Applied Acoustics 101 (2016) 156–167

Contents lists available at ScienceDirect

Applied Acoustics

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

Release from masking of speech intelligibility due to fluctuating ambient noise in open-plan offices

P. Chevret

Institut National de Recherche et de Sécurité, INRS, Rue du Morvan, CS 60027, 54519 Vandœuvre Les Nancy Cedex, France

ARTICLE INFO

Article history:
Received 5 June 2015
Received in revised form 3 August 2015
Accepted 20 August 2015
Available online 3 September 2015

Keywords:
Open plan office
Intelligibility
Indicator

ABSTRACT

The sensitivity of intelligibility to ambient noise modulation is analyzed in this paper by means of two laboratory tests. The first one consists of measuring release from masking using acoustic samples composed of speech noise synthesized in the laboratory, but also real signals measured in two open plan offices. The Speech Reception Threshold values obtained from the experiment for the synthetic signals are compared with the literature. Those obtained for the real samples show significant differences as a function of the open-plane office. Secondly, a new indicator based on the calculation of the speech transmission index (IEC 60268-16, 2011) is proposed, by taking into account the modulation of the ambient noise in the 4 Hz octave band. Simultaneously, a second intelligibility experiment is performed, resulting in a very good correlation between the intelligibility scores and the new indicator. To complete the evaluation, the sound samples of the first test are reused to demonstrate that the indicator is a good descriptor of release from masking due to the modulation of ambient noise for the synthetic signals and that it is capable of classifying work spaces in terms of level of distraction.

1. Introduction

It is now accepted that noise is one of the main disturbance factors in open-plan offices [43]. In most cases, the disturbance is not sufficient to totally stop employees from doing a task. However, it can reduce their job satisfaction [31], and their performance [4]. In addition, surveys conducted on employees working in open-plan offices show that noise annoyance can have consequences on their health because it acts either directly as a fatigue factor, or indirectly as a co-factor of other causes of malaise or of stress [37,20].

Such surveys also reveal that although ambient noise in an open-plan workspace is made up of a variety of sound sources such as noise from services and equipment (air-conditioning, printers, and telephones), and noise from footsteps or doors banging, the main source of distraction is the mixture of speech from conversations between colleagues or from telephone conversations [39]. Reducing intelligibility in an open space office is an objective that has direct positive consequences on the disturbance generated by ambient noise on a person who wishes to concentrate on a cognitive task, but it will also improve the level of privacy, meaning the possibility of maintaining a certain level of confidentiality in conversations. It was to this end that standard ISO 3382-3 [24], based in particular on the works of Virjonen et al. [49] and Bradley [5] defined single number indicators which are the main operational tool for dealing with open office acoustics. The level of distraction and privacy is generally defined using an intelligibility indicator, whether it be the Articulation Index [1] as suggested by standard ASTM E1130-08 [3] or the *STI* (Speech Transmission Index: [23]), recommended by standard ISO 3382-3 [24].

Indeed, the evaluation of intelligibility through a single number indicator can be considered in several ways. The *SII* (Speech Intelligibility Index: [2]) like the Articulation Index [1], is based on the idea that low intelligibility is mainly due to the energy of the speech signal being masked by the mean ambient noise, this masking being described by the signal-to-noise ratio in the relevant speech bands. The method of computing this indicator was inspired directly by the work of French and Steinberg [17], Fletcher and Galt [16], or Kryter [29,30]. The *STI* is based on the work by Steeneken and Houtgast [45,46] which shows that the intelligibility of a speech signal depends on the decrease in modulation of the signal through a transmission system. That indicator is particularly suitable for telephony, for architectural acoustics, or for studying the consequences of hearing loss on intelligibility in hearing-impaired people.

Even though those indices are frequently used for evaluating the quality of workspaces [50,6], they have so far yielded only incomplete information. By way of illustration, Virjonen et al. [50], who recommended using the *STI*, deliberately ignored the

E-mail address: patrick.chevret@inrs.fr

http://dx.doi.org/10.1016/j.apacoust.2015.08.010
0003-682X/

presence of people in the room: only the background noise from services and equipment (whose level is generally much lower than the occupancy noise), was taken into account. In such a situation, the *STI* merely characterizes the intelligibility potential of the room, and in no way characterizes intelligibility during work. For their part, Bradley and Gover [6] proposed using standard ambient noise spectra. This approach is relevant insofar as the aim is to make comparisons between rooms for identical uses or occupancies. However, the use of the mean noise spectrum is not sufficient for characterizing intelligibility correctly.

The presence of successive periods of silence and energy peaks in the ambient noise gives rise to a phenomenon of release from masking, whereby the energy of the conversations is unmasked, thus enhancing intelligibility. Rhebergen and Versfeld [41] and Rhebergen et al. [42] proposed an approach for objectifying this phenomenon through a short-term version of the *SII* that they call the "*ESII*" for "Extended Speech Intelligibility Index". The principle is to define an analysis window whose size depends on the temporal resolution of the human ear, and to apply processing similar to the processing in the standard ANSI S3.5-1997 [2] to obtain a value for the *SII* at each position of the time window. The overall intelligibility index is the mean of the *SII* s over the duration of the signal processed. George et al. [19] used this approach in order to evaluate the release from masking due to reverberation and non-stationary noise. Cooke [10] adopted the same approach, but generalized to the time–frequency domain, by defining "glimpses". He showed that the percentage of occurrence of such glimpses in ambient noise has a direct effect on consonant intelligibility. Rhebergen et al. [42] managed to show that the *ESII* does indeed describe release from masking in the presence of synthesized ambient noise (saw-tooth, sinusoidal, and square-wave noise) and of speech noise, except for low modulations (4 Hz), but those authors attributed the differences observed to their procedure for determining the Speech Reception Threshold (*SRT*). Basically, the criticism leveled at these methods is that they are not capable of taking account of non-linear changes that might be made by the transmission system, such as envelope compression or spectral subtraction. In addition to this restriction, their main limitation lies, in our opinion, in the choice of analysis window. These authors often based this choice on the results of Moore [34–36], but they admitted that it was difficult to commit themselves to an indisputable value: Rhebergen et al. proposed a size of 9.4 ms at low frequencies and as large as 35 ms at high frequencies. However, it was stated that if a fixed size is desired, 12 ms is a good compromise. For their part, George et al. [18] adjusted these values from 1.9 ms to 5.8 ms, while Cooke [10] proposed a constant window of 8 ms.

There is another family of methods that could lead to incorporating the effects of ambient noise on intelligibility with more precision. These methods suggest modeling based on calculating the power spectrum of the envelope of the noisy signal and of the ambient noise [11,26]. An extension to fluctuating ambient noise was proposed by Jørgensen et al. [27] (mr-sEPSM, for the "Multiresolution Speech-Based Envelope Power Spectrum Model"), by using an analysis window over the envelope, whose size is inversely proportional to the modulation frequency. The same philosophy was applied by Schwerin and Paliwal [44] for the development of the quasi-stationary speech transmission index (QSTI), except that the inputs of the analysis are the degraded and clean signals. The advantage of such methods is that they are robust to any non-linear modifications that might be made to the signals (spectral subtraction, envelope compression, etc.) They are particularly useful for applications to fitting hearing-impaired people with hearing aids, but such applications lie well beyond the scope of this article. A drawback with such approaches is that they require either that the waveforms of the speech signals and noise signals be known, or else that it be possible to establish statistics on the basis of a set of waveforms. Given that limitation and the fact that they are particularly advantageous for complex operations on the signals, these methods were not used in this study.

The goal of this paper is to quantify the effect of ambient noise modulation on intelligibility in open-plan offices. To this end, we based our approach on Standard IEC 60268-16 [23] for calculating the *STI* in the presence of steady-state ambient noise, and we adapted it while explicitly taking account of ambient noise modulation, which is a physical descriptor directly related to a characteristic of the room, such as density of occupancy. Using field sound samples, we set up laboratory intelligibility tests in order to determine not only the effect of such modulation but also the relevance of this new indicator.

In the first part of this paper, we describe how we conducted a *SRT* test in order to evaluate the sensitivity of the speech reception threshold to fluctuations in ambient noise, in particular on the basis of sound samples collected in the field. Although this type of test had already been conducted several times in the past with different languages and different types of presentation, it provides several new elements. Firstly, it supplemented the study by Jansen et al. [25], which is certainly the most comprehensive study on intelligibility of French phrases in steady-state noise. Secondly, this test allowed highlighting a significant difference of intelligibility between two different open-space offices, justifying a second series of laboratory tests to obtain a relation between a physical indicator and intelligibility scores.

This second series of test is developed in the second section, in which we also present a new approach for the development of an intelligibility indicator that could be used in open-plan offices. After having shown that this indicator is strongly correlated with the intelligibility of phrases, its capacity to predict release from masking by modulated ambient noise is studied on the basis of the signals of the first experiment. In the case of synthetic signals, the masking values measured during the first experiment are predicted successfully. In the case of real samples, we are able to demonstrate its capacity to rank open-space offices in terms of level of distraction.

2. Evaluation of *SRT* sensitivity to ambient noise modulation

Many factors contribute directly to the intelligibility of a conversation in a sound environment. If we omit the temporal modifications affecting the target signal during its passage through a transmission system (communication system, rooms with high reverberation capacity), it is customary to group these factors into two categories: energy masking which groups factors contributing to placing the target signal in competition with the masking signal in the spectral domain and in the temporal domain, and the informational masking that occurs when similarities exist between the target signal and the masking signal (for example, the competition between two voices of the same gender are a priori marked by strong informational masking). Many works can be found in the literature on the effects of these two types of masking on intelligibility. In particular, the effects of modulating ambient noise on release from masking have been the subject of numerous studies with sometimes different results, depending in particular on laboratory conditions (see the synthesis presented in Appendix A).

However, since most of these works are performed for parametric purposes, they rely on using signals synthesized in the laboratory (white noise, two-band uncorrelated noise, sine wave, speech from a trained reader, etc.) However, under real conditions, especially in an open-plan office, the sound environment is very complex as it results from a mixture of speech noises produced by men and women whose number and level depend in particular