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

Anthony J. Brammer,[b)] Gongqiang Yu, Eric R. Bernstein, Martin G. Cherniack, and Donald R. Peterson
Ergonomic Technology Center, University of Connecticut Health Center, Farmington Avenue, Farmington, Connecticut 06030-2017

Jennifer B. Tufts
Department of Communication Sciences, University of Connecticut, Storrs, Connecticut 06269

(Received 25 September 2013; revised 23 April 2014; accepted 30 May 2014)

An adaptive, delayless, subband feed-forward control structure is employed to improve the speech signal-to-noise ratio (SNR) in the communication channel of a circumaural headset/hearing protector (HPD) from 90 Hz to 11.3 kHz, and to provide active noise control (ANC) from 50 to 800 Hz to complement the passive attenuation of the HPD. The task involves optimizing the speech SNR for each communication channel subband, subject to limiting the maximum sound level at the ear, maintaining a speech SNR preferred by users, and reducing large inter-band gain differences to improve speech quality. The performance of a proof-of-concept device has been evaluated in a pseudo-diffuse sound field when worn by human subjects under conditions of environmental noise and speech that do not pose a risk to hearing, and by simulation for other conditions. For the environmental noises employed in this study, subband speech SNR control combined with subband ANC produced greater improvement in word scores than subband ANC alone, and improved the consistency of word scores across subjects. The simulation employed a subject-specific linear model, and predicted that word scores are maintained in excess of 90% for sound levels outside the HPD of up to ∼115 dBA.
 [http://dx.doi.org/10.1121/1.4883385]

PACS number(s): 43.50.Hg, 43.66.Vt, 43.50.Ki, 43.72.Dv [CYE] Pages: 671–681

I. INTRODUCTION

Understanding speech when attempting to communicate in a noisy environment becomes progressively more difficult as the intensity of noise increases. There have been numerous studies that have addressed this issue (see, for example, Suter, 1992). Most solutions have attempted to increase the speech signal-to-noise ratio (SNR) either directly or by influencing the upward spread of masking (Rankovic *et al.*, 1992; Schwander and Levitt, 1987; Shields and Campbell, 2001; van Dijkhuizen *et al.*, 1991). Understanding speech when wearing a communication headset/hearing protector (HPD) in a noisy environment introduces additional considerations (Abel *et al.*, 2011; Bockstael *et al.*, 2011; Giguère *et al.*, 2011). There are two typical but fundamentally different situations. The first involves face-to-face communication, while the second involves communicating with a remote person by means of an electronic communication channel, whereby speech is reproduced by a miniature loudspeaker mounted within the HPD. The latter problem is the subject of this contribution.

When listening in a noisy environment, the most common solution to the problem of understanding speech in the communication channel is simply to increase the speech SNR by adjusting the volume control. This strategy will have limited consequences until the combined environmental and communication sound pressures at the ear approach the limits for safe exposure. While there have been many guidelines for occupational exposure to noise, the goal adopted here is not to exceed an 8-h energy equivalent sound level of 85 dBA (National Institute for Occupational Safety and Health, 1998). A recent review of studies measuring the speech SNR chosen by users of HPDs in environmental noise found the preferred speech SNR to be on the order of 12 to 15 dB (Giguère and Dajani, 2009; Giguère *et al.*, 2012). Thus, the HPD will need to attenuate the environmental noise so that the exposure attributable to noise and communication, the latter being set by the speech SNR established by the user and the fraction of time communication occurs, falls below the chosen exposure guideline. With the noise reduction rating of most HPDs less than 15 dB in the field (Berger, 2003), it is evident that the strategy of the user adjusting the volume control can lead to overexposure to noise in some situations. Clearly, a different control strategy is required in these circumstances.

The problem does not seem to have attracted much attention in the literature. Some commercial electronic devices have introduced amplification of sounds outside the HPD at low sound levels, with amplitude compression or limiters to restrict the maximum sound level at the ear (Casali, 2010). This strategy may improve the audibility of sounds outside the HPD but does not affect the audibility of sounds in the communication channel. There have been attempts to improve speech understanding in electronic HPDs by introducing active noise control (ANC), that is, electronic

[a)]An earlier version of part of this work was presented at the 21st International Congress on Acoustics.

[b)]Author to whom correspondence should be addressed. Electronic mail: brammer@uchc.edu

cancellation of noise by phase opposition (Brammer *et al.*, 2008; Simshauser *et al.*, 1956; Wheeler, 1986), which can improve the speech SNR at low frequencies. However, the compromises in earphone performance needed to implement a stable analog feedback control system as well as reproduce speech restrict the improvement obtainable. This has led some commercial devices to employ two miniature loudspeakers—one for active control and another to reproduce speech. An alternative feed-forward control structure, which overcomes this limitation, has shown promise (Brammer *et al.*, 2005), as have devices employing both feed-forward and feedback control structures (Rafaely and Jones, 2002; Ray *et al.*, 2006). However, none of these adaptively control the speech signal level.

The desirability of improving speech SNR in limited frequency ranges, for example, octave bands, has been demonstrated when the noise is predominantly low frequency and/or of restricted bandwidth (Rankovic *et al.*, 1992; van Dijkhuizen *et al.*, 1991). The approach has been developed for assisting face-to-face communication by persons wearing hearing aids (Shields and Campbell, 2001), and hearing aid algorithms for this purpose have been evaluated for use in electronic HPDs (Chung, 2007). In this paper, an automated method for optimizing the speech SNR is described for situations in which an HPD containing a communication channel is operated in environmental noise. A signal processing strategy is employed in which the total bandwidths of the environmental noise and communication signal are divided into restricted bands of contiguous frequencies ("subbands"), and processed separately. Subbands have been described in the literature for ANC but do not appear to have been applied to HPDs or communication headsets (Morgan and Thi, 1995; Shields and Campbell, 2001; Toner and Campbell, 1993), where the short time available to generate the cancellation signal presents challenges to a real-world implementation. In principle, the subbands employed for processing environmental noise and speech may be identical or different. In practice, because the anticipated frequency ranges of environmental noise at the ear (effectively reduced by the passive attenuation of the HPD) and speech differ, subbands with different bandwidths are used for this application. The task becomes one of optimizing the speech SNR for each communication channel subband, subject to the constraints of noise exposure and the anticipated speech SNR preferred by the user, while maintaining speech quality.

The purpose of the study was to explore the performance of a proof-of-concept device that could be worn by human subjects to evaluate the potential of the proposed signal processing strategy. The study builds on the system development and related work conducted by Bernstein (2013), and Bernstein *et al.* (2010, 2013), details of which will be published elsewhere.

An implementation is described that is constructed from the mechanical components of a commercially produced circumaural HPD and from electronic components custom developed for the application. Results are presented when the proof-of-concept device was worn by human subjects in an anechoic chamber, where environmental noise was reproduced. Noises were from a local paper-making factory and the commander's position of a military tank. Speech understanding in the communication channel was assessed psychophysically using the Modified Rhyme Test (MRT) (ANSI, 1989). The sound levels of the environmental noise and speech were constrained so as not to pose a hazard to the subjects' hearing.

The performance of the proof-of-concept device was evaluated in simulation for other noise exposure conditions. The simulation is first validated by comparing predictions of word scores with those obtained by a human subject under identical listening conditions when the sound level of environmental noise outside the HPD was ∼90 dBA. The simulations are used to predict the performance of the proof-of-concept device in environmental noise at sound levels outside the HPD of up to ∼120 dBA.

II. DEVICE, METHODS, AND SUBJECTS

A. Algorithm

A simplified block diagram of the signal processing is shown in Fig. 1. The transducers, microphones R and E, and miniature loudspeaker S, are mounted on circumaural earmuffs, one of which is sketched in cross-section at the top of

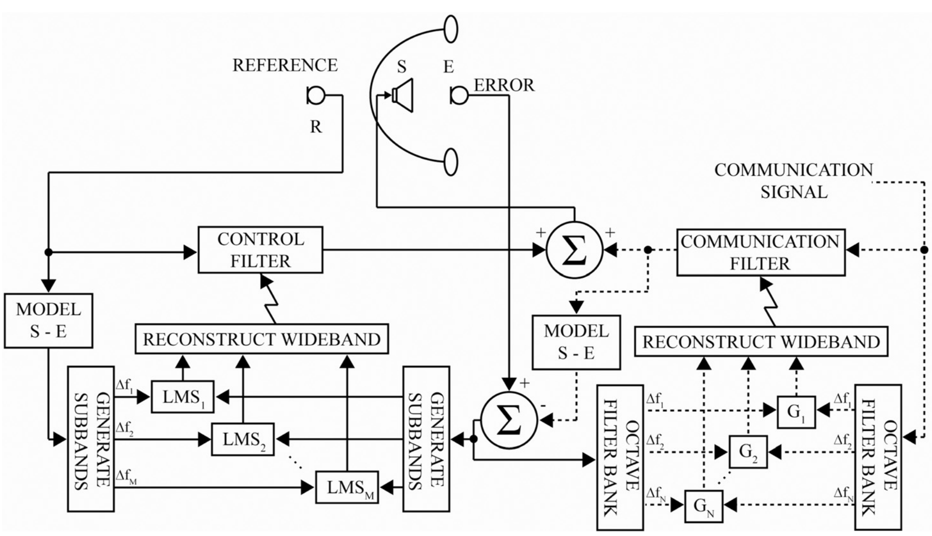

FIG. 1. Block diagram of algorithm with sketch of ear cup, reference (R) and error (E) microphones, and secondary source loudspeaker (S). The signal paths for ANC are shown by the solid lines and for communication control by the dashed lines.

the diagram. Microphone R is positioned on the outer surface of the ear cup approximately on the inter-aural axis. The miniature loudspeaker is similarly positioned within the ear cup, with the plane formed by its diaphragm normal to this axis, while microphone E is attached to the front surface of the loudspeaker housing.

The signal processing elements can be separated conceptually into two parts: The first, a subband ANC subsystem, shown to the left of the diagram with interconnections drawn as solid lines, and the second, a subband communication signal processor, shown to the right of the diagram with interconnections drawn as dashed lines. Both subsystems employ a feed-forward control structure, with outputs derived from inputs that have been passed through adaptive filters (labeled "control filter" and "communication filter" in Fig. 1). The input for the noise controller is derived from the miniature microphone attached to the outer surface of the ear cup, R, and that for the communication signal processor from an electronic signal characterizing a remote sound source (e.g., a remote talker). The outputs of the two subsystems are summed (indicated by Σs) and fed to the loudspeaker, S, which, as already noted, is located within the ear cup. The two subsystems will be described separately.

The ANC subsystem possesses a so-called "delayless" subband structure (Morgan and Thi, 1995), in which the time delay imposed by signal processing is minimized by feeding the control signal from the microphone sensing the environmental noise as directly as possible to the secondary source producing the controlling sound, S. In the present algorithm, the environmental noise is sensed by a reference microphone, R, processed by a control filter, which has been adjusted so as to minimize the residual sound in the volume enclosed by the ear cup and external ear, and thence to S after summation with the communication signal. The adjustment of the control filter is performed in parallel with this operation as is customary in digital feed-forward ANC, in order to minimize the time delay from input to output (Kuo and Morgan, 1996). However, the adjustment is performed after splitting the spectrum of the environmental noise into subbands (Lee *et al.*, 2009). Each of the M subbands, with frequency bandwidths $\Delta f_1, \Delta f_2, \ldots, \Delta f_M$, form independent, adaptive, digital active noise controllers, which compare the signals derived from R and the microphone sensing the residual sound or "error" in the volume between the ear cup and external ear, E, in order to optimize the control filter characteristics within their respective frequency bands. In the implementation described in this paper, optimization employed the least mean squares (LMS) algorithm to adapt the filter coefficients of individual subbands (Kuo and Morgan, 1996). The "filtered-*X*" LMS algorithm used requires the reference input to the subband generator to be filtered by a model of the error path, that is, by a filter with transfer function that represents the conversion of electrical signals to sound by S, the transmission of sound from S to E, and the conversion of sound to an electrical signal by E—shown as "model S–E" in the block diagram. In order to implement the algorithm, error path filter coefficients are determined when an HPD is worn by a human subject (see Sec. II C).

The subband coefficients are finally combined to reconstruct the complete bandwidth of the control filter ("reconstruct wideband"), and the finite impulse response (FIR) filter coefficients so obtained are used to generate the control filter. In addition, it is advantageous, but not essential, for convergence to remove the communication sounds reaching E (Brammer and Pan, 1998), which is done here by passing the communication signal output through a filter modeling the error path before subtracting it from the error signal. The ANC subsystem serves to decorrelate sound at the reference microphone from that at the error microphone, so the presence of an additional uncorrelated sound at the error microphone (e.g., speech) may perturb the rate of convergence.

The communication signal processor, shown to the right of the block diagram, shares some features with the ANC subsystem. As already noted, it has a feed-forward delayless subband control structure. The processor requires a communication signal input and a measure of the residual environmental noise at the ear, which is obtained from the error signal after removal of communication sounds at the ear. In contrast to the noise controller, which employs so-called linear subbands (Lee *et al.*, 2009), octave frequency subbands are used for the communication signal processor, as these are commonly employed to predict speech intelligibility in environmental noise (Yu *et al.*, 2010). By transforming both the environmental noise at the ear and communication signal into octave frequency bands, the communication SNR can be determined in each of the N subbands ($SNR_1, SNR_2, \ldots, SNR_i, \ldots, SNR_N$). An optimization routine is used to compute the amplification to be applied to each subband ($G_1, G_2, \ldots, G_N$), with constraints that the maximum overall sound level at the ear does not exceed a prescribed maximum set for hearing conservation, the individual SNR_i do not exceed a preset target, and large differences in gain between subbands are reduced to improve speech quality. The values of the SNR_i are initially taken to be the preferred overall SNR chosen by users when manually adjusting the communication channel gain of active HPDs and communication headsets (Giguère and Dajani, 2009; Giguère *et al.*, 2012). For each subband, the sound level at the ear obtained from the error signal at E is used to set a maximum sound level for hearing conservation. If this is reached, the gain of the subband and the corresponding value of SNR_i are reduced to minimize the increase in sound level. When the gains of all subbands have been determined, the gain of the subband with least gain is increased to the mean gain of the other subbands to avoid excessive spectral distortion of the communication signal, and any additional gain readjustments made to comply with the prescribed maximum sound level at the ear. Finally, as before, the individual octave band gains are combined to create the complete bandwidth of the communication filter, which is also implemented as an FIR filter.

B. Device

A proof-of-concept device has been constructed from readily available components to evaluate the performance of the proposed signal processing. The primary components are earmuffs, headbands, miniature loudspeakers, and microphones. Single core, floating point, digital signal processors (DSPs) are used to implement the algorithm (TMS320C6713, Texas Instruments, Dallas, TX), with in-house custom-designed analog interface circuits.

For convenience, and to avoid the need to design and fabricate the mechanical components of a circumaural HPD, commercial passive HPDs were evaluated for their suitability for conversion into an active device (Bernstein *et al.*, 2010; Bernstein, 2013). The evaluation focused on the passive noise reduction (PNR) of the earmuff and the predicted maximum active noise reduction (ANR) obtainable. Both were measured on human subjects in a white noise diffuse sound field within a reverberation room. A custom-built probe microphone enabled the sound pressure level (SPL) at the eardrum to be reconstructed with a precision of ± 2 dB up to a frequency of 6 kHz (Brammer *et al.*, 2009). The probe microphone was used to determine the PNR, by sensing the difference in sound pressure at the eardrum when the HPD was worn and when it was not worn.

The maximum ANR attainable was estimated by measuring the coherence between the sound field outside the earmuff and underneath the ear cup at the entrance to the ear canal when the HPD was worn (Kuo and Morgan, 1996). Miniature microphones were selected as sensors for the measurements so that their locations could be changed, and were attached to the outer surface of the ear cup and to an ear plug at the entrance to the ear canal. The location of the microphone on the outer surface of the ear cup was adjusted to optimize the predicted ANR.

The results of these measurements are summarized for the selected earmuff in Fig. 2, and are adapted from the measurements of Bernstein *et al.* (2010) and Bernstein (2013). Reference to this diagram confirms that the PNR increases with frequency up to a maximum of approximately 40 dB, as expected (Shaw and Thiessen, 1958), while the predicted ANR is close to 20 dB at low frequencies, and decreases to about 10 dB at 500 Hz, and to no more than 2 to 3 dB at 1 kHz. The combined noise reduction is close to 40 dB at frequencies from 100 Hz to 1 kHz, and is somewhat less than 40 dB at higher frequencies (not shown). Thus, for the purposes of the present work, the selected earmuff (Optime 98, Peltor, St. Paul, MN) provided adequate attenuation of environmental noise at all speech frequencies.

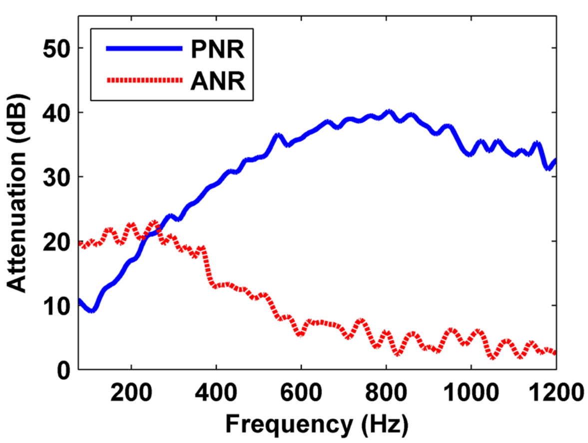

FIG. 2. (Color online) Measured PNR and predicted ANR of earmuff.

Miniature loudspeakers were then mounted sequentially within the selected ear cup and evaluated for frequency response and sensitivity. The transducer selected was from a commercial headphone (HD580, Sennheiser Electronic Corp., Old Lyme, CT). Its frequency response at speech frequencies was equalized at the error microphone. The microphones for the proof-of-concept device were selected for small size, adequate sensitivity, and flat frequency response from 50 Hz to 10 kHz (WM-61A, Panasonic, Secausus, NJ). Analog circuits for preamplifiers, power amplifiers, and anti-aliasing, reconstruction, and bandpass filters were designed and assembled in our laboratories. One DSP and associated interface circuitry was dedicated to operate the subband system for each ear cup.

C. Simulation

A simulation of the proof-of-concept device has been undertaken in MATLAB, both to assist its development and to provide a means for evaluating its performance in situations in which it is ethically unacceptable to expose human subjects (Bernstein *et al.*, 2013). The simulation involves all electronic elements (i.e., filters, amplifiers, analog to digital converters), acoustic elements (i.e., acoustic filters), and electro-acoustic elements (microphones and loudspeaker). Elements are modeled by the transfer function from the (electrical) input to the (electrical) output. For the acoustic elements, sound pressure transformations from the "input" to "output" are adjusted for the microphone characteristics. The transfer functions of the acoustic, and combined acoustic and electro-acoustic, elements were measured directly on the physical proof-of-concept device when worn by a human subject, and so embody the geometry and configuration of the components selected as influenced by the fit of the earmuff to the head. An example of such a transfer function is the error path model from S–E in Fig. 1. By using transfer functions for the acoustic and electro-acoustic elements determined on a real earmuff worn by a subject, together with measured transfer functions for the purely electronic elements, the combination provides a credible simulation of the device. When an algorithm implementing the signal processing summarized in Fig. 1 is introduced, the approach permits the effects of changes to the elements and algorithm on the performance of the proof-of-concept to be readily assessed. The influence of environmental noise on the active noise controller and its ability to maintain acceptable sound levels at the ear can thus be simulated by introducing environmental noise at R and estimating the sound level at the ear. Speech intelligibility in environmental noise is evaluated by introducing, in addition, speech into the communication channel and predicting the intelligibility from a speech model, such as the speech-stimulus Speech Transmission Index (STI) (Payton and Braida, 1999).

The simulation as previously reported, however, omits several wave properties of sound fields in the vicinity of the head. In the absence of other limitations, the coherence between the environmental noise at R and S (the former as reproduced by S) will restrict the ANR that can be achieved in the real world. This degradation in performance cannot be eliminated by the control algorithm. The decreasing magnitude of the coherence as frequency increases will progressively reduce the ANR for a given distance from R, as can be seen from Fig. 2. The spatial coherence of the sound field also decreases with distance. Thus the ANR at the entrance to the ear canal will differ substantially from that recorded at E. A correction is therefore introduced into the simulation in the form of a frequency-dependent adjustment to the ANR. An approximate magnitude for the correction was obtained from measurements conducted on an ANC earmuff mounted on a flat-plate coupler, by comparing the ANR at E, which was attached to the miniature loudspeaker forming the secondary source, with that at a microphone positioned at the entrance to the "ear canal" (Pan *et al.*, 1995). It is known that sound propagation from the entrance to the ear canal to the eardrum can be described by a plane wave mode at frequencies below about 4 kHz (Stinson and Daigle, 2005), which is above the maximum frequency of ANC. Hence the sound pressure recorded at the entrance to the ear canal will reflect that reaching the eardrum for the frequencies of interest.

A discrepancy in speech SNR arises from the differences between sound pressures at E and at the eardrum. This is a consequence of the complex interaction between a circumaural earmuff containing a miniature loudspeaker and the external ear (Shaw, 1997). Moreover, the sound pressure transformation from the free field to the eardrum (in the absence of an earmuff) is known to depend on the anatomy of the external ear and so differs across subjects (Hammershøi and Møller, 1996). Hence the interaction between the earmuff, loudspeaker, and ear can be expected to be subject dependent. A discrepancy of approximately 6 dB was observed across subjects who participated in the current study, which was attributed to these individual differences.

Accordingly, to apply the simulation, the sound pressure transformation from E to the eardrum was measured on a human subject wearing the proof-of-concept device. White noise was produced by the secondary source S and the transfer function was recorded between the error microphone and the probe microphone at the eardrum by the DSP. A Least Squares algorithm was developed in MATLAB to adapt the coefficients of an FIR filter for system identification using the method described by Bernstein *et al.* (2013). The resulting sound pressure transformation was applied to the speech signal at E to permit comparison with the predicted environmental noise at the eardrum. This method of system identification was also used to establish the error-path response for each subject, shown as "model S–E" in Fig. 1, by determining the transfer function from the secondary source S to microphone E, as well as to equalize the frequency response of S for speech reproduction.

The intelligibility of speech is predicted from the signals at the eardrum using a speech-stimulus model for the STI that has been summarized elsewhere (Yu *et al.*, 2010; Brammer *et al.*, 2011). A relationship between the STI and the mean word scores recorded by subjects with normal hearing listening to speech presented in broadband speech-spectrum shaped, white, and -3 dB/octave noise has been obtained using the MRT (Yu *et al.*, 2010). This relationship enables a word score to be predicted from the STI, and hence enables the results of the simulation to be compared directly with psychoacoustic measurements on human subjects. As the relationship between mean word scores and the STI was derived from measurements of sound pressure at the center-head position (in the absence of the subject), a mean sound pressure transformation from the eardrum to the center-head position was also employed to confirm the results of the simulation (Shaw and Vaillancourt, 1985). While this was considered acceptable for ratios of sound pressures, or differences in SPL such as the speech SNR, it was not employed to estimate sound levels for hearing conservation at the center-head position from those recorded at the ear (e.g., by microphone E), because of individual differences in the sound pressure transformation.

D. Psychoacoustic assessment of speech intelligibility

The MRT was used to establish the intelligibility of similar sounding words in environmental noise (ANSI, 1989; House *et al.*, 1965). Stimuli were produced, and the subject's responses recorded, under computer control. The procedures were coded in MATLAB, and controlled the instruments used to perform the test including reproducing speech (Tucker-Davis Technologies System 3). The test words and carrier phrase were commercial recordings with a male talker (Auditec Inc., St. Louis, MO).

Environmental noise was reproduced by four loudspeaker towers positioned at the corners of a distorted horizontal "square." Each tower consisted of a JBL SRX715 woofer and tweeter and a SRX718S sub-woofer (JBL, Stamford, CT). Transducers, audio signal processors (SP2030, Yamaha Corp. of America, Buena Park, CA), and amplifiers (2B, 3B, and 4B, Bryston, Peterborough, ON) with sufficient power to reproduce the sound levels required for the study in our anechoic chamber were integrated with computer control of all signals using a digital sound card (AES16, Lynx Studio Technology Inc., Costa Mesa, CA). A pseudo-diffuse sound field was generated in the horizontal plane at the position of the subject's head from two- or four-channel sound source recordings, by delaying the output from one loudspeaker tower relative to another and by introducing reverberation. The signal manipulations were performed within the psychoacoustic constraint that a subject experienced only one sound image. The extensive signal processing, together with the electrical power required (8 kW), resulted in a sound source with "flat" frequency response from 40 Hz to 10 kHz (± 3 dB). The SPL of environmental noise so produced was sensed at each eardrum by a probe microphone (Brammer *et al.*, 2009).

Measurements were conducted when the proof-of-concept device was worn by subjects who were subjected to noise recorded either near dryers in a local paper-making factory or at the commander's position of a Leopard military tank. One-third octave-band spectra of the noise sources are shown in Fig. 3. Relative levels of the sources are shown, as the overall sound levels of the environmental noises, and speech, were adjusted so as not to pose a hazard to the subjects' hearing.

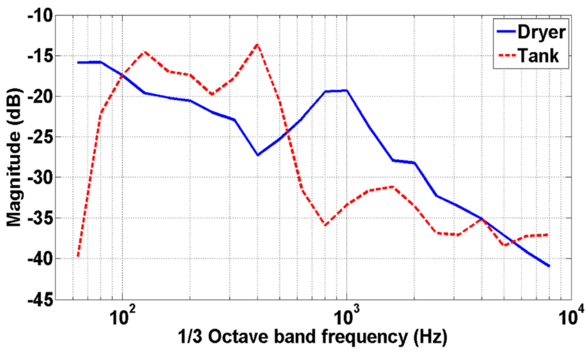

FIG. 3. (Color online) Relative one-third octave-band SPLs of the noise of an industrial dryer, and a military tank (at the commander's position).

The MRT test words were reproduced by the miniature loudspeaker within the proof-of-concept device, which could be operated as a conventional passive circumaural HPD, or as an active HPD with subband ANC either with, or without, subband communication channel gain processing. Three measurement conditions were evaluated: Conventional passive HPD; active HPD with subband ANC and no communication signal processing; and active HPD with subband ANC and communication channel gain processing. To perform measurements, subjects sat with a computer-controlled touch screen at a convenient height in front of them and initiated each test by pressing a "PLAY" button. One of six words displayed on the touch screen was randomly presented within a carrier phrase, e.g., "Circle the… (insert test word)… again." Subjects were instructed to choose one word by touching the screen, and initiate the next trial when ready. The procedure was fully automated. There were 50 trials using randomly selected, standardized, word ensembles in each test (ANSI S3.2-1989), which was repeated twice, so that there were 150 trials from which a word recognition score was derived for the preset speech SNR.

E. Subjects

Six healthy volunteers with normal hearing (mean age of 29.5 ± 8.5 yrs) participated in the experiments, and were paid for their time (4 male and 2 female). Hearing thresholds were determined at study induction and were better than 20 dB hearing level (HL) re ANSI S3.6-2010 (ANSI, 2010) at audiometric frequencies from 500 Hz to 8 kHz. In addition, the difference in HL between a subject's left and right ears was less than 10 dB. All were native speakers of American English. Subjects gave their informed consent to participate in the study, which was conducted according to the provisions of the University of Connecticut Health Center's ethics review board.

III. RESULTS

Measurements were conducted when the proof-of-concept device was worn by subjects in sound levels of environmental noise and speech that were adjusted so as not to pose a hazard to hearing. The performance of the proof-of-concept device was evaluated separately in simulation for other sound levels of the noise sources and speech.

A. Word recognition scores

The performance of the proof-of-concept device has been compared under three operating conditions when worn by subjects. Each used the same ear cups, headbands, and ear cushions. The results are summarized for the six subjects by the mean MRT scores [± 1 standard deviation (SD)], and are shown in Fig. 4. All word recognition scores are greater than would be expected by chance (shown by the horizontal dashed line). The data were obtained when the environmental noise outside the HPD was approximately 90 dBA at the center-head position. In this way, risks to the hearing of subjects from short-duration noise exposure should the HPD fail to introduce sufficient attenuation, or be accidentally dislodged or removed during an experiment, were effectively eliminated. When appropriately worn, the passive attenuation of the proof-of-concept device reduced environmental noise by $\sim$30 dBA. For these measurements, the long-term average A-weighted sound level of speech at the eardrum was first set to 48 dBA, and the environmental noise sound level was adjusted so that the speech SNR at the eardrum was -12 dB for each subject. The conditions were chosen to produce, on average, a word score of approximately 50% correct for the conventional passive HPD, so that both increases and decreases in word scores accompanying the introduction of signal processing could be tracked.

Inspection of Fig. 4 reveals that the MRT scores depend on the operating condition of the device covering the ears as well as the noise source. Subjects wearing the passive HPD (labeled "passive/fixed gain") and the proof-of-concept subband ANC device without adaptive communication gain processing (labeled "ANC/fixed Gain") produced similar mean MRT scores for these noise sources. The small differences between the mean MRT scores are consistent with the results of previous experiments on the speech intelligibility of communication headsets equipped with active adaptive feed-forward ANC (Brammer *et al.*, 2005). Higher MRT scores (i.e., improved speech understanding) resulted when subjects wore the proof-of-concept subband ANC device

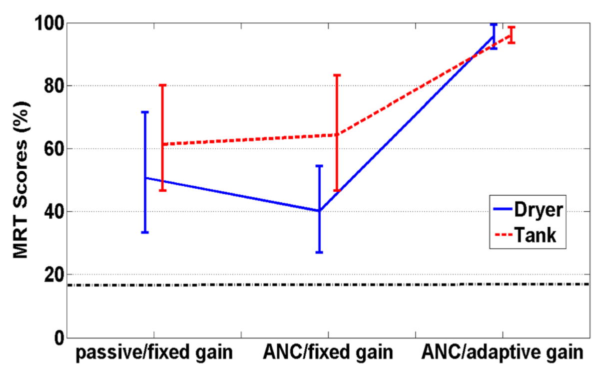

FIG. 4. (Color online) Mean MRT scores ($\pm$SD) in environmental noise (dryer, and tank) for six subjects when wearing the proof-of-concept device operated as a passive HPD with fixed communication channel gain (labeled passive/fixed gain), a subband ANC HPD with fixed communication channel gain (labeled ANC/fixed gain), or a subband ANC HPD with adaptive communication channel signal processing (labeled "ANC/adaptive gain"), when the speech SNR is set to -12 dB before signal processing. The horizontal dotted line shows a chance response.

with adaptive communication channel signal processing (labeled "ANC/adaptive gain").

A one-way within subjects analysis of variance (ANOVA) showed a significant difference in word recognition scores among the three device conditions for the military noise, $F(2,10) = 25.8$, $p < 0.001$, $\eta^2 = 0.62$. Tukey's Honestly Significant Difference (HSD) comparisons indicated significant differences between the adaptive ANC and communication channel gain vs passive conditions, and between the adaptive ANC and communication channel gain vs adaptive ANC and fixed communication channel gain conditions ($p < 0.05$), but not a significant difference between the passive vs adaptive ANC and fixed communication channel gain conditions ($p > 0.05$). Similarly, a one-way within subjects measures ANOVA showed a significant difference among the three device conditions for the industrial noise, $F(2,10) = 65.4$, $p < 0.001$, $\eta^2 = 0.81$. Tukey's HSD comparisons indicated significant differences between the adaptive ANC and communication channel gain vs passive conditions, and adaptive ANC and communication channel gain vs adaptive ANC and fixed communication channel gain conditions ($p < 0.05$), but not a significant difference between the passive vs adaptive ANC and fixed communication channel gain conditions ($p > 0.05$).

There were, however, considerable differences in word scores across subjects. Word recognition scores are shown for each subject in Fig. 5 when the subjects were exposed to the military tank noise and in Fig. 6 when they were exposed to the noise from the industrial dryer. Inspection of Fig. 5 reveals that the word scores ranged from 47% to 80% correct for the passive HPD, that is, when the sound level of the speech SNR was set to -12 dB and there was no signal processing to enhance speech intelligibility. Similarly, the word scores ranged from 22% to 62% correct for the passive HPD when the device was worn in the industrial noise (Fig. 6). The lowest score (S4) is only slightly greater than chance ($\sim$17%). The improvement in word scores when the proof-of-concept subband ANC device was worn and adaptive subband communication signal processing was enabled can be seen to be substantial for most subjects. Reference to Figs. 5 and 6 reveals that the word recognition scores for all subjects have increased (labeled ANC/adaptive gain), and are more than 90% for all subjects when exposed to the military tank noise and more than 85% when exposed to the industrial dryer noise.

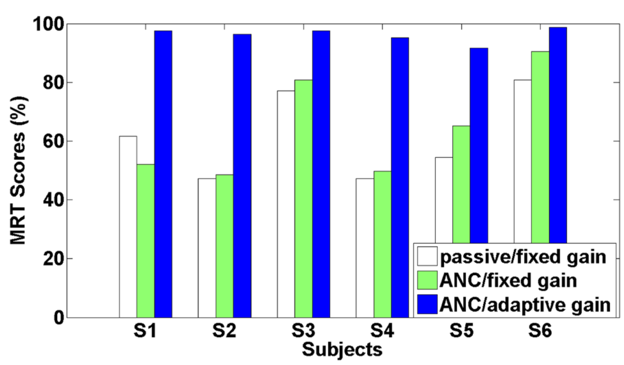

FIG. 5. (Color online) MRT scores in Leopard tank noise for individual subjects (S1–S6) when wearing the proof-of-concept device operated as a passive HPD with fixed communication channel gain (labeled passive/fixed gain), a subband ANC HPD with fixed communication channel gain (labeled ANC/fixed gain), or a subband ANC HPD with adaptive communication channel signal processing (labeled ANC/adaptive gain), when the speech SNR is set to -12 dB before signal processing.

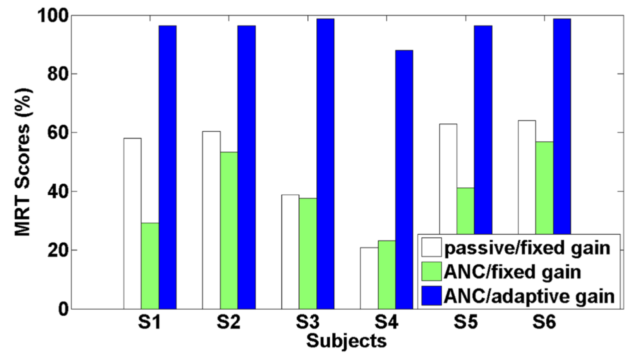

FIG. 6. (Color online) MRT scores in industrial dryer noise for individual subjects (S1–S6) when wearing the proof-of-concept device operated as a passive HPD with fixed communication channel gain (labeled passive/fixed gain), a subband ANC HPD with fixed communication channel gain (labeled ANC/fixed gain), or a subband ANC HPD with adaptive communication channel signal processing (labeled ANC/adaptive gain), when the speech SNR is set to -12 dB before signal processing.

B. Simulation

One purpose of simulating the proof-of-concept device was to predict its performance under noise exposure conditions that could not be used in experiments involving human subjects. It is therefore first necessary to establish the accuracy with which the simulation replicates results on human subjects under exposure conditions that are ethically acceptable. Since the SPLs of environmental noise and speech were sensed at the subjects' eardrums, the comparisons between word scores predicted by the simulation from the SPLs and those observed must be made for an individual, as sound pressure transformations from near the head to the eardrum will be subject specific, as already noted. Word scores obtained by simulation for one subject, S6, are compared with the observed MRT word scores in Fig. 7. The comparison is for both the industrial and military noises when the speech SNR at the eardrum is set to -12 dB, that

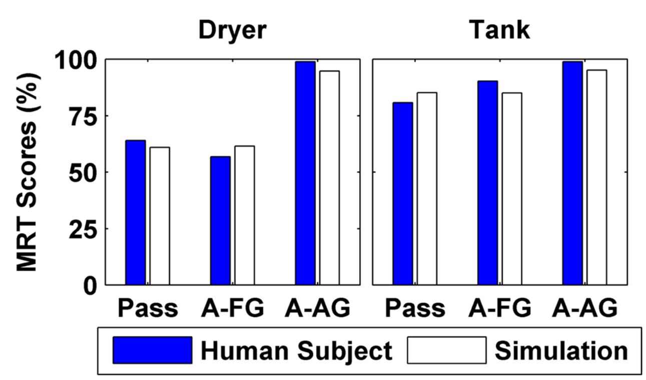

FIG. 7. (Color online) Predicted and observed MRT scores in environmental noise (industrial dryer, or tank) for subject S6 when wearing the proof-of-concept device operated as a passive HPD (labeled "pass"), a subband ANC HPD with fixed communication channel gain (labeled "A-FG"), or a subband ANC HPD with adaptive communication channel signal processing (labeled "A-AG"), when the speech SNR is set to -12 dB before signal processing.

is, the conditions used for the human subject testing. There are no adjustable parameters.

Inspection of Fig. 7 reveals that the observed word recognition scores are closely predicted by the simulation for each noise and for each operating condition of the proof-of-concept device (within ±5%). The simulation involved the following measurements, all conducted on subject S6, as well as the transfer functions for the signal processing components shown in Fig. 1.

(1) Transfer function from reference microphone to error microphone and to the eardrum when subject wearing proof-of-concept device.
(2) Transfer function from miniature loudspeaker in ear cup to error microphone when subject wearing proof-of-concept device.
(3) Transfer function from error microphone to eardrum when subject wearing proof-of-concept device.

Reference to Figs. 5 and 6 shows that the changes in word recognition scores between devices is smallest for this subject when exposed to the military noise and typical of those experienced by other subjects when exposed to the industrial noise. Overall, the agreement between the performance of the device, as predicted in simulation, and that obtained when the physical implementation is worn by the subject provides confidence that its performance can be predicted under other environmental noise conditions.

An important requirement for the device is that the total noise at the eardrum does not exceed a pre-established maximum sound level irrespective of the environmental noise level and the presence or absence of speech in the communication channel. A series of simulations has been performed to estimate the speech SNRs and A-weighted sound levels in each octave band for subject S6 as the environmental noise increases from ~90 to ~120 dBA (outside the earmuff). The results are shown for selected octave bands in Fig. 8 when the subject is exposed to the noise of the industrial dryer.

Inspection of Fig. 8 reveals that the speech SNR (solid line, and left ordinate) is maintained at nominally +10 dB in each octave band as the octave-band A-weighted sound level under the earmuff is increased from 60 to 70 dBA, or more (abscissa), depending on the band center frequency (shown above each panel). As communication channel gain control is applied separately to each octave band, the speech SNR is maintained at +10 dB in a band until the sound level of the combined environmental noise and speech under the earmuff reaches an octave-band level of ~80 dBA (dotted line, and right ordinate). For higher band environmental noise levels, the speech SNR is reduced so that the combined environmental noise and speech octave-band sound levels plateau at close to 80 dBA. The process is continued until the speech SNR reaches a preset minimum (−10 dB). Under these conditions the speech contributes little to the overall sound level, and further increases in the environmental noise outside the earmuff result in increased overall sound levels under the earmuff (right ordinate). This can be seen to be occurring in the 1 kHz octave band when the band environmental noise levels under the earmuff exceed 80 dBA.

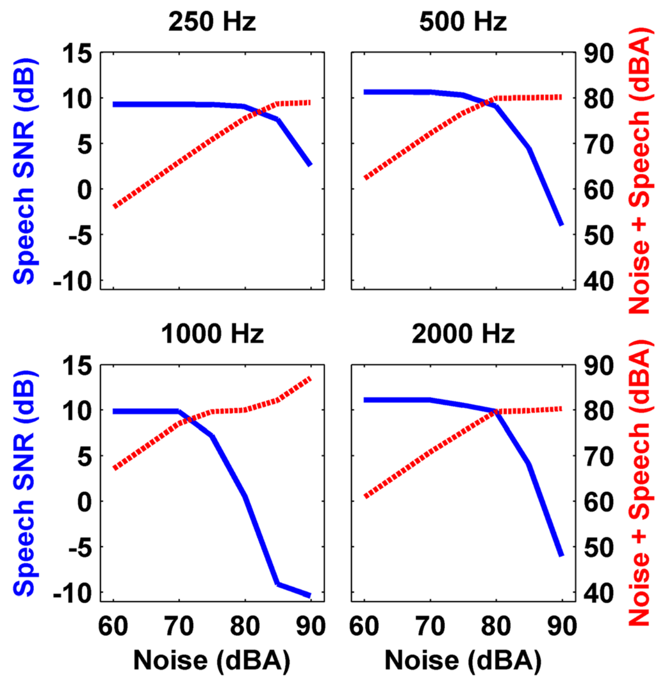

FIG. 8. (Color online) Predicted octave-band speech SNRs (solid line, and left ordinate) as a function of the A-weighted octave-band SPL of the industrial dryer under the earmuff before signal processing (abscissa), and the predicted A-weighted octave-band SPLs under the earmuff of the combined environmental noise and speech after processing (dashed line, and right ordinate). Predictions are for the 250, 500, 1000, and 2000 Hz octave bands.

The consequences of the signal processing are evident from the word scores predicted for subject S6 in Fig. 9. In this diagram, the word scores (solid line, and left ordinate) are shown as a function of the overall A-weighted sound level of the environmental noise under the earmuff before signal processing (abscissa). The combined environmental noise and speech is also shown after processing (dotted line, and right ordinate). It can be seen that the predicted word score is close to 95% correct for environmental noise levels under the earmuff from 60 to almost 85 dBA. The

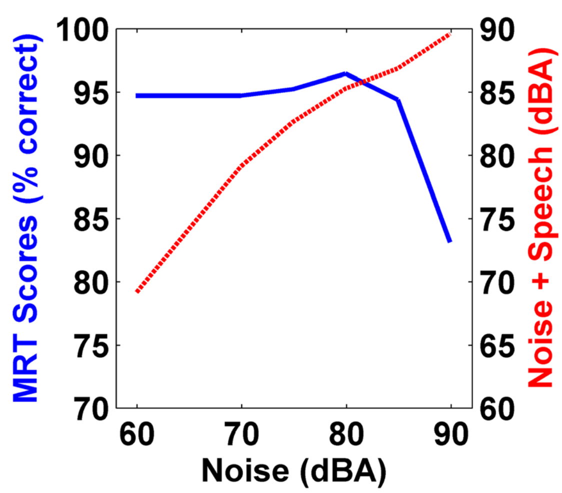

FIG. 9. (Color online) Predicted MRT scores (solid line, and left ordinate) as a function of the A-weighted sound level of the industrial dryer under the earmuff before signal processing (abscissa), and the predicted A-weighted sound level under the earmuff of the combined environmental noise and speech after processing (dashed line, and right ordinate).

corresponding sound levels of environmental noise outside the earmuff are between ∼90 and ∼115 dBA, and represent exposure conditions in which the proof-of-concept device could be expected to find application in industry. For all these sound levels, the combined environmental noise and speech sound levels under the earmuff after signal processing range from 70 to close to 85 dBA (right ordinate). It should be noted that the word score shown in this diagram for the lowest sound level is confirmed by the MRT word recognition score recorded by subject S6 (shown in Fig. 7). Figure 9 also shows that the combined environmental noise and speech sound levels under the earmuff after signal processing exceed 85 dBA when the environmental noise outside the earmuff exceeds 115 dBA.

IV. DISCUSSION

For the environmental noises employed in this study, which were chosen to possess time histories and frequency spectra representative of those likely to be encountered during use of the proof-of-concept device, subband speech SNR control produced greater improvement in word scores than subband ANC (see, for example, Figs. 4–6). This observation is a consequence of several factors. Most information in speech is contained at frequencies from about 250 Hz to 4 kHz, while the effectiveness of ANC decreases with increasing frequencies above 400 Hz. Hence the direct improvement in speech SNR from ANC will be greatest at frequencies that contain some, but not all, speech information. Nevertheless, the maximum ANC achievable for the earmuff employed in this study is predicted to be ∼2 to 3 dB at frequencies up to more than 1 kHz (see Fig. 2), which could, in principle, improve the discrimination of speech components at these frequencies. A commensurate improvement in word recognition score was not, however, observed. This may have been a consequence of failure of the proof-of-concept device to achieve the predicted maximum ANR when worn by a subject. It may also have been a consequence of the speech test employed, which evaluates consonant discrimination and includes distinguishing similar sounds containing higher frequencies, for example, "path" from "pass." The MRT was chosen for this study for its relevance to communications in which it is necessary to identify individual words or sounds, such as in air traffic control (e.g., confusing aircraft call signs "AA123" and "UA123"), rather than understand conversational speech.

The potential magnitude of the change in word score will, of course, depend on the initial condition chosen for the measurements, which was a speech SNR of −12 dB for each noise source. This was expected to result in a mean word score of ∼50% in the absence of signal processing (i.e., the passive/fixed gain condition of Fig. 4), and hence enable both substantial increases and decreases in word recognition scores to be recorded. This is evident from the word scores of individuals, which, as expected, differ. Inspection of Figs. 5 and 6 reveals the inconsistency of the word scores across individuals in the passive/fixed gain condition, even when the same speech SNR was initially set at the eardrum for each subject. Of more interest is the dramatic improvement in the consistency of the word scores, in addition to the increase in magnitude of the word scores, when subband ANC and adaptive communication channel signal processing are employed.

The speech SNR of +10 dB used as the target for communication channel signal processing will provide little improvement in MRT word score beyond that obtained when SNR = +4 dB for normal-hearing listeners (House *et al.*, 1965). The higher SNR was employed in this study as it reflected values apparently commonly selected by users of communication headsets and HPDs (Giguère and Dajani, 2009; Giguère *et al.*, 2012). It is well known that listeners commonly choose a greater speech SNR than necessary for intelligibility to reduce listening effort (Sarampalis *et al.*, 2009). The mean speech SNR chosen by users, which was in fact reported by Giguère and Dajani (2009) to be almost +14 dB, may reflect anticipating the sound level needed to maintain intelligibility when the environmental noise becomes more intense, listening while performing another task (Hodgetts *et al.*, 2009), or may possibly reflect temporary or permanent hearing loss in some cases. However, it is not necessarily an appropriate value when the speech SNR is automatically adjusted to maintain speech intelligibility. A value that gives confidence to users that all critical communications will be understood would seem more desirable. The approach adopted here of optimizing speech SNR in separate subbands while avoiding signal compression eliminates the possibility of introducing electronic distortion that could compromise intelligibility (Bockstael *et al.*, 2011). Moreover, a subband approach could be extended to persons with mild hearing loss, who could benefit from increased speech SNR (Plomp, 1986).

The range of sound levels over which the performance of the proof-of-concept device could be determined on human subjects in a laboratory setting without risk to hearing constrained establishing its performance under the conditions likely to be found in an industrial setting. The approach adopted here to overcome this limitation was to assess the performance of the device in simulation. A common criticism of simulations is the accuracy and range of performance parameters over which they reproduce the performance of the device they purport to represent. These criticisms have been addressed in the following ways.

First, considerable effort was expended to reproduce in simulation the word recognition score obtained by one subject in the two noise environments. As already noted, no adjustable parameters are used in the simulation to predict the MRT word scores, which agree closely with those obtained by the subject (Fig. 7). An essential component of the simulation was to employ sound pressures measured at the eardrum of the subject and sound pressure transformations recorded when the subject wore the proof-of-concept device. The use of averaged or generalized sound pressure transformations from several subjects or those reported in other studies (e.g., Hammershøi and Møller, 1996), or sound pressures for environmental noise or speech measured at locations other than the eardrum, failed to replicate the word recognition scores obtained by this subject. Similarly, the mean word scores obtained by the subjects could not be predicted by a simulation built on averaged sound pressure transformations and sound pressures not recorded at the eardrum. It is not clear the extent to which the disagreement arises from the absence of physical measurements of sound pressure obtained on each individual in the simulation or for other reasons, such as the failure to include a subject's hearing thresholds in the STI model for predicting speech intelligibility, or the intrinsic differences between individuals in auditory abilities (Kidd *et al.*, 2007). The need to obtain physical measurements of sound pressure on an individual for the simulation was only recognized sometime after the completion of human subject testing, when only one subject could be recalled (S6). Some success in predicting mean word scores in groups of subjects exposed to noise has been obtained using the Speech Intelligibility Index when individual hearing thresholds are included (Dubno *et al.*, 2005).

Second, the prediction of word scores at higher sound levels than those employed for subject testing has only been attempted for subject S6, and the algorithm used for the simulation was that used in the physical implementation (Sec. II A). With the exception of the reference microphone (R in Fig. 1), which will be subjected to the intense environmental noise, all transducers operate at moderate SPLs (typically from 70 to 90 dBA—see right ordinate of Fig. 9) and so are unlikely to introduce substantial distortion. A microphone designed to operate with low distortion at the sound levels of the environment is required for R. Thus a linear model for the physical components of the system, as employed in the simulation, would appear appropriate. However, the speech sound level will increase as the environmental noise increases under the earmuff, reaching ∼85 dBA when the sound level of noise outside the earmuff is ∼115 dBA. At these sound levels the word score obtained in practice will be up to 10% less than that predicted owing to the upward spread of masking (Studebaker *et al.*, 1999; Dubno *et al.*, 2005). It may be necessary to reduce this effect in a future device by, for example, maintaining speech sound levels at the ear to less than ∼75 dB, though the consequent reduction in speech SNR at high environmental noise levels will also reduce intelligibility (e.g., see Fig. 9 for environmental sound levels under the earmuff of 85 dBA, and above). A solution would be to further increase the attenuation of the device.

V. CONCLUSIONS

For the environmental noises employed in this study, which were chosen to possess time histories and frequency spectra representative of those likely to be encountered during use of the device, subband speech SNR control produced greater improvement in word scores than subband ANC. A primary benefit of the subband ANC with adaptive communication channel gain signal processing was the improved consistency of word recognition scores across subjects, in addition to the general increase in magnitude of the word scores.

Sound pressures measured at the eardrum and sound pressure transformations recorded when a subject wore the proof-of-concept device were required to predict the word recognition scores recorded by the subject. The use of averaged or generalized sound pressure transformations from several subjects, or sound pressures for environmental noise or speech measured at locations other than the eardrum, failed to replicate the observed word scores.

A linear model was employed in the simulation for the physical components of the system, to predict the performance of the proof-of-concept device at sound levels outside the earmuff of up to ∼120 dBA. The model predicts word scores will be maintained in excess of 90% correct under these conditions. However, it does not account for the upward spread of masking, which is expected to reduce the word score by up to 10% at environmental sound levels in excess of ∼100 dBA.

The overall performance of the proof-of-concept device is considered sufficiently encouraging to warrant miniaturization of the electronics, to produce a body-mounted unit with battery life sufficient to operate for at least a working day. A low-powered version of the DSP used in the study is now available with appropriate power consumption for a portable device: The rapid development of electronic components for portable applications will enable further miniaturization and reduction in power requirements. While the goal of the present device is to remove the user's volume control, the device would benefit from providing the user some control of speech SNR, to adjust for individual differences in speech understanding, and for different noise environments. Additionally, a simpler device could be developed involving only subband communication signal processing. Alternatively, a different control structure could be employed for ANC to increase the ANR at higher speech frequencies. Extending the performance of the device to include more generally the problems of face-to-face communication could well encourage convergence with technologies being developed for hearing aids.

ACKNOWLEDGMENTS

This work was supported by the National Institute for Occupational Safety and Health under research Grant No. R01 OH0086690-05.

Abel, S. M., Nakashima, A., and Saunders, D. (**2011**). "Speech understanding in noise with integrated in-ear and muff-style hearing protection systems," Noise & Health **13**, 378–384.

ANSI (1989). ANSI S3.2-1989. (R2009) American National Standard—Method for Measuring the Intelligibility of Speech over Communication Systems (American National Standards Institute, New York), pp. 1–16.

ANSI (2010). ANSI S3.6-2010. American National Standard Specification for Audiometers (American National Standards Institute, New York), pp. 1–55.

Berger, E. H. (**2003**). "Hearing protection devices," in *The Noise Manual*, 5th ed., edited by E. H. Berger, L. H. Royster, J. D. Royster, D. P. Driscoll, and M. Layne (American Industrial Hygiene Association, Fairfax, VA), pp. 379–454.

Bernstein, E., Brammer, A., Yu, G., Cherniack, M., and Peterson, D. (**2010**). "Ear cup selection for feedforward active noise reduction hearing protectors," Can. Acoust. **38**(3), 82–83.

Bernstein, E. R. (**2013**). "Development of a delayless subband active noise reduction hearing protector for improvement of speech intelligibility and detection of auditory warning sounds," Ph.D. dissertation, University of Connecticut, Storrs, CT.

Bernstein, E. R., Brammer, A. J., and Yu, G. (**2013**). "Improving speech understanding in communication headsets: Simulation of adaptive subband processing for speech in noise," Int. J. Industr. Ergonom. **43**, 526–535.

Bockstael, A., De Coensel, B., Botteldooren, D., D'Haenens, W., Keppler, H., Maes, L., Philips, B., Swinnen, F., and Bart, V. (**2011**). "Speech recognition in noise with active and passive hearing protectors: A comparative study," J. Acoust. Soc. Am. **129**, 3702–3715.

Brammer, A. J., and Pan, G. J. (**1998**). "Opportunities for active noise control in portable communication devices," *Noise Effects 98 Congress Proceedings* (Noise Effects 98 Ltd., Sydney), pp. 181–186.

Brammer, A. J., Peterson, D. R., Cherniack, M. G., and Gullapalli, S. (**2005**). "Improving the effectiveness of communication headsets with active noise reduction: Influence of control structure," in *RTO-MP-HFM-123, New Directions for Improving Audio Effectiveness* (North Atlantic Treaty Organization, Neuilly-sur-Seine, France), pp. 6-1–6-8.

Brammer, A. J., Yu, G., Bernstein, E. R., Cherniack, M. G., Tufts, J. B., and Peterson, D. R. (**2011**). "Intelligibility of speech corrupted by nonlinear distortion," in *Proceedings of the 10th International Congress on Noise as a Public Health Problem* (London), pp. 287–292.

Brammer, A. J., Yu, G., Bernstein, E. R., Peterson, D. R., Cherniack, M. G., and Tufts, J. B. (**2009**). "Field measurement of sound pressure at the eardrum," *Inter-Noise2009* (Inst. Noise Control Eng., Ames, IA), pp. 1–7.

Brammer, A. J., Yu, G., Peterson, D. R., Bernstein, E. R., and Cherniack, M. G. (**2008**). "Hearing protection and communication in an age of digital signal processing: Progress and prospects," in *Proceedings of the 9th International Congress on Noise as a Public Health Problem* (Foxwoods, CT), pp. 1–8.

Casali, J. G. (**2010**). "Powered electronic augmentations in hearing protection technology circa 2010 including active noise reduction, electronically-modulated sound transmission, and tactical communications devices: Review of design, testing, and research," Int. J. Acoust. Vib. **15**, 168–186.

Chung, K. (**2007**). "Effective compression and noise reduction configurations for hearing protectors," J. Acoust. Soc. Am. **121**, 1090–1101.

Dubno, J. R., Horwitz, A. R., and Ahlstrom, J. B. (**2005**). "Word recognition scores in noise at higher-than-normal levels: Decreases in scores and increases in masking," J. Acoust. Soc. Am. **118**, 914–922.

Giguère, C., Behar, A., Dajani, H. R., Kelsall, T., and Keith, S. E. (**2012**). "Direct and indirect methods for the measurement of occupational sound exposure from communication headsets," Noise Control Eng. J. **60**, 630–644.

Giguère, C., and Dajani, H. R. (**2009**). "Noise exposure from communication headsets: The effects of external noise, device attenuation and effective listening signal-to-noise ratio," *Inter-Noise2009* (Inst. Noise Control Eng., Iowa State University, Ames, IA), pp. 1–8.

Giguère, C., Laroche, C., Brammer, A. J., Vaillancourt, V., and Yu, G. (**2011**). "Advanced hearing protection and communication: Progress and challenges," in *Proceedings of the 10th International Congress on Noise as a Public Health Problem* (London), pp. 225–233.

Hammershøi, D., and Møller, H. (**1996**). "Sound transmission to and within the human ear canal," J. Acoust. Soc. Am. **100**, 408–427.

Hodgetts, W., Szarko, R., and Rieger, J. (**2009**). "What is the influence of background noise on the listening level of iPod users?," Int. J. Audiol. **48**, 825–832.

House, A. S., Williams, C. W., Hecker, M. H. L., and Kryter, K. D. (**1965**). "Articulation testing methods: Consonant differentiation with a closed-response set," J. Acoust. Soc. Am. **37**, 158–166.

Kidd, G. R., Watson, C. S., and Gygi, B. (**2007**). "Individual differences in auditory abilities," J. Acoust. Soc. Am. **122**, 418–435.

Kuo, S., and Morgan. D. (**1996**). *Active Noise Control Systems: Algorithms and DSP Implementations* (John Wiley & Sons, New York), pp. 1–388.

Lee, K., Gan, W., and Kuo, S. (**2009**). *Subband Adaptive Filtering: Theory and Implementation* (John Wiley & Sons, New York), pp. 1–133.

Morgan, D., and Thi, J. (**1995**). "A delayless subband adaptive filter architecture," IEEE Trans. Signal Process. **43**, 1819–1830.

National Institute for Occupational Safety and Health (**1998**). DHHS (NIOSH) 98-126. "Criteria for a Recommended Standard: Occupational Exposure to Noise" (National Institute for Occupational Safety and Health, Cincinnati), pp. 1–122.

Pan, G. J., Brammer, A. J., Zera, J., and Goubran, R. (**1995**). "Application of active feed-forward noise control to a circumaural hearing protector," ACTIVE 95 (Noise Control Foundation, New York), pp. 1319–1326.

Payton, K. L., and Braida, L. D. (**1999**). "A method to determine the speech transmission index from speech waveforms," J. Acoust. Soc. Am. **106**, 3637–3648.

Plomp, R. (**1986**). "A signal to noise ratio model for speech reception threshold of the hearing impaired," J. Speech Hear. Res. **29**, 146–154.

Rafaely, B., and Jones, J. (**2002**). "Combined feedback-feedforward active noise-reducing headset—The effect of the acoustics on broadband performance," J. Acoust. Soc. Am. **112**, 981–989.

Rankovic, C. M., Freyman, R. L., and Zurek, P. M. (**1992**). "Potential benefits of adaptive frequency-gain characteristics for speech reception in noise," J. Acoust. Soc. Am. **91**, 354–362.

Ray, L., Solbeck, S., Streeter, A., Collier, R. (**2006**). "Hybrid feedforward-feedback active noise reduction for hearing protection and communication," J. Acoust. Soc. Am. **120**, 2026–2036.

Sarampalis, A., Kalluri, S., Edwards, B., and Hafter, E. (**2009**). "Objective measures of listening effort: Effects of background noise and noise reduction," J. Speech Lang. Hear. Res. **52**, 1230–1240.

Schwander, G. H., and Levitt, H. (**1987**). "Effect of two-microphone noise reduction on speech recognition by normal hearing listeners," J. Rehab. Res. Dev. **4**, 87–92.

Shaw, E. A. G. (**1997**). "Acoustical features of the human external ear," in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. H. Gilkey and T. H. Anderson (Lawrence Erlbaum Associates, Mahwah, NJ), pp. 25–47.

Shaw, E. A. G., and Thiessen, G. J. (**1958**). "Improved cushion for ear defenders," J. Acoust. Soc. Am. **30**, 24–36.

Shaw, E. A. G., and Vaillancourt, M. M. (**1985**). "Transformation of sound pressure level from the free field to the eardrum presented in numerical form," J. Acoust. Soc. Am. **78**, 1120–1123.

Shields, P. W., and Campbell, D. R. (**2001**). "Improvements in intelligibility of noisy reverberant speech using a binaural subband adaptive noise-cancellation processing scheme," J. Acoust. Soc. Am. **110**, 3232–3242.

Simshauser, E. D., Meeker, W. F., and Balakrishnan, A. V. (**1956**). "The noise-canceling headset—An active ear defender," J. Acoust. Soc. Am. **28**, 773.

Stinson, M. R., and Daigle, G. A. (**2005**). "Comparison of an analytic horn equation approach and a boundary element method for the calculation of sound fields in the human ear canal," J. Acoust. Soc. Am. **118**, 2405–2411.

Studebaker, G. A., Sherbecoe, R. L., McDaniel, D. M., and Gwaltney, C. A. (**1999**). "Monosyllabic word recognition at higher-than-normal speech and noise levels," J. Acoust. Soc. Am. **105**, 2431–2444.

Suter, A. H. (**1992**). "Communication and job performance in noise: A review," Monograph 28 (American Speech-Language-Hearing Association, Rockville, MD), pp. 1–84.

Toner, E., and Campbell, D. R. (**1993**). "Speech enhancement using subband intermittent adaptation," Speech Commun. **12**, 253–259.

van Dijkhuisen, J. N., Festen, J. M., and Plomp, R. (**1991**). "The effect of frequency-selective attenuation on the speech reception threshold of sentences in conditions of low-frequency noise," J. Acoust. Soc. Am. **90**, 885–894.

Wheeler, P. D. (**1986**). "Voice communication in the cockpit noise environment—The role of active noise reduction," Ph.D. thesis (Institute of Sound and Vibration Research, University of Southampton, Southampton, UK), pp. 1–97.

Yu, G., Brammer, A. J., Swan, K., Tufts, J., Cherniack, M., and Peterson, D. (**2010**). "Relationships between the modified rhyme test and objective metrics of speech intelligibility," J. Acoust. Soc. Am. **127**, 1903.