Edingburgh, Scotland
Euronoise 2009
October 26-28

Speech recognition in noise with active and passive hearing protectors: a comparative study

Annelies Bockstael[a]
Bart Vinck[b]
Ghent University
De Pintelaan 185 2P1
9000 Gent, Belgium

Dick Botteldooren[c]
Ghent University
Sint-Pietersnieuwstraat 41
9000 Gent, Belgium

ABSTRACT

A common argument for not wearing passive hearing protection is that they create unsafe situations by hampering communication. Making sound attenuation more comfortable is one of the incentives for the development of so-called 'augmented' hearing protectors. However, several studies fail to clearly demonstrate their benefits on signal perception in noise over standard passive protectors. Apparently, the positive effects of these protectors are not clear-cut and testing of new prototypes is essential. Within the class of augmented protectors 'passive', i.e. without electronics, and 'active', i.e. with electronics, can be distinguished; in this project, speech perception with custom-made active earplugs is compared to passive custom-made earplugs. The active protectors amplify the incoming sound to the maximum safe level, or to a user defined fraction of this level. For the experiment minimal and maximal settings are compared. In the concrete 20 different speech-in-noise samples are presented to 60 normal-hearing subjects and the speech recognition is scored. The background noise is selected from realistic industrial noise samples with different loudness, frequency and temporal characteristics recorded with a Head And Torso Simulator (HATS). The speech is chosen from standardized word lists used for speech audiometry. The speech-in-noise samples are created by presenting the mixture of speech and noise to the HATS with unoccluded simulators and with passive and active protectors. The results suggest that the performance of the hearing protectors strongly depends on the kind of background noise. The active protectors with minimal amplification outclass the other protectors for the most difficult and the easiest noise conditions, but they also limit binaural listening. For the remaining noise conditions, the passive protectors clearly surpass their active counterparts.

1. INTRODUCTION

A common concern when wearing hearing protectors is that they will hamper verbal communication and the perception of warning signals[1]. When making attempts to handle this issue, one must always bear in mind that signal detection and understanding with hearing protectors depends on a complex of factors[2]. Therefore it is not surprising that different studies yield to sometimes contradictory conclusion with respect to the influence of hearing protectors on speech intelligibility in

[a]Email address: annelies.bockstael@ugent.be
[b]Email address: bart.vinck@ugent.be
[c]Email address: dick.botteldooren@ugent.be

noise[3;4]. Moreover, the wide variety in types of hearing protectors increases the variability among research results. In this regard one can roughly distinguish on the one hand the 'standard' or 'classical' protectors that solely block the sound path to the eardrum whereas on the other 'augmented' protectors actually process the incoming sound in a passive, i.e. without any electronics, or active way. The latter hearing protectors have been developed[5] to make sound attenuation more comfortable[6] and to diminish the masking effect of noise on signals.

This project tests prototypes of active hearing protector designed to be worn in similar conditions as their standard passive counterparts. Different types of background noise are selected and in each sound environment speech fragments are recorded without any hearing protectors, with passive earplugs and active ones. The speech intelligibility for each sound fragment is then determined from the intelligibility scores of normal-hearing subjects.

2. MATERIAL AND METHODS

A. Hearing protectors

In this research project, two types of hearing protectors are included; passive custom-made acrylic hearing protectors (PC) and active custom-made acrylic hearing protectors with volume control. For the latter, two settings are selected; full (AC 3) and minimal amplification (AC 1). These settings are chosen because they provide clearly different output levels for the input levels used in this study.

The passive earplugs with ST35 filter simply block the ear canal and thus reduce the sound with a fixed amount, regardless of the input level. The active protectors on the other hand contain a microphone and a loudspeaker so that the attenuation can be adapted to the incoming sound level. In this regard the brickwall limiter imposes a hard 'ceiling' on the loudspeaker output - once the signal reaches the threshold, it can go no further.

B. Sound environment

This study aims to assess speech recognition in noise for different listening conditions. Therefore, realistic noise fragments are recorded using a Brüel & Kjær Head And Torso Simulator (HATS) type 4128 C with left and right ear simulator. Additionally, the same HATS is used to record the speech material in an anechoic room. The characteristics of both speech and noise are discussed in this section.

Noise material The influence of noise on speech largely depends on three major components; signal-to-noise ratio, frequency spectrum[7] and temporal pattern[8]. In accordance with this knowledge, three types of noise are chosen for this study (see Figure 1). First recordings are made inside the alternators and turbines hangar of a power station where the noise is more or less equally diffused in space. In contrast to this, a more focused sound source is included in the form of a bottle filling machine. Finally, moving fork-lift trucks are selected because they produce more fluctuating noise. All recordings are made with the HATS facing the sound source. In addition, the possible beneficial effect on speech recognition of spatial separation between signal and noise is included by recording the noise of the bottle filling machine recorded with the left ear pointing at the sound source.

Speech material For the speech material standardized recordings of the 'Brugse Lijst' are chosen. This material - read by a professional female speaker - is especially designed to perform speech audiometry and consists of consonant-vocal-consonant words spread among 20 lists with 17 words per list[9].

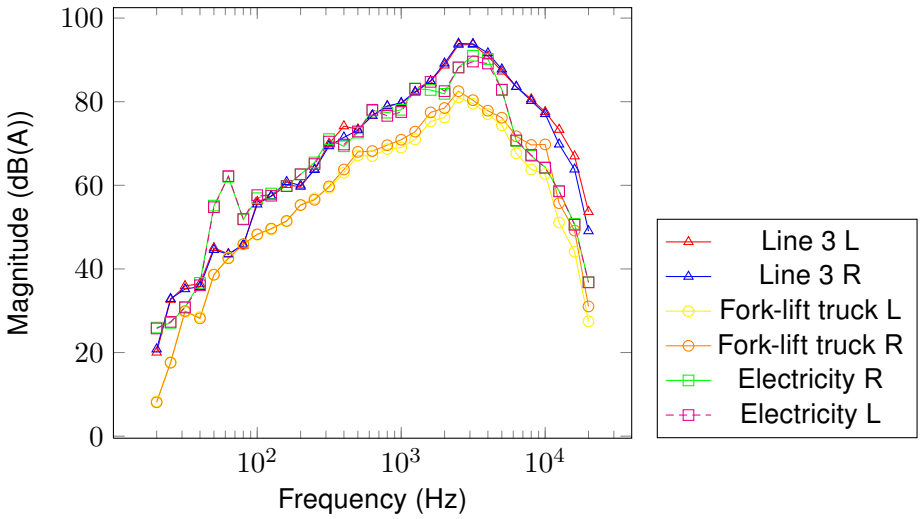

Figure 1: A-weighted $\frac{1}{3}$-octave band spectra of the recordings with the HATS facing the sound source; 'Electricity' refers to the noise of alternators and turbines, 'Line 3' to the bottle filling machine and 'Fork-lift trucks' to moving fort-lift trucks. 'L' are the recordings made by the left ear simulator while 'R' refers to the right ear simulator.

The standardized recordings of the 'Brugse Lijst' are not suitable to be directly mixed with the noise fragments because the former do not include the head-related transfer functions whereas the latter do. Thus, the 'Brugse Lijst' is recorded in an anechoic room with the HATS placed at 1 m from the sound source. For the first 16 lists, the HATS faces the loudspeaker so that the right and the left ear simulator receive the same signal. For the last 4 lists, the HATS is turned with its right ear toward the sound source.

The level of the speech is set at approximately 72 dB(A) measured at the HATS's eardrum, comparable to 68 dB(A) measured in free field[10]. This level is chosen to approximate a normal communication situation where a female person would speak at free-field levels between 63 dB(A) (raised) and 71 dB(A) (loud)[11]. The calibration of the speech signals is done with continuous speech noise, especially developed for this particular set of speech material[9].

C. Listening conditions

Recodings Test material is built by first recording speech and noise separately with unoccluded ear simulators, subsequently mixing speech and noise to create several speech-in-noise fragments and finally recording the fragments with the HATS under headphone in different listening conditions, i.e. with an unoccluded ear canal and with the different protectors. This way, combining five different sound environments with four different listening conditions yields to 20 test fragments.

Before the final recordings of the material, unwanted influence of the HATS is filtered out. This is necessary because the composing of the test material includes two recordings with the HATS, once with open ears at the workfloor (for noise) and in the anechoic room (for speech) and once under headphone for the different listening conditions. This implies that the HATS's ear canal and pinna will influence the test material twice which is of course undesirable. Moreover, the headphone used for the final recordings might also unwantedly mark the test fragments. Hence, the effect of the ear simulators and headphone is minimized by filtering noise and speech so that recordings under headphone made by the HATS with open ears resemble the original recordings as close as possible, especially between 500 Hz and 4000 Hz, i.e. the frequency range most

important for speech recognition. These filtered fragments, and not the original recordings, are used to create the speech-in-noise listening items.

Audio equipment All recordings and the presentation of the listening material to the test subjects is carried out with the same audio equipment. The fragments are played on a laptop PC using the free available Audacity software and then the signal is sent to an open circum-aural Philips headphone (type SBCHP890) via a Pioneer A-607 R direct energy MOS amplifier.

Post-processing It is well-known that measurements with a HATS tend to overestimate the attenuation of passive hearing protectors. To compensate for this effect, the approach proposed by Hiselius[12] is followed by taking into account the bone conduction thresholds. Since an absolute prediction of the attenuation is not the scope of this work, this procedure seems allowed to make the test set-up more realistic.

D. Speech intelligibility test

Test subjects Only native Dutch-speakers who have at least successfully finished high school are included. Further, normal hearing is required with tonal hearing thresholds of 25 dB or better for all octave frequencies between 250 Hz and 8000 Hz[13], normal tympanometric results and normal speech perception. These criteria yield to a final group of 60 test subjects (30 female and 30 male) who are on average 27.6 years old without any gender related significant differences in hearing level or in speech recognition. All participated voluntarily and signed an informed consent.

Test set-up At the beginning of each test day, the audio equipment is calibrated so that the noise level of the speech reference noise under headphone is 72 dB(A) at both HATS's ear simulators. For each subject, the 20 speech-in-noise fragments are presented through the headphones in random order and the instruction are in conformity with classical speech audiometry[14]. In the concrete, the subject has to repeat the words and the correct phonemes are per word marked by the investigator. Counting the identified phonemes yields to a phoneme score per fragment.

Statistical analysis The phoneme scores for the 20 fragments are mutually compared with the free statistical software R.

3. RESULTS

The mean speech recognition scores and standard deviations are depicted in Figure 2 for the 20 test fragments. On this data a mixed model regression analysis is carried out to track influential independent variables[15]. For this analysis, the variables 'sound environment' and 'listening condition' are included as fixed factors whereas the variable 'subject' is included as a random factor. The most important conclusion from this analysis is the statistically significant interaction effect ($p < 0.05$) between the two fixed factors and therefore a pairwise Tuckey-post hoc test is carried out with the interaction effect as independent variable[16]. The most striking results are discussed below.

A. Listening conditions in sound environments

From Figure 2 it becomes clear that the performance of the different hearing protectors strongly depends on the type of background noise. First, passive protectors seem to give the best results for the noise at the power plant and the noise from the bottle filling machine recorded from the left with speech coming from the right. These results are in agreement with previous studies suggesting that passive protectors might enhance communication for normal hearing subjects

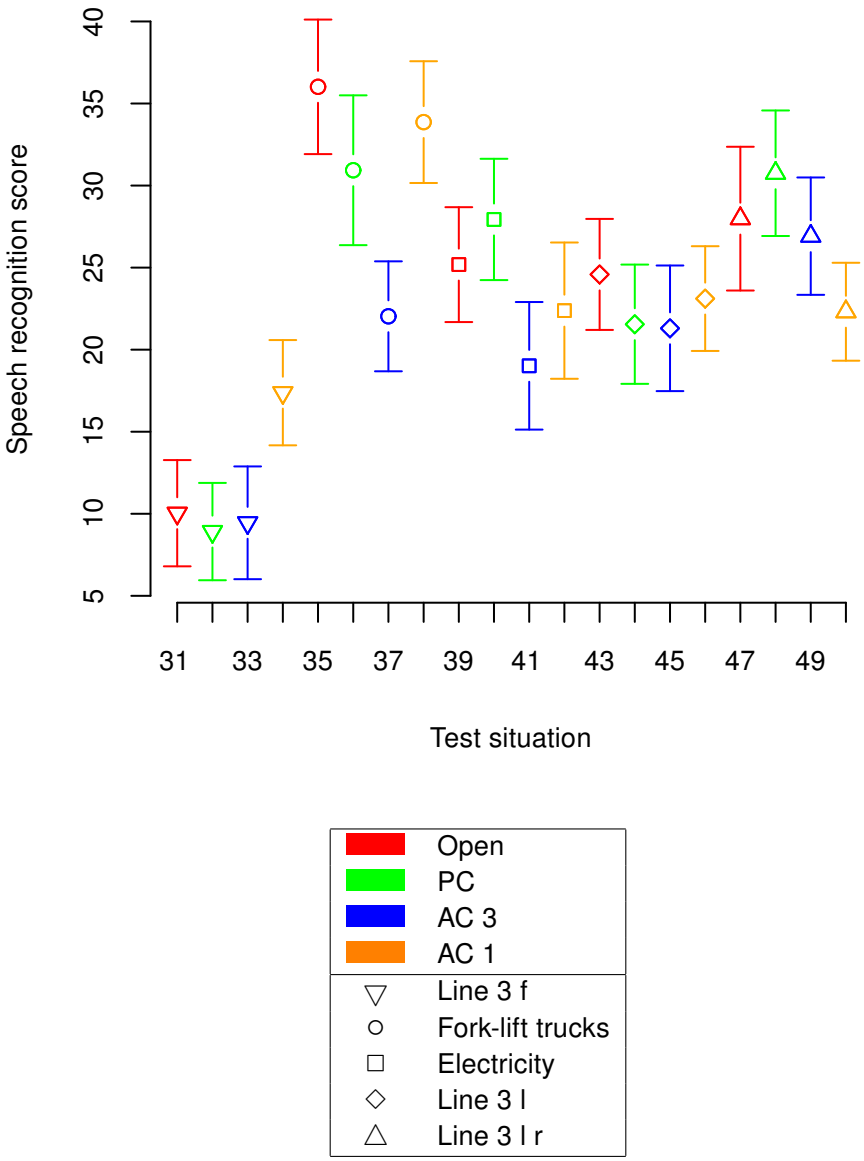

Figure 2: Error bar plot for the speech recognition scores of the 20 test fragments represented by the track number of the different speech lists. The center of the error bars is given by the mean speech recognition score whereas the width of the bars equals one standard deviation. The colors represent the different listening conditions; with 'open' ear canal, with passive protectors (PC) and with active protectors with maximal (AC 3) and mininal (AC 1) gain. The different symbols represent the sound environments, namely the noise from the power plant (Electricity), from the moving fork-lift trucks (Fork-lift trucks), the frontally recored bottle filling machine with frontally recorded speech (Line 3 f) and the same machine recorded from the left combined with frontally (Line 3 l) and sideways recorded speech (Line 3 l r).

when the noise levels exceed 85 dB(A)[3]. For the noise of the fork-lift trucks, the passive earplugs appear to hamper speech recognition more than the unoccluded situation. In this environment, the noise level is quite soft compared to the other conditions and therefore it is not unthinkable that the attenuation of the passive earplug is overprotective, involving more difficulties in communication.

The performance of the active protectors clearly depends on the settings of the volume control. Whereas maximal gain seems to lead to the least performing listening condition for most sound environments, minimal gain enhances recognition for the frontal recordings of the bottle filling machine and is the best occluded listening condition for the fork-lift trucks.

To illustrate the effect of the listening conditions on the sound, the spectra of the four fragments with the fork-lift truck noise are depicted in Figure 3. One of the most striking features is that the active protector with maximal amplification seems to emphasize strongly the frequencies between 1000 Hz and 4000 Hz. This region is indeed important for speech perception, but an excessive amplification might distort the balance between the different frequencies and hence hamper speech intelligibility instead of improving it. Further, Figure 3 reveals the substantial interaural difference for low-frequency attenuation of the active protector with minimal gain.

B. Sound environments in listening conditions

The Tuckey-post hoc test confirms the results obvious from Figure 2; the bottle filling machine frontally recorded is clearly the most difficult listening situation whereas the fork-lift trucks appear to be the least disturbing for the unoccluded situation and the active protector with minimal gain. For the other sound environments, the ranking depends on the listening condition.

The better results for the noise of the fork-lift trucks is little surprising since Figure 1 clearly shows that this noise fragment has the lowest overall sound pressure level. Because of the fixed speech level in all sound environments, a lower noise level leads automatically to a better signal-to-noise ratio and hence more favorable listening conditions.

As for the bottle filling machine Figure 1 also suggests that this environment yields to the lowest signal-to-noise ratio for the frontally recorded noise fragments. The difference with the noise from the power plant is quite small, but the bottle filling machine produces higher levels above 1000 Hz (see Figure 1). The fact that this relation is reversed around 50 Hz is less important since high-frequency noise appears more disturbing[7].

Further, spatial segregation between speech and noise seems to enhance speech recognition for all listening conditions if the sound environments with frontally recorded noise are compared to the sideways recordings. Changing the direction of the speech from frontal to sideways also induces a positive effect on speech perception for all listening conditions except the active protector with minimal gain.

4. DISCUSSION

The ongoing evolution in signal processing gradually relieves the technological constraints in the development of active hearing protectors. In theory these protectors might alleviate the perceived negative influence of hearing protectors on speech intelligibility. Since communication difficulties are often reported as an argument for not wearing hearing protection[17], active protectors can enhance the safety at the workfloor in two ways; by preventing hearing loss and by lowering the risk at accidents due to malcommunication[5;18]. The key question is of course whether these protectors can actually come up to the expectations.

It should be noted that the study's method is not entirely representative for communication in real working environments. There are for instance no competing tasks to perform[8], the level of the speech is independent of the noise level[19] and the listening material is clearly read by a professional speaker but is at the same time much less redundant than normal sentences[14].

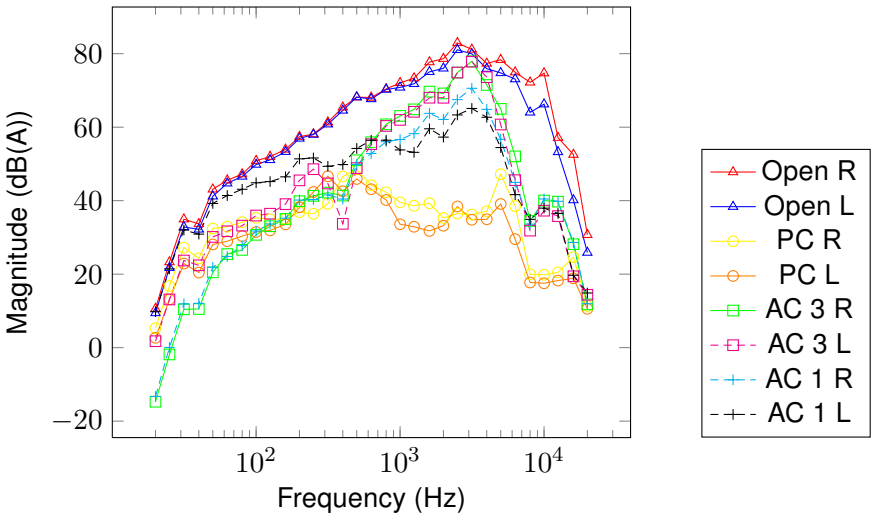

Figure 3: A-weighted $\frac{1}{3}$-octave band spectra of the fork-lift trucks. All abbreviations are similar to the other figures.

The somehow unrealistic listening conditions are not a drawback since the major aim is a direct comparison of speech perception under different protectors. This requires a very controlled test design and therefore measurements are carried out with the HATS and under headphone.

The test fragments are kept as realistic as possible by a careful selection of speech and noise. This implies that the speech material provided by the developers of the 'Brugse lijst' is not used directly, but recordings are made in an anechoic room with the HATS placed at an appropriate distance of the sound source[11]. Further, a realistic speech level is chosen. This level is fixed for all test fragments to enhance mutual comparability.

To create the sound environments, speech fragments are combined with a variety of industrial noise samples. Different noise features are selected to include these aspects that mostly determine the masking effect of noise on speech perception. The fact that the different sound environments indeed yield to statistically significant differences in speech recognition confirms that the protectors are tested in certain variety of situations.

Statistical analysis also reveals that variation in speech intelligibility across listening conditions depends on the sound environment under study. In most environments the active protector with maximal gain seems to decrease speech intelligibility, possibly because the original signal is distorted. This is of course a highly unwanted side-effect of active hearing protectors[18].

Leaving aside the distorting effect, the active protectors with minimal amplification clearly do a good job in the most silent condition. Although they do not completely alleviate the hinder of wearing protectors, they seem to be a better choice than passive protectors when sound levels barely exceed safety limits. In these circumstances, passive protectors can be overprotective and therefore might enhance inconsistent use which in its part clearly increases the risk of noise-induced hearing loss[20].

For the noise coming from the alternators and turbines, the passive protectors give the best results for speech intelligibility for normal-hearing subjects. By contrast, it is not unthinkable that they would degrade communication if a hearing loss is present[21].

In the sound environment with the lowest overall speech recognition, the active protectors with minimal gain provide better communication, but they fail to let the user fully benefit from binaural listening. This might be caused by apparent unequal attenuation in the lower frequencies, rather than being a shortcoming of the signal processing itself. However, this idea is somewhat

contradicted by the fact that in the most difficult sound environment the increased difference between the left and the right ear seems to alleviate communication more than hampering it. This is reassuring because at the workfloor it is probably even more difficult to achieve a perfectly balanced fitting.

In contrast with the active protectors with minimal amplification, the maximal gain setting takes full advantage of spatial separation between speech and noise, as does the passive protector. In this regard it must be noted that preservation of the binaural unmasking effect does not necessarily guarantee good sound localization under hearing protection, despite the fact that they depend upon the same acoustical cues[22].

Thus far, the acoustical analysis is limited to the spectrum and global sound pressure level of the test fragments. A more thorough assessment including binaural aspects and loudness might yield to a deeper insight in the obtained speech recognition scores. Once it is verified that the signal processing is adequate for normal hearing subjects, tests with hearing impaired subjects might be a worthwhile extension.

5. CONCLUSION

Technological progress in signal processing opens in theory the door for active hearing protectors that not only protect against noise-induced hearing loss but also noticeably alleviate communication in noise. Nevertheless the question arises to what extend the implementation of new electronics will effectively result in better hearing protectors.

This paper clearly shows that the performance of different protectors depends on the characteristics of the noise environment; passive protectors yield in general to better results in louder environments whereas active protectors perform better in lower and fluctuating noise. However, one should notice that spatial unmasking might be compromised by these active protectors and that distortion by excessive amplification should be excluded.

The current study confirms that speech intelligibility in noise is a complex process depending on a host of factors. This strongly complicates the prediction of hearing protectors on communication and therefore controlled listening tests with human subjects seem indispensable.

REFERENCES

1. O. Hong, D. Samo et al., Perception and attitudes of firefighters on noise exposure and hearing loss, *J. Occup. Environ. Hyg.*, **5**(3), pp. 210–215 (2008).

2. C. Giguère, C. Laroche et al., Modelling the effect of personal hearing protection and communications devices on speech perception in noise. (2008).

3. S. Abel, N. Armstrong and C. Giguère, Auditory perception with level-dependent hearing protectors: the effect of age and hearing loss, *Scandinavian audiology*, **22**, pp. 71–85 (1993).

4. T. Dolan and D. O'Loughlin, Amplified earmuffs impact on speech intelligibility in industrial noise for listeners with hearing loss, *American Journal of Audiology*, **14**(1), pp. 80–85 (2005).

5. J. Casali, G. Robinson et al., Effect of electronic ANR and conventional hearing protectors on vehicle backup alarm detection in noise., *Hum. Factors*, **46**(1), pp. 1–11 (2004).

6. S. Dantscher, Hearing protectors with electronic components - selection and use, in Noise at Work 2007 - Proceedings, Lille (2007), pp. 1345–1349.

7. G. Studebaker, R. Taylor and R. Sherbecoe, The effect of noise spectrum on speech recognition performance-intensity functions, *Journal of Speech, Language and Hearing Research*, **37**(2), pp. 439–448 (1994).

8. C. Laroche, S. Soli et al., An approach to the development of hearing standards for hearing-critical jobs, *Noise and Health*, **6**(21), pp. 17–37 (2003).

9. W. Damman, De Brugse lijst voor spraaakaudiometrie - versie 1989, *Tijdschrift voor logopedie en audiologie*, **24**(3), pp. 21–37 (1994).

10. D. Hammershøi and H. Møller, Determination of noise immission from sound sources close to the ears, *Acta Acust. United Ac.*, **94**(1), pp. 114–129 (2008).

11. W. Olsen, Average speech levels and spectra in various speaking/listening conditions: A summary of the Pearson, Bennett, & Fidell (1977) report, *American Journal of Audiology*, **7**(2), pp. 21–25 (1998).

12. P. Hiselius, Attenuation of earplugs - objective predictions compared to subjective reat measurements, *Acta Acust. United Ac.*, **91**, pp. 764–770 (2005).

13. R. J. Roeser, K. A. Buckley and G. S. Stickney, Audiology diagnosis, Thieme, chapter Pure tone tests, pp. 227–251 (2000).

14. W. Damman, Spraakaudiometrie in de praktijk, *Tijdschrift voor logopedie en audiologie*, **23**(1), pp. 15–38 (1993).

15. D. Bates, Package lme4, Technical Report 0.999375-28, The R Project for Statistical Computing, www.r-project.org (2009).

16. T. Hothorn, F. Bretz and P. Westfall, The multcomp package, Technical Report 1.0-6, The R Project for Statistical Computing, www.r-project.org (2009).

17. A. M. Nakashima, M. J. Borland and S. M. Abel, Measurement of noise and vibration in Canadian forces armoured vehicles, *Ind. Health*, **45**(2), pp. 318–327 (2007).

18. J. Casali and H. Berger, Technology advancements in hearing protection circa 1995: active noise reduction, frequency/amplitude-sensitivity and uniform attenuation, *Am. Ind. Hyg. Assoc. J.*, **57**, pp. 175–185 (1996).

19. H. Hörmann, G. Lazarus-Mainka et al., The effect of noise and the wearing of ear protectors on verbal communication, *Human Factors: The Journal of the Human Factors and Ergonomics Society*, **23**(2), pp. 69–77 (1984).

20. J. Voix and F. Laville, Problematiques associees au developpemenr d'un bouchon d'oreile intelligent., *Pistes*, **7**(2) (2005).

21. S. Abel, P. Alberti et al., Speech intelligibility in noise: Effects of fluency and hearing protector type, *J. Acoust. Soc. Am.*, **71**(3), pp. 708–715 (1982).

22. B. E. Edmonds and J. F. Culling, The role of head-related time and level cues in the unmasking of speech in noise and competing speech, *Acta Acust. United Ac.*, **91**(3) (2005).