INTELLIGIBILITY IN AUDITORY DISPLAYS

Durand R. Begault, NASA Ames Research Center

1. INTRODUCTION

The design of an auditory display is crucial for safe and efficient operations conducted in a high-stress human-machine environment, for example, in an aircraft flight deck or in a virtual environment teleoperation activity. The binaural hearing system's advantage over one-ear listening can be demonstrated not only in a laboratory context but also in practical applications for improving auditory intelligibility. Although these operational environments are necessarily dependent on intelligibility measures for determining the quality of speech communications, there are other criteria that can be equally important for assessing, measuring or predicting an improved auditory display design. These criteria apply not only to speech communications but also to other auditory or multi-modal forms of information related to the operational state of the machine. Evaluation of the auditory display may be conducted in terms of measurement and evaluation of human performance, in terms of error rates; time for task completion; discriminability between simultaneous streams of information; recognizability; and reaction time. Evaluation of perceived quality may also be germane. Human performance within the display can in turn be predicted by evaluating distortion levels, spectrum of masking background noise, and the configuration in perceived auditory space of multiple information streams Examples of applications and research pertinent to auditory display design at the Spatial Auditory Display laboratory at NASA Ames Research Center will be given.

2. AUDITORY DISPLAYS: A DICHOTOMY

There is a dichotomy that shouldn't exist between the standard of audio quality we demand for entertainment as compared to the one used in high-stress human interfaces. For example, a commercial aircraft 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 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, non-integrated design.

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 US air carrier was originally designed for telephone operators in 1961; it has an effective bandwidth of approximately 0.3–3.5 kHz. By non-integrated 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.

3. INTEGRATED AUDIO DISPLAYS AND 3-D AUDIO

By delivering both communications and warning signals through supra-aural 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 3-D audio techniques. Using either circumaural or supra-aural 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. Additionally, the use of supra-aural headsets allows inter-cockpit communication to occur without an intercom system linking the pilot and first officer. Finally, the masking effects of noise are minimized by eliminating the use of loudspeakers for warning signals and instead using headphone delivery. This allows warning signals to be designed to be noticeable by virtue of factors other than loudness (relative to an assumed receiver position). Loud, "obnoxious" 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 [1]. 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 previously measured, or that of a "good localizer" for whom non-individualized localization data had been gathered previously [2]. Warning signals can be "pre-spatialized" 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 amongst 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, e.g., the engines; and (4) display of information according to priority. This last criterion refers to the use of a special virtual location for messages of the highest priority, e.g., in the center of the head. Figure 1 summarizes these criteria in terms of psychoacoustic research areas and how they are combined into an integrated auditory display. “Binaural localization” refers to 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 towards a direction by an aural cue, and/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 active, conscious localization *per se* 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 ability.

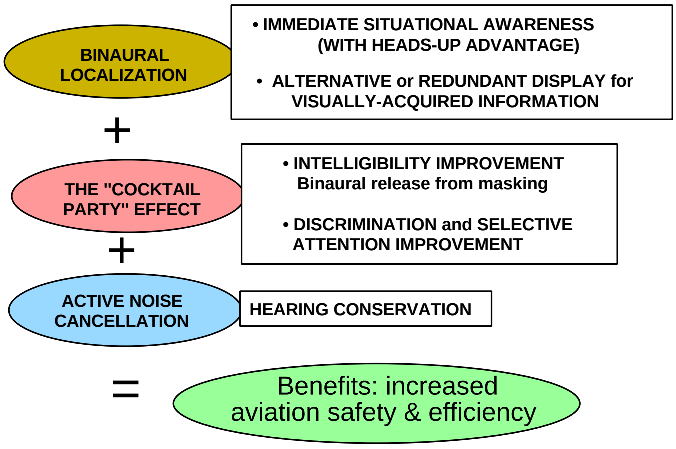

FIGURE 1. Overview of the benefits of an integrated auditory display.

4. AURALLY-GUIDED VISUAL SEARCH

Aurally guided visual search is a form of a “head-up” auditory display, which keeps directional situational awareness intelligible while keeping a pilot’s gaze out-the-window. Two experiments are summarized below where 3-D audio techniques are used for acquiring an out-the-window aircraft that would activate a Traffic Collision and Avoidance (TCAS) advisory warning.

Commercial aircraft in the US 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, e.g., “climb...climb.” Currently, a TCAS advisory warning activates a verbal announcement over the cockpit loudspeaker, a loud “traffic....traffic.” The pilot determines the location both by looking out the window and referring to a visual radar display map. The goal of these experiments 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 method for 3-D audio presentation of auditory warning and guidance signals is based upon non-real–time digital filtering of a warning signal at selected azimuths (e.g., 0 degrees, and left 30, 60 and 90 degrees) and as many as 3 different elevations (0 degrees and ± 45 degrees). Figure 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 "pre-alert" played at a relatively low volume, followed by a spoken alert. No head tracking was used, as in [7]; spatial positions were instead referenced to the nose of the plane. Head tracking would be advantageous if the 3-D audio cueing needed to be referenced to the pilot's head position, or if sound source reversals were an issue, since head 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 degrees azimuth). Pilots have reported informally in post-experiment 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.

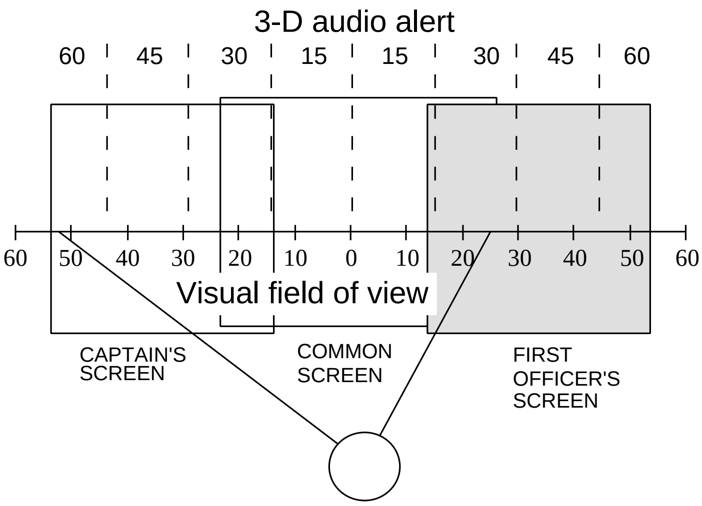

FIGURE 2. The horizontal field-of-view in the simulator, from the perspective of the left seat (Captain's position). The corresponding azimuth position of the 3-D sound cue is shown above for the second TCAS experiment. A different positional scheme was used in the first TCAS experiment; the visual range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

.

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 pre-alert 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 cueing (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 (e.g., visual targets at 15 degrees azimuth would correspond to spatialized stimuli at 30 degrees 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 sec).

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 sec). The advantage of aurally guided visual search is in line with previous results where stimuli presentation occurred under laboratory conditions where the visual stimuli were more controlled [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. Multivariate data analysis and post-experiment interviews from the experiment suggested that elevation cues were not effectively utilized by the pilots; a verbal cue would probably have been better (e.g., "traffic, high"). Successful perception of elevation cues from HRTFs tends to be more difficult than azimuth or externalization cues [12].

As shown in Figure 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.

5. GROUND OPERATIONS

In future airport ground operations, it is possible that global positioning sensor (GPS) or other methodologies may be used for improving situational awareness in both the cockpit and tower, regarding the location of aircraft, gates, and taxiway markers. Such a system may be especially useful under low-visibility conditions, which currently hampers 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 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. We hypothesized that there would be a significant preference for such a system to be included in the flight deck. The results were favorable for such a system, as might be expected. Another objective 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 significantly reduce the time required for taxiing to the gate, compared to the unaided condition, where the crew is dependent solely on a paper map and/or radio communications. The dependent variable for 3-D audio guidance system was the time necessary to complete taxi route under aided and unaided conditions. There was no overall significant effect for the guidance system, although some individual crews were able to use it to advantage.

6. AUDITORY FEEDBACK AND REDUNDANCY

The application of 3-D auditory cues for feedback and redundancy in the use of tactile-visual controls was also evaluated informally under simulation conditions by use of questionnaires [1]. Spatialized cues were used to add redundancy to messages related to the spatial orientation of the aircraft's engines; e.g., "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 non-operational 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.

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 screen. Audio feedback is used commonly in touch screen monitors that simulate cash-register–type interfaces, e.g., 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. 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. Finally some advantage was informally observed in terms of coordination ("cockpit resource management") in that the crew was aware of each other's operation of switches without moving their visual attention from another task.

7. 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 to both 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 around a 6 dB improvement in intelligibility, depending on the experimental condition (see Figure 3). The system has also been found to be advantageous with low-pass–filtered speech typical of most communication systems [22].

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 termed "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 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.

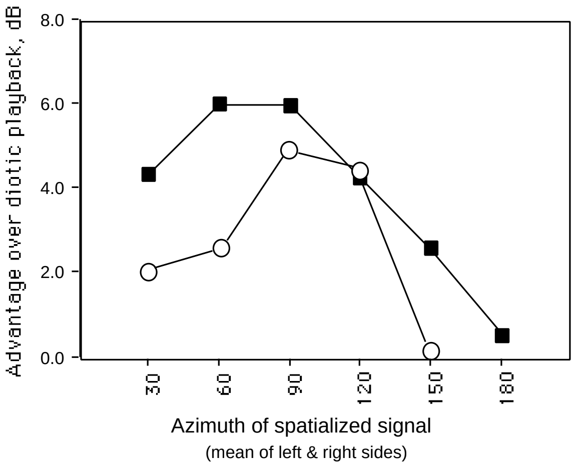

FIGURE 3. Speech intelligibility advantage of a 3-D auditory display compared to two-ear monaural (diotic) listening, as a function of virtual azimuths (mean of left and right sides). Mean values for fifty-percent intelligibility shown for five different subjects in each study. Filled squares: Intelligibility 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 using the Modified Rhyme Test (ANSI, 1989), using multiple speakers for the signal and speech-spectrum noise as the masker [22].

8. HEARING LOSS

It may be feasible that the design of an integrated audio display should be tailored or even compensate for the possibility of hearing loss amongst users. A starting point for assessing the extent of hearing loss was to gather questionnaire data and conduct audiograms on professional pilots that were available as subjects from other studies. A hearing questionnaire was prepared and filled out by 64 professional pilots (of these, 20 were audiometrically evaluated). The pilots who responded to the questionnaire had a median age of 53 years, ranging from 35–64 years. Figure 4 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, 1990-91. While presbycusis (loss of hearing as a result of old age) may have been a factor, Figure 4 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 steadily improved 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 (e.g., firing guns without hearing protection) were considered, but for the most part were not reported by subjects.

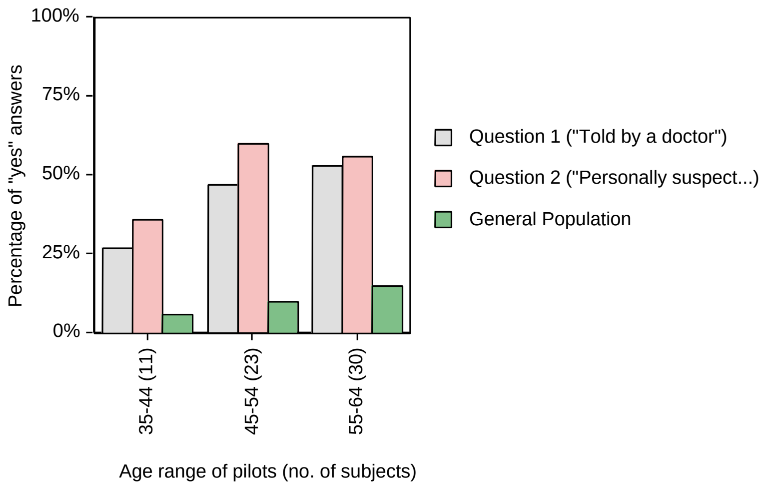

FIGURE 4. Answers to questions from a pilot hearing loss survey.

A significant loss of sensitivity for speech is generally defined as a loss greater than 25 dB HL, in the range of 125–4 kHz; 0 dB HL is a reference based on the hearing ability of a "normal" young adult [26]. Figure 5 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 (especially 6 and 8 kHz); 55% of the pilots evaluated had a hearing sensitivity loss > 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, Figure 5 also indicates that 70% of the pilots tested had a hearing loss at 6 kHz, and 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 amongst 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 15–100% of the time.

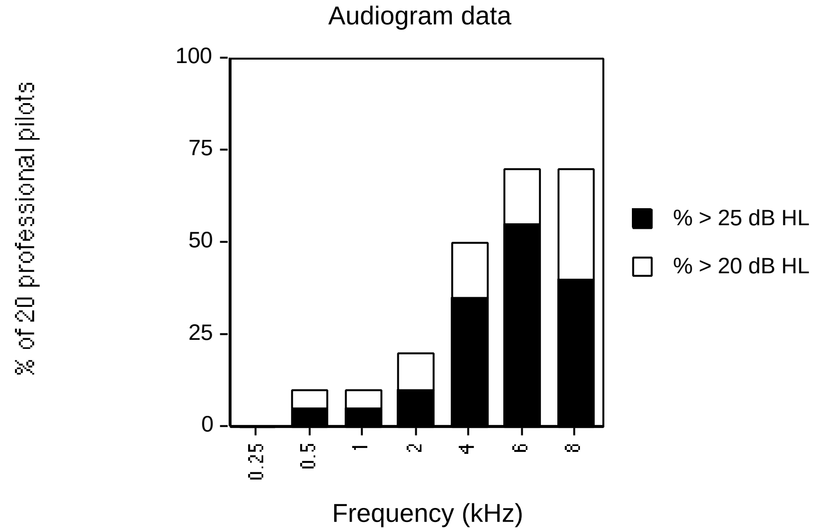

FIGURE 5. Hearing loss as a function of frequency, for 20 professional pilots. Filled bars show percentage of pilots tested with a loss of greater than 25 dB HL; the filled and open bars combined shows the percentage of pilots tested with a loss of greater than 20 dB HL.

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 because of 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 SPL levels, and that newer designs can be 6–9 dB quieter than the 737-300. On the other hand, the noise exposure for flight attendants in the rear of the plane is about 3–9 dB higher than in the cockpit, depending on the location of the engines in the particular plane.

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 cueing! 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].

9. MAXIMALLY DETECTABLE ALERTS

The following research was concerned with "how can the auditory display of alarms be made as detectable as possible within an auditory display?" Design methodologies for insuring human ability to detect the presence of a non-speech auditory alert (alarm) that is part of a warning system have for the most part concerned with an analytical approach to the amplitude spectrum of the alert and the background noise. International Standard 7731 covers the formation of auditory alerts for danger signals and states that certain frequency components be >= 13 dB above the masked threshold within 1/3 octave bands from 300-3000 Hz [6]. It is well understood from the auditory literature that, by making spectral components of an alert substantially louder than the measured background noise level, one can insure for the audibility or “detection” of such a signal [29, 30]. Such detection is referred to as “release from masking” in that spectral components of the signal are sufficiently greater in amplitude such that they may be heard. However, the technique of unmasking an auditory alert masking by means of spatial manipulation of the signal is unexplored. In addition there has not been a method taught where existing signals can be enhanced.

In high-stress environment such as an airline flight deck, an approach to insuring alert detection via emphasis on the amplitude of spectral components as opposed to spatial manipulation methods is primarily due the relatively poor quality of communication equipment and the use of monaural loudspeaker or single-earpiece headset playback systems. Eventually, binaural headphone or loudspeaker systems may become incorporated into flight deck, 911 consoles, machine operator transportation, media, and communications devices that will require the human to process a great deal of information. Also, the background noise of these types of environments will be continually reduced, allowing for presentation of informative or alerting types of signals without the need for levels that have the potential to create a “startle” effect [31]. The design of an auditory alert using spectral amplitude as a criterion is potentially problematic because other desirable signals as opposed to noise may be masked. There is a compelling motivation to provide critical alert information in a manner that allows the alert to be both audible and recognizable, but in as “quiet spoken” a manner as is possible.

In light of the above, a method has been described in a patent application that provides for the synthesis and two-channel playback of an auditory alert that is relatively more detectable against a background of noise, relative to one-channel “monaural” or “diotic” playback of the same auditory alert [32]. There are three components to this method, of which the “spatial modulation” technique is described here. The technique of spatial modulation (“jitter”) of an auditory alert along the auditory azimuth involves taking a one-channel alert signal and processing it into a two-channel signal for headphone playback, moving from a central location at 0 degrees azimuth to a position 45-90 degrees azimuth to the right, and then to the mirror image position at least 45-90 degrees azimuth to the left, and then back to 0 degrees: see Figure 6. In the presence of steady state background noise, which is relatively unvarying in its spatial properties, it was hypothesized that a spatially jittered alert is more detectable than one which is not spatially jittered.

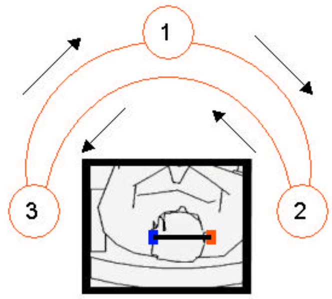

FIGURE 6. A “jittered” spatial modulation. The virtual sound is moved between positions 1-2-3-1 over the period of the alarm at either 1.6 or 3.3 Hz.

From a design standpoint, the alert should not bring attention to any particular spatial location during its excursion. As exemplified by everyday experience with fast-moving insects and vehicles, attention may be stimulated by the activation of the auditory system’s response to dynamic changes in interaural relationships, but within certain ranges of source velocity it is difficult to actually localize a moving source. Evidence has also been gathered recently for an auditory motion detector, independent of a mechanism for localization, that is potentially analogous to a visual motion detection mechanism in the cererbal cortex [33-36]. For the technique described here, the spatial modulation frequency (which corresponds to the frequency of amplitude or time delay modulation) should be in the region of 2-4 Hz, and no greater than 10 Hz. For interaural time delay manipulation, this range is where the auditory system perceives movement, but is relatively bad at following the location, a phenomenon known as “binaural sluggishness” [37-39].

There are many methods by which to implement spatial jitter. It can be done by means of linear or exponential amplitude panning, or by continuously varying a time delay to each ear in the range 0 – ca. 0.8 ms. Finally, binaural variations in time and amplitude as a function of HRTF convolution can be implemented via a 3-D sound interface that allows movement of a virtual source to a listener. The latter technique was used in the study from which the results are presented below.

The advantage of spatially modulating an existing alert is that it is still inherently recognizable as the alert itself, despite the spatial modification. Although spectral components are modified in amplitude by dynamic Head-Related Transfer Function (HRTF) filtering, the temporal-spectral gestalt of the alert remains recognizable to a pilot or other machine operator.

A comparison was made to determine how the “existing” condition for auditory alerts on the flight deck, where the alert is provided through a monaural loudspeaker, compares to a spatially jittered, headphone delivered alert. Figure 7, left, shows the data for these conditions (mean and standard deviation for all participants), with the threshold plotted relative to the r.m.s level of the background noise. The advantage of mechanical noise attenuation provided by the headphone, combined with the jittered alert, was on average 13.4 dB compared to the loudspeaker alert.

Further analyses were made to test the effect of alert jittering and to determine the effect of trajectory velocity rate. The independent variable was trajectory velocity (0.0 Hz, or no jitter; 1.6 Hz trajectory; and 3.3 Hz trajectory). Figure 7, right, indicates the mean and standard deviation values of the results across the fourteen subjects tested. The mean thresholds were –13.3 dB for no jittering; -21.2 dB for the 1.6 Hz trajectory, and –21.0 for the 3.3 Hz trajectory. Analyses of variance (ANOVAs) were run for each pair wise comparison. The no movement condition (0 Hz trajectory) was significantly different than either trajectory condition ([$F(1,13) = 203$, $p < .000$], [$F(1,13) = 274$, $p < .000$]), but there was no significant difference between the two moving trajectory velocities tested. The overall reduction in threshold caused by spatial jittering at either rate is about 7.8 dB.

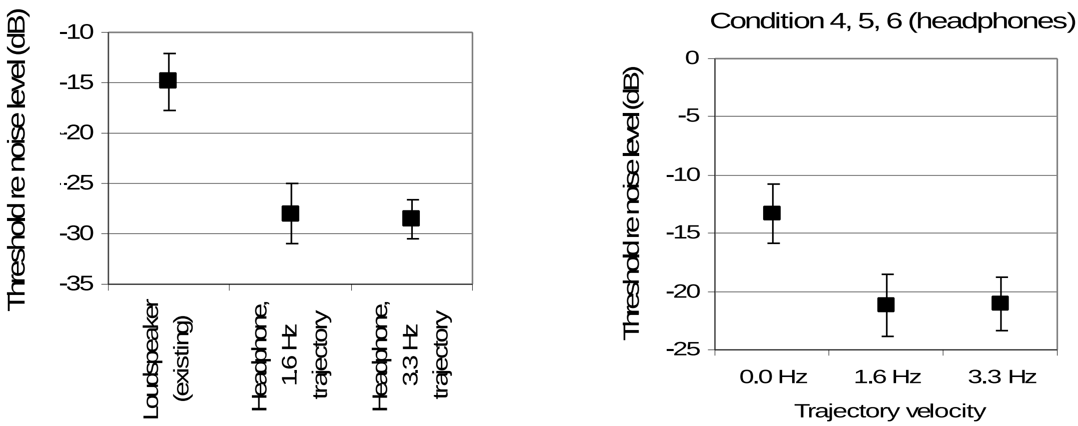

FIGURE 7. Left: loudspeaker alert versus headphone alert. Right: headphone alert thresholds at different target velocity rates.

7. CONCLUSION

This paper 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 work environment should be improved in terms of “auditory display intelligibility” it will be possible to begin to improve the safety and efficiency of commercial aircraft operations.

8. ACKNOWLEDGEMENTS

Material in this paper has appeared previously in [40, 41]. Many thanks for the support of the personnel at the Spatial Auditory Display Laboratory at NASA Ames Research Center. Funding for this work was provided by the PPSF-AOS Project of NASA’s Airspace Systems Program.

10. REFERENCES

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

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

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

[4] Cherry, E.C., “Some experiments on the recognition of speech with one and two ears.” *Journal of the Acoustical Society of America*, vol. 25, pp. 975 - 979 (1953).

[5] Cherry, E.C. and W.K. Taylor, “Some further experiments on the recognition of speech with one and with two ears”. *Journal of the Acoustical Society of America*, vol. 26, pp. 549 - 554 (1954).

[6] ISO Standard 7731-1986 (E), *Danger signals for work places-Auditory danger signals*. Geneva: International Organization for Standardization (1995).

[7] Sorkin, R.D., Wightman, F. L., Kistler, D. S., and Elvers, G. C. “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] Wenzel, E.M., “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* . New York: IEEE Press (1995).

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

[10] Begault, D.R. and M.T. Pittman, “Three-Dimensional Audio Versus Head-Down Traffic Alert and Collision Avoidance System Displays”. *International Journal of Aviation Psychology*, vol. 6, pp. 79-93 (1996).

[11] Perrott, D.R., *et al.*, “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] Wenzel, E.M., Arruda, M., Kistler, D.J. and Wightman, F.L. “Localization using non-individualized head-related transfer functions”. *Journal of the Acoustical Society of America*, vol. 94, pp. 111-123 (1993).

[13] Begault, D.R., Wenzel, E. M., Shrum, R., and Miller, J. “A virtual audio guidance and alert system for commercial aircraft operations”. In *Proceeding of the 1996 International Conference on Auditory Display* (In Press).

[14] Palmer, E. and A. Degani, “Electronic Checklists: Evaluation of Two Levels of Automation”. In *Proceedings of the Sixth International Symposium on Aviation Psychology* Aviation Psychology Laboratory, Ohio State University (1991).

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

[16] Begault, D.R. and E.M. Wenzel, "Techniques and applications for binaural sound manipulation in human-machine interfaces". *International Journal of Aviation Psychology*, vol. 2, pp. 1-22 (1992).

[17] Begault, D.R. and E.M. Wenzel, *Technical aspects of a demonstration tape for three-dimensional auditory displays* (technical memorandum TM 102286). Moffett Field, CA: NASA-Ames Research Center (1990).

[18] Bronkhorst, A.W. and R. Plomp, "Effect of multiple speechlike maskers on binaural speech recognition in normal and impaired hearing". *Journal of the Acoustical Society of America*, vol. 92, pp. 3132-3139 (1992).

[19] Bronkhorst, A.W. and R. Plomp, "The effect of head-induced interaural time and level differences on speech intelligibility in noise." *Journal of the Acoustical Society of America*, vol. 83, pp. 1508-1516 (1988).

[20] Genuit, K., *A model for the description of outer-ear transmission characterstics* (translation of *Ein Modell zur Beschreibung von Außenohrübertragungseigenschaften)* Dissertation, Rhenish-Westphalian Technical University, Aachen, Germany (1984).

[21] Begault, D.R. and T. Erbe, "Multichannel spatial auditory display for speech communications". *Journal of the Audio Engineering Society*, vol. 42, pp. 819-826 (1994).

[22] Begault, D.R., "Virtual acoustic displays for teleconferencing: Intelligibility advantage for "telephone grade" audio". In *AES 98th Convention* preprint 4008 (1995).

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

[24] Junqua, J.-C., "The Lombard reflex and its role on human listeners and automatic speech recognizers". *Journal of the Acoustical Society of America*, vol. 93, pp. 510-524 (1993).

[25] NCHS, *Data from the National Health Interview Survey*, Series 10, Number 188, Table 1. National Center for Health Statistics (1994).

[26] Miller, M.H. and L.A. Wilber, "Hearing Evaluation." In *Handbook of Acoustical Measurements and Noise Control (3rd ed.)*, C.M. Harris, Editor. New York: McGraw-Hill (1991).

[27] Gabriel, K.J., J. Koehnke, and H.S. Colburn, "Frequency dependence of binaural performance in listeners with impaired binaural hearing". *Journal of the Acoustical Society of America*, vol. 91, pp. 336-347 (1991).

[28] Colburn, H.S. and C. Trahiotis, *Effects of Noise on Binaural Hearing*, in *Noise-induced hearing loss*, A. Dancer, D. Henderson, R. Salvi, and R. Hamernik, Ed.s. St. Louis. MO: Mosby (1991).

[29] Moore, B.C.J., ed. *Hearing. 2nd ed. "*Handbook of Perception and Cognition*".* 1995, Academic Press: San Diego.

[30] Zwicker, E., and Fastl H., *Psychoacoustics.* "Facts and Models." 1990, Berlin: Springer-Verlag.

[31] Patterson, R., "Guidelines for auditory warning systems on civil aircraft." 1982, *Civil Aviation Authority*, London.
[32] Begault, D.R., "Making audible alarms more noticeable in noisy environments*." NASA Tech Briefs*, 2002. 26(9).
[33] Rashbass, C., "The relationship between saccadic and smooth tracking eye movements." *Journal of Physiology* (London), 1961. 159: p. 326-338.
[34] Krukowski, A.E., Begault, D.R., Wenzel, E.M., and Stone, L.S., "Human oculomotor responses to virtual auditory motion." *Society for Neuroscience 31st Annual Meeting*. 2001. San Diego, CA.
[35] Krukowski, A.E., Begault, D.R., Wenzel, E.M., and Stone, L.S., "Human smooth pursuit eye movement responses to visual, auditory, and imagined target motion (abstract)." *The Society for the Neural Control of Movement Twelfth Annual Meeting.* 2002. Naples, Florida.
[36] Krukowski, A.E., Begault, D.R., Wenzel, E.M., and Stone, L.S., "Pursuing the invisible: Human oculomotor response to auditory motion." *Nature* (submitted), 2003.
[37] Grantham, D.W., "Spatial Hearing and related phenomenon," *Hearing*, B.C.J. Moore, Editor. 1995, Academic Press: San Diego.
[38] Grantham, D.W., and Wightman, F.L., "Detectability of a pulsed tone in the presence of a masker with time-varying interaural correlation." *Journal of the Acoustical Society of America*, 1979. 65: p. 1509-1517.
[39] Grantham, D.W., and Wightman, F.L., "Detectability of varying interaural temporal differences." *Journal of the Acoustical Society of America*, 1978. 63: p. 511-523.
[40] Begault, D.R., "Virtual acoustics, aeronautics and communications." *Journal of the Audio Engineering Society*, 1998. 46(6): p. 520-530.
[41] Begault, D. R., Anderson, M. R., & McClain, B. U. (2003). *Spatially-modulated auditory alerts.* Paper presented at the Proceedings of the 2003 International Conference on Auditory Display, Boston, MA July 6-9 2003.