Virtual Acoustics, Aeronautics, and Communications*

DURAND R. BEGAULT, *AES Member*

Human Factors Research and Technology Division, NASA Ames Research Center, Moffett Field, CA 94035-1000, USA

An optimal approach to auditory display design for commercial aircraft would utilize both spatialized (3-D) audio techniques and active noise cancellation for safer operations. Results from several aircraft simulator studies conducted at NASA Ames Research Center are reviewed, including Traffic alert and Collision Avoidance System (TCAS) warnings, spoken orientation "beacons" for gate identification and collision avoidance on the ground, and hardware for improved speech intelligibility. The implications of hearing loss among pilots is also considered.

0 INTRODUCTION

There is a dichotomy that should not exist between the standard of audio quality we demand for entertainment and the one used in high-stress human interfaces. Consider that a pilot can commute to work in a modern car with the advantages of full-bandwidth stereo digital sound in a quiet environment. In the cockpit, the pilot is within an audio environment that is held to a much different and, arguably, lower standard than the audio environment within a car. The environment is much noisier; the communication and audio warning systems (collectively termed the auditory display) are based on a low-fidelity, nonintegrated design, as summarized in Table 1.

The low-fidelity aspect of the communication technology is inherited from telephony, where "acceptable" intelligibility of a single channel of sound fulfilled the design criteria. For example, one commonly used headset (insert-ear type) especially popular within a major U.S. air carrier was originally designed for telephone operators in 1961. It has an effective bandwidth of approximately 0.3–3.5 kHz. By nonintegrated design is meant that sounds are distributed between headsets and loudspeakers, and that the location of the loudspeakers (and consequently the sound source) is not by design but dictated by available space. As a result, the overall signal at the pilot's ears—the sum and individual characteristics of the signals, relative to background noise—cannot be predicted. An integrated audio display by contrast allows the human factors engineer to optimally control all types of audio in its presentation to the user.

The overall goal of engineering these integrated displays, based on research both in the laboratory and in the simulator, is to make commercial aviation operations safer (for both the flight crew and the passengers) and more efficient. In this engineering report an approach to an integrated audio display is described, along with a review of some of the experiments conducted at NASA Ames Research Center from 1989 to 1997.

1 INTEGRATED AUDIO DISPLAYS AND 3-D AUDIO

By delivering both communications and warning signals through supraaural headsets, it is possible to predict and control the relative signal levels at the ear relative to noise. With stereo headsets, the perceived spatial location of signals can be predicted and varied using spatialized (3-D) audio techniques. Using supraaural (on the ear) headsets, it is also possible to include active noise cancellation. In commercial jet airliners it is not necessary to include the passive noise reduction of relatively heavy circumaural headsets. Supraaural headsets are preferred because they allow intercockpit communication to occur without an intercom system linking the pilot and the first officer. Finally, the masking effects of noise are minimized by eliminating the use of loudspeakers for warning signals and using headphone delivery instead. This allows warning signals to be designed to be noticeable by virtue of factors other than loudness

* Presented at the 101st Convention of the Audio Engineering Society, Los Angeles, CA, 1996 November 8–11; revised 1998 March 9.

Table 1. Comparison of audio standard for pilot's car (entertainment audio) and cockpit (high-stress, high-workload human interface).

Criteria	Pilot's Home/Car Audio System	Cockpit Audio Display
Factors driving standard of audio quality	Immersion; state-of-the-art technology; realism of virtual audio; reconfigurable	Intelligibility; cost; dispersed component design; hardware reliability; replacement; time-proven performance
Frequency range	Human hearing (20 Hz to 20 kHz)	Intelligible speech (200 Hz to 5 kHz)
Typical dynamic range of signals	≈40–60 dB (music)	≈10–20 dB (compressed speech)
Environmental context	Quiet, acoustically treated for low background noise	Noisy enclosure with multiple signals from dispersed locations
Relationship of multiple audio sources	Integrated within a single audio display design	Dispersed between many locations and different devices
Transducers	High-quality loudspeakers	Typically, low-quality communication headsets

(relative to an assumed receiver position). Loud, "obnoxious" loudspeaker alarms can cause a "fright" reaction as well as mask desired communications.

The spatial processing of a warning or communication signal results in a two-channel binaural signal for each desired location that can be delivered over a headset. (See, for example, [1] for more information on 3-D audio techniques.) 3-D audio techniques impose the interaural spectral, intensity, and time differences of head-related transfer functions (HRTFs). These can be of the individual pilot who has been measured previously, or of a "good localizer" for whom nonindividualized localization data had been gathered [2]. Warning signals can be "prespatialized" into two-channel binaural versions and digitally stored for real-time playback. Communication signals on the other hand must be processed in real time.

The virtual audio locations for each signal stream are guided by one or more of the following criteria: 1) optimizing intelligibility among multiple signals; 2) matching the virtual location of a signal to the actual location of an external object that poses a safety hazard; 3) matching the virtual location to the location of a flight system, such as the engines; and 4) displaying information according to priority. This last criterion refers to the use of a special virtual location for messages of the highest priority, for example, in the center of the head. Fig. 1 summarizes these criteria in terms of psychoacoustics and how they are combined into an integrated auditory display. Binaural localization refers to the ability to predict human localization of virtual sound sources. This ability is utilized for "aurally guided visual search," where the head and eyes are guided toward a direction by an aural cue, or for situational awareness of spatial location. The "cocktail party" effect refers here to the binaural advantage over one-ear listening for intelligibility of a signal against noise, and the ability to discriminate or "stream" multiple signals through selective attention [3]–[5]. Although the cocktail party effect is a result of binaural localization, note that an active localization of the sound sources is not necessary for the binaural intelligibility advantage to help separate signal from noise. Active noise cancellation allows audio signals to be played back at a relatively lower amplitude, and may help to mitigate noise exposure levels and conserve hearing.

The method for 3-D audio presentation of auditory warning and guidance signals is based on non-real-time

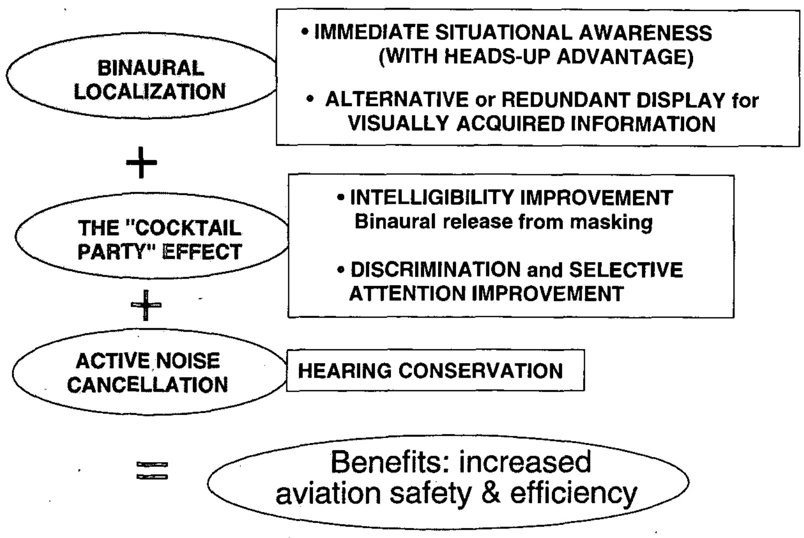


Fig. 1. Overview of benefits of integrated auditory display.

digital filtering of a warning signal at selected azimuths (for example, 0° and left 30°, 60°, and 90°) and as many as three different elevations (0° and ±45°). Fig. 2 shows the relationship between visual and auditory stimuli for a typical display. The warning signal typically consists of two or three iterations of a broad-band electronic "prealert" played at a relatively low volume, followed by a spoken alert. Fig. 3 shows an example from one of the three Traffic alert and Collision and Avoidance System (TCAS) studies conducted at NASA Ames.

The spoken "traffic–traffic" is used for the TCAS warning. In ground operations studies, we have used instructions (such as "hold short") or taxiway turn-off names (such as "P4") as the alert. The purpose of the prealert is to allow a maximal localization advantage (relative to speech) through the use of a broad-band sound source with a rapid amplitude envelope onset. The frequency content, sequencing, and duration of the prealert signal were designed to meet ISO standard 7731 for auditory danger signals [6]. Fig. 4 shows a sonogram of a prealert designed for an upcoming ground collision avoidance study, consisting of three tones. Note the presence of broad-band spectral energy up to 11 kHz.

No head tracking was used, as in the study by Sorkin, Wightman, Kistler, and Elvers [7]. Spatial positions were instead referenced to the nose of the plane. Head tracking would be advantageous if the 3-D audio cuing needed to be referenced to the pilot's head position, or if sound source reversals were an issue, since head

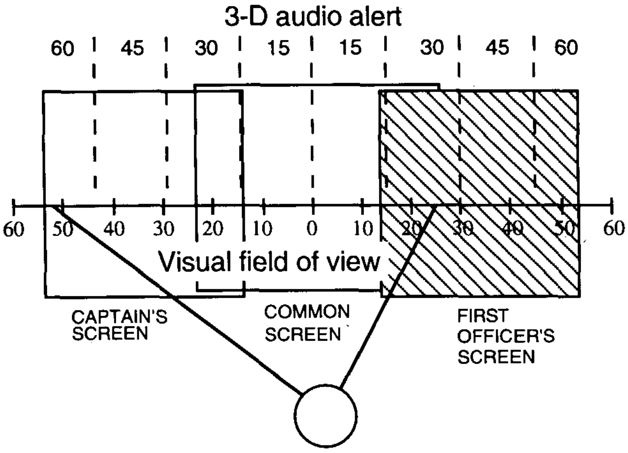


Fig. 2. Horizontal field-of-view in simulator from perspective of left seat (captain's position). The corresponding azimuth position of the 3-D sound cue is shown above for the second TCAS experiment. A different positional scheme was used in the first TCAS experiment; the visual range was divided differently, and audio azimuths were exaggerated by up to a factor of 2.

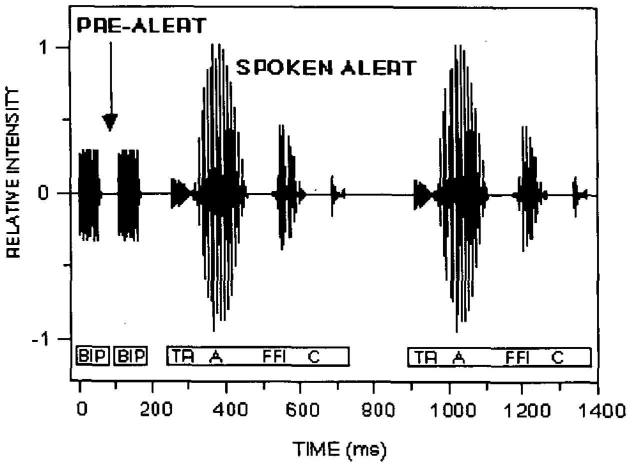


Fig. 3. Arrangement of prealert and voice alert used in second TCAS study. Prealert is composed of a broad-band signal; note its lower amplitude level relative to the speech signal.

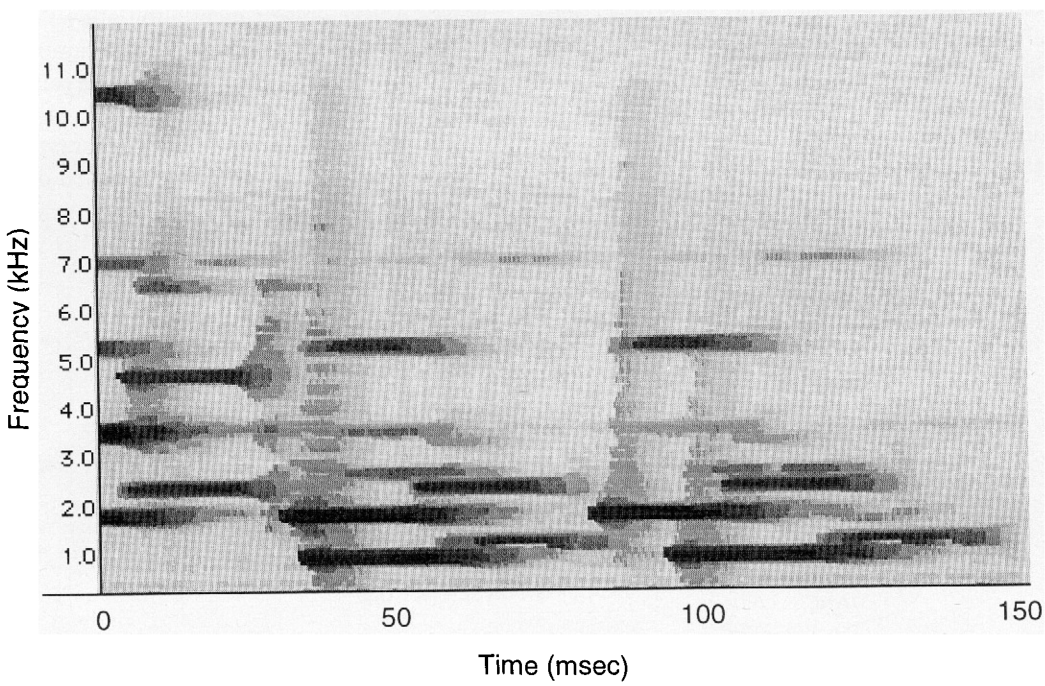


Fig. 4. Sonogram of three-tone prealert signal, designed for optimal localization and recognition. First tone is an octave higher than second and third tones. Note the presence of significant spectral energy across the entire spectrum up to 11 kHz, and the rapid rise time of the spectral envelope. Recognizable broad-band tones such as these can be played over headphones at lower sound pressure levels than band-limited alarms or speech, but with equivalent loudness.

tracking helps to eliminate virtual source reversals [8]. But reversals were never mentioned as a problem, probably because the cues related to visual objects in the front hemisphere (−90° to +90° azimuth). Pilots have reported informally in postexperiment interviews that they are comfortable with an "aircraft-centric" as opposed to egocentric frame of reference. Finally, magnetic tracker information can be distorted by the metal surroundings of the cockpit.

2 AURALLY GUIDED VISUAL SEARCH

Three experiments have been conducted to determine the advantage, if any, of 3-D audio techniques for acquiring an out-the-window aircraft that would normally activate a TCAS advisory warning. We summarize two of the experiments here. The data for the third experiment will appear in a future publication.

Commercial aircraft in the United States are equipped with TCAS transponders that allow other similarly equipped aircraft to see their location on a radar map display. The system also gives instructions when a certain safety envelope is violated, for example, "climb . . . climb." Currently a TCAS advisory warning activates a verbal announcement over the cockpit loudspeaker, a loud "traffic . . . traffic." The pilot determines the location by both looking out the window and referring to a visual radar display map. The use of 3-D sound described below was to reduce the time to locate the traffic visually by use of aural guidance, in place of the head-down time necessary to find the traffic on the radar display.

The experiments took place at NASA Ames' Crew–Vehicle Systems Research Facility under full-mission simulation conditions, within a generic "glass cockpit" flight simulator. Scenario software generated multiple visual targets along a flight path from San Francisco to Los Angeles. Nighttime simulations were used, with traffic at a distance of 2–3 miles. Both acquisition time and the number of targets captured were dependent variables. No significant difference was found in the number of targets located between conditions in any of the studies.

Twelve commercial airline crews were tested in the first TCAS experiment [9]. The stimuli consisted of the prealert described earlier and the words "traffic . . . traffic," corresponding to a TCAS aural advisory. A between-subjects design was used, with half the crews hearing one-ear (monophonic) playback of the advisory, and the other half receiving spatialized cues over stereo headphones. No TCAS map display was supplied for either group. Only seven positions were used for azimuth cuing (corresponding to clock positions from 9 to 3 o'clock), all at eye-level elevation. In addition, the spatialized audio stimuli were exaggerated in relationship to the visual stimuli by a factor of up to 2 (for example, visual targets at 15° azimuth would correspond to spatialized stimuli at 30° azimuth). Results of the study found a significant reduction in visual acquisition time when spatialized sound was used to guide head direction (4.7 versus 2.5 s).

In the second TCAS experiment, five crews used a head-down visual map display with standard TCAS symbology, and five crews used 3-D audio TCAS presentation with no map display [10]. Results showed a significant difference in target acquisition time between the two conditions, favoring the 3-D audio TCAS condition by 500 ms (2.6 versus 2.1 s). The first and second TCAS experiments are not directly comparable for the following reasons. First, in this second experiment there was no exaggeration factor between audio and visual stimuli (see Fig. 2). Second, three categories of elevation cues were provided: "eye level," "above" (alert spatialized to 36° above eye level), and "below" (36° below eye level). Third, a prealert was used (Fig. 3). Finally, the position of the out-the-window targets was held constant, independent of the movement of the pilot's aircraft.

The time advantage may simply be due to the fact that pilots can keep their heads looking out the window for traffic without needing to move their heads downward to the visual display and then back up. But the advantage of aurally guided visual search is in line with previous results, where stimulus presentation occurred under more controlled conditions [11]. Although 500 ms may appear to be only a modest improvement, it does suggest that, in an operational setting, an aural 3-D TCAS display may be desirable in addition to a standard TCAS display. This was evaluated in our third TCAS experiment, using a within-crew design. The same combination of prealert and spoken alert was used as in the second experiment, but elevation cues were not supplied. Multivariate data analysis and postexperiment interviews from the second experiment suggested that elevation cues were not effectively utilized by the pilots; a verbal cue would probably have been better (for example, "traffic, high"). Successful perception of elevation cues from HRTFs tends to be more difficult than azimuth or externalization cues [12].

As shown in Fig. 2, the field of view in the simulator is such that the person sitting on the left side cannot see beyond 25° to the right, and the person on the right side cannot see beyond 25° to the left. It may be that the spatial auditory cue was used for determining whether the pilot or first officer conducted the visual search, in order to transcend the limitations of the simulator environment. This is indicative of a cooperative task delegation procedure for effective management of the crew workload, and is probably relevant to the actual field of view in an actual aircraft. The fact that an advantage was found with both exaggerated and unexaggerated azimuth cues suggests that the audio cue might have been utilized for determining a general head direction (left, front, right) for localization that was subsequently refined by visual search.

3 GROUND OPERATIONS

In future airport ground operations it is possible that the locations of aircraft, gates, and taxiway markers in the runway–taxiway area are identifiable with global positioning sensor (GPS) equipment. This information can be collected at a central computer and then transmitted to both the ground control tower and the aircraft. Such a system may be especially useful under low-visibility conditions, which currently hamper airport efficiency and safety. The goal of the study described in [13] was to link this information to 3-D auditory cues, since keeping the head directed out the window is especially important under low-visibility conditions.

In this study, the same techniques as those used for TCAS advisories were applied to a prototype 3-D audio ground collision avoidance warning system (GCAW). An alarm was designed for alerting pilots to the direction of a potential incursion (see Fig. 5). We hypothesized that there would be a significant preference for such a system to be included in the flight deck. The dependent variable for the GCAW study was a measurement of preference on a Likert scale questionnaire, given during the pilot debriefing session. Another objective of this study was to determine ground taxi time from landing to the gate under aided and unaided conditions. The aided condition featured a 3-D audio guidance system for orientation and guidance that announced specific taxiway turnoffs on the route. We hypothesized that this system would reduce the time required for taxiing to the gate significantly, compared to the unaided condition, where the crew is dependent solely on a paper map and/or radio communications. The dependent variable for the 3-D audio guidance system was the time necessary to complete the taxi route under aided and unaided conditions.

The experiment was conducted within a 747-400 simulator at NASA Ames; we used 12 crews from a U.S. carrier as subjects. A total of seven routes (six orientation routes and one incursion route) were designed based on normal routes used at O'Hare Airport (Chicago). Each crew taxied three of the six orientation routes twice (once with and once without the guidance system) to enable a within-subject evaluation. The order and the assignment of routes were randomized such that each route was run an equal number of times by all crews. In addition each crew ran the same seventh route, which included a potential conflict from another aircraft, to evaluate preference for GCAW. To ensure a realistic condition that included the element of surprise, the crews had no previous knowledge of the total number of routes or which route included a potential incursion.

A prealert signal was synthesized for the GCAW system that was designed to be noticeable without being as loud as typical alarms, through 1) use of significant frequency energy from 0.3 to 13 kHz and 2) inclusion of a "transient" (fast rise time) amplitude envelope. A female voice followed the "prealert" signal for guidance information; it would call out taxiway crossing names, such as "A 15."

All GCAW questions yielded significant results in the preference evaluation. Given the possibility of these responses:

strongly disagree___	disagree___	neither agree nor disagree___	agree___	strongly agree___

the following questions were among those that yielded a significant response of "agree" (chi-square test, $\alpha = 0.05$):

- A system using an audio incursion alert like that heard in the last run would be useful for avoiding a potential incursion under low visibility (300-foot runway visual range) conditions.
- The audio incursion alert would also be useful under normal visibility conditions.

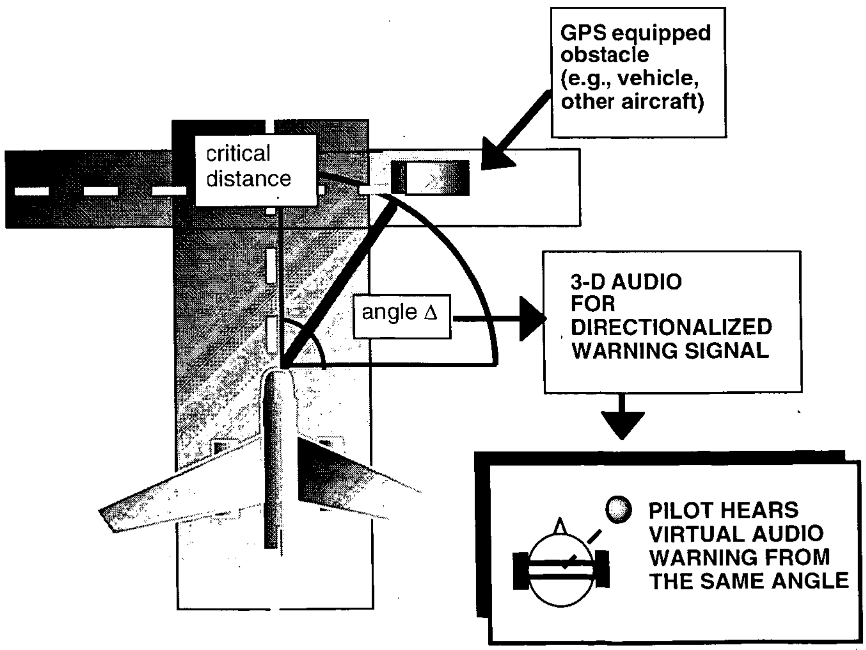


Fig. 5. Illustration of 3-D audio GCAW system. When a critical distance is reached by an oncoming vehicle that has not been "cleared," an alert is sounded from the virtual location of the vehicle through the pilot's headset. The system could also be extended to an audio system within the vehicle. The guidance system worked in a similar way, using a model where strategic points (such as turnoffs) on the runway are indicated via a global positioning sensor (GPS) located near the existing taxiway signs. "RVR 300" refers to a runway visual range of 300-ft visibility.

Interestingly, a significant number of pilots responded "neither agree nor disagree" regarding whether or not the spatial quality, in particular, of the alert helped them to visually locate the other aircraft. This is probably because each crew only experienced one possible incursion, and several were taxiing unusually fast (27–30 knots) due to the lack of visual cues under low visibility. Since our implementation was based on *distance* and not *time until impact* (like TCAS systems), it was possible to thwart the utility of the alert by taxiing at high speeds. Our future implementation of the GCAW system will be redesigned to use an estimation of time to impact to determine when to activate the alarm. Unlike TCAS, the primary task upon receiving a GCAW alert would be to stop the plane—a nonspatial task. In other words, acquisition follows stopping. Whether or not the spatial quality of the alert is useful for GCAW will require evaluation of a redesigned time-until-impact type of system, which is currently being evaluated.

Two questions related to the design of the auditory alert resulted in less decisive group opinion. The data came out midway between "neither agree nor disagree" and "agree" for the question: "An incursion alert using speech would have been preferable to the use of a nonspeech alarm." But some pilots during the debriefing strongly desired a verbal STOP command in conjunction with the alarm (as if a third crew member had noticed the potential incursion). We have included this in our current implementation, in a manner similar to Fig. 3.

No significant difference was found in the time needed to complete taxi routes with or without 3-D audio. Fig. 6 shows the similarity of taxi times under both conditions for each crew. Overall, the mean duration for completion of the taxi routes was 5 min 48 s and 5 min 44 s under unassisted and assisted (3-D audio) conditions, respectively. Table 2 summarizes the mean duration (in seconds) and standard deviations as a function of the individual routes, separated by condition. The large standard deviations indicate the wide variability between crews in running the same route.

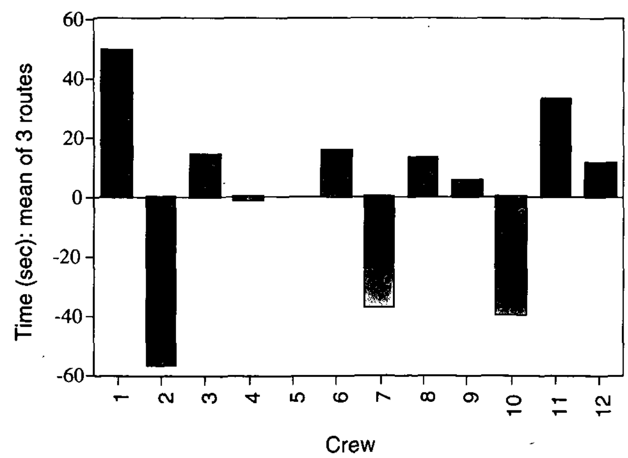


Fig. 6. Reduction in taxi time—data from 3-D audio guidance study. Time difference between unassisted and 3-D audio assisted conditions shown by crew. Positive values indicate an advantage when using 3-D audio. No statistically significant difference was found between unassisted and 3-D audio conditions. Time values were measured from first turnoff of high-speed runway until parking brake was set at last turnoff (pilots were instructed to "hold short").

On the basis of both the data and the comments during the debriefing sessions, an automated audio guidance system is probably best implemented so that one's current location and the next gate are given "on demand" in order to lessen auditory "clutter." On the other hand, the use of a GCAW system for warning of a potential incursion is indicated as desirable by pilots. We are currently running a follow-up experiment using a more sophisticated GCAW system design, which accounts for time until impact. This will enable the evaluation of crew avoidance of collisions in greater detail, including whether or not the spatial aspect of the audio cues contributes anything to improved performance or safety.

4 AUDITORY FEEDBACK AND REDUNDANCY

The application of 3-D auditory cues for feedback and redundancy was evaluated informally in the simulator. No formal data beyond postexperiment interviews have been gathered. Spatialized cues were used to add redundancy to messages related to the spatial orientation of the aircraft's engines; for example, "left engine fire" and "right engine fire" alerts came from left and right. By accessing two perceptual modalities simultaneously with the same information, documented mistakes such as shutting down the left engine when the right engine is nonoperational may be avoided.

Spatialized audio cues were found to be useful for feedback on touch-panel-equipped monitors for "electronic checklists" that replaced the normal paper checklist used in the cockpit [14]. A problem found at the outset was that pilots were uncertain when they had positively engaged the "virtual switch" on the screen. (Similar problems have been reported anecdotally to the author with the aircraft data link screens currently in use.) A solution was to link the touch screen's virtual buttons to very quiet audio feedback cues heard through the headsets, spatialized to the actual location of the

Table 2. Data from 3-D audio guidance study.**

Route	Condition	Mean	Standard Deviation
1	Unassisted	410.2	94.8
	3-D audio	416.1	56.2
2	Unassisted	365.1	52.0
	3-D audio	408.6	72.7
3	Unassisted	384.2	55.6
	3-D audio	374.4	81.5
4	Unassisted	297.9	65.0
	3-D audio	304.6	34.7
5	Unassisted	234.4	39.8
	3-D audio	234.1	53.7
6	Unassisted	394.1	60.7
	3-D audio	369.1	69.6

** Means and standard deviations (seconds) for six different routes tested. No statistically significant difference was found between conditions. Note high standard deviations for each combination of route–condition and similarity of means between conditions for each route.

screen. Audio feedback is used commonly in touch screen monitors that simulate cash-register-type interfaces, for example, as used in the restaurant industry, but the sound palette and fidelity are rather limited and the sounds are loud.

Recordings of actual aircraft switches and several symbolic sounds were spatialized and activated in relation to switch functions on the touch panel, as shown in Fig. 7. The majority of pilots favored the use of auditory feedback, and several users reported an experience akin to synesthesia: that one could feel the virtual switch being engaged in and out.

5 SPEECH INTELLIGIBILITY

As opposed to monotic (one ear) listening—the typical situation in communications operations—binaural listening allows a listener to use head-shadow and binaural interaction advantages simultaneously for unmasking a signal against noise [15]. The binaural advantage for speech intelligibility is well known as the "cocktail party effect" [4]. A cocktail party 3-D audio display consists of real-time spatial processing of incoming communication signals to selected locations. In an ultimate version of an integrated auditory display, information would be presented both to maximize intelligibility of multiple sources and to prioritize information according to its urgency by moving sounds to particular positions.

Approaches to distributing multiple speech sources have been reported on previously [16]–[20]. The Ames Spatial Auditory Display places five communication channels at fixed virtual auditory positions about the listener in real time [21]. This results in an approximately 6-dB improvement in intelligibility, depending on the experimental condition (see Fig. 8). The system has also been found to be advantageous with low-pass-filtered speech typical of most communication systems [22]. We plan to evaluate the system for both pilots and controllers, especially for ground operations where the auditory communication load is at its highest.

Spatial distribution of sounds also enables selective attention, as evidenced by the way a listener can pick out various sections of an orchestra at a concert; this is a type of "auditory streaming" or "auditory grouping" [3], [23]. For radio communications, this results in a type of "hands-free operation." Normally communication personnel bring the volume up for a desired channel in order to hear it over undesired channels, but with a spatial display, one can direct attention to the desired stream at will. But another important advantage is that

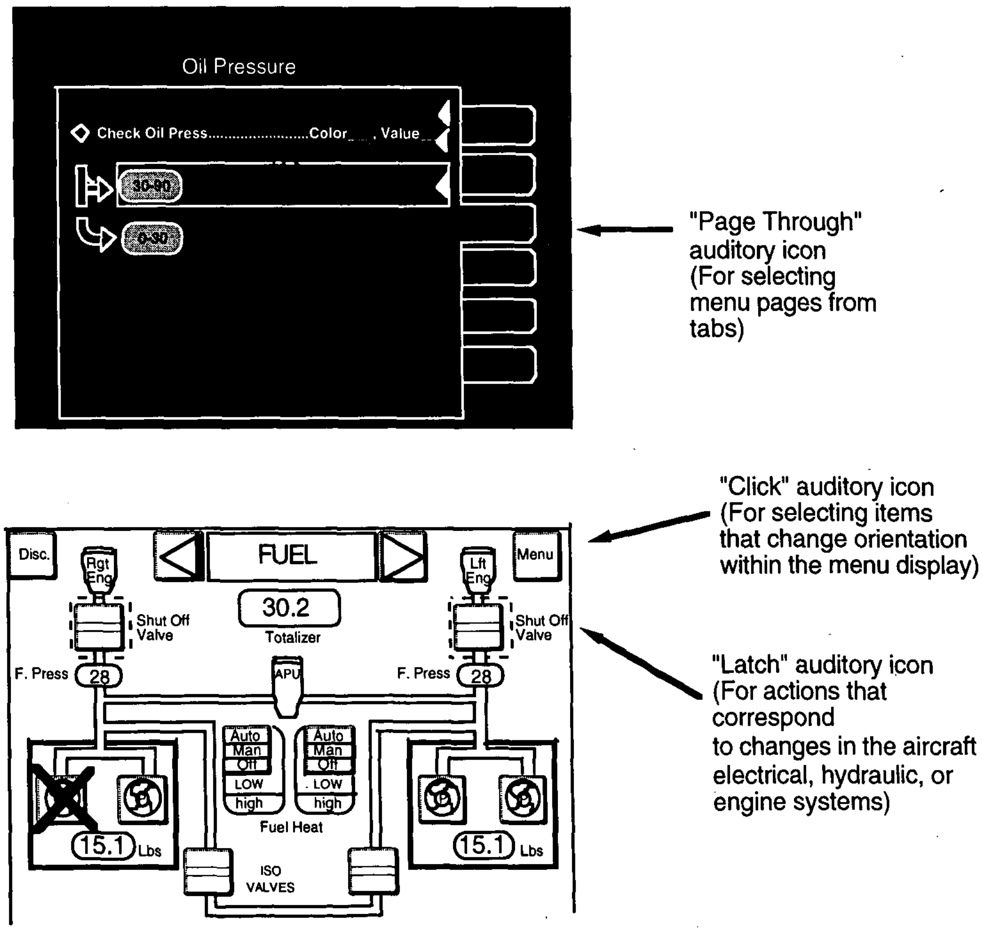


Fig. 7. Touch panel pages from advanced cab cockpit simulator showing how different auditory icons are associated with specific actions. The "latch" icon is distinguished perceptually from the "page through" and "click" icons by use of a louder, more noticeable sound, because it corresponds to an action that changes the operating state of the aircraft. There were two of these screens, located in front of the pilot and first officer; the direction of the cuing was spatialized relative to each person's position to allow awareness of the actions of the other pilot.

the overall intensity level at the headset can remain lower for an equivalent level of intelligibility, an important consideration in light of the stress and auditory fatigue that can occur when one is "on headset" for long periods of time. Lower listening levels over headphones could possibly alleviate raising the intensity of one's own voice (known as the "Lombard reflex"; see [24]), reducing overall fatigue and thereby enhancing both occupational and operational safety.

6 HEARING LOSS

In assessing hearing loss, the greatest focus of the investigation is on speech intelligibility as a function of speech frequency sensitivity between 200 Hz and 4 kHz. Most commercial aircraft cockpits will be within occupational health and safety regulations for the prevention of permanent threshold shifts if the time-weighted average exposure of the pilot is limited in duration; for example, an A-weighted level of 85 dB for an 8-h exposure can cause permanent hearing loss, but not a 4-h exposure. In Fig. 9 measurements of the noise level near the pilot's head are shown for two separate Boeing 737-300 cockpits (time-weighted average over 60 s). The level is about 82 dB when C-weighted (78 dB when A-weighted) at 33 000-ft cruise, and increasing by about 5 dB during descent (unfortunately, no spectral data were obtained). Take-off and landing levels are much higher, but relatively brief. An accurate assessment would require accounting for the sound pressure level and noise attenuation of the communications headset.

It may be feasible that the design of an integrated audio display should be tailored to or even compensate for the possibility of hearing loss among users. A starting point for assessing the extent of hearing loss was to gather questionnaire data and conduct audiograms on professional pilots who were available as subjects from other studies. A hearing questionnaire was prepared and filled out by 64 professional pilots (of these, 26 were evaluated audiometrically). The age of the pilots who responded to the questionnaire ranged from 35 to 64 years, with a median age of 53 years. Fig. 10 shows the percentages of pilots responding "yes" to the following yes-or-no questions: 1) Have you ever been told by a doctor that you may have any sort of a permanent hearing loss in one or both ears? and 2) Do you personally suspect that you have a hearing loss in one or both ears? For comparison, the three age groups used were matched to data published by the National Center for Health Statistics [25] for general population statistics on hearing impairment in the United States, 1990–1991. While presbycusis (loss of hearing as a result of old age) may have been a factor, Fig. 10 suggests that pilot hearing loss is not necessarily age dependent. Many pilots have had military experience that could have contributed to hearing loss, although hearing protection practices have improved steadily since the 1950s. Furthermore, for many pilots the total time exposed to noise in the military is small relative to the time exposed to noise in the commercial airline cockpit. Other factors that could contribute to hearing loss (such as firing guns without hearing protection) were considered, but for the most part were not reported by subjects.

A significant loss of sensitivity for speech is generally defined as a loss greater than 25 dB HL, in the range of 125 Hz to 4 kHz; 0 dB HL is a reference based on the hearing ability of a "normal" young adult [26]. Fig. 11 shows a summary of the audiogram data for sensitivity loss from 20 of the 64 pilots questioned. The greatest loss in sensitivity was found to be at higher frequencies

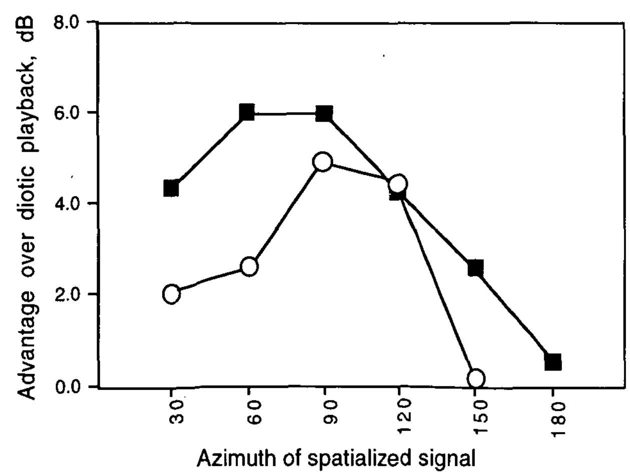


Fig. 8. Speech intelligibility advantage of 3-D auditory display compared to two-ear monaural (diotic) listening, as a function of virtual azimuths. Mean values for 50% intelligibility shown for five different subjects in each study. Filled squares—intelligibility advantage for spatialized call signs against diotic speech babble, one male speaker [21]. Call signs used were based on a subset of those used by spacecraft launch personnel at NASA Kennedy Space Center Circles—intelligibility advantage with Modified Rhyme Test (ANSI S3.2-1989), using multiple loudspeakers for signal and speech-spectrum noise as masker [22].

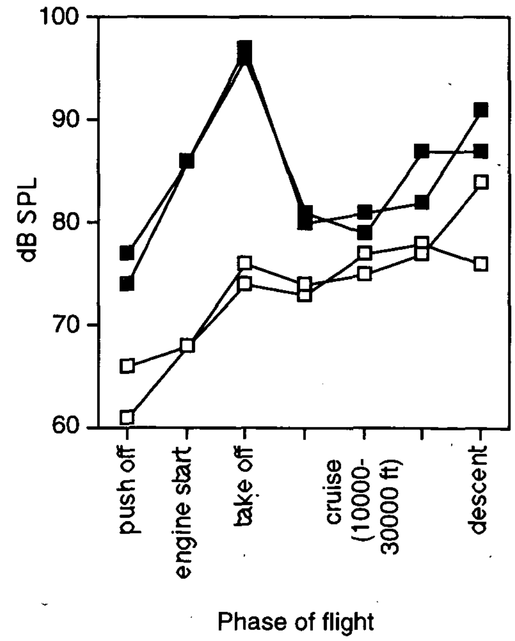


Fig. 9. Measured sound pressure levels in 737-300 cockpit from two flights. ■—C-weighted values; □—A-weighted values.

(especially 6 and 8 kHz); 55% of the pilots evaluated had a hearing sensitivity loss greater than 25 dB HL at 6 kHz. This hearing loss would generally not be regarded by occupational health and safety regulations as a severe impairment for speech intelligibility, but there is certainly an implication for overall hearing quality of life. Assuming that a loss of more than 20 dB in hearing sensitivity (>20 dB HL) is important, Fig. 11 also indicates that 70% of the pilots tested had a hearing loss at 6 kHz as well as at 8 kHz.

In addition to hearing loss, noise exposure can also cause tinnitus, the phenomenon often called "ringing" or "buzzing" in the ears. It is manifested as a constant, disturbing tone and varies between individual cases. While among the general population 15–32% have had tinnitus at some point and 5% are disabled by it, 18 (29.5%) of the 61 responding pilots reported tinnitus "occasionally" or "frequently." Those who reported "occasionally" or "frequently" indicated their tinnitus to be present anywhere from 15 to 100% of the time.

These data do not definitively prove occupational hearing loss, but they amplify and document the informal observations made by many pilots. They suggest a need for improvement in hearing protection for pilots as part of an improved auditory display of any type. Obviously, diminished hearing capacity results in diminished ability to utilize auditory cuing. However, there is no established connection between localization ability and mild hearing loss. People with presbycusis apparently adapt to the alteration in spatial cues over the course of a lifetime [27], [28].

Active noise cancellation introduces noise attenuation benefits only at frequencies below approximately 1 kHz, but could still contribute to hearing conservation in that the overall level of an integrated audio display could be reduced. This is due to the potential reduction of the "upward masking" of higher frequencies caused by low-frequency noise. It should be noted that different cockpits range widely in measured sound pressure levels, and that newer designs are generally quieter. On the other hand, the noise exposure for flight attendants in the rear of the plane is about 3–9 dB higher than it is in the cockpit, depending on the location of the engines in the particular plane.

7 CONCLUSION

This engineering report reviewed the concept of an integrated auditory display, where the total sum of signals presented to the ears of a pilot could be predicted and engineered in an optimal manner. Utilizing 3-D audio techniques in such a display adds the advantages of immediate situational awareness, source discrimination, selective attention, and redundancy. By recognizing that the standard for audio quality in the commercial aviation

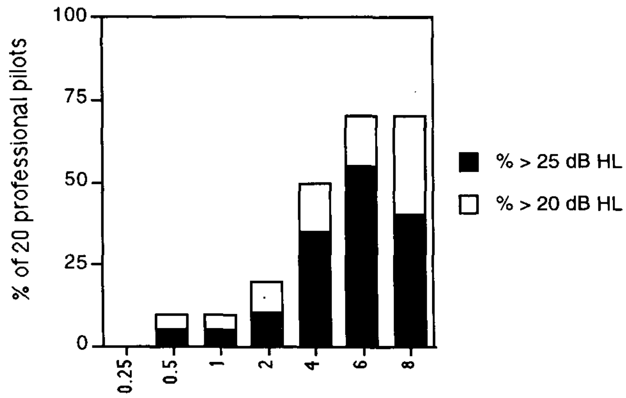


Fig. 11. Hearing loss as a function of frequency for 20 professional pilots. Filled bars—percentage of pilots tested with loss greater than 25 dB HL; filled and open bars combined—percentage of pilots tested with loss greater than 20 dB HL.

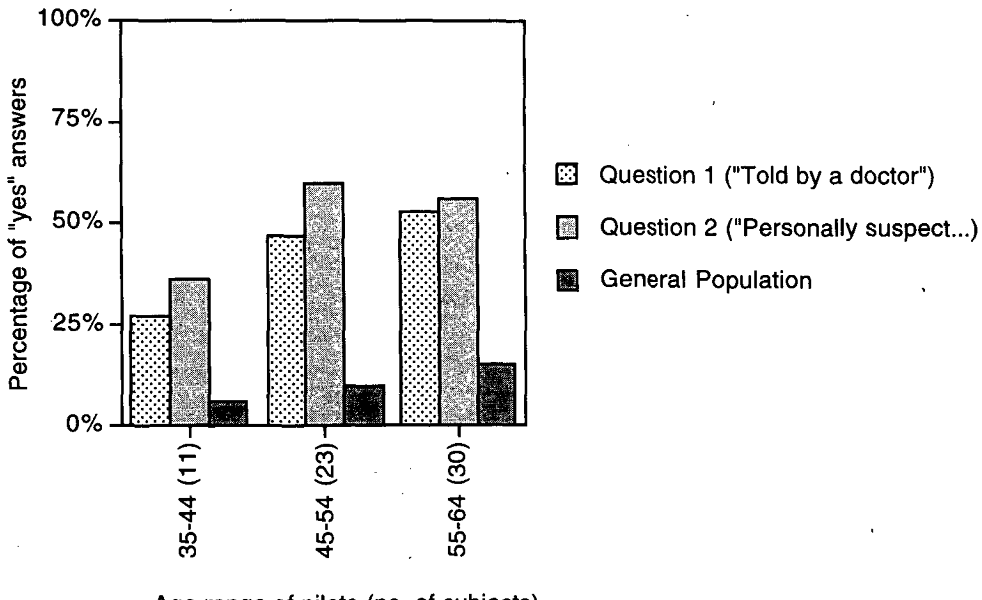


Fig. 10. Answers to questions from pilot hearing loss survey.

work environment should be improved in terms of frequency response, noise immunity, and overall integration and coordination of auditory information, it will be possible to begin to improve the safety and efficiency of commercial aircraft operations.

8 ACKNOWLEDGMENT

Many thanks are due to Elizabeth M. Wenzel, Joel Miller, Rick Shrum, Mark Pittman, Norm Stein, Mark Anderson, Laura Tran, and the team at the Crew–Vehicle Systems Research Facility (CVSRF) at NASA Ames Research Center. Support for this research was in part provided by NASA–San José State University Foundation cooperative agreement NCC-2-327.

9 REFERENCES

[1] D. R. Begault, *3-D Sound for Virtual Reality and Multimedia* (Academic Press Professional, Cambridge, MA, 1994).

[2] D. R. Begault and E. M. Wenzel, "Headphone Localization of Speech," *Human Factors*, vol. 35, pp. 361–376 (1993).

[3] A. S. Bregman, *Auditory Scene Analysis: The Perceptual Organization of Sound* (MIT Press, Cambridge, MA, 1990).

[4] E. C. Cherry, "Some Experiments on the Recognition of Speech with One and Two Ears," *J. Acoust. Soc. Am.*, vol. 25, pp. 975–979 (1953).

[5] E. C. Cherry and W. K. Taylor, "Some Further Experiments on the Recognition of Speech with One and with Two Ears," *J. Acoust. Soc. Am.*, vol. 26, pp. 549–554 (1954).

[6] ISO Std. 7731-1986(E), "Danger Signals for Work Places—Auditory Danger Signals," International Organization for Standardization, Geneva, Switzerland (1995).

[7] R. D. Sorkin, F. L. Wightman, D. S. Kistler, and G. C. Elvers, "An Exploratory Study of the Use of Movement-Correlated Cues in an Auditory Head-Up Display," *Human Factors*, vol. 31, pp. 161–166 (1991).

[8] E. M. Wenzel, "The Relative Contribution of Interaural Time and Magnitude Cues to Dynamic Sound Localization," in *1995 IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics* (IEEE Press, New York, 1995).

[9] D. R. Begault, "Head-Up Auditory Displays for Traffic Collision Avoidance System Advisories: A Preliminary Investigation," *Human Factors*, vol. 35, pp. 707–717 (1993).

[10] D. R. Begault and M. T. Pittman, "Three-Dimensional Audio versus Head-Down Traffic Alert and Collision Avoidance System Displays," *Int. J. Aviation Psychol.*, vol. 6, pp. 79–93 (1996).

[11] D. R. Perrott, T. Sadralodabai, K. Saberi, and T. Strybel, "Aurally Aided Visual Search in the Central Visual Field: Effects of Visual Load and Visual Enhancement of the Target," *Human Factors*, vol. 33, pp. 389–400 (1991).

[12] E. M. Wenzel, M. Arruda, D. J. Kistler, and F. L. Wightman, "Localization Using Non-individualized Head-Related Transfer Functions," *J. Acoust. Soc. Am.*, vol. 94, pp. 111–123 (1993).

[13] D. R. Begault, E. M. Wenzel, R. Shrum, and J. Miller, "A Virtual Audio Guidance and Alert System for Commercial Aircraft Operations," in *Proc. 1996 Int. Conf. on Auditory Display*, ICAD 1996 (http://www.Santafe.edu/~kramer/icad/ICAD96/proc96/begault.htm).

[14] E. Palmer and A. Degani, "Electronic Checklists: Evaluation of Two Levels of Automation," in *Proc. 6th Int. Symp. on Aviation Psychology* (Aviation Psychology Laboratory, Ohio State University, 1991).

[15] P. M. Zurek, "Binaural Advantages and Directional Effects in Speech Intelligibility," in *Acoustical Factors Affecting Hearing Aid Performance*, G. A. Studebaker and I. Hochberg, Eds. (Allyn and Bacon, Needham Heights, MA, 1993).

[16] D. R. Begault and E. M. Wenzel, "Techniques and Applications for Binaural Sound Manipulation in Human–Machine Interfaces," *Int. J. Aviation Psychol.*, vol. 2, pp. 1–22 (1992).

[17] D. R. Begault and E. M. Wenzel, "Technical Aspects of a Demonstration Tape for Three-Dimensional Auditory Displays," Tech. Memo TM 102286, NASA-Ames Research Center, Moffett Field, CA (1990).

[18] A. W. Bronkhorst and R. Plomp, "Effect of Multiple Speechlike Maskers on Binaural Speech Recognition in Normal and Impaired Hearing," *J. Acoust. Soc. Am.*, vol. 92, pp. 3132–3139 (1992).

[19] A. W. Bronkhorst and R. Plomp, "The Effect of Head-Induced Interaural Time and Level Differences on Speech Intelligibility in Noise," *J. Acoust. Soc. Am.*, vol. 83, pp. 1508-1516 (1988).

[20] K. Genuit, "A Model for the Description of Outer-Ear Transmission Characteristics," trans. of "Ein Modell zur Beschreibung von Außenohrübertragungseigenschaften," Ph.D. dissertation, Rhenish-Westphalian Technical University, Aachen, Germany (1984).

[21] D. R. Begault and T. Erbe, "Multichannel Spatial Auditory Display for Speech Communications," *J. Audio Eng. Soc. (Engineering Reports)*, vol. 42, pp. 819–826 (1994 Oct.).

[22] D. R. Begault, "Virtual Acoustic Displays for Teleconferencing: Intelligibility Advantage for 'Telephone Grade' Audio," presented at the 98th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 43, p. 401 (1995 May), preprint 4008.

[23] C. J. Darwin and R. P. Carlyon, "Auditory Grouping," in *Hearing*, B. C. J. Moore, Ed. (Academic Press, San Diego, CA, 1995).

[24] J.-C. Junqua, "The Lombard Reflex and Its Role on Human Listeners and Automatic Speech Recognizers," *J. Acoust. Soc. Am.*, vol. 93, pp. 510–524 (1993).

[25] National Center for Health Statistics, "Data from the National Health Interview Survey," ser. 10, no. 188, table 1 (1994).

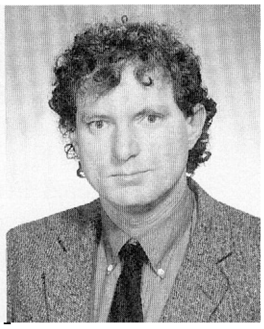
[26] M. H. Miller and L. A. Wilber, "Hearing Evaluation," in *Handbook of Acoustical Measurements and*

Noise Control, 3rd ed., C. M. Harris, ed. (McGraw-Hill, New York, 1991).

[27] K. J. Gabriel, J. Koehnke, and H. S. Colburn, "Frequency Dependence of Binaural Performance in Listeners with Impaired Binaural Hearing," *J. Acoust. Soc. Am.*, vol. 91, pp. 336–347 (1991).

[28] H. S. Colburn and C. Trahiotis, "Effects of Noise on Binaural Hearing," in *Noise-Induced Hearing Loss*, A. Dancer, D. Henderson, R. Salvi, and R. Hamernik, Eds. (Mosby, St. Louis, MO, 1991).

THE AUTHOR



Durand R. Begault received his Ph.D. from U.C. San Diego (Computer Audio Research Laboratory). He has been a senior researcher at the Auditory Display Laboratory at NASA Ames Research Center since 1988, under the auspices of the National Research Council and San Jose-State University Foundation. His research specialties include acoustic displays for aviation, psychoacoustics of spatial hearing, digital signal processing, and applied acoustical engineering.

Dr. Begault has authored two books, *3-D Sound for Virtual Reality and Multimedia, The Sonic CD-ROM for Desktop Audio Production. An Electronic Guide to Producing Computer Audio for Multimedia*, both published by Academic Press Professional.