Reprint of Canadian Acoustics, Vol. 28, No. 1 (2000) pp. 3-13 article with corrections — ***Review article / Article de revue***

BINAURAL TECHNOLOGY FOR APPLICATION TO ACTIVE NOISE REDUCTION COMMUNICATION HEADSETS: DESIGN CONSIDERATIONS

Christian Giguère(1), Sharon M. Abel(2) and G. Robert Arrabito(3)

(1) Programme d'audiologie et d'orthophonie, Université d'Ottawa, Ottawa (ON), K1N 6N5

(2) The Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto (ON), M5G 1X5

(3) Defence and Civil Institute of Environmental Medicine, Toronto (ON), M3M 3B9

SUMMARY

This article examines the fundamental basis and the technical aspects involved in integrating two emerging technologies in the design of communication headsets for use in noisy environments. The first technology, active noise reduction (ANR), can improve signal detection and speech intelligibility by reducing the amount of interfering noise from the environment. The second technology, known as binaural technology, allows the creation of 3D auditory displays, which can improve signal detection and speech intelligibility in noise, and situational awareness, over monaural listening. For an optimal integration of binaural technology into ANR headsets, digital devices are preferred over analog devices. The complexity of the integrated system, particularly the features of the binaural simulation, is found to be largely dependent on the specific demands of the application targeted. Two extreme cases relevant to an aircraft cockpit environment are analyzed. The greatest benefit is likely to be found in situations of divided attention listening in relatively low signal-to-noise environments.

SOMMAIRE

Cet article examine les principes de base et les aspects techniques nécessaires à l'intégration de deux technologies émergentes dans la conception de casques d'écoute pour les milieux bruyants. La première technologie, le contrôle actif du bruit, permet d'améliorer la détection de signaux et l'intelligibilité de la parole en réduisant l'interférence causée par le bruit environnant. La deuxième technologie, la technologie binaurale, permet de créer un environnement d'écoute 3D, ce qui en retour permet d'améliorer la détection de signaux et l'intelligibilité de la parole dans le bruit, ainsi que la vigilance en situation d'écoute, par rapport à l'écoute monaurale. L'utilisation de casques actifs numériques est préférable aux casques actifs analogiques pour assurer une intégration optimale avec la technologie binaurale. La complexité du système total, tout particulièrement les caractéristiques de la simulation d'écoute binaurale, dépend en grande partie des exigences de l'application ciblée. Deux situations extrêmes appliquées à un environnement de cockpit d'avion sont analysées. L'avantage d'appliquer la technologie binaurale aux casques actifs sera le plus important en situation d'écoute où l'attention doit être partagée entre plusieurs signaux dans des milieux dont le rapport signal au bruit est faible.

1.0 INTRODUCTION

This study was undertaken to evaluate the feasibility of applying binaural technology to the design of active noise reduction (ANR) communication headsets. The long-term objectives of combining both technologies are the improvement of the intelligibility of competing spoken messages presented simultaneously in the presence of noise, and the enhancement of situational awareness in complex auditory listening environments. ANR technology (Steeneken and Verhave, 1996) can improve speech intelligibility by reducing the amount of interfering noise from the environment. This is accomplished by electronic sound wave cancellation of the environmental noise inside the earcups of the device. Binaural technology (Moller, 1992), on the other hand, allows the transfer of coincident messages to different virtual spatial positions by filtering the incoming communication signals with the head-related transfer functions of the user. This processing generates interaural time difference (ITD) and interaural level difference (ILD) cues for each message. Variation in these cues normally signifies real-world differences in spatial location (Blauert, 1997), and impacts on the intelligibility of speech in noise (Bronkhorst and Plomp, 1988).

In this article, we begin by reviewing the fundamental research on binaural speech intelligibility in noise and binaural technology. The practical aspects involved in the creation of directional audio signals through ANR headsets are then discussed in terms of the required technical characteristics of the devices, and the application requirements. The process of integrating binaural technology into ANR headsets is illustrated through two different listening scenarios relevant to an aircraft cockpit environment. The potential benefits are assessed.

2.0 Binaural Signal Detection And Speech Intelligibility in Noise

Incident acoustic signals are transformed by the complex geometry of the human torso, head and external ear (Shaw, 1974). This filtering produces direction-dependent sound spectra at the ears, and encodes time and level differences in the sound across the left and right ears. Binaural analysis of these cues provides the basis for localizing sound sources in space (Blauert, 1997).

In addition, the detection, discrimination and recognition of a sound signal in the presence of other signals or noises can sometimes be markedly improved when listening binaurally rather than monaurally (Yost, 1997). It has long been suggested that the binaural hearing cues could play a major role in separating sound sources perceptually (Cherry, 1953). Data on the benefits of binaural over monaural listening have been collected over the past decades.

2.1 Headphone studies

A representative early study was conducted by Levitt and Rabiner (1967a). They presented speech signals interaurally out-of-phase over headphones in the presence of a broadband white noise masker interaurally in-phase ($S_\pi N_0$). They found a masked threshold for speech detection about 13 dB lower than if both signal and noise were presented interaurally in-phase ($S_0 N_0$). Further experiments showed that this release from masking for detection was determined primarily by interaural phase opposition in the low-frequency region of the speech signal, typically below 0.5 kHz. In contrast, the maximum binaural gain at the 50% intelligibility level, i.e., the maximal decrease in speech reception threshold (SRT) with respect to the $S_0 N_0$ condition, was about 6 dB and required interaural signal phase opposition over a wide frequency region. Presenting the speech signal in-phase with an interaurally uncorrelated noise masker ($S_0 N_u$) led to a small decrease of about 3 dB in the masked detection threshold, but no advantage for speech intelligibility. In all conditions investigated, the binaural gain was substantially lower than the corresponding decrease in masked threshold. Levitt and Rabiner (1967b) predicted that even lower binaural gains could be expected for reference intelligibility levels greater than 50%.

2.2 Sound-field studies

Binaural speech intelligibility has also been investigated in rooms using spatially-separated loudspeakers for signal and noise sources (Figure 1). For example, Plomp and Mimpen (1981) measured the SRT for normal listeners in an anechoic room for a frontal speech signal, as a function of the azimuthal position θ of a speech noise masker. They found a general decrease in the SRT when the noise source was displaced from frontal to lateral positions. A maximal decrease in SRT, or binaural gain, of about 9-11 dB was found for a noise azimuth close to $\theta = 90°$.

Santon (1986) performed a similar experiment and found a maximal decrease in SRT of about 8 dB when a broadband white noise masker of moderate level was displaced from a frontal to a lateral position. If the broadband masker was divided into two noise bands, above and below 1.4 kHz, then the maximal decrease in SRT for each band was limited to 3 dB. Further experiments (Santon, 1987) showed that the variations in SRT for low (0.125-0.8 kHz) or mid (1-2 kHz) frequency noise bands were smaller than the corresponding variations in detection threshold for pure tones near the centre frequency of the bands, but followed the same trends as a function of noise azimuth. In the case of a high-frequency (2.5-6.3 kHz) noise band masker, the variations in SRT were also smaller than the corresponding variations in detection threshold, but the trends as a function of noise azimuth differed.

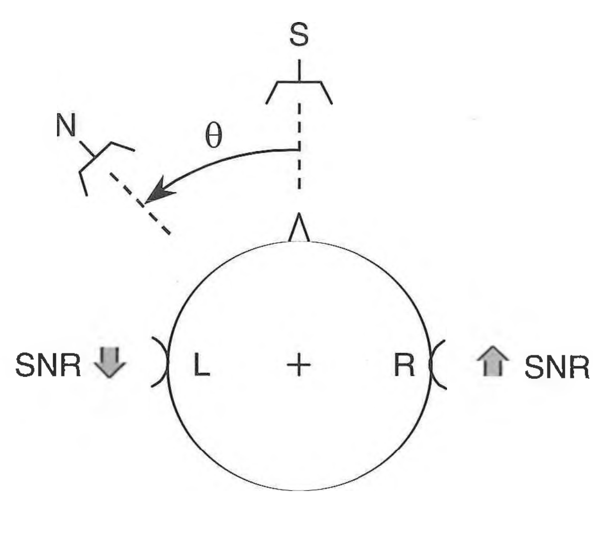


Figure 1: Schematic representation of binaural listening in sound field. L: left ear, R: right ear, S: signal source, N: noise source, θ: azimuth angle, SNR: signal-to-noise ratio.

Januška (1983) measured speech intelligibility for a frontal speech signal as a function of the level and spatial location (frontal, lateral, behind, above) of different noise maskers. Data were obtained for broadband white noise and octave bands of noise centred on 0.125 kHz to 8 kHz. The effect of varying the reverberation time of the listening space (anechoic, 0.4 s and 2.0 s) was also investigated. A coincident position of speech and noise sources was always the most unfavourable condition. Both the masking effect and the benefit of displacing the masker away from the frontal position were always greater for the broadband than the octave-band noises. Speech intelligibility gains were found for all conditions of the listening space, but were most evident in the anechoic environment. Typically, for broadband white noise, the maximum binaural gain at the 50% speech intelligibility level was about 14 dB in the anechoic environment, and 8 dB and 6 dB in the rooms with reverberation times of 0.4 s and 2.0 s respectively.

2.3 Simulated sound-field studies

In the sound-field studies discussed above, the speech and noise levels were defined with respect to the free field, typically at the head position in absence of the listener. Due to the direction-dependent transfer function of the external ear (Shaw, 1974), the actual signal-to-noise ratio (SNR) at the listener's ears will vary when the speech signal and/or noise sources are spatially displaced, and will be different across the two ears. For example, in Figure 1, when the noise source is displaced laterally, the noise level increases in the ipsilateral ear (SNR decreases) and decreases in the contralateral ear (SNR increases). Thus, the speech intelligibility benefit of spatially separating signal and noise sources from a common position may include a monaural contribution from the ear with the best SNR, as well as the contribution from binaural processing per se. Also, interaural time and level differences cannot be independently controlled in sound-field experiments, and thus their respective roles cannot be separated in the interpretation of speech intelligibility results.

To address these questions, Bronkhorst and Plomp (1988) simulated free-field conditions over headphones by presenting speech and noise signals recorded a-priori on a KEMAR manikin in an anechoic room. The speech signal recordings corresponded in all conditions to a frontal sound incidence. The noise recordings were made at several azimuth angles θ in the horizontal plane, and from each recording two additional noise signals were derived by computer processing, one containing only ITDs and one containing only ILDs. The results for normal-hearing listeners showed, as in Plomp and Mimpen (1981), a gain of about 10 dB in SRT, when the noise containing both ITDs and ILDs was presented laterally relative to the frontal position. In the same conditions, the noise containing ITDs alone provided a gain of about 5 dB, and the noise containing ILDs alone provided a gain of about 7 dB. Thus, the effects of ITDs and ILDs were not additive. The benefit of the ITD cues was essentially unaffected by simulating a one-sided attenuation of 20 dB at either ear. Also, the effect of the ILD cues was entirely dependent on monaural processing and not on binaural processing per se, since the same gain in intelligibility could be obtained by listening only through the ear with the best SNR. Overall, for a frontal speech signal and a lateral noise masker, the minimum and maximum gains observed for binaural listening compared to monaural listening were 2.5 dB and 13.2 dB respectively. The higher value is the binaural gain compared to monaural listening through the ear with the worst SNR, and the lower value is the binaural gain against the ear with the best SNR.

Bronkhorst and Plomp (1989) extended their experiments to hearing-impaired listeners. These listeners had a 2.5 dB higher SRT than normal-hearing listeners when the speech signal and noise masker were presented from the front, and a 2.6-5.1 dB smaller binaural intelligibility gain when the noise masker was displaced laterally depending on the configuration of the hearing loss. The shortfall in binaural gain for hearing-impaired listeners was mainly due to an inability to take full advantage of ILD cues. This was especially pronounced for listeners with asymmetrical high-frequency hearing losses when the noise source was displaced contralaterally to their best ear. In contrast, the gain in speech intelligibility due to ITD cues was less affected by hearing impairment. It was about 4-5 dB for normal-hearing listeners and listeners with symmetrical losses, but 2.5 dB for listeners with asymmetrical losses. When ITD cues were introduced in a noise already containing ILD cues, the resulting gain was 2-2.5 dB for both groups of hearing-impaired listeners.

Bronkhorst and Plomp (1992) further investigated binaural speech intelligibility in simulated free-field conditions with a frontal speech signal source in the presence of one to six mutually-uncorrelated noise sources located in the horizontal plane in various configurations. Over all conditions, the hearing-impaired listeners needed a 4.2-10 dB better SNR than normal listeners for equal intelligibility. The binaural advantage arising when the noise maskers were displaced from the frontal position to symmetrical or asymmetrical spatial configurations around the listeners varied from 1.5 to 8 dB for normal listeners, and from 1 to 6.5 dB for hearing-impaired listeners. The higher value corresponds to a single masker moved laterally to the side of the listeners, and the lower value corresponds to a configuration of six maskers located symmetrically around the listeners at 60° intervals. Comparison of binaural listening with monaural listening results through the best ear showed a fairly constant binaural advantage of about 3 dB across noise masker configurations and listener groups.

In summary:

- the advantage of binaural over monaural listening in noise is greater for detection than intelligibility tasks;
- only the ITD cues provide a true benefit for speech intelligibility in noise;
- the effects of ILD cues can be fully accounted by monaural SNR considerations alone;
- the maximum speech intelligibility benefit derived from binaural listening over monaural listening through the ear with the best SNR is limited to 4-6 dB under the most favourable conditions (anechoic environment, normal hearing or symmetrical hearing loss, single noise source spatially separated from the speech signal, broadband noise, low overall SNR); and
- the benefit of binaural hearing decreases with increasing signal level above masked threshold, and is very small when the signal is relatively easy to detect (Yost, 1997).

Most experiments described to date have been devoted to measures of selective attention, where the listener is asked to focus on a particular signal source and ignore all others (Yost, 1997). There is very little information on situations of divided attention, where the listener must attend to several or all the sound sources in the environment.

3.0 Binaural technology

The input signals to the hearing system are the sound pressure waves inside the left and right ear canals. Three-dimensional auditory environments or displays could thus be simulated through headphones if the directional transfer functions of the human head and external ears were known. The methods and techniques necessary to create virtual auditory environments are together referred to as binaural technology (e.g., Moller, 1992; Blauert, 1997).

3.1 General methodology

There are two main steps involved in a typical binaural technology application (Moller, 1992). The first is the derivation of the head-related transfer functions (HRTFs) of the listener from binaural measurements. The second is the creation of binaural signals by filtering the desired acoustic source input with the HRTFs, and the playback of these signals to the listener using headphones.

HRTFs measurements:

An important aspect to consider in the derivation of the HRTFs is the selection of a reference position for the binaural measurements. The reference position should allow the recording of all the spatial information available to the ear. Moller (1992) investigated three possible positions: at the eardrum, at the entrance to the open ear canal, and at the entrance to the blocked ear canal. The entrance to the blocked ear canal offered several advantages (Moller, 1992; Moller et al., 1995b). Firstly, the blocked ear canal method is easier to implement because it is less prone to microphone fitting and stability problems, measurement noise, sound field interference and other artifacts. Secondly, recordings at the blocked ear canal entrance are free from individual subject differences in ear canal transmission that are not related to the spatial characteristics of the ear as such. HRTFs derived from the blocked ear canal method possess less interindividual variation than other reference positions. This was demonstrated theoretically (Moller et al., 1996) and verified experimentally (Moller et al., 1995b). In the latter study, blocked ear canal HRTFs measured in 40 human subjects showed a clear common structure with small interindividual variations up to about 8 kHz.

Headphone equalization:

In a practical application, the derivation of HRTFs is not an end in itself. Binaural signals must be created by convolution or filtering with the HRTFs and they must be played back to the listener. However, the electroacoustic transfer function of the headphones contributes to the total sound transmission to the ear, and thus require equalization for the correct playback of binaural signals.

Moller (1992) examined the correction functions required in the headphone equalization step. The first correction compensates for the electroacoustical pressure transfer function of the headphone (or PTF in the terminology of Moller et al., 1995a) from the electrical input terminals of the heaphone to the sound pressure at the reference position. There is no other correction needed when a location in the open ear canal is chosen for the reference position. When the blocked ear canal is used as the reference position, the equalization step also requires an extra correction to account for the different acoustical source impedance loading of the ear when listening through headphones instead in the free-field. This correction term is referred to as the pressure division ratio (PDR) (Moller et al., 1995a). It reduces to unity if the radiation impedance looking outwards from the ear canal entrance is unchanged by fitting of the headphone, or if this impedance is much smaller than the input impedance looking inwards from the ear canal entrance (Moller, 1992). In this case, the headphone is said to provide a free-air equivalent coupling (FEC) to the ear.

Moller et al. (1995a) measured the PTF and PDR functions of 14 commercial headphones at the blocked ear entrance of 40 human subjects. The PTF functions were found to be far from flat for all headphones and, in general, none of the headphones tested was deemed suitable for the playback of binaural signals without proper equalization. All PTF functions also showed considerable interindividual variations, especially above 8 kHz. A blocked ear canal reference position led to smaller interindividual variations than a position in the open ear canal, and was easier to implement from a methodological standpoint. The PDR functions were found to be much smaller than the PTFs for all headphones. In general, headphone constructions closely mounted to the ear canal had larger PDF functions than those mounted further away. Up to 2 kHz, all headphones gave flat PDR functions close to 0 dB. Above 2 kHz, the PDRs showed fluctuations with some degree of interindividual variations but with a common structure for each headphone.

Individualized versus generic binaural signals:

Perfect reproduction of the binaural signals can only be guaranteed only if binaural measurements and headphone equalization steps are realized with the target listener's own ears. Thus, the question of the possible errors introduced by using another subject, or artificial head, in either or both steps of the binaural technique must be considered. Moller et al. (1996) conducted an error analysis of the reproduced binaural signals, and compared the sensitivity of different reference positions to the use of non-individualized binaural measurements and/or headphone equalization. For the smallest possible error, they proposed: (1) a blocked ear canal reference position for the binaural measurements, (2) the use of an FEC headphone, and (3) the use of individualized PTF headphone equalization whether the binaural measurements were individualized or not.

3.2 Psychoacoustical evaluation

Wightman and Kistler (1989) studied the localization performance of 8 subjects over a set of 72 source directions for broadband white noise bursts presented either by loudspeakers in the free field or by headphones. The headphone stimuli were derived from individualized HRTFs and were individually equalized using a reference position at the eardrum of the open ear canal. Overall, the localization performance with headphone stimuli was nearly identical to that of the reference condition in the free field for each subject. The only noticeable differences that emerged were a greater percentage of front/back confusions (almost double) and a slightly poorer perception of elevation in the headphone condition than in the free field.

Moller et al. (1996) studied the localization performance of 8 subjects over 19 source directions and distances for speech stimuli presented either by loudspeakers in a listening room with reverberation time of 0.4 s or by headphones. The headphone conditions were meant to reproduce binaural recordings made in the same room with the same loudspeaker arrangement. These recordings were made at the blocked ear canal of several subjects. Subjects listened to their own recordings (i.e., individualized), or to those of another subject or a mixture of subjects (i.e., generic). The headphones were always individually equalized to the target listener for the localization experiment. The results for the loudspeaker condition and the headphone condition with individualized recordings were not significantly different. However, the headphone condition with generic recordings led to a significantly greater percentage of errors for sources in the median plane (approximately double), including front/back confusions, and a slight increase in the number of distance errors. Out-of-cone errors were very rare in all conditions tested. None of the subjects reported in-the-head perception of localization in any of the headphone conditions, regardless of whether or not the recordings had been individualized.

In the headphone experiments above, the binaural signals were not synchronized with the head movements of the subjects. Subjects were instructed to keep their heads fixed. Head movements may reduce localization ambiguities, especially front/back confusions and within cone-of-confusion errors (Wallach, 1940; Wightman and Kistler, 1999), and may facilitate the perception of externalization (Durlach et al., 1992). Head movements can be taken into account in a binaural technology application by using a head-tracking device to update the location of the source(s) with respect to the listener's head coordinate system (Blauert, 1997). The externalization of signals, and distance perception, may also be greatly facilitated by reflections and reverberant energy in the listening room (Durlach et al., 1992). The simulation of room acoustics for binaural technology applications requires sound-field modelling (Blauert, 1997).

4.0 Integrating Active Noise Reduction and Binaural Technologies

4.1 General concept

Communication headsets with sound attenuation capabilities are often used in situations where an individual must be in contact with others at a remote location while operating in a noisy environment. The most common design is based on a passive circumaural hearing protection device fitted with earphones inside the earcups and a boom microphone in front of the mouth. ANR communication headsets can provide a significantly higher amount of low-frequency attenuation compared to passive headsets. The potential benefits of this additional attenuation are a reduced noise exposure for the user, and improvements in speech intelligibility and signal detection through the communication channels. If the speech signals are spatialized and separated from the environmental noise using binaural technology (Section 3), further improvements in intelligibility and signal detection can be expected (Section 2). In addition, virtual auditory displays and 3D models of the listening space can be created through binaural technology, which can greatly facilitate the

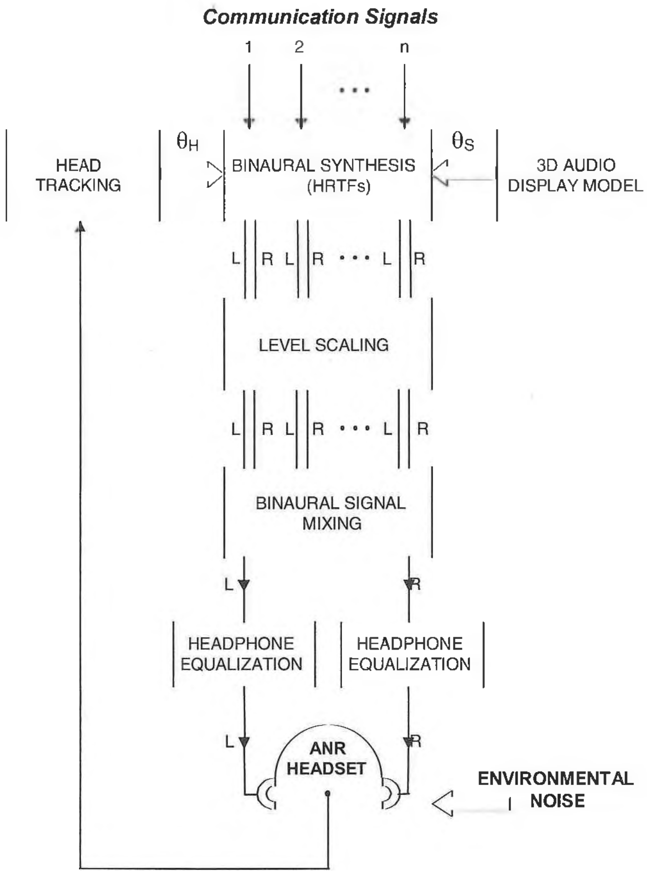


Figure 2: Sketch of an ANR-binaural communication headset

monitoring and interpretation of the various sources of information presented to the listener (Begault, 1993; McKinley et al., 1994; Bronkhorst et al., 1996).

Figure 2 illustrates the main components in a complete system integrating binaural technology to ANR communication headsets. A typical use of such a system would be inside a noisy aircraft cockpit. The different communication signals are individually spatialized on the basis of the HRTFs of the target listener, the desired spatial position θ_S of each signal source, and the current head position θ_H. The coordinates θ_S express the virtual display model to be created for the particular listening task. The right and left ear signals from each spatialized source are then scaled in level to compensate for head movement effects, as appropriate. The signals are then mixed to sum all left and right components. The resulting two signals are equalized for headphone sound transmission, and fed into the left and right communication channels of the ANR headset. The ANR headset, itself, reduces the external noise from the environment with little or no effect on the transmission of the communication signals.

There are several aspects to consider when combining ANR and binaural technologies as in Figure 2. Firstly, there is the hearing protection performance of the ANR device for the given environmental noise conditions. Secondly, there are the electroacoustical characteristics of the communication channels of the device that would be necessary for an adequate reproduction of binaural signals. Thirdly, there are the requirements of the application itself, which determine the complexity of the binaural simulation, and the design of a suitable virtual auditory display.

4.2 ANR devices for hearing protection and speech transmission

ANR technology provides a means of increasing the low-frequency attenuation in communication headsets or hearing protectors for use in high-level environmental noise (Casali and Berger, 1996). A miniature microphone housed within the earcup samples the incoming waveform. An inverted copy is created and added to the original for the purpose of cancellation. Components of the two waveforms that are out-of-phase will cancel, thereby reducing the overall sound pressure inside the earcup. ANR systems mounted on earmuffs are currently limited to frequencies below 0.5-1 kHz, where they add to the passive attenuation provided by the earcup (McKinley et al., 1996). Attenuation at higher frequencies is achieved by passive means only. Maximum active low-frequency attenuation in the order of 10-20 dB has been measured around 0.125-0.25 kHz over the passive mode (McKinley et al., 1996; Abel and Spencer, 1997; Abel and Giguère, 1997).

The additional low-frequency noise reduction achieved by ANR headsets over passive devices points to improvements in auditory perception for signals transmitted through the communication channels. Objective predictions based on the Articulation Index (Nixon et al., 1992) and the Speech Transmission Index (Steeneken and Verhave, 1996) procedures have demonstrated the speech intelligibility gains that can be realized. However, this has not always been achieved in practice (Gower and Casali, 1994). The frequency response of the communication channels and the effects of the ANR circuitry on the speech transmission quality are important determinants of intelligibility (Steeneken and Verhave, 1996). Several studies have also shown that ANR devices fail to operate when noise levels saturate the ANR circuitry, typically in the range of 120-135 dBA (Brammer et al., 1994). Other characteristics of the device affecting performance are the presence of transients or shut down periods after overload, and the comfort during use (Crabtree, 1996; Steeneken and Verhave, 1996).

4.3 ANR devices for binaural technology

Previous work:

ANR headsets have not been used extensively in binaural technology applications. Ericson and McKinley (1997) from the Armstrong Laboratory at the Wright-Patterson AFB (Ohio) reported a virtual audio presentation of speech communication signals over a Bose AH-1A headset, an ANR headset, configured for binaural operation. The HRTFs from the KEMAR manikin were measured at a 1° spacing in azimuth angle and used to simulate the virtual audio sources. The elevation angle of the sources was maintained fixed in the horizontal plane and distance cues were essentially absent. A head-tracking system measured the orientation of the listener's head and was used to maintain the virtual sources fixed in space.

Currently, the research group at the Armstrong Laboratory is utilizing the blocked ear canal method to derive the HRTFs and headphone equalization functions with the microphone inserted about 2-3 mm inside the canal entrance (McKinley, 1997). The critical factor is the repeatability of the microphone/plug location during the measurements. However, once a consistent fitting of the headset and microphone/plug assembly can be ascertained, they find no particular problems in equalizing their ANR headsets. To facilitate the equalization process, they choose headsets with matched left/right earphone drivers, typically within 2 dB in sensitivity. Tracking and integrating the head movements of the listener into the binaural simulation is found important for the perception of externalization of signals, and to maintain the highest possible speech intelligibility (McKinley, 1997).

Design criteria:

Commercial ANR headsets have not been specifically designed for binaural technology applications. The following criteria are proposed for the selection or design of a suitable ANR headset for experimentation with virtual audio signals. A minimal set of recommended electroacoustic specifications is given.

Listening mode — The ANR headset must support stereo communication signals for dichotic listening as a pre-condition to a binaural technology application.

Volume control — Headsets equipped with a single knob to control the signal volume under both earcups are preferred over headsets equipped with dual controls to independently adjust the volume in the left and right ears. Independent volume control interferes with the correct reproduction of ILDs from the virtual audio sources.

Cross-talk attenuation — The amount of cross-talk attenuation between the left and right communication channels should exceed the maximum possible ILD in the free field plus a safety margin of about 10 dB. The maximum ILD arises for lateral incidence, and is typically 20-25 dB around 5 kHz and in excess of 30-35 dB above 8 kHz (Moller et al., 1995a; Wightman and Kistler, 1997).

Earphone linearity — The earphone frequency response from the electrical input of the communication channel to the sound pressure output in the earcup must be linear over the full dynamic range of signal levels, and must be insensitive to environmental noise conditions. Two conditions are to be met: (1) the sound pressure output of the signal must grow linearly with the electrical input signal under constant environmental noise presented at different levels, and (2) the sound pressure output of the signal must remain constant for a constant electrical input signal under variable environmental noise levels.

Earphone frequency response — The earphone frequency response must be as uniform and smooth as possible to simplify the headphone equalization procedure. The upper frequency limit depends on the location of the virtual audio sources in the intended application. In the case of virtual sources positioned in the horizontal plane and frontally, an upper frequency limit of 4-5 kHz may be adequate. In the case of sources distributed both in front and at the back, in the median plane, or elevated from the horizontal plane, an upper frequency limit of 8 kHz or more may be required.

Coupling to the ear — The headset selected should provide a free-air equivalent coupling (FEC) to the ear, as defined in Moller et al. (1995a), to simplify the headphone equalization process when the blocked ear canal method is used. This could be verified by measuring the pressure difference ratio (PDR) of the headsets (Moller et al., 1995a), or by measuring the radiation impedance looking outwards from the ear canal using an artificial head with and without headset fitted (Schroeter and Poesselt, 1986).

Interaural earphone matching — The interaural amplitude and phase matching in the earphone frequency response is obtained by equalizing independently the left and right sides at the headset. In cases where a single generalized equalization function is required, the left and right earphone transmission should be closely matched in frequency response and sensitivity.

Device selection:

The characteristics of nine commercial ANR headsets were assessed against the proposed design criteria (Abel and Giguère, 1997). The devices surveyed are listed in Table I. The manufacturers differ greatly in their methods of presenting the communication signals through the devices. For example, the Bose Aviation approach (Gauger, 1995) is based on the conventional feedback servosystem where the output, the sound pressure wave inside the earcup, is tracking a desired input, the electrical communication signal, while minimizing interfering noise. Thus, the effect of the ANR feedback loop on the communication signal must be compensated for by an equalization filter to flatten the transmission response. The communication signal of the Telex ANR Headset System is injected electronically just before the earphone transducer, but is subtracted from the sensing microphone output. The communication signal is in effect removed from the ANR feedback loop and its transmission becomes essentially insensitive to the operation of the ANR circuitry. A similar approach is used in the Sennheiser NoiseGard (Crabtree, 1997). The David Clark H1013X uses two earphone transducers, one for the communication signal and one for the ANR cancellation procedure (Crabtree, 1997).

Table I: Analog ANR communication headsets surveyed

Peltor ANR Aviation Headset
Sennheiser NoiseGard
Bose Aviation Headset
Bose Aviation Series II
David Clark DCNC Headset
David Clark H1013X
Telex ANR Headset System
Telex ANR 4000
TechnoFirst NoiseMaster

The survey showed that the most likely candidates for binaural technology applications are the Peltor, Sennheiser and TechnoFirst devices (see Abel and Giguère (1997) for additional technical details). They all support stereophonic listening, have a single control knob for volume control in both earcups, and provide good sound attenuation properties. Unfortunately, the manufacturers' specifications do not provide sufficient information to assess adequately all the technical characteristics necessary for binaural technology on any device. In particular, the amount of cross-talk attenuation, interaural earphone matching and type of coupling to the ear are essentially unspecified. In practice, earmuff-type ANR devices are likely to behave as FEC or near-FEC headphones because of their relatively large earcup volumes that are necessary for good low-frequency passive attenuation. Indeed, Schroeter and Poesselt (1986) found that the radiation impedance looking outwards from the ear canal is essentially unchanged above 0.4 kHz by fitting an earmuff-type hearing protector. Below 0.4 kHz, these hearing protectors do affect the radiation impedance of the ear, but then, this impedance is much smaller than the impedance looking into the ear canal. Under these conditions (Moller, 1992), earmuff-type ANR headsets could be considered FEC.

The commercial devices surveyed in Table I were all based on analog ANR technology. Prototype ANR devices based on digital technology have been tested in research laboratories in the past few years (Pan et al., 1995), and the first commercial digital ANR headsets have been recently introduced (e.g. Telex ANR-1D). Since binaural technology applications are also based on digital signal processing, digital ANR headsets could lead to more completely integrated and compact ANR-binaural systems than analog ANR headsets would allow. A particularly attractive digital ANR design for use with binaural technology is based on adaptive feedforward noise control. The feedforward control structure does not perturb the communication signals, and so offers the potential for higher fidelity reproduction than the commonly used feedback control structure (Brammer and Pan, 1998)

4.4 Aircraft cockpit application

The complexity of the binaural simulation depends on the requirements of the application at hand. In an aircraft cockpit application, very different listening situations could arise. Two extreme scenarios are detailed below.

Simple selective attention task:

In this task, the pilot must focus on the speech of one and only one speaker through the communication channel of the headset, in the presence of the environmental noise in the cockpit. Using binaural technology, the speech communication signal could be externalized and positioned in space to provide an angular separation with the environmental noise. The goal would be either to (1) maximize the speech intelligibility score for a given signal level, or (2) minimize the signal level for a given speech intelligibility score. For listeners with normal hearing or symmetrical hearing losses, a gain at the 50% speech intelligibility level up to about 4-6 dB with respect to diotic listening can be expected under the most favourable noise conditions (Section 2). Under conditions of reverberation, band-limited noise, or multiple noise sources, the speech intelligibility gain will be smaller. For listeners with asymmetrical hearing losses, the speech intelligibility gain due to ITDs is typically half that of normal-hearing listeners. Nonetheless, given the very steep slope of the intelligibility function near the 50% level, typically 15% per dB for sentence material, a gain of only a few decibels could give rise to substantial intelligibility improvements for all classes of listeners, but only in communication systems with low signal-to-noise ratios. It is also under conditions of low SNRs that the greatest benefits of ANR over passive communication headsets are anticipated for speech intelligibility.

In the simple listening application above, there is no localization task involved per se. Thus, the design of the binaural system could be simplified by the use of generic HRTFs (Section 3), particularly if a direction of incidence in the hor-

izontal plane is selected for the virtual speech signals. Likewise, the benefits of synchronizing the binaural signal simulation with the head movements of the user may be minimal in this application, so the binaural signal procesing could be further simplified.
The selection of an optimal direction of incidence for the spatialized communication signal will depend on the characteristics of the environmental noise sound field at the location of the pilot's head. Cockpit noise spectra and levels are dependent on the type of aircraft, and the speed and altitude of the aircraft, among other factors (Rood, 1988). To achieve the maximal binaural speech intelligibility gain, the speech signal should be spatially separated from the noise by about 45° or more. However, several difficulties can arise because there are in general more than one source of noise in an aircraft, and the noise field in a typical cockpit is not free field. Another problem is that earmuff-type devices can severely disrupt the localization cues from external sounds (Abel and Hay, 1996). In practice, the sources of noise are large and distributed in typical aircrafts, and there is minimal or no acoustical treatment in the cockpit. Under these conditions, the environmental noise at the pilot's head could be classified as diffuse or quasi-diffuse, and thus the selection of a speech signal incidence would not be too critical.

An important related aspect to consider is the scaling of the binaural signal level at the ears. In a system where the HRTFs are synchronized with the head movements of the user, the sound exposure arising from the communication signal will vary with the selected direction of incidence. Moreover, if at any time the speech incidence lies outside the median plane, exposure from the speech signal will be asymmetric across the two ears, typically larger on the ipsilateral ear than the contralateral ear. A possible solution to maintain a constant and symmetric exposure is to scale the HRTFs with a direction-dependent gain so that the total speech energy becomes independent of sound incidence and equal in each ear. Another possibility, suggested by Bronkhorst and Plomp's (1988) experiments, is to scale the amplitude spectrum of each HRTF to a common reference amplitude spectrum, such as that corresponding to a frontal incidence, while keeping the phase spectrum intact. These solutions are based on the observation that it is the ITDs alone that provide a true binaural benefit for speech intelligibility (Section 2), and that accurate localization of the speech signal is secondary in this task.

Complex divided attention task:

At the other extreme, in a complex divided attention task, the pilot must attend to several speakers (e.g., co-pilot, pilots in other aircrafts, ground crew, etc.) through the communication channel of the headset, in the presence of cockpit noise. The pilot may also need to be alert to various visual targets in his/her environment that are cued to characteristic warning sounds. In this case, the speech signals from the different speakers and the other sounds to attend to would be externalized using binaural technology and positioned in space on the basis of ergonomic considerations. The goal would be to provide the user with a model of his/her complex acoustic environment in order to facilitate the interpretation of the various sources of information (Figure 2). The actual display design would depend on the specific demands placed on the user (Mack et al., 1998).

Localization errors, particularly elevation errors and front/back confusions, would be very detrimental in this application, because of the need to maintain a consistent spatial model of the environment. To maximize localization performance, this application would likely require individualized HRTFs and headphone equalization. It would also be highly desirable for ergonomic considerations and for optimizing accuracy in sound localization to track the head movements of the user and update the binaural simulation synchronously, so that the acoustic sources of information remained fixed in space. Because of the localization needs, both the ITD and ILD binaural cues are important in this task. This prevents manipulations of the amplitude spectrum of the HRTFs, other than applying a direction-dependent gain to each pair of left/right HRTFs.

5.0 Conclusions

This article reviewed the fundamental research and several practical aspects relevant to the integration of ANR and binaural technologies in the design of improved communication headsets, with particular attention to an aircraft cockpit application. ANR technology can reduces the interfering noise from the environment. Binaural technology allows the creation of 3D auditory displays to transfer coincident messages to different spatial positions.
In a simple selective attention task, the requirements of the binaural simulation are not very stringent as far as localization performance and the tracking of head movements are concerned, but careful consideration must be given to the direction of incidence of the environmental noise and to the scaling of the binaural signal levels. Under the most favorable conditions, a speech intelligibility improvement equivalent to a gain of 4-6 dB in SNR can be expected for this task with a ANR-binaural headset system over a system with ANR capabilities alone.

In a complex divided attention task, the binaural simulation system must provide for accurate sound localization performance, but the scaling of the binaural signals is less critical. The greatest benefits of 3D virtual auditory displays may well be found for this type of task, when there are more than two speakers or signals to attend to simultaneously (Ericson and McKinley, 1997). However, more fundamental research is needed to quantify the real advantage gained in terms of improved speech intelligibility, total information transfer, increased situational awareness or reduced workload fatigue (Begault, 1993; McKinley et al., 1994; Bronkhorst et al., 1996)

Commercial analog ANR communication headsets are not designed for binaural applications and would require extensive testing before making firm recommendations on specific devices. A list of features relevant to binaural technology includes the listening mode and volume control options, the cross-talk attenuation across the two channels, the earphone linearity, the earphone frequency response and degree of interaural matching, and the type of coupling to the ear. Newly developped digital ANR headsets may faciltate the integration with binaural technology.

6.0 Acknowledgements

This research was supported by a contract from the Defence and Civil Institute of Environmental Medicine (Canada). The comments of Brian Crabtree and Sharon McFadden on earlier drafts of the contract report were greatly appreciated. In early 1997, the authors benefited from site visits to several research laboratories. The authors are particularly indebted to Adelbert W. Bronkhorst, Herman J.M. Steeneken, J.A. Verhave, Nöel Château, Guido F. Smoorenburg, Jens Blauert, Klaus Hartung, Jörg Sahrhage, Anthony J. Brammer, Richard L. McKinley, Charles W.Nixon, Timothy R. Anderson and Robert H. Gilkey.

7.0 References

Abel, S.M. and Hay, V.H. (1996). "Sound localization: The interactive effects of aging, hearing loss and hearing protection," Scand. Audiol. 25, 3-12.

Abel, S.M. and Giguère, C. (1997). "A review of the effect of hearing protective devices on auditory perception: The integration of active noise reduction and binaural technologies," Contract Report No. W7711-6-7316/001 SRV, Defence and Civil Institute of Environmental Medicine, Department of National Defence (Canada), 50 pp.

Abel, S.M. and Spencer, D.L. (1997). "Active noise reduction vs. conventional hearing protection: The relative benefits for normal-hearing and impaired listeners," Scand. Audiol. 26, 155-167.

Begault, D.R. (1993).). "Head-up auditory displays for traffic collision avoidance system advisories: A preliminary investigation," Human Factors 35(4), 707-717.

Blauert, J. (1997). Spatial Hearing: The Psychophysics of Human Sound Localization, Revised Edition (MIT Press, Cambridge).

Brammer, A.J., Ryan, J.G., Zera, J. and Pan, G.J. (1994). "Attenuation of active noise reducing headset at high sound pressure." Vth Int. Symposium on the Effects of Noise on Hearing, Gothenburg, Sweden (May).

Brammer, A.J. and Pan, G.J. (1998). "Opportunities for active noise control in communication headsets," Canadian Acoustics 26(3), 32-33.

Bronkhorst, A.W. and Plomp, R. (1988). "The effect of head-induced interaural time and level differences on speech intelligibility in noise," J. Acoust. Soc. Am. 83, 1508-1516.

Bronkhorst, A.W. and Plomp, R. (1989). "Binaural speech intelligibility in noise for hearing-impaired listeners," J. Acoust. Soc. Am. 86, 1374-1383.

Bronkhorst, A.W. and Plomp, R. (1992). "Effect of multiple speechlike maskers on binaural speech recognition in normal and impaired hearing," J. Acoust. Soc. Am. 92, 3132-3139.

Bronkhorst, A.W. Veltman, J.A. and van Breda, L. (1996). "Application of a three-dimensional auditory display in a flight task," Human Factors 38(1), 23-33.

Casali, J.G. and Berger, E.H. (1996). "Technology advancements in hearing protection circa 1995: Active noise reduction, frequency/amplitude-sensitivity, and uniform attenuation," Am. Ind. Hyg. Assoc. J. 57, 175-185.

Cherry, E.C. (1953). "Some experiments on the recognition of speech, with one and with two ears," J. Acoust. Soc. Am. 25, 975-979.

Crabtree, R.B. (1996). "Constraints in the application of personal active noise reduction systems." Proceedings of the 82nd AGARD Meeting, Copenhagen, Denmark (October).

Crabtree, R.B. (1997). Personal communication.

Durlach, N.I., Rigopulos, A., Pang, X.D., Woods, W.S., Kulkarni, A., Colburn, H.S. and Wenzel, E.M. (1992). "On the externalization of auditory images," Presence 1, 251-257.

Ericson, M.A. and McKinley, R.L. (1997). "The intelligibility of multiple talkers separated spatially in noise," In: Binaural and Spatial Hearing in Real and Virtual Environments, edited by R.H. Gilkey and T.R. Anderson (Erlbaum, Mahwah, NJ), pp. 701-724.

Gauger, D. (1995). "Active noise cancellation for military aircraft and vehicles." Military Med. 19, 2-7.

Gower, D.W. Jr., and Casali, J.G. (1994). "Speech intelligibility and protective effectiveness of selected active noise reduction and conventional communication headsets," Human Factors 36(2), 350-367.

Januška, I. (1983). "Influence of the mutual position of speech and noise sources on speech intelligibility in noise," In: Proceedings of the 11th Int. Congress on Acoustics, Paris, France, Vol. 4, pp. 53-56.

Levitt, H. and Rabiner, L.R. (1967a). "Binaural release from masking for speech and gain in intelligibility," J. Acoust. Soc. Am. 42, 601-608.

Levitt, H. and Rabiner, L.R. (1967b). "Predicting binaural gain in intelligibility and release from masking for speech," J. Acoust. Soc. Am. 42, 820-829.

Mack, C.I., Churchill, L.L. and Arrabito, G.R. (1998). "Results of a survey on audio and visual warning systems for military helicopters, " Canadian Acoustics 26(3), 42-43.

McKinley, R.L, Erickson, M.A. and D'Angelo, W.R. (1994). "3-Dimensional auditory displays: Development, Applications, and Performance," Aviation, Space, and Environmental Medecine, May, A31-A38.

McKinley, R.L., Steuver, J.W. and Nixon, C.W. (1996). "Estimated reductions in noise-induced hearing loss by application of ANR headsets," In: Scientific Basis of Noise-Induced Hearing Loss, edited by A. Axelsson, H. Borchgrevink, R.P. Hamernik, P.-A. Hellstrom, D. Henderson and R.J. Salvi (Thieme, New York), pp. 347-360.

McKinley, R.L. (1997). Personal communication.

Moller, H. (1992). "Fundamentals of binaural hearing," Appl. Acoust. 36, 171-218.

Moller, H., Hammershoi, D., Jensen, C.B. and Sorensen, M.F. (1995a). "Transfer characteristics of headphones measured on human ears," J. Audio Eng. Soc. 43, 203-217.

Moller, H., Sorensen, M.F., Hammershoi, D. and Jensen, C.B. (1995b). "Head-related transfer functions of human subjects," J. Audio Eng. Soc. 43, 300-321.

Moller, H., Sorensen, M.F., Jensen, C.B. and Hammershoi, D. (1996). "Binaural technique: Do we need individual recordings?" J. Audio Eng. Soc. 44, 451-469.

Nixon, C.W., McKinley, R.L., and Steuver, J.W. (1992). "Performance of active noise reduction headsets," In: Noise-Induced Hearing Loss, edited by A.L. Dancer, D. Henderson, R.J. Salvi, and R.P. Hammernik (Mosby, St. Louis), pp. 389-400.

Pan, G.J., Brammer, A.J., Zera, J. and Goubran, R. (1995). "Application of adaptive feed-forward active noise control to a circumaural hearing protector," In: Proceedings of Active 95, Newport Beach, California (July), pp. 1319-1326.

Plomp, R. and Mimpen, A.M. (1981). "Effect of the orientation of the speaker's head and the azimuth of a noise source on the speech-reception threshold for sentences," Acustica 48, 325-328.

Rood, G.M. (1988). "Noise and communication", In: Aviation Medicine, 2nd ed., edited by J. Ernstig, and P. King (Butterworths, London), pp. 353-382.

Santon, F. (1986). "Intelligibilité, de la parole masquée et rôle de la direction de la source masquante," Acustica 61, 67-74.

Santon, F. (1987). "Détection d'un son pur dans un bruit masquant suivant l'angle d'incidence du bruit. Relation avec le seuil de réception de la parole," Acustica 63, 222-228.

Schroeter, J. and Poesselt, C. (1986). "The use of acoustical test fixtures for the measurement of hearing protector attenuation. Part II: Modeling the external ear, simulating bone conduction, and comparing test fixture and real-ear data," J. Acoust. Soc. Am. 80, 505-527.

Shaw, E.A.G. (1974). "Transformation of sound pressure level from the free field to the eardrum presented in the horizontal plane," J. Acoust. Soc. Am. 56, 1848-1861.

Steeneken, H.J.M. and Verhave, J.A. (1996). "Personal active noise reduction with integrated speech communication devices: development and assessment," Proceedings of the 82nd AGARD Meeting, Copenhagen, Denmark (October).

Wallach, H. (1940). "The role of head movements and vestibular and visual cues in sound localization," J. Exp. Psychol. 27, 339-368.

Wightman, F.L. and Kistler, D.J. (1989). "Headphone simulation of free-field listening. II: Psychophysical validation," J. Acoust. Soc. Am. 85, 868-878.

Wightman, F.L. and Kistler, D.J. (1997). "Factors affecting the relative salience of sound localization cues," In: Binaural and Spatial Hearing in Real and Virtual Environments, edited by R.H. Gilkey, and T.R. Anderson (Erlbaum, Mahwah, NJ), pp. 1-23.

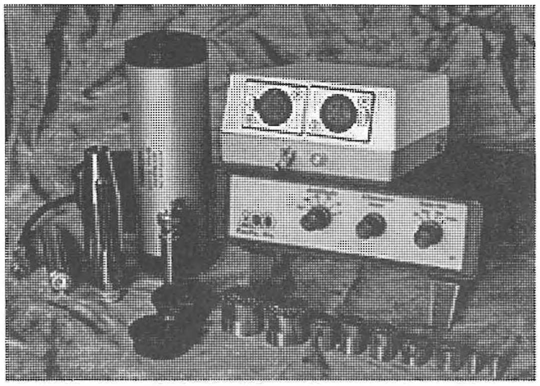
Wightman, F.L. and Kistler, D.J. (1999).). "Resolution of front-back ambiguity in spatial hearing by listener and source movement," J. Acoust. Soc. Am. 105, 2841-2853.

Yost, W.A. (1997). "The cocktail party problem: forty years later." In: Binaural and Spatial Hearing in Real and Virtual Environments, edited by R.H. Gilkey and T.R. Anderson (Erlbaum, Mahwah, NJ), pp. 329-347.

ACO Pacific, Inc.

2604 Read Ave.
Belmont, CA 94002 U.S.A.
Tel:650-595-8588 FAX:650-591-2891
e-mail: acopac@acopacific.com

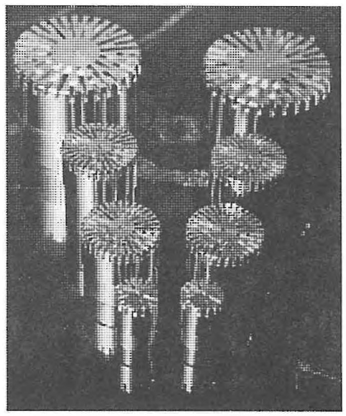
Established in 1978 ACO Pacific, Inc. is chartered to serve the needs of both End-users and Equipment Manufacturers. Headquartered in the San Francisco Bay Area, ACO Pacific, Inc. has manufacturing facilities and suppliers both in California and internationally.



Our OEM relationships include Hewlett Packard, Audio Precision, SPS, DRA Labs and many other large and small instrumentation and process control manufacturers. Our End-user customers include: manufacturers of loudspeakers; computers and peripherals; heavy equipment; automobiles and automotive parts - tires, brakes, engines; universities; aerospace; and an alphabet soup of government agencies US and foreign.

ACO Pacific's Product Family

Measurement Microphones
Microphone Preamplifiers
Microphone Power Supplies
SPL Calibrators
Simple Intensity™ Sound Intensity Systems
Very Random™ Noise Generators



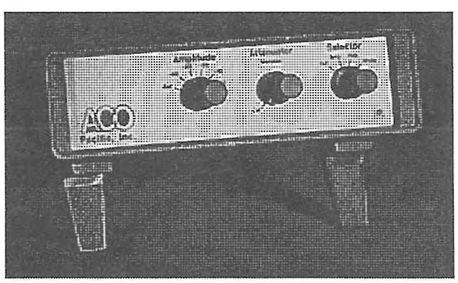
The "Alternative"

Type 1 Measurement Microphones
1, 1/2 and 1/4 Inch Models
Stainless Steel
and
Titanium Diaphragms
Quartz Insulators
Frequency Responses to 120 kHz
Noise Floors below 10 dBA
Meets or Exceeds IEC and ANSI Standards

Model 3024

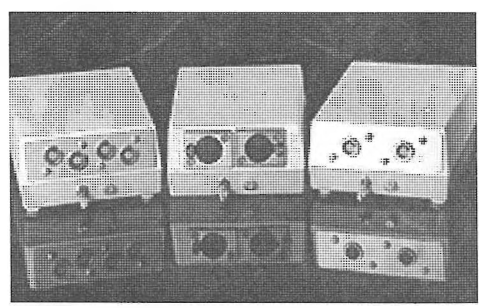
Very Random™ Noise Generator

Pink and White Noise, 1kHz Sine Outputs
1.6 Hz to 39 kHz (-3dB)
Portable - Battery and AC Power



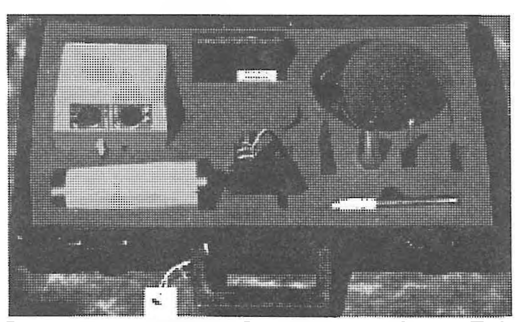
ACOustical Interface™

Precision Microphone Power Supplies
2 and 4 Channels - Gain Available
XLR and Lemo™ Connectors



PS9200KIT

Includes: PS9200 (XLR) Power Supply
AC Adaptor
4012 Preamplifier w/CA4012-5 Cable
Selection of 1/2 Inch Type 1 Microphone
WS1 - 3 inch Windscreen
SC1 Die Cut Storage Case (SC2 optional)
Options: 511E SPL Calibrator (shown)
"G" Gain Stage



ACOustics Begins With ACO™