Applied Acoustics 126 (2017) 68–80

Contents lists available at ScienceDirect

Applied Acoustics

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

Auditory distraction in open-plan office environments: The effect of multi-talker acoustics

Manuj Yadav *, Jungsoo Kim, Densil Cabrera, Richard de Dear

School of Architecture, Design and Planning, The University of Sydney, Australia

ARTICLE INFO

Article history:
Received 11 January 2017
Received in revised form 27 April 2017
Accepted 8 May 2017

Keywords:
Open-plan office acoustics
Irrelevant sound effect
Multi-talker simulation
Psychoacoustics
Indoor environmental quality

ABSTRACT

Within the soundscapes of open-plan offices, irrelevant speech has consistently been reported as the most distracting, and causing performance decrements for workers. Notwithstanding this generalization, the 'babble' created by multiple simultaneously active talkers can sometimes provide beneficial sound masking, but due to spatial release from masking (SRM), speech may still be sufficiently intelligible up to a certain number of talkers (estimated to be about four). This was explored within a highly-realistic office simulation, where the cognitive performance, and subjective distraction of participants were tested. The experimental design was a 4×2 factorial (4 talker numbers, 2 levels of broadband sound masking, as the factors). The results indicated that within lower sound pressure level (SPL) of broadband sound masking, multi-talker sound environments degraded cognitive tasks performance more than those with a single talker, suggesting SRM effects. For higher SPL broadband sound masking, the cognitive test scores were similar within the different talker numbers. The subjective distraction increased monotonically with the number of talkers, with higher distraction within lower SPL broadband sound masking. Overall, the results call into question the single talker assumption (being the most distracting) within the international standard for measuring open-plan office acoustic environments (ISO 3382-3:2012). Soundscapes with 4 simultaneous talkers were still not adequately providing beneficial 'babble' masking, and were more distracting than 1 active talker. In conclusion, it is suggested that the acoustics environment of open-plan offices needs better characterization by incorporating some of the complexity and psychoacoustics of multi-talker scenarios.

Crown Copyright © 2017 Published by Elsevier Ltd. All rights reserved.

1. Introduction

1.1. Background

When compared to cellular or private office configurations, there is mounting evidence against open-plan offices, in the form of demonstrable dissatisfaction amongst its occupants, and a decline in the task-based performance [1–4]. Overall, one of the main attractions of the open-plan layout – the presumed increase in productivity due to increased ease of interaction between the workers, has been shown to be substantially offset by a number of factors such as distraction due to noise and privacy issues (see reviews in [5–7]). From the perspective of a worker engaged in cognitively intensive tasks, it seems ironic that this major benefit of the shared work environment and ease of interaction, in fact, becomes the root of the problem. Intelligible speech-based communication affects both the talker and the listener (active or passive) as it has the potential to distract a *listener* who may need to concentrate on a task, and make a *talker* anxious about lack of speech privacy. This reflects in the findings of studies showing that, of all the indoor environmental quality (IEQ) factors, perceived quality of acoustics scores lowest in open-plan offices, where both intelligible speech and non-speech stimuli [7–12] have been reported as the contributing factors.

Given the significance of acoustics in the overall perception of IEQ within open-plan offices, the standardisation of some speech-related room acoustics measures in ISO 3382-3 (based largely on [13]) represents a useful development. Briefly, ISO 3382-3 considers the acoustic environment in the form of speech intelligibility and background noise distribution around workstations [14].

Abbreviations: IEQ, Indoor environmental quality; ISE, Irrelevant sound effect; SNR, Signal-to-noise ratio; SRM, Spatial release from masking; SRT, Speech reception threshold; STI, Speech transmission index; SI, Speech intelligibility; mr-sEPSM, Multi-resolution speech based envelope power spectrum model; SPL, Sound pressure level; NC, Noise criterion; RNC, Room noise criterion.

* Corresponding author at: Room 589, 148, City Road, The University of Sydney, 2006 NSW, Australia.

E-mail address: manuj.yadav@sydney.edu.au (M. Yadav).

http://dx.doi.org/10.1016/j.apacoust.2017.05.011
0003-682X/Crown Copyright © 2017 Published by Elsevier Ltd. All rights reserved.

However, there appear to be a few fundamental issues with ISO 3382-3, arising mainly from certain simplifications and assumptions, which may limit its applicability. Specifically, the parameters in ISO 3382-3 are based around the assumption that the most distracting scenario in an open-plan office is one active talker, compared to when more than one talker is simultaneously active. This assumption seems to contradict extant literature regarding the psychoacoustics of auditory distraction in multi-talker environments (see reviews in [15,16]) such as an open-plan office. Studies have shown that a combination of few voices that are spatially seperated can be similarly distracting (in terms of subjective ratings and cognitive performance) as a single voice [17–19]. This paper explores the validity of single talker assumption within ISO 3382-3, by first presenting a literature review of multi-talker psychoacoustics, as it applies to open-plan offices. This is followed by reporting an experiment in which hypotheses that address the assumptions of ISO 3382-3 were tested within an open-plan office simulation that had a realistic multi-talker sound environment.

1.2. Auditory distraction due to the irrelevant sound effect

The degree of auditory distraction and the associated decline in cognitive performance within an open-plan office sound environment has been shown to be affected by the degree of uncontrolled (or, to-be-ignored) audition of irrelevant sounds: the so-called *irrelevant sound effect* (ISE; first studied by [20], see review in [21]). In this regard, what makes an irrelevant sound *stream* distracting, consistently over time, has been encapsulated within the so-called *changing-state* hypothesis (first advanced in [22]). Within the changing-state hypothesis, the sound stream can be thought in terms of segments, which vary in their acoustic-perceptual properties over time (speech being a prime example). For such segments, the extent or the degree of distraction increases with the extent to which the segments change state within the sound stream [23]. As an often cited illustration of the changing-state hypothesis, a sound stream that is composed of segments 'a b c b c a b...' has been shown to be more distractive than steady-state sound streams 'a a a a a a' (e.g., a repeating tone, or steady-state broadband noise profiles). Taking this example further, it can be argued, as was done by Macken et al. [18] (p. 45), that segments from individual voices in a multi-talker environment (each voice registering certain acoustic-perceptual attributes) can exhibit a changing-state to elicit the ISE.

It is interesting that in the presence of an irrelevant sound stream, the auditory system does not habituate to selectively limit the distraction. This is presumably due to evolutionary reasons, as a shift in relevance of the ambient sounds may sometimes require an immediate response. This applies not only in cases where there is an imminent danger, but also in more benign cases where the shift implies that the hitherto unwanted sounds begin carrying task-specific or important information. This has been typified within the duplex theory for the irrelevant sound effect. The duplex theory highlights the important differences between the *functional mechanisms* by which the irrelevant changing-state sound stream is assumed to cause distraction: either by *interference-by-process*, or *attention capture* (details in [24]). The latter relates to a momentary shift in attention from the focal task due to an unexpected change in the sound environment, which could contain either specific (e.g., their name being called) or generic cues (e.g., phone ringing) to the listeners. The former (interference-by-process) has been shown to occur when the irrelevant sounds, or segments, compete for the same cognitive processing resources as those required by the voluntary task within a certain cognitive domain, e.g., serial short-term memory [25], semantic processing [26,27]; or more typically, a combination of cognitive domains.

Since the focus in this paper is on speech-based distraction, it is worthwhile to briefly describe speech as a signal within the context of duplex theory of ISE; how its changing-state nature varies with the number of talkers; and the psychoacoustical literature relating to multi-talker environments. To begin with, speech can be considered as a 'special' signal with a high degree of redundancy in its information coding, which renders it robust against many types of system and environmental distortions, at least from the point of view of *speech intelligibility* ([15], pp. 1467). Let us consider a simple example of speech from one talker, where the energy is generally concentrated in discrete speech tokens, or segments. Such segments exhibit natural envelope modulations (or, rhythms) in the temporal and spectral domains. These spectrotemporal envelope modulations, along with the gaps between these segments typify the changing-state nature of a stream of speech, which, of course, incorporates the audition of both relevant and irrelevant speech.

Moving on from the simple case of one talker, adding more voices to the signal can have the effect of 'filling in' (or, *energetic* masking [28]) these aforementioned gaps [29]. Depending on the extent of such filling in, adding more voices can reduce the advantage (alternatively, reduce the disadvantage) of 'listening in the gaps', which, in effect, may reduce the changing-state effect. This is similar to adding spectrally shaped broadband noise, which has also been shown to be effective in reducing the 'listening in the gaps', or changing-state effect of speech [23]. Such broadband noise treatment with different kinds of spectral manipulations is generally also referred to as 'sound masking' in many commercial and scientific applications [30]. The effect of adding such 'sound masking' was also explored in this paper, as seen in Section 1.4. Another perspective in terms of adding more voices to the one talker signal can be seen as the averaging out of the modulations in individual voices, e.g., speech-on-speech *masking* over time (example of *energetic* masking), leading to reduced signal-to-noise ratio (SNR), and hence, reduced speech intelligibility. Other types of masking may include *informational* (i.e., semantic) [28] and, perhaps, *perceptual* masking ([31] p. 119), and collectively, a mixture of many voices essentially leads to the condition of speech 'babble' in the form of multi-talker speech where the streams from individual talkers may not be segmented enough to be considered distracting, and are in fact beneficial [17]. In other words, speech 'babble' has the potential to, in fact, reduce the ISE. In the context of the current experiment design, primarily the interference-by-process aspect of the ISE was more relevant, compared to the attention capture mechanism. This will be further explained in Section 2.2. However, what is important to consider here is that the conceptualisation of speech babble, thus far, has been from a largely signal perspective, while ignoring several environmental and psychoacoustical factors; for instance, the spatial arrangement of talkers in the multi-talker babble, familiarity with voices in a real scenario that could aid segmentation, etc.

Within real listening environments, such as an open-plan office, the aforementioned speech-on-speech, or babble masking, and the purported decline in the changing-state nature of the sound stream leading to babble needs careful qualification (the following ignores any contextual issues such workplace dynamics and cultural factors, which are also likely to be quite important). One of the main reasons being that, due to the spatial separation of talkers, speech-on-speech masking is still subject to *spatial release from masking* (SRM) (see review in [31] pp. 120–124) before the sound reaches the ears (i.e., peripheral masking) [15,32]. SRM, which is essentially a monaural effect, generally means that the extent of masking decreases with a spatial separation between speech sources. The *binaural advantage* (or, disadvantage in relation to distraction), and head movements can further enhance the signal-to-noise ratio,

دریافت فوری ← متن کامل مقاله

ISIArticles
مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات