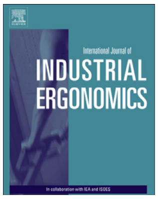
International Journal of Industrial Ergonomics 53 (2016) 372–379

Contents lists available at ScienceDirect

International Journal of Industrial Ergonomics

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



Echo threshold between passive and electro-acoustic transmission paths in digital hearing protection devices

Narimene Lezzoum, Ghyslain Gagnon, Jérémie Voix*

École de technologie supérieure, 1100 Notre Dame West, Montréal, QC H3C 1K3, Canada

ARTICLE INFO

Article history:
Received 17 March 2015
Received in revised form
22 February 2016
Accepted 7 April 2016
Available online 24 April 2016

Keywords:
Hearing protection device
Echo threshold
Just noticeable difference
Delay

ABSTRACT

Electronic hearing protection devices are increasingly used in noisy environments. Theses devices feature a miniaturized external microphone and internal loudspeaker in addition to an analog or digital electronic circuit. They can transmit useful audio signals such as speech and warning signals to the protected ear and can reduce the sound pressure level using dynamic range compression. In the case of a digital electronic circuit, the transmission of audio signals may be noticeably delayed because of the latency introduced by the digital signal processor and by the analog-to-digital and digital-to-analog converters. These delayed audio signals will hence interfere with the audio signals perceived naturally through the passive acoustical path of the device. The proposed study presents an original procedure to evaluate, for two representative passive earplugs, the shortest delay at which human listeners start to perceive two sounds composed of the signal transmitted through the electronic circuit and the passively transmitted signal. This shortest delay is called the *echo threshold* and represents the delay between the time of perception of *one fused sound* from *two separate sounds*. In this study, a transient signal, a clean speech signal, a speech signal corrupted by factory noise, and a speech signal corrupted by babble noise are used to determine the echo thresholds of the two earplugs. Twenty untrained listeners participated in this study, and were asked to determine the echo thresholds using a test software in which attenuated signals are delayed from the original signals in real-time. The findings show that when using hearing devices, the echo threshold depends on four parameters: (a) the attenuation function of the device, (b) the duration of the signal, (c) the level of the background noise and (d) the type of background noise. Defined here as the shortest time delay at which at least 20% of the participants noticed an echo, the echo threshold was found to be 8 ms for a bell signal, 16 ms for clean speech and 22 ms for speech corrupted by babble noise when using a shallow earplug fit. When using a deep fit, the echo threshold was found to be 18 ms for a bell signal and 26 ms for clean speech and 68 ms for speech in factory. No echo threshold could be clearly determined for the speech signal in babble noise with a deep earplug fit.

1. Introduction

Occupational hearing loss is the most common work injury in North America with approximately 22 million workers exposed daily to hazardous noise (NIOSH, 1998). To prevent hearing loss, wearing Hearing Protection Devices (HPD) becomes a necessity in industrial workplaces. In fact, wearing HPDs is also required nowadays for professional musicians since they too are exposed to loud sounds and thus vulnerable to hearing loss (Macdonald et al., 2008).

HPDs come in various forms. There are earplugs, which must be placed within or against the entrance of the ear canal, and earmuffs, which either fit around the ear, or in the form of helmets, encasing the entire head (Berger, 2003). HPDs can be grouped in two types of operating mode: passive HPDs and active (or electronic) HPDs (Casali, 2010).

Passive HPDs are the traditional HPDs. They reduce the background noise mechanically based on their shape and material composition, while electronic HPDs are equipped with an external microphone to capture the signals, an internal loudspeaker to playback the signals under the protected ear and an analog circuit or a Digital Signal Processor (DSP) in order to process the incoming signals in real-time (Casali, 2010). Electronic HPDs are increasingly used by workers, musicians, and the military for their high

---

* Corresponding author.
E-mail address: jeremie.voix@etsmtl.ca (J. Voix).

http://dx.doi.org/10.1016/j.ergon.2016.04.004

flexibility and multiple functionalities such as active noise control (Brimhall et al., 2002) or adaptive gain control (Hotvet, 1996).

Recently, some advanced functionalities have been developed for electronic HPDs as listed in (Voix, 2014), such as background noise reduction (Chung, 2007) and (Chung et al., 2009), warning signals detection (Carbonneau et al., 2013), and voice activity detection (Lezzoum et al., 2014) for the development of a smart HPD (S-HPD) or smart earphones to guarantee protection and to discriminate between speech and noise, allowing the transmission of enhanced speech signals to the ear.

Electronic HPDs process the incoming signals in real-time for retransmission to the ear. Real-time processing is defined as the continuous generation of an output signal within time constraints (Kuo et al., 2014). These time constraints depend on the targeted application for which the processing is dedicated. For example, in Voice over IP (VOIP) communications, the time that elapses between the moment the talker utters the words and the moment the listener hears them is referred as the *mouth to ear delay* (Janssen et al., 2002), and represents the maximum delay between the input and the output signals. As mentioned in the ITU-T recommendation (ITU-T, 2003), mouth to ear delays of less than 150 ms for the transmission of speech or non-speech signals will experience essentially transparent interactivity. However, in other applications where visual information is also available in addition to the audio, such as teleconferencing, the audio signal should never be delayed by more than 45 ms from the video signal, while the video signal should never be delayed by more than 15 ms from the audio signal as demonstrated in (Cooper, 2003) and (Younkin and Corriveau, 2008) to avoid introducing lip-sync errors.

Digital HPDs may introduce a delay between the signals transmitted through the passive path of the HPD and the signals processed and transmitted through the internal loudspeaker. The passively transmitted signals reach the protected ear through the bone conduction and HPD material. When the processing delay increases, the processed signal will be heard as an echo of the passively transmitted signal, thus two signals will be heard. The delay at which the perception of *one fused sound* becomes *two separate sounds* is called the *echo threshold* (Litovsky et al., 1999), or the delay of the *Just Noticeable Difference* (JND) (Quené, 2007), which are widely used psycho acoustic metrics. In (Haas, 1972), the influence of a single echo on the audibility of clean speech has been studied depending on different parameters such as the intensity, the timbre, the angle of incidence and the room reverberation, concluding that when the echo sound is at the same intensity as the original sound, the critical delay (the delay where 10–20% of participants felt disturbed) is about 68 ms, while when the echo sound is attenuated by 3 dB, the critical delay rises to 108 ms, and when the echo sound is attenuated by 10 dB, no echo is felt. Furthermore (Haas, 1972), showed that the attenuation of the high frequencies of the echo increases the tolerable delay.

The echo threshold can also be determined when a sound from one direction is followed by the same sound coming from another direction (Yang and Grantham, 1997). This phenomenon is known as the *precedence effect*. The precedence effect has been widely studied in the last decades and the influence of an echo on the audibility of clicks (transient signals) coming from different spatial locations has also been studied such as in (Freyman et al., 1991), (Yang and Grantham, 1997), and (Saberi and Antonio, 2003). These studies showed that when the click sound echo has equal intensity as the original click sound, the echo threshold is around 5–10 ms.

Studies and experiments reported to date on the determination of the echo thresholds have been conducted with clean speech (Haas, 1972), or with transient signals (Yang and Grantham, 1997), (Litovsky et al., 1999), (Saberi and Antonio, 2003). However, non-ideal real-world conditions such as noisy speech signals have not been investigated yet. For transient signals, the echo threshold was only determined with equal intensities. In addition, the motivation of almost all the previous studies was to understand how the auditory system processes and perceives the same signal coming from different directions such as reverberant spaces. However, the determination of the echo threshold for applications such as electronic HPDs, including the effect of their specific frequency response and resonances, has not been addressed yet, despite the fact that these electronic devices inevitably generate a processing delay.

The current study investigates the influence of frequency-dependent attenuation functions obtained from two representative fits of a custom earplug to evaluate the echo threshold dependence on the attenuation function. Furthermore, this study tends to mimic real-world environments using clean speech signals and speech corrupted by two types of noise environments: factory and babble noise. In addition, a bell ringing sound is used as a transient signal.

This study was conducted on 20 human participants. Each participant was asked to determine the echo threshold between the passively and digitally transmitted signals using a real-time test software where the delay between the two signals could be user-controlled.

The present paper is organized as follows: Section 2 models the sound transmission paths in digital HPDs. Section 3 describes the materials and methods used for the attenuation functions calculation, stimuli generation, and subjective test protocol. Section 4 presents the analysis of the stimuli signals using the spectrograms and the results from the subjective test and Section 5 discusses the findings and concludes this work.

2. Digital hearing protection device

2.1. Sound transmission paths

The digital HPD is a traditional passive HPD in which electro-acoustic hardware is embedded (Fig. 1). To capture signals, a miniature external microphone is connected to the audio input of an ultra-low power DSP. The DSP output is connected to a miniature loudspeaker to transmit the desired signals to the ear.

In addition to the digital path, the external sound is also transmitted through the HPD's material and, to a lesser extent, through bone conduction. Fig. 2 illustrates the three sound transmission paths for a digital HPD.

The transmission through the HPD material highly depends on the fit of the earplug. As an example, Fig. 3 shows the attenuation function of a shallow and deeply fitted HPDs, where differences of up to 20 dB can be observed.

The signal path through the human skull (bone conduction) is highly attenuated (from 45 to 55 dB) making it a negligible

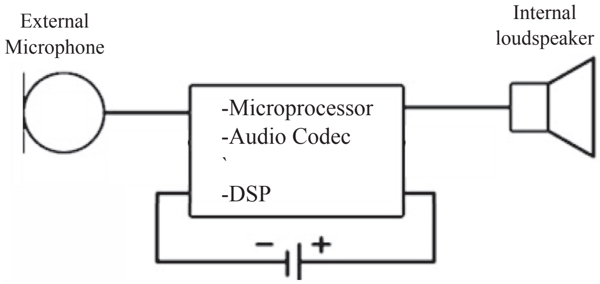


Fig. 1. The hardware resources embedded in the digital hearing protection device.