Comparison of speech intelligibility measures for an electronic amplifying earmuff and an identical passive attenuation device

David C. Byrne,[1] Catherine V. Palmer[2]

[1]National Institute for Occupational Safety and Health, Cincinnati, OH; [2] Department of Communication Science and Disorders, University of Pittsburgh, Pittsburgh, PA, USA

Abstract

The purpose of this study was to identify any differences between speech intelligibility measures obtained with MineEars electronic earmuffs (ProEars, Westcliffe, CO, USA) and the Bilsom model 847 (Sperian Hearing Protection, San Diego, CA, USA), which is a conventional passive-attenuation earmuff. These two devices are closely related, since the MineEars device consisted of a Bilsom 847 earmuff with the addition of electronic amplification circuits. Intelligibility scores were obtained by conducting listening tests with 15 normal-hearing human subject volunteers wearing the earmuffs. The primary research objective was to determine whether speech understanding differs between the passive earmuffs and the electronic earmuffs (with the volume control set at three different positions) in a background of 90 dB(A) continuous noise. As expected, results showed that speech intelligibility increased with higher speech-to-noise ratios; however, the electronic earmuff with the volume control set at full-on performed worse than when it was set to *off* or the lowest *on* setting. This finding suggests that the maximum volume control setting for these electronic earmuffs may not provide any benefits in terms of increased speech intelligibility in the background noise condition that was tested. Other volume control settings would need to be evaluated for their ability to produce higher speech intelligibility scores. Additionally, since an extensive electro-acoustic evaluation of the electronic earmuff was not performed as a part of this study, the exact cause of the reduced intelligibility scores at full volume remains unknown.

Correspondence: David C. Byrne, National Institute for Occupational Safety and Health, Robert A. Taft Laboratories, 4676 Columbia Parkway, Cincinnati, OH 45226, USA.
Tel. +1.513.533.8414. E-mail: dbyrne@cdc.gov

Key words: noise, hearing protection, speech intelligibility.

Contributions: DCB, initiated and designed the project, and acquired, analyzed, and interpreted the data; drafted the article; and takes overall responsibility for final approval of the version to be published; CVP, provided substantial contributions to conception/design of the project and data analysis and interpretation; revised the article critically for important intellectual content; and provides approval of the version to be published.

Conflicts of interests: the authors declare no potential conflicts of interests.

Received for publication: 14 September 2011.
Revision received: 3 January 2012.
Accepted for publication: 4 January 2012.

Audiology Research 2011;1:e5
doi:10.4081/audiores.2011.e5

Introduction

Hearing loss prevention is listed among the 21 Priority Research Areas, as described in NIOSH's National Occupational Research Agenda.[1] Occupational hearing loss is the most common occupational disease in the United States, and is specifically identified as a problem in the mining industries. Efforts to prevent occupational hearing loss appear to be hindered because the problem is insidious and occurs without causing pain in affected individuals. One consequence of noise-induced hearing loss is a reduced quality of life due to the inability to communicate with family, friends, and the general public. However, this normally occurs after the hearing loss has progressed significantly and the damage is irreversible. This problem can have other serious repercussions, considering that a study of older workers with disabilities indicated hearing loss as a risk factor for occupational injury.[2]

In January 1995, the Physical Agents Effects Branch, located in the National Institute for Occupational Safety and Health (NIOSH), Division of Biomedical and Behavioral Science, began collaboration on a project with the Mine Safety and Health Administration (MSHA) that was designed to determine the prevalence of hearing loss among miners. Two reports were forwarded to MSHA: one for coal miners,[3] and one for metal/non-metal miners.[4] After removing potentially invalid audiograms through a quality assurance process, the first report contained an analysis of 17,260 audiograms for 2871 coal miners, and the second report reviewed 22,488 audiograms on 5244 metal/non-metal miners. For comparison purposes, hearing thresholds were calculated for a similar-aged population of non-exposed individuals by using Annex A from ISO-1999.[5] The noise levels that would be predicted to cause the amount of hearing loss observed for the miners were also calculated from ISO-1999.

The results of these investigations showed that miners developed hearing loss much more quickly than those in the non-occupational noise-exposed database used by ISO-1999, and that the miners experienced a greater severity of hearing loss than would be expected for non-occupational noise-exposed persons of the same age and gender. Using hearing thresholds at 4000 Hz as an indicator, coal miners experienced hearing loss two-and-a-half to three times greater than would be expected for persons not exposed to occupational noise. At 50 years of age, 90% of the coal miners and 49% of the metal/non-metal miners were found to have a hearing impairment. In comparison, only 9% of the non-occupationally exposed group had a hearing impairment at the same age.

A new MSHA noise standard was published on September 13, 1999 and became effective on September 13, 2000.[6] This rule closely resembles the pre-existing Occupational Safety and Health Administration (OSHA) Occupational Noise Exposure Standard and Hearing Conservation Amendment (29 CFR 1910.95),[7] and replaced the different standards for occupational noise exposure in coal mines and in metal/non metal mines with a single new standard applicable to all mines. Based on information such as that described above, the previous standards (which had been promulgated in the early 1970s) were found to be inadequate to prevent the occurrence of noise-induced hearing loss among miners. Although the proposed noise exposure limits would not totally eliminate the risk of material impairment, it is expected to reduce by two-thirds the number of miners currently projected to suffer a material impairment of their hearing.

Both the OSHA and MSHA regulations require employers to use all feasible engineering and administrative controls to reduce noise exposures to acceptable levels. Engineering controls are the ideal solution for most circumstances, although they may be prohibitively expensive to implement. Administrative controls usually involve an attempt to lower daily noise exposures by rotating employees through *a less noisy* workstation; however, this practice is often not feasible due to production constraints and/or employee training requirements. If engineering and administrative noise controls fail to sufficiently reduce exposure levels, personal hearing protection (along with training on proper fitting, care, and use) must be provided to workers.

Over 50 manufacturers have developed and sold at least 241 different hearing protection devices.[8] Traditional hearing protectors usually take the form of either insert-type earplugs that are inserted into the ear and seal against the ear canal walls, or muff-type devices that seal against the head around the ear. There also are concha-seated protectors that provide an acoustic seal right at the entrance to the external ear canal. More technologically advanced *non-conventional* hearing protectors include active noise reduction protectors, level-dependent (non-linear) protectors, and sound restoration (amplification) devices. There is no single *best* type for all individuals or situations. However, some types are better than others for use in specific noise environments, for some work activities, or for some environmental conditions.

In March 2001, the MineEars electronic hearing protector received Schedule 2G approval (for use in gassy mines) from MSHA, and was added to their list of permissible equipment. Anecdotal reports indicate that these devices are being used in some underground coal mines. Similar commercially available devices have been around for many years, although none had previously applied for 2G approval. The primary difference between the MineEars earmuff and competing products is that MineEars employs an amplitude compression circuit while others are peak-clipping devices. Previous studies have discovered the potential for peak-clipping earmuffs to produce distortion and interfere with speech communication.[9] All manufacturers of electronic-amplifying hearing protectors claim that a user will be protected from hazardous noise levels while still being able to hear important sounds in their environment; however, the actual benefit that current state-of-the-art models provide under occupational noise conditions has not been quantified. The specific research question in this study was to determine whether speech understanding differs between the MineEars electronic earmuffs (with the volume control set at three different positions) and the Bilsom model 847 passive earmuffs as a function of speech-to-noise ratio, as measured in a constant speech-shaped background noise of 90 dB(A). This study was primarily a head-to-head comparison of the speech intelligibility scores obtained with the different devices. The purpose was not to identify potential electro-acoustical issues or provide specific solutions to any problems that might be discovered. Additional testing would be required before recommendations for improvements could be made.

Materials and Methods

Subjects

Subjects (9 male, 6 female; age range 21-60 years) were recruited from employees at NIOSH's Pittsburgh campus. The test protocol was reviewed and approved by the NIOSH Human Subjects Review Board (approval code HSRB 03-PRL-02XP). Participants were provided with the required assurances of confidentiality, and an informed consent form was signed prior to performing any screening/testing procedures. Subjects were excluded from this study if excessive cerumen was found (*i.e.* an amount that prohibited visualization of the eardrum). To eliminate the potentially confounding effects a pre-existing hearing loss might have on speech intelligibility measures, subjects were required to have normal hearing sensitivity (less than 20 dB HTL) at all audiometric test frequencies from 125 to 8000 Hz in each ear. The final eligibility requirement was that subjects demonstrate a normal signal-to-noise ratio (SNR) loss as determined by using version 1.3 of the QuickSIN™ speech-in-noise test.[10,11] On average, normal-hearing individuals require speech to be 2 dB louder than the background noise (*i.e.* a +2 dB signal-to-noise ratio) to identify 50% of the key words in sentences on the QuickSIN test. Subjects were considered eligible for participation if they had an SNR loss between 0-7 dB, as obtained from the average of two standard equivalent 6-sentence lists. Similar to the normal hearing sensitivity stipulation, this requirement was intended to control for extraneous variables by enrolling only subjects with normal speech recognition ability.

Acoustical test environments

Two different test rooms were used for this study, one for initial subject screening and the other for the actual experimental testing. Subject screening took place in a standard Industrial Acoustics Company (Bronx, New York, USA) double-wall audiometric test booth. Speech intelligibility testing was conducted in a specially designed reverberant test chamber (interior dimensions: 9' long by 11' wide by 8' high) located in the NIOSH-Pittsburgh Research Laboratory. This room has a solid sheet metal floor, and the walls and ceiling are lined with an acoustically reflective surface laminated to 3/4-inch high-density particle board. The chamber was outfitted with three (3) ElectroVoice T251+ loudspeakers driven by separate channels of a Sherbourn 5/1500A power amplifier. Each loudspeaker was oriented in a different plane (*i.e.* up/down; left/right; and front/back) for maximum sound dispersion throughout the space. Sound levels up to 115 dB SPL can be generated inside the chamber. This facility meets the requirements for sound field uniformity and non-directionality as specified in ANSI S12.42-1995 (R2004).[12] A random incidence (*i.e.* diffuse) sound field is produced, which provides no directional cues to the listener, and all sounds (*i.e.* both speech and noise) are always intermixed.

Instrumentation

Subject screening/qualification was conducted with the subject seated in the audiometric test booth. First, hearing thresholds were measured in each ear using a WelchAllyn GSI 61 Clinical Audiometer and TDH-50P earphones. Next, QuickSIN testing was performed with a Panasonic DVD-RA60 audio/video player connected to the *external A* input of the GSI 61 Clinical Audiometer. Tracks 3 and 4 (standard lists 1 and 2) of the QuickSIN compact disc were played binaurally through the TDH-50P earphones.

Speech intelligibility testing was conducted using the HINT - Hearing in Noise Test.[13] Version 2.0 of the HINT audio CD was presented via loudspeakers in the diffuse field chamber. Each channel of a Panasonic DVD-RA60 audio/video player was routed to a separate Tucker-Davis Technologies System II PA4 Programmable Attenuator, to enable independent level control of the speech and noise channels. The outputs of the attenuators were connected to a Rane SM 26B Splitter/Mixer, which combined the speech and noise signals and provided three monaural output signals. These three outputs were then directed to three channels of the power amplifier, which were then fed to the three loudspeakers in the test chamber. As mentioned above, due to the room construction and loudspeaker placement, sounds cannot be localized to any particular source/direction in this chamber, while precise control of the overall level and signal-to-noise ratio of the test materials is possible by adjusting the attenuators.

Earmuff test conditions

Two types of earmuffs were used in this study: the MineEars electronic hearing protector and the Bilsom model 847, which is a conventional passive-attenuation earmuff. The MineEars device consisted of a Bilsom 847 earmuff with the addition of an electronic amplification circuit in each earcup. Testing was conducted with the volume control on the electronic earmuffs set at three different positions: *off*, at the lowest *on* setting, and full-on (designated as OFF, LOW, and HIGH in subsequent sections).

Procedures

The Hearing in Noise Test (HINT) consists of 24 equivalent 10-sentence lists that may be presented with speech-shaped background noise. The test materials can also be presented as 12 equivalent 20-sentence lists. Each HINT sentence contains between three and seven words (the majority are either five or six words), and each list of 10 sentences has a total of 49-57 words. The sentences were recorded by a male talker, and each list is phonemically balanced. The HINT speech-shaped noise was suitable for this study because its low-frequency emphasis produces the masking effect most likely to be encountered in a work environment.

For this study, the HINT sentence tests were administered using a fixed-level protocol, as described in the HINT Audio CD Operating Instructions.[13] While wearing each pair of earmuffs, subjects were instructed to listen and repeat back (without the benefit of visual cues or other assistance) each sentence that they heard. For each sentence list, the number of words correctly repeated divided by the total words in each list (with the result multiplied by 100) indicated the percentage of correct words.

The lists of sentences were presented at levels of 85, 90, and 95 dB(A), which produced a -5, 0, and +5 dB signal-to-noise ratio when the background noise was set at 90 dB(A). The electronic earmuffs were treated as three separate devices, and were tested with the volume control *off*, at the lowest *on* setting, and at the maximum (highest) setting. The passive earmuff was considered as one additional device, so, in effect, four different earmuff conditions were tested. A total of 12 repetitions of the HINT (*i.e.* 12 equivalent 20-sentence lists) were administered to each subject (3 speech-to-noise ratios×4 earmuff conditions).

Experimental design and control

Repeated-measures ANOVA was selected to examine the effects of earmuff test condition, signal-to-noise condition, and the interaction between these two conditions. The sample size (*i.e.* number of subjects required) was determined by a power analysis calculation using the following information: within-subject standard deviation was estimated to be 5%; α was set at 0.05; and β was set at 0.2. This design allows a main effect size of 0.8 to be identified with 10 subjects, assuming a within-conditions correlation of 0.6. A large effect size was chosen because a large main effect would be necessary for the findings to have practical importance. Power calculations also showed that a sample size of 15 was required to detect a medium interaction effect. A medium effect size (0.25) was posited for the interaction because if there is an interaction effect, the expectation is that it would be smaller than the main effect, yet it would be important to detect it for a complete understanding of the variables involved. To account for both the main effect and interaction effect sizes, a sample size of 15 was selected for this study.

To guard against practice effects, each subject was trained on the HINT test procedure using the three 10-sentence practice lists (tracks 25-27) on the HINT Audio CD. This training was conducted once in a quiet environment at a 75 dB(A) presentation level, and twice at a 95 dB(A) presentation level with 90 dB(A) background noise (+5 dB signal-to-noise ratio) while wearing the two different earmuffs. To prevent sequencing effects, the measurement trials were administered according to an incomplete counterbalancing technique.

Performance-intensity (P-I) functions for sentences have steeper slopes, which means that better performance (*i.e.* a higher score) is obtained with smaller increases in the presentation (intensity) level, as compared to the P-I function for word lists.[14] This is partly due to contextual cues that make recognition of test materials, such as sentences more predictable (and consequently easier) at lower presentation levels or signal-to-noise ratios.[15] Although the test stimuli in the present experiment were presented in sentences, the analysis used to determine significant differences between two scores (from a clinical/practical perspective) was based on a binomial model, which is normally reserved for scoring tests of words in isolation.[16] The analysis was conducted this way because administering the HINT test according to the fixed-level procedure (rather than using the adaptive protocol to determine sentence speech reception thresholds) is similar to conducting a speech recognition test with single words. Specifically, the P-I function for the HINT sentences more closely approximates the P-I functions for lists of single words than the P-I functions typical of sentence materials.[17] The HINT sentences were judged to be written at a first or second grade reading level, which might also contribute to their similarity to single word speech tests.[18]

Results and Discussion

Speech intelligibility (HINT) scores for all 15 subjects are contained in Table 1. Each of the subjects' scores presented in Table 1 represent the average score obtained on two 10-sentence lists, which equates to approximately 100 test items.

Statistical analysis

The primary analysis examined the effects of muff test condition, signal-to-noise condition, and the interaction between muff condition and signal-to-noise condition using repeated-measures ANOVA with two within-subject factors (muff and signal-to-noise ratio). The analysis was conducted with the Geisser-Greenhouse conservative *F*-test in case the spherical assumption was violated. Despite the fact that multiple hypothesis tests were conducted, no adjustment to the level of significance (to account for a potential inflated Type I error risk) was made. The practical consequence of making a Type I error (*i.e.* finding a difference between muff types when no real difference exists) is much less severe than making a Type II error (*i.e.* reporting that there is no significant difference between muff types, and missing the opportunity to recommend a muff that produces better speech intelligibility).

As expected, there was a significant main effect ($F(2,28)=1014.50$,

Table 1. HINT test scores for all 15 subjects (SD, standard deviation).

	Passive Earmuff			Electronic OFF			Electronic LOW			Electronic HIGH		
	-5 dB	0 dB	+5 dB	-5 dB	0 dB	+5 dB	-5 dB	0 dB	+5 dB	-5 dB	0 dB	+5 dB
S-1	20%	64%	96%	15%	65%	87%	17%	56%	91%	10%	37%	79%
S-2	15%	91%	98%	22%	80%	91%	19%	89%	97%	16%	60%	88%
S-3	21%	71%	96%	2%	73%	95%	11%	65%	97%	5%	51%	96%
S-4	20%	78%	97%	33%	80%	94%	22%	72%	97%	3%	36%	86%
S-5	20%	78%	98%	11%	71%	98%	19%	81%	98%	6%	66%	95%
S-6	13%	65%	95%	14%	51%	88%	11%	64%	98%	8%	44%	73%
S-7	23%	82%	96%	21%	79%	99%	11%	48%	99%	3%	57%	86%
S-8	27%	71%	93%	16%	68%	97%	19%	81%	90%	4%	64%	91%
S-9	13%	75%	97%	5%	67%	95%	6%	63%	94%	0%	42%	77%
S-10	31%	80%	99%	11%	77%	99%	17%	87%	94%	18%	64%	99%
S-11	10%	60%	89%	7%	64%	99%	7%	41%	96%	1%	35%	85%
S-12	12%	65%	97%	7%	66%	92%	8%	74%	99%	12%	32%	84%
S-13	8%	38%	96%	7%	55%	85%	17%	44%	81%	0%	22%	63%
S-14	18%	83%	100%	11%	75%	97%	12%	70%	94%	6%	40%	94%
S-15	21%	80%	100%	15%	81%	98%	21%	63%	95%	8%	44%	86%
Mean	18%	72%	97%	13%	70%	94%	14%	67%	95%	7%	46%	85%
S.D.	6%	13%	3%	8%	9%	5%	5%	15%	5%	5%	13%	10%

MS_{error}= 96.29, P <0.0001) for signal-to-noise condition; the highest scores were obtained with a +5 dB signal-to-noise ratio, while the lowest scores were obtained with a -5 dB signal-to-noise ratio. A significant main effect also was found for the earmuff condition ($F(3,42)$=57.19, MS_{error}= 39.832, P <0.0001). Subjects obtained the highest speech intelligibility scores while wearing the passive earmuff, and the lowest scores when wearing the electronic earmuff at the HIGH volume control setting. Additionally, the interaction effect was significant ($F(6,84)$=6.94, MS_{error}= 38.304, P <0.0001), as illustrated by the non-parallel lines in Figure 1.

A full set of pairwise comparisons was conducted to fully investigate the interaction effect. Separate charts for each signal-to-noise ratio were constructed, where each cell mean from Table 1 was contrasted against each of the other cell means (Tables 2, 3 and 4). Table 2 (-5 dB signal-to-noise ratio) indicates that all muff conditions were different from each other, except for the electronic muff at the OFF and LOW volume control settings. For the other two signal-to-noise conditions (0 dB and +5 dB), the electronic muff with the volume control set to HIGH produced significantly different (worse) HINT scores than any of the other earmuff conditions (Tables 3 and 4).

Test administration considerations

To be certain that the statistical analysis of the data revealed findings that were practically and clinically significant from a test administration viewpoint, the test results were compared against the critical differences developed by Thornton and Raffin.[16] Their table of 95% critical differences (page 515, Table 4) for a 100-item test was used to determine whether the individual speech intelligibility scores were within (or outside of) calculated confidence intervals.

For a test score of 18%, the lower limit of the critical differences is 9%, and the upper limit is 29%.[16] This indicates a significant difference between the average test scores of the passive earmuff (18%) and the electronic earmuff on HIGH (7%) at the -5 dB signal-to-noise ratio. At the -5 dB signal-to-noise ratio, mean speech intelligibility scores for the electronic earmuff turned OFF (13%) and the electronic earmuff on LOW (14%) fell within the 9-29% range, and are not considered different from the passive earmuff score of 18%. Considering that the average HINT scores obtained at the -5 dB signal-to-noise ratio were 18% or less for all tests (*i.e.* all earmuff conditions yielded poor speech intelligibility scores), the statistically significant findings in Table 2 do not appear to have any practical significance.

A mean score of 46% was obtained at the 0 dB signal-to-noise ratio with the electronic earmuff on HIGH. The corresponding range of the 95% critical differences was 33-59%, according to the Thornton and Raffin critical differences table.[16] At the 0 dB signal-to-noise ratio, mean scores for the passive earmuff and the electronic earmuff at the OFF and LOW volume settings were 72%, 70%, and 67%, respectively, which were outside (above) the range of critical differences. This indicates that the electronic earmuff on HIGH may be considered significantly different (*i.e.* provide worse speech intelligibility scores) than the other earmuff test conditions at the 0 dB signal-to-noise ratio.

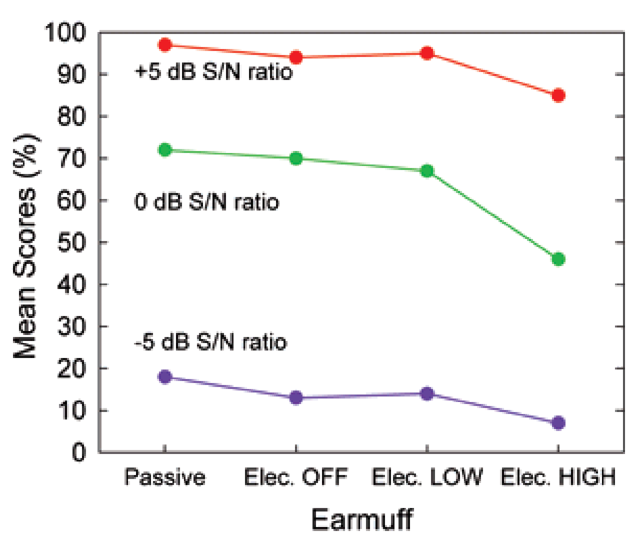

Figure 1. Graphic presentation of mean speech intelligibility scores for each earmuff condition with speech-to-noise ratio as the parameter.

Table 2. Pairwise comparisons with a -5 dB signal-to-noise ratio.

Source	Degrees of freedom	Sum of squares	Mean square	*F* value	Pr >*F*
Model	17	2169.663376	127.627257	4.75	<0.0001
Error	42	1128.181188	26.861457		
Corrected total	59	3297.844564			
Source	**Degrees of freedom**	**Type III SS**	**Mean square**	***F* value**	**Pr >*F***
Muff type	3	1043.899574	347.966525	12.95	<0.0001
Difference	**Estimate**	**Standard error**	**P value**	**95% Confidence Interval of Difference**	
Passive vs Elec OFF	5.075*	1.8925	0.010	(1.256, 8.894)	
Passive vs Elec LOW	3.854*	1.8925	0.0480	(0.035, 7.674)	
Passive vs Elec HIGH	11.583*	1.8925	<0.0001	(7.764, 15.403)	
Elec OFF vs Elec LOW	-1.221	1.8925	0.5224	(-5.040, 2.598)	
Elec OFF vs Elec HIGH	6.508*	1.8925	0.0013	(2.689, 10.327)	
Elec LOW vs Elec HIGH	7.729*	1.8925	0.0002	(3.910, 11.548)	

*As judged in an LSD test (DF=42, least significance difference=3.8192), the means differ significantly at the α=0.05 level.

Table 3. Pairwise comparisons with a 0 dB signal-to-noise ratio.

Source	Degrees of freedom	Sum of squares	Mean square	*F* value	Pr >*F*
Model	17	12472.84387	733.69670	11.07	<0.0001
Error	42	2784.43167	66.29599		
Corrected total	59	15257.27554			
Source	**Degrees of freedom**	**Type III SS**	**Mean Square**	***F* value**	**Pr >*F***
Muff type	3	6265.065089	2088.35503	31.50	<0.0001
Difference	**Estimate**	**Standard error**	**P value**	**95% Confidence Interval of Difference**	
Passive vs Elec OFF	2.064	2.9731	0.4914	(-3.936, 8.064)	
Passive vs Elec LOW	5.496	2.9731	0.0716	(-0.504, 11.496)	
Passive vs Elec HIGH	25.679*	2.9731	<0.0001	(19.679, 31.679)	
Elec OFF vs Elec LOW	3.432	2.9731	0.2549	(-2.568, 9.432)	
Elec OFF vs Elec HIGH	23.615*	2.9731	<0.0001	(17.615, 29.615)	
Elec LOW vs Elec HIGH	20.183*	2.9731	<0.0001	(14.183, 26.183)	

*As judged in an LSD test (DF=42, least significance difference=6), the means differ significantly at the α=0.05 level.

Table 4. Pairwise comparisons with a +5 dB signal-to-noise ratio.

Source	Degrees of freedom	Sum of squares	Mean square	*F* value	Pr >*F*
Model	17	2199.883417	129.404907	5.56	<0.0001
Error	42	977.880648	23.282873		
Corrected Total	59	3177.764065			
Source	**Degrees of freedom**	**Type III SS**	**Mean square**	***F* value**	**Pr >*F***
Muff type	3	1119.344792	373.114931	16.03	<0.0001
Difference	**Estimate**	**Standard error**	**P value**		**95% confidence interval of difference**
Passive vs Elec OFF	2.179	1.7619	0.2231		(-1.377, 5.734)
Passive vs Elec LOW	1.184	1.7619	0.3092		(-1.742, 5.369)
Passive vs Elec HIGH	11.122*	1.7619	<0.0001		(7.566, 1.678)
Elec OFF vs Elec LOW	-0.365	1.7619	0.8368		(-3.921, 3.191)
Elec OFF vs Elec HIGH	8.943*	1.7619	<0.0001		(5.387, 12.499)
Elec LOW vs Elec HIGH	9.308*	1.7619	<0.0001		(5.753, 12.864)

* As judged in an LSD test (DF=42, least significance difference=3.5557), the means differ significantly at the α=0.05 level.

The mean score for the electronic earmuff at the +5 dB signal-to-noise ratio with the volume control HIGH was 85%, and the range of critical differences around this score is 74-93%. Average scores for the other three earmuff conditions were 94% (electronic OFF), 95% (electronic LOW), and 97% (passive earmuff). Similar to the 0 dB signal-to-noise results, mean scores for the passive, electronic OFF, and electronic LOW earmuff test conditions were outside the range of critical differences for the electronic earmuff on HIGH. This indicates that the average electronic earmuff score with the volume control on HIGH (at the +5 dB signal-to-noise ratio) was significantly different (again, worse) from the other three earmuff test conditions.

Acoustical analysis

The primary finding in this study was that the electronic earmuff (at any of the three volume control settings used) did not provide a better speech intelligibility score than the otherwise identical passive earmuff in a 90 dB(A) background noise. Testing was conducted with the volume control knob at three positions that were easily defined and were intended to represent the settings that a typical wearer might use. Volume settings other than the ones used would be difficult to replicate and test, since the volume control knob on this particular set of earmuffs is continuously variable, and does not have discreet detents. Considering this, an assumption was made that a typical naïve user might set the control as high as it will go in an attempt to gain the most benefit from the electronic circuitry. However, human subject testing revealed that the highest volume control setting produced speech intelligibility scores no better (and in many cases worse) than those obtained with a pair of similar muffs without the amplification circuits.

Additional laboratory testing was conducted to investigate why the speech intelligibility test scores were different for the different earmuff conditions. An acoustical test fixture was used to measure the attenuation provided by the passive earmuff, as well as the attenuation/gain provided by the electronic earmuff at the same three volume control settings used during human subject testing. The test fixture was built by the French-German Research Institute of Saint-Louis (ISL), and consists of a solid block of material molded in the shape of a human head, with a Brüel and Kjær Type 4157 ear simulator and ½-inch microphone (Brüel and Kjær Type 4165) mounted inside (right ear only), and HEAD Acoustics® ITU-T P.57 Type 3.4 pinnas attached to the outside. The microphone output was connected to a Bruel and Kjaer Type 2807 power supply, and delivered to a National Instruments PCI-4462 24-bit data acquisition card for data storage/analysis.

A measurement of the 90 dB(A) noise used in this study taken with a precision measurement microphone system (G.R.A.S. model 40HF) at the head-center location in the diffuse field chamber is shown in Figure 2. A measurement with the acoustical test fixture and the resulting transfer function (*i.e.* difference between these two measurements) are also shown. The upper curve from Figure 2 is shown again in Figure 3, and the remaining curves in this figure indicate the levels present under each earmuff condition with only the 90 dB(A) speech-shaped background noise being played into the test chamber. The difference between the upper curve and any of the others represents the amount of attenuation provided by that particular earmuff condition.

Amplification of the background noise occurred with the electronic earmuffs at the LOW and HIGH volume control settings. As seen in Figure 3, the amount of amplification (gain) at each frequency is shown by the increase in level above the OFF condition.

The measurements in Figure 3 revealed that essentially the same background noise spectrum/levels were being heard with the passive earmuffs and the electronic earmuffs at the OFF and LOW volume control settings. Attenuation values for the electronic muff OFF and passive muff are similar but not identical, because the physical presence of the electronics/batteries inside the earcups causes the acoustical conditions to be different. A small amount of amplification is observed from 315 to 800 Hz and from 3150 to 5000 Hz with the electronic earmuff on LOW.

At least 25 dB of attenuation was observed at each frequency tested, with some attenuations approaching 45 dB for three of the four earmuff conditions (passive earmuffs, electronic earmuffs OFF, and electronic earmuffs LOW). With the electronic earmuff set to its full-on (HIGH) position, 20 dB or more attenuation was measured only at the lowest and highest frequencies. Twelve to 18 dB of attenuation was measured in the mid-frequencies (400 through 3150 Hz). A prominent peak occurred at 4000 Hz, where the least amount of attenuation (7 dB) was observed.

The peak at 4000 Hz with the electronic earmuffs on HIGH corresponds to subjective reports from the human listeners, who indicated that a high-pitched noise was evident with the electronic earmuffs on the highest volume control setting. Without further testing, it is unclear whether the lower speech intelligibility scores (for the electronic earmuff on HIGH) were due to distortion of the speech signal, a masking effect, or some other reason.

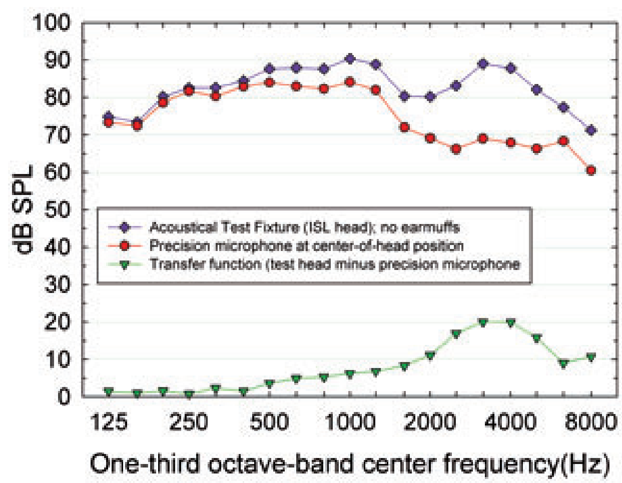

Figure 2. Measurements of HINT speech-shaped noise in the diffuse-field test chamber [overall level=90 dB(A)].

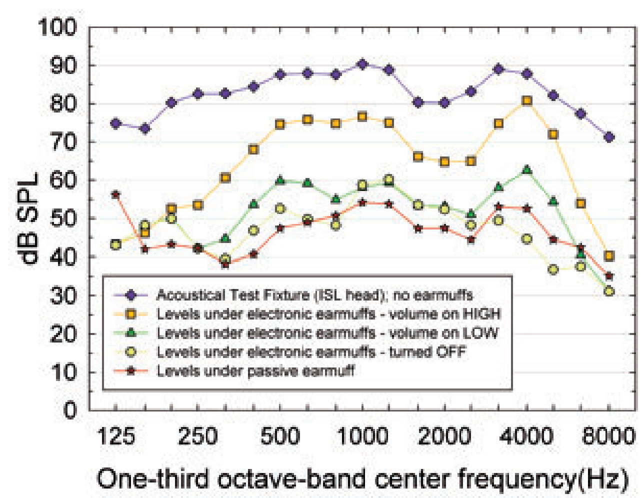

Figure 3. Measurements of HINT speech-shaped noise in the diffuse-field test chamber with the acoustical test fixture.

Speech intelligibility predictions

To determine whether the experimental results could have been predicted, the Speech Intelligibility Index, or SII (ANSI S3.5-1997, R2002)[19] was computed for each of the four earmuff conditions at all three signal-to-noise ratios (Table 5). The range of possible SII values is from zero to one. Calculated SII values will not necessarily match the score obtained from any particular speech intelligibility test; instead, the SII may be considered as the proportion of the total number of speech cues available to the listener. An SII value of 0.0 indicates that none (*i.e.* 0%) of the speech cues reach the listener, while an SII value of 1.0 indicates that all (*i.e.* 100%) of the speech cues are available to the listener. Similarly, a value of 0.5 represents half (50%) of the speech cues reaching the listener.

Pearson's product-moment correlation coefficients were computed to determine whether the calculated SII values corresponded with the actual speech intelligibility scores. Overall, the SII values and HINT scores obtained in this study were found to be highly correlated (r=0.977). The highest correlation was observed with the results from the 0 dB and +5 dB signal-to-noise ratios (r=0.981 and 0.962, respectively), while results from the -5 dB signal-to-noise ratio had a lower correlation coefficient (r=0.635). The measured speech intelligibility scores show good association with the calculated SII values (*i.e.* the rank ordering is highly correlated); however, the magnitude of each pair of SII values and test scores is obviously not the same (Table 5).

The HINT scores and the SII values can each be considered as a rater of speech intelligibility, and the intraclass correlation (ICC) can be used to quantify inter-rater reliability. In this case, the ICC was computed to obtain a better estimate of the agreement between the calculated SII values and the observed HINT scores. Similar to other measures of reliability or consistency, the ICC is a coefficient where the values range from 0.0 (consistency is totally absent) to 1.0 (complete consistency). For this situation, the ICC type (3,1) procedure was used, which means that the third ICC model and a one-time measurement was assumed. This procedure is a 2-way ANOVA design where the same two raters evaluated each case. This is the appropriate method to use considering that only two ratings were involved in the study (*i.e.* HINT scores and SII calculations) and generalization to other rating schemes was not intended. The ICC shows a moderate degree of agreement (ICC=0.67; 95% Confidence Interval: 0.19-0.89) between the two rating methods, which indicates a limited amount of predictability for this particular test scenario.

Study limitations

A comprehensive electro-acoustical evaluation of the electronic earmuffs was not conducted as a part of this study. Therefore, the exact cause of the poorer speech intelligibility test scores with the electronic earmuff on the HIGH volume setting is unknown. An evaluation of the amount of amplification delivered at various ranges of input levels/frequencies would provide some insight into this question. Another item to check is the phasing and level balance of the speaker output inside the electronic earmuffs.

Additionally, not all possible volume control settings were evaluated, so there is the potential for the electronic earmuff to allow wearers to perform better than the subjects involved in this study. Follow-up testing could be conducted with the volume knob at different positions between the lowest and highest settings in the event that the setting that would produce the highest speech intelligibility score was not evaluated in the current study.

Conclusions

Speech intelligibility testing was conducted in a constant speech-shaped background noise of 90 dB(A) while subjects wore conventional passive-attenuation earmuffs and identical devices with electronic amplification circuitry added. As anticipated, speech understanding was found to differ as a function of the speech presentation level. At a -5 dB signal-to-noise ratio, poor speech intelligibility results (average test scores less than 18%) were obtained for all earmuff test conditions. Average speech intelligibility scores ranged from 46% to 72% at a 0 dB signal-to-noise ratio. Average scores for all earmuff test conditions were 85% or higher at a +5 dB signal-to-noise ratio.

The most practically relevant finding was that for each of the three signal-to-noise test conditions, better test scores were obtained with the set of passive (non-electronic) earmuffs or the electronic amplifying earmuffs with the user-adjustable volume control set to OFF or LOW, as compared to the electronic earmuffs set to HIGH. This finding was consistent with predicted values of the Speech Intelligibility Index. Higher noise levels were found in acoustical measurements taken underneath the electronic earmuffs at the maximum volume setting, although exactly how this related to the lower test scores was not determined.

From a speech intelligibility perspective, this particular type of electronic earmuff (with the volume control full-on) would not be a desirable choice for use in a constant background noise of 90 dB(A), although it may prove to be suitable in other ambient noise conditions.

Table 5. SII calculations using the one-third octave-band procedure (HINT speech intelligibility scores for the same test conditions from Table 1 are shown in parenthesis).

	Passive earmuff	Electronic off	Electronic low	Electronic high
-5 dB S/N ratio	0.27 (18%)	0.27 (13%)	0.25 (14%)	0.25 (7%)
0 dB S/N ratio	0.43 (72%)	0.43 (70%)	0.42 (67%)	0.40 (46%)
+5 dB S/N ratio	0.60 (97%)	0.60 (94%)	0.59 (95%)	0.55 (85%)

References

1. National Occupational Research Agenda. Cincinnati (OH): National Institute for Occupational Safety and Health; DHHS (NIOSH) Pub. No. 96-115; 1996.
2. Zwerling C, Whitten PS, Davis CS, Sprince NL. Occupational injuries among workers with disabilities: the National Health Interview Survey, 1985-1994. J Am Med Assn 1997;278:2163-66.
3. Franks JR. Analysis of audiograms for a large cohort of noise-exposed miners. Cincinnati (OH): National Institute for Occupational Safety and Health; 1996. pp 1-7.
4. Franks JR. Prevalence of hearing loss for noise-exposed metal/non-metal miners. Cincinnati (OH): National Institute for Occupational Safety and Health; October 7, 1997. pp 1-5.
5. International Standards Organization. Acoustics - determination of occupational noise exposure and estimation of noise-induced hearing impairment [ISO 1999.2]. Geneva: Switzerland; 1989.
6. Health Standards for Occupational Noise Exposure, 30 C.F.R. Parts 56, 57, 62, 70, and 71, 1999.
7. Occupational Noise Exposure: Hearing Conservation Amendment, 29 CFR Sect. 1910.95, 1983.
8. The NIOSH compendium of hearing protective devices. Cincinnati, OH: National Institute for Occupational Safety and Health; DHHS (NIOSH) Pub N. 95-105; 1994.
9. Casali JG, Berger EH. Technology advancements in hearing protection circa 1995: active noise reduction, frequency/amplitude-sensitivity, and uniform attenuation. Am Ind Hyg Assn J 1996;57:175-85.

10. The QuickSin Speech-in-Noise-Test Version 1.3 [audio CD]. Etymotic Research, Elk Grove Village, IL, 2001.
11. Killion MC, Niquette PA, Gudmundsen GI, Revit LJ, Banerjee S. Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. J Acoust Soc Am 2004;116:2395-405.
12. American National Standards Institute. American National Standard Microphone-in-Real-Ear and Acoustic Test Fixture Methods for the Measurement of Insertion Loss of Circumaural Hearing Protection Devices [ANSI S12.42-1995 (R2004)]. New York; 2004.
13. Hearing in Noise Test - HINT version 2.0 [audio CD]. Eden Prairie, MN: Maico Diagnostics and House Ear Institute; 2004.
14. O'Neill JJ. Recognition of intelligibility test materials in context and isolation. J Speech Hear Dis 1957;22:87-90.
15. Wilson RH, McArdle RM. Speech signals used to evaluate functional status of the auditory system. J Rehab Res Dev 2005;42:79-94.
16. Thornton AR, Raffin MJM. Speech-discrimination scores modeled as a binomial variable. J Speech Hear Res 1978;21:507-18.
17. Nilsson MN, Soli SD, Sumida A. Development of norms and percent intelligibility functions for the HINT. Los Angeles, CA: House Ear Institute; 1995.
18. Nilsson M, Soli SD, Sullivan JA. Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise. J Acoust Soc Am 1994;95:1085-99.
19. American National Standards Institute. American National Standard Methods for Calculation of the Speech Intelligibility Index [ANSI S3.5-1997 (R2002)]. New York; 2002.