3D-Sonification for Obstacle Avoidance in Brownout Conditions

Dr. M. Godfroy-Cooper
Senior Research Associate, SJSU/NASA ARC, Human System Integration Division, Moffett Field, CA

J. D. Miller
Research Associate, SJSU/NASA ARC, Human System Integration Division, Moffett Field, CA

Z. Szoboszlay
Researcher
U.S. Army Aeroflightdynamics Directorate, Ames Research Center, Moffett Field, CA

Dr. E. M. Wenzel
Senior Research Psychologist, NASA ARC, Human System Integration Division, Moffett Field, CA

ABSTRACT

Helicopter brownout is a phenomenon that occurs when making landing approaches in dusty environments, whereby sand or dust particles become swept up in the rotor outwash. Brownout is characterized by partial or total obscuration of the terrain, which degrades visual cues necessary for hovering and safe landing. Furthermore, the motion of the dust cloud produced during brownout can lead to the pilot experiencing motion cue anomalies such as vection illusions. In this context, the stability and guidance control functions can be intermittently or continuously degraded, potentially leading to undetected surface hazards and obstacles as well as unnoticed drift. Safe and controlled landing in brownout can be achieved using an integrated presentation of LADAR and RADAR imagery and aircraft state symbology. However, though detected by the LADAR and displayed on the sensor image, small obstacles can be difficult to discern from the background so that changes in obstacle elevation may go unnoticed. Moreover, pilot workload associated with tracking the displayed symbology is often so high that the pilot cannot give sufficient attention to the LADAR/RADAR image. This paper documents a simulation evaluating the use of 3D auditory cueing for obstacle avoidance in brownout as a replacement for or compliment to LADAR/RADAR imagery.

NOTATION

3D	Three-Dimensional
AATD	Aviation Applied Technology Directorate
ADAS	Advanced Driver Assistance Systems
ADD	Aviation Development Directorate
AFRL	Air Force Research Laboratory
AMRDEC	Aviation and Missile Research Development and Engineering Center
ANOVA	Analysis of Variance
DVE	Degraded Visual Environment
FM	Frequency Modulation
FOV	Field-of-View
HDD	Head-Down Display
HMP	Horizontal Median Plane
HRIR	Head-Related Impulse Response
HRTF	Head-Related Transfer Function
ILD	Interaural Level Difference
IPD	Interaural Phase Difference
ITD	Interaural Time Difference
LADAR	Laser Detection and Ranging
NASA	National Aeronautics and Space Administration
PLSD	Protected Least Significant Difference
RADAR	Radio Detection and Ranging
RT	Response Time
SMP	Sagittal Median Plane
TCAS	Traffic alert and Collision Avoidance System
USGS	U.S. Geological Surveys
VAS	Virtual Auditory Scene
VE	Virtual Environment

Presented at the AHS International 73rd Annual Forum & Technology Display, Fort Worth, Texas, USA, May 9-11, 2017.

INTRODUCTION

This simulation study was a joint effort by the U.S. Army Aviation and Missile Research Development and Engineering Center (AMRDEC) and National Aeronautics and Space Administration Ames Research Center (NASA ARC).

This study evaluates the use of *spatial auditory cueing* (3D sound) for the representation of natural and man-made obstacles, as a replacement for or compliment to LADAR/RADAR imagery and aircraft state symbology (Ref. 1, Ref. 2). More specifically, we propose a new use of looming/receding auditory warning signals for collision avoidance during the approach/hover phases of flight in degraded visual environment (DVE) conditions. Four participants made repeated open-loop egocentric localizations of auditory targets presented randomly across the frontal hemifield. The results are reported in terms of variable error, constant error, and local distortion. The results provide a baseline for the use of 3D audio in UH-60 background noise. They demonstrate the usability and acceptability of spatialized sound to provide a dynamic representation of obstacles during a simulated helicopter drift.

Background

Motivation for the use of sound for in-vehicle technologies

The auditory channel generally provides information in terms of speech and sounds. Auditory display of information through simulated non-speech sound, i.e., *sonification*, has become a new area of research over the last two decades (Ref. 3). Sonification can be defined as the mapping or transformation of data streams onto auditory dimensions for the purposes of facilitating communication or interpretation (Ref. 4). Sonification includes auditory icons, earcons and audification. Auditory icons represent a sound "image" of the object or motion to which it is referring. This is a direct comparison to visual icons. E.g., a heartbeat sound can be used for monitoring pulse information (Ref. 5). Earcons are abstract audio messages used in the user interface to provide information and feedback to the user about user interface entity state (Ref. 6). In contrast to auditory icons, earcons are harder to remember and learn because they have no natural link or mapping to the objects or events they represent. On the other hand, they are highly structured and can easily represent families and hierarchies of objects and actions with very simple audio messages. This type of sonification has better results in desktop interfaces, alarms and warning systems such as vehicle collision detection and immersive visual environments (VEs). Sonification has been used successfully in advanced driver assistance systems (ADAS) with high priority warnings such as forward collision warnings, lane or road departure warnings, and blind spot and back-up warnings. Last, audification is a specific type of auditory data analysis in which data samples are isomorphically mapped to time or frequency domain audio data. Audification is the most direct form of sonification, as all data samples are preserved and spectral features within the original data will be present as timbral components in the resulting sound file.

The rationale and motivation for displaying information using sound (rather than visual information) have been discussed extensively in the literature (Ref. 7). Because auditory displays exploit the superior ability of the human auditory system to recognize temporal changes and patterns (Ref. 8), they may be the most appropriate modality when the information being displayed has complex patterns, changes in time, including warnings, or call for immediate action. In controlled conditions, auditory cues presented at the location of their visual counterparts can be used to exogenously capture a participant's visual attention and thereby facilitates the performance of a variety of visual tasks (Ref. 9). In critical domains such as low-level flight where unintentional drift, changes in altitude, and sink rates require immediate counteractive measures to avoid flight into terrain, auditory cues have the ability to capture pilot's attention and elicit orientation responses regardless of head position or eye fixation (Ref. 10). Interestingly, and ecologically valent, Rummukainen (Ref. 11) reported task-relevant auditory cues to aid in orienting to and detecting a peripheral but not central visual target.

Recent studies have suggested that looming sounds (sounds that rapidly increase in amplitude) may provide a particularly salient stimulus with multisensory implications (Ref. 12, Ref. 13). Humans and other primates show a particular responsiveness to looming sounds (Ref. 14, Ref. 15), possibly because these might indicate a potential threat. Leo et al. (Ref. 16) reported a specific benefit in visual orientation discrimination sensitivity, i.e. multisensory effects, when using structured looming sound. In contrast, receding, static and white noise sounds produced no such "spatial congruence" effect on visual orientation sensitivity.

Static cues

Three-dimensional (3D) auditory display makes use of the natural sound localization ability of humans. The localization of an auditory stimulus in the horizontal dimension (azimuth, defined by the angle between the source and the sagittal plane) results from the detection of left-right interaural differences in time (interaural time differences, ITDs, or interaural phase differences, IPDs) and differences in the received intensity (interaural level

Presented at the AHS International 73rd Annual Forum & Technology Display, Fort Worth, Texas, USA, May 9-11, 2017.

differences, ILDs, Ref. 17). To localize a sound in the vertical dimension (elevation, defined by the angle between the source and the horizontal plane containing the listener's ears) and to resolve front-back confusions, the auditory system relies on the spectral cues provided by the detailed geometry of the pinnae. Pinna features cause acoustic waves to diffract and undergo direction-dependent reflections (Ref. 18; Ref. 19). The two different modes of indirect coding of the position of a sound source in space (as compared to the direct spatial coding of visual stimuli) result in differences in spatial resolution in these two directions. Carlile (Ref. 19) studied localization *accuracy* for sound sources on the sagittal median plane (SMP), +/- 20º about the auditory-visual horizon (the SMP being the vertical plane passing through the midline). Using a head pointing technique, he reported constant errors (*CE*s) as small as 2º-3º for the horizontal component and 4º and 9º for the vertical component (see also Ref. 20; Ref. 21; Ref. 22 for similar results). For frontal sound sources (0º position in both the horizontal and vertical plane), Makous and Middlebrooks (Ref. 20) reported *CE*s of 1.5º in the horizontal plane and 2.5º in the vertical plane. The smallest errors appear to occur for locations associated with the audio-visual horizon, also referred to as horizontal median plane (HMP) while locations off the audio-visual horizon were shifted towards the audio-visual horizon, resulting in a compression of the auditory space that is exacerbated for the highest and lowest elevations (Ref. 21). Such a bias has not been reported for locations in azimuth. Recently, Pedersen and Jorgensen (Ref. 23) reported that the size of the *CE*s in the SMP depends on the actual sound source elevation and is about +3° at the horizontal plane, 0° at about 23° elevation, and becomes negative at higher elevations (e.g., -3° at about 46°) (see also Ref. 19). For *precision*, variable errors (*VE*s) are estimated to be approximately 2° in the frontal horizontal plane near 0° (directly in front of the listener) and 4° to 8° in elevation (Ref. 24; Ref. 25). The magnitude of the *VE* was shown to increase with sound source laterality (eccentricity in azimuth) to a value of 10° or more for sounds presented on the sides or the rear of the listener, although to a lesser degree than the size of the *CE*s (Ref. 26). For elevation the *VEs* are minimum at frontal location (0°, 0°) and maximum at the extreme positive and negative elevations.

Dynamic cues

In addition to the static localization cues (ITD, ILD and anatomical scattering), humans use dynamic cues to reinforce localization. These arise from active, sometimes unconscious, motions of the listener, which change the relative position of the source (Ref. 27).

Virtual sound sources

The technique used to create such spatial auditory displays is based on real-time filtering of the sound with head-related transfer functions (HRTFs) that simulate the acoustic effects of the listener's shoulders, head and pinnae. If the HRTF is known, one can synthesize a virtual auditory scene (VAS) that gives the listener the impression of the sound sources being presented in exocentric space. In a static anechoic environment, filtering of the source signal with the HRTF for a given direction delivers to the listener's eardrums the same acoustic pressure waves as the true source in the same environment. By including reverberation and motion cues due to ego-motion of the listener, one can synthesize more realistic environments (Ref. 28). However, individual differences in anatomy, especially the shape of the pinnae does not allow a good match of the same HRTF to all users. The pinna, head, and torso sizes can vary greatly from one person to the next. Thus, the spectral characteristics that convey sound location also varies and the HRTF of one individual can create a significantly distorted perception for another.

Perceptually veridical localization over headphones is possible if the spectral shaping by the pinnae as well as the interaural difference cues (ITDs and ILDs) can be adequately reproduced in a 3D sound system or virtual acoustic display (Ref. 21). Using broadband white noise sound bursts in complete darkness with a ±30° range in azimuth and elevation, head fixed, and using eye-movement responses, Hoffman reported individualized azimuth errors in the range of 2.9° to 6.6° in azimuth and 4.2° to 7.7° in elevation. With altered pinna cues, simulating the effect of non-individualized HRTFs, azimuth errors ranged from 3.8° to 8.3° and elevation errors from 16.3° to 23.0°.

Several flight simulator studies have investigated the use of 3D audio for the aural Traffic alert and Collision Avoidance System (TCAS) warning, which is installed in most commercial aircraft (Ref. 29). All studies showed that out-the-window visual search time for the intruding aircraft was reduced with 3D audio, compared to monaural warnings. Bronkhorst (Ref. 30) examined the application of 3D audio to indicate the location of a target jet in a fighter intercept task. They observed that the fastest target acquisition times were obtained with the combination of the visual HDD and the 3D auditory display. No difference was found between the conditions with only the visual display or the 3D auditory display. The application of 3D audio can also be extended to other types of auditory signals in the cockpit. For instance, Haas (Ref. 31) used 3D audio as a warning display for system malfunctions in helicopters, where the spatial source of the 3D audio warning corresponded to the location of a system malfunction of the aircraft or to the location of a visual indicator light inside the cockpit. The results showed faster response times to the warnings when they were presented with 3D audio (i.e., 3.6 sec on average) compared to the condition when only visual warning signals were present (i.e., 5 sec).

The present study

The present study investigates the usability and acceptability of 3D sound as a way to depict obstacle location as a

helicopter drifts during a hover in DVE. A part-task simulation was conducted that evaluated the localization capabilities of moving sounds in the presence of UH60 cockpit noise without visual cues. The effects of sonification type (son1 vs. son2), condition of presentation (Continuous or Memory), and Generic versus Individualized HRTFs were tested between participants.

METHOD

NASA Simulation Facility

The experiment was conducted in the Advanced Controls and Displays (ACD) laboratory within the Human Systems Integration Division at NASA Ames Research Center. Participants were placed in a double-walled soundproof booth and seated on a height-adjustable rotating stool positioned at the center of a large, curved projection surface (Elumens VisionStation Display) (Figure 1). The display surface has a 5'-wide (1.5m) projection area with a 33" (84cm) spherical radius of curvature. Head tracking was performed using a Polhemus Fastrak® sampled at 120Hz. Virtual acoustic environment state and data collection were both updated at a 60Hz update rate. Participants wore Sennheiser Precision HD 580 headphones for the presentation of spatial sonifications and background noise. All sound was rendered using the slab3d spatialization engine.

Figure 1. NASA ACD Soundproof Booth.

The slab3d Spatial Audio Engine

slab3d (Ref. 32, Ref. 33, Ref. 34) is an Open-Source real-time virtual acoustic environment rendering system originally developed in the ACD laboratory in the late 1990s (Ref. 35). Since then, it has also been developed and maintained by the AFRL and the U.S. Army. The slab3d release includes a collection of applications, libraries, documentation, MATLAB utilities, and source code for virtual acoustic environment development and real-time audio signal processing (Figure 2).

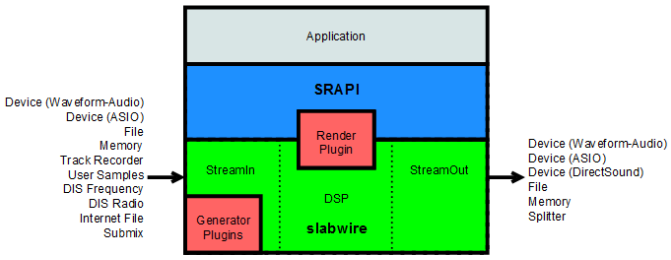


Figure 2. slab3d Software Architecture.

AvadeServer

AvADE (Aviation Auditory Display Engine) was developed to enable rapid prototyping and experimental study of advanced auditory display concepts for aviation applications. It is a subset of the slab3d release that allows a Windows PC to be dedicated as a networked spatial audio and TTS (text-to-speech) server controlled via UDP or TCP/IP command strings. The AvADE software comprises three applications: AvadeServer, the main AvADE server application, AvadeClient, a demo, testing, and scripting client application, and WinSpeak, a SAPI 5.4 TTS test utility. AvadeServer was designed with a lean layered approach to ease maintainability and extensibility (Figure 3). It can be easily modified to support a wide variety of spatial audio, sonification, TTS, and communication server applications.

AvadeServer Architecture	
Component / Layer	**Functionality**
AVF AvadeFile (XML)	Source and Stream State
Form GUI (C#, EXE)	User Interface
SlabServer (C#, EXE)	UDP, TCP/IP, SAPI
SlabSharp (C#, DLL)	Sonifiers
MSRAPI (C++/CLI, DLL)	Managed SRAPI
SRAPI (C++, LIB)	Slab3d Render API
slabwire (C++, LIB)	Stream I/O, DSP

Figure 3. AvadeServer Software Architecture.

AvadeServer (Figure 4) supports a variety of features organized around the concept of sound "sources", sample "streams", and "Render Plugins" (i.e., providing a GUI and network interface for the slab3d engine shown in Figure 2). Render Plugins process stream input to stream output using plugin-specific state information encapsulated in sources. Two plugins are supported by AvadeServer, "Spatial", an HRTF-based virtual acoustic environment, and "Mixer", a virtual mixing console. For the Spatial Render Plugin, a

source is a virtual environment sound emitter, i.e., a virtual entity with an azimuth, elevation, and range relative to a listener. For the Mixer Render Plugin, a source is a mixer channel strip, i.e., output channel gains and a binaural HRTF effects processor. For both plugins, a stream is the sound sample stream that plays from the sound source, e.g., analog input, wave files, TTS speech, DIS Radio communications, signal generators (potentially controlled by "Sonifiers"), etc. A “Sonifier” is a module that maps input parameters and/or virtual environment state to signal generator parameters. Although input streams can be at arbitrary sample rates (they are resampled as needed), the HRTF database and sound output are presently limited to 44,100 samples/s. This experiment used the Spatial Render Plugin and the discussion that follows assumes its use.

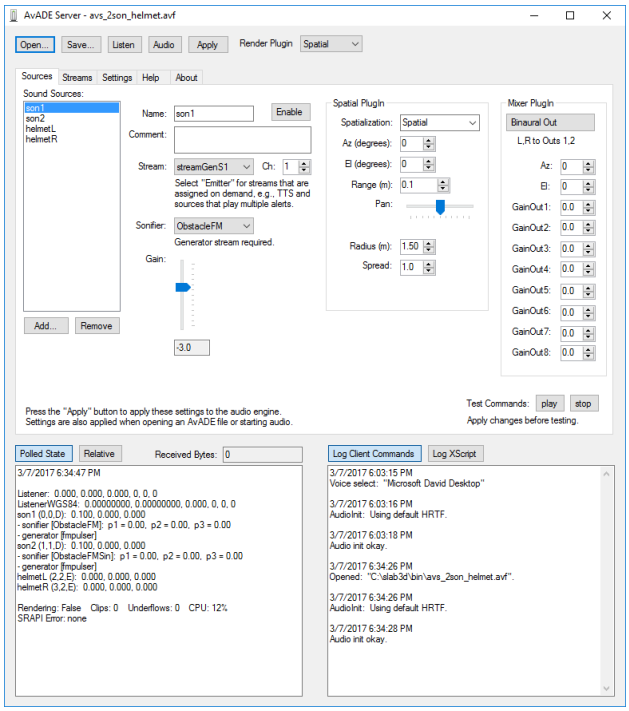


Figure 4. AvadeServer Application.

HRTF Databases

Two Listener HRTF Databases were used in the experiment. Slab3d’s default jdm.slh database (Figure 6) was used as both the non-individualized database and as an individualized database. jdm.slh was measured using the ACD lab’s Snapshot single-speaker HRTF measurement system. The mg.slh database (Figure 7) used as only as individualized database was measured using the ACD lab’s Headzap (Ref. 36) multi-speaker system. Both Snapshot and HeadZap used Golay codes as the test signal (Ref. 37) and equalized the HRTFs using free-field mic and speaker measurements.

Drift Trajectories and Alert Rings

Each trial simulated helicopter drift using one of sixty-nine predesigned trajectories relative to an obstacle placed at the origin (Figure 5). The trajectories were selected based on alert ring radii, forward hemifield destination distribution, and obstacle-ownship azimuth sweep. The alert rings proceed outwards from a blade collision radius of 25 ft. (red circle, Figure 5) as listed in Table 1. All trajectories began 10 feet beyond the Advisory ring. Nine trajectories were placed on axis so that no azimuth sweep occurred. The trajectories began at the Figure 5 diamonds and proceeded towards the squares at a speed of 10 ft./s.

Table 1. Alert Rings.

Alert Ring	Alert Level	Obstacle-Ownship Distance	Pulse Period
ring1	Warning	40 feet	200 ms
ring2	Caution	60 feet	400 ms
ring3	Advisory	90 feet	600 ms

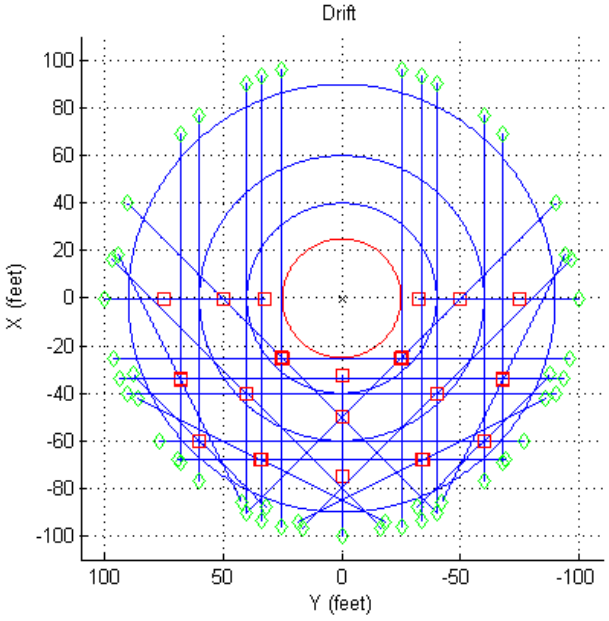


Figure 5. Helicopter drift trajectories with an obstacle at the origin in the local slab3d coordinate system. Trajectories start at the green diamonds and proceed towards the red squares at a speed of 10 ft/s. The helicopter blade radius of 25 ft is shown in red around the origin. The large blue circles are the alert rings defined in Table 1. Though the rings are, by design, around ownship, the ownship and obstacle are idealized dimensionless objects in regards to collision distance calculations. Thus, the alert rings can be visualized around either entity.

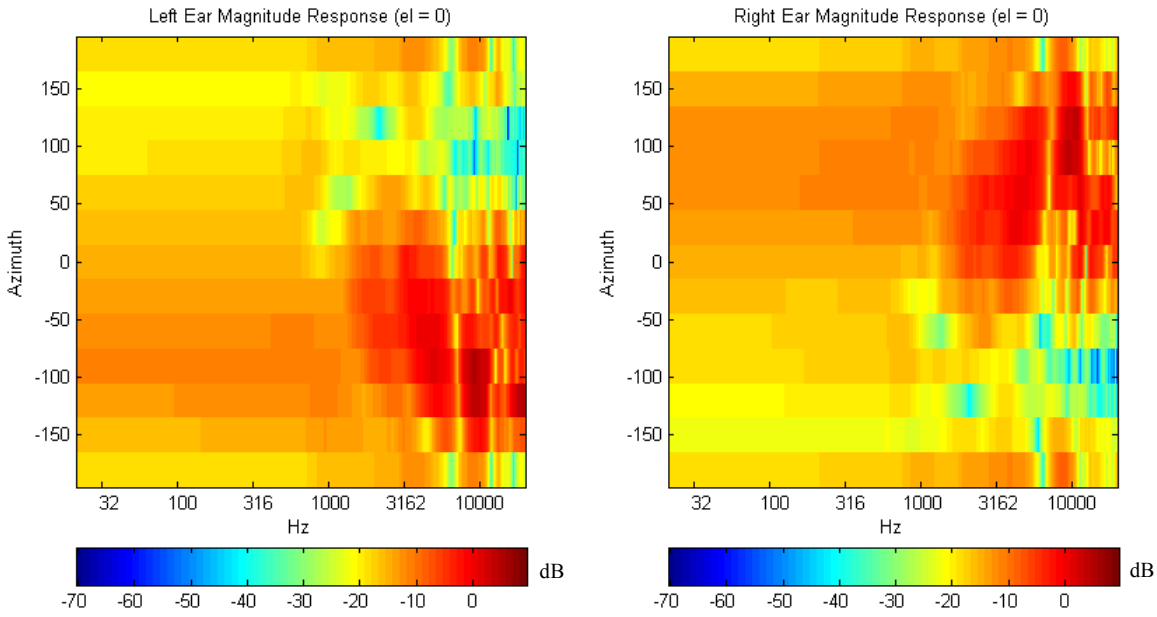


Figure 6. The HRTF database jdm.slh is the default slab3d HRTF database and was measured using the ACD lab's Snapshot measurement system. It was measured with an azimuth increment of 30° (-180° to 180°) and an elevation increment of 18° (-36° to +54°) and interpolated to a full sphere. All azimuths for the horizontal plane are shown.

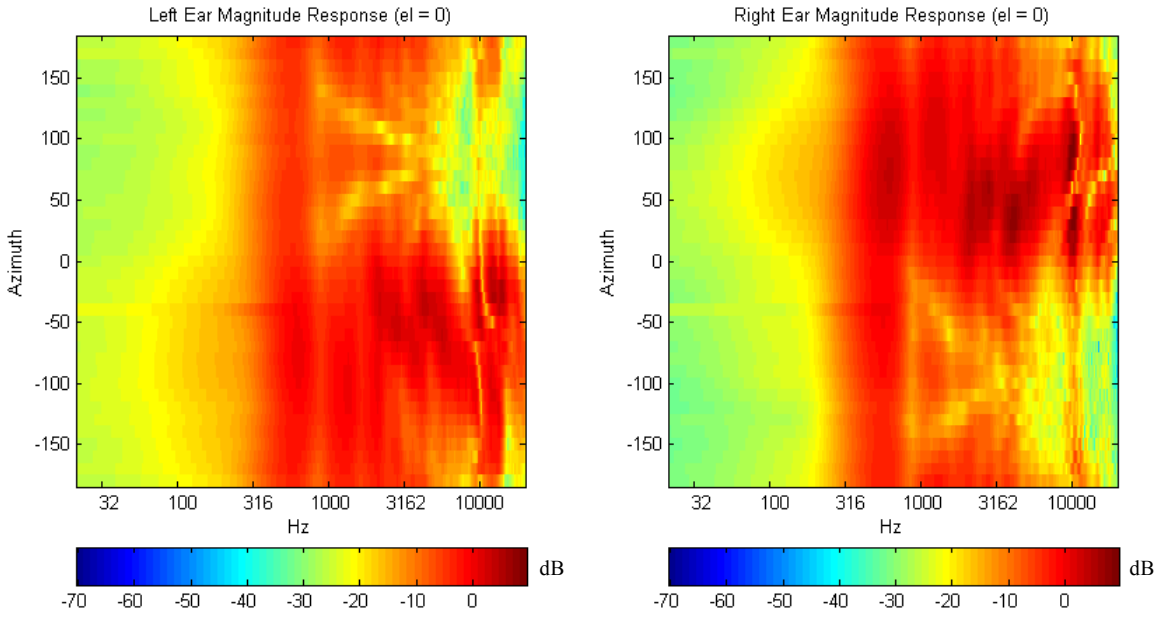


Figure 7. The individualized HRTF database mg.slh was measured using the ACD lab's HeadZap measurement system. It was measured with an azimuth increment of 10° (-180° to 180°) and an elevation increment of 10° (-40° to +70°) and interpolated to a full sphere. All azimuths for the horizontal plane are shown.

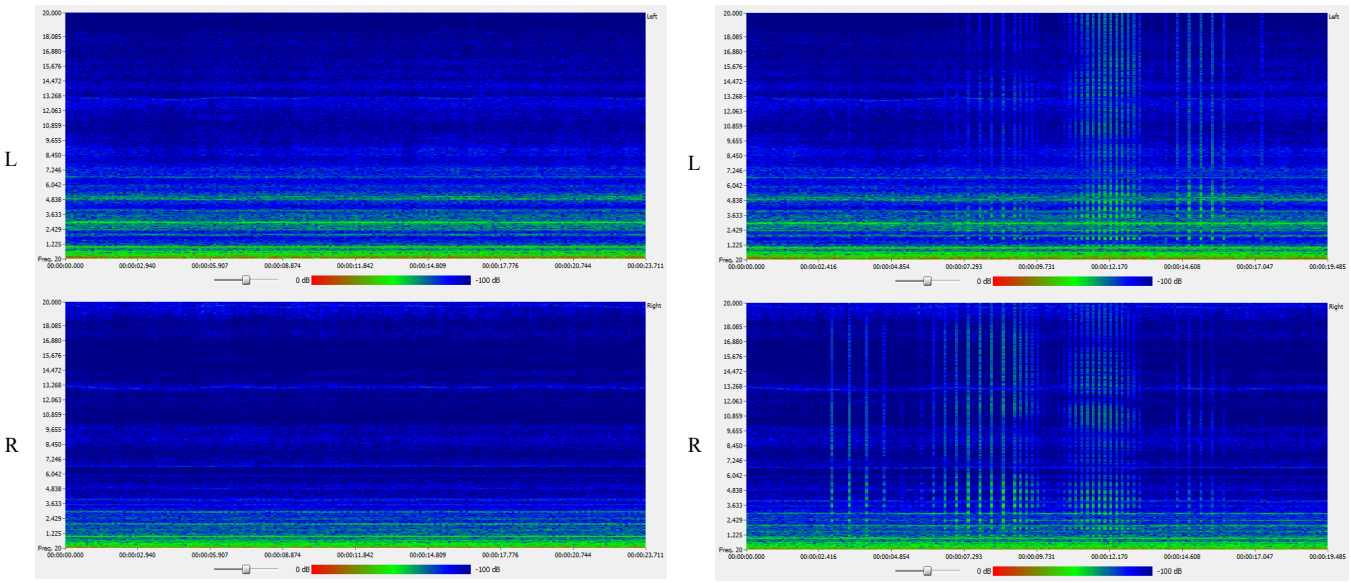


Figure 8. Left and right-ear sonograms of the binaurally-recorded helicopter background noise, 20 Hz – 20 kHz, alone, Figure 8 left, and with the spatially-rendered son1 sonification, Figure 8 right. The Figure 8 right sonograms are of the headphone output generated by the sample drift trajectory in the text. The background noise sonograms are shown for a similar period of time.

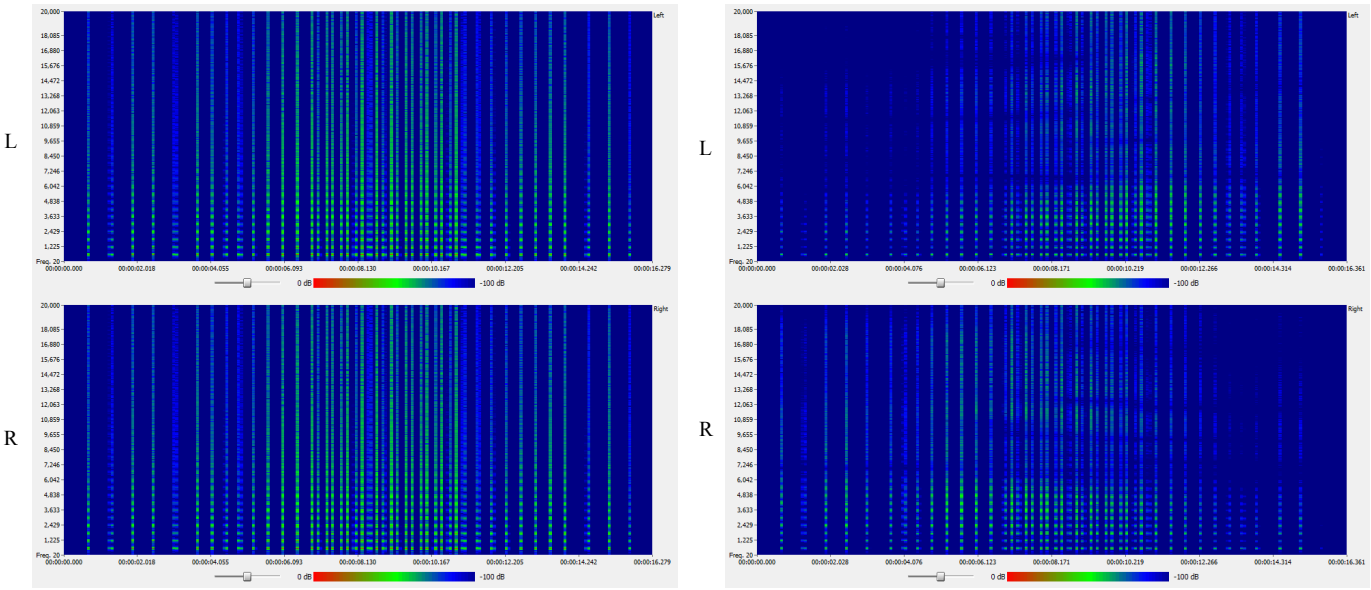


Figure 9. The son1 sonification sonograms without HRTF processing, 20 Hz – 20 kHz, Figure 9 left, and with HRTF processing, Figure 9 right. The wave files were generated using the sample drift trajectory in the text with the background noise disabled. The Figure 9 left sonograms show the integer harmonics of the fundamental created by using a sawtooth waveform as the son1 carrier. The Figure 9 right sonograms show multiple horizontal frequency notch paths imposed by neighboring HRTFs in the HRTF database as the obstacle-ownship azimuth angle changes along the sample trajectory.

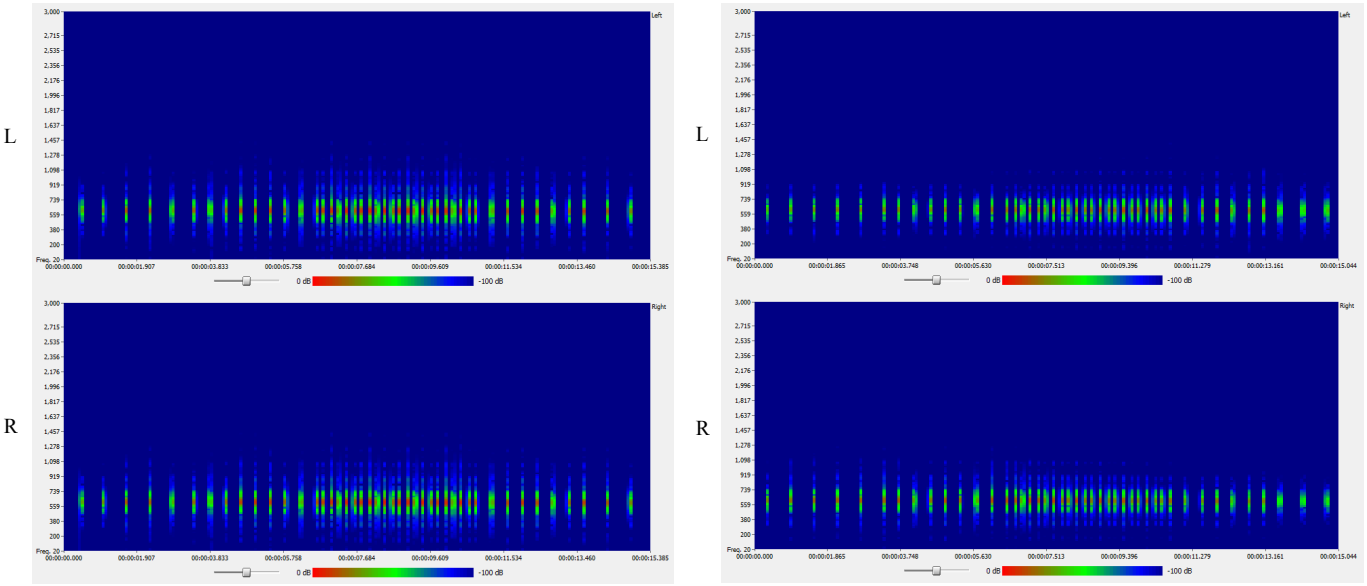


Figure 10. The son2 sonification sonograms without spatial processing, Figure 10 left, and with spatial processing, Figure 10 right, 20 Hz – 3 kHz. The wave files were generated using the sample drift trajectory in the text with the background noise disabled. The Figure 10 left sonograms show the fundamental frequency and lack of harmonics using a sinewave as the son2 carrier. The Figure 10 right sonograms display some spectral shaping due to HRTF processing along the sample trajectory, but greatly diminished relative to son1.

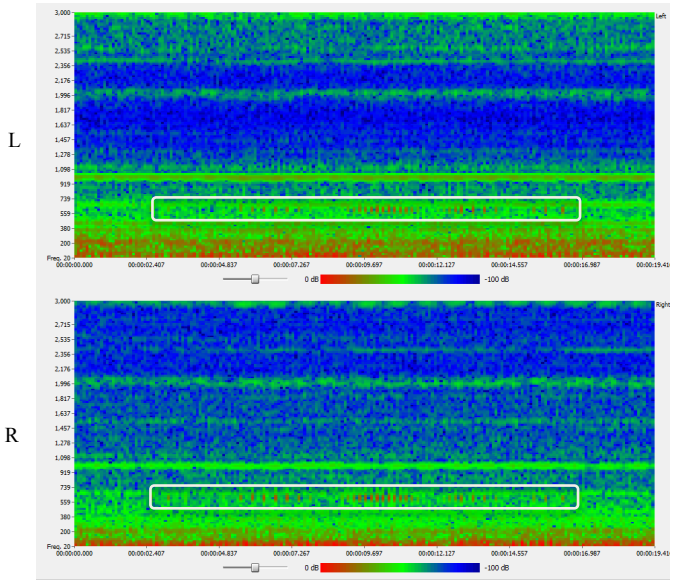


Figure 11. The son2 sonification sonograms for the headphone output corresponding to the sample drift trajectory in the text, 20 Hz – 3 kHz. The sonification is visible as the red dots in the highlighted regions. Although the son2 sonification looks somewhat masked in the sonograms, it is still quite audible.

The Auditory Stimuli

Helicopter Background Noise

A binaural recording in the UH-60 cockpit was conducted to characterize flight environmental noise, communications, and alerts. A Brüel & Kjær Type 4101-A Binaural Microphone was worn beneath an HGU-56/P helmet with the binaural signal recorded to a Sound Devices 702 Flash Recorder. A calibration recording was made using a 94 dB SPL 1000 Hz Brüel & Kjær Type 4231 Acoustical Calibrator coupled to the mics using a DP0978 coupler. This calibration recording allows the experiment playback system to be calibrated to the sound pressure level of the environmental noise recording. A 42 second segment of the 2-hour long binaural recording was selected that minimized mic movement noise, communications, and alerts. This segment was looped for the entire duration of an experimental session. A sonogram of the helicopter background noise loop is shown in Figure 8, left.

Sonifications

A series of earcons (abstract sounds with no prior association with referents, Ref. 8) were designed to enable audibility in the UH-60 cockpit background noise. A selection of the earcons most robust to masking was made in a preliminary phase of the study. Two of these earcons were chosen for the sonifications presented during the experiment. Both sonifications consisted of pulsed frequency-modulated waveforms with square-wave modulators (using the slab3d "fmpulser" signal generator). The pulse period was determined by obstacle-ownship distance relative to the drift alert rings in Table 1. The pulse shape had a duration of 80 ms with a 30 ms fade-in and a 30 ms fade-out. The carrier frequency was 600 Hz with a modulator amplitude of 110 Hz and a modulator frequency of 90 Hz. The modulator was band-limited to 11 harmonics and the resultant carrier waveform time index was parameter tracked (performed using a leaky integrator with a time constant of 15ms) so that the modulation was subtle and smooth. The only difference between the two sonifications was the choice of carrier waveform. The "son1" sonification (aka "Mosquito") used a non-band-limited sawtooth as the carrier (Figure 12, top). The "son2" sonification (aka "UFO") used a sinewave as the carrier (Figure 12, bottom). These two carrier waveforms represent two extremes in regards to frequency content. A sawtooth wave contains all integer harmonics of the fundamental frequency (Figure 9, left) whereas a sine wave contains no harmonics in addition to the fundamental frequency (Figure 10, left). If a digital sawtooth is not band-limited to half the sampling rate, the higher order harmonics can alias, providing additional frequency content in the audible range. For son1, since the resultant sound was not harsh or unpleasant, the choice to use a non-band-limited sawtooth was intentional so that the additional frequency content would better illuminate the spectral cues of the HRTF (Figure 9, right). For both sonifications, subtle frequency modulation was applied to further enrichen the timbral characteristics of their sound. Both sonifications stood out well relative to the UH-60 background noise (Figure 8, right, and Figure 11).

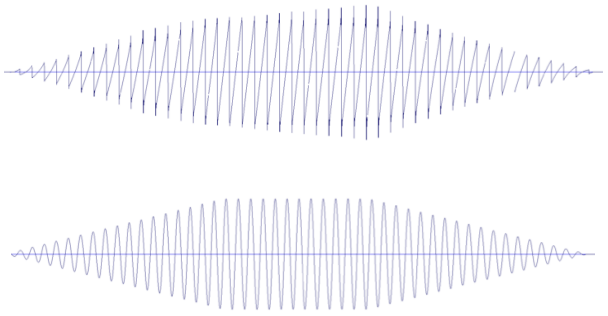


Figure 12. The son1 sonification with sawtooth carrier, top, and the son2 sonification with sinewave carrier, bottom. The pulse duration is 80 ms with a 30 ms fade-in and a 30 ms fade-out.

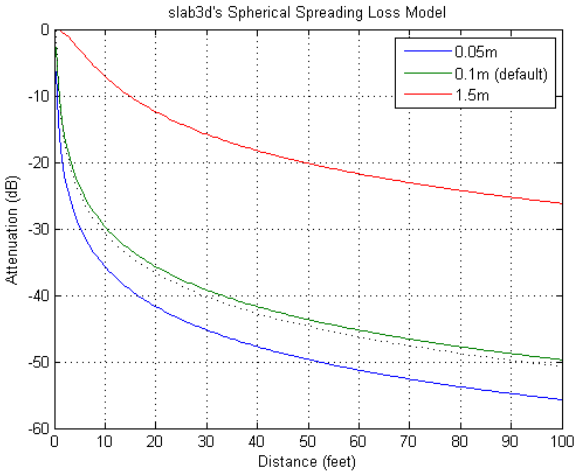


Figure 13. slab3d's source-listener distance gain model for the following source radii in meters (bottom curve up): 0.05, 0.1 (slab3d default), and 1.5 (drift scenario). Also shown is a dotted inverse-distance gain curve referenced to a 0.09 m interaural radius.

Looming Effect

Visual looming refers to the rate of change in the size of an approaching object's retinal image. A corresponding auditory "Looming Effect" occurs with an oncoming sound's increase in intensity over time. Therefore, it is advantageous for a visual object's sonification to share an overall stimulus

energy profile with the visual object (when visible). A sonification Looming Effect was achieved using slab3d's native spherical spreading loss model. This model computes a spreading loss gain attenuation of $(1 + d^2 / r^2)^{-s/2}$, where d is the source-listener (obstacle-ownship) distance, r is the source radius, and s is the spread exponent. When s is 1, this characteristic closely approximates that of a planar baffled cylindrical piston of radius r (Ref. 38). When the radius is set to the interaural radius, the model approximates point-source inverse-distance gain behavior. Values of s other than 1 lessen or exaggerate attenuation. For the drift dimensions under investigation (Figure 5), a source radius of 1.5 m (Figure 13, top curve) yielded sonification Looming Effects that were both audible and realistic for all drift trajectories.

Presentation

All sound presentation was managed by AvadeServer and displayed to the subject via an RME Fireface UFX USB audio peripheral and Sennheiser Precision HD 580 headphones. An experiment control application managed the trials, prompted the subject, and recorded subject responses and head tracker data. It communicated with AvadeServer via UDP command strings that selected the sonification type and updated obstacle and ownship locations using (latitude, longitude, height, yaw, pitch, roll) coordinates.

Latency

One reason for the selection of the RME Fireface as a playback device was to keep audio playback latency at a minimum. Given that three experiment applications were running on the same PC (Fastrak server, experiment control application, and AvadeServer), the typical AvadeServer RME buffer size of 64 samples was increased to 128 samples to improve playback robustness. This resulted in an estimated API latency (command-to-audio-result) of 7-10ms (Ref. 39). Since the size of the audio buffer is often the weakest link in the latency chain, it is important to keep its contribution to a minimum. Low total system latency, i.e., head motion to audio playback, is important for head-tracked virtual acoustic environments (Ref. 40).

Calibration

The AvadeServer pan option was used to calibrate headphone output to UH-60 background noise conditions using the 94 dB calibration recording described earlier. The headphones were placed on a Brüel & Kjær Sound Quality Head and Torso Simulator Type 4100 dummy head. The dummy head ear canal sound pressure level was measured using an Agilent 35670A Dynamic Signal Analyzer. The headphone gain of the RME Fireface was adjusted so that the calibration recording playback measured 94 dB SPL on the analyzer. The AvadeServer pan option was also used to present the binaural background noise unprocessed with an AvadeServer gain of 0 dB.

Headphone Levels

With the experiment playback system calibrated, the analyzer was used to measure the A-weighted SPL levels of the binaural background noise. Given that Sennheiser Precision HD 580 headphones employ an open design, this measurement also included leakage from the sound booth ambient noise (minimal) and the dome projector used to project the experiment reticle and subject instructions (more significant). The resulting values were left 77 dB SPL and right 75 dB SPL, values consistent with in-situ helmet-attenuated background noise in a UH-60 helicopter. These values were, however, considered a bit loud for extended subject exposure and were lowered 8 dB resulting in an experiment background noise level of left 69 dB SPL and right 67 dB SPL. A trajectory with the minimum obstacle-ownship distance was used to measure the loudest sonification levels. This yielded maximum sonification plus background noise levels on the order of 81 dB SPL.

A Sample Drift Trajectory

The behavior of the auditory display is illustrated in Figure 14 with one of the four predesigned drift trajectories that traveled the greatest distance, 161 ft.

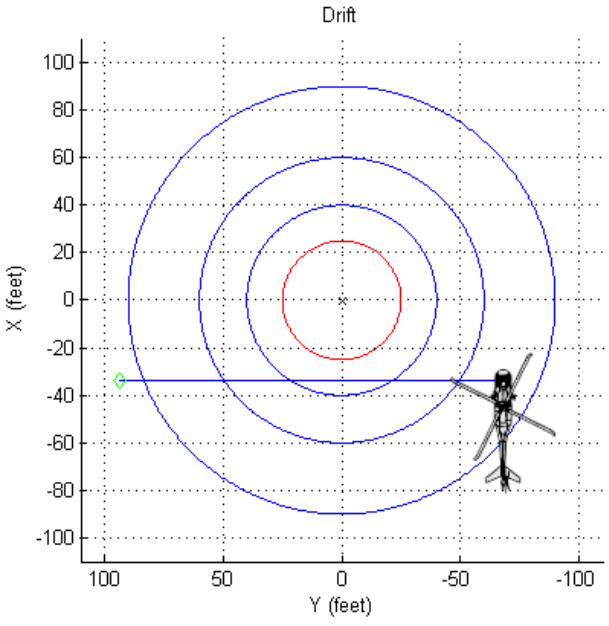


Figure 14. The sample drift trajectory under consideration is one of four that traveled the greatest distance, 161 ft.

The location and rendering parameters that follow were captured during an experiment trial. They are used to illustrate their corresponding effect upon the perceptual cues of: sonification pulse period, spatialization spectral shaping, ILD, and ITD, Looming Effect, and Doppler Shift.

Sonification Pulse Period

The sample drift trajectory headphone displays for the son1 and son2 sonifications, without background noise, are shown in Figure 15. At this scale, the most obvious characteristic is the sonification pulse period as the ownship drifts through the alert rings of Warning, Caution, Advisory, and Caution, stopping in Warning. The corresponding pulse periods are, in ms, 600, 400, 200, 400, and 600. The contributions of spherical spreading loss and ILD are also visible as a trial-length amplitude envelope. The spreading loss creates a visible Looming Effect as the ownship approaches the obstacle.

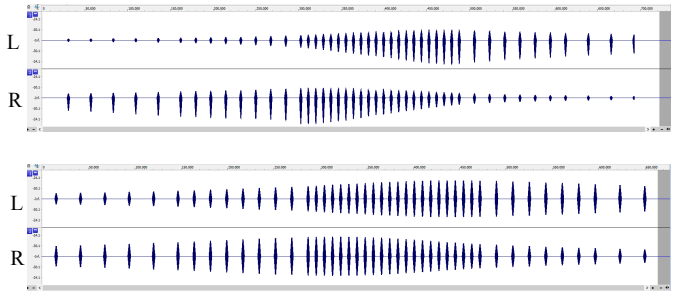


Figure 15. The headphone displays (without background noise) of the sonifications son1, top, and son2, bottom, for an entire trial of the sample drift trajectory in Figure 14.

Spatialization

The obstacle-listener bearing angles and head orientation angles during the sample trial are shown in Figure 16. Since the subject is told to remain motionless during the drift trajectory, head-tracked yaw and pitch should remain close to zero until drift completion. At drift completion, the subject orients towards the obstacle in a localization phase of the trial. Since the obstacle and ownship are both placed on the horizontal plane, elevation should remain near zero until localization. Thus, during drift, with the head mostly stationary, source azimuth will depend primarily on ownship location. For memory conditions (the condition shown), after drift completion, the azimuth and elevation are no longer updated because the sonification is not present. For continuous conditions, the azimuth and elevation are updated to reflect both final obstacle-ownship geometry and the orientation of the subject's head.

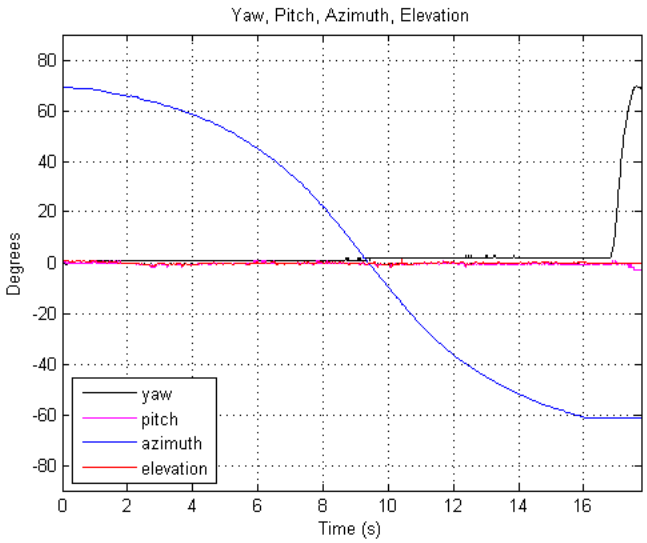


Figure 16. The yaw and pitch of the subject during the sample trial and the corresponding obstacle-listener azimuth and elevation bearing angles.

While the sonification is present, obstacle-listener azimuth and elevation are used to spatially render the sonification by performing an HRTF and ITD lookup and interpolation (see HRTF magnitude response in Figures 6 and 7 and ITDs in Figure 17). The resulting HRTF pair and ITD impart spectral shaping, level, and time delay differences between the left and right ears (time domain in Figure 15 and frequency domain in Figures 9 and 10, right). Given son1's frequency content, the spectral shaping is most evident in the son1 sonogram (Figure 9, right). Note, this spectral shaping is also responsible for son1's trial-length amplitude envelope appearing more pronounced than son2's in Figure 15, top.

Frequency-Independent ILD and ITD

For sound sources in the horizontal plane and for narrow-band signals, e.g., son2, the frequency-independent effects of ILD and ITD play a critical role in spatialization. The spatial rendering performed by slab3d uses a frequency-independent ITD extracted at measurement and saved in the HRTF database. Frequency-independent ILD can be visualized by examining the RMS values of the corresponding HRIRs. The horizontal-plane values for jdm.slh and mg.slh, and the database ITDs, are shown in Figure 17. These ITD and ILD cues allow the son2 sonification to be an effective spatial sonification even though little spectral content is available for spectral shaping.

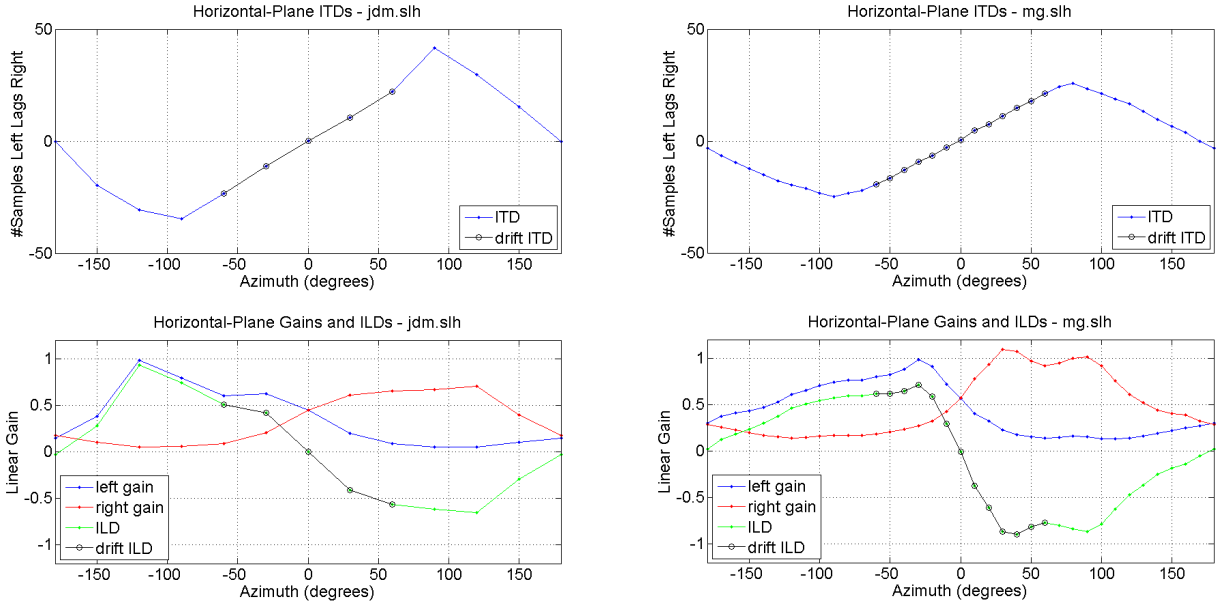


Figure 17. Frequency-independent ITD and ILD values for the default (and individualized) HRTF database jdm.slh, left, and the individualized HRTF database mg.slh, right. The sample drift trajectory results in the roughly 60° to -60° azimuth span highlighted by the circled portions of the ITD and ILD plots.

Looming Effect and Doppler Shift

In addition to bearing angles, another scene geometry parameter calculated during the drift trajectory is *obstacle-ownship distance* (Figure 18). As discussed earlier, this distance is used to determine slab3d's spherical spreading loss.

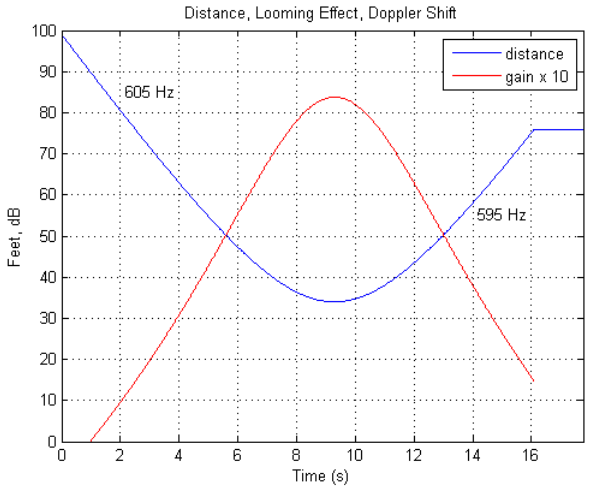


Figure 18. Obstacle-ownship distance (blue curve) and impact on Looming Effect (red curve) and Doppler Shift.

In the context of drift trajectory sonification, the resulting gain along the trajectory relative to the initial gain at alert threshold crossing allows calculating the Looming Effect. In slab3d, obstacle-ownship distance also indexes a sound source propagation delay line. The corresponding compression and rarefaction in air is mirrored in the delay line such that the constant drift velocity of 10 ft./s results in a Doppler Shift of ±5 Hz relative to the sonification fundamental frequency of 600 Hz.

The Participants

Four participants took part in the study (two men, two women, age 25 to 54, with normal audiometric capacities, allowing for typical age-related differences). They include two of the authors (*MGC* and *JDM*).

Each participant experienced both Memory and Continuous conditions and the two types of sonifications, son1 and son2 (repeated measures design). In the Continuous condition, the participants had to localize the obstacle while the sonification remained present at the end of the simulated drift. In the Memory condition, the sonification stopped at the end of the simulated drift, before localization is permitted. The two conditions were run separately.

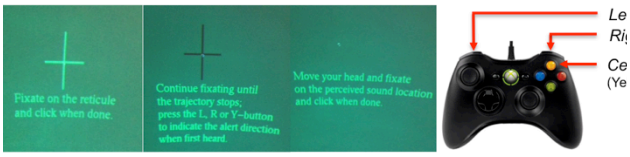
Each block contained 690 trials, corresponding to 2 sonification types (son1 vs. son2) * 5 off-axis destinations

(Figure 5) * 6 off-axis directions (North, South, East, West, Left-Diagonal, Right-Diagonal) * 2 hemispheres (Left, Right) + 3 axis destinations (*+X, -X, -Y*) * 3 ranges (32.5ft, 50ft, 75ft)=138 trials repeated 5 times=690 trials per condition (Memory vs. Continuous). Final drift trajectory endpoints were associated with obstacle azimuths (in relation to the observer) of -90°, -63°, -45°, -27°, 0°, 27°, 45°, 63° and 90°. A trial lasted on average 10 sec, and a block took on average 28 minutes to be completed (138 trials). Participants were instructed to take a break or interrupt a session (one block) at their convenience.

There were five repetitions per condition and the order of presentation was counterbalanced to minimize the effect of experience. Experimental sessions were limited to 45 minutes/session. The written instructions for participants were reviewed with the experiment proctor prior to starting the study. Practice trials were administrated before beginning the formal study as needed for each participant.

The task

Figure 19 illustrates the succession of events occurring during a trial. Instructions were projected on the hemispherical surface of the projection dome. An Xbox® controller allowed the participants to trigger the succession of the events constituting a trial that were depicted on the half-spherical dome. A boresight was performed by having the participants fixate on a reticle displayed at the center of the dome surface for a random duration of 1 to 2.5 sec. After this period, one of the two sonifications originated from one of the directions, within one of the three distance rings. The participants were instructed to keep fixation the reticle, which changed color from white to red to indicate the new state. At the end of the trajectory, the reticle disappeared and the participant were instructed to move their heads/eyes toward the perceived sound location and keep fixating until they validated their response by pressing the L, R or Y control buttons. This validation ended a trial.



1. *Fixate on the reticle and click when done* (either L, R or Y control)
2. *Continue fixating until the trajectory stops; press the L, R or Y-button to indicate the alert direction when first heard* (origin of the sound)
3. *Move your head and fixate on the perceived sound location and click when done* (either L, R or Y control)

Figure 19. Succession of events constituting a trial.

The auditory stimuli (UH-60 background noise + sonifications) were delivered over headphones. As noted above, the maximum sonification plus background noise levels were on the order of 81 dB SPL. These levels are generally not hazardous to hearing since they are below OSHA requirements for noise exposure and noise dosage for the daily exposure durations in this study[1]. Care was taken to limit the dB levels at which sounds are presented via calibration prior to each testing session. The localization responses were recorded both in head-tracked and eye-tracked egocentric coordinate systems, allowing for collection and cross-verification of each response measure.

The Measures of Performance

The measures of precision and accuracy

The raw data consisted of the 2D coordinates of the terminal position of the head/gaze relative to target. Outliers (+/- 3 *SD* from the mean) were removed for each target location, each condition, each sonification type and each subject to control for intra-individual variability (3.27% of the raw data). To test the hypothesis of colinearity between the *x* and *y* components of the localization responses, a hierarchical multiple regression analysis was performed. Tests for multicollinearity indicated that a very low level of multicollinearity was present [variance inflation factor (*VIF*) = 1 for the both Memory and Continuous conditions]. Results of the regression analysis provided confirmation that the data were governed by a bivariate normal distribution (i.e., 2 dimensions were observed). To analyze the endpoint distributions, we determined the covariance matrix of all the 2D responses (*x* and *y* components) for each target eccentricity, each sonification type (son1 vs. son2), each condition (Memory vs. Continuous) and each HRTF type (Generic vs. Individualized). The 2D variance (σ_{xy}^2) represents the sum of the variances in the two orthogonal directions ($\sigma_{xy}^2 = \sigma_x^2+\sigma_y^2$). The distributions were visualized by 95% confidence ellipses. We calculated ellipse orientation (θ_a) as the orientation of the main eigenvector (*a,* Eigmax), which represents the direction of maximal dispersion, while *b* represents the variance in the perpendicular axis (minimum dispersion, Eigmin). Because an axis is an undirected line where there is no reason to distinguish one end of the line from the other, the data were computed within a +/-180º range, 90º corresponding to a vertical orientation and +/-180º to an horizontal orientation. A measure of *anisotropy* of the distributions, ε, was provided, a ratio value close to 1 indicating no preferred

[1] https://www.osha.gov/pls/oshaweb/owadisp.show_document?p_table=standards&p_id=9735

direction, and a ratio value close to 0 indicating a preferred direction:

$$\varepsilon = \sqrt{1 - (b/a)^2} \qquad \text{(Equation 1)}$$

where a and b represent respectively the maximum (main) and the minimum eigenvectors.

For the measure of localization accuracy, the difference between the actual 2D target position and the centroid of the distributions was computed, providing an error vector $\vec{a}$ that can be analyzed along its length (or amplitude, r) and angular direction (α). The traditional mean (μ), standard deviation (σ^2) and standard error (σ) are reported for all the measures.

The Statistical Analyses

Univariate and repeated measures analyses of variance (ANOVAs) were used to test for the effects of condition, sonification type, HRTFs type and obstacle eccentricity in azimuth. A custom post-experiment questionnaire was administrated to evaluate subjective system usability and acceptability. All of the effects described here were statistically significant at p <.05 or better.

RESULTS

Quantitative Measures

The local characteristics of precision and accuracy are illustrated in Figures 20 and 21 and summarized in Table 2. Generic and Individualized HRTF types were analyzed separately before being compared for participant *MGC*. The effects of sonification type (son1 vs. son2) and condition (Memory vs. Continuous) were assessed for each HRTF type, generic vs. individualized.

Generic HRTFs

Precision

It can be seen from Figure 20 that auditory localization was generally characterized by anisotropic response distributions (ε: μ=.65, σ =.01, isotropic responses would have a value of 1) oriented mostly vertically (126º, 90º representing a true vertical orientation) for obstacles with a 0º azimuth (in the Cartesian coordinate system) and mostly oriented horizontally (165º, 180º representing a true horizontal orientation) for the most peripheral targets.

These scatter properties emphasize the fact that azimuth and elevation localization are dissociate processes (see Introduction) that are associated with independent variations in spatial resolution as a function of a sound location in space.

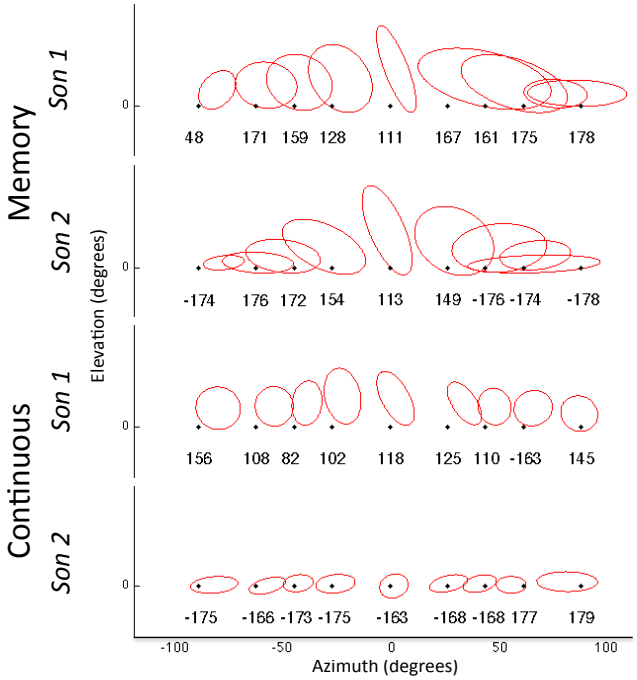


Figure 20: Generic HRTFs. Precision: the distributions of the responses are visualized by 95% confidence ellipses.

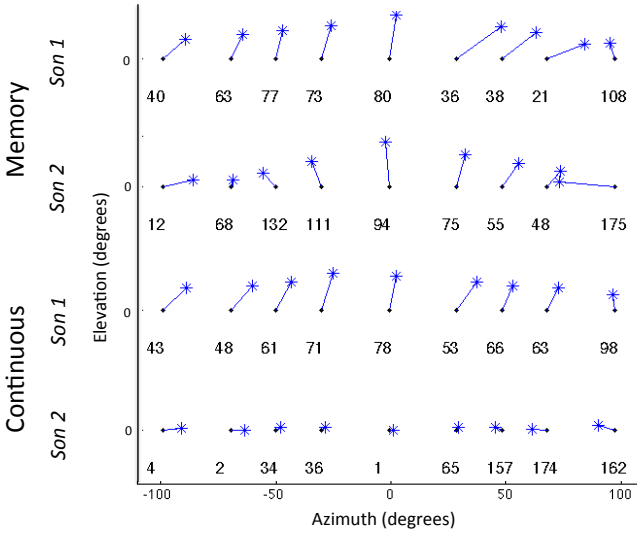


Figure 21: Generic HRTFs. Accuracy (CE) as a function of the condition of presentation (Memory vs. Continuous) and as a function of sonification type (son1 vs. son2).

For targets located straight ahead (0º azimuth), precision was greater in the horizontal than in the vertical dimension. Conversely, for the most peripheral targets (+/-90º azimuth), the opposite pattern was generally observed, with a greater precision in elevation than in azimuth. The distributions (ellipses) were more anisotropic for son1 than for son2 (ε: son1: μ=.55, σ =.04; son2: μ=.73, σ =.02; son1, son2: t=-.17, p=.03), and in the Memory than in the Continuous condition (ε: Memory: μ=.71, σ =.03; Continuous: μ=.57, σ =.02; son1, son2: t=.13, p=.04).

Table 2. Generic HRTFs. Precision (VE, [μ (σ^2)]), Accuracy (CE, [μ (σ^2)]) and Eigenvalues [μ (σ^2)] as a function of the condition of presentation (Memory vs. Continuous) and as a function of sonification type (son1 vs. son2)

	Memory		*Continuous*	
	Son1	Son2	Son1	Son2
Precision VE	29.16 (8.11)	27.68 (7.41)	18.46 (1.89)	13.21 (2.47)
Eigmin	9.62 (3.15)	7.93 (3.74)	7.80 (1.30)	4.15 (.63)
Eigmax	19.54 (6.77)	19.57 (5.84)	10.66 (1.94)	9.05 (2.39)
Accuracy CE	14.05 (8.11)	11.80 (5.74)	12.05 (2.68)	4.05 (2.44)

In agreement with ellipses properties, obstacle localization was marginally more precise with son2 than with son1 (son1: μ=23.82, σ =.96; son2: μ=20.55, σ =1.70; son1, son2: t=3.26, p=.05) essentially the result of a reduction of the variance in the elevation component of the responses (Eigmin: son1: μ=8.70, σ =.43; son2: μ=5.79, σ =.76; son1, son2: t=2.90, p=.03, Eigmax: son1: μ=14.63, σ =1.39; son2: μ=14.01, σ =1.14; son1, son2: t=.61, p=.71). This difference was exacerbated in the Continuous condition, with a 3.4% improvement in precision for Memory vs. a 2.9% improvement in precision for Continuous. Similarly, obstacle localization was more precise in the Continuous than in the Memory condition (Memory: μ=28.48, σ =2.59; Continuous: μ=15.89, σ =.35; Memory, Continuous: t=12.58, p=.01), an performance improvement essentially attributable to a reduction of the variance along the main Eigenvector (Eigmin: Memory: μ=8.77, σ =1.03; Continuous: μ=5.97, σ =.21; Memory, Continuous: t=2.79, p=.05, Eigmax: Memory: μ=19.65, σ =1.89; Continuous: μ=9.86, σ =.43; Memory, Continuous: t=9.79, p=.001). The overall gain in localization precision in the most efficient combination (son2 * Continuous) represented 54% of the variance in the less efficient combination (son1 * Memory). There was no effect of interaction between condition and sonification ($F_{1,4}$ =5.36, p=.08). The overall gain in localization precision in the most efficient combination (son2 * Continuous) represented 54% of the precision reported in the less efficient combination (son1 * Memory).

There was no systematic effect of eccentricity on the magnitude of the precision ($F_{4,4}$ =1.02, p=.49). Note however, that the magnitude of the variable error was maximum for 27º and 45º azimuths in the Memory condition, a phenomenon that was not apparent in the Continuous condition (see Figure 22 below).

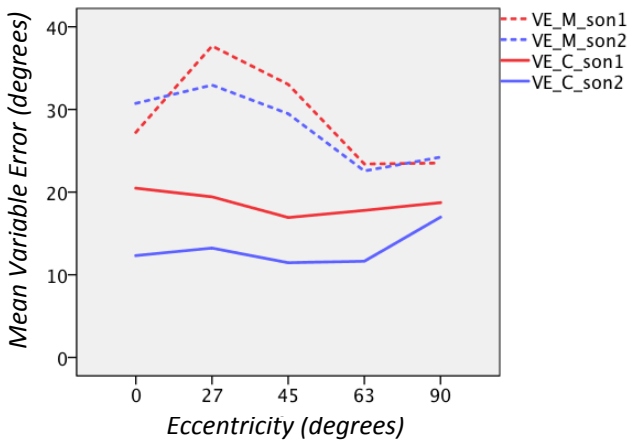


Figure 22: Generic HRTFs. Precision as a function of obstacle eccentricity.

Accuracy

Auditory localization accuracy in the HMP was characterized by an overshoot of the responses in elevation, as seen in Figure 21, where the error vector directions are oriented upward relative to the initial fixation point (μ=69.76, σ =14.28), irrespective of condition, sonification type or azimuth. This result is congruent with previous research suggesting that the A and the V "horizons" may not coincide, as was previously reported, though not discussed (Ref. 13, Ref. 41). Note that this effect vanished in the Continuous/ son2 condition. It can be seen from Figure 21 that a systematic rightward shift of the error vectors essentially with son1, both in the Continuous and in the Memory conditions (vectors directions comprised between 90º and 270º represent leftward shifts in relation to the target, vectors directions comprised between 270º and 90º represent leftward shifts in relation to the target). For son2, we observed a slight azimuth overshoot in the Memory condition (i.e. the error vector directions are in the direction of the targets relative to the initial fixation point). In the Continuous condition, conversely, the accuracy was

characterized by a slight undershoot in azimuth (i.e. the vectors direction was opposite to the direction of the target).

Auditory localization (μ=10.72, σ =.43) was significantly more accurate in the Continuous than in the Memory condition (Memory: μ=13.40º, σ =1.11; Continuous: μ=8.03º, σ =.49; Memory, Continuous: t=5.37, p=.02). Son2 was associated with a greater accuracy than son1 (son1: μ=13.31º, σ =.86; son2: μ=8.12º, σ =.74; son1, son2: t=5.19, p=.01). We observed a significant effect of interaction between condition and sonification type (condition * sonification: $F_{1,4}$ =15.91, p=.01). Indeed, sonification type did not not affect auditory accuracy in the Memory condition (son1, son2: t=2.01, p=1). Conversely, in the Continuous condition, auditory localization was more accurate by a factor of 3 when using son2 as compared to son 1(son1, son2: t=8.38, p=.005).

As for precision, there was no systematic effect of eccentricity on localization accuracy ($F_{4,4}$ =2.16, p=.23). Although the interaction didn't reach significance, it can be seen from Figure 23 that accuracy performance with son2 decreases as eccentricity increases, an effect opposite with son1. The overall gain in localization accuracy in the most efficient combination (son2 * Continuous) represented 71% of the accuracy reported in the less efficient combination (son1 * Memory).

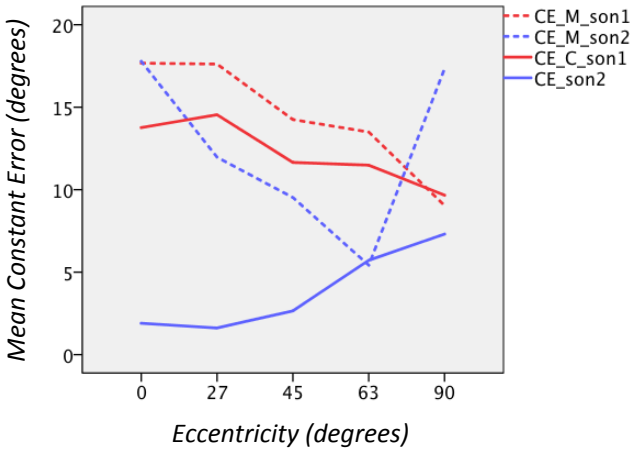


Figure 23: Generic HRTFs. Accuracy as a function of obstacle eccentricity.

Individualized HRTFs

Precision

The topological properties of the localization response distributions are differed to those observed in the generic condition, with ellipses mostly oriented horizontally (175º, 180º representing a true horizontal orientation), as seen in Figure 24.

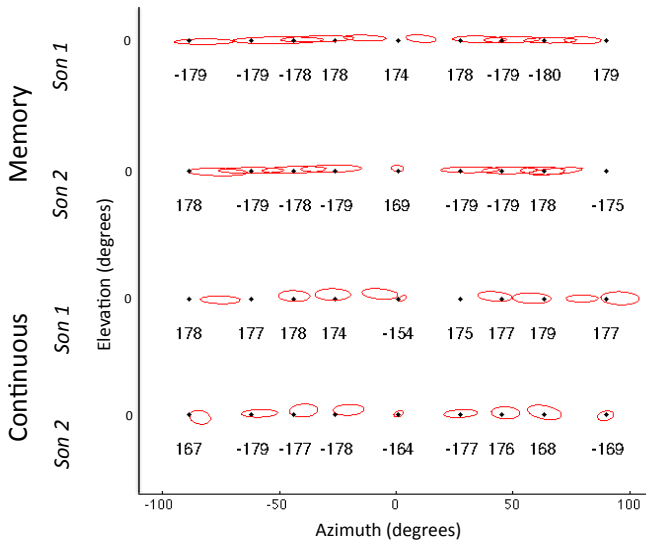


Figure 24: Individualized HRTFs. Precision: the distributions of the responses are visualized by 95% confidence ellipses.

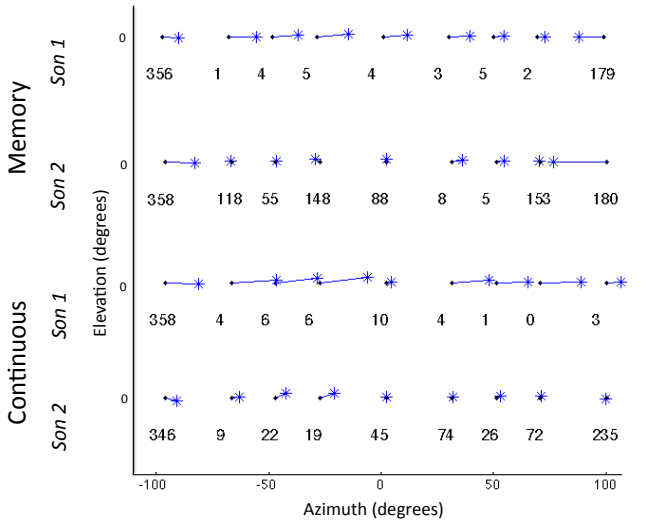


Figure 25: Individualized HRTFs. Accuracy (CE) as a function of the condition of presentation (Memory vs. Continuous) and as a function of sonification type (son1 vs. son2).

The ellipses were significantly more anisotropic in the Continuous than in the Memory condition (ε: Memory: μ=.91, σ =.005; Continuous: μ=.76, σ =.007; t=-.14, p<.0001) and marginally narrower for son2 than for son1 (ε: son1: μ=.86, σ =.009; son1: μ=.81, σ =.01; t=.04, p=.05). Localization was significantly more precise with son2 than

with son1 (son1: μ=13.06, σ =1.12; son2: μ=8.54, σ =.69; son1, son2: t=4.51, p<.0001. A decomposition of the variance as a function of the eigenvectors reveals a paradoxical effect: while the variance along the min eigenvector is significantly reduced with son2 compared to son1, the opposite effect is observed along the main eigenvector, where the variance is greater with son2 than with son1 (Eigmin: son1: μ=1.38, σ =.01; son2: μ=2.14, σ =.14; son1, son2: t=-.76, p=.001, Eigmax: son1: μ=11.67, σ =1.13; son2: μ=6.39, σ =.58; son1, son2: t=5.27, p<.001).

Similarly, precision was not quite statistically different between conditions (Memory: μ=12.37, σ =.93; Continuous: μ=8.00, σ =.08; Memory, Continuous: t=.80, p=.10). The separate analysis of the variance error along the two eigenvectors showed a marginal reduction of the variance for the main eigenvector in the Continuous condition (Eigmin: Memory: μ=1.69, σ =.09; Continuous: μ=1.84, σ =.12; Memory, Continuous: t=-.14, p=.39, Eigmax: Memory: μ=9.51, σ =.85; Continuous: μ=8.56, σ =.85; Memory, Continuous: t=.94, p=.05). There was no effect of interaction between condition and sonification ($F_{1,4}$ =2.62, p=.18). The effect of eccentricity (Figure 26) on the magnitude of the precision was marginally significant ($F_{4,4}$ =6.31, p=.05).

See Table 3 for a detailed summary of the dependent variable descriptive statistics as a function of condition and sonification.

The overall gain in localization precision in the most efficient combination (son2 * Continuous) represented 40% of the precision reported in the less efficient combination (son1 * Memory).

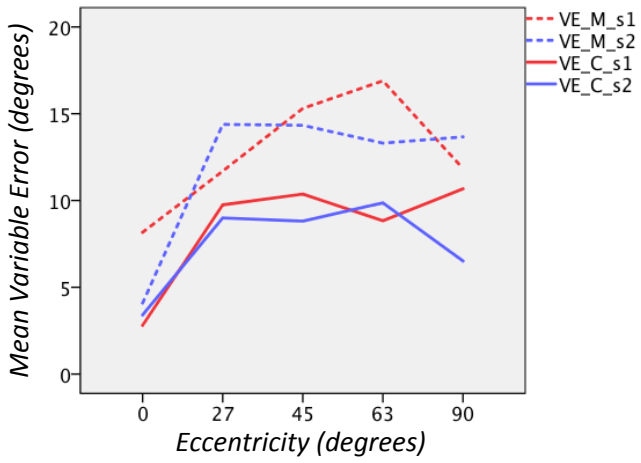


Figure 26: Individualized HRTFs. Precision as a function of obstacle eccentricity.

Accuracy

Auditory localization accuracy in the HMP was characterized by an horizontal bias, i.e. by error vectors oriented mostly horizontally, as seen in Figure 25.

Table 3. Individualized HRTFs. Precision (VE, [μ (σ^2)]), Accuracy (CE, [μ (σ^2)]) and Eigenvalues [μ (σ^2)] as a function of the condition of presentation (Memory vs. Continuous) and as a function of sonification type (son1 vs. son2)

	Memory		*Continuous*	
	Son1	Son2	Son1	Son2
Precision VE	13.29 (3.98)	12.83 (3.51)	9.11 (2.48)	7.97 (2.17)
Eigmin	1.34 (.13)	1.43 (.15)	2.04 (.52)	2.25 (.59)
Eigmax	11.95 (4.00)	11.40 (3.48)	7.07 (2.10)	5.27 (1.93)
Accuracy CE	8.53 (3.31)	5.13 (7.25)	13.39 (5.82)	2.60 (2.21)

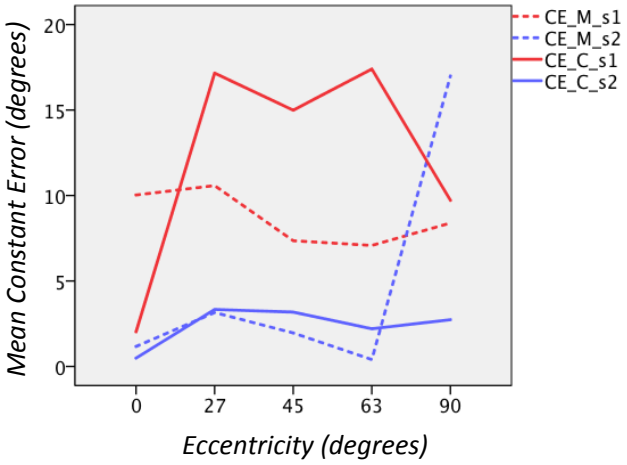


Figure 27: Individualized HRTFs. Accuracy as a function of obstacle eccentricity.

One can see a systematic rightward shift of the error vectors in particular with son 1 (vectors directions comprised between 90º and 270º represent leftward shifts in relation to the target, vectors directions comprised between 270º and 90º represent leftward shifts in relation to the target). However, the magnitude of the error vectors was significantly smaller with son2 than with son1, i.e. localization was more accurate with son2 than with son1 (son1: μ=10.47º, σ =1.08; son2: μ=3.56º, σ =.41; son1, son2: t=6.91, p=.002). Localization accuracy was not

significantly different between condition (Memory: μ=6.71º, σ =98; Continuous: μ=7.32º, σ =1.08; Memory, Continuous: t=-.61, p=.72). However, we observed a significant interaction effect between condition and sonification (condition * sonification: $F_{1,4}$=11.43, p=.02).

The overall gain in localization accuracy in the most efficient combination (son2 * Continuous) represented 69% of the accuracy reported in the less efficient combination (son1 * Memory).

Generic vs. Individualized

One participant (S3) ran the experiment twice, once with her own HRTFs (mg.slh), which were used for the general analyses, and once with the generic HRTFs (jdm.slh). These two datasets provide the opportunity to evaluate the benefits associated with HRTF individualization.

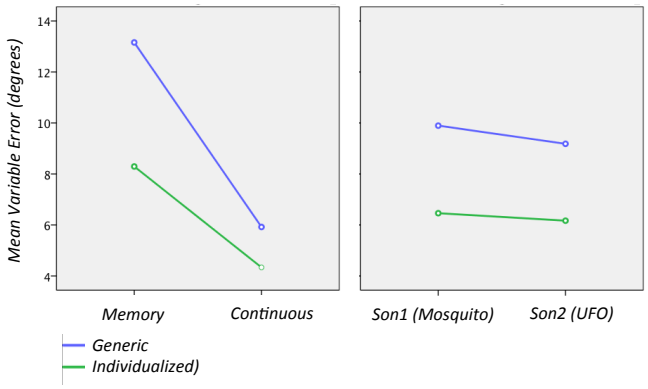


Figure 28: Generic vs. Individualized HRTFs. Precision (VE) as a function of the condition of presentation (Memory vs. Continuous) and as a function of sonification type (son1 vs. son2).

Table 4. Generic vs. Individualized HRTFs. Precision (VE) and Accuracy (CE) as a function of the condition of presentation (Memory vs. Continuous) and as a function of sonification type (son1 vs. son2)

		Memory		*Continuous*	
		Son1	Son2	Son1	Son2
Generic	Precision VE	14.26 (4.54)	13.45 (5.76)	6.23 (1.20)	5.70 (.98)
	Accuracy CE	13.97 (5.27)	8.28 (5.70)	15.59 (6.87)	4.80 (3.71)
Individual	Precision VE	8.57 (1.52)	8.52 (1.75)	4.63 (.72)	4.22 (.97)
	Accuracy CE	10.00 (5.78)	8.41 (10.37)	12.08 (4.76)	3.03 (2.23)

Precision

HRTF individualization was associated with a significant reduction in localization variance (Generic: μ=9.95, σ =.86; Individualized: μ=6.31, σ =.30; Generic, HRTF: t=3.22, p=.01). There was no significant effect of interaction (HRTF * Condition: $F_{1,4}$=.85, p=.55; HRTF * son: $F_{1,4}$=.20, p=.67; see Figure 28; HRTF * eccentricity: $F_{1,4}$=1.79, p=.29).

The overall gain in localization precision in the most efficient combination (son2 * Continuous) represented 60% of the precision reported in the less efficient combination (son1 * Memory).

Accuracy

Similarly, auditory localization was more accurate in the Individualized than in the Generic condition (Generic: μ=10.37, σ =1.91; Individualized: μ=8.06, σ =1.29; Generic, HRTF: t=2.30, p=.08, marginal significance). There was no effect of interaction between HRTF type and Condition (HRTF * Condition: $F_{1,4}$ =.26, p=.63), HRTF type and Sonification (HRTF * Sonification: $F_{1,4}$ =4.12, p=.11), as seen in Figure 29. The interaction between HRTF type and eccentricity was not significant ($F_{1,4}$=1.67, p=.31).

The overall gain in localization accuracy in the most efficient combination (son2 * Continuous) represented 50% of the accuracy reported in the less efficient combination (son1 * Memory).

See Table 4 for a detailed summary of descriptive statistics for precision and accuracy as a function of HRTF type, condition, and sonification.

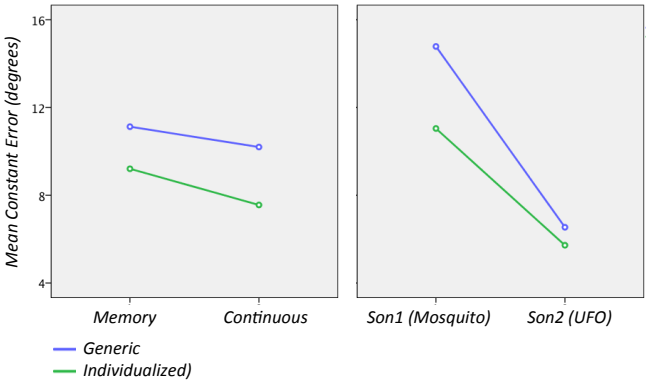


Figure 29: Generic vs. Individualized HRTFs. Accuracy (CE) as a function of the condition of presentation (Memory vs. Continuous) and as a function of sonification type (son1 vs. son2).

Interindividual Differences

Generic

Overall, auditory localization precision was not statistically different between the two participants (S1: μ=15.57, σ =1.35; S2: μ=18.96, σ =1.35; S1, S2: t=-3.38, p=.09). However, S1's precision was independent of the sonification type (son1: μ=14.16, σ =.62; son2: μ=16.98, σ =2.05; $F_{1,8}$ =1.55, p=.24) while greater in the Continuous than in the Memory condition (Memory: μ=18.19, σ =1.83; Continuous: μ=12.95, σ =.79; $F_{1,16}$=6.90, p=.03).

For participant S1, the overall gain in localization precision in the most efficient combination (son1 * Continuous) represented 39% of the precision reported in the less efficient combination (son2 * Memory). For participant S2, the overall gain in localization precision in the most efficient combination (son1 * Continuous) represented 66% of the accuracy reported in the less efficient combination (son2 * Memory).

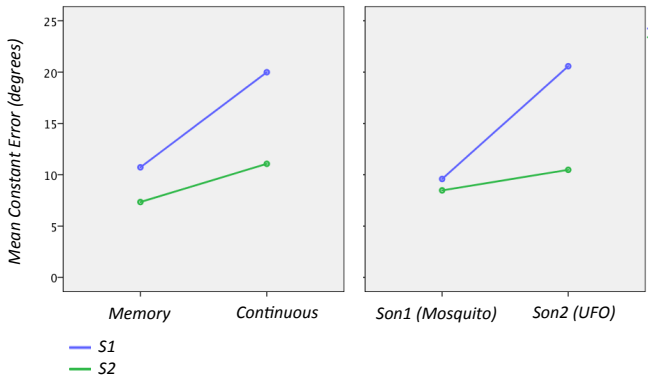


Figure 30: Generic HRTFs. Interindividual variability. Accuracy (CE) as a function of the condition of presentation (Memory vs. Continuous) and as a function of sonification type (son1 vs. son2).

The difference in accuracy between the two participants was significant (S1: μ=9.03, σ =1.34; S2: μ=15.52, σ =1.34; S1, S2: t=-6.49, p=.004). It can be seen from Figure 30 that the magnitude of the differences between conditions and between sonification type was greater for S1 than for S2 (Participant * Condition: $F_{1,16}$ =5.94, p=.02; Participant * Sonification: $F_{1,16}$ =7.31, p=.01). However, we observe an inversion in the polarity of the differences in the son2 Continuous condition (Participant * Condition * Sonification: $F_{1,16}$=10.35, p=.005).

See Table 5 for a detailed summary of descriptive statistics for precision and accuracy as a function of subject (listening to generic HRTFs), condition and sonification.

Table 5. Precision (VE) and Accuracy (CE) as a function of the condition of presentation (Memory vs. Continuous) and as a function of sonification type (son1 vs. son2)

		Memory		*Continuous*	
		Son1	Son2	Son1	Son2
S1	Precision VE [μ (σ)]	15.80 (2.91)	20.59 (3.13)	12.51 (.59)	13.86 (.92)
	Accuracy CE [μ (σ)]	11.68 (2.30)	9.77 (3.42)	7.50 (1.20)	7.18 (.87)
S2	Precision VE [μ (σ)]	28.66 (2.91)	23.14 (3.13)	14.52 (.59)	9.51 (.92)
	Accuracy CE [μ (σ)]	22.14 (2.30)	17.84 (3.43)	19.01 (1.20)	3.12 (.87)

Individualized

Localization responses were significantly more precise for S3 than for S4 (S3: μ=6.48, σ =.89; S4: μ=12.01, σ =.89; S3, S4: t=-5.52, p<.0001), a relative superiority that remains true for both conditions and sonification types (Participant * Condition: $F_{1,16}$ =.72, p=.40; Participant * Sonification: $F_{1,16}$ =1.25, p=.28), as seen in Figure 31.

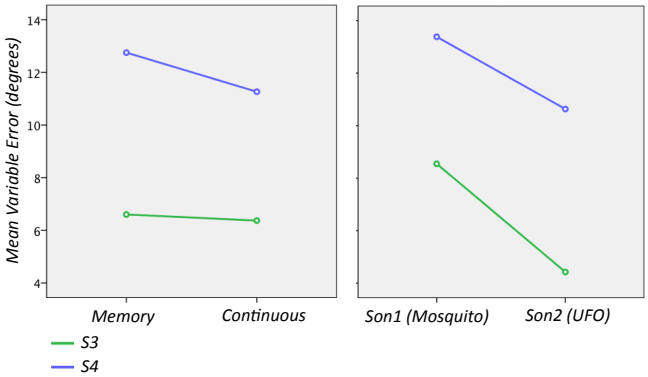


Figure 31: Individualized HRTFs. Interindividual variability. Precision (VE) as a function of the condition of presentation (Memory vs. Continuous) and as a function of sonification type (son1 vs. son2).

For participant S3, the overall gain in localization precision in the most efficient combination (son1 * Continuous) represented 57% of the precision reported in the less efficient combination (son2 * Memory). For participant S2, the overall gain in localization precision in the most efficient combination (son1 * Continuous) represented 30% of the precision reported in the less efficient combination (son2 * Memory).

For accuracy (Figure 32), there was no overall significant difference in precision between participants (S3: μ=8.38, σ =1.07; S4: μ=9.06, σ =1.07; S3, S4: t=-.68, p=.66). There was no interaction with Condition (Participant * Condition: $F_{1,16}$ =2.28, p=.15) or Sonification type (Participant * Sonification: $F_{1,16}$ =1.58, p=.22). For participant S3, the overall gain in localization accuracy in the most efficient combination (son2 * Continuous) represented 75% of the accuracy reported in the less efficient combination (son1 * Continuous). For participant S4, the overall gain in localization accuracy in the most efficient combination (son1 * Continuous) represented 77% of the accuracy reported in the less efficient combination (son2 * Memory).

See Table 6 for a detailed summary of descriptive statistics for precision and accuracy as a function of subject (listening to individualized HRTFs), condition, and sonification.

Table 6. Precision (VE) and Accuracy (CE) as a function of the condition of presentation (Memory vs. Continuous) and as a function of sonification type (son1 vs. son2)

		Memory		*Continuous*	
		Son1	Son2	Son1	Son2
S3	Precision VE [μ (σ)]	8.57 (1.52)	8.52 (1.75)	4.63 (.72)	4.22 (.97)
	Accuracy CE [μ (σ)]	10.00 (5.78)	8.41 (10.37)	12.08 (4.76)	3.03 (2.23)
S4	Precision VE [μ (σ)]	14.15 (7.29)	12.62 (5.36)	11.35 (1.84)	9.91 (3.27)
	Accuracy CE [μ (σ)]	8.81 (2.89)	6.27 (3.91)	17.21 (4.45)	3.96 (4.37)

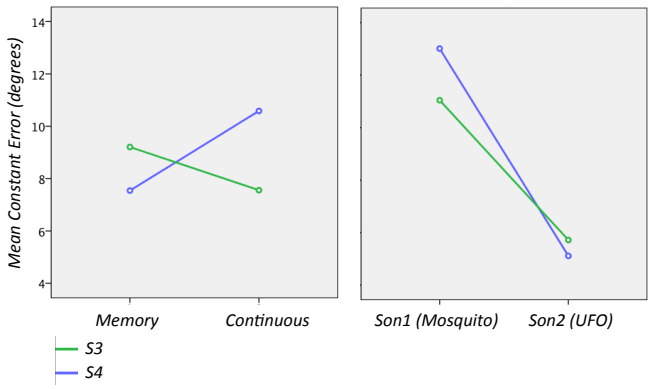


Figure 32: Individualized HRTFs. Interindividual variability. Accuracy (CE) as a function of the condition of presentation (Memory vs. Continuous) and as a function of sonification type (son1 vs. son2).

Qualitative Measures

Post-Trial Questionnaire

The questionnaire used a 5-points scale ranging form strongly disagree (coded 1) to strongly agree (coded 5).

All the participants agreed that the task was easy to complete (4.75), that it was easy to discriminate the sonifications from the background noise (5) and that the sonifications were adequate to represent obstacles (4). They also agreed that they experienced some form of front-back reversal (5) and to a lesser extent, elevation illusions (2.25). The participants reported that is was relatively easy to determine the position of the sound at the beginning of the trial (4.25), to follow the sound trajectory (4.5) and to determine the position of the sound at the end of the trajectory (4.25). They were also relatively confident in their localization responses, along both the azimuth and the elevation components (azimuth: 4; elevation: 4.25). Overall, son1 ("Mosquito": 4.5) was perceived to convey more urgency, a difference that was marginally significant ($z = -1.89$, p=.059). Conversely, son1 was judged more agreeable than son1, a difference again just marginally significant (son1: 3.75, son2: 5, $z = -1.89$, p=.059). Last, son1 was considered easier to localize than son2 (3.6), a subjective perception that contradicts the objective measures of performance

CONCLUSIONS

The present research provides a proof of concept for the usability of spatialized sonification of helicopter drifts in actual cockpit background noise. An ecological egocentric two-dimensional localization task was designed to avoid traditional perceptual biases associated with constrained response along one spatial dimension (typically azimuth) and requiring the user to perform a reference frame transformation. An example would be when the task requires the observer to provide his/her response within an allocentric representation of the environment but the task itself is egocentric.

Establishing a baseline for the auditory localization performance with earcons when visual cues are not available is essential for the development of more sophisticated sonifications such as auditory icons and their integration with other sensory modalities, in particular, visual and tactile.

Overall, the results demonstrated that a single obstacle presented in the frontal hemifield in the horizontal median plane can be localized with a minimum accuracy of 14 degrees, and a maximum of less than 3 degrees. Different factors were shown to influence the performance.

First, as expected, a continuous presentation of the sonification was a significant factor in the improvement of localization precision and accuracy. The main purpose of the so-called memorized condition was to quantify the effect of target eccentricity on localization performance. For targets located in the SMP (0º eccentricity), we observed a degradation in the performance along the main eigenvector, i.e. in the vertical dimension. As the target eccentricity increased, the direction of the error shifted to its horizontal component.

Secondly, and less expected, son2 (sine carrier, no harmonics present) contributed more than son1 (non-band-limited sawtooth carrier, many harmonics present) to performance enhancement. This apparently paradoxical effect must relate, by virtue of their sole difference, to the high frequency content of son1 versus son2. This content illuminates the spectral shaping features of the HRTF, introducing elevation cues not provided by son2. These features can also introduce error when not well-matched to the subject (non-individualized condition). The overshoot in elevation observed in the memory condition is consistent with previous research suggesting that the A and the V "horizons" may not coincide. The overshoot in elevation observed with son1 (but also present with son2 in the Memory condition) may reflect the fact that, in humans, high pitch is consistently mapped to high positions in space (frequency elevation mapping, FEM) in a wide range of cognitive (Ref. 42) and attentional functions (Ref. 43). This effect is particularly evident in the middle range of the spectrum, between 1 and 6 kHz (Ref.44). Potential origins of this mapping include the fact that at higher elevations, more energy is generated in high frequencies (e.g., leaves on the trees rustle in a higher frequency range than the footsteps on the floor), or that the absorption of the ground is frequency dependent in a way that it filters out more of the high-frequency spectrum. These results highlight the possibility of using sound spectral frequency to simulate the vertical elevation of sound sources.

Thirdly, as expected, individualized HRTFs led to a significant performance enhancement, both for precision and accuracy, and with a similar magnitude for all tested conditions (Memory vs. Continuous, son1 vs.son2).

Finally, individualized HRTFs' quality, as expressed here by interindividual differences in the Individualized HRTF condition in the sampling resolution, strongly influenced localization precision, for all conditions tested (Memory vs. Continuous, son1 vs. son2). For accuracy, the superiority of the most "detailed" HRTF (mg.slh) was only significant in the Continuous condition.

FUTURE DIRECTIONS

As noted above, the present research provides a proof of concept for the usability of spatialized sonification of helicopter drifts in actual cockpit background noise. Overall, the results demonstrated that a single obstacle presented in the frontal hemifield in the horizontal median plane can be localized with a minimum accuracy of 14 degrees, and a maximum of less than 3 degrees. Different factors were shown to influence performance and the results provide some guidance regarding the future development of improved helicopter warnings using spatial sound and sonification techniques.

For example, the son1 results potentially emphasize the need for HRTF matching and/or measurement to match the spectral characteristics of the HRTF database to the listener's spectral characteristics. The tendency for horizontal spatial sounds to appear elevated could argue for techniques to counteract this effect, but also to take advantage of it by pre-processing the HRTF database and/or adding an additional filter that specifically emphasizes elevation cues. The perception of front-back reversals also argues for HRTF pre-processing and/or additional real-time processing steps.

Other aspects of the stimulus content could potentially be manipulated to enhance perceptual performance. For example, one could manipulate pulse widths to experiment with different timbres, resulting in more of an auditory icon approach where the timbre of the incurring obstacle is conveyed and is audible in order to identify which obstacle corresponds to the alarm (e.g., a large obstacle would have a deeper timbre). This approach could also employ natural sounds (excerpts or granular synthesis) or physical models. It would be illustrative to compare their benefit relative to other synthesis topologies and the ability of the core synthesis technique to create an overall consistent and complementary auditory display ecology. This would be evaluated in the context of determining the best frequency ranges and content relative to the background noise and simultaneous alerts.

For simulating moving sound sources, one could also explore alternative "looming curves" either based on the existing spherical spread roll-off model as well as extending it or modifying it to account for various ownship/obstacle speeds and distances. For example, the current model used here is referenced to the interaural radius, the closest virtual environment placement, but is not necessarily the best 0 dB reference for this or other scenarios.

Given the increased localization performance observed with son2, one could potentially eliminate pinna-based spectral processing altogether and test frequency-independent ILDs while retaining the other aspects of virtual environment

processing, e.g., listener-relative geometry for ILD and ITD lookup and interpolation, propagation delay lines, and spherical spreading loss. Investigating other distances and speeds would also likely impact various processing steps. For example, to simulate greater distances, the sound propagation delay line would have to be unmanageably long and would introduce long delays between an event and the auditory display. Techniques for getting around this issue would need to be developed.

Author contact: Martine Godfroy-Coooper martine.godfroy-1@nasa.gov, Joel D. Miller joel.d.miller@nasa.gov

ACKNOWLEDGMENTS

This work is a joint effort between the U.S. Army and NASA Ames Research Center. Special thanks to Dr. Edward Bachelder, John Archdeacon, Dr. Durand Begault, and Melinda Zavala.

REFERENCES

[1]Szoboszlay, Z., McKinley, R. A., Braddom, S. R., Harrington, W. W., Burns, H. N., and Savage, J. C. "Landing an H-60 helicopter in brownout conditions using 3D-LZ displays," American Helicopter Society 66th Annual Forum Proceedings, Phoenix, Arizona, May 2010, pp. 1-30.

[2]Harrington, W., Braddom, S. R., Savage, J. C., Szoboszlay, Z. P., McKinley, R. A., and Burns, H. N. "3D-LZ Brownout Landing Solution," American Helicopter Society 66th Annual Forum Proceedings, Phoenix, Arizona, May 2010.

[3]Sanderson P., Anderson J., and Watson M., "Extending ecological interface design to auditory displays," Proceedings of the 10th Australasian Conference on Computer-Human Interaction, Dec. 4th 2000, pp. 259-266.

[4]Kramer, G., Walker, B., Bonebright, T., Cook, P., Flowers, J. H., Miner, N., & Neuhoff, J. (2010). Sonification report: Status of the field and research agenda.

[5]Sanderson, P. M., Watson, M. O., & Russell, W. J. (2005). Advanced patient monitoring displays: tools for continuous informing. *Anesthesia & Analgesia*, *101*(1), 161-168.

[6]Blattner, M. M., Sumikawa, D. A., & Greenberg, R. M. (1989). Earcons and icons: Their structure and common design principles. *Human–Computer Interaction*, *4*(1), 11-44.

[7]Nees, M. A., and Walker, B. N., *Auditory interfaces and sonification*, the universal access handbook, 2009, pp. 507-521.

[8]Bregman, A. S. *Auditory Scene Analysis: The perceptual organization of sound*, MIT Press, Cambridge, MA, 1994.

[9]Vidulich, M. A., Wickens, C. D., Tsang, P. A., & Flach, J. M. (2010). Information processing in aviation. *Human factors in aviation*, 175-215.

[10]Rummukainen, O., & Mendonça, C. (2016). Task-relevant spatialized auditory cues enhance attention orientation and peripheral target detection in natural scenes. *Journal of Eye*

[11]Bach, D. R., Neuhoff, J. G., Perrig, W., & Seifritz, E. (2009). Looming sounds as warning signals: The function of motion cues. *International Journal of Psychophysiology*, *74*(1), 28-33.

[12]Romei, V., Murray, M. M., Cappe, C., & Thut, G. (2009). Preperceptual and stimulus-selective enhancement of low-level human visual cortex excitability by sounds. *Current biology*, *19*(21), 1799-1805.

[13]Neuhoff, J. G. (1998). Perceptual bias for rising tones. *Nature*, *395*(6698), 123-124.

[14]Cappe C., Thelen A., Romei V., Thut G., and Murray MM. "Looming signals reveal synergistic principles of multisensory integration," *The Journal of Neuroscience*. Vol. 32(4), Jan. 25, 2012, pp. 1171-82.

[15]Leo, F., Romei, V., Freeman, E., Ladavas, E. and Driver, J., "Looming sounds enhance orientation sensitivity for visual stimuli on the same side as such sounds," *Experimental Brain Research*, Vol. 213 (2-3), Sept. 2011, pp. 193-201.

[16]Middlebrooks, J. C., and Green, D. M. (1991). Sound localization by human listeners. *Annual review of psychology*, *42*(1), 135-159.

[17]Blauert, J. (1997). *Spatial hearing: the psychophysics of human sound localization*. MIT press.

[18]Hofman, P.M., and Van Opstal, A.J. (2003). Binaural weighting of pinna cues in human sound localization. *Experimental brain research*, 148(4), 458-470.

[19]Carlile, S., Leong, P., and Hyams, S. (1997). The nature and distribution of errors in sound localization by human listeners. *Hearing research*, 114(1), 179-196.

[20]Makous, J. C., and Middlebrooks, J. C. (1990). Two□dimensional sound localization by human listeners. *The journal of the Acoustical Society of America*, 87(5), 2188-2200.

[21]Hofman, P. M., Van Riswick, J. G., & Van Opstal, A. J. (1998). Relearning sound localization with new ears. *Nature neuroscience*, *1*(5), 417-421.

[22]Oldfield, S. R., and Parker, S. P. (1984). Acuity of sound localization: a topography of auditory space. I. Normal hearing conditions. *Perception*, 13(5), 581-600.

[23]Pedersen, J. A. and Jorgensen, T. (2005). Localization Performance of Real and Virtual Sound Sources. Proceedings of the NATO RTO-MP-HFM-123 New Directions for Improving Audio Effectiveness Conference, pp. 29-1 to 29-30. Neuilly-sur-Seine, France, NATO.

[24]Bronkhorst, A. W. (1995). Localization of real and virtual sound sources. *The Journal of the Acoustical Society of America*, *98*(5), 2542-2553.

[25]Perrott, D. R., Ambarsoom, H., & Tucker, J. (1987). Changes in head position as a measure of auditory localization performance: auditory psychomotor coordination under monaural and binaural listening conditions. *The Journal of the Acoustical Society of America*, *82*(5), 1637-1645.

[26]Zotkin, D. N., Duraiswami, R., & Davis, L. S. (2004). Rendering localized spatial audio in a virtual auditory space. *IEEE Transactions on Multimedia*, *6*(4), 553-564.

[27]Väljamäe, A., Larsson, P., Västfjäll, D., & Kleiner, M. (2008). Sound representing self-motion in virtual environments enhances linear vection. *Presence: Teleoperators and Virtual Environments*, *17*(1), 43-56

[28]Wenzel, E. M., Arruda, M., Kistler, D. J., and Wightman, F. L., "Localization using non-individualized head□related transfer functions," *The Journal of the Acoustical Society of America*, Vol. 94(1), Jul. 1st 1993, pp.111-23.

[29]Begault, D. R., and Pittman, M. T., "Three-dimensional audio versus head-down traffic alert and collision avoidance system displays," *The International journal of aviation psychology*, Vol. 6(1), Jan. 1st 1996, pp. 79-93.

[30]Bronkhorst, A. W., Veltman, J. H., and Van Breda, L. "Application of a three-dimensional auditory display in a flight task," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, Vol. 38(1), Mar 1st 1996, pp. 23-33.

[31]Haas, E. C., "Can 3-D auditory warnings enhance helicopter cockpit safety?" Human Factors and Ergonomics Society Annual Meeting Proceedings, Oct. 1st 1998, Vol. 42, (15), pp. 1117-1121). SAGE Publications.

[32]Miller, J. D. and Wenzel, E. M., "Recent Developments in SLAB: A Software-Based System for Interactive Spatial Sound Synthesis," Proceedings of the International Conference on Auditory Display, ICAD 2002, Kyoto, Japan, pp. 403-408, 2002.

[33]http://slab3d.sonisphere.com/, http://humansystems.arc.nasa.gov/SLAB/, http://sourceforge.net/projects/slab3d

[34]Wenzel, E. M. and Miller, J. D. "Sound Lab: A real-time, software-based system for the study of spatial hearing," 108th Audio Engineering Society Proceedings, Paris, France, February 2000, New York: Audio Engineering Society, Preprint 5140.

[35]Miller, J.D., Abel, J.S. and Wenzel, E.M. (1999) Implementation issues in the development of a real-time, Windows-based system to study spatial hearing. Journal of the Acoustical Society of America, 105, 1193.

[36]Begault, D.R., Godfroy, M., Miller, J.D., Roginska, A., Anderson, M.R., and Wenzel, E.M. (2006) Design and Verification of HeadZap, a Semi-Automated HRIR Measurement System. 120th Convention Audio Engineering Society, Paris, France, 20-23 May 2006.

[37]M. J. E. Golay, "Complementary series," IRE Trans. Info. Th., vol. 7, pp. 82-87 (1961).

[38]Pierce, A. (1989) Acoustics. Acoustical Society of America: New York, p. 219.

[39]Miller, J.D., Anderson, M.R., Wenzel, E.M., and McClain, B.U. (2003) Latency Measurement of a Real-Time Virtual Acoustic Environment Rendering System. Proceedings of the International Conference on Auditory Display, ICAD 2003, Boston, MA, pp. 111-114.40[3]E.

[40]M. Wenzel (1999) Effect of increasing system latency on localization of virtual sounds, Proceedings of the Audio Engineering Society 16th International Conference on Spatial Sound Reproduction, Rovaniemi, FIN. New York: Audio Engineering Society, 42-50.

[41]Godfroy-Cooper, M., Sandor, P. M., Miller, J. D., & Welch, R. B. (2015). The interaction of vision and audition in two-dimensional space. *Frontiers in neuroscience*, *9*.

[42]Rusconi, E., Kwan, B., Giordano, B. L., Umilta, C., & Butterworth, B. (2006). Spatial representation of pitch height: the SMARC effect. *Cognition*, *99*(2), 113-129.

[43]Chiou, R., & Rich, A. N. (2012). Cross-modality correspondence between pitch and spatial location modulates attentional orienting. *Perception*, *41*(3), 339-353.

[44]Parise, C. V., Knorre, K., & Ernst, M. O. (2014). Natural auditory scene statistics shapes human spatial hearing. *Proceedings of the National Academy of Sciences*, *111*(16), 6104-6108.

APPENDIX

Post-Study Questionnaire

INSTRUCTIONS: Place an "X" over the circle in the box that best describes your opinion. If you do not understand a question, please ask for clarification. At the end, a space is provided for you to make additional comments pertaining to individual questions or for your general observations.

		Strongly disagree	Somewhat disagree	Neither disagree nor agree	Somewhat agree	Strongly agree	Not applicable/ Did not use
1.	Overall, the task was too difficult.	○	○	○	○	○	○
2.	Overall, the experiment was too long.	○	○	○	○	○	○
3.	The audio stream was set to a comfortable listening level.	○	○	○	○	○	○
4.	The comfort of the headset was good.	○	○	○	○	○	○
5.	The comfort of the eye-tracker was good.	○	○	○	○	○	○
6.	It was easy to isolate the sonifications from the noise	○	○	○	○	○	○
7.	The starting location of the sonifications was easy to determine	○	○	○	○	○	○
8.	It was easy to follow the sonifications' trajectories	○	○	○	○	○	○
9.	I was confident in my reported localization of the obstacle azimuth	○	○	○	○	○	○
10.	I was confident in my reported localization of the obstacle elevation	○	○	○	○	○	○
11.	The final location of the sonifications was easy to determine	○	○	○	○	○	○
12.	I experienced some front-back reversals	○	○	○	○	○	○
13.	I sometimes perceived the sound above my head	○	○	○	○	○	○
14.	The "Mosquito" (high-pitch) sonification conveyed a sense of urgency	○	○	○	○	○	○
15.	The "UFO" (low-pitch) sonification conveyed a sense of urgency	○	○	○	○	○	○
16.	The sonifications were adapted to the representation of obstacles	○	○	○	○	○	○
17.	The "Mosquito" (high-pitch) sonification was agreeable to listen to	○	○	○	○	○	○
18.	The "UFO" (low-pitch) sonification was agreeable to listen to	○	○	○	○	○	○
19.	The "Mosquito" (high-pitch) sonification was easier to localize than the "UFO" (low-pitch) sonification	○	○	○	○	○	○

Comments/suggestions on the current concept of spatial sonification for obstacle warning/avoidance and possible future implementations: