Thomas Behrens[1,2] & Florian Kramer[1,3]

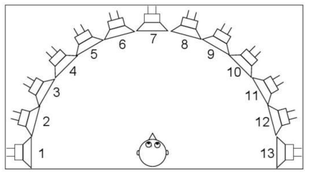
Simultaneous localization and identification of environmental sounds in auditory scenes by normal hearing and hearing impaired listeners

Aim

In many everyday listening situations, it is necessary to simultaneously localize and identify concurrent sound sources. This type of dual task has only been scarcely used in formal testing. Also, many localization experiments either use artificial sounds or speech. Therefore, a test paradigm was developed to investigate the abilities of normal hearing and aided hearing impaired listeners in simultaneous identification and localization of a broad range of environmental sounds.

Setup

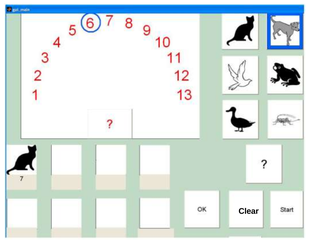
The test took place in the anechoic room at the Eriksholm Research Centre. Thirteen loudspeakers were placed equidistantly with a spacing of 15º in the frontal horizontal plane, as shown on the illustration given below.



Schematic of the placement of speakers in the horizontal plane

Procedure

Test persons were asked to localize and identify sounds of an auditory scene. An example of a user interface is given here:



Example user interface for scene ”Park”

The test person completed a run by listening to the presented sounds and then entering sounds by means of the UI. This was done in a three step procedure: selecting a perceived sound by means of the icons, selecting a perceived location, and pressing ok.

A block consisted of 30 runs. Each run was a selection of 3, 4, or 5 sounds from the current scene. Sounds were presented in the following way:

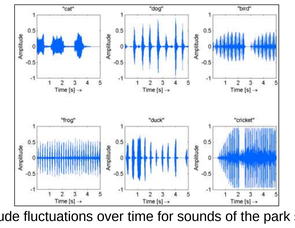
Number of sounds	Presentation location
3	5, 7, 9
4	4, 6, 8, 10
5	3, 5, 7, 9, 11

Stimuli

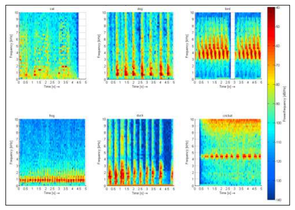
The stimuli chosen for the experiment were 5 second extracts from high quality audio libraries and sound archives. They were used to create the following scenes:

Scene	Elements	Used for
Zoo	Elephant, Chimpanzee, Rattlesnake, Parrot, Lion	Training
Home	TV (Weather report), Steps on stairs, Door bell, Clock ticking, Door opening & closing, Phone ringing	Test
Restaurant	Dishes clattering, Soda poured into glass, Sizzling pan, Music, Newspaper (turning pages), Man talking	Test
Park	Dog, Cat, Frog, Duck, Bird, Cricket	Test
Farm	Cow, Horse, Sheep, Rooster, Pig, Chicken	Test

Each stimuli was constructed to have the characteristic acoustic event of each sound element being repeated a number of times during presentations. This was done to give the listener the chance of glimpsing all sounds in the presented background. Analyses of amplitude fluctuations and spectrograms as shown below, were used to assure this.



Amplitude fluctuations over time for sounds of the park scene.

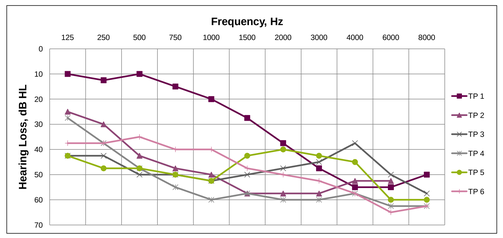


Spectrograms of the six sounds of the park scene.

Overall signal levels were adjusted so that the long term average level did not differ by more than +/- 1 dB. Presentation level was 65 dB SPL.

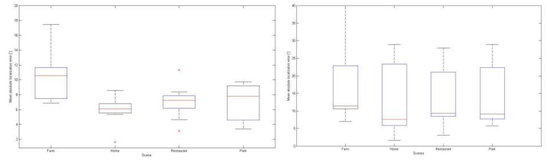
Test Subjects

Nine normal hearing and six hearing impaired (HI) test persons took part in the experiment. Pure tone thresholds of HI listeners were:



Results

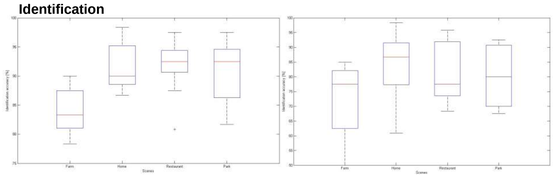
Localization error and identification accuracy from normal hearing (NH) and hearing impaired (HI) listeners is reported below.



Mean absolute localization errors for the four test scenes. Results for NH listeners to the left and HI listeners to the right. Note factor two difference in y-axis range.

Localization errors shown are larger and vary more for HI listeners than for NH listeners. Error is greater for the Farm environment than the others, across both groups of listeners. For NH listeners spread in error is very small for Home and Restaurant, whereas it varies substantially more for the Farm and Park environments. For HI listeners spread in error is large across all used sound scenes.

Identification



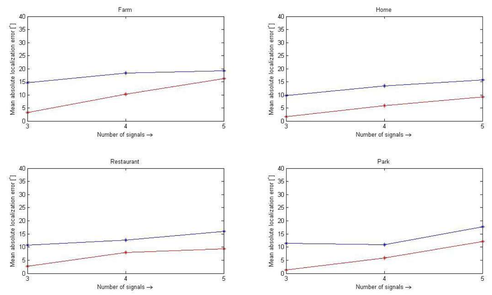
Overall identification accuracy for the four test scenes. Results for NH listeners to the left and HI listeners to the right. Note again factor two difference in y-axis range.

Identification accuracy for NH listeners vary considerably across environments, with highest performance for the Restaurant and Park environments and poorest performance for the Farm environment. Spread in performance also vary considerably, with the smallest spread found for the Restaurant environment and the largest found for the Park and Farm environments.

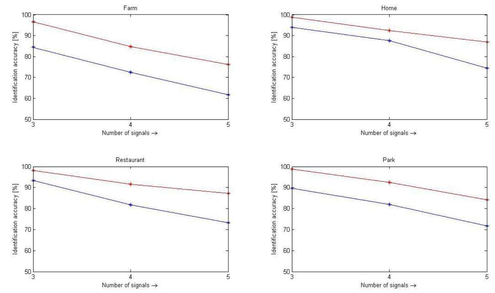
For the HI listeners, performance is poorer than that found for NH listeners. Also, spread in performance again is much larger than that found for NH listeners. It is interesting to note that the scene that produced the highest performance for HI listeners does not coincide with the ones that produced the highest performance for NH listeners.

Influence of number of signals presented

Task complexity was not only affected by the type of sound scene presented. Also, the number of signals presented had a large influence on accuracy, as shown in the following.



Mean localization error for each of the four test scenes and as a function of number of presented signals. NH results in red and HI in blue.



Mean identification accuracy for each of the four test scenes and as a function of number of presented signals. NH results in red and HI in blue.

Observing performance across HI and NH groups as the number of signals increases suggests that localization performance differences between the groups decreases with number of signals. In contrast, the difference in identification accuracy between the two groups increases with the number of signals.

This hints at a difference in listening strategy between the groups. As the listening environment becomes more complex, HI listeners focus on a lower number of signals and try to localize them correctly, whereas NH listeners try to localize and identify all signals correctly.

Summary

Localization error for normal hearing listeners was on the order of 2-4 degrees in the three signal condition dependent on the scene. For the hearing impaired group mean localization error of 10-15 degrees was typical for the three signal condition. The mean identification performance of the normal hearing and hearing impaired test persons was around 90% and 80% correct in the three signal condition.

The performance decreased with increasing number of signals for both groups by 7-12 degrees (localization) and 10-20% (identification). Furthermore, factors such as scene, number of signals and test-person were found to have a highly significant influence on the results for both localization and identification ($p<0.01$).

Acknowledgements

The work presented here was part of the Bachelor thesis work of Florian Kramer. The work was done at the Eriksholm research centre. The authors wish to thank Louise Kragelund for help with HI testing.

Selected references

1. Eramudugolla, R., Irvine, D., McAnally, K., Russell, M. and Mattingley, J. (2005). Directed Attention Eliminates ’Change Deafness’ in Complex Auditory Scenes. Current Biology, 15:1108–1113.
2. Hawley, M., Litovsky, R. and Colburn, H. (1999). Speech intelligibility and localization in a multi-source environment. Journal of the Acoustical Society of America, 105(6):3436–3448.
3. Gygi, B. and Shafiro, V. (2007). Effect of auditory context on the identification of environmental sounds. 19th International Congress of Acoustics.

1: Oticon A/S, Research Centre Eriksholm, Snekkersten, Denmark
2: Oticon A/S, Team Audiology and Algorithms, Smørum, Denmark
3: Jade Hochschule, Institute of Hearing Science and Audiology, Oldenburg, Germany

oticon
PEOPLE FIRST