Effects of Sound Type on Recreating the Trajectory of a Moving Source

Graham Wilson
Stephen Brewster
Glasgow Interactive Systems Group
School of Computing Science
University of Glasgow
Glasgow G12 8QQ, UK
{first.last}@glasgow.ac.uk

Hector Caltenco
Charlotte Magnusson
Department of Design Sciences
Lund University
PO Box 118
221 00 Lund
{first.last}@certec.lth.se

Sara Finocchietti
Gabriel Baud-Bovy
Monica Gori
Robotics Brain & Cognitive Science
Instituto Italiano di Tecnologia
16163 Genova
{first.last}@iit.it

CHI'15 Extended Abstracts, Apr 18-23, 2015, Seoul, Republic of Korea
ACM 978-1-4503-3146-3/15/04.
http://dx.doi.org/10.1145/2702613.2732821

Abstract
The ABBI (Audio Bracelet for Blind Interaction) device is designed for visually impaired and blind children to wear on the wrist and produce sound based on the movement of the arm through space. The primary function is to inform a child (or adult) about his/her own movements to aid spatial cognition rehabilitation. However, the device could also be worn by friends and family and be used to inform the visually impaired person of others' movement in the environment. In this paper, we describe an initial experiment that measured how well blindfolded sighted individuals could track a moving sound source in 2D horizontal space and then walk the same route to the same end position. Six sounds, including natural sounds, abstract sounds, Earcons and speech, were compared to identify which type of sound produced more accurate route recreation.

Author Keywords
Sound perception; visual impairment; spatial cognition

ACM Classification Keywords
K.4.2. [Social Issues]: Assistive technologies for persons with disabilities

Introduction
The ABBI (Audio Bracelet for Blind Interaction) project is developing technologies and procedures to rehabili-

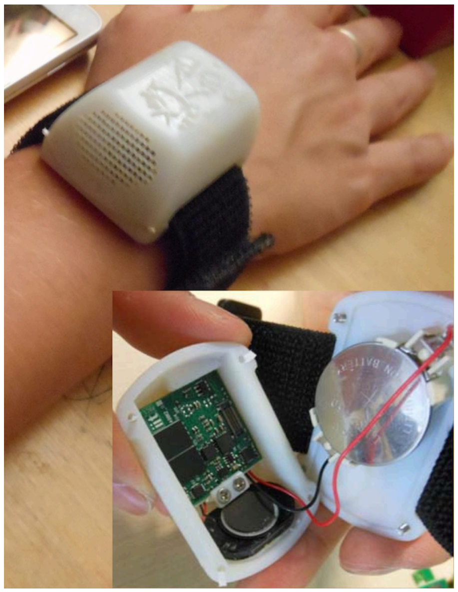

Figure 1: The ABBI device shown on the wrist, with internal components, including circular battery, microcontroller and speaker (which outputs through grill).

tate spatial cognitive brain processes in visually impaired people through natural audio-motor associations. While the sighted rely on vision to learn and calibrate motor skills and spatial cognition [2], ABBI aims to replace vision with audio feedback. The ABBI device (Figure 1) detects movement (acceleration and tilt) and produces sound based on the nature of that movement. ABBIs will be placed on the wrists or ankles of children (and adults) with visual impairments so, rather than seeing their arms or legs move through space, they can hear the movement instead, through inherently spatialised sound. Combined with kinaesthetic and proprioceptive information, these sound sources may support the visually impaired child or adult to build a better representation of his/her movement in space, and improve or expedite motor learning, by associating movements with feedback that conveys spatial information in a natural and direct way.

In addition, and what is the focus of this study, sound sources could also be placed on other people and objects to provide a better sense of the events taking place in the environment (extra-personal space), and to improve the mobility and social skills of blind children and adults. Relevant examples could include 1) parents of visually impaired children, or 2) children of visually impaired adults wearing sound-emitting ABBIs to indicate to the other where they are moving (Figure 2). The successful design of ABBI partly depends on identifying suitable sounds to provide the necessary information reliably. The sounds should not only be statically localisable, but the motion of the sounds must be tracked and laid out through a representation of space, to allow the blind person to understand where the sound source now lies.

As a means of determining what types of sound might be suitable, this paper describes an experiment that tested how well blindfolded sighted individuals could track the movement of different sounds in horizontal space through a room before walking the same route. Although participants were sighted, the results form an initial baseline and performance indicator to inform ABBI's auditory design to take forward into trials arranged with visually impaired people. This initial study compared 6 sounds: 2 natural, 2 abstract, 1 musical Earcon [3] and 1 speech sound. Our future research will test more sounds and include other tests, including walking to a static sound and recreating 2D shape trajectories.

Related Research

The development of a child's spatial cognition and spatial understanding of its own actions develops based on both visual information and its related motor signals [2]. Beyond the development of motor skills, vision also impacts the child's spatial awareness or understanding of extra-personal space. Because of this, congenitally and early-blind individuals develop motor behaviours much more slowly [2] and tend to understand the world in relation to an egocentric, rather than exocentric, reference frame [7]. Blind children are less likely to engage with nearby objects and peers [2] and orient in space [7], and particularly self-initiate action.

The ABBI device primarily aims to improve spatial cognition and motor behaviour in children by substituting the visual information with naturally spatialised audio feedback during rehabilitation. The child can hear their arm or leg moving through space, instead of seeing it moving. Parents and rehabilitators can also wear ABBI to assist and encourage movements in the child, such as mimicking arm movements and reaching for moving

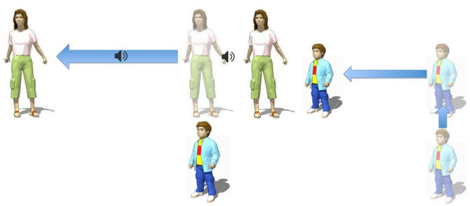

Figure 2: A blind child could hear a parent or friend wearing ABBI moving through space and make their way to them.

Figure 3: Workshop with blind and visually impaired children to identify suitable sounds for the ABBI device.

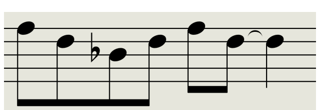

Figure 4: Melody used in the musical Earcon.

arms. When others wear ABBI, the device could aid in improving the child's perception and understanding of extra-personal space. If the child hears a parent or friend moving through the room this could encourage the child to move to them. This may also support a more independent analysis of activity in space, rather than other people explicitly announcing their movement/location. To be able to facilitate this, we needed to know how well the movement of sounds is perceived and recreated.

Initial Exploratory Workshop

Because ABBI is envisioned in use throughout childhood, it is important to take personalisation into account in the audio design. If the child does not like the way ABBI sounds they are much less likely to use it. The sounds that are, empirically, most easily localisable in space may not be enjoyable enough to be used for long periods. Therefore, we ran a workshop to get an understanding of what kinds of sounds blind and visually impaired children find enjoyable and acceptable. The workshop was run with 9 blind and 8 visually impaired children visiting the Chiossone Institute in Genoa (Figure 3). 32 natural sounds and 39 synthetic sounds were played to both groups and they were asked to indicate whether they "liked" or "disliked" each sound. Overall, 34 sounds were more liked than disliked, with the most liked sounds including crashing waves, bubbling water, birdsong and a synthetic, rhythmic dropping sound.

Auditory Tracking Experiment

The study in this paper tested whether personalisation of ABBI sounds and support for rehabilitation (through accurate perception of sound motion) can be reconciled: do different sound types lead to different accuracy in judging the movement of a sound? Other research in our project is testing the localisation of sounds and the recreation of arm movements, while here we look at how well the movement of sounds in 2D horizontal space can be perceived and their trajectories recreated in the absence of visual input.

Type	*Natural*		*Abstract*		*Musical*	*Voice*
Sound	Birds	Waves	Pulse	Dropping	Earcon	Speech

Table 1: Sounds used in the experiment[1]

Sounds

For this initial test we took four of the most liked sounds from the workshop and added two more, based on different strands of HCI and perceptual science[1]. The sounds were classified into four categories: *natural, abstract, musical* and *speech* and can be seen in Table 1. The natural sounds were liked in the workshop: *birds singing* and *waves* (light crashing). Birdsong can be difficult to localise so, while it was enjoyed during the workshop, it may be difficult to follow the trajectory accurately [1]. The abstract *pulse* and synthetic rhythmic *dropping* sounds (pitch C3, 130.81Hz) were also liked in the workshop. *Pulse* was a fuzzy, 2-sec C3 note with quick attack and slow decay.

For the musical sound we used an *Earcon* [3,6], structured sounds that can be used to convey multidimensional information. We used a grand piano playing the melody shown in Figure 4 (notes between A#3 and F4), one of those tested in previous research [3]. Speech is used as a means of testing hearing impairment [5] and is a primary output modality for assistive technology. We used a sample of recorded male speech taken from free audio book recordings. The voice spoke the first

[1] Sounds available at www.dcs.gla.ac.uk/~gawilson/ABBI.html

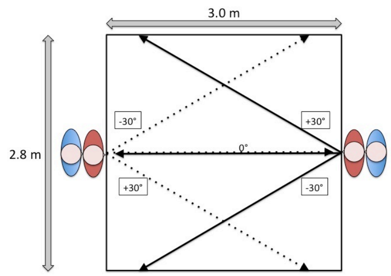

Figure 5: Layout of experimental space, including trajectories (lines), experimenter (red) and participant (blue) positions. Trials alternated between each side.

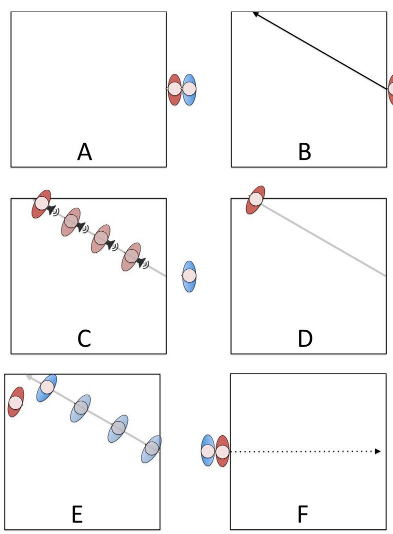

Figure 6: Trial procedure: A – Experimenter and participant at start position; B – Trajectory is chosen; C – Experimenter walks trajectory while sound plays; D – Sound stops; E – Participant walks trajectory; F – Experimenter guides Participant to next start position.

sentence of "Alice in Wonderland" by Lewis Carroll. All sounds were set to an equal level of subjective volume.

All sounds were started 5 seconds before the experimenter began walking the required route, to avoid perceptual bias where the onset of an immediately moving sound is mis-localized in the direction of movement [4]. The sounds were looped continuously during the route before being stopped two seconds after finishing the route. The sounds were stored on, and played from, an Android mobile phone with the sound being emitted from a Bose SoundLink Mini Bluetooth speaker held at head height by the experimenter.

Experimental Space and Trajectories
The experiment took place in a lab housing a 12-camera Vicon motion tracking system. The experimental space (illustrated in Figure 5) was a 2.8m x 3m rectangle in the centre of the room. The three walking trajectories used for testing (also in Figure 5) started from the middle of either side of the space (trials alternated sides) and ran at -30° (30° to the left of centre), 0° (centre) and 30° (to the right of centre). All trajectories were 3 meters in length and were in straight lines, walked at a rehearsed speed of 0.6m/sec. Each was performed five times in a random order. Participants were not told the trajectory shape (straight line), simply that the sound will follow "a route through the room" and they were to recreate that route.

Participants and Experimental Design
6 male participants aged between 27 and 31 (mean 29.7) took part. All were sighted and each was paid £10. The within-subjects study was split into 6 conditions, based on the sound being tested. Due to the number of sounds it wasn't possible to fully counterbalance the condition order and so they were completed in a random order. The Independent Variables were *Sound* and *Movement Direction* (-30°, 0°, 30°). We recorded the Vicon movement trace of both the experimenter, which provided the reference trajectory, and the participant. The speaker had two Vicon markers placed on top (and was held on top of the head) and the participant wore a Vicon-specific cap with two markers: one at the forehead and one at the crown.

Procedure
A diagram of the procedure can be seen in Figure 6. Each trial started with the participant standing immediately behind the experimenter at either end of the movement space. The experimenter held the speaker on top of his head, facing back towards the participants. The sound was started and the experimenter then walked the reference trajectory. Upon reaching the end of the trajectory, the sound was stopped and this removal of the audio was the participant's cue to start to recreate the trajectory. Once the participant had stopped at the perceived final position, they were guided to the starting position on the near side of the space, to walk the other direction for the next trial. This continued for 15 trials (each trajectory x 5) per sound.

Results
The analysis compared the final positions of the reference trajectory and the participant trajectory, with the error (distance) between the positions being an index of goodness of movement. We also analysed the overall deviation between the reference and participant trajectories as well as the total distance travelled. These would give an indication of end-point accuracy, walking route accuracy and depth perception accuracy. Average values for all 3 measures are in Table 2.

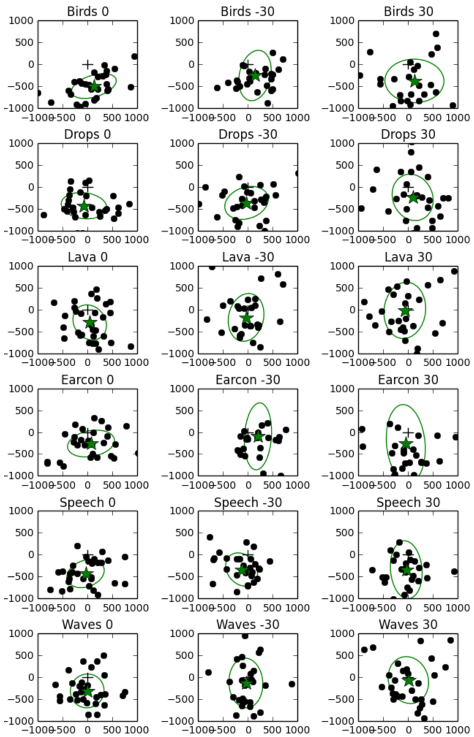

Figure 7: Distribution of end points for each sound (row) and trajectory (column). Both axes are in mm, and indicate distance from reference end-point (central cross-hair). Star = mean, ellipse = covariance.

End Point Distance

The average distance between the reference and participant end points was 62.99cm (SD = 348.71). A repeated-measures ANOVA found a significant effect of Sound on end point distance ($F_{(5,145)} = 2.30$, $p < 0.05$). *Post hoc* Bonferroni pairwise comparisons showed that the Bird sound resulted in a significantly higher distance than the *pulse*, *speech* and *waves* sounds. The mean end point distances were: Birds = 69.6cm (SD = 33.36cm); Dropping = 62.36cm (30.06cm); Earcon = 68.72cm (46.51cm); Pulse = 60.26cm (31.15cm); Speech = 59.48cm (35.45cm); Waves = 57.50cm (28.77cm).

The distributions of all end points are shown in Figure 7, arranged by sound (rows) and trajectory (columns). It shows that the distributions for 0° trajectories (left columns) are generally less spread along the y-axis (distance) than the ±30° trajectories, suggesting less variability in distance perception at 0°. However, points are consistently below the reference point (crosshair), so participants generally did not walk far enough. The distributions for -30° are more widely spread than the others, reflecting the significantly greater end-point distance in this direction.

These patterns are confirmed by looking at the average distance along the individual X- and Y-axes, where an average end point away from 0 would indicate a bias (+X = right, -X = left; +Y = too far, -Y = too short). There were small biases in the x-axis: *birds* = +13.58cm, *Earcon* = +7.87cm, *speech* = -6.36cm. Along the Y-axis, however, all sounds had an average negative value (i.e., they fell short of the end point), ranging from -16.48cm for *pulse*, to -37.57cm for *birds*.

Trajectory Deviation

The overall deviation from the reference trajectory was 32.69cm (SD = 18.32). A repeated-measures ANOVA found no effect of Sound on trajectory deviation. The average deviations were: Birds = 33.09cm (SD = 17.39cm); Dropping = 33.42cm (15.63cm); Earcon = 36.36cm (23.81cm); Pulse = 31.75cm (17.68cm); Speech = 31.11cm (17.62cm); Waves = 30.42cm (16.50cm). Overall, *waves* had the lowest deviation and *Earcon* the highest.

Total Distance Travelled

The average difference in the distance travelled by the experimenter vs. participant was 35.58cm (SD = 27.62cm). A repeated-measures ANOVA found no effect of Sound on the distance difference, with means of 39.27cm (SD = 28.76cm; Birds), 30.15cm (25.27cm; Dropping), 36.16cm (28.99cm; Earcon), 39.95cm (27.34cm; Pulse), 30.88cm (24.38cm; Speech) and 37.06cm (29.77cm; Waves). *Birds* and *pulse* produced the largest differences, while *speech* and *dropping* had the smallest.

We also looked to see whether participants tended to walk too far or not far enough, similar to the individual axis analysis of the end point distance. The average distance travelled compared to the reference was small, at -6.96cm, and suggested a tendency to not walk far enough. *Pulse* and *waves* had averages near 0 but the other sounds had negative values, from -7.02cm (*dropping*) up to -17.98 (*birds*). The finding that most sounds led to both a shorter distance and an end point nearer the start shows that participants did not walk sufficiently far enough.

Sound	**End Point Distance**	**Trajectory Deviation**	**Distance Difference**
Birds	69.60cm	33.09cm	39.27cm
Dropping	62.36cm	33.42cm	30.15cm
Earcon	68.72cm	36.36cm	36.16cm
Pulse	60.26cm	31.75cm	39.95cm
Speech	59.48cm	31.11cm	30.88cm
Waves	57.50cm	30.42cm	37.06cm

Table 2: Mean participant trajectory values for each sound, relative to reference trajectory, for each experimental measure

Discussion & Future Work

Overall, this first test suggests that the 6 sounds facilitated similarly accurate trajectory recreation, with a small number of exceptions. In one sense this is promising, as personalisation of sounds can be better supported if multiple sounds provide the same performance. However, the average error is quite high, at 63cm, suggesting it is hard to precisely move to the end location of a moving sound.

In particular, the *bird* sound, while liked by the workshop attendees, was difficult to follow, leading to inaccuracy in reaching the end point and errors in following the trajectory appropriately. Several people noted that the sound subjectively appeared to move horizontally (outside of the trajectory movement), unlike the other sounds. These results are consistent with previous research [1]. Also, all sounds resulted in participants stopping to the near side of the end point and most led to insufficient distance being walked. This contrasts with the bias mis-localizing the end point of a moving sound *beyond* its true position [4]. The difference could have come from a mismatch between where the participants believed they had walked and where they had intended to walk. In general, the *speech* and *waves* sounds facilitated slightly more accurate trajectory recreation. *Waves* had the lowest end point distance and trajectory deviation, while the *speech* sound had close to the lowest values for all measures.

These values were not significantly lower than others, but we will test if they also facilitate accurate movement or perception in our future research. This study only tested blindfolded sighted participants, and so our next experiment will test visually impaired and blind individuals. In other future work we will test trajectory recreation with more sounds, as well as testing how accurately participants can walk to a static sound and walk 2D shape trajectories.

Conclusions

The ABBI device is targeted to support sensorimotor rehabilitation of blind children, by associating physical movement with spatial audio feedback. Our preliminary test suggests that all of the initial sounds facilitate recreation of 2D horizontal movement trajectories similarly well, although *birdsong* was problematic and *speech* and *waves* were more promising. This may mean that personalisation of ABBI sounds is possible while retaining their positive effects for rehabilitative support.

References

[1] Blauert, J. *Spatial Hearing.* MIT Press, Massachusetts.

[2] Brambring, M. Divergent Development of Gross Motor Skills in Children Who Are Blind or Sighted. *Journal of Visual Impairment & Blindness 101*, 12, 620–634.

[3] Brewster, S., Wright, P., and Edwards, A. Experimentally derived guidelines for the creation of earcons. *Adjunct Proceedings of HCI 1995*, pp 155–159.

[4] Getzmann, S. and Lewald, J. Localization of moving sound. *Perception & Psychophysics 69*, 6, 1022–1034.

[5] Macleod, A. and Summerfield, A. A procedure for measuring auditory and audio-visual speech-reception thresholds for sentences in noise. *British Journal of Audiology 24*, 1, 29–43.

[6] McGookin, D. and Brewster, S. Dolphin: The Design and Initial Evaluation of Multimodal Focus and Context. *Proceedings of ICAD 2002*, pp Article 2.

[7] Röder, B., Kusmierek, A., Spence, C., and Schike, T. Developmental vision determines the reference frame for the multisensory control of action. *Proceedings of the National Academy of Sciences 104*, 11, 4753–4758.