Verbal Short-term Memory Performance is often but not always reduced by Background Speech and Music - Experiments exploring the Irrelevant Sound Effect (ISE)

Sabine J. Schlittmeier [1], Jürgen Hellbrück [1] & Maria Klatte [2]

[1] Catholic University of Eichstätt-Ingolstadt; [2] University of Oldenburg

Theoretical Background and Research Interest

Performance in immediate serial recall of a list of visually presented unrelated digits is disrupted by certain simultaneously presented background sounds which are irrelevant for the task. This phenomenon is known as Irrelevant Sound Effect (ISE) and is explained in terms of short-term memory models.
Most frequently use is made of the Phonological Loop Model proposed by Alan Baddeley (1986). It comprises a phonological store, in which memory traces exist, and a subvocal rehearsal-process. Heard speech and periodical non-speech sounds are supposed to have obligatory access to the phonological store. Written language however, has to be transformed for storage in the phonological store by means of a graphem-to-phonem translation achieved by the subvocal rehearsal-process. The ISE is attributed to interferences arising in the phonological store between the representations of the items to-be-remembered and those of the background sound.
An alternative model for the explanation of the ISE has been suggested by Dylan Jones: the Object-Oriented Episodic Record Model (O-OER Model; Jones, 1993). In this model, all sensory input is deposited directly in an abstract and amodal code in one unitary store. The ISE for visually presented items, is considered to the consequence of impaired maintenance of order information caused by the irrelevant background sound. However, this applies only to sounds with distinct temporal-spectral variations (changing-state features). Sounds without these characteristics (steady-state sounds) do not cause an ISE.
For visually presented items, the occurrence of an ISE can be considered as an empirically robust finding which has been repeatedly shown for different background sounds (e.g., Klatte, Kilcher & Hellbrück, 1995, for speech and music with changing-state characteristics). However, the influence of background sound on auditorily presented items has hardly been investigated. The same is true for the role of item presentation rate. In our experiments we vary the modality of item presentation (visual versus auditory presentation) as well as presentation rate (fast versus slow item presentation).

Empirical Studies

Task and Experimental Design

Subjects had to perform the immediate serial recall of unrelated digits, the standard task for investigating short-term memory. Digits from 1 to 9 were presented in randomized order. Following by a 10 s retention interval all digits appeared simultaneously in random order on the computer screen. Using the computer mouse the task consisted in clicking the digits in exactly the same order previously presented to subjects. Each digit not recalled on the serial position it had been presented was classified as an error.
In case of auditory presentation digits were spoken by a female voice and presented with 64 dB(A) via one loudspeaker located in front of the subjects. Visual items appeared one after the other in the middle of the screen. Fast item presentation was achieved by presenting the digits with a rate of one digit per second. So item presentation as a whole lasted 9 s, in contrast to 16 s for slowly presented digits. All background sounds were played with 60 dB(A) via one loudspeaker placed behind the subjects.
Subjects of speech-groups learned under four different sound conditions silence, pink noise, steady-state speech (vowel A spoken by a male voice and looped) and changing-state speech (letters G M K P F L Q spoken by a male voice and looped). Subjects of music groups were tested under silence, legato music (meditation music; low temporal-spectrale dynamics) and staccato music (baroque music; distinct temporal-spectral dynamics). Background sound presentation started 2 s before the first item of a list was presented and stopped after the subject completed his response. 20 trials had to be performed under each sound condition. The sequence of sound conditions was randomized for each subject.

Experiment 1: Fastly presented auditory and visual items under irrelevant background speech and music

In experiment 1 two experimental groups each consisting of 20 subjects were tested under irrelevant background speech: one group learning visually presented items and the other learning auditorily presented items. Likewise, under the exposure of irrelevant background music, two further groups learned either visually (*n*=20) or auditorily presented digits (*n*=25).
For each of the four experimental groups means and standard errors of error rates are depicted in Figure 1 and 2.

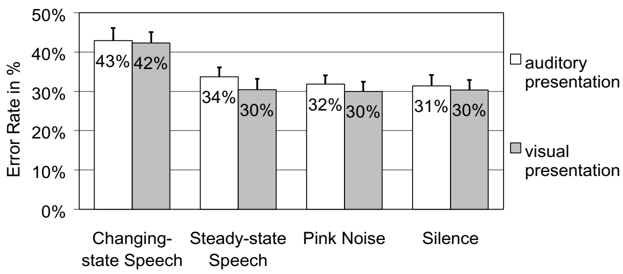

Figure 1: Fastly presented auditory and visual items under irrelevant background speech (*N*=40)

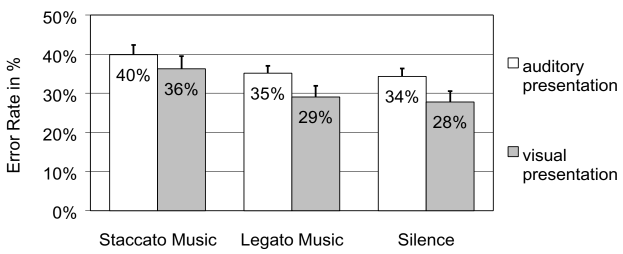

Figure 2: Fastly presented auditory and visual items under irrelevant background music (*N*=50)

All four groups showed an ISE just for changing-state sounds (speech-group learning auditorily presented items: $F(3,57)=20.37$, $p<.001$; speech-group learning visually presented items: $F(3,57)=23.19$, $p<.001$; music-group learning auditorily presented items: $F(2,48)=7.41$, $p=.002$; music-group learning visually presented items: $F(2,48)=16.20$, $p<.001$). Performance under steady-state speech and under pink noise was not significantly different from performance under silence for any of the groups.

Experiment 2: Slowly presented auditory items under irrelevant background speech and music

In experiment 2 serial recall performance for slowly presented auditory items has been examined. One group was tested under irrelevant background speech (*n*=20) and the other under irrelevant background music (*n*=25).
In figure 3 means and standard errors of error rates are plotted for the speech group. Analogously Figure 4 shows the data for the music group. Data of the corresponding groups learning fastly presented auditory items (9 s) in experiment 1 has been added to both figures for a direct comparison.

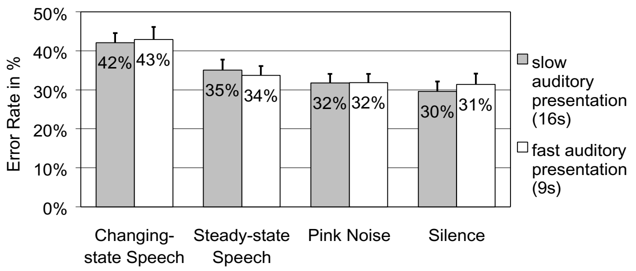

Figure 3: Slowly presented auditory items under irrelevant background speech (*N*=40)

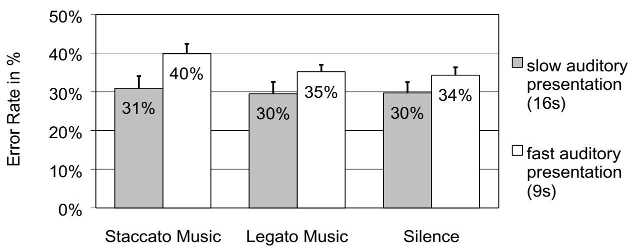

Figure 4: Slowly presented auditory items under irrelevant background music (*N*=50)

In the speech group only changing-state speech gave rise to an ISE ($F(3,57)=21.23$; $p<.001$), thus closely resembling the result of the rapid presentation condition. For the music group however, the results deviated strongly from the fast presentation condition. Here staccato music did not reduce the immediate serial recall of slowly presented auditory items ($F(2,48)<1$), although the same background sound caused an ISE for fastly presented items (experiment 1).

Experiment 3: Variation of the onset of irrelevant background speech for fastly presented auditory items

Experiment 3 was carried out to answer the question if the observed performance reduction under changing-state speech for auditorily presented items in experiment 1 and 2 has to be attributed to a disturbance of the encoding of the material to-be-remembered (e.g., masking, increased hearing effort). For this, one more group (*n*=20) was examined under irrelevant background speech, now with background sound presentation starting after the presentation of the ninth and last item of a list. Therefore, item presentation always took place under silence.
Figure 5 shows means and standard errors of error rates for each learning condition. For comparison, error rates of the auditory group receiving background sound all throughout the trial (during presentation, retention interval and recall) in experiment 1 are included into the figure.
Post-hoc tests did not find any differences between the two groups on any of the learning conditions. So the ISE caused by changing-state speech persists even if background sound presentation starts after item presentation and the items themselves are encoded under silence ($F(3,57)=22.89$, $p<.001$).

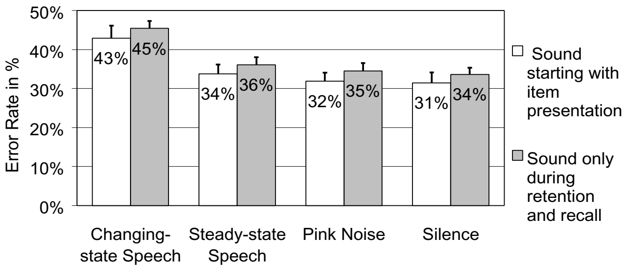

Figure 5: Variation of the onset of irrelevant background speech for fastly presented auditory items (*N*=40)

Summary and Discussion

It has been shown that background speech and background music influences performance in immediate serial recall irrespective of input modality of the items to-be-remembered (experiment 1). On the one hand this stands in line with the Phonological Loop Model. This model does not suppose any difference between auditorily and visually presented items concerning their susceptibility to irrelevant background sound, because items are processed independently of presentation modality (Baddeley, 1986). On the other hand we would have expected that auditorily presented items are not effected by irrelevant background sound if the items and the background signal are perceived as two different auditory streams. It can be assumed that this has been the case because of at least two reasons. Firstly, the items to-be-remembered and the background sounds were presented via two loudspeakers spatially distinctly separated and secondly, they were spoken in two different voices.
For slowly presented auditory items only changing-state speech lead to an ISE, whereas staccato music did not have an effect on recall performance (experiment 2). This ISE persisting under irrelevant background speech cannot be explained in terms of masking or increased hearing effort, but has to be located in short-term memory (experiment 3). Since presentation rate is not implemented in neither of the two mentioned short-term memory models, especially the observed interaction between presentation rate (fast versus slow presentation) and kind of background sound (speech versus music) cannot be explained within the scope of these models.
It has to be emphasized here that speech is the most disturbing background sound for cognitive tasks. This has, for example, practical implications for office workplaces even if the disturbing impacts of background sounds can not yet be fully explained from a cognitive psychological point of view.

Literature

Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.
Jones, D. M. (1993). Objects, streams and threads of auditory attention. In A. D. Baddeley & L. Weiskrantz (Eds.), *Attention: Selection, awareness and control* (S. 87-104). Oxford: Clarendon Press.
Klatte, M., Kilcher, H. & Hellbrück, J. (1995). Wirkung der zeitlichen Struktur von Hintergrundschall auf das Arbeitsgedächtnis und ihre theoretischen und praktischen Implikationen. *Zeitschrift für Experimentelle Psychologie, 40*, 91-98.