition were averaged for each participant. These error rates were collapsed over the whole corresponding experimental group and an error median was obtained for every sound condition. In the following descriptions these error medians will be referred to in terms of error percentages.

Background sounds were presented binaurally with defined sound pressure levels either via loudspeaker or head phones. All background sounds were of moderate level (35-60 dB(A)). The sound pressure level was measured with a Brüel & Kjær 2231 sound pressure level meter or a NoiseBook 2.0 (HEAD acoustics GmbH) in the case of loudspeaker presentation, and with a Brüel & Kjær 4153 artificial ear in the case of presentation via headphones.

3 The algorithm

In the behavioral experiments, the highest disturbance impact was found during irrelevant background speech. Narration is typically characterized by four syllables per second, which corresponds to fluctuations of 4 Hz in the temporal envelope [6]. At the modulation frequency of 4 Hz the hearing sensation "fluctuation strength" is at its maximum. The fluctuation strength describes modulated sound signals with modulation frequencies lower than 20 Hz. Our auditory system can perceive these fluctuations as single events, they are therefore perceived as level variations [7]. Since the fluctuation strength of background sounds correlates with their detrimental impact on performance, this psycho-acoustical quantity shows promise for predicting the ISE of many sounds. It was thus chosen to model the ISE.

For this purpose, the sound files tested in the behavioral experiments were limited to 30 representative seconds, corresponding to the duration of the trial. All sound files were adjusted to the sound pressure level used in the ISE experiments. The fluctuation strength of each sound file was measured using the software PAK® (Müller-BBM VibroAkustik Systeme GmbH) and its arithmetic mean was calculated over time. As a reference sound condition for the algorithm, a background sound with a small interquartile range was chosen. Short-term memory performance during this reference sound condition along with its fluctuation strength were used to normalise the algorithm. Although the ISE has been shown from a qualitative perspective to be independent of the presentation modality of the items to be remembered [2], the magnitude of the ISE seems to depend on it. Thus two algorithms are presented, one for visual items and one for auditory items. These algorithms differ exclusively in one factor while otherwise being identical from a mathematical perspective.

3.1 Visual items

The database for visual items encompasses the behavioral results of 46 experimental measurements of verbal short-term memory performance carried out during different background sound conditions. Music with prominent staccato passages (sound no. 15 in figure 1) was chosen as the reference sound condition. The fluctuation strength of this sound is F = 0.68 vacil and its disturbance impact (ISE) is 7.5 %, i.e. the mean error rates in the verbal short-term memory task were raised by 7.5 % during staccato music in comparison to performance during silence. Using this, the ISE for visual items (ISE_v) can be estimated for any background sound according to the following Eq. (1):

$$ISE_v = \frac{\mathrm{F}}{0.68\ \mathrm{vacil}} \cdot 7.5\ \ [\%] \qquad (1)$$

The expected enhancement in error rates during the different background sound conditions compared to silence was calculated for all 46 experimental measurements within the visual items data base using Eq. (1). Figure 1 depicts the algorithmic calculated error rates as well as the corresponding experimentally verified behavioral error rates for each experimental measurement. It should be noted that these two error rates correlate significantly (Spearman's rank correlation: $r_s = .72$; $p < .01$).

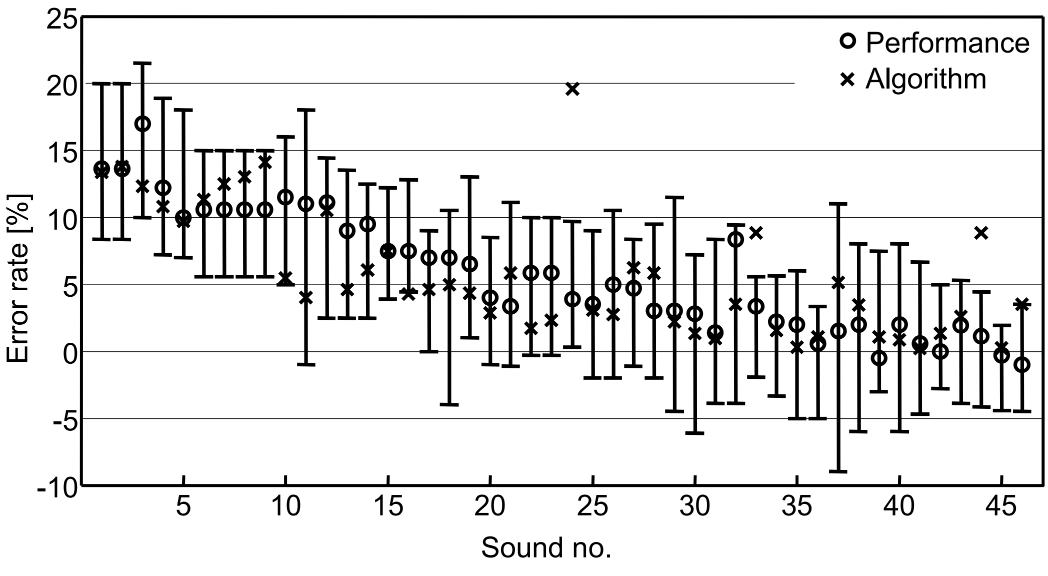

Figure 1: Verbal short-term memory performance for visually presented items in 46 experimental measurements during different background sounds. Relative error rates (baseline: performance during silence) within interquartile ranges are plotted. Circles represent the relative error rates found in behavioral experiments; crosses represent the algorithmically calculated relative error rates.

This algorithm reproduced the behavioral results of 43 out of 46 experimental measurements within the interquartile ranges. The inefficiency of sound nos. 24, 33 and 44 with respect to short-term memory performance is not estimated adequately by the algorithm. To note, these are synthetic sounds with periodicity in the temporal and/or spectral perspective, i.e. these sounds are characterized by a periodically reoccurring auditory-perceptive token. For example, sound nos. 33 and 44 represent the repeated presentation of a vowel or consonant, respectively. Background sounds which are constituted by the repetition of one auditory-perceptive token have been shown to reduce short-term performance significantly less than changing-state sounds (see Introduction). Since the algorithm accounts exclusively for the fluctuation strength of a given background sound the disturbance impact (i.e. the ISE) of such artificial steady-state sounds is noticeably overestimated.

3.2 Auditory items

The data base for auditorily presented items encompasses the behavioral results of 24 experimental measurements of verbal short-term memory performance during different background sounds. A speech signal (sound 3 in figure 2) was selected as a reference sound condition for normalisation. It is characterised by a fluctuation strength of F = 0.95 vacil and a disturbance impact (ISE) of 11.5 %. With this the ISE for auditory items (ISE_a) can be estimated for any background sound according to the following Eq. (2):

$$ISE_a = \frac{F}{0.95 \text{ vacil}} \cdot 11.5 \ [\%] \qquad (2)$$

Using Eq. (2), the ISE, i.e. the expected enhancement in error rate due to different background sounds during the task performance, was calculated for all 24 experimental measurements. Figure 2 depicts the calculated error rates and the corresponding experimentally verified behavioral error rates. Here, these two error rates were also found to correlate significantly (Spearman's rank correlation: r_s = .87; p < .01).

The present algorithm adequately models the performance results of 21 out of the 24 experimental measurements for auditorily presented items, i.e. within the interquartile ranges. Here, the error rates for certain steady-state sounds (sounds no. 9, and in particular 14 and 15) were calculated as significantly too high by the algorithm, too.

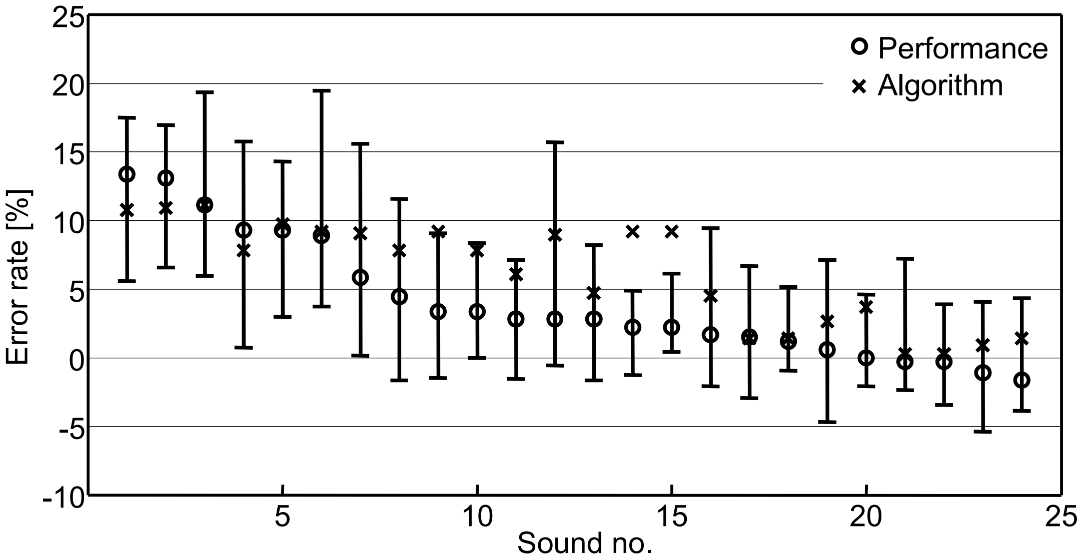

Figure 2: Verbal short-term memory performance for auditorily presented items in 24 experimental measurements during different background sounds. Relative error rates (baseline: performance during silence) within interquartile ranges are plotted. Circles represent the relative error rates found in behavioral experiments; crosses represent the algorithmically calculated relative error rates.

4 Conclusions

The present algorithm satisfactorily models the detrimental impact of background sounds on short-term memory performance (ISE) based on the hearing sensation fluctuation strength. Altogether, the algorithm has reproduced more than 90 % of the behavioral results within the interquartile ranges. With this, the hearing sensation fluctuation strength can be considered as appropriate to describe the physical-perceptive aspects that define a sound as a changing-state sound able to significantly impair cognitive performance.

The only sounds for which the algorithm does not provide adequate estimations, are synthetic sounds characterized by the repeated presentation of one auditory-perceptive token (steady-state sounds). Thus, the algorithm needs further refinement with respect to these steady-state sounds. This might be accomplished by implementing a periodicity identifier to allow for steady-state correction.

Yet steady-state sounds are not found in real world settings but are only presented in cognitive psychological experiments for exploring basic research questions. For evaluations of background sounds occurring in the field, e.g. background speech, office noise, background music etc., the algorithm provides very good results. The algorithm is able to estimate the differing performance effects induced by background speech differing in intelligibility as found in office environments due to altered sound transmission situations (see [4]). With this, the algorithm proves to be extremely promising in applied perspective, e.g. for evaluating noise abatement measures with respect to their performance effects in the planning phase.

References

[1] Banbury, S. P., Tremblay S., Macken, W. J., Jones, D. M., “Auditory distraction and short-term memory Phenomena and practical implications”, *Human Factors* 43(1), 12-29 (2001).

[2] Schlittmeier, S. J., Hellbrück, J., Klatte, M., “Does irrelevant music cause an irrelevant sound effect for auditory items?” *European Journal of Cognitive Psychology* 20(2), 252 – 271 (2008).

[3] Macken, W. J., Mosdell, N., Jones, D. M., “Explaining the irrelevant sound effect. Temporal distinctiveness or changing state?” *Journal of Experimental Psychology: Learning, Memory & Cognition* 25(3), 820-824 (1999).

[4] Schlittmeier, S. J., Hellbrück, J., Thaden, R., Vorländer, M., “The impact of background speech varying in intelligibility: Effects on cognitive performance and perceived disturbance”, *Ergonomics* (in press).

[5] Weissgerber, T., Schlittmeier, S., Kerber, S., Fastl, H., Hellbrück, J. „Ein Algorithmus zur Vorhersage des Irrelevant Sound Effects“, DAGA 2008, Dresden, Germany (2008).

[6] Fastl, H., “A background noise for speech audiometry”, Audiol. Acoustics 26, 2-13 (1987).

[7] Fastl, H, Zwicker, E., “Psychoacoustics – Facts and Models”, 3rd Edition, Springer Verlag, Heidelberg (2007).