Applied Acoustics 110 (2016) 170–175

Contents lists available at ScienceDirect

Applied Acoustics

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

The effects of speech intelligibility and temporal–spectral variability on performance and annoyance ratings

Andreas Liebl [a,*], Alexander Assfalg [b], Sabine J. Schlittmeier [b,c,d]

[a] Fraunhofer Institute for Building Physics, Nobelstraße 12, 70569 Stuttgart, Germany
[b] Catholic University of Eichstaett-Ingolstadt, Ostenstraße 25, 85072 Eichstaett, Germany
[c] University of Kaiserslautern, Erwin-Schrödiger-Straße, 67663 Kaiserslautern, Germany
[d] HSD Hochschule Döpfer, Waidmarkt 3 und 9, 50676 Cologne, Germany

ARTICLE INFO

Article history:
Received 20 June 2015
Received in revised form 19 February 2016
Accepted 21 March 2016
Available online 31 March 2016

Keywords:
Noise
Office
Performance
Annoyance

ABSTRACT

Ambient sound can impair verbal short-term memory performance. This finding is relevant to the acoustic optimization of open-plan offices. Two algorithmic approaches claim to model the impairment during a given sound condition. One model is based on the Speech Transmission Index (STI). The other approach relies on the hearing sensation fluctuation strength (*F*). Within the scope of our consulting activities the approach based on *F* can hardly be applied and the model based on the STI is often misinterpreted in terms of semanticity. Therefore we put to test the two models and elucidate the relevance of temporal–spectral variability and semanticity of background sound with regard to impairment of performance. A group of 24 subjects performed a short-term memory task and rated perceived annoyance during eight different speech and speech-like noise conditions, which varied with regard to STI and *F*. The empirical data is compared to the model predictions, which only partly cover the experimental results. Speech impairs performance more than all other sound conditions and variable speech-like noise is more impairing than continuous speech-like noise. Sound masking with continuous speech-like noise provides relief from the negative effect of background speech. This positive effect is more pronounced if the signal to noise ratio is −3 dB(A) or even lower.

1. Introduction

Ambient speech is known to impair individual task performance at silent, concentrated work [1,2]. The detrimental impact of background speech on verbal short-term memory was reported for the first time by Colle and Welsh [3]. Since then the phenomenon has been object of a mere uncountable number of studies and is termed [4] as the Irrelevant Sound Effect (ISE). It describes a significant impairment of verbal short-term memory during presentation of certain sound conditions, even so the material to be remembered is presented visually and the sound condition is supposed to be ignored. Most of these studies were motivated by a basic research perspective and intended insight into automatic and obligatory aspects of speech and sound processing in verbal short-term memory. Consequently they focused on necessary and sufficient characteristics of sound conditions for eliciting a detrimental impact. By doing so, a prerequisite for an ISE to arise has been revealed which is that the sound condition is a so-called changing-state sound [5]. A changing-state sound is characterized by a distinct temporal structure with different auditory-perceptive tokens varying successively (for further information cp. e.g. [6]).

Performance of verbal short-term memory is usually operationalized by the serial recall task where visual items (digits, letters) are presented sequentially and have to be recalled afterwards in the strict order of presentation. The performance decrement in this task during changing-state sound is explained by the interference by-process principle [7–9]. A decline of performance occurs if automatic processing of the sound condition and voluntary task processing call for the same cognitive resources. Correspondingly, the ISE is assumed to be due to a conflict between the intentional serial processing of the items of the serial recall task and the automatic and involuntary serial processing of the successively changing auditory-perceptive tokens of the sound condition. Such interference effects, which are highly specific to the correspondence of sound and task characteristics, have been verified recently to be qualitatively different from the so-called deviation effect. The latter is described by the distraction of attention from the task at hand due to unexpected changes in

* Corresponding author.
E-mail addresses: andreas.liebl@ibp.fraunhofer.de (A. Liebl), a.assfalg@gmail.com (A. Assfalg), sabine.schlittmeier@sowi.uni-kl.de (S.J. Schlittmeier).

http://dx.doi.org/10.1016/j.apacoust.2016.03.019

the acoustic background, for example by a telephone starting to ring or a change in background speaker voice [10]. This differentiation of specific interference processes and auditory distraction effects is the core assumption of the duplex-mechanism account of auditory distraction, a recently published framework of cognitive noise effects [11,12].

Yet, the detrimental impact of sound conditions on cognitive performance is not only a matter of basic research in the field of cognitive psychology but also an aspect of everyday real life and relevant in many contexts, e.g. in open-plan and group offices, in noisy classrooms, at production lines and construction sites. Therefore the question of which sound condition characteristics are decisive for an impairment of performance to occur is also of economic interest. Consequently, the performance impact of sound conditions have been investigated during the last years in a series of applied studies focusing on open-plan and group offices on the one hand and background speech on the other hand [13–15].

In each of these two lines, basic and applied research, an algorithmic model has been developed to estimate the expectable performance decrement during a given sound condition. From an applied perspective, an algorithmic model is helpful for designing and evaluating working environments with the objective to guarantee for employees optimum performance and well-being under consideration of the cost-value ratio. In basic research, an algorithmic model adds to the understanding of basic cognitive functions and provides quantifiable hypotheses as compared to making only qualitative predictions. The two different approaches will be outlined in the following and applied in the present experiment.

1.1. Theoretical background

The first approach to be considered here has been presented by Hongisto [1] and is driven by an applied perspective. It was motivated by the fact that employees in open-plan and group offices judge irrelevant background speech produced by colleagues talking on the phone or with other colleagues to be one of the severest problems in their working environment as it is perceived to be annoying and disturbing (e.g. [15]). However, talking colleagues cannot be banned from open-plan offices, so that the main focus of most interventions is to alter room acoustics in a way that irrelevant background speech becomes less intelligible on its transmission path from a talker to an involuntarily listening co-worker (cp. [15]). This stems from empirical results verifying that reducing background speech's intelligibility reduces its detrimental impact on task performance and perceived annoyance [1,2,13–19].

Against this background, the approach by Hongisto [1] modeling the detrimental impact of sound conditions has its foundations in room acoustics and focuses on background speech. Hongisto [1] provided a model according to Eq. (1) that predicts how much performance is reduced due to speech of varying Speech Transmission Index (STI). The STI is an instrumental measure of speech intelligibility in rooms [20]. In principle it is based on the estimated degradation of a given speech signal on its transmission path from the source (talker) to the listener's ear (receiver). Its value varies from 0 (completely unintelligible) to 1 (perfect intelligibility). According to the model increasing speech intelligibility – reflected by increasing STI values – causes a decrease of performance (DP).

$$\mathrm{DP(STI)} = \frac{-7}{1+\exp[(\mathrm{STI}-0.4)/0.06]} + 7 \quad (1)$$

Even so the model claims to predict cognitive performance in applied settings the data base relies on two laboratory studies and only one field study, focusing on verbal short-term memory performance, proofreading and self-estimated daily waste of working time due to the presence of disturbing sound conditions. For more detailed information the reader is referred to Hongisto [1].

The second approach modeling the disruption of task performance during irrelevant sound conditions has been proposed by Schlittmeier et al. [6]. It differs in several perspectives from the foresaid approach by Hongisto [1]. To start with, the model applies not only to speech but also to non-speech sounds while explicitly focusing on the decrease of short-term memory performance during irrelevant sound. This phenomenon was described above as the ISE. Accordingly the approach presented by Schlittmeier et al. [6] is driven by a basic research perspective. Its defining characteristic is to estimate the performance decrement during a certain sound based on the hearing sensation fluctuation strength (F) as expressed in Eq. (2). This quantity, which can also be instrumentally measured, is named after the hearing percept induced when listening to sounds which are slowly modulated regarding to frequency or amplitude ($f_{\mathrm{mod}} < 20$ Hz). The data base from which Eq. (2) is derived incorporates 70 behavioral outcome measures for 40 different sounds, e.g. background speech, music, traffic noise, tone sequences and office noise. For more detailed information the reader is referred to Schlittmeier et al. [6].

$$\mathrm{DP}(F) = \mathrm{ISE} = \frac{F}{0.68\ \mathrm{vacil}} \cdot 7.5 \quad (2)$$

1.2. Research intent

The two foresaid models claim to predict the decline of individual task performance due to the presence of sound but they rely on different parameters, namely the STI [1] and the hearing sensation F [6]. Both models have strengths and shortcomings, which need to be addressed. In our practical consultative work we have learned that the approach based on F can hardly be applied in real offices since the signal to noise ratio between disturbing background speech and neutral steady background sound is often too low. The approach based on the STI is often misinterpreted by facility managers and office owners who equate the physical quality measured by the STI with speech intelligibility and especially with semanticity of background speech. As a consequence it is argued that disturbance is mainly a matter of the employees curiosity and that they could just ignore the background speech of their colleagues. We conducted a laboratory experiment to put to test the two models and to elucidate the relevance of temporal–spectral variability and semanticity of background sound. In particular, we wanted to find out about the effects of manipulations of intelligibility and temporal–spectral variability with regard to the decline of individual task performance. Therefore, the effects of several speech and speech-like noise signals on verbal short-term memory were tested and compared to a silent reference condition. In addition to objective performance scores, subjective annoyance ratings were collected for each sound condition.

2. Methods and materials

2.1. Sample

24 students of the Catholic University of Eichstaett-Ingolstadt took part in the experiment (Age: 20–33 years; Median: Md = 25 years; Gender: 4 male, 20 female). All participants were native German speakers and reported normal hearing ability. A small allowance was paid for participation.