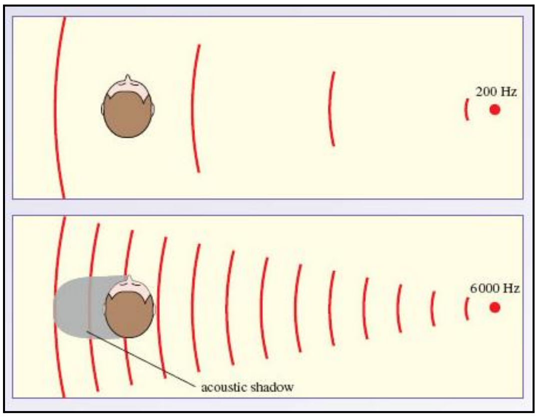
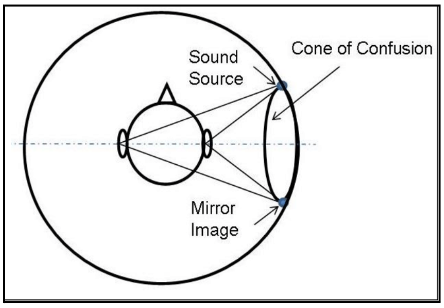
Introduction
To localize sounds humans are mainly using three different sounds localization cues. The interaural time difference (ITD), the interaural level difference (ILD) and the spectral shape cues.
The first one is the interaural time difference (ITD). This cue is based on the fact that sounds that are coming from the left reacht the left ear first with some time difference ΔT. This time difference is based on the distance Δd from a point to both ears. ITDs are working for frequencies below 1.5 kHz and you can only distinguish azimuth with them.
The second one is the interaural level difference. This is based on the head-shadow effect as seen below. For high frequencies that have a wavelength shorter than the head an acoustic shadow is cast such that the intensity of the wave behind the ear is lower than before it. This results in a level difference between the ears. In the figure below this is shown. ILDs are working above 1.5 kHz and you can only distinguish azimuth with them.



12.4 Interaural intensity differences. (z.d.). OpenLearn. Geraadpleegd op 9 oktober 2022, van https://www.open.edu/openlearn/science-maths-technology/biology/hearing/content-section-12.4

The problem with ITDs and ILDs alone is that they still give rise to ambiguity. In the figure below you can see that points in a circle around the line perpendicular to the midsagittal plane have the same head-shadow effect and time difference to both ears. This is called the cone of confusion and is shown in the figure below. ITDs and ILDs need both ears and are binaural cues.



Auditory Spatial Perception: Auditory Localization - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/The-concept-of-the-cone-of-confusion_fig2_265033553 [accessed 7 Oct, 2022]

To still be able to localize sounds the spectral shape cues are needed. These are based on all kinds of body parts such as the head and shoulders, but are most influenced by the pinnae. With this you can in the end pinpoint very accurately where sounds are coming from. The HRTF is the function that gives a print of how someone changes sound intensity for each frequency and elevation. You can use this for frequencies above 1.5 kHz and for frequencies above 4 kHz this really starts to differ, because the pinnae starts to be important there. The HRTF is something only one ear can do using monaural cues.

Methods

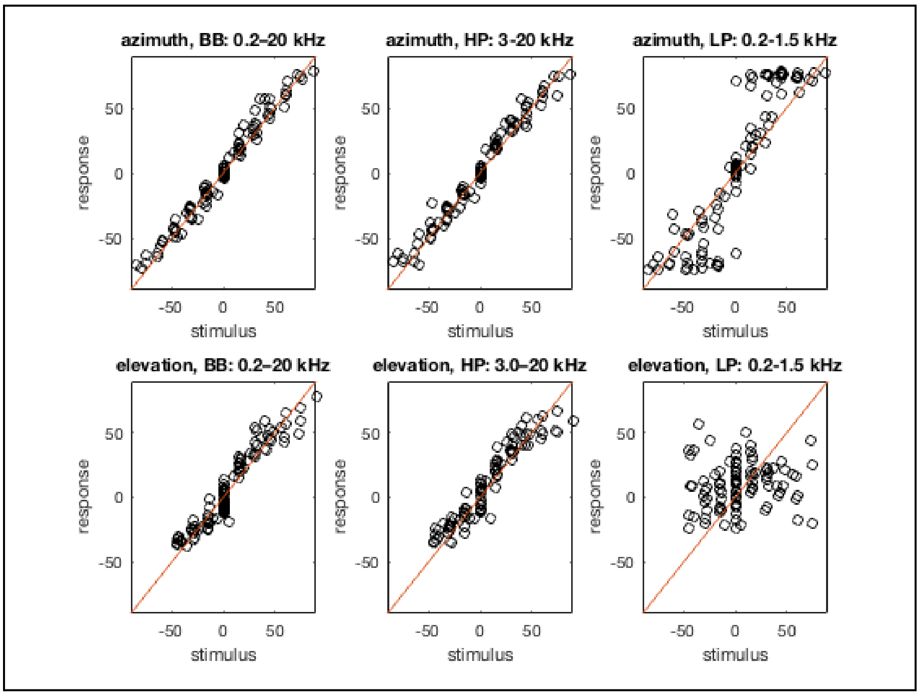
The setup of the experiment is that a test subject tries to locate sounds of different frequencies that are presented. The subject sits on a chair with in front of him a centered light that is presented during the experiment. First a calibration takes place such that it is known how the head of the subject is located in space.

Then the lights are turned off and for 15 minutes sounds are presented around the subject in the hemisphere in front of him. The sounds are grouped in three sorts of noise, namely broadband filtered (0.-20 kHz), high-pass filtered (0.2-1.5 kHz) and low-pass filtered (0.2-1.5 kHz).

The test subject is Sam(=me!) and has no hearing disorders of affected motor responses.

Results

The figure below shows the results of the experiment where the stimulus location and the response location in degrees are plotted against each other. This is done for azimuth and elevation. The orange line indicates where the response would be exactly at the stimulus place. As can be seen for azimuth and most of the measurements are around the line, but for LP sounds and elevation this is really not the case.



This can also be shown more numerically by plotting the slope and offset of a linear regression through the points, which is done in the table below.

Slope	BB	HP	LP
azimuth	0.9880	0.9696	1.2478
elevation	0.9436	0.8859	0.0748

Offset	BB	HP	LP
azimuth	2.3989	2.3162	1.8528
elevation	1.8125	3.5107	9.8491

Discussion

The results show that the azimuthal sounds can be localized quite well. For the elevations for high frequencies and broadband frequencies this is a little less, but still quite well. This is because the HRTF needs to be used here, but is in the frequency range that it can predict quite well. For low frequencies, specifically below 1.5 kHz, the HRTF is not valid and therefore the slope that needs to be around 1 is very much off as is the offset which should be around 0.