INTELLIGIBILITY IN ACTIVE COMMUNICATION HEADSETS:
Role of Error Path in Active Noise Reduction and Speech Reproduction

A.J. Brammer,[1,2] R.B. Crabtree,[3] D.R. Peterson,[2] M.G. Cherniack[2]
[1]National Research Council, Canada; [2]University of Connecticut Health Center, U.S.A.;
[3]DRDC-Toronto, Canada

Introduction A primary expectation of communication headsets is to maintain speech intelligibility under all operational conditions, and especially in circumstances in which a loss of intelligibility may have serious consequences, e.g., military operations, air-traffic control. This requirement may be difficult to maintain in noisy environments and for persons with hearing loss, and also when the communication system is operated at sound levels sufficient to induce temporary threshold shift from the speech or spurious electronic signals.
Effective application of active control to maintain or improve speech intelligibility in communication headsets requires an understanding of the limitations imposed by physical acoustics, control theory, signal processing, and electronics, as well as the physiological and psycho-acoustic mechanisms influencing speech perception. Almost all of the current generation of commercial communication headsets equipped with active control to reduce environmental noise employ a miniature microphone and loudspeaker mounted in a circumaural earmuff, with the electro-acoustic components connected to a feedback control system. The device is usually designed to provide active noise reduction (ANR) at frequencies below 1 kHz, as the passive attenuation of the circumaural earmuff provides sufficient noise reduction at higher frequencies.
The interaction between the performance of the control system and speech communication in ANR headsets has been infrequently examined.(Gower et al., 1994; Nixon et al., 1992; Steeneken et al., 1997; Brammer et al., 1998) The purpose of this paper is to explore this relationship with particular reference to the control structure and, more specifically, the error path. While the role of most factors in the performance of a communication headset with ANR has been extensively discussed, the role of the error path has received comparatively little attention.

Control Structure and Error Path Simplified block diagrams for one circumaural earmuff of an active headset containing the essential elements of a single-input, single-output, feedback, or feedforward, control system are shown in Figs. 1A and 1B, respectively. A fixed-filter feedback control structure is shown, as it is commonly employed commercially, and an adaptive-filter feedforward control structure that has been applied to a communication headset. The latter has been described in detail elsewhere.(Brammer et al., 1997) The complete headset consists of two such earmuffs with cushions to provide an air seal between the earmuff and the head, connected by a sprung headband. Each ear cup contains an independent ANR system. The block diagrams show the signal paths (continuous lines for the control system, and dashed lines for the communication channel) with directions (arrows), and signal summation and subtraction (Σ^{+}, and Σ^{-}).
In Fig. 1A, the control filter processes the input signal in a prescribed manner intended to reduce environmental noise at the location of the microphone. An integral part of the process of sound cancellation is the transformation of the electrical signal to sound by the loudspeaker, the propagation of sound from the loudspeaker, S, to the microphone, E, and the transformation

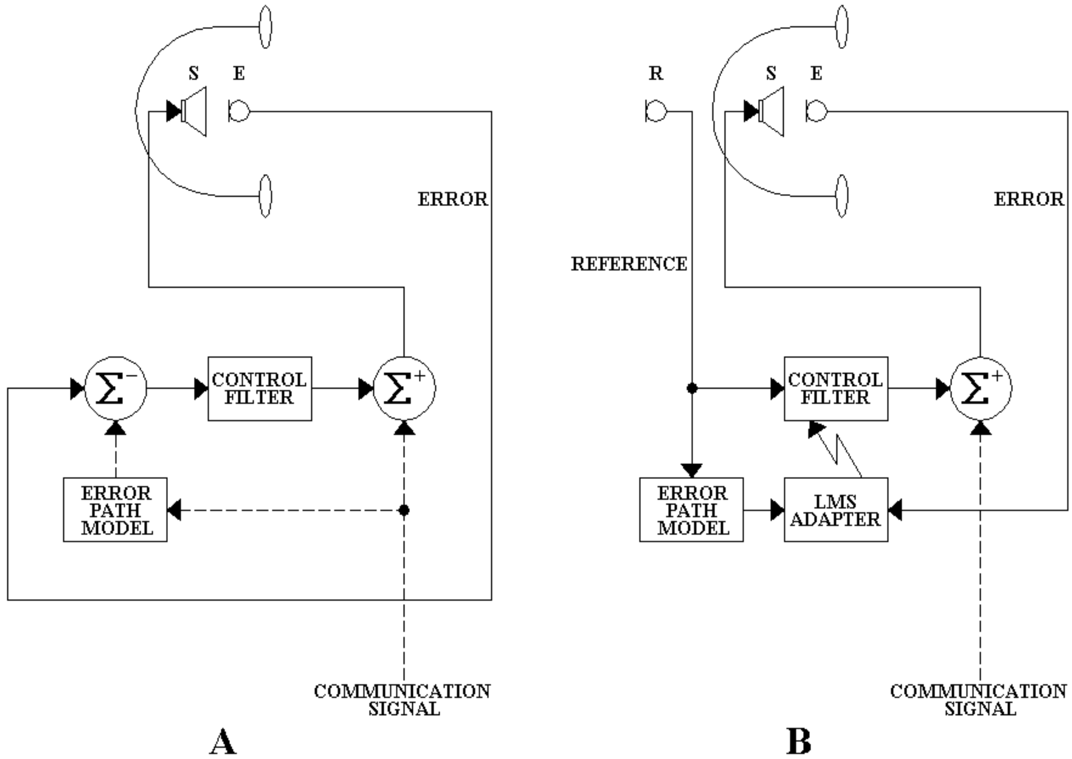

Figure 1: Communication headset with: A – feedback, and B – feedforward active control system, with LMS algorithm to adjust the control filter in B.

of sound into an electrical signal by the microphone. These processes together define the transfer function from S to E, which is termed the *error path*. In essence, the microphone detects the error in the sound produced by the loudspeaker (i.e., the combination of the environmental noise and the canceling sound at E), and hence is known as the *error* microphone. It should be noted that the error path will be influenced by the presence of the head, and an air gap between the cushion of the earmuff and the head (i.e., earmuff sealed or unsealed), and so is a variable and difficult to control component of the system.(Brammer et al., 1999; Steeneken et al., 1997)

Examples of error-path transfer functions determined using multiple length sequence stimuli applied to the loudspeaker, S, are shown in Fig. 2, when the headset is worn by a human subject. The method of measurement has been described elsewhere.(Brammer et al., 1999) The results shown by the continuous line, A, give the relative magnitude of the transfer function when the cushion of the earmuff is well sealed to the head. When the seal is broken by, in this case, introducing an 8 mm diameter rod between the cushion and the side of the head, the magnitude of the error-path transfer function is changed to that shown by the dashed line, B. It is evident from these measurements that the primary effect of the change in contact between the earmuff and the head occurs at frequencies below 700 Hz. A similar result was obtained when the earmuff was suspended above, but not in contact with, a flat-plate coupler.(Brammer et al., 1999) Clearly, the changes in transfer function occur at frequencies at which the active control system is required to function effectively, and so may compromise the ANR.

The introduction of a communication signal to the control system of Fig. 1A may be performed in several ways. Some commercial active headsets feed the communication signal directly into the feedback control loop employing only the signal summation path at Σ^{+}. A more desirable control structure is shown in the diagram, which introduces a representation of the error-path

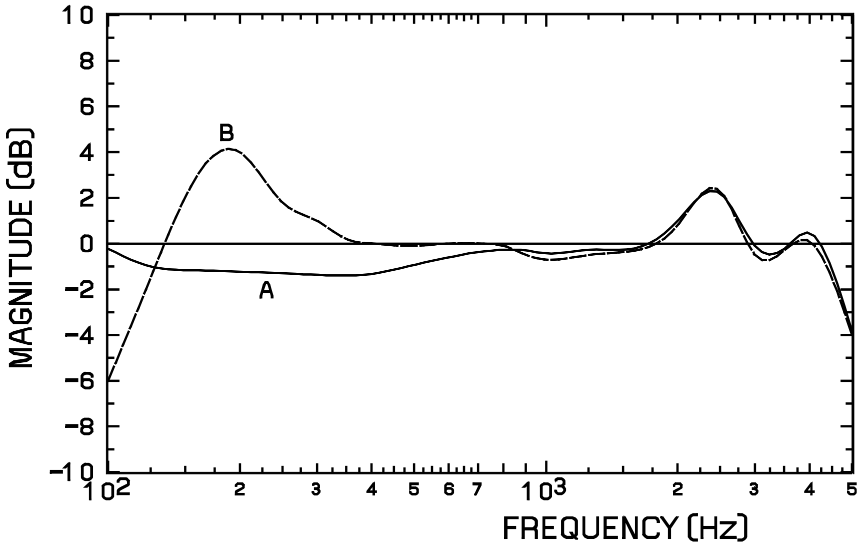

Figure 2: Relative magnitude of error-path transfer function: A – with good seal between head and earmuff, and B – seal broken by 8 mm diameter rod.

transfer function (error-path model). The communication signal is directly summed with the signal produced by the control filter at Σ^+ and fed to the loudspeaker. However, in these circumstances the speech sounds are mixed with the environmental sounds sensed by the error microphone, and would be cancelled by the controller unless removed prior to signal processing. This is achieved in Fig. 1A by subtracting the speech signal from the error signal at Σ^-. The success of this separation between the signal representing the residual noise to be controlled and that representing the speech, and thus ultimately the quality of the speech signal subsequently fed to the loudspeaker, depends on the precision of the error-path model as well as the fidelity of sound reproduction by the loudspeaker. An alternate, and equivalent, approach to that shown in Fig. 1A is to remove the error path model from the communication channel and insert an inverse model of the error path into the control path after the summation, Σ^+. As already noted, the error path is dependent on the subject and the fit of the earmuff to the side of the head. In a fixed-filter design, it is inevitable that a typical error-path model will be employed, which must compromise speech quality.

The same loudspeaker and error microphone are to be found within the earmuff of a headset employing a feedforward control system (Fig. 1B). In this case, however, an additional microphone, R, is used to sense the sound field external to the earmuff. This *reference* microphone provides the input signal to the controller, which must then model the transfer function from the location of the reference microphone to that of the error microphone (i.e., the transmission of sound through the ear cup, as well as leakage the earmuff cushion and the head), and the transformation of the electrical signal to sound by the loudspeaker. This process is done by successively modifying the filter constructed by the controller, and is implemented digitally by an adaptive filter. The adjustment requires comparing the reference signal, filtered by an error-path model, with the error signal, using an algorithm designed for this purpose (e.g., LMS algorithm, in Fig. 1B).(see Nelson et al., 1993) The communication signal is directly summed with the signal produced by the control filter at Σ^+ and fed to the loudspeaker, as before. Note that in this control structure the output of the error microphone does not enter

the control filter, and so cannot influence the speech signal. No compensation for the presence of speech in the error signal is thus required, and no degradation of the speech signal by the

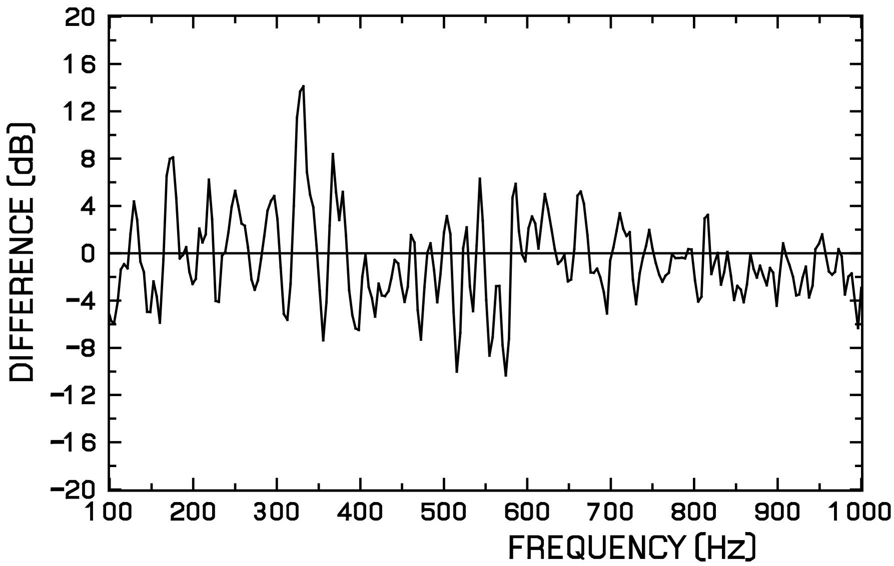

Figure 3: Difference in ANR between earmuff with, and earmuff without, speech signal.

control system is expected to occur. The quality and intelligibility of the speech signal may then be expected to depend solely on the fidelity of sound reproduction by the loudspeaker, assuming the electronics introduce insignificant distortion.

Experiments designed to confirm the extent to which the noise reduction and speech intelligibility of ANR headsets are degraded by the interaction between the speech and environmental noise are described in the following sections for the two control structures.

Influence of Speech Signal on Active Noise Reduction For a feedback active control system with control structure shown in Fig. 1A, the ANR will be influenced by the introduction of a communication signal if the error path model is inaccurate. A similar conclusion may be drawn for a feedback controller that either omits the error path model from Fig. 1A or includes an inverse model in the control path. The influence of a speech signal on the performance of an adaptive, feedforward active control system is believed to be insignificant, based on the control structure and the lack of coherence between the speech and the environmental noise. To confirm this expectation, the interaction between speech and noise has been measured for a headset employing the feedforward control structure shown in Fig. 1B when worn by a human subject.(Brammer et al., 1998) The headset contained two, identical, independent ANR systems, one for each ear, with loudspeaker and microphones mounted on circumaural earmuffs. For this experiment the subject was immersed in a simulated diffuse sound field representing the noise inside a Leopard tank, which produced a sound level at the ear comparable to that of speech from the communication channel when the ANR system was operating. The measurement relies on the premise that the rate of adaptation of the control system can be set to be much longer than the pause between utterances, but shorter than the total duration of several utterances. In these circumstances, the control system will adapt to the situation in which the error signal contains the speech signal. The speech consisted of

sentences recorded by a male speaker with a brief pause (approximately 5 s) between each. In order to compare the performance of the control system in the presence of speech with that occurring in the absence of speech, the experiment was performed with the communication channel operating only in the ANR system mounted on the right ear. The ANR system for the

Table 1: STI Values obtained with feedforward, and feedback, ANR headsets, averaged across two measurement replications and five subjects (N=10). The resultant sound levels at the ear due solely to the reverberant noise field are also given. The speech stimulus level was 70 dBA in all cases.

	STI Full	STI -6 dB	STI Quiet	Noise Level Full	Noise Level -6 dB	Noise Level Quiet
Feedforward ANR – ON	0.69	0.85	0.98	65.0 dBA	59.0 dBA	29.6 dBA
Feedback ANR – ON	0.39	0.55	0.85	66.6 dBA	60.8 dBA	28.4 dBA
Feedforward ANR – OFF	067	0.83	0.99	67.5 dBA	61.4 dBA	33.6 dBA
Feedback ANR – OFF	0.41	0.58	0.85	71.3 dBA	65.2 dBA	34.8 dBA

left ear thus continued to operate as an unperturbed active control system. The residual sound pressures at the error microphones with the two control systems operating were determined simultaneously during pauses between utterances (i.e. when there were no speech sounds at the error microphone in the right earmuff).

The difference between the sound pressures recorded at the two error microphones under these circumstances is shown in Fig. 3 for the speech frequencies at which the ANR systems operated (ANR at right error microphone minus ANR at left error microphone). In this diagram, a deviation from zero in excess of approximately 2 dB indicates a change in performance between the ANR systems. It can be seen that the average difference between the sound pressures in the two earmuffs is close to zero, but there are narrow frequency bands where the difference appears to deviate from zero, with cyclical trends both above and below zero. At frequencies at which the largest deviations from zero occur (below 400 Hz), each active control system is producing at least 10 dB of noise reduction, which is comparable to the magnitude of the some of the deviations. Thus, the expectation that the control system will operate independently of the presence or absence of the communication signal, as previously reported, can only be given qualified confirmation from close inspection of the results.

Influence of Active Noise Reduction on Speech Intelligibility The influence of ANR on speech intelligibility was determined using an objective procedure that involves the measurement of the Speech Transmission Index, or STI.(see Houtgast et al., 2002) STI is a figure of merit for the communication link under test that varies from zero (no intelligibility) to unity (ideal intelligibility). The method employed was the STITEL variant of STIDAS (Speech Transmission Device using Artificial Signals) intended for assessing the speech quality of telecommunication equipment.

Seven contiguous octave bands of noise at voice frequencies (center frequencies from 125 to 8000 Hz) were modulated simultaneously by low-frequency sine waves and the depth of modulation in each band of the return signal was computed, weighted in accordance with the contribution to intelligibility and then combined to form the STI.(IEC 60268-16, 1998) A miniature microphone was placed at the entrance of the ear canal of five human subjects. The STITEL signal level was set to 70 dBA through the communication channel of the ANR headset with the feedforward control system, and of a typical ANR headset with a feedback control system containing an approximate inverse error path model and both signal summation and subtraction, Σ^+ and Σ^-. A noise spectrum shaped to correspond to that of speech with a sound level of 90 dBA was established at the subject position within a reverberant room. The

three noise conditions, full, 6 dB below full, and quiet, were intended to produce a range of STI values. In all cases, the miniature microphone was used to register the combination of the STITEL test signal and the remaining confounding noise both with, and without, active control. The results are given in Table 1. It can be seen from the Table that the different measurement conditions produced a range of STI values, as intended, with the largest STI values being obtained with no interfering noise (quiet condition) and the smallest values when the most intense interfering noise was employed (full condition). For each noise condition the sound level of the interfering noise at the ear differed by less than 2 dB between headsets when the ANR systems were operating. However, in each case the STI recorded by the feedforward control system was much greater than that recorded by the feedback system (viz: 0.69 versus 0.39, and 0.85 versus 0.55). Moreover, a comparison of the STI values when the ANR systems were operating versus not operating, i.e., "ON" versus "OFF" in Table 1, reveals that the STI is marginally increased by the feedforward control system and marginally decreased by the feedback control system. The minimum intelligibility rating recommended for critical person-to-person communications, and alert and warning situations where correct understanding of critical words is required, is termed "fair" by ISO/DIS 9921, 2001. This rating corresponds to a STI value of from 0.45 to 0.6, and is only consistently achieved by the ANR headset with the feedforward control structure.

Conclusions Experiments to demonstrate the extent to which the noise reduction and speech intelligibility of ANR headsets depend on the control structure have largely confirmed expectations based on the influence of the error path. The ANR of a headset with a feedforward control system appears to be slightly perturbed by the presence of a speech signal, while the speech intelligibility of this system appears to be significantly greater than that of an ANR headset with a typical feedback control system. The results seem to support the notion that feedforward implementation of ANR is intrinsically less disruptive of communication signals when compared with the more common feedback circuits.

References

Brammer AJ, Pan GJ (1998). Opportunities for active control in portable communication devices. *Noise Effects 98 Congress Proc*., Noise Effects Ltd., Sydney, 181-186.

Brammer AJ, Pan GJ (1999). Application of adaptive feedforward active noise control to communication headsets: Use of subject-specific and synthesized error-path models. *Proc. Active 99*, Fort Lauderdale, U.S.A., 695-704.

Brammer AJ, Pan GJ, Crabtree RB (1997). Adaptive feedforward active noise reduction headset for low-frequency noise. *Proc. Active 97,* Budapest, Hungary, 365-372.

Gower DW, Casali JG (1994). Speech intelligibility and protective effectiveness of selected active noise reduction and conventional communication headsets. Human Factors, 36: 350-367.

Houtgast T, Steeneken HJM (2002). *Past, Present and Future of the Speech Transmission Index*. Soesterberg, The Netherlands, TNO Human Factors Publication ISBN 90-76702-02-0.

Nelson PA, Elliot SJ (1992). *Active Control of Sound*. London, Academic Press.

Nixon CW, McKinley RL, Steuver JW (1992). Performance of active noise reduction headsets, in: *Noise-Induced Hearing Loss*, Dancer AL, Henderson D, Salvi RJ, Hamernik RP, eds. St. Louis, Mosby, 389-400.

IEC 60268-16 (1998). Sound system equipment – Part 16: Objective rating of speech intelligibility by the speech transmission index. Geneva, International Electrotechnical Commission.

ISO/DIS 9921 (2001). Ergonomics: Assessment of speech communication. Geneva, International Organization for Standardization.

Steeneken HJM, Verhave JA (1997). Personal active noise reduction with integrated speech communication devices: Development and assessment, in: *Audio Effectiveness in Aviation.* Neuilly-sur-Seine, France: AGARD-CP-596, 18-1 - 18-8.