Reaching to Sound Accuracy in the Peri-personal Space of Blind and Sighted Humans

Marc J.-M. Macé[1,2,*], Florian Dramas[1,2], and Christophe Jouffrais[1,2]

[1] IRIT, University of Toulouse, Univ. Paul Sabatier, 31062, Toulouse cedex 9, France
[2] IRIT, CNRS, Univ. Paul Sabatier, 31062, Toulouse cedex 9, France
CA: `marc.mace@irit.fr`

Abstract. With the aim of designing an assistive device for the Blind, we compared the ability of blind and sighted subjects to accurately locate several types of sounds generated in the peri-personal space. Despite a putative lack of calibration of their auditory system with vision, blind subjects performed with a similar accuracy as sighted subjects. The average error was sufficiently low (10° in azimuth and 10 cm in distance) to orient a user towards a specific goal or to guide a hand grasping movement to a nearby object. Repeated white noise bursts of short duration induced better performance than continuous sounds of similar total duration. These types of sound could be advantageously used in an assistive device. They would provide indications about direction to follow or position of surrounding objects, with limited masking of environmental sounds, which are of primary importance for the Blind.

Keywords: Sound localization, Blindness, assistive device, augmented reality.

1 Introduction

Humans are able to localize distant [1 - 2] and proximal [3 - 4] sound sources with a fair accuracy. This capability has been used previously in assistive devices for the Blind to guide them with virtual beacons [5] along a path. In the present study, the issue of sound localization is included in the context of an assistive device for the Blind [6, 7] that could assist both in navigation tasks and object localization tasks. As this assistive device is intended to be used in daily life, the sounds it produces must be designed to minimally interfere with the sounds of the environment. This could be achieved by using sounds as short as possible while still allowing good localization performance. Another important feature of this assistive device is that it should be able to orientate users with distant sounds as well as to guide reaching movements towards close-by objects with proximal sounds.

The usability of this device relies heavily on the sound localization capabilities of the Blind. A preliminary step before generating virtual sounds to indicate a position in the proximal space would be to verify that blind people are able to localize close-by

---

[*] Corresponding author.

K. Miesenberger et al. (Eds.): ICCHP 2012, Part II, LNCS 7383, pp. 636–643, 2012.

sound sources as accurately as sighted people despite the lack of vision to calibrate their acoustic space. Some studies were conducted on this issue and no definitive answer emerges as blind people show better [8, 9], identical [10] or worse [11] performance compared to sighted people depending on the experimental protocol. Also, most of these sound localization studies with blind persons had concentrated on the localization of distant [12 - 13] or continuously presented [14] sounds and little is known of their localization capabilities with very short stimuli presented in the peri-personal space.

As an assistive device for the blind based on sound localization should be able to indicate directions precisely, the first objective of this study was to determine if blind persons are able to localize brief nearby sounds and if their accuracy is comparable to the accuracy of sighted persons [4]. The second objective was to determine the most important characteristics for a sound to be correctly localized in the proximal space -to guide reaching movements- while being the least intrusive, an aspect of the auditory stimuli which had been largely overlooked in the past.

We tested a group of blind subjects and a group of sighted subjects with seven auditory stimuli, varying the number of bursts and stimulus duration. The task for the subject consisted in pointing at the perceived location of the sound with the index finger.

2 Material and Methods

2.1 Subjects

8 legally blind subjects (mean age 40.7) and 9 sighted subjects (mean age 25.8) were involved in this experiment. Half of the blind subjects were born blind; the four others lost vision between the age of 1 and 8 years. All subjects were right handed and fourteen out of seventeen subjects had normal hearing. The three subjects whose audiogram showed a deficit of 20 to 50 dB at one or two frequency bands above 4000 Hz were still included in the study as the stimuli used were all broad-band. These subjects showed no statistical differences in performance compared to the others.

2.2 Protocol

Sighted subjects didn't see the setup before the experiment and were blindfolded to follow the exact same protocol as blind subjects. An experimental platform (half-disc, radius 1 m) was equipped with 35 loudspeakers (ref.: CB990, 8 Ohm, 3 W) disposed on 5 semi-circular rows of 7 speakers each (See Fig 1), covering 180°.

The subjects were seated in the hollow part at the center of the platform, in front of the second column of speakers (0°). Subjects were pointing to the targets with the right hand. The orientation of the head was monitored with a magnetic sensor (Flock of Bird, Ascension Technology) and the pointing movements were measured with a home-made video-based tracking system (maximum RMS error inferior to 3 cm).

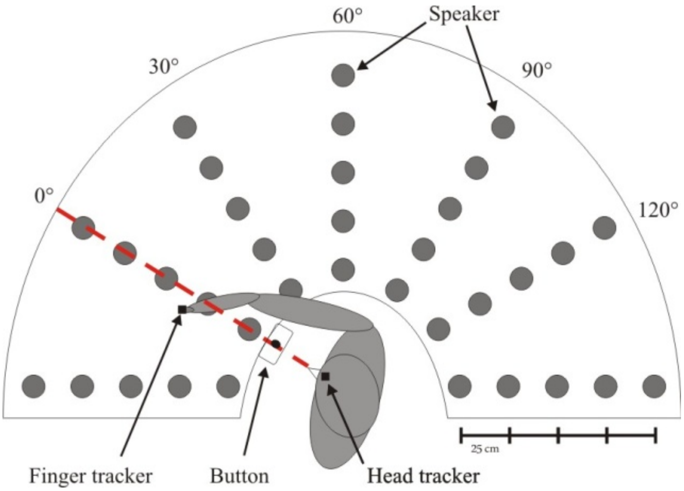

Fig. 1. The semi-circular platform was covered by 7 columns and 5 rows of loudspeakers. The subject was seated in front of the column labeled 0° (dotted line) and had to press a button to start the trial. After a delay, one of the loudspeakers emitted a sound. The subject had to point with its forefinger to the perceived sound and bring back his hand to the start button.

A trial started when the subject pressed the button of a mouse (at the starting position, Fig 1), with the additional condition that his head was orientated towards the 0° column (with +/- 2.5°). A brief auditory stimulus was then presented via a single loudspeaker. The subject had to point with the forefinger at the perceived location of the sound source before coming back to the starting position to trigger the next trial.

There were seven auditory conditions according to the number and the duration of the bursts within the sound stimulus: 1*10 ms, 1*25 ms, 1*50 ms, 1*200 ms, 2*25 ms with a 30 ms pause (50 ms of sound / 80 ms of total stimulus duration), 3*25 ms with a 30 ms pause (75 ms / 135 ms) and 4*25 ms with a 30 ms pause (100 ms / 190 ms). Each burst in the stimuli consisted of a Gaussian white noise covering 20 to 20000 Hz at approximately 60 dB.

To avoid border effects found in a previous experiment, only the results from speakers between 0 and 120° were used. Each subject performed 4 pointing movements to each stimulus on each loudspeaker, corresponding to a total of 700 trials. These trials were arranged in a complete random design within and across subjects. Subjects never had feedback on their performance during the experiment. Before the recording session, they performed a trial session with 10 pointing movements towards some of the stimuli used during the experiment.

2.3 Data Analysis and Statistics

After removing the front-back errors which were analyzed separately (2.5% of the trials; data not shown), we measured the pointing accuracy across two components: azimuth error and distance error relative to the subject. Azimuth error was computed as the absolute value (in degrees) of the difference between the azimuth of the sound source and the pointed azimuth. Distance error was computed as the absolute value (in mm) of the difference between the distance of the sound source and the pointed distance.

Analysis of variance was performed on the two groups of subjects (BLIND and SIGHTED) using the azimuth error and distance error as data. We also calculated 2-factors ANOVAs: CONDS*GROUPS for the same two measurements, to assess the performance across stimulus conditions for BLIND and SIGHTED groups. Tukey post-hoc tests were performed to assess the significant differences within factors. Significance level was set at 0.05 for all the analysis.

The data of one blind subject was removed from the analysis as her performance was at least 3 standard deviations below the performance of other blind subjects for most of the measurements.

3 Results

3.1 Azimuth and Distance Error across Groups

The average error in azimuth was significantly different between blind and sighted GROUPS (F(1, 11532)=8.96, p=0.003) with blind subjects being in average half a degree more precise in azimuth than sighted subjects (Fig 2A).

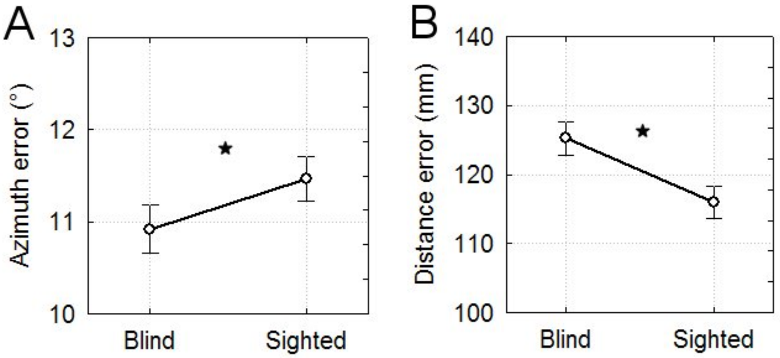

Fig. 2. Mean azimuth error in degrees (A) and mean distance error in mm (B) for Blind and Sighted subjects for all stimulus conditions together. Error bars are IC95. Blind subjects had a better accuracy than sighted subjects to evaluate the azimuth of a sound source but were less accurate to evaluate its distance.

The average error in distance was also significantly different between the two groups of subjects (F(1, 11532)=28.63, $p<10^{-5}$): sighted subjects were in average slightly better (close to 10 mm) to evaluate the distance of the sound source than blind subjects (Fig 2B).

3.2 Azimuth Error across Stimulus Conditions

For the two groups, the azimuth error depended both on sound duration and number of repetitions (Fig 3) and there was no GROUP*CONDITION interaction (F(6, 11520)=0.36, p=0.90) as the azimuth accuracy across conditions was similar for blind and sighted subjects.

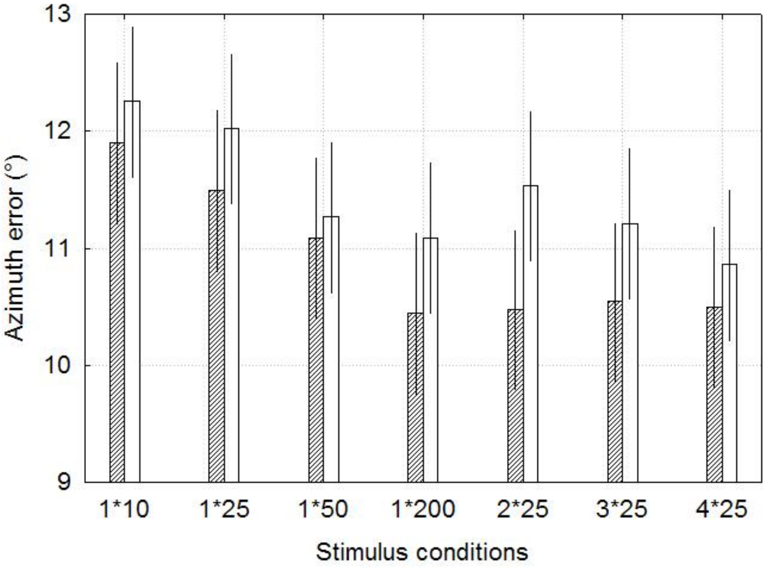

Fig. 3. Mean azimuth error in degrees for blind (hatched) and sighted (white) subjects across stimulus conditions. From conditions 1*10 to 1*200, the azimuth error decreased with sound duration. The repetition of a sound burst also resulted in an accuracy increase from 1*25 to 2*25 before reaching a plateau with 3 and 4 repetitions of the sound, especially for blind subjects. Error bars are IC95.

When considering conditions with 1 burst only (1*Sound duration), azimuth error linearly decreased with sound duration from 12.1° (1*10) to 10.8° (1*200) for both groups, meaning that pointing accuracy significantly increases with stimulus duration (F(3, 6564)=5.78, p=0.0006). Pointing accuracy also significantly increased with the number of 25 ms bursts (F(3, 6615)=3.94, p=0.008), but a Tukey post-hoc analysis revealed that the only significant difference was between 1*25 ms and the other N*25 conditions, which indicates that pointing accuracy in azimuth was already at best with only two repetitions of the sound.

Interestingly, the stimulus in condition 2*25 had a total duration of 80 ms (2*25 ms + 30 ms of silence) and induced similar accuracy in azimuth than the longer 1*200 ms stimulus, especially for blind subjects.

3.3 Distance Error across Stimulus Conditions

For the two groups, the distance error depended both on sound duration and the number of repetitions (see Figure 5) and there was no GROUP*CONDITION interaction (F(6, 11520)=0.83, p=0.55), revealing that distance estimation across conditions was similar for the two groups of subjects.

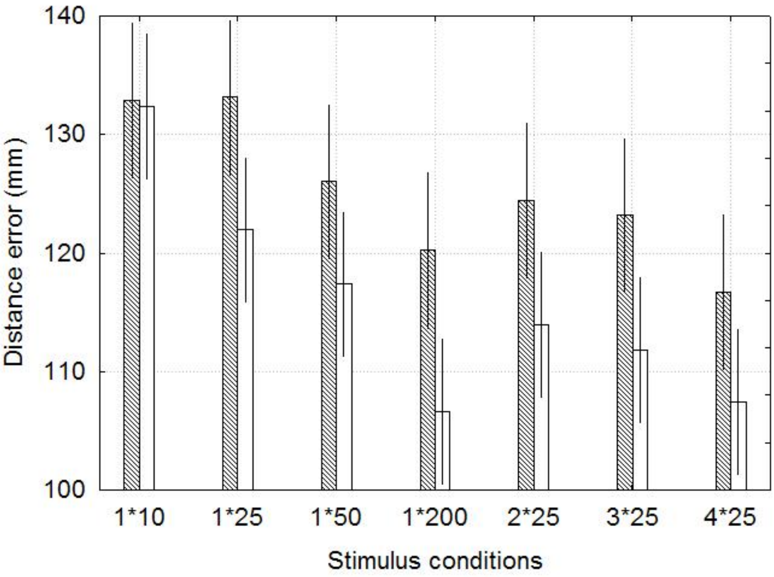

Fig. 4. Mean distance error in degrees for blind (hatched) and sighted (white) subjects across stimulus conditions. From conditions 1*10 to 1*200, the distance error decreased with sound duration. The repetition of a sound burst also resulted in an accuracy increase from 1*25 to 2*25 before reaching a plateau with 3 and 4 repetitions of the sound. Error bars are IC95.

When considering conditions with 1 burst only (1*Sound duration), distance error linearly decreased with sound duration from 133 mm (condition 1*10) to 113 mm (condition 1*200) for both group together, meaning that pointing accuracy significantly increased with stimulus duration (F(3, 6564)=13.06, $p<10^{-5}$). The number of 25 ms bursts did also significantly increased the pointing distance accuracy (F(3, 6615)=7.76, p=0.00004). However a Tukey posthoc analysis revealed that the only significant difference was between 1*25 ms and the other N*25 conditions, which indicates that pointing accuracy in distance was already at best or close to best with only two repetitions of the sound.

4 Discussion

In average, the performance of blind subjects to localize nearby short sounds matched the performance of sighted subjects. For sounds with a unique burst, pointing accuracy of blind subjects increased with sound duration up to a plateau around 10.5° and 120 mm when pointing to a 200 ms sound. Interestingly, the same performance was reached when only two bursts of 25 ms were presented. The total duration of this double burst stimulus was 80 ms (2*25 ms + 30 ms silence), compared to the 200 ms of the unique burst. This advantage of multiple sound bursts over continuous sounds can be attributed to the localization cues specifically present in the onset and offset of a sound [15]. As a direct conclusion of this result, the sounds used in an assistive device for the Blind should be composed of short bursts to maximize the localization performance while preserving normal hearing of the environment. Preserving the surrounding sounds is especially important as blind persons heavily rely on auditory cues for Orientation and Mobility skills.

In conclusion, blind subjects were able to locate short sounds in peri-personal space with a good accuracy. The average accuracy for the optimal condition (2*25 ms white noise burst) was around 10° in azimuth and 12 cm in distance, which is precise enough to orientate the subjects towards a specific goal and could even guide a grasping movement until tactile feedback occurs when the object is reached. Coupled with an artificial vision system, this approach, where nearby targets could be sonified with good accuracy, could well complement navigation systems [5] where only far-field targets and landmarks are indicated. It could also be extended by using short sounds designed to convey additional information such as earcons [16] or spearcons [17].

This performance with real sounds is a first step towards designing an assistive device where virtual sounds will be used instead of real sounds. It is possible to generate virtual sounds that will be perceived at any spatial location by filtering the binaural signal in a specific way that depends on each person's morphology [18]. This set of transfer function are called HRTF's (Head Related Transfer Functions) and are obtained by recording different sounds with a microphone inserted inside the ear. The next step in the development of our assistive device will be to evaluate the sound location capabilities of nearby and distant short virtual sounds by blind people.

Acknowledgements. This work has been supported by the French Foundation for Medical Research (FRM-INE2005) and by the French National Research Agency (ANR) through TecSan program (project NAVIG n°ANR-08-TECS-011). The authors thank BFG Katz for helpful comments on the design of the setup and experiment.

References

1. Middlebrooks, J.C., Green, D.M.: Sound localization by human listeners. Annual Review of Psychology 42, 135–159 (1991)
2. Blauert, J.: Spatial hearing: the psychophysics of human sound localization. MIT Press (1997)
3. Brungart, D.S.: Auditory localization of nearby sources. III. Stimulus effects. The Journal of the Acoustical Society of America 106, 3589–3602 (1999)

4. Brungart, D.S., Durlach, N.I., Rabinowitz, W.M.: Auditory localization of nearby sources. II. Localization of a broadband source. The Journal of the Acoustical Society of America 106, 1956–1968 (1999)
5. Loomis, J.M., Golledge, R.G., Klatzky, R.L., Speigle, J.M., Tietz, J.: Personal guidance system for the visually impaired. In: Proceedings of the First Annual ACM Conference on Assistive Technologies, Assets 1994, pp. 85–91. ACM Press, New York (1994)
6. Katz, B.F.G., Dramas, F., Parseihian, G., Gutierrez, O., Kammoun, S., Brilhault, A., Brunet, L., Gallay, M., Oriola, B., Auvray, M., Truillet, P., Thorpe, S., Jouffrais, C.: NAVIG: Guidance system for the visually impaired using virtual augmented reality. Technology and Disability (in press, 2012)
7. Kammoun, S., Parseihian, G., Gutierrez, O., Brilhault, A., Serpa, A., Raynal, M., Oriola, B., Macé, M.J.-M., Auvray, M., Denis, M., Thorpe, S.J., Truillet, P., Katz, B.F.G., Jouffrais, C.: Navigation and space perception assistance for the Visually Impaired: The NAVIG project. BioMedical Engineering and Research 33, 182–189 (2012)
8. Röder, B., Teder-Sälejärvi, W., Sterr, A., Rösler, F., Hillyard, S.A., Neville, H.J.: Improved auditory spatial tuning in blind humans. Nature 400, 162–166 (1999)
9. Lessard, N., Paré, M., Lepore, F., Lassonde, M.: Early-blind human subjects localize sound sources better than sighted subjects. Nature 395, 278–280 (1998)
10. Zwiers, M.P., Van Opstal, A.J., Cruysberg, J.R.: Two-dimensional sound-localization behavior of early-blind humans. Experimental Brain Research 140, 206–222 (2001)
11. Zwiers, M.P., Van Opstal, A.J., Cruysberg, J.R.: A spatial hearing deficit in early-blind humans. The Journal of Neuroscience 21, RC142, 1–5 (2001)
12. Lewald, J.: Vertical sound localization in blind humans. Neuropsychologia 40, 1868–1872 (2002)
13. Voss, P., Lassonde, M., Gougoux, F., Fortin, M., Guillemot, J.-P., Lepore, F., Postale, C.: Early- and late-onset blind individuals show supra-normal auditory abilities in far-space. Current Biology 14, 1734–1738 (2004)
14. Wanet, M.C., Veraart, C.: Processing of auditory information by the blind in spatial localization tasks. Perception And Psychophysics 38, 91–96 (1985)
15. Rakerd, B., Hartmann, W.M.: Localization of sound in rooms, III: Onset and duration effects. The Journal of the Acoustical Society of America 80, 1695 (1986)
16. Blattner, M., Sumikawa, D., Greenberg, R.: Earcons and Icons: Their Structure and Common Design Principles. Human-Computer Interaction 4, 11–44 (1989)
17. Walker, B.N., Nance, A., Lindsay, J.: Spearcons: speech-based earcons improve navigation performance in auditory menus. In: Proceedings of the International Conference on Auditory Display, pp. 95–98 (2006)
18. Katz, B.F.G., Parseihian, G.: Perceptually based head-related transfer function database optimization. JASA 131 (2012)