BACKGROUND SPEECH VARYING IN INTELLIGIBILITY – EFFECTS ON COGNITIVE PERFORMANCE AND PERCEIVED DISTURBANCE

PACS 43.50

Schlittmeier, Sabine J.[1]; Liebl, Andreas[1]; Hellbrück, Jürgen[1]; Thaden, Rainer[2]; Vorländer, Michael[2]
[1]Work-, Environmental- and Health Psychology, Catholic University of Eichstätt-Ingolstadt, D-85071 Eichstätt, Germany; sabine.schlittmeier@ku-eichstaett.de
[2]Institute of Technical Acoustics, Aachen University, D-52056 Aachen, Germany; mvo@akustik.rwth-aachen.de

ABSTRACT

We present three experiments, which tested the impact of background speech varying in intelligibility and/or level on basic cognitive functions (verbal short-term memory, sustained attention, verbal-logical reasoning). Variations of background speech intelligibility were obtained by auralisation based on room- and building acoustics. In each experiment, performance was measured during low background speech (35 dB(A)) of good and poor intelligibility, respectively, and compared to silence as well as to highly intelligible speech not lowered in level (55 dB(A)). Besides performance data subjective disturbance ratings were collected.

Our results show that highly intelligible background speech of 55 dB(A) impairs cognitive performance. In order to diminish its disturbance impact significantly, a sole level reduction does not suffice – an additional reduction of speech intelligibility is necessary. Subjective disturbance ratings, however, value lowering of the background speech's level to a greater extent than reducing its intelligibility. Furthermore, although cognitive performance coincided during soft and poorly intelligible speech with those during silence, it was still rated as significantly disturbing. Thus, our experiments demonstrate that objective performance tests and subjective ratings complement but cannot substitute each other with regard to a comprehensive evaluation of acoustic office environments and their alterations.

INTRODUCTION

Office employees frequently complain about noise, in particular about background speech such as conversations among colleagues or the noise of other people talking on the phone (e.g. [1]). A reduction of background speech's intelligibility is considered a solution to this problem (e.g. [2]). The idea that background speech's high intelligibility might be crucial to its disturbance impact on cognitive performance – and not only regarding subjective assessments – can be derived from psychological research on the irrelevant sound effect (ISE).

The ISE describes the significant impairment of verbal short–term memory during irrelevant speech and certain non–speech sounds, and is attributed to the loss of information laid down in short-term memory (see e.g. [3]). The effect occurs even though background sounds are irrelevant with respect to the required task, and even though participants are instructed to ignore them. One essential background sound feature for ISE evocation is that the irrelevant sound is characterised by distinct temporal–spectral variations that allow for perceptual segmentation of the irrelevant sound, while successive perceptual tokens vary in acoustic–perceptive perspective. These so–called changing–state features (e.g. [4]) distinguish highly intelligible background speech. Consequently, reducing its perceptual cues for segmentation should reduce the disturbance impact of background speech. In fact, this is actually done by sound insulation measures implemented in offices to reduce speech intelligibility; these predominantly lower frequencies above approximately 400 Hz and with this, the decisive cues for speech segmentation, namely consonants.

However, verbal short-term memory is not the sole brick of human cognition. Within the cognitive psychological approach of human information processing, several additional basic cognitive

functions – like perception, attention, reasoning, inhibition of automated processes and long-term storage – are identified (cf., e.g. [5]). The interplay of these functions allows us to accomplish complex cognitive tasks (e.g. text comprehension). Thus, evaluating whether a reduction of background speech intelligibility corresponds with enhanced cognitive performance compared to working conditions in which background speech is highly intelligible requires systematic testing of basic cognitive functions. At the very least, regarding short-term memory performance as basic cognitive function and a more complex proof reading task, empirical evidence has been presented that demonstrates that reduced speech intelligibility leads to a diminished disturbance effects on cognitive performance [6, 7]. These authors, however, reduced intelligibility by partially masking the background speech with continuous noise, which results in qualitatively different acoustic environmental conditions than that induced by sound insulation measures. (For example, if the overall level of background sound is higher due to additionally played-back partial masker, subjective evaluations of the working conditions as well as physiological aspects such as arousal can be expected to be affected.)
To determine whether reducing background speech's intelligibility by sound insulation measures in fact diminishes the disturbance impact, the effects of background speech on cognitive performance were investigated in three experiments. Besides verbal short-term memory (Exp. 1), sustained attention (Exp. 2) and verbal-logical reasoning (Exp. 3) as further basic cognitive functions were explored during background speech of high and low intelligibility. Additionally to performance measures, participants in our experiments subjectively evaluated the perceived disturbance impact of each sound condition.

EXPERIMENT 1: VERBAL SHORT-TERM MEMORY

Method

Using auralisation based on room- and building acoustics (cf. [8]) speech signals were generated, as they would be hearable when the speaker sits in the neighbouring office. Different shapes of sound insulation curves (Figure 1) representing unequal building situations, led to two soft speech signal (35 dB(A)), one of good intelligibility (S35_G) and one of bad intelligibility (S35_B). Listening tests requiring the repetition of short, relatively easy and semantically meaningful sentences proved the differing intelligibility of the two auralised speech signals (S35_G: 98% ratio of understood words to the total number of words; S35_B: 75%).

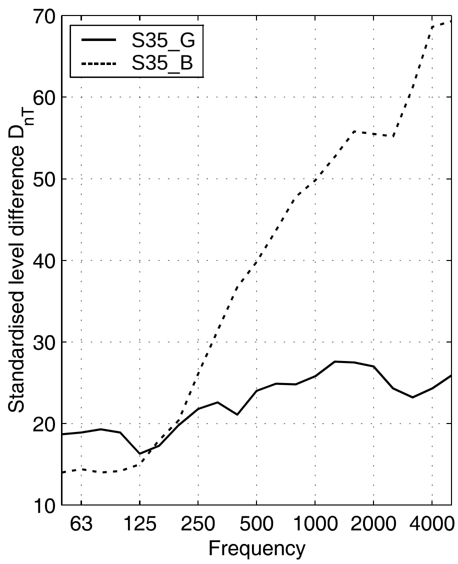

Figure 1.- Sound insulation curves of the two used conventional building situations used to auralise the two speech signals differing clearly in speech intelligibility. The standardised level difference D_{nT} illustrates the filtering of the original speech signal.

Verbal short-term memory performance was measured during the two auralised speech signals of good and poor intelligibility (S35_G, S35_B), the speech signal of good intelligibility originally used for auralisation (S55; 55 dB(A)) and a silence condition represented by very soft pink noise

(25 dB(A)). Serial recall, the standard procedure for measuring verbal short-term memory capacity was used. Digits from1 to 9 were successively presented visually in the middle of a computer screen (700 ms on, 300 ms off) in randomised order and had to be recalled in exact presentation order. 20 successive trials had to be worked out during each of the four sound conditions (sound condition sequence was balanced over participants). After each of these blocks, participants had to answer the question ‘How disturbing were the background sounds in this experimental block to you?’ on a 5–point scale bearing the verbal anchors ‘not at all’ (0), ‘a little’ (1), ‘middle’ (2), ‘rather’ (3) and ‘extremely’ (4).
The experiment was carried out in single sessions in a double–walled IAC sound–proof booth with binaural sound presentation via headphones. 20 participants took part in Experiment 1.

Results

Each digit not recalled in its previously presented serial position was scored as an error (Figure 2). T–tests revealed that significantly more errors were made during highly intelligible speech of 55 dB(A) in comparison to silence (p<.01). Its disturbance impact was not significantly reduced by an exclusive lowering of noise level while leaving intelligibility high (S35_G; p=.19), only a reduction of both level and intelligibility led to significantly less errors (S35_B; p<.01). Furthermore, S35_B did not reduce serial recall performance significantly and only tended to do so when compared with silence (p=.07).

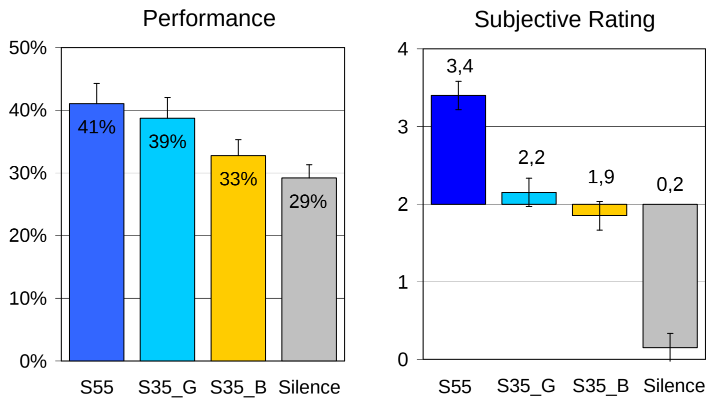

Figure 2.- Verbal short–term memory and subjectively rated disturbance (0=not at all; 4=extremely) in Exp. 1 (n=20). Means of error rates and ratings with standard errors are plotted.

The pattern of mean disturbance ratings follows the succession of error rate means. An important aspect is that soft auralised speech of bad intelligibility (S35_B) was clear-cut rated as more disturbing than silence (p<.01) whereas performance was just in tendency reduced during this sound condition. All other comparisons were also significant (p<.05).

EXPERIMENT 2: SUSTAINED ATTENTION (CONCENTRATION)

Method

In Experiment 2 a modified version of the Konzentrations–Leistungs–Test (KLT, Concentration Performance Test, [9]) was used. Here, three single-digit addition and subtraction problems, each consisting of three summands, were presented visually and successively. A certain calculation rule had to be applied to the intermediate results to obtain the final result. Performance in this task is interpreted as an indicator for sustained attention (concentration): If attention cannot be upheld and information processing is not coordinated adequately, the correct final result can only be obtained by chance. 24 trials had to be worked out during each sound condition. Background sounds and their presentation conditions, testing procedure and intermediate questioning remained unchanged from Experiment 1. 24 participants took part in the experiment.

Results

As Figure 3 indicates, significantly more errors were made during highly intelligible speech condition S55 than during silence (p<.01). Solely lowering the level was not sufficient to significantly reduce its detrimental impact (S35_G: p=.34). Instead, additionally reducing speech intelligibility was necessary (S35_B: p=.02). Error rates during soft and badly intelligible speech (S35_B) were not significantly raised compared to silence (p=.10). Subjective disturbance ratings, however, significantly differentiated between all four sound conditions (p<.01). Thus, in contrast to performance data, soft speech of high intelligibility (S35_G) was rated as significantly less disturbing than the original speech signal S55, and soft speech of bad intelligibility (S35_B) was rated as significantly more disturbing compared to silence.

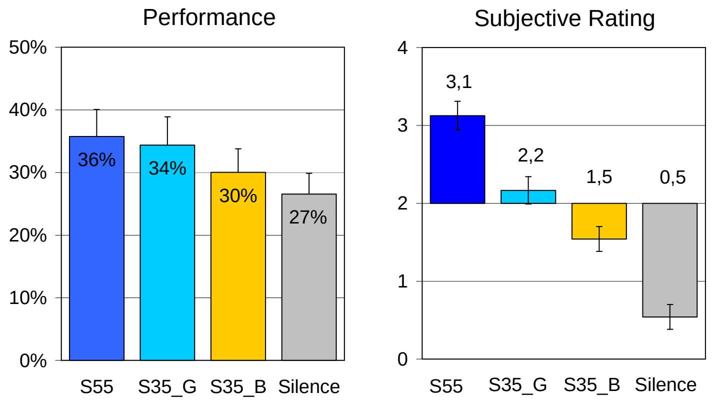

Figure 3.- Mental arithmetic performance and subjectively rated disturbance (0=not at all; 4=extremely) in Exp. 2 (n=24). Means of error rates and ratings with standard errors are plotted.

EXPERIMENT 3: VERBAL-LOGICAL REASONING

Method

The impact of speech sounds differing in intelligibility and level on verbal–logical reasoning was tested using a modified version of the grammatical reasoning test taken from the AGARD STRES Battery [10]. In this task, two visually presented declarative sentences were to be compared with a symbol order. This task requires mental translation of the partly negative and passive declarative sentences into requirements the symbol sequence is compared with.
28 participants worked out 25 trials during each sound condition. Background sounds and their presentation conditions, testing procedure as well as intermediate questioning were identical to Experiments 1 and 2.

Results

Regarding performance, two subgroups were given in this experiments: 13 participants made more errors during highly intelligible speech at 55 dB(A) (S55) than during silence (Speech Subgroup), whereas 12 participants made fewer errors during S55 compared to silence (Silence Subgroup). (Equal error rates during S55 and silence were produced by three participants who were excluded from further analysis.)
En detail, the Speech Subgroup made significantly more errors during the original speech signal (S55) in comparison to silence (p<.01; Figure 4). An exclusive reduction of sound level (S35_G) did not significantly decrease the detrimental impact (p=.11). During soft and badly intelligible speech (S35_B), however, significantly less errors were made than during the original speech signal S55 (p=.02). S35_B did not reduce performance significantly but still in tendency (p=.08). The Silence Subgroup, however, performed significantly better during highly intelligible speech at 55 dB(A) (S55) compared to the other three sound conditions ($p\leq$.02), which did not differ significantly ($p\geq$.30).
In contrast to differing patterns of performance, both the Speech Subgroup and the Silence Subgroup rated the disturbance impact evoked by the sound conditions analogously. Solely reducing level did not enhance subjective disturbance ratings significantly (S55 compared to S35_G: $p\geq$.10), but a reduction of level *and* intelligibility did (S55 compared to S35_B; $p\leq$.04). Disturbance ratings for the two soft speech conditions did not differ significantly depending on speech intelligibility ($p\geq$.13).
Since participants in the Silence Subgroup rated highly intelligible speech at 55 dB(A) (S55) also as the most disturbing sound condition, although they performed best during it, we

presume reactive effort enhancement [11] to underlie the performance pattern of this subgroup. Feeling disturbed may have encouraged these participants to activate and invest additional cognitive resources in terms of enhanced effort in task–solving and thus, during S55, overcompensation may have occurred.

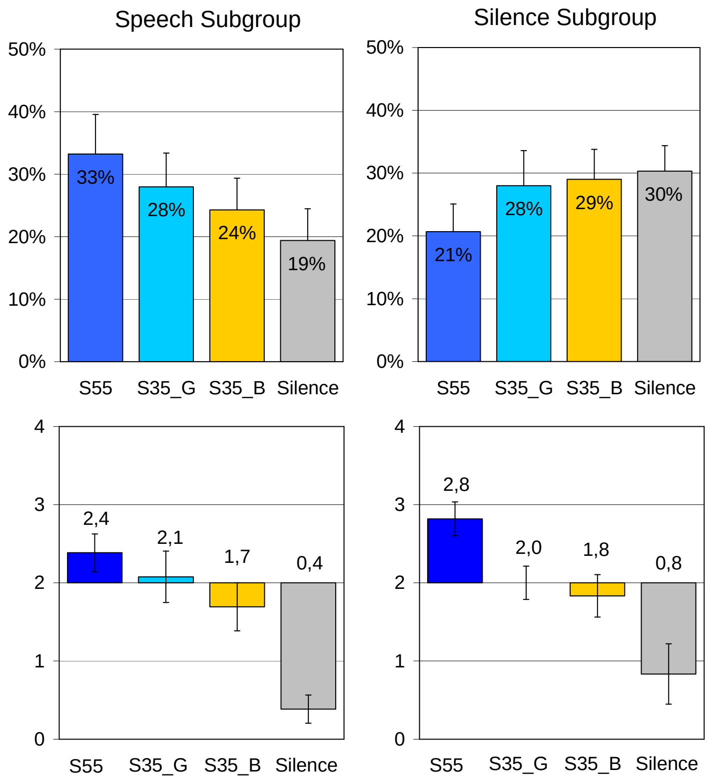

Figure 4.- Verbal-logical reasoning performance (upper panels) and subjectively rated disturbance (lower panels; 0=not at all; 4=extremely) in Exp. 3 (Speech Subgroup: *n*=13; Silence Subgroup: *n*=12). Means of error rates and ratings with standard errors are depicted.

Note that two alternative explanations for the subgroup split in this experiment can be ruled out. Firstly, the two subgroups did not differ regarding strategy use as a post-experimental interview revealed. And secondly, enhanced error rates during highly intelligible speech in the Silence Subgroup are not the result of a speed–accuracy trade–off [12]. This phenomenon describes a strategy to produce faster responses while making more errors. Reaction times, however, did not differ significantly during the four sound conditions for any of the two subgroups ($p \geq .21$), contrary to error rates.

CONCLUSIONS

The present experiments verify the basic cognitive functions verbal short-term memory (Exp. 1), sustained attention (Exp. 2) and verbal–logical reasoning (Exp. 3) to be significantly impaired by untreated and highly intelligible background speech at 55 dB(A). Its disturbance impact is significantly diminished by a combined reduction of level *and* intelligibility while a sole reduction of level does not suffice. These performance effects of the original speech signal and its alterations are consonant with the changing–state concept (see Introduction). This holds for the detrimental impact of unaltered speech, for the less negative effects when intelligibility of ambient speech is diminished and for the inefficiency of exclusively reducing overall level, as far as background speech remains audible [13]. The latter aspect is because, despite lowering the overall level, the prominence of the perceptual segmentation cues persists relative to the remaining sound signal, and with this, the changing-state characteristic of the background speech signal.

Even though a subgroup performed best during highly intelligible speech at 55 dB(A) in the verbal-logical reasoning task, at least two aspects have to be considered. First, compensating for the disturbance impact of background speech by putting more effort into task solving cannot be a solution if performance has to be sustained over a workday of several hours. Second, this subgroup rated highly intelligible speech of 55 dB(A) as the most disturbing sound condition, too. That the experimental groups produced inhomogeneous results in the verbal-logical reasoning task only illustrates that testing one, single, basic cognitive function can suffice neither to evaluate noise abatement measures in cognitive perspective nor to predict noise effects on a wide range of complex cognitive tasks. In contrast to the sustained attention test or the verbal short-term memory task, the verbal-logical reasoning task required untrained and non–automated mental transformation of the presented information and was thus sensitive to the amount of mental resources participants mobilise for task performance [14].

Besides performance data, subjective ratings were collected. These honoured a reduction of level mostly to a greater extent than a reduction of speech intelligibility, in contrast to performance data. Even soft and poorly intelligible speech was rated as significantly disturbing when compared with silence, although cognitive performance during this sound condition was not significantly diminished. Since noise abatement measures can only be successfully implemented if perceived as beneficial by the office employees involved, the soft and badly intelligible speech signal S35_B could be considered as a target sound for background speech that remains audible in an office environment, but holds the potential of further optimisation.

Taken together, the present data underlines the necessity to reduce background speech intelligibility with respect to cognitive performance since its reduction corresponds with enhanced task performance compared to working conditions in which background speech is highly intelligible. On the other hand, the present data illustrates that subjective ratings and objective performance data can differ greatly. Consequently, the combination of objective performance tests and subjective ratings is indispensable for the comprehensive evaluation of acoustic office environments and their alterations.

References

[1] S.P. Banbury, D.C. Berry: Office noise and employee concentration: Identifying causes of disruption and potential improvements. Ergonomics **48(1)** (2005) 25-37.

[2] J.S. Bradley, B.N. Gover (2004): Criteria for acoustic comfort in open–plan offices. In Proceedings of the inter–noise 2004, the 33rd International Congress and Exposition on Noise Control Engineering, Prague, Czech Republic.

[3] A. Baddeley: Working memory and language: An overview. Journal of Communication Disorders **36(3)** (2000) 189-208.

[4] D.M. Jones, C. Madden, C. Miles: Privileged access by irrelevant speech to short-term memory: The role of changing state. Quarterly Journal of Experimental Psychology: Human Experimental Psychology **44(4)** (1992) 645-669.

[5] J.R. Anderson (2004): Cognitive psychology and its implications. Palgrave Macmillan.

[6] W. Ellermeier, J. Hellbrück: Is level irrelevant in 'Irrelevant Speech'? Effects of loudness, signal-to-noise ratio, and binaural unmasking. Journal of Experimental Psychology: Human Perception & Performance **24(5)** (1998) 1406-1414.

[7] N. Venetjoki, A. Kaarlela-Tuomaala, E. Keskinen, V. Hongisto, V.: The effect of speech and speech intelligibility on task performance. Ergonomics **49(11)** (2006) 1068-1091.

[8] M. Vorländer, R. Thaden: Auralisation of airborne sound insulation in buildings. ACUSTICA united with ACTA ACUSTICA **86(2)** (2000) 76-89.

[9] H. Lukesch, S. Mayrhofer (2001): KLT-R: Konzentrations–Leistungs–Test, Revidierte Fassung. Göttingen: Hogrefe.

[10] Advisory Group for Aerospace Research and Development (1989): Human performace assessment methods. Neuilly–sur–Seine, France: NATO.

[11] D. Kahneman (1973): Attention and effort. Englewood Cliffs, NJ: Prentice Hall.

[12] R. Hockey (1984): Varieties of attentional state: The effects of environment. In R. Parasuraman & D.R. Davies (eds.): Varieties of attention, pp. 449–483. Orlando: Academic Press.

[13] S. Tremblay, D.M. Jones: Change of intensity fails to produce an irrelevant sound effect: Implications for the representation of unattended sound. Journal of Experimental Psychology: Human Perception & Performance **25(4)** (1999) 1005-1015.

[14] G. Matthews, T.J. Sparkes, H.M. Bygrave: Attentional overload, stress, and simulated driving performance. Human Performance **9** (1996) 77–101.