In: Johan Stahre & Lena Mårtensson (Ed.): **Automated Systems Based on Human Skill** : Symposium - Joint Design of Technology and Organization 2003. Elsevier, 2004. ISBN 0-08-044291-9 (IPV - IFAC Proceedings Volume, ISSN 1474-6670). pg. 129-134

EVALUATION OF AUDITORY DISPLAYS SUPPORTING AIRCRAFT APPROACH AND LANDING

Bernd-Burkhard Borys

Laboratory for Human-Machine Systems
University of Kassel, Kassel, Germany

Abstract: The paper describes an auditory interface using directional sound as a possible support for pilots during approach in an instrument landing scenario. Several ways of producing directional sounds are illustrated. One using speaker pairs and controlling power distribution between speakers is evaluated experimentally. Results show, that power alone is insufficient for positioning single isolated sound events, although discrimination in the horizontal plane performs better than in the vertical. Additional sound parameters to compensate for this are proposed.

Keywords: Auditory Displays; Human-Machine Interface; Human-Centred Design; Transportation; Aircraft Control

1. INTRODUCTION

Developments in process control and instrumentation resulted in workplaces that are clean, calm, and in safe distance from the process. Information is presented mainly visually and other modalities – sound, vibration, temperature, smell – are lost although these sensations give valuable support to an experienced operator. The upcoming of multimedia add-ons for the standard PC provided us with affordable equipment to make additional non-visual information available. In the case of two-dimensional tracking tasks like guiding a vehicle through airspace, this may be easier accessible and understandable than numerical and textual information and even visual information.

Since five years, we evaluate the use of auditory displays to convey information in human-machine interaction. We performed already several experiments and demonstrated the usability of the equipment. We currently evaluate the use of directional sound to support aircraft guidance in different scenarios.

In general, a human-machine system comprises of three parts, the human, the technical system or machine, and the interface. For the human, we can distinguish different roles the human may take over: Control, supervision, planning, diagnosis, learning, and more. The technical system can be, for example, a vehicle (air, land, sea, space), a plant or its control room, a computer, a robot. For the interface, we can deal with aspects of presentation and interaction, dialogue design and adaptability, or support aspects. Presentation and interaction can be visual, auditory, gestural, or multi-modal. Within these dimensions, this paper deals with the following specific aspect of human-machine systems:

- The *technical system* is a simulated aircraft.
- The *human* is working in a manual control task, guiding the approach to an airport for landing.
- In the *human-machine interface*, the research concentrates on presentation of deviations from the approach path.
- The *presentation* is given visual and auditory, with emphasis on directional auditory stimuli to signal amount and direction of deviation.

A series of experiments is currently in progress to determine favourable solutions for several aspects of this type of presentation. One is the necessary physical set-up – arrangement of speakers, individual sound pressure levels to generate directional cues. Others are the type of auditory stimuli to use and which parameters may carry the information. A large number of options exist and several decisions have been made based on the available equipment and on restrictions of the task environment, the aircraft cockpit. Other decisions how to reduce the number of options and how to get a feasible solutions are based on the outcome of the experiments described later.

2. AUDITORY DISPLAYS

Imagine, while driving your car you hear the following rhythm from your radio:

Today, more urgent information from the radio station about congestion or adverse road conditions would interrupt the music from your radio, while attention getting sounds from your car's systems just overlay the music. Now imagine a workplace in a vehicle where radio only means communication and not entertainment and you are busy to operate controls and to observe the environment (visually) and at the same time watch system state presented by visual indicators. Visual perception dominates and when focusing on essential cues from the environment necessary to keep the vehicle on the correct path these may block out peripheral, but important information from vehicle instruments. Hartz (1997) designed instruments specially for peripheral vision. These, however, still use the visual sensory channel while the auditory channel is underused except for attention getting warning signals.

Imagine the same sound telling you about your system's state with small, but noticeable changes in some parameters when system parameters change. Sound parameters could be musical, physical, or perceptional. Musical parameters are the instruments used, timbre, chord, the rhythm, its speed and timing, while physical properties comprise the overall power and power relations as well as spectral properties. Perceptional parameters include the perceived position of a sound in azimuth, elevation, and distance. Specially these directional and spatial parameters have an intrinsic meaning and one of our hypotheses was that a human operator would react faster and more correct on a directional auditory cue than on a visual indicator.

Other sounds than music may provide the auditory stimulus ("*… even elephant cries and women's screams have not escaped study*." Fidell 1978) and carry state and directional information. We know, however, from previous experiments (Borys 2001) that subjects preferred soft and complex sounds (like the sound of turning a sheet of paper or the clicking of a gear-shifting mechanical apparatus) over simple tone burst, beeps, chirps, and pings (like filtered noise or exponentially shaped sine waves). A field worth studying to find complex sounds is nature with wind and waves, water, trees, and birds. Conversy (1998) gives an example how to generate realistic wind and wave sounds with two parameters (*strength* and *shape*) to carry the information. Sonnenschein (2001) provides technical information for recording outdoor natural sound and how to integrate it into a complex (cinema) scene to give a consistent meaning. In an industrial scene, artificial soundscapes driven by plant parameters could return the information that is missing in modern plants. Here the need to provide safe, clean, and silent workplaces consequently reduces the exposition to unhealthy noise and in the same time removes important indications for changes in plant state. We will, of cause, keep workplaces safe; however, an operator should be able to experience his work with all senses.

The Laboratory for Human-Machine Systems of the University of Kassel started the *Multi-Media Process Control Room* project in 1997. Our equipment comprises of several PCs, audio-, video-, and musical hardware and software from the fields of virtual reality, video, graphics, sound, and music processing and analysis. We described the beginning of this project in Borys and Johannsen (1997).

3. DIRECTIONAL SOUND FOR AIRCRAFT GUIDANCE

During an instrument (ILS) approach of a modern, glass-cockpit aircraft, indicators integrated into the primary flight display show vertical and lateral deviations from the glide path. According the technical equipment of aircraft and airport and the operational state of the equipment, different categories of instrument approaches apply. In the highest categories (III A, B, C), approach guidance until touch down is possible without visual reference to the runway. In lower categories (I, II), the pilot needs to see the runway below a specified height above ground and to decide, whether a safe landing is possible or not (decision height – DH). Thus, for the last approx. 300 feet of descent the pilot must share attention between the instrument and the view outside the front window. The ILS indicators are not useful for peripheral vision and the solution proposed by Hartz (1997, see above) does not support continuous information in the visual periphery. Therefore, auditory support continuously indicating deviation when the pilot is working head-up with the visual focus outside the cockpit would be useful. Category-I- and -II-approaches would become more important when augmentation systems for the Global Positioning System (like EGNOS in Europe, under test since spring this year) are operational and GPS approach guidance becomes possible.

Directional information either may be given intrinsically by spatialized sound or the information can be coded into the sound. A *spatialized* sound is generated or processed such that it apparently comes from a source in a specific direction. The simplest *coding* would be spoken commands (*up*, *down*, *left*, *right*) to express a direction. A previous experiment, however, showed that subjects could not control a tracking target effectively based on verbal information in a two-dimensional tracking experiment (Borys, 2003; Linge, 2001; Jedamski, 2001). We assume that perception and discrimination of the spoken commands introduce too much delay into the tracking loop. In particular, the German words used in the experiment (*hoch*, *runter*, *rechts*, *links*) take more time to speak than the English translation. However, in aviation, the ground controlled approach (GCA) procedure exists and is successfully used, in which the radar controller directs the aircraft to landing using voice commands. Thus, this method may work with slower systems and may be considered for future experiments.

In other experiments, the intended movements of a robot have been coded into a melody, three notes for each motion direction (Johannsen 2002; Johannsen 2001). Rhythm was used for one axis (left, right), pitch for the other (up, down on the experimental display). In the diagonal axes, pitch and rhythm were combined. Results show, that subjects are able to learn this code. Assigning meaning to the pitch axis is straightforward, raising pitch means, e.g., *up*. The meaning of the rhythm could well be remembered translating it to Morse code (letters D and U). However, those subjects with a musical education or practice in singing performed better.

Many other ways of coding directional information can be found. The user, however, must learn the code and decode the signal. Directional hearing is already learned, continuously trained in daily life and the decoding process hardwired into the brain.

Different technical means can provide directional auditory stimuli; among these are spatializing filtering techniques, one speaker at any desired position, or two or more speakers at suitable positions and methods manipulating sound power and/or phase relations across these speakers. The techniques show different capabilities in generating front vs. back, distance and elevation cues and in the need to measure or fix head position.

Human discrimination of sound direction is assumed to be based on phase and level differences between the ears and on filtering effects of head and pinnae. All these effects together can be measured, modelled as head-related transfer function (HRTF), and implemented in signal processing hardware. Crystal River's *Acoustetron* provides this in real-time with signal-processing hardware in a standard PC. The problem with this solution is that headphones are necessary and that azimuth and elevation need to be captured with respect to the head and – for a fixed direction of the stimulus – must change with rotational head movements. Thus, additional equipment is necessary to capture head movements without interfering with the human's task. A promising solution uses cameras to determine head position and orientation (Smart Eye AB, 2003).

The position of the pilot's head, however, is fairly constant in the aircraft cockpit and the cockpit gives sufficient opportunities to place loudspeakers around the head. Thus, we use a speaker array positioned around the listener's head in the experiments described in this paper. Wenzel et al. (1991) compared the performance of HRTF with speaker arrays, using filtered noise as cues. They found comparable results in both versions; however, subjects performed better in the horizontal axis than in the vertical. Like others, Wenzel also reported that subject mixed up front and back stimuli.

In previous experiments, we evaluated directional sound for aircraft guidance in two different scenarios. The first was the identification of the intruder's direction following a Traffic Advisory of the Traffic Alert and Collision Avoidance System (TCAS TA). The *Traffic, Traffic!* announcement was given from one of eight speakers placed around the subject's head (Gudehus, 2002). The second was an instrument approach in minimum weather while the deviation information from the Instrument Landing System (ILS) triggered sound signals in one, two, or four speakers in front of the subject. Direction of deviation was articulated as direction of the sound by routing the signal to the appropriate speakers. This experiment used sine waves of different frequency, duration, and repetition time to express amount of deviation in three levels (Quittkat, 2003). Results from both experiments showed that subjects could perform their task using the auditive information alone. However, in terms of reaction time and correctness in the TCAS experiment and in terms of deviation from the glide path for the approach we found no significant advantage over the visual displays.

The latest experiments, described in this paper, use the approach scenario with a more sophisticated method to generate directional cues and to signal amount of deviation. When two speakers are placed to the left and right in front of a listener and play the same signal with equal phase and power, the listener perceives a virtual sound source in the midpoint between the speakers. The virtual source moves horizontally out of the centre when phase or power is different. Phase manipulation (A/B-Stereophony) is, again, very delicate with respect the head movements. Manipulation of power (X/Y-Stereophony) is more robust and is in modern entertainment systems extended to even more speakers (e.g., five in Dolby Digital, seven in Sony's SDDS, both with an additional low-frequency speaker). However, these entertainment systems all arrange speakers in a horizontal plane. Thus, one open question is whether it will be possible to shift the virtual sound source vertically using a vertical speaker arrangement. The experiments described below should give answers to this and the following questions:

- What are the stimuli that enable perception of a direction?
- How many speakers are necessary for both azimuth and elevation cues?
- What is the best arrangement of speakers around the listener?
- What power distribution across speakers provides the desired effect?

4. THE PHYSICAL ARRANGEMENT

The current experiment – as well as two of the experiments mentioned above – is centred around a flight simulator comprising of a cockpit mock-up (see Fig. 1) and a PC with joystick and throttle. The PC runs Microsoft's Flight Simulator 2002 with an add-on scenery of Frankfurt Rhein-Main Airport (FRA / EDDF).

A digital flat panel display shows the instruments in the cockpit, a second video output drives a video beamer throwing the outside vision against the wall in front of the cockpit. Using Microsoft's multiplayer feature, a tracking programme on a second PC receives position, speed, orientation, and aircraft state information from the simulator. The tracking programme calculates deviations from the flight path and controls the sound equipment, using the MIDI (Musical Instrument Digital Interface) protocol.The sound equipment consists of a Roland JV2080 synthesizer, a Yamaha 01v digital mixer with 8-channel ADAT optical output, and a Korg SoundLink 880DA de-multiplexer, that converts the digital ADAT signal to eight analogue lines and drives up to eight speakers positioned in and around the cockpit.

Fig. 1. Flight Simulator

The tracking programme sends MIDI commands to the synthesizer to trigger the sound and to the mixer to distribute the sound to one or more speakers. Speakers are placed left and right from the cockpit (L, R), above the subjects head (T) and under the chair (B), as shown on Fig. 2.

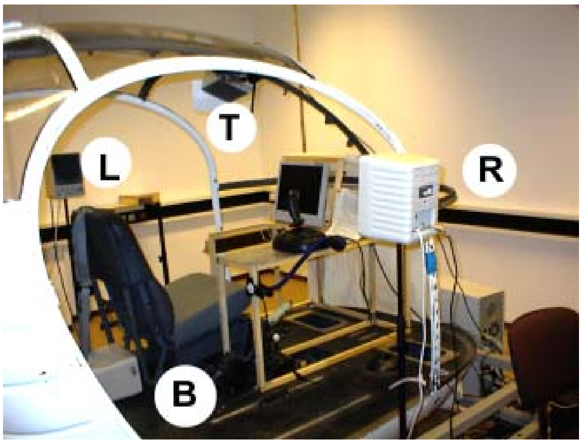


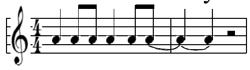
Fig. 2. Speaker Arrangement, Four Speakers: Left-Right-Top-Bottom

MIDI commands to the mixer use 128 steps to adjust the gain in each channel. The useful range of gain reaches from +6 to -30 decibels (dB); a non-linear relation allows adjustments on tenth of decibels around 0 dB. Before starting experiments, the absolute power level of each speaker is adjusted using pink noise to the same value of 60 dB(A), measured with a Brüel&Kjær 2238 sound level meter at the position of the subject's head. This should also compensate the different distances from speakers to the listener.

We can generate sound controlling several parameters. First, it can be – when restricting to the General MIDI instruments (GM) – one of the 128 chromatic instruments (piano, guitar, trumpet, …) or of the 64 percussion rhythm sets (e.g., tambourine, triangle, …) in the MIDI 'drums' Channel no. 10. Next, we can select pitch and volume (note and velocity in terms of MIDI) in 128 steps. Duration can be controlled from about 1/700 of a quarter note to infinite. Several additional MIDI features like pitch bend, fine-tuning, tremolo will not be used in this stage of the experiment.

Not all combinations produce useful results. In theory, A/B-Stereophony uses low, X/Y-Stereophony high frequency components, compared to the frequency of about 1500 Hz, when wavelength is comparable to head dimensions. High pitch percussion instruments were expected to be good candidates to produce suitable stimuli.

5. SELECTING THE STIMULI

Signals produced by the synthesizer are optimised for good musical sound, not for direction detection. To find those sounds optimal for direction detection, a series of simple trials was performed. A chromatic or percussion instrument was randomly chosen, a short sequence played  and the signal randomly routed to one or two speakers and the subject was asked to indicate the perceived position. A position close to the speaker used or close to the mid-point between two speakers used was counted as correct. Instruments that showed significantly more false than correct identifications were successively omitted from the trials, leaving a smaller set to choose from for consecutive trials.

The first instruments omitted were from the General MIDI sets 'Sound Effects', 'Synthesizer Effects' (rain drops, birds, helicopters), the sets of eight different pianos and of eight different organs, as well as the soft instruments like oboe, saxophone, and clarinet. The chromatic instruments remaining after several series of trials are from the 'chromatic percussion' set, especially xylophone, marimba and tubular bells. In general, percussion instruments performed better but not all remained in the selection set. When we stopped the selection process, the Xylophone from the chromatic percussion set triggered with the note-value of 57 (an >a< in the third octave) and 11 instruments in the 'drums' channel remained in the set of useful stimuli. The percussion instruments were General MIDI nos. 38 (Ballad Snare), 42 (Closed High Hat 1), 44 (Pedal High Hat), 48 (High-Mid Tom), 56 (Cowbell), 58 (Vibraslap), 64 (Low Conga), 90 (Old Kick), 94 (808 Snare), 96 (Brush Swiss), and 97 (Brush Roll).

6. DETERMINING POWER DISTRIBUTION

The next experiment tried to find the correct distribution of power between two speakers in each axis. Using the rhythm and the twelve remaining instruments described in the previous section, stimuli were routed to two speakers (either left and right or top and bottom speaker). Subjects indicated with a mouse click in a rectangle on the screen from which direction in the forward hemisphere (left to right and top to bottom) they heard the sound. An algorithm with a self-adapting and a random component described below determined the power of each speaker. In several trials before the final experiment, we recorded power setting of the two speakers used and position of the decision on the related axis. Approximations with second-order polynomials with a constraint (power not above 0 dB) were made to get a relation between desired position and necessary speaker power. Consecutive trials used the previous approximations plus random variations to set the power, collected new power/position pairs and calculated new approximations. The final approximations shown in Fig. 3 and 4 are used for the final experiment. Although the absolute power of each single speaker was adjusted for the same sound pressure level at the listener's position on the 0 dB relative setting, the approximations are not symmetric: The right speaker (farther away at the opposite side of the cockpit) and the bottom speaker (hidden below the seat) need more power. We did not compensate phase relations, thus, wave fronts from left and top speaker reach the listener first. Excessive power seems to be necessary to compensate for this.

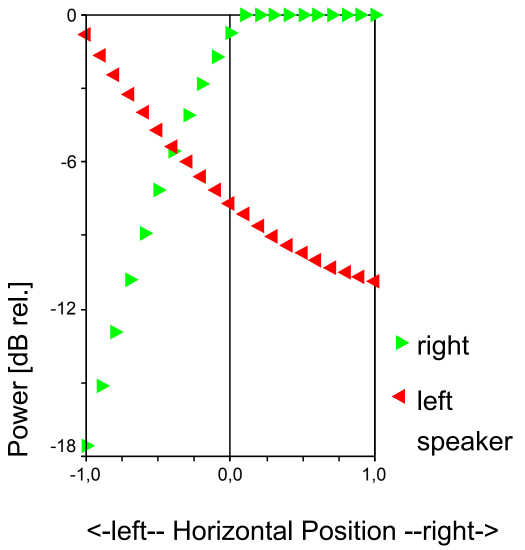


Fig. 3. Desired Position and Power Setting, Horizontal Axis

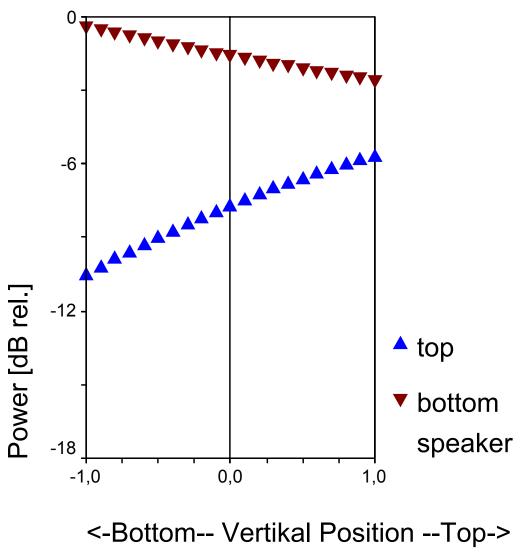


Fig. 4. Desired Position and Power Setting, Vertikal Axis

12 students with interest in auditory displays participated in the final experiment. The experiment was self-paced, shortly after the click a new stimulus was presented with instrument and axis randomly chosen, desired position equally distributed on the axis, and power setting calculated with the approximations described above and a random component. Due to the random component, the experiment did not test perceived vs. desired position but perceived position vs. power relations. Within a 20 minutes trial, the subjects judged from 250 to 440 sounds each, this (and an outlier with 660) totalling in 4164 responses.

To estimate to what extend power-driven stereophony may be used to force a direction experience we plot perceived position against power difference (difference in dB is ratio of absolute sound pressure). Fig. 5 shows 1500 pairs for stimuli on the horizontal axis, Fig. 6 2100 pairs for the vertical axis.

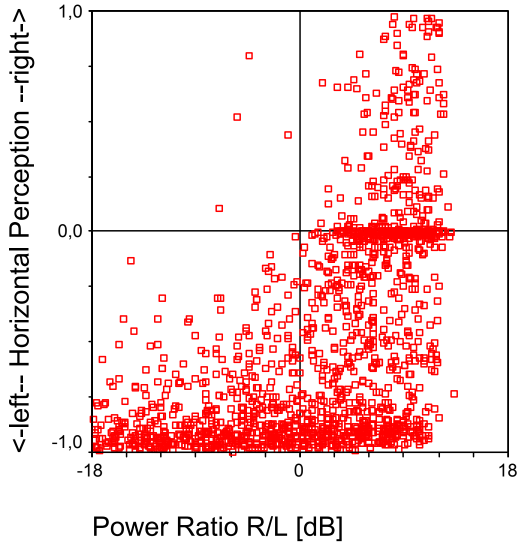


Fig. 5. Perceived Position and Power Difference, Horizontal Axis

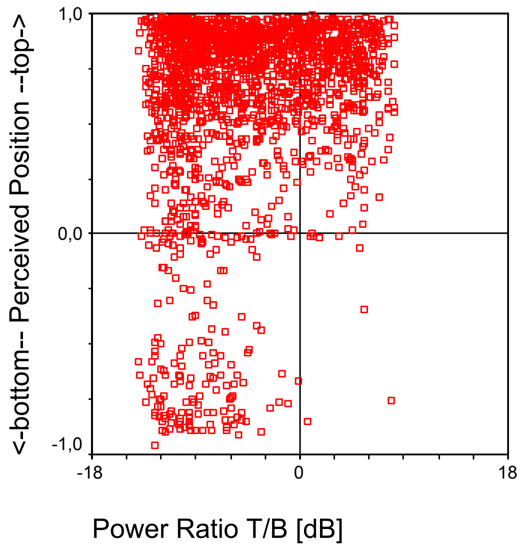


Fig. 6. Perceived Position and Power Difference, Vertical Axis

For the vertical axis, Fig. 5 shows the dominance of the left speaker. When the right speaker's power is higher than the left, sound is perceived more often in the centre. For the vertical axis, Fig. 6 shows that subjects always hear sounds above the vertical and close to the top speaker. It was not possibly to consistently shift a virtual sound source in the vertical axis by power changes. (SPSS Vers. 11 was used to produce statistics and figures.)

In contrast to the results for single independent and random stimuli, a continuous sound shifted slowly across the horizontal axis implemented for demonstration purpose gives a useful impression of a moving virtual sound source that can be tracked and deviation compensated. However, this does not hold for the vertical axis.

7. CONCLUSION

We found, that X/Y stereophony gives a good impression of the position of a continuous virtual sound source moving in the horizontal axis. For single, unrelated sound events, as used in the experiments, we found large random errors and a dominance of the speaker closer to the listener. The latter may be explained by phase relations or just by the knowledge of the subjects about the speaker arrangement. In the vertical axis, we found also a dominance of the closer speaker; we were, however, not able to generate a virtual source that moves and can be tracked vertically and vertical position estimation showed even more errors. Percussion instruments, especially those with high pitch, perform better in signaling a direction.

Therefore, for coming experiments we will design a percussion rhythm that accompanies initial and final approach. Directional sound will signal glide path intercept and direction of lateral deviation from glide path. An additional parameter, pitch of a chromatic instrument being a good candidate, needs to be present to indicate vertical deviation. Amount of deviation will change characteristics of the rhythm and additional accents will show events like gear and flap movement or reaching decision height.

REFERENCES

Borys, B.-B. (2001): Hardware, Software, and Experiments for Auditory Displays. In: *Human Decision Making and Manual Control* (Proceedings of the 20th European Annual Conference on..., Lind, M. (Ed.)) Kgs. Lyngby, Denmark: ØRSTED*DTU, Automation. ISBN 87-87950-89-8.

Borys, B.-B. (2003): 15 Years Research and Experiments in Aeronautics in the Laboratory for Human-Machine Systems. In: *From Muscles to Music : A Festschrift to Celebrate the 60th Birthday of Gunnar Johannsen.* (Borys, B.-B. and Wittenberg, C. (Eds.)) Kassel: Kassel Univ. Press, 2003. ISBN 3-933146-87-9, pp 41-51.

Borys, B.-B.; Johannsen, G. (1997): An Experimental Multimedia Process Control Room. In: Advances in Multimedia and Simulation (*Proceedings of the 1997 Annual Conference of the Human Factors and Ergonomics Society Europe Chapter Annual Conference*) Bochum: Human Factors and Ergonomics Society Europe Chapter. ISBN 3-00-002435-2.

Conversy, S. (1998): Wind and Wave Auditory Icons for Monitoring Continuous Processes. Student Paper, *CHI'98*. ACM; ISBN 1-58113-028-7, pp. 351-352.

Fidell, S. (1978): Effectiveness of Audible Warning Signals for Emergency Vehicles. *Human Factors*, **20** (l), pp. 19-26.

Gudehus, T. C. (2002): *Drei-dimensionale akustische Anzeigen zur Signalisierung von Kollisionswarnungen im Flugzeugcockpit* <Three-Dimensional Acoustic Displays to Signal Collision Warnings in the Aircraft Cockpit, in German>. Diplomarbeit II, Fachgebiet Systemtechnik und Mensch-Maschine-Systeme, Universität Kassel (Master's Thesis).

Hartz, J.-O. (1997): *Darstellung von Flugparametern in der visuellen Peripherie am Beispiel der Entwicklung von Hubschraubercockpits* <Display of Flight Parameters in the Peripheral Visual Field, in German>. PhD Dissertation; Universität Kassel.

Jedamski, J. (2001) *Auditive Anzeige für eine Tracking-Aufgabe – Versuchsplanung, Durchführung und Interpretation*. <Auditory Display for a Tracking Task - Experimental Design, Experiments, and Interpretation, in German>. Studienarbeit, Fachgebiet Systemtechnik und Mensch-Maschine-Systeme, Universität Kassel (Student's Thesis).

Johannsen, G. (2001) Auditory displays in human-machine interfaces of mobile robots for non-speech communication with humans. *Journal of Intelligent & Robotic Systems* **32** (2), pp. 161-169 Oct 2001. ISSN: 0921-0296.

Johannsen, G. (2002) Auditory Display of Directions and States for Mobile Systems. *Proceedings of the 2002 International Conference on Auditory Display*, Kyoto, Japan, July 2-5, 2002.

Linge, S. (2001) *Auditive Anzeige für eine Tracking-Aufgabe – Programmierung des Experiments*. <Auditory Display for a Tracking Task - Programming the Experiment, in German>. Studienarbeit, Fachgebiet Systemtechnik und Mensch-Maschine-Systeme, Universität Kassel (Student's Thesis).

Quittkat, M. (2003) *Experimentelle Untersuchung auditiver Anzeigen im Landeanflug* <Experimental Evaluation of Auditory Displays during Approach for Landing, in German>. Diplomarbeit I, Fachgebiet Systemtechnik und Mensch-Maschine-Systeme, Universität Kassel (Bachelor's Thesis).

Smart Eye AB (2003): *Technology*. Internet page at www.smarteye.se

Sonnenschein, D. (2001): *Sound design : the expressive power of music, voice, and sound effects in cinema.* Studio City, Calif. : Wiese. ISBN 0-941188-26-4.

Wenzel, E. M.; Wightman, F. L.; Kistler, D. J. (1991). Localization with non-individualized virtual acoustic display cues. *Human Factors in Computing Systems Conference Proceedings on Reaching Through Technology*, New Orleans, Louisiana, United States; Pages: 351 – 359; ISBN: 0-89791-383-3