# 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

## 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.

© 2016 Published by Elsevier Ltd.

## 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 et al., 2003b). 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 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

---

* Corresponding author.
Email address: jeremie.voix@etsmtl.ca (J. Voix).

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
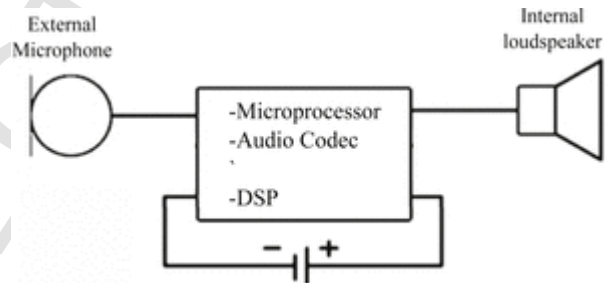


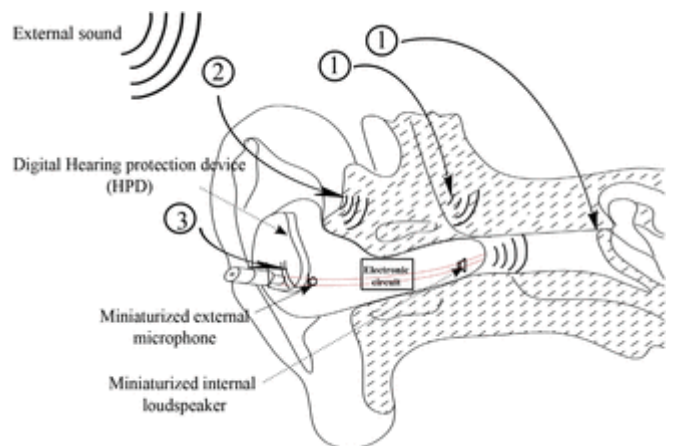**Fig. 1.** The hardware resources embedded in the digital hearing protection device.



**Fig. 2.** Sound transmission pathways through a digital HPD: (1) bone conduction path, (2) passive transmission through the earplug material, and (3) digital transmission through the active path of the earplug. This figure has been adapted from (Voix and Laville, 2009).
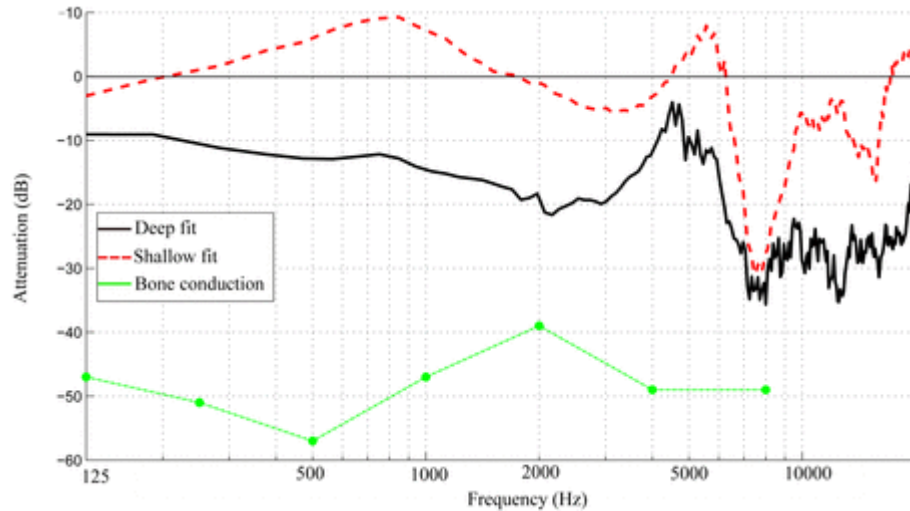
**Fig. 3.** Attenuation functions of two custom-molded earplugs: shallow fit represents the low attenuation function, and deep fit represents the high attenuation function.

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 secondary path (Berger et al., 2003a,b). Therefore, in the rest of this paper, bone conduction is ignored and passively transmitted signals denote only the signals transmitted by means of HPD material.

### 2.2. HPD characteristics

Custom-molded earplugs are a type of HPD that fit instantly to the user's ear canal, by injecting a soft expendable medical silicon rubber agent between a rigid core and an expendable envelope (Voix and Laville, 2009).

The prediction of the attenuation function of these earplugs is conducted as described in (Voix and Laville, 2009): an internal microphone is used to capture the passively transmitted sound to the ear, and an external microphone is used simultaneously to capture the external sound. The attenuation functions are computed from the internal and external sound pressure level.

In a previous study (Nadon et al., 2015), eight human participants were fitted with custom-molded earplugs. The corresponding transfer functions were assessed using white noise. From this dataset, two transfer functions have been selected for the purposes of the current study. These transfer functions represent two extreme cases: the first transfer function has a low attenuation and was obtained from a participant with a shallow fitted earplug, while the second has a high attenuation and was obtained from a participant with a deeply fitted earplug. The magnitudes of these transfer functions are illustrated in Fig. 3. This figure shows the frequency-dependent attenuations that both (shallow and deep) fits exhibit. It also shows that the shallow fit has two resonance frequencies, the first one corresponds to a Helmholtz resonator resulting from the leaking earplug, while the second one corresponds to the longitudinal resonance of the occluded ear canal. Fig. 3 shows that the attenuation function corresponding to the deep fit attenuates the signal by 10 dB below 3500 Hz, and by 5 dB around 5000 Hz, while it attenuates the high frequencies by about 30 dB.

## 3. Methodology

In the first part of this section, we present the stimuli signals used for this study, while in the second part, subjective tests conducted with 20 untrained human participants using the generated stimuli signals and a test software are presented.

### 3.1. Stimuli generation

#### 3.1.1. Types of signals

Two types of signals are used in this study. These signals are considered as desired signals for an S-HPD application use case, thus their unaltered transmission through the digital earplug to the protected ear is important. These signals are:

- Speech signals: one speech sentence uttered by a male speaker in Canadian French from the HINT (Hearing in Noise Test) database (Lamothe et al., 2002) was used. The length of this clean speech signal is around 2 s, and the sampling frequency is 22 kHz. Two different scenarios are considered: the first, consists of presenting clean speech to the participants. In the second scenario, noisy speech signals were presented to the participants by artificially adding, to the same clean speech signal, babble and factory noise obtained from the Aurora database (Hirsch and Pearce, 2000) with a 5 dB signal to noise ratio. This situation mimics noisy environments, such as workplaces or restaurants, in which wearing HPDs or other smart in-ear devices is beneficial.
- Transient signals: transient signals are characterized by their abrupt high energy peaks with a period varying between 5 and 10 ms followed by decaying oscillations with a longer period. Hearing an echo of the transient signal can be annoying to the HPD wearer. For this purpose, a bell ring obtained from a free online database (FreeSound, 2014) is used. The sampling frequency is 44 kHz.

#### 3.1.2. Signal processing

The two attenuation functions, corresponding to two fits of the earplugs, are applied to the four signals presented in Section 3.1.1 for the generation of the passively transmitted signals $y(n)$:

$$y(n) = x(n) * h(n) \tag{1}$$

with * for the convolution, $x(n)$ for the original signal, $h(n)$ the impulse response of the attenuation function, and $n$ the sample number. The impulse response of the attenuation function $h(n)$ is adjusted to 44 $\mu$s, corresponding to the delay of the passively transmitted signal through the earplug (15 mm traveled at 340 m/s).

Thus, eight signals are generated: four signals for the shallow fit, and four signals for the deep fit.

Fig. 4 illustrates a block diagram for stimuli generation. The stimulus $s(n)$ is generated by adding the passively transmitted signal $y(n)$ to the digitally transmitted signal $x(n)$:

$$s(n) = y(n) + x(n - d) \tag{2}$$

with $n > d$ and $d$ is the number of taps that represents the delay difference between the original signal transmitted via the digital path of the earplug and the passively transmitted signal.

Varying the delay $d$ between the two signals will lead to the determination of the echo threshold using a subjective tracking procedure described in the next subsection.

### 3.2. Subjective test protocol

The test was conducted in an ANSI S3.1 compliant audiometric booth with 20 French speaking and normal hearing participants: 17 males and 3 females aged between 22 and 35 years of age with an average age of 25 years. All signed a consent form prior to participation. The subjective tests presented in this paper were approved by the internal review board of ETS (Comité d'Éthique de la Recherche of École de technology superieure) (CER, 2014).

This subjective test is conducted with untrained participants as the electronic HPDs are aimed to be used by a large population.

Participants were outfitted with professional headphones and placed in front of a computer screen equipped with a test interface which allows the user to change the delay between the passively attenuated and the non-attenuated signals in real-time to determine the echo threshold. Fig. 5 illustrates a picture of the experimental setup.

The test interface was created using the open source software PureData (PureData, 2014). Fig. 6 illustrates a screen shot of this interface which has been developed in our labs.

Before the test, participants were instructed to find the echo threshold that corresponds to the delay at which they start to hear an echo of the first signal. To do so, participants were instructed to vary the delay by moving the wheel of the computer mouse: when the wheel is moved up, the delay increased by 2 ms, when moving it down, the delay decreased by 2 ms steps. The delay could vary between 0 and 1000 ms.
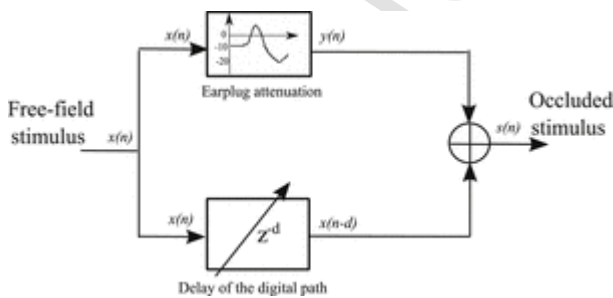


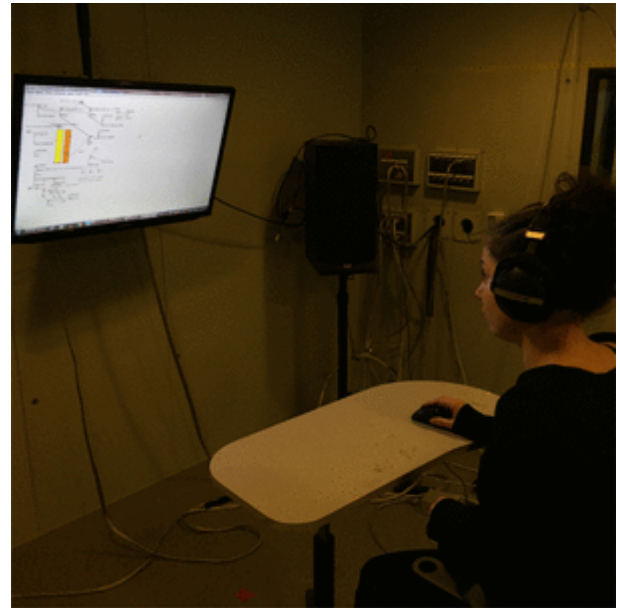**Fig. 4.** Block diagram for the stimuli generation.



**Fig. 5.** Picture of the experimental setup: a participant outfitted with headphones and placed in front of the screen varies the delay using the mouse wheel.

Before validating their response, participants were asked to decrease the delay to be more accurate and detect the threshold of the just noticeable difference. Once this threshold was found, they were asked to validate their response by pressing a button and pass to the following stimuli.

Participants were instructed to fix the threshold to the maximum value (1000 ms) if the passively transmitted signal could not be distinguished from the digitally transmitted signal, i.e. if no echo was perceptible.

The test signals were presented to all the participants in the same order: starting with the stimuli generated from the shallow earplug fit then the stimuli generated from the deep earplug fit following this order: the bell ring signal, the clean speech, the speech corrupted by factory noise, then the speech corrupted by babble noise.

## 4. Data analysis and results

Before analyzing the collected data, a spectral and temporal analysis of the passively and digitally transmitted signals is conducted using spectrograms to understand how the two signals are delayed in time and frequency.

Afterwards, the data collected from this test is subjected to statistical analysis.

### 4.1. Spectrogram analysis

Fig. 7 illustrates the spectrograms of the bell signal, clean speech signal, and speech corrupted by factory noise (each with the two fits) with no delay ($d = 0$ ms) and with a delay of 80 ms. This figure shows that for the bell signal, a difference is observed between the two spectrograms ($d = 0$ and $d = 80$ ms) for the shallow and deep fits. The same observations are also noticed for the clean speech with the two fits. For speech corrupted by factory noise with deep fit, we notice that there is no difference between the two spectrograms. This is due to the low energy of the passively transmitted speech signal (as shown in Fig. 7), which is masked by the factory noise. However, with the shallow fit, we observe a difference in the spectrograms between $t = 0.6$
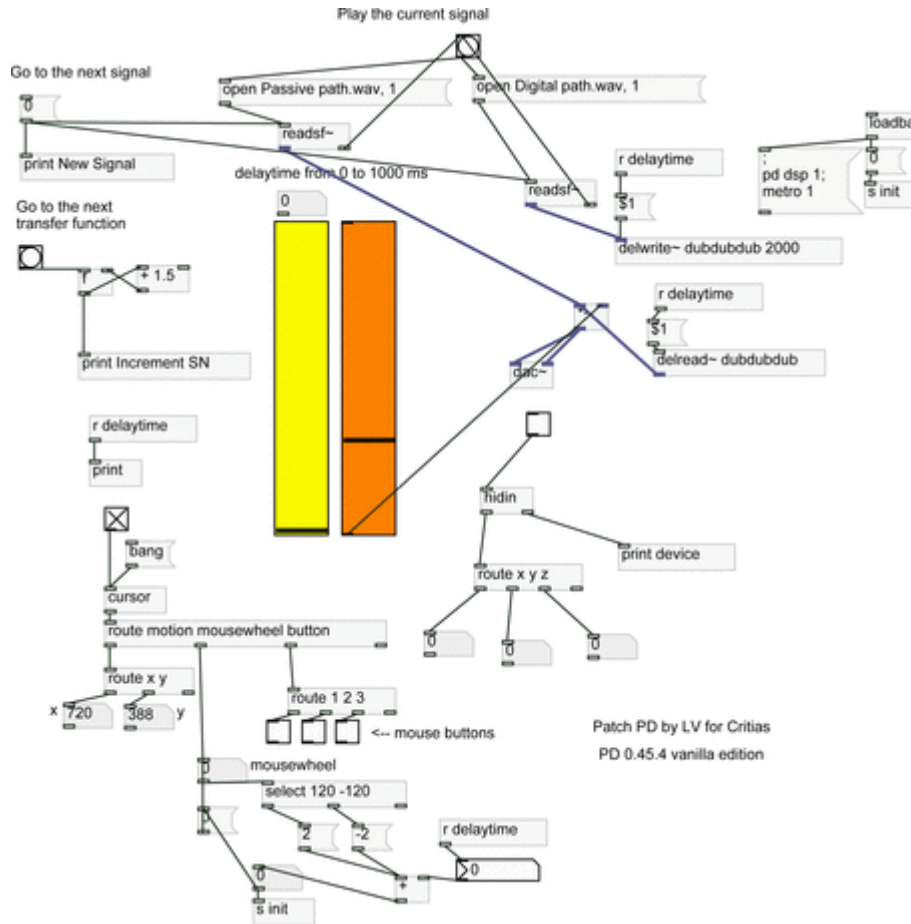
**Fig. 6.** Screen shot of the test interface designed with PureData for real-time signal delaying.

and $t = 0.8$ s where there is an obvious redundancy of the speech segment.

### 4.2. Descriptive statistics

The minimum, the first quartile, the median, the third quartile, and the maximum were calculated upon the echo threshold determined by the participants depending on the stimuli and are illustrated in the box-and-whisker plot in Fig. 8.

Fig. 8 shows that for the bell ring the echo threshold median for the shallow fit (16 ms) is close to the median of the deep fit (26 ms). It also shows that for the clean speech, the median is almost the same for the shallow (36 ms) and deep fits (39 ms). However, for the speech corrupted by factory noise, the echo threshold medians are distant for the two fits (39 ms for the shallow fit and 76 ms for the deep fit), knowing that during the test, four participants notified that they did not perceive an echo even if the maximum value was reached (1000 ms) for the deep fit, which is due to the background noise which masks the passively transmitted speech signal.

With the last stimuli (speech corrupted by babble noise), a big difference is noticed between the shallow and deep fit stimuli. The median echo threshold for the shallow fit was found at 43 ms, while for the deep fit, 15 subjects among the 20 did not notice any difference between the passively and numerically transmitted signals. With the speech corrupted by babble noise (deep fit), one participant fixed the echo threshold at 46 ms. This participant is a musician and is very sensitive to changes in the frequency components of a signal. He commented that his choice was very influenced by frequency components change in the signal.

### 4.3. Analysis of the variance

In order to assess significant differences between the echo thresholds obtained with each attenuation function, signal and participant, we subjected the echo threshold determined by the 20 participants to statistical analysis. For this purpose, a three way analysis of variance (ANOVA) was conducted in Matlab™ (Mathworks, MA). The model used is not with repeated measurements and the alpha value is 0.05 which corresponds to 95% confidence. In this analysis, two factor interactions with three levels have been used and consist of: first, the interaction between the fit type and the signal type; second, the interaction between the fit type and the participants, and third, the interaction between the signal type and the participant. Table 1 illustrates the details of the ANOVA analysis with the p-values.

Results from the ANOVA confirm the previous obtained results and show that there was a significant interaction between the type of the fit and the type of signal ($F(1.5) = 10.55$, $p < 0.05$), while no significant interaction was found between the type of the fit and the participant ($F(1.5) = 0.97$) as well as between the type of the signal and participant ($F(1.5) = 0.79$).
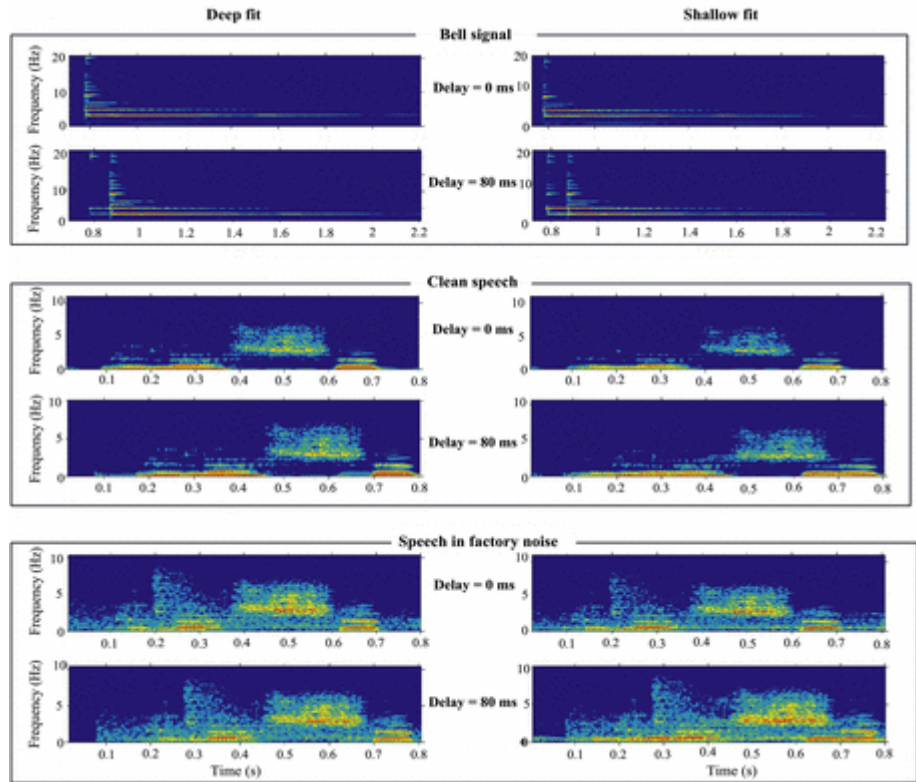
**Fig. 7.** Spectrograms of the bell, clean speech, and speech corrupted by factory noise with a delay of 0 and 80 ms between the passively and the digitally transmitted signals for the shallow and deep fits.
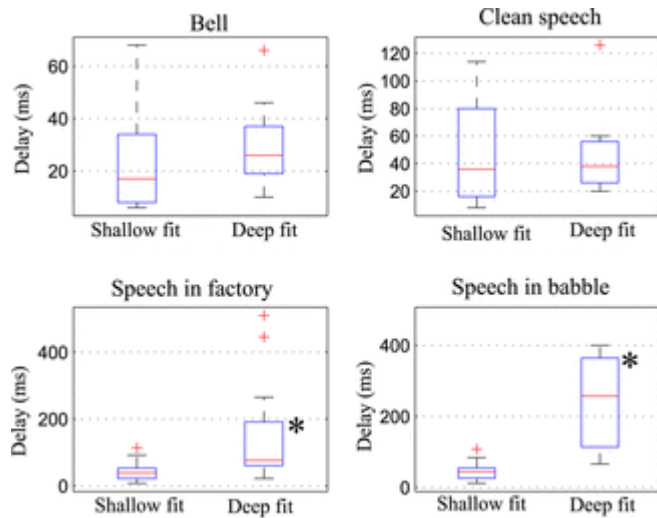


**Fig. 8.** Echo threshold box and whisker plot. The asterix (*) symbol upon the deep fit results of the speech corrupted by factory noise reflects that 20% of the subjects did not perceive any difference between the passively and digitally transmitted signals respectively, and the asterix (*) symbol upon the deep fit results of the speech corrupted by babble noise reflects that 75% of the subjects did not perceive any difference between the passively and digitally transmitted signals respectively.

### 4.4. Determination of the echo threshold

The echo threshold is defined here as the minimum delay at which at least 20% of the participants perceive two distinct signals, as was

**Table 1**
Results of the Analysis of variance ANOVA.

| Source | $F$ | $p$ values |
|---|---|---|
| Fit type | 35.03 | 0 |
| Signal type | 13.93 | 0 |
| Participants | 2.12 | 0.0237 |
| Fit type × Signal type | 10.55 | 0 |
| Fit type × Participants | 0.97 | 0.5087 |
| Signal type × Participants | 0.79 | 0.7945 |

done in (Haas, 1972). The echo threshold for each stimuli was determined by plotting the Cumulative Density Functions (CDFs) which are shown in Fig. 9. The results are summarized in Table 2, which highlights the dependence on the fit of the earplug (from 8 to 18 ms for the bell signal, and 16–26 ms for the clean speech). With a deep fit, most participants could not distinguish the echo from the original sound for the speech signals in babble noise.

## 5. Discussions and conclusions

As described in this paper, with the miniaturization of microelectronic devices, it is now possible to include a DSP in a HPD to perform real-time signal processing on incoming audio signals. However, this signal processing introduces some delay which can be annoying to the user. Determining the echo threshold in real-world conditions allows to set the allowable processing delay of the DSP in such devices.

The allowable processing delay for the electronics represents the time difference between the echo threshold and the delay of the acoustic path through the earplug. The delay of the acoustic path through the earplug is around 44 $\mu$s (15 mm traveled at 340 m/s),
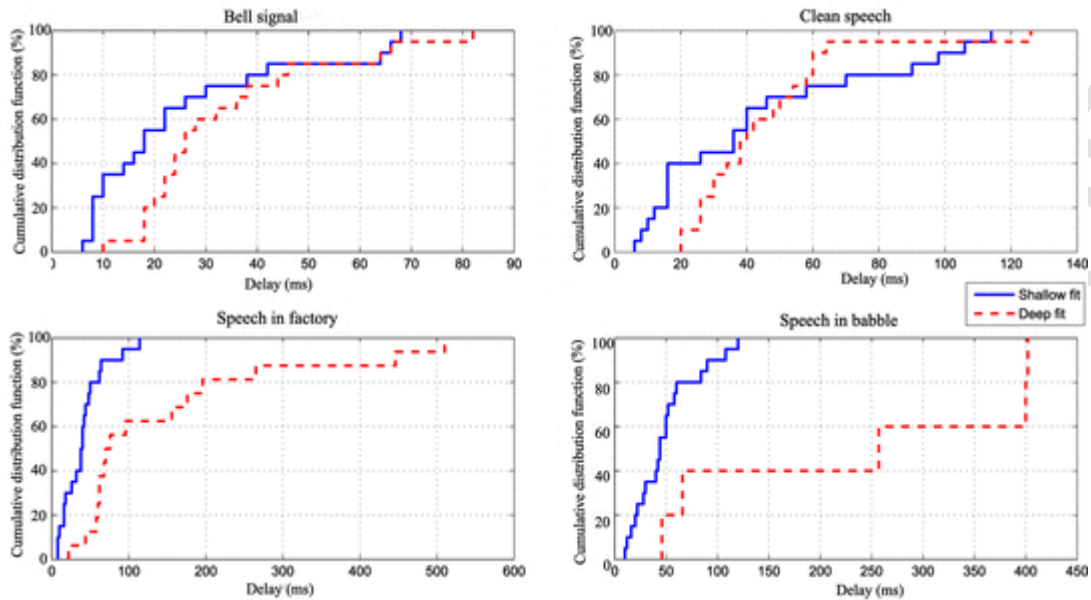
**Fig. 9.** Cumulative Density Functions for the eight stimuli.

**Table 2**

The echo thresholds for the eight stimuli. For the deep fit with factory noise, 20% of the subjects did not perceive any difference between the passively and digitally transmitted signals. In babble noise 75% of the subjects did not perceive any difference between the passively and digitally transmitted signals.

| Signal | Echo thresholds for the two fits | |
|---|---|---|
| | Shallow (ms) | Deep (ms) |
| Bell | 8 | 18 |
| Clean speech | 16 | 26 |
| Speech in factory | 16 | 68 |
| Speech in babble | 22 | * |

which is nonsignificant when compared to the echo threshold reported in the study. Therefore, we conclude that the delay introduced by the entire electronic path (from the microphone to the loudspeaker) should be made lower than the echo threshold.

The subjective results presented in this paper showed that when the earplug has a shallow fit and presents a resonance frequency in the critical frequency range of speech (between 200 and 1000 Hz), the echo threshold of clean speech stimuli is almost the same as the echo threshold of speech corrupted by a stationary (factory) or non-stationary speech-shaped noise (babble) (for 20% of the participants the echo threshold for the three stimuli is 16, 16, and 22 ms respectively). However, when the earplug has a deep fit without resonance frequency in the critical frequency range of speech, the echo threshold of clean speech stimuli is lower than the echo threshold of speech corrupted by factory noise, while when the speech is corrupted by speech-shaped noise (babble), there is no perceptible difference between the passively and digitally transmitted signals.

From the current study, we conclude that the echo threshold between the passively and digitally transmitted signals depends on four parameters:

- The attenuation function: the amount of attenuation of the in-ear device is a very important parameter for the determination of the echo threshold between the passively and the digitally transmitted signals. The higher the attenuation is, the higher the delay.

- The duration of the signal: the delay depends on the duration of the signal, if the signal has a short duration such as transients, the delay is low and it increases when the incoming signal duration increases.

- The presence of background noise: the current study showed that when background noise is present, the echo threshold increases compared to clean speech conditions. For instance, with a deep fit, the clean speech stimuli gave a median echo threshold of 38 ms, while when speech is corrupted by factory noise, the median echo threshold was found at 96 ms.

- The type of background noise: when the incoming signal is corrupted by background noise, the delay increases since the background noise masks the passively transmitted signal. The delay not only depends on the presence of background noise, but also on the type of noise: if the background noise is non-stationary such as babble noise, the delay is higher than when the background noise is stationary such as factory noise.

The delay between the passively and digitally transmitted signals depends not only on one criterion but on the combination of the four criteria.

Our findings suggest that manufacturers of electronic HPDs and the next generation of digital in-ear devices should set the processing time depending first on the attenuation function of the device. In addition, the processing time should be chosen as a function of the type of the desired signal to be sent through the digital path of the HPD: if the electronic HPD is designed to transmit signals with short periods such as transient signals, then the processing time should be lower (between 8 and 18 ms) than if the device was designed to transmit other signals such as speech signals (between 16 and 26 ms). Furthermore, if the device is developed to be used in noisy environments, the processing time can be higher and depend on the nature of the background noise. In situations where the processing time can be sufficiently long, other sophisticated modules such as speech recognition, speaker recognition, signal identification or background noise classification can be embedded in the in-ear devices. Nevertheless, in situations where the visual information is also provided to the in-ear device wearer, the processing time should be determined as a function of this information and should

not exceed a certain amount of time (45 ms) to avoid generating a lip-sync error.

## Acknowledgments

## References

Berger, E.H., Kieper, R.W., Gauger, D., 2003a. Hearing protection: surpassing the limits to attenuation imposed by the bone-conduction pathways. J. Acoust. Soc. Am. 114 (4), 1955–1967.
Berger, E.H., Royster, L.H., Royster, J.D., Driscoll, D.P., Layne, M., 2003b. The Noise Manual, fifth ed., .
Brimhall, OD., Collotzi, CM., Koshowich, GN., 2002. Attenuation, protective hearing devices with multi-band automatic amplitude control and active noise.
Carbonneau, M.-A., Lezzoum, N., Voix, J., Gagnon, G., Sep. 2013. Detection of alarms and warning signals on a digital in-ear device. Int. J. Ind. Ergon. 43 (6), 503–511.
Casali, J.G., 2010. Powered electronic augmentations in hearing protection technology circa 2010 including active noise reduction, electronically-modulated sound transmission, and tactical communications devices : review of design, testing, and Research. Int. J. Acoust. Vib. 15 (4), 168–186.
CER, 2014. Comité d'ethique et de la recherche. Ecole de technologie superieure. N H20121004. Tech. rep.
Chung, K., 2007. Effective compression and noise reduction configurations for hearing protectors. J. Acoust. Soc. Am. 121 (2), 1090–1101.
Chung, K., Tufts, J., Nelson, L., May 2009. Modulation-based digital noise reduction for application to hearing protectors to reduce noise and maintain intelligibility. Hum. Factors J. Hum. Factors Ergon. Soc. 51 (1), 78–89.
Cooper, J.C., 2003. A Short Tutorial on Lip Sync Errors, the Sources and Solutions. Tech. Rep. 408.
FreeSound, 2014. https://www.freesound.org/.
Freyman, R.L., Clifton, R.K., Litovsky, R.Y., Aug. 1991. Dynamic processes in the precedence effect. J. Acoust. Soc. Am. 90 (2 Pt 1), 874–884.
Haas, H., 1972. The influence of a single echo on the audibility of speech. J. Audio Eng. Soc. 20 (2), 146–159.
Hirsch, H.-g., Pearce, D., 2000. The Aurora experimental framework for the performance evaluation of speech recognition systems under noisy conditions. In: ISCA ITRW ASR2000 "Automatic Speech Recognition: Challenges for the Next Millennium". .
Hotvet, DA., 1996. Directional ear device with adaptive bandwidth and gain control.
ITU-T, 2003. Transmission Systems and Media, Digital Systems and Networks: One-way Transmission Time. Tech. rep.
Janssen, J., Vleeschauwer, D.D., Windey, R., Petit, G.H., Leroy, J.-M., 2002. Delay bounds for voice over IP calls transported over satellite access networks. J. Mob. Netw. Appl. (MONET), Special Issue Satell. Based Inf. Serv. 7 (1), 79–89.
Kuo, S.M., Lee, B.H., Tian, W., 2014. Real-time Digital Signal Processing: Fundamentals, Implementations and Applications, third ed. Wiley.
Lamothe, J., Gascon, C., Lariviere, M., Laroche, C., 2002. Hearing in normalisation du hearing in noise test (HINT) auprès d'une population francophone bilingue et d'une population anglophone. Revue d'orthophonie et d'audiologie 26 (2), 81–89. 2.
Lezzoum, N., Gagnon, G., Voix, J., 2014. Voice activity detection system for smart earphones. IEEE Trans. Consumer Electron. 60 (4), 737–744.
Litovsky, R.Y., Colburn, H.S., Yost, W.A., Guzman, S.J., Oct. 1999. The precedence effect. J. Acoust. Soc. Am. 106 (4 Pt 1), 1633–1654.
Macdonald, E.N., Behar, A., Wong, W., Kunov, H., 2008. Noise exposure of opera musicians. Can. Acoust. 36 (4), 11–16.
Nadon, V., Bockstael, A., Botteldooren, D., Lina, J.-M., Voix, J., Mar. 2015. Individual monitoring of hearing status: development and validation of advanced techniques to measure otoacoustic emissions in suboptimal test conditions. Appl. Acoust. 89, 78–87.
NIOSH, 1998. Occupational Noise Exposure. U.S department of Health and Human Services.
PureData, 2014. http://puredata.info/.
Quené, H., Jul. 2007. On the just noticeable difference for tempo in speech. J. Phon. 35 (3), 353–362.
Saberi, K., Antonio, J.V., 2003. Precedence-effect thresholds for a population of untrained listeners as a function of stimulus intensity and interclick interval. J. Acoust. Soc. Am. 114 (1), 420.
Voix, J., 2014. Did you say "bionic ear"?. Can. Acoust. 42 (3), 68–69.
Voix, J., Laville, F., Jul. 2009. Prediction of the attenuation of a filtered custom earplug. Appl. Acoust. 70 (7), 935–944.
Yang, X., Grantham, D.W., Oct. 1997. Echo suppression and discrimination suppression aspects of the precedence effect. Percept. Psychophys. 59 (7), 1108–1117.
Younkin, A.C., Corriveau, P.J., Sep. 2008. Determining the amount of audio-video synchronization errors perceptible to the average end-user. IEEE Trans. Broadcast. 54 (3), 623–627.