

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

Author

Bockstael, Annelies, De Coensel, Bert, Botteldooren, Dick, D'Haenens, Wendy, Keppler, Hannah, Maes, Leen, Philips, Birgit, Swinnen, Freya, Bart, Vinck

Published

2011

Journal Title

Journal of the Acoustical Society of America

Version

Version of Record (VoR)

DOI

[10.1121/1.3575599](https://doi.org/10.1121/1.3575599)

Rights statement

© 2011 Acoustical Society of America. The attached file is reproduced here in accordance with the copyright policy of the publisher. Please refer to the journal's website for access to the definitive, published version.

Downloaded from

<http://hdl.handle.net/10072/401174>

Griffith Research Online

<https://research-repository.griffith.edu.au>

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

Annelies Bockstael, Bert De Coensel, Dick Botteldooren, Wendy D'Haenens, Hannah Keppler, Leen Maes, Birgit Philips, Freya Swinnen, and Vinck Bart

Citation: *The Journal of the Acoustical Society of America* **129**, 3702 (2011); doi: 10.1121/1.3575599

View online: <https://doi.org/10.1121/1.3575599>

View Table of Contents: <https://asa.scitation.org/toc/jas/129/6>

Published by the [Acoustical Society of America](#)

---

### ARTICLES YOU MAY BE INTERESTED IN

[Methods of measuring the attenuation of hearing protection devices](#)

*The Journal of the Acoustical Society of America* **79**, 1655 (1986); <https://doi.org/10.1121/1.393228>

[Speech intelligibility in noise: Effects of fluency and hearing protector type](#)

*The Journal of the Acoustical Society of America* **71**, 708 (1982); <https://doi.org/10.1121/1.387547>

[Understanding speech when wearing communication headsets and hearing protectors with subband processing](#)

*The Journal of the Acoustical Society of America* **136**, 671 (2014); <https://doi.org/10.1121/1.4883385>

[Hybrid feedforward-feedback active noise reduction for hearing protection and communication](#)

*The Journal of the Acoustical Society of America* **120**, 2026 (2006); <https://doi.org/10.1121/1.2259790>

[Speech production in noise with and without hearing protection](#)

*The Journal of the Acoustical Society of America* **114**, 1069 (2003); <https://doi.org/10.1121/1.1592165>

[Effective compression and noise reduction configurations for hearing protectors](#)

*The Journal of the Acoustical Society of America* **121**, 1090 (2007); <https://doi.org/10.1121/1.2409859>

---



Advance your science and career  
as a member of the

ACOUSTICAL SOCIETY OF AMERICA

LEARN MORE



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

Annelies Bockstael,<sup>a)</sup> Bert De Coensel, and Dick Botteldooren

*Ghent University, Department of Information technology, Sint-Pietersnieuwstraat 41, 9000 Ghent, Belgium*

Wendy D'Haenens, Hannah Keppler, Leen Maes, Birgit Philips, Freya Swinnen, and Vinck Bart

*Ghent University, Department of Oto-rhino-laryngology, De Pintelaan 185, 9000 Ghent, Belgium*

(Received 18 December 2009; revised 31 January 2011; accepted 17 March 2011)

The perceived negative influence of standard hearing protectors on communication is a common argument for not wearing them. Thus, “augmented” protectors have been developed to improve speech intelligibility. Nevertheless, their actual benefit remains a point of concern. In this paper, speech perception with active earplugs is compared to standard passive custom-made earplugs. The two types of active protectors included amplify the incoming sound with a fixed level or to a user selected fraction of the maximum safe level. For the latter type, minimal and maximal amplification are selected. To compare speech intelligibility, 20 different speech-in-noise fragments are presented to 60 normal-hearing subjects and speech recognition is scored. The background noise is selected from realistic industrial noise samples with different intensity, frequency, and temporal characteristics. Statistical analyses suggest that the protectors’ performance strongly depends on the noise condition. The active protectors with minimal amplification outclass the others for the most difficult and the easiest situations, but they also limit binaural listening. In other conditions, the passive protectors clearly surpass their active counterparts. Subsequently, test fragments are analyzed acoustically to clarify the results. This provides useful information for developing prototypes, but also indicates that tests with human subjects remain essential.

© 2011 Acoustical Society of America. [DOI: 10.1121/1.3575599]

PACS number(s): 43.50.Hg, 43.66.Vt, 43.72.Dv [BSF]

Pages: 3702–3715

## I. INTRODUCTION

It is well-known that exposure to excessive noise levels might cause noise-induced hearing loss. Personal hearing protectors are a popular safety measure in occupational settings, but they require consistent use since the actual background level remains unaltered. By contrast, the concern of missing verbal cues and warning signals is often a reason for incorrect partial insertion or even temporary removal of the protector. Apart from adequate hearing conservation as such, this implies that the implementation of personal protection is only worthwhile if the protectors sufficiently preserve environmental awareness and—not to forget—if the users actually judge communication and signal detection under protectors positively.

Signal detection and understanding with hearing protectors depend on a complex of factors connected with listener, speaker, message, and environment. Hence, it is not surprising that different studies report contradictory conclusions with respect to the influence of hearing protectors on speech intelligibility in noise. 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 and, on the other hand, “augmented” protectors that actually convert the incoming sound with or without electronics, i.e., in an active or passive way.

For normal-hearing subjects, [Abel \*et al.\* \(1993\)](#) found that wearing classical hearing protectors is beneficial for speech perception in noise. However, studies designed to simulate on-the-job listening conditions have reported poorer audibility with standard earplugs ([Casali \*et al.\*, 2004](#)). Regarding hearing impaired subjects, [Abel \*et al.\* \(1982, 1993\)](#) demonstrate that wearing earmuffs or earplugs deteriorates word recognition in noise when a high-frequency or flat sensorineural hearing loss is present.

Whether or not standard protectors (slightly) improve speech intelligibility, communication in background noise remains difficult, even without any hearing loss. In addition, listeners with protectors might benefit less from relatively silent periods in fluctuating noise because of momentary overprotection. As a consequence, augmented hearing protectors have been developed ([Casali and Berger, 1996](#)), offering different possibilities to make sound attenuation more comfortable and to diminish the masking effect of noise on signals.

Nowadays, a wide variety of augmented protectors is available, each with their own operating mechanism. To the authors’ knowledge, only limited research has been published concerning communication with these protectors and the studies usually fail to univocally establish their benefits on signal perception in noise over standard passive protectors ([Abel \*et al.\*, 1993](#); [Abel and Spencer, 1997](#); [Casali \*et al.\*, 2004](#); [Dancer \*et al.\*, 1999](#); [Dolan and O’Loughlin, 2005](#)).

The current paper addresses prototypes of active hearing protectors designed to be worn in similar conditions as

---

<sup>a)</sup>Author to whom correspondence should be addressed. Electronic mail: bockstael@ugent.be

standard passive protectors. Different types of realistic background noise are selected and in each sound environment speech fragments are recorded using a head-and-torso-simulator (HATS) without any hearing protectors, with passive earplugs, and with active ones. The speech intelligibility for each speech-in-noise fragment is then determined from the scores of normal-hearing subjects. Additionally, efforts have been made to determine those acoustical features of the signals under hearing protector influencing speech recognition.

## II. MATERIALS AND METHODS

### A. Hearing protectors

The current research initially includes the three types of hearing protectors shown in Fig. 1: passive custom-made acrylic hearing protectors, active custom-made acrylic hearing protectors with volume control, and active foam earplugs. For the hearing protectors with volume control full and minimal amplification are selected. These settings are chosen because they provide clearly different output levels for the input levels under study.

The acrylic passive earplugs [Fig. 1(a)] have an inner bore drilled over the total length of the protector via which the actual attenuation can be modified depending on the inserted acoustical filter. Here, an ST35 filter is used offering 35 Lohm attenuation. The unit “Lohm” is used by [The LEE Company \(2006\)](#) to reflect flow resistance of gasses and is calculated by the following equation:

$$\text{Lohms} = [Kf_T P/Q] \quad (1)$$

with  $Q$  representing the gas flow (in standard liters per minute),  $K$  the gas units constant (to prevent the need to convert pressure and flow parameters into specific units),  $f_T$  a temperature correction factor, and  $P$  the upstream absolute pressure [in psi (absolute)]. Assessment by the German Institute for Occupational Safety and Health (BGIA) following [ISO 4869-1 \(1990\)](#) and [ISO 4869-2 \(1994\)](#) reveal that the attenuation amounts to 25 dB SNR (single number rating); the assumed protection values (APVs) per octave band can be found in Table I. The attenuation can be considered constant over the range of input levels used in this study.

Active protectors for their part contain a microphone and a loudspeaker so that the attenuation can be adapted to the incoming sound level. The active foam earplugs [Fig. 1(b)] amplify the incoming sound with constant gain whereas the ones with volume control [Fig. 1(c), battery compartment opened] have a brickwall limiter that imposes a hard “ceiling” on the loudspeaker output—the signal cannot amplify any further once it reaches the safety threshold and this regardless of the amplification settings.

### 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 HATS type 4128 C with left and right ear simulator [see Fig. 1(d)]. Additionally, the same HATS is used to record the speech material in an anechoic room. The rationale behind this approach and the characteristics of both speech and noise are discussed in this section.

#### 1. Noise material

The influence of noise on communication largely depends on four major components. First, the signal-to-noise ratio plays a very important role. The louder the noise compared to the speech signal, the more speech recognition is hampered. Secondly, the frequency spectrum of the noise also influences speech intelligibility. In this regard [Studebaker et al. \(1994\)](#) have shown that noise with relatively more energy in the higher frequencies might have a more

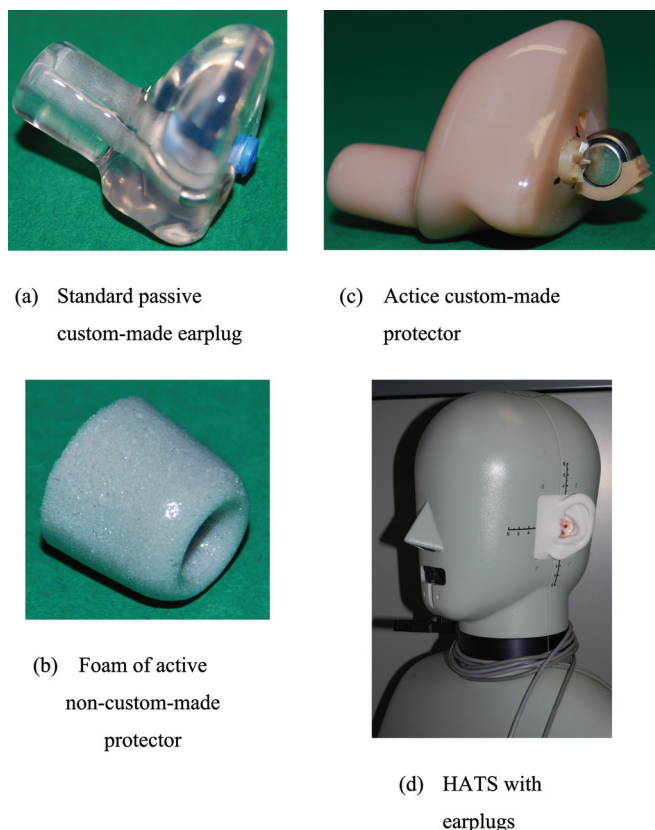


FIG. 1. (Color online) Pictures of the different hearing protectors and HATS used in this study.

TABLE I. Assumed protection values (APVs) in function of the octave bands center frequencies for the custom-made passive earplugs with ST35 filter measured by the German Institute for Occupational Safety and Health (BGIA).

Frequency (Hz)	APV (dB)
63	14.4
125	14.4
250	17.6
500	20.3
1000	21.9
2000	27.6
4000	26.3
8000	24.0

detrimental influence on speech recognition compared to other types of noise. Thirdly, the temporal characteristics of the noise have to be taken into account since for normal-hearing subjects intensity modulated background noise helps to unmask speech compared to steady-state noise (Festen and Plomp, 1990). Fourthly, the overall sound pressure level is of interest because elevated levels increase cochlear distortion, leading to reduced frequency discrimination (Casali and Berger, 1996).

In accordance with this knowledge, three types of noise are chosen for this study. First, recordings are made of alternators and turbines operating inside a power station hangar where the noise incidence is more or less diffuse. In contrast to this, a more directional 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. The device is preferred over a pressure microphone because binaural recordings, including the influence of head and body on the sound field, make it possible to reproduce proper spatial images thus allowing for spatial unmasking to be observed (see below). Naturally the ear-drum values thus obtained will differ from standard free-field recordings.

The time averaged spectra depicted in Fig. 2 reveal that the alternators and turbines produce more low-frequency energy whereas the bottle filling machine has more energy in the frequency region between 1000 and 10 000 Hz. The fork-lift trucks produce an overall lower sound pressure level. Only the spectra of the left ear simulator are shown when recordings are very similar for the left and the right ear simulator.

Naturally, speech intelligibility in noise is not only determined by the acoustical characteristics of the signals but also by the way they are listened to. In normal condi-

tions, the nervous system combines information from both ears. This binaural processing can enhance the detectability of signals in a background of maskers or detractors and hence improve the intelligibility of speech, especially if speech and noise come from different directions (Bronkhorst and Plomp, 1988).

To include the possible effect of spatial separation in the conducted experiments, the noise of the bottle filling machine is also recorded with the left ear pointing at the sound source. This particular machine is chosen because its directional characteristics produce the largest interaural difference if the HATS is no longer frontally facing the source. Figure 2 reveals that the sound pressure levels at the left and the right simulator differ most distinctly between 2000 and 5000 Hz.

## 2. Speech material

For the speech material, standardized recordings of the Dutch-language “Brugse Lijst” are chosen. This material—read by a professional female speaker—is especially designed to perform speech audiometry (see Sec. II D 2 for further details) and consists of consonant-vowel-consonant (the last consonant equal or not to the first) words spread among 20 lists with 17 words per list. Care is taken that each list is equally difficult, hence the specific list chosen for each test situation should not influence the result (Damman, 1994). The advantage of this type of words is their very low redundancy and the fact that the recognition of individual monosyllables in noise appears to be hardly improved by repetition (Sust *et al.*, 2009).

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 played with the audio equipment described in Sec. II C 2 and recordings are made in an anechoic room with a Renkus-Heinz (model CM 81) loudspeaker at 1 m from the HATS. 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 four lists, the HATS is turned with its right ear toward the sound source.

## C. Listening conditions

### 1. General considerations

The recorded speech and noise fragments are electronically mixed to create different sound environments, these fragments are then presented via headphones to the HATS with and without hearing protectors. Making two subsequent recordings—first with the unoccluded HATS in the field and then under headphones with and without hearing protectors—might seem needlessly time-consuming since speech and noise fragments could also be from the start recorded separately in the different listening conditions. However, speech and background noise have to be played back simultaneously because the active hearing protectors are level-dependent, which implies that the amplification is based on the overall sound pressure level. Since the speech is noticeably

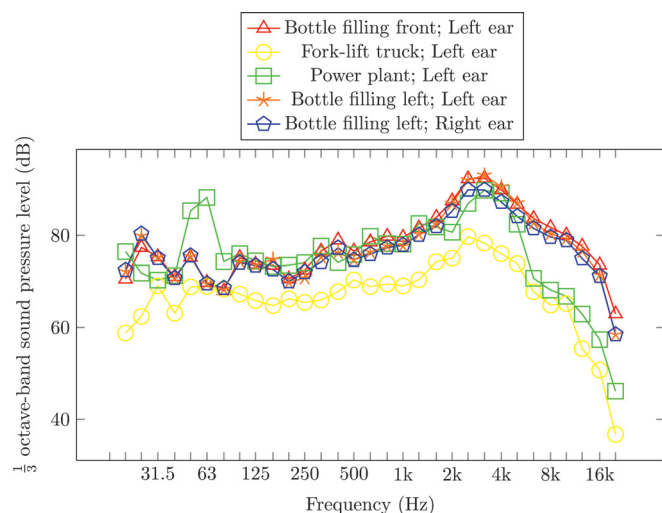


FIG. 2. (Color online) 1/3-octave band spectra ( $L_{eq}$  averaged over 80 s) of the recordings with HATS; “bottle filling front” refers to the bottle filling machine frontally recorded, “fork-lift trucks” to moving fork-lift trucks, “power plant” to the noise of alternators and turbines and “bottle filling left” to the bottle filling machine recorded with the HATS left ear facing the machine. “Left ear” are the recordings made by the left ear simulator while “right ear” refers to the right ear simulator.

softer than the noise, the processing of the speech will be clearly different if the words are presented alone or in combination with noise. Although the current approach requires extra processing, it has the additional benefit that the sound environments are essentially identical for the different listening conditions.

When making the second recordings with the HATS and for actual speech recognition tests, headphones are preferred over free-field stimulation to rule out unwanted influence from the test space, head movements or minor variations in the subject's position (Bronkhorst and Plomp, 1988). It is experimentally verified that the protectors operate similar—both in time and frequency domain—under this transducer and in real industrial settings.

## 2. Audio equipment

All recordings and the presentation of the listening material to the test subjects are carried out with the same audio equipment. The fragments are played on a laptop PC using AUDACITY software and then the signal is sent to an open circumaural Philips headphone (type SBC HP890) via a Pioneer A-607 R direct energy MOS amplifier.

## 3. Pre-processing

The project aims to test speech intelligibility in sound environments that are as realistic as possible, thus different issues have to be considered in the processing of the material.

First, the characteristics of the headphone and the double influence of the HATS are filtered out of the sound material. To establish the necessary filters, all original recordings are played again to the HATS via the headphones used in this study. Then, the spectra of the original recordings and the headphone recordings are compared, focusing on the frequency range most important for speech recognition, i.e., between 500 and 4000 Hz. Band-pass filters are designed with frequency response close to the difference between the recorded spectra.

The filters' gain is about 10 dB for 500 Hz and from this frequency on the spectra monotonously decrease with increasing frequency. This shows that the lower frequencies are softer than the original when played back through headphones. At 3000 Hz, the recordings under headphones slightly exceed the original (approximately 2 dB)—due to resonance of the outer ear canal—and from 3000 to 4000 Hz, the spectrum declines back to 0 dB gain.

Besides filtering, the test setup is kept as realistic as possible by selecting appropriate input levels for speech and noise. The level of the speech is set to 74 dB, measured at the HATS's eardrum, comparable to 68 dB(A) measured in free field (Hammershøi and Møller, 2008). This is in accordance to a normal communication situation where a female person would speak at 1 m free-field levels between 63 dB(A) (raised) and 71 dB(A) (loud) (Olsen, 1998). The calibration of the speech signals is done with continuous speech noise, especially developed for this particular set of speech material (Damman, 1994).

The noise fragments equal the original intensity at the work floor; only the loudest are slightly attenuated to protect

the participants' hearing. This is done by simply decreasing the overall level, preserving the spectral characteristics of the noise. Since there are only five recordings with open ears and each recording lasts about 75 s, it is possible to respect the safe exposure level with the sound pressure levels for open ear recordings tabulated in Table II. In this table, A-weighting is applied to allow comparison with the safety limits. Nevertheless, this type of weighting is less applicable for measurements at the eardrum since it also accounts for the influence of the outer and middle ear.

## 4. Recordings in different listening conditions

Prior to the recordings under headphones in different listening conditions, the filtered fragments of both speech and noise are set at an appropriate level (see Sec. II C 3) and mixed. The first 16 lists of the “Brugse Lijst”—frontally recorded—are combined with the noise from the alternators and turbines, from the fork-lift trucks and from the bottle filling machine recorded both frontally and sideways. The latter noise fragment is also combined with the last four speech lists, such that in these conditions noise predominantly

TABLE II. Overall A-weighted sound pressure level [dB(A)] for the 25 different test fragments recorded with the left (“left ear”) and right (“right ear”) ear simulator. “Sound environment” refers to the type of background noise (abbreviations similar to Fig. 2) and “listening condition” are the five different listening conditions: unoccluded, with passive protection (“passive”), augmented foam protectors (“active foam”), and custom-made augmented protectors with maximal (“active max”) and minimal (“active min”) amplification. Speech is, in general, recorded with the HATS frontally facing the loudspeaker, except for the last sound environment where the right ear simulators points at the loudspeaker (“speech right”).

Sound environment	Listening condition	Left ear	Right ear
Bottle filling front	Unoccluded	92.7	93.0
	Passive	58.0	58.4
	Active max.	81.7	82.3
	Active foam	96.1	96.2
	Active min.	76.0	79.3
Fork-lift trucks	Unoccluded	85.6	87.9
	Passive	51.2	54.0
	Active max.	81.3	81.1
	Active foam	94.0	91.0
	Active min.	69.4	73.2
Power plant	Unoccluded	91.9	91.2
	Passive	56.2	55.9
	Active max.	83.3	83.4
	Active foam	94.0	94.3
	Active min.	77.2	79.1
Bottle filling left	Unoccluded	94.0	92.0
	Passive	56.4	56.5
	Active max.	81.8	83.1
	Active foam	97.0	96.1
	Active min.	78.5	79.2
Bottle filling left; Speech right	Unoccluded	94.0	92.0
	Passive	56.0	56.2
	Active max.	82.3	83.3
	Active foam	97.0	96.0
	Active min.	78.2	78.9

comes from the left and speech from the right. In this way, five global sound environments are created. Within each sound environment, one particular noise fragment is combined with four successive speech lists to establish in total 20 unique speech-in-noise fragments.

Subsequently, for each of the five sound environments, recordings with the HATS are made under headphones in five listening conditions; without hearing protectors, with passive earplugs, with active foam earplugs, and with active custom-made plugs, the latter once at minimal and once at maximal amplification. Since there are only four unique fragments per sound environment, each time one speech list was used twice.

Before each series of recordings, it is ensured that the unoccluded sound pressure level of the speech reference noise—i.e., the noise used for the calibration of the speech signals—under headphone equals 74 dB at the left and right ear simulator. For both passive and active hearing protectors, it is also verified that the right and left ear simulator measure the same overall intensity level when the speech reference noise of 74 dB is presented as input signal.

## 5. 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 sound pressure level under the hearing protector is adapted, including bone and tissue conduction as a secondary pathway for sound transmission. The attenuation of the earplug ( $A_{\text{earplug}}$ ) with the HATS is calculated in 1/3 -octave band from the level difference between unoccluded and occluded recordings. Following [Hiselius \(2005\)](#), the effective attenuation ( $A_{\text{total}}$ ) in 1/3-octave bands is then calculated as the sum of two uncorrelated sound sources

$$A_{\text{total}} = -20 \log\left(10^{-A_{\text{earplug}}/20} + 10^{-A_{\text{bone}}/20}\right) \quad (2)$$

with  $A_{\text{bone}}$  the equivalent attenuation of the bone conduction reported by [Hiselius \(2005\)](#), in close agreement with bone conduction thresholds reported elsewhere ([Berger and Kerivan, 1983](#); [Berger et al., 2003](#)). Similar to the pre-processing (Sec. II C 3), a band-pass filter with frequency response equaling the difference between  $A_{\text{earplug}}$  and  $A_{\text{total}}$  is subsequently applied to the speech-in-noise fragments under passive protectors. As expected, the filter provides no gain below 1000 Hz, reaches at 2000 Hz (where the lowest bone conduction threshold is reported) a maximum close to 10 dB and then declines back to few decibels gain in the higher frequencies. It is experimentally confirmed that the correction is unnecessary for the active protectors by verifying that sound conduction via electronic amplification is much more prominent than possible contribution of bone and tissue conduction.

Apart from excluding bone and tissue conduction, the ear canal wall of the HATS might also be not representative for real human ears ([ISVR, 2003](#)). This becomes important for the passive earplugs when the sound is transmitted mechanically—by movements of the earplug in the ear canal—instead of acoustically. However, previous research

([Bockstael et al., 2008](#)) has shown that for these particular earplugs, the acoustical pathway clearly dominates and hence possible artificial ear canal boundary conditions are of less importance.

## 6. Selection of listening conditions

A summary of all test situations with the sound pressure level measured at the HATS ear simulators can be found in Table II. This table clearly shows that the active foam hearing protector amplifies the sound too strongly, even beyond the unoccluded level. Free-field measurements at the work floor confirm the clearly elevated levels under the foam earplug—between 90 and 95 dB—never seen for the other protectors.

Apart from the active signal processing, the unexpected results may be also caused by bad placement of the protector in the ear canal. Although the protectors are always inserted very carefully and the elevated levels are consistent over all measurements, it is possible that the foam earplug is simply not suited for the HATS ear canal.

To investigate this issue, the attenuation of the foam earplug is compared to that of the passive custom-made protector, both measured with the HATS in seven free-field conditions at the work floor. If the sound pressure level is dominated by leakages around and through the protector, one would expect a merely constant amount of attenuation, quasi-independent of the input level. This holds indeed true for the passive protector where there is no correlation between the levels measured with unoccluded ear canals and the attenuation (Spearman's rank correlation  $\rho = 0.06$ ;  $p = 0.83$ ). However, for the foam earplug the attenuation increases with increasing input level (Spearman's rank correlation  $\rho = 0.78$ ;  $p = 0.001$ ), suggesting that the active mechanism does adapt the amplification to the incoming sound, but not sufficiently to reduce the sound pressure level below safety levels. Because of this clear malfunction, the foam active earplug is removed from the listening tests.

## D. Speech intelligibility test

### 1. Test subjects

The projects aims to compare speech intelligibility for different controlled listening situations, and thus it is important that the general communication skills of the listeners do not influence the outcome of tests. Therefore only native Dutch-speakers who have at least successfully finished high school are included. All participated voluntarily and signed an informed consent.

Further, normal hearing is required with tonal hearing thresholds of 25 dB or better for all octave frequencies between 250 and 8000 Hz, as are normal tympanometric and speech audiometric results. The latter is tested monaurally in silence using the standardized Dutch consonant-vowel-consonant NVA-list ([Damman, 1993](#)).

These criteria yield to a group of 60 test subjects (30 female and 30 male) who are on average 27.6 years old. According to [Abel et al. \(1993\)](#) age is an insignificant factor with respect to consonant discrimination and word recognition, at least for the working population under study ([George et al.,](#)

2006). Further, no gender related significant differences in hearing level or in speech recognition are present in the test group.

## 2. Test setup

At the beginning of the experiment, each subject is told that he/she will hear fragments of speech in noise of approximately 1 min 15 s. Although none of the levels is harmful because of the short duration, it is stressed that the test can be stopped at any time if the subject feels that he cannot stand the noise.

Further instructions to the test subject are in conformity with classical speech audiometry. He/she is told that consonant-vowel-consonant words will be presented, but the words themselves are not revealed beforehand as this is an open test. Moreover, the subject is encouraged to repeat as much speech material as possible, even though he may have to guess or can only repeat one or two phonemes instead of a complete word. The investigator notes per word the correct phonemes (consonants and vowels) on a score form and afterwards the number of correctly understood phonemes is counted for each list, yielding to a specific speech recognition score. Since each list consists of 17 words and each word has three phonemes, the maximal speech recognition score per list is 51. There is always a short break between the different listening fragments and the subject has the option to take a longer break if necessary. In this way, the concentration level is kept up during the entire test.

The 20 speech-in-noise fragments are presented in random order. At the beginning, a test list is played so that the participants become used to the whole test concept.

## 3. Statistical analysis

The phoneme scores for the 20 fragments are mutually compared with analysis of variance (ANOVA) and appropriate post-hoc comparisons. The analyses are discussed in detail in the next section.

### III. RESULTS

#### A. Comparison of speech recognition scores

The mean speech recognition scores and standard deviations are depicted in Fig. 3 for the 20 test fragments. On this data a mixed model ANOVA is carried out with three crossed factors, using the statistical software SPSS. The variables “sound environment” and “listening condition” are included as fixed factors whereas the variable “subject” is included as a random factor. The three-way interaction effect is assumed insignificant based on visual inspection of the error bar chart (Kutner *et al.*, 2004) and the ANOVA is performed with two-way interaction and main effects. All two-way interaction effects are significant ( $\alpha=0.05$ ) (Kutner *et al.*, 2004), i.e., the interaction effect between sound environment and subject [ $F(236, 647) = 1.2; p = 0.032$ ], between listening condition and subject [ $F(162, 647) = 1.3; p = 0.009$ ] and finally between sound environment and listening condition [ $F(12, 647) = 77.7; p < 0.0001$ ].

Before this model is interpreted, it is checked that the residuals come from a normal distribution (Kutner *et al.*,

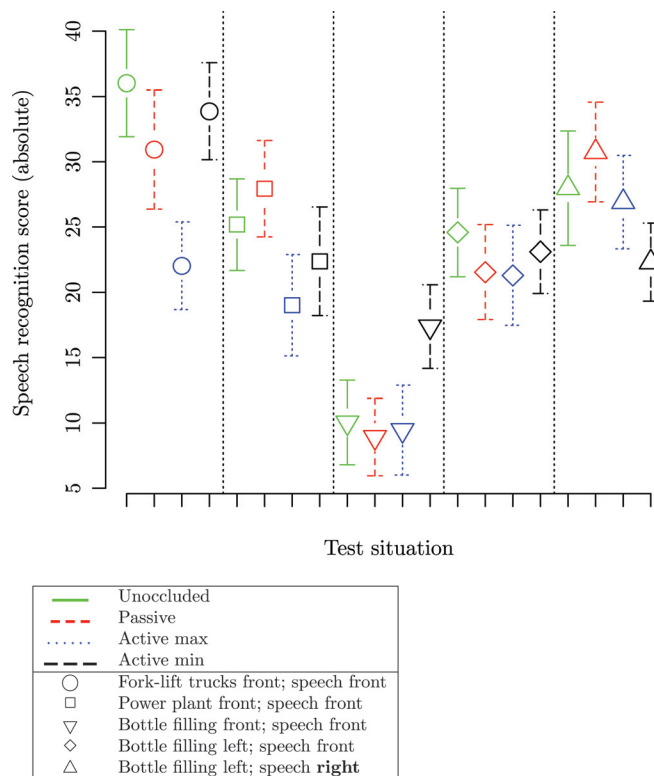


FIG. 3. (Color online) Bar plot of the speech recognition scores (absolute numbers) of the 20 test fragments. The center of the error bars is given by the mean speech recognition score whereas the width of the bars equals one standard deviation. Vertical dashed lines are added to separate the sound environments. Different colors refer to the listening conditions (for abbreviations see Table II). For the sound environments (different symbols—for abbreviations see Fig. 2), the input direction of the speech is indicated behind the semicolon, coming from the front or from the right.

2004). Both the Kolmogorov–Smirnov test and the Shapiro–Wilk normality test yield to insignificant  $p$ -values ( $p > 0.01$ ) and the residuals turn out to be randomly distributed around zero, showing that the mixed model approach is valid for this dataset.

Apparently, the influence of sound environment and listening condition does not apply to the same degree to all levels in the population of the random factor ‘subject’ because both random interaction effects are significant. Since the factor levels of these interactions effects constitute a sample from a larger population of factor levels, they are not of intrinsic interest in themselves (Kutner *et al.*, 2004). Conversely, the levels of the fixed interaction effect between the variables “sound environment” and “listening condition” are important for further analysis and therefore a pairwise Tukey-post hoc test is carried out with the interaction effect as independent variable (see Tables III and IV). The most striking results are discussed below.

#### 1. Listening conditions in sound environments

In Table III results are summarized in function of the sound environment. It becomes clear that the performance of the different hearing protectors strongly depends on the type of background noise. First, passive protectors appear to give the best results for the noise at the power plant and the noise



TABLE III. Summary of the pairwise comparisons between the listening conditions tabulated per sound environment. For each table, the “best” condition is put in the first column, the second best in the second and so on. In the cells, the differences between the absolute speech recognition scores of the respective column and row is obtained from the post-hoc statistical pairwise comparison. The result is printed bold if the difference is statistically significant. In addition, footnotes are used to indicate the magnitude of the associated  $p$ -value. All abbreviations are similar to Fig. 3. For instance, the value 7.1 in the first cell of this table means the average absolute speech recognition score under active protectors with a minimal amplification that is 7.1 higher than in the unoccluded situation and this difference is highly statistically significant.

(a) Bottle filling front; speech front			
	Active min.	Unoccluded	Active max.
Unoccluded	<b>7.1<sup>c</sup></b>		
Active min.	<b>7.7<sup>c</sup></b>	0.6	
Passive	<b>8.1<sup>c</sup></b>	1.1	0.5
(b) Fork-lift trucks; speech front			
	Unoccluded	Active min.	Passive
Active min.	<b>2.2<sup>a</sup></b>		
Passive	<b>5.1<sup>c</sup></b>	<b>3.9<sup>c</sup></b>	
Active max.	<b>14.0<sup>c</sup></b>	<b>11.8<sup>c</sup></b>	<b>8.9<sup>c</sup></b>
(c) Power plant front; speech front			
	Passive	Unoccluded	Active min.
Unoccluded	<b>2.8<sup>c</sup></b>		
Active min.	<b>5.5<sup>c</sup></b>	<b>2.7<sup>b</sup></b>	
Active max.	<b>8.9<sup>c</sup></b>	<b>6.2<sup>c</sup></b>	<b>3.4<sup>c</sup></b>
(d) Bottle filling left; speech front			
	Unoccluded	Active min.	Passive
Active min.	1.4		
Passive	<b>3.0<sup>c</sup></b>	1.6	
Active max.	<b>3.3<sup>c</sup></b>	1.9	0.3
(e) Bottle filling left; speech right			
	Passive	Unoccluded	Active max.
Unoccluded	<b>2.8<sup>c</sup></b>		
Active max.	<b>3.8<sup>c</sup></b>	1.1	
Active min.	<b>8.2<sup>c</sup></b>	<b>5.4<sup>c</sup></b>	<b>4.3<sup>c</sup></b>

<sup>a</sup> $p < 0.1$ .  
<sup>b</sup> $p < 0.05$ .  
<sup>c</sup> $p < 0.001$ .

from the bottle filling machine recorded from the left with speech coming from the right. For the noise of the fork-lift trucks, the passive earplugs appear to hamper speech recognition more than the unoccluded situation.

The performance of the active protectors clearly depends on the settings of the volume control. Whereas maximal gain leads 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.

## 2. Sound environments in listening conditions

In Table IV results are summarized in function of the listening conditions. From these analyses, two conclusions can be drawn directly; 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.

Further, spatial segregation between speech and noise enhances speech recognition for all listening conditions if the sound environments with frontally recorded bottle filling machine 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 for the active protector with minimal gain. Here, the status quo for the active protectors is certainly not due to the ceiling-effect—speech recognition simply reaching its highest point—because the unoccluded conditions have even higher scores.

## B. Acoustical analysis

### 1. Speech Intelligibility Index

Linking speech intelligibility with objective acoustical characteristics is naturally a long-existing challenge and to-date different procedures are available. Here, the Speech Intelligibility Index (SII) is calculated in accordance to the 1/3-octave band procedure from ANSI S3.5-1997. The first issue is creating suitable separate speech and noise input fragments since they have to be inherently mixed in the current test setup due to the operating mechanism of the active protectors. Additionally, the separating procedures proposed in the ANSI S3.5-1997 cannot be applied since the active protectors are clearly not linear.

For the background noise spectra the original speech-in-noise fragments are used because the noise level greatly exceeds the speech level in all frequency bands so that the final spectra are solely determined by the noise. A close comparison between noise and speech-in-noise fragments confirms this.

The speech spectra are estimated by comparing the corresponding unoccluded and occluded speech-in-noise spectra. For each of the 1/3-octave bands between 160 and 8000 Hz the difference is made between the sound pressure level (in dB) registered by the ear simulators with open ear canals and with protectors. This difference is then applied to the spectrum of the continuous speech noise fragment—originally provided by the developers of the “Brugse Lijst” (Damman, 1994)—recorded with unoccluded ear simulators. Hence, a corrected speech spectrum has been obtained for each background noise spectrum and thus the SII can be calculated for all test fragments, separately for the left and right ear simulator.

The outcome is in good agreement with the speech recognition scores across sound environments: sound environments with a higher speech recognition score have higher SII than the ones with lower speech recognition. However, results are less convincing for the different listening conditions within one sound environment. For the latter, the SII is almost consistently the highest for the active protectors (in both amplification settings) and the unoccluded conditions lead to the lowest SII.

The problem probably lies in the parameters  $K$  and  $L$  used to calculate the SII summed over all frequency bands  $i$

TABLE IV. Summary of the pairwise comparisons between the sound environments tabulated per listening condition. The result is printed bold if the difference is statistically significant. For the fork-lift trucks and the power plant, the recording direction is always frontal for speech and noise and therefore this is not noted explicitly in the tables.

(a) Unoccluded				
	Fork-lift trucks	Bottle filling left; Speech right	Power plant	Bottle filling left; Speech front
Bottle filling left; Speech right	<b>8.0<sup>a</sup></b>			
Power plant	<b>10.8<sup>a</sup></b>	<b>2.8<sup>a</sup></b>		
Bottle filling left; Speech front	<b>11.4<sup>a</sup></b>	<b>3.4<sup>a</sup></b>	0.6	
Bottle filling front; Speech front	<b>26.0<sup>a</sup></b>	<b>18.0<sup>a</sup></b>	<b>15.2<sup>a</sup></b>	<b>14.6<sup>a</sup></b>
(b) Passive				
	Fork-lift trucks	Bottle filling left; Speech right	Power plant	Bottle filling left; Speech front
Bottle filling left; Speech right	0.2			
Power plant	<b>3.0<sup>a</sup></b>	<b>2.8<sup>a</sup></b>		
Bottle filling left; Speech front	<b>9.4<sup>a</sup></b>	<b>9.2<sup>a</sup></b>	<b>6.4<sup>a</sup></b>	
Bottle filling front; Speech front	<b>22.0<sup>a</sup></b>	<b>21.8<sup>a</sup></b>	<b>19.0<sup>a</sup></b>	<b>12.6<sup>a</sup></b>
(c) Active max.				
	Bottle filling left; Speech right	Fork-lift trucks;	Bottle filling left; Speech front	Power plant
Fork-lift trucks	<b>4.9<sup>a</sup></b>			
Bottle filling left; Speech front	<b>5.6<sup>a</sup></b>	0.7		
Power plant	<b>7.9<sup>a</sup></b>	<b>3.0<sup>a</sup></b>	<b>2.3<sup>a</sup></b>	
Bottle filling front; Speech front	<b>17.5<sup>a</sup></b>	<b>12.6<sup>a</sup></b>	<b>11.9<sup>a</sup></b>	<b>9.6<sup>a</sup></b>
(d) Active min.				
	Fork-lift trucks	Bottle filling left; Speech front	Bottle filling left; Speech right	Power plant
Bottle filling left; Speech front	<b>10.7<sup>a</sup></b>			
Bottle filling left; Speech right	<b>11.2<sup>a</sup></b>	0.6		
Power plant	<b>11.2<sup>a</sup></b>	0.7	0.1	
Bottle filling front; Speech front	<b>16.7<sup>a</sup></b>	<b>6.0<sup>a</sup></b>	<b>5.4<sup>a</sup></b>	<b>5.3<sup>a</sup></b>

<sup>a</sup> $p < 0.001$ .

$$SII = \sum_{i=1}^k I_i L_i K_i \quad (3)$$

with  $I_i$  the band importance function given by the ANSI standard.  $K$  accounts for the difference between the equivalent speech and disturbance spectrum for each frequency band and is limited between 0 and 1.  $L$ , on the other hand, is the level distortion factor, per frequency band based on the difference between the equivalent speech and the standard speech spectrum at normal vocal effort.

In the current dataset, the disturbance largely exceeds the speech level and as a result  $K$  becomes zero for a considerable part of the frequency range, especially above 1000 Hz. This means that the final SII is mainly determined by a few discrete points, excluding a substantial amount of spectral information that might be important to detect subtle differences between listening conditions. Moreover, the level distortion factor  $L$  is naturally the lowest for the unoccluded conditions where the speech is the loudest. It also decreases with increasing frequency for the active protectors, but these points are not taken into account because there  $K$  equals zero. This can explain why the SII for the active protectors almost consequently exceeds the values in unoccluded condition.

It can be concluded that for the current assessment of listening conditions within different sound environments, calculating the SII is less illuminating to link acoustical characteristics with speech recognition, due to the aggregation steps necessary to obtain a single numeric outcome. Therefore possible important acoustical variables are studied separately.

## 2. Global sound pressure level

Assessing the overall sound pressure level with the HATS is the first step to ensure that a hearing protector can effectively prevent noise-induced hearing loss. In this, the malfunction of the foam active earplug is immediately clear.

Comparing sound environments, the post-hoc analysis reveals that the noise from the fork-lift trucks having the lowest overall sound pressure level is the least interfering with speech. 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. Table II also suggests that the bottle filling machine would yield to the lowest signal-to-noise ratio for the frontally recorded noise fragments, although the overall level difference with the noise from the power plant is quite small. This will be discussed further in Sec. III B 4.

Furthermore, the influence of the overall level at the eardrum is compared for the different listening conditions. In this matter, the effect of the passive protectors is definitely related to the unoccluded noise level. If the latter condition is well-above 85 dB(A)—as is the case for the bottle filling machine and the noise at the power plant—the passive protectors seem to improve or at least retain speech intelligibility compared to the unoccluded condition; which is in agreement with previous findings from Abel *et al.* (1993). For the softer noise from the fork-lift trucks, the passive earplugs appear to hamper speech recognition more than the unoccluded situation. It could be that the attenuation of the passive earplug is overprotective in this environment, introducing more difficulties in communication.

This clearly illustrates the importance of comparing speech recognition scores within one sound environment, instead of within one listening condition across sound environments. If one would solely compare the scores for the passive protectors across the tested sound environments, the less good performance for the fork-lift trucks would not be revealed because the absolute scores are still high compared to the bottle filling machine and the noise at the power plant. However, within the fork-lift truck noise, the results are poorer with passive earplugs than without any protection. This means that in real working conditions removing the protectors would actually enhance communication, but at the same time increase the risk of noise-induced hearing damage.

For the active protectors, the link between sound pressure level and speech recognition is much less clear, suggesting that other mechanisms govern the outcome. However, the active protectors with minimal gain do a good job for the noise from the fork-lift trucks, possibly because they keep the sound pressure level higher than their passive counterparts.

### 3. Time pattern

As explained previously in Sec. II B 1, the fluctuating character of the noise from moving fork-lift trucks improves speech intelligibility compared to the constantly working bottle filling machine and alternators and turbines. In this regard, Fig. 4 shows that the active hearing protector with maximal amplification does not preserve the temporal structure of the noise. Artificially equalizing the noise levels naturally cancels the benefits from silent gaps and might partially explain the low intelligibility scores for this particular test fragment.

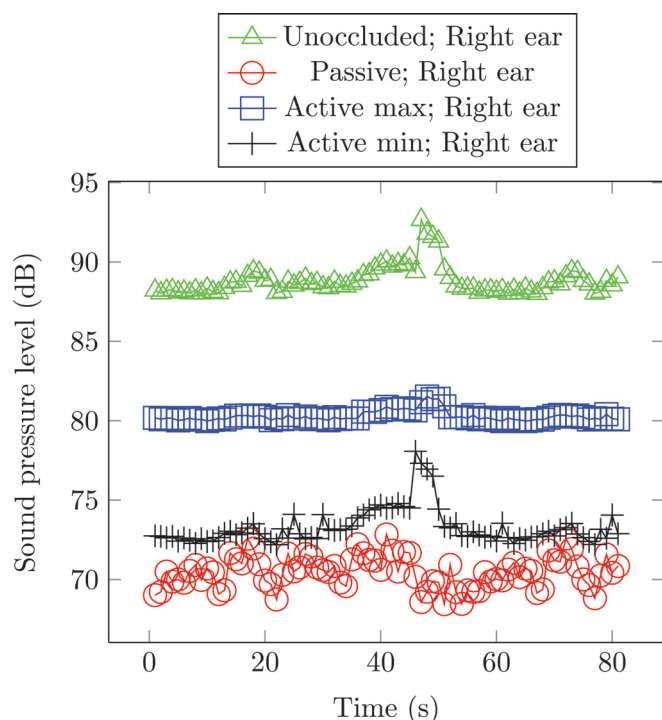


FIG. 4. (Color online) Overall sound pressure level per second ( $L_{eq1s}$ ) for the noise of the fork-lift trucks recorded with the right ear simulator. All abbreviations are similar to Fig. 3.

It might also be observed that the time pattern under the passive protector differs from the unoccluded situation. In this regard, it is verified that the levels under the passive protectors are correctly measured and not due to the noise floor of the measurement equipment. By contrast, the temporal differences result from different spectral characteristics. In general, the spectra under passive protectors contain more low-frequency energy—because their attenuation is proportional to the frequency—whereas frequencies above 1000 Hz are more present in unoccluded conditions. As a result, sound pressure fluctuations due to higher frequency sounds will be less prominent under passive earplugs. For instance, the clear increase around 50 s due to the high-frequency reverse alarm of a fork-lift truck is heavily suppressed by the passive protectors.

### 4. Spectral analysis

Apart from the overall sound pressure level and the time pattern, the spectra of the test fragments also influence speech intelligibility. For instance, the difference between the overall sound level at the bottle filling machine and at the power plant is quite small, but the speech recognition scores clearly differ. Figure 2 shows that the bottle filling machine indeed produces higher levels above 1000 Hz, i.e., in the region important for speech recognition. As far as communication is concerned, the low-frequency noise at the alternators and turbines is much less disturbing (Studebaker *et al.*, 1994).

To illustrate the effect of spectral contents on the listening conditions, the spectra of the four listening conditions with the fork-lift truck noise are depicted in Fig. 5. One of the most striking features is the interaural difference for low-frequency attenuation of the active protector. This is most prominent for the situation with minimal gain and will be discussed in Sec. III B 7. Further, the active protector with maximal amplification emphasizes strongly the frequencies between 1000 and 4000 Hz. The region is indeed important for speech perception, but excessive amplification might introduce distortion.

### 5. Distortion

Active hearing protectors might distort the signals in two ways; first they can excessively amplify one particular frequency region, thus altering the original relation between the different frequency bands. Secondly, strong initial amplification will require strong compression in (highly) elevated background noise, possibly introducing nonlinear distortion.

To quantify the first form of linear spectral distortion, the amplitude of speech-in-noise fragments is investigated more closely between 1000 and 4000 Hz. More specifically the magnitude in the unoccluded condition is subtracted from each occluded condition within the same sound environment and this for each 1/3-octave band in the frequency region of interest. The spread of the differences is assessed by calculating the coefficient of variance (cv) as

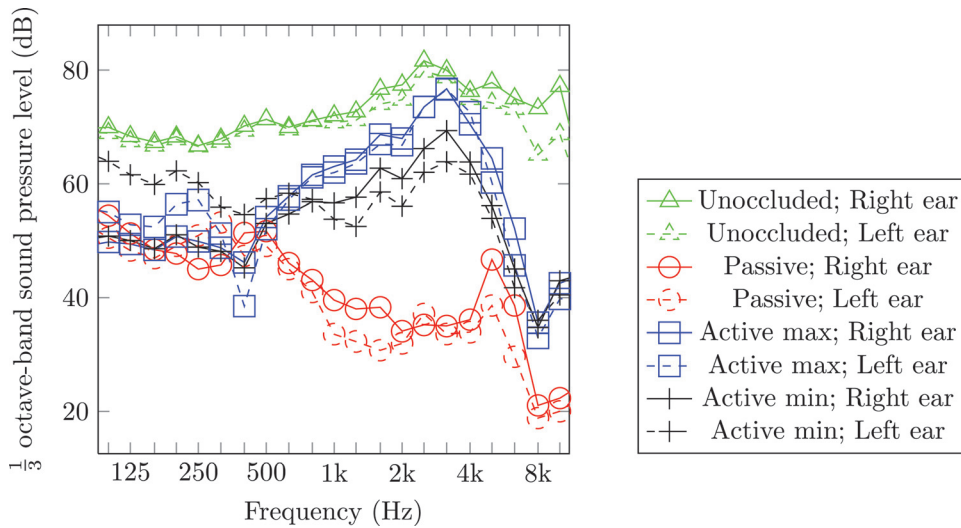


FIG. 5. (Color online) 1/3-octave band spectra of the noise from the fork-lift trucks in different listening conditions as recorded by the left and right ear simulator. All abbreviations are similar to Fig. 3.

$$cv = \frac{s}{\bar{y}}(100\%) \quad (4)$$

with  $s$  the standard deviation of the observations and  $\bar{y}$  the mean.

If the hearing protectors respect the relation between the different frequency bands seen in the original signal, the amplification or attenuation should be more or less constant across this frequency range and therefore the coefficient of variation should be quite low. This is the case for the passive protectors and the active protectors with minimal amplification; over the different sound environments the coefficient of variance is maximal 16%. By contrast, for the active protectors with maximal amplification, the coefficient is minimal 20%, and even reaches 44% for the noise of the fork-lift trucks. These results suggest that the low intelligibility scores for maximal amplification might partially be associated with linear distortion of the original spectrum.

In addition, nonlinear distortion might also play its role. Recall that the active protectors are equipped with a brick-wall limiter to control the output level. This type of limiter largely compresses the signal so that the output level do not exceed the safety threshold, but hearing aid practice learns that this kind of signal processing might introduce harmonic and intermodulation distortion (Dillon, 2001), making brick-wall limiting detrimental for speech recognition if often applied. In this regard, the temporal pattern of the active protectors with maximal amplification suggest that they operate almost constantly at their output limit (see Fig. 4), thus strong compression—and possible nonlinear distortion—occurs during the whole fragment.

The effect can be quantified by calculating the coherence of broadband signals, taking both harmonic and intermodulation distortion into account (Dillon, 2001). For the calculation, the approach described by Kates (1992) is followed as closely as possible, including careful temporal alignment and smoothing over 1/3-octave bands. Kates (1992) defines the coherence  $\gamma(\omega)$  as the normalized spectral density

$$\gamma(\omega) = \frac{S_{xy}(\omega)}{\sqrt{S_{xx}(\omega)S_{yy}(\omega)}}, \quad (5)$$

where  $S_{xy}(\omega)$  is cross-spectral density between two stationary random processes  $x(n)$  and  $y(n)$  having auto-spectral densities  $S_{xx}(\omega)$  and  $S_{yy}(\omega)$ , respectively. The signal-to-distortion ratio SDR is then calculated as

$$SDR = \frac{|\gamma(\omega)|^2}{1 - |\gamma(\omega)|^2}. \quad (6)$$

In the analysis, the original test fragments are used as output signals to ensure that the actual test situation and settings of the hearing protectors are addressed. However, this implies that—compared to the original input signal—the frequency response of the headphone and the HATS will also be taken into account. To eliminate these factors as much as possible, the output signals are not compared to the original input signal but to the corresponding unoccluded test fragment. In addition, the focus does not lie on the SDR as such, but rather on the difference for minimal and maximal amplification. This allows to directly assess the influence of the amplification settings because the identical test setup cancels out other possible sources of nonlinear distortion.

The calculations reveal that the signal-to-distortion ratio is systematically higher between 2000 and 4000 Hz in case of minimal amplification and this for all sound environments except the fork-lift trucks, suggesting that the active protectors with maximal amplification introduce more nonlinear distortion in this frequency range. Indeed, the amplification is mostly applied in the higher frequencies, increasing the need for compression—and the risk of distortion—for higher input levels. This also explains why the effect is not observed for the fork-lift trucks; the sound environment with the lowest input level will probably demand less compression.

## 6. Loudness

A key element when studying the influence of noise on speech is the amount of energetic masking. This is most

commonly assessed by comparing the A-weighted sound pressure level of speech to that of noise yielding to a certain signal-to-noise ratio. However, much more relevant information can be gained.

For instance, the beneficial effect of passive protectors on speech recognition in noise might be explained with the generally accepted theory that by lowering the total incident energy of both speech and noise, passive protectors alleviate cochlear distortion that occurs at high sound levels (Casali and Berger, 1996). This effect and other important variables like bandwidth, frequency content and duration of sounds can be included by working with “loudness” (in sone) instead of the A-weighted sound pressure level (Zwicker and Fastl, 1999).

The loudness of speech and noise is calculated following the ISO 532-1975 (1975) standard completed with improvements proposed by Zwicker and Fastl (1999) for sounds with strong low-frequency components. From the loudness calculated for the left and right ears, the binaural loudness can be derived following the approach of Moore and Glasberg (2007). Again, speech and noise are made separately available similar to Sec. III B 1. This procedure gives for each test fragment the loudness of speech and noise in function of the critical band-rate (in bark) approximating the frequency selectivity of the hearing system (Zwicker and Fastl, 1999). The signal-to-noise ratio expressed in loudness per bark can then be derived easily by dividing the loudness of speech and noise.

Comparing these ratios between the five sound environments confirms the conclusions drawn earlier. When the ratios are mutually compared between the different listening conditions, the most striking results are found for the noise at the power plant where the ratio for the passive protectors (on average 0.54) clearly exceeds the ratio for the others in the higher critical band rates (on average 0.47 for open condition, 0.27 for active protectors with maximal gain, and 0.31 for the same protectors with minimal gain).

## 7. Interaural differences

Research has clearly demonstrated that (speech) signals can be more easily detected and understood if the input of the left and right auditory system is combined (Hafer and Trahiotis, 1997). This binaural unmasking mechanism is found to rely on two important cues, interaural time difference (ITD), and interaural level difference (ILD).

Changing the recording position at the bottle filling machine from frontal to sideways (left) significantly increases the speech recognition scores for all listening conditions. For most of these conditions, moving the noise source indeed increases the ILD by lowering the noise level at the right ear due to head-shadow (Stern *et al.*, 2006). Since the level of the speech remains constant, the signal-to-noise ratio will become more advantageous. Further relocation of the speech source to the right will similarly increase the speech level and signal-to-noise ratio at this ear. However, the ILD cannot explain why in case of the active protectors with maximal amplification, the speech recognition does increase when moving the noise source, despite an increase

in noise level (instead of decrease) at the ear where a head shadow would be expected.

Calculating the interaural cross-correlation function confirms that all listening conditions—including the active protectors with maximal amplification—do preserve the ITD that marks the relocation of speech and noise source (Hafer and Trahiotis, 1997). This might explain the amelioration in speech intelligibility for the active protectors with maximal amplification, but it does not make clear why active protectors with minimal amplification do not yield to better speech recognition since they do not markedly distort changes in ITD or ILD compared to the other sound environments.

However, Table II and Fig. 5 clearly show interaural differences for this particular protector in sound environments where speech and noise are frontally recorded; the spectra under the left protector have somewhat more low-frequency energy compared to the right protector whereas the opposite is true for the higher frequencies. These inequalities introduced by the earplugs themselves might obscure changes in ILD and ITD when noise and/or speech sources are actually moved, hence reducing the spatial unmasking effect on speech intelligibility.

A possible cause for these findings is a different fit inside the ear canal. Prior to the recordings, it is ensured that the left and right ear simulator register under protector the same overall output level for the reference input signal, this by adapting the amplification of the earplug with the adjustable rotary knob. Although the output is clearly dominated by the electronic pathway—see for instance the difference in sound pressure level under similar passive and active protectors—it is not unthinkable that a fraction of the input is transmitted via leakages around the protector, which indeed would be more prominent in the lower frequencies (Apfel, 1997). Especially for lower levels of electronic amplification, the extra low-frequency contribution would allow for slightly lower volume settings to obtain the same global output as earplugs with a tighter fit; this would then result in somewhat lower energy for the frequencies between 1000 and 4000 Hz.

## IV. 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 protection on speech intelligibility; the key question is, of course, whether they can actually come up to the expectations.

The first task of any hearing protector is the prevention of noise-induced hearing loss. This implies that active protectors with electronic circuits must be able to fulfill their safety function at all times, even when the electronic components fail (Buchweiller *et al.*, 2003). Measurements with the active foam earplug in this study reveal that the fulfillment of this requirement cannot be taken for granted.

Once it is verified that the protectors reduce the sound pressure level sufficiently, their influence on speech perception can be addressed. Here this is done by recording samples of speech and noise with and without hearing protectors

and then scoring the speech recognition for a group of volunteers. It should be noted that the chosen method is not representative for communication in real working conditions. Nevertheless, the somewhat unrealistic listening conditions are a necessity since the study's major aim is a direct comparison of speech perception under different protectors. A fair assessment of the signal processing quality assumes that the observed variations can be attributed to true differences between the protectors, not to other confounding factors influencing communication.

To respect this requirement, equality of input signals is a key factor. Therefore lists for speech audiometry are preferred over realistic utterance since they are designed to be equally difficult (Damman, 1994). This is of major importance because each test fragment naturally consists of a different set of words.

When it comes to the noise sets, recordings are made with a HATS instead of presenting the material directly to a test subject wearing the different protectors. Despite the artificial character of the HATS, the followed approach averts that intersubject variation in the protectors' fit either suggests or masks the effects of a particular listening condition. Another advantage is that the different noise recordings can be easily used for other types of (active) hearing protectors, making new results comparable to those of the current study.

Within all these constraints of a controlled test design, the test fragments themselves are kept as close to reality as possible. First this implies that speech and noise have to be mixed before sending them to the level-dependent active protectors. Although the procedure somehow complicates the acoustical analysis, it does ensure that the influence of the active protectors is for certain correctly included. Conversely if speech and noise are recorded as individual signals in different listening conditions, precise knowledge about the signal processing is indispensable for the mixing and the accuracy of the assumptions can never be fully guaranteed.

The recordings of the speech are made in an anechoic room with the HATS placed at an appropriate distance (1 m) of the sound source (Olsen, 1998). The free-field conditions in the anechoic room mimic the situation where communication partners stand relatively close to each other, minimizing the influence of environmental reverberations on the speech signal. In addition, possible effects of head-and-torso transfer functions and of binaural listening are included.

For the level of the speech, it is clear that a speaker has at least to raise his voice in all selected sound environments; therefore an elevated fixed level is chosen (Olsen, 1998). Alternatively, the actual speech signal could have been estimated more accurately, taking into account the complex of acoustic-phonetic changes known as the Lombard effect. This can be modeled to a certain extent under unoccluded conditions (Corthals, 2004), but things will be different with hearing protectors and might even depend on the actual type—i.e., standard or augmented. Tufts and Frank (2003) have shown that the Lombard effect is much less pronounced when people wear passive protectors since the attenuation and the occlusion effect alter the perception of the background and their own voice. It is possible that active protectors preserve the effect better because the transmitted noise

levels are higher; on the other hand, the occlusion effect will still be present (although the net effect will probably depend on the speech level heard via air-conduction) and the active protectors might also emphasize the higher frequency content of background noise which is potentially less favorable to invoke the Lombard effect (Lu and Cooke, 2009). The importance of all this can be discussed (Lu and Cooke, 2009; Junqua, 1996) but it is clear that an accurate estimation of the Lombard effect on speech production under hearing protectors is a very complex task. Naturally, the type of protector worn by the listener will also influence his benefits of adapted speech. Here, the energy shift to higher frequencies (Lu and Cooke, 2009) might be more perceived under the active protectors whereas the standard ones might cross the effect due to increased attenuation with increasing frequencies.

Making assumptions on the type of protectors worn by speaker and listener together with their individual acoustical environment (all of which are not necessarily the same) would largely complicate the comparison between sound environments whereas the gain in accuracy would be questionable. Therefore, a fixed speech level without further acoustical modifications is considered to be a more valid approach for the research questions at hand. Including the Lombard effect would nevertheless be very interesting in a later stage.

To create the sound environments, speech fragments are combined with a variety of industrial noise samples. Because the focus lies on noise conditions for which the hearing protectors are effectively suitable, no extremely loud or soft situations are included nor are any impulsive sounds. The fact that the different sound environments have different acoustical characteristics and that they yield to statistically significant differences in speech recognition confirms that the protectors are tested in a sufficiently wide variety of situations.

Statistical analysis also reveals that variation in speech intelligibility across listening conditions depends on the sound environment under study. In this regard, objective acoustical parameters and speech recognition have been linked by different indexes like the Speech Intelligibility Index (SII) (George *et al.*, 2006) and the Speech Transmission Index (STI) (Payton *et al.*, 1994), but Sec. III B 1 suggests that aggregation steps involved might mask subtle differences between the studied listening conditions. Instead, acoustical parameters are directly used to elucidate the performance of test subjects in different listening conditions. Some findings are quite striking and should be kept in mind for further development, but one should be aware that there exist a lot of possible interesting acoustical parameters, all interacting closely. Therefore, conclusions upon the effects of the particular features addressed in this study are drawn with caution.

First, the active protector with maximal gain decreases speech intelligibility due to distortion and over-equalization of temporal fluctuations; phenomena known as highly unwanted side-effects of active hearing protectors (Casali and Berger, 1996). The active protectors with minimal amplification clearly do a good job in the most silent

condition, but passive protectors are beneficial in the louder noise conditions. One must bear in mind that these positive effects are found for normal-hearing subjects and this excludes by no means that they could degrade communication if a hearing loss is present (Abel *et al.*, 1982).

Further, the passive earplugs and the active protectors with maximal gain let the user benefit from the unmasking effect of binaural listening. In this regard it must be noted that preservation of the binaural unmasking effect does not necessarily guarantee good sound localization under hearing protection. Several lines of research indicate that the relationship between sound localization and binaural unmasking is not as tight as one might expect, despite the fact that they depend upon the same acoustical cues (Brungart *et al.*, 2004).

Comparing ILD and ITD across sound environments corresponds to a good extent to the observed variation in speech recognition scores due to spatial separation of speech and noise. However, they do not provide clarification of the results for active protectors with minimal gain. A possible cause is that ITD and ILD do not simply add up (Bronkhorst and Brungart, 2005) and although the cues are a valid starting point to understand binaural listening, only much more complicated models can fully cover the effects (Stern *et al.*, 2006).

In this respect, the apparent unequal attenuation of the active protectors with minimal gain might also contribute to the limited effect of binaural unmasking by distorting the interaural differences. If this is indeed due to different fitting in the ear canal, one could argue that the issue is more a methodological error of the current research than a true deficit of the earplugs. Nevertheless, the authors are convinced that these problems might also occur in practice. First, the current rotary knob has no scale, hence interaural equalization has to be based on level comparison by measuring in the laboratory or, in practice, by listening. Furthermore, it is virtually impossible to adjust the volume precisely when the earplug is placed in the (HATS's) ear canal. Hence, it could not first be ensured that both protectors have the same fitting (by comparing the passive attenuation) and subsequently adapt the volume setting. This implies that an air-conducted pathway remains always possible with all its consequences for protectors with minimal amplification. Although the interaural differences introduced by the protectors do not necessarily hamper communication—given the positive results of the active protector for the fork-lift trucks and for the frontally recorded noise of the bottle filling machine—these considerations beyond purely electronic signal processing should be taken into account when implementing active protectors.

The distinct connection between acoustical parameters and speech recognition described above could open the door to a more formalized relationship, for instance based on regression analysis. If a standard set of sound material is used, this approach would allow one to predict the influence of new protectors on communication from recordings with a HATS. In this research, there are insufficient fragments to establish such a model because of the complex interaction of the influential acoustical parameters. However, the test setup does permit to enlarge the existing data with speech recogni-

tion scores for new protectors using the same input signals. These efforts are advantageous because once a valid model is built, the effect of innovations in hearing protectors on speech perception can be assessed more easily before time-consuming tests with human subjects are to be carried out. Nevertheless, tests with human subjects should always be the last step before new protectors are put into practice because predictive models will never fully capture the complexity of human communication.

## V. CONCLUSION

From the research conducted in this paper, it becomes clear that speech recognition with hearing protectors strongly depends on the specific sound environment under study. In this matter, passive protectors might improve speech intelligibility if the unoccluded levels are sufficiently high. By contrast, they can be (momentarily) overprotective in more silent and fluctuating conditions and in those cases the active custom-made protectors appear in this study a more preferable alternative if their amplification is set sufficiently low. However, a critical issue with the active protectors is to preserve the beneficial effects of spatial unmasking along with the risk of distortion if the gain is maximized.

The apparent variation among the performance of hearing protectors strongly suggests that newly developed augmented protectors should be rigorously tested beforehand in different sound environments. The protocol proposed here allows such testing; in addition it permits one to reuse the input material for similar recordings with other protectors so that more protectors can be mutually compared. The conducted acoustical analysis might initiate the establishment of a predictive model for speech recognition, making it easier to judge the benefits of certain innovations. However, the complexity of communication suggests that tests with human subjects are always required in the final stage of the developing process.

## ACKNOWLEDGMENTS

The authors would like to thank HearingCoach and Stefan Voortmans for their help in supplying the necessary equipment. B.D.C. and A.B. are postdoctoral fellows of the Research Foundation—Flanders; the support of this organization is gratefully acknowledged.

- Abel, S., Alberti, P., Haythornthwaite, C., and Riko, K. (1982). "Speech intelligibility in noise: Effects of fluency and hearing protector type," *J. Acoust. Soc. Am.* **71**, 708–715.
- Abel, S., Armstrong, N., and Giguère, C. (1993). "Auditory perception with level-dependent hearing protectors: The effect of age and hearing loss," *Scand. Audiol.* **22**, 71–85.
- Abel, S., and Spencer, D. (1997). "Active noise reduction versus conventional hearing protection relative benefits for normal-hearing and impaired listeners," *Scand. Audiol.* **26**, 155–167.
- ANSI S3.5–1997 (1997). *Methods for Calculation of the Speech Intelligibility Index* (American National Standards Institute, New York).
- Apfel, R. E. (1997). *Encyclopedia of Acoustics* (Wiley, New York), Chap. Acoustic lumped elements from first principles, pp. 161–164.
- Berger, E. H., and Kerivan, J. E. (1983). "Influence of the physiological noise and the occlusion effect on the measurement of real-ear attenuation at threshold," *J. Acoust. Soc. Am.* **74**, 81–94.
- Berger, E. H., Kieper, R. W., and Gauger, D. (2003). "Hearing protection: Surpassing the limits to attenuation imposed by the bone-conduction pathways," *J. Acoust. Soc. Am.* **114**, 1955–1967.

- Bockstael, A., de Greve, B., Van Renterghem, T., Botteldooren, D., D'haenens, W., Keppler, H., Maes, L., Philips, B., and Vinck, B. (2008). "Verifying the attenuation of earplugs in situ: Method validation using artificial head and numerical simulations," *J. Acoust. Soc. Am.* **124**, 973–981.
- Bronkhorst, A., and Brungart, D. (2005). "Advances in research on spatial and binaural hearing," *Acta Acust. Acust.* **91**, V–XII.
- Bronkhorst, A. W., and Plomp, R. (1988). "The effect of head-induced interaural time and level differences on speech intelligibility in noise," *J. Acoust. Soc. Am.* **83**, 1508–1516.
- Brungart, D., Kordik, A., and Simpson, B. (2004). "The effects of single and double hearing protection on the localization and segregation of spatially-separated speech signals (L)," *J. Acoust. Soc. Am.* **116**, 1897–1900.
- Buchweiller, J.-P., Mayer, A., Klein, R., Iotti, J.-M., Kusy, A., Reinert, D., and Christ, E. (2003). "Safety of electronic circuits integrated into personal protective equipment (PPE)," *Safety Sci.* **41**, 395–408.
- Casali, J., and Berger, H. (1996). "Technology advancements in hearing protection circa 1995: active noise reduction, frequency/amplitude-sensitivity and uniform attenuation," *Am. Ind. Hyg. Assoc. J.* **57**, 175–185.
- Casali, J., Robinson, G., Dabney, E., and Gauger, D. (2004). "Effect of electronic ANR and conventional hearing protectors on vehicle backup alarm detection in noise," *Hum. Factors* **46**, 1–11.
- Corthals, P. (2004). "Sound pressure level of running speech: percentile level statistics and equivalent continuous sound level," *Folia Phoniatr. Logop.* **56**, 171–181.
- Damman, W. (1993). "Spraaaudiometrie in de praktijk (Speech audiometry in practice)," *Tijdschrift voor logopedie en audiologie* **23**, 15–38.
- Damman, W. (1994). "De Brugse lijst voor spraaudiometrie - versie 1989 (The Brugse lijst for speech audiometry - version 1989)," *Tijdschrift voor logopedie en audiologie* **24**, 21–37.
- Dancer, A., Buck, K., Hamery, P., and Parmentier, G. (1999). "Hearing protection in the military environment," *Noise Health* **5**, 1–15.
- Dillon, H. (2001). *Hearing Aids* (Thieme Medical, New York), Chap. Electroacoustic performance and measurement, pp. 74–116.
- Dolan, T., and O'Loughlin, D. (2005). "Amplified earmuffs impact on speech intelligibility in industrial noise for listeners with hearing loss," *Am. J. Audiol.* **14**, 80–85.
- Festen, J., and Plomp, R. (1990). "Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing," *J. Acoust. Soc. Am.* **88**, 1725–1736.
- George, E. L. J., Festen, J. M., and Houtgast, T. (2006). "Factors affecting masking release for speech in modulated noise for normal-hearing and hearing-impaired listeners," *J. Acoust. Soc. Am.* **120**, 2295–2311.
- Haftner, E. R., and Trahiotis, C. (1997). *Encyclopedia of Acoustics* (Wiley, New York), Chap. Functions of the binaural system, pp. 1461–1479.
- Hammershøi, D., and Møller, H. (2008). "Determination of noise immission from sound sources close to the ears," *Acta Acust. Acust.* **94**, 114–129.
- Hiselius, P. (2005). "Attenuation of earplugs—objective predictions compared to subjective REAT measurements," *Acta Acust. Acust.* **91**, 764–770.
- ISO 4869–1 (1990). *Acoustics—Hearing Protectors—Part 1: Subjective Method for the Measurement of Sound Attenuation* (International Organization for Standardization, Geneva, Switzerland).
- ISO 4869–2 (1994). *Acoustics—Hearing Protectors—Part 2: Estimation of Effective A-weighted Sound Pressure Level when Hearing Protectors are Worn* (International Organization for Standardization, Geneva, Switzerland).
- ISO 532–1975 (1975). *Acoustics—Method for Calculating Loudness Level* (International Organization for Standardization, Geneva, Switzerland).
- ISVR (2003). "Background information on head and torso simulators and relevant standards," Consultancy Report (Institute of Sound and Vibration Research Consulting, University of Southampton, Southampton).
- Junqua, J. (1996). "The influence of acoustics on speech production: A noise-induced stress phenomenon known as the Lombard reflex," *Speech Commun.* **20**, 13–22.
- Kates, J. (1992). "On using coherence to measure distortion in hearing aids," *J. Acoust. Soc. Am.* **91**, 2236–2244.
- Kutner, M. H., Nachtsheim, C. J., Neter, J., and Li, W. (2004). *Applied Linear Statistical Models*, 5th ed. (McGraw-Hill, New York).
- Lu, Y., and Cooke, M. (2009). "Speech production modifications produced in the presence of low-pass and high-pass filtered noise," *J. Acoust. Soc. Am.* **126**, 1495–1499.
- Moore, B., and Glasberg, B. (2007). "Modeling binaural loudness," *J. Acoust. Soc. Am.* **121**, 1604–1612.
- Olsen, W. (1998). "Average speech levels and spectra in various speaking/listening conditions: A summary of the Pearson, Bennett, & Fidell report," *Am. J. Audiol.* **7**, 21–25.
- Payton, K. L., Uchanski, R. M., and Braid, L. D. (1994). "Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing," *J. Acoust. Soc. Am.* **95**, 1581–1592.
- Stern, R. M., Brown, G. J., and Wang, D. (2006). *Computational Auditory Scene Analysis* (IEEE Press, Piscataway), Chap. Binaural sound localization, pp. 147–185.
- Studebaker, G., Taylor, R., and Sherbecoe, R. (1994). "The effect of noise spectrum on speech recognition performance-intensity functions," *J. Speech Lang. Hear. Res.* **37**, 439–448.
- Sust, C. A., Lazarus, H., Steckel, R., Kulka, M., and Kurtz, P. (2009). "Assessing speech comprehension in noise – acoustic quality of speech for near real communication conditions," *Acta Acust. Acust.* **95**, 86–96.
- The LEE Company (2006). "Laws for gas – how to calculate flow resistance for gasses," [www.theleeco.com](http://www.theleeco.com) (Last viewed September 27, 2010).
- Tufts, J., and Frank, T. (2003). "Speech production in noise with and without hearing protection," *J. Acoust. Soc. Am.* **114**, 1069–1080.
- Zwicker, E., and Fastl, H. (1999). *Psychoacoustics* (Springer, Germany), Chap. Loudness, pp. 203–238.