



Norwegian University of
Science and Technology

Localization of Sounds in the Horizontal Plane using the QuietPro System

Øyvind Tønsberg

Master of Science in Communication Technology

Submission date: August 2009

Supervisor: Odd Kr. Pettersen, IET

Co-supervisor: Olav Kvaløy, SINTEF

Problem Description

QuietPro is a combined hearing protection and communications terminal produced by NACRE. The system has been very well received in the market. One central feature of the system is the realistic hearing experience in a quiet environment.

The assignment consists of designing and running an experiment to quantify the sound localization performance of the QuietPro system compared to natural hearing.

Assignment given: 02. March 2009
Supervisor: Odd Kr. Pettersen, IET

Abstract

This report studies the effect of electronic hearing protectors on directional hearing in the horizontal plane. 11 subjects participated in a sound localization test, sitting in the center of a circle of 24 speakers. The test was run once with open ears, and once wearing QuietPro earplugs from Nacre. The subjects were presented with 195 trials consisting of a randomly located 150 ms broadband noise burst, and instructed to identify the source location.

The results show that there was a significant increase in localization errors when wearing the electronic hearing protectors, particularly due to an increase in source reversals. Large individual differences between subjects was also observed in this occluded condition.

Contents

Abstract	i
1 Introduction	1
2 Background	3
2.1 The ear	3
2.1.1 The outer ear	4
2.1.2 The middle ear	4
2.1.3 The cochlea	5
2.1.4 Hearing damage	6
2.2 The neural system	8
2.2.1 Neural signal transmission	8
2.2.2 Central auditory pathway	9
2.3 Localization of sound	11
2.3.1 Interaural Time Differences	11
2.3.2 Interaural Level Differences	12
2.3.3 The Pinna and other factors	13
2.4 Similar studies	14
2.5 Stimuli	15
2.5.1 Spectral contents	15
2.5.2 Duration	16
2.6 The QuietPro system	16
3 Methods	17
3.1 Subjects	17
3.2 Lab setup	18
3.3 Experimental protocol	20

3.4	Data analysis	21
4	Analysis and Results	23
4.1	Preliminary analysis	23
4.2	Individual results	24
4.3	Averaged results	27
5	Discussion	30
5.1	Localization errors	30
5.1.1	General accuracy	30
5.1.2	Source reversals	31
5.1.3	Comparison to other studies	32
5.2	Causes of localization errors	32
5.3	Sources of error in current experiment design	33
5.4	Suggestions for further studies	34
5.4.1	Spherical localization	34
5.4.2	Training	35
6	Conclusions	36
A	Audiograms	37
B	Subject-specific scatter plots	39
C	Subject-specific <i>t</i>-test results	44
D	Subject-specific averaged results	51
E	Sound spectrum measurements	58
	References	61

1 Introduction

We rely on sensory input to understand the world around us. While vision is our primary, and best developed sense, the sense of hearing also has a central role. It registers sound waves, mechanical vibrations in the air around us, and then attempts to translate them into an auditory event, or perception of this sound.

This process is astoundingly complex. Take the following thought experiment: standing on the shore of a lake, you dig two narrow parallel canals. Now, from only observing the wave patterns at the end of these two canals, describe the shape of the lake, how many boats are sailing on it, where they are going, and so forth. This is roughly similar to what our auditory system does, only in three dimensions. In addition, we are able to interpret and understand the complex wave patterns of speech, and even pick out and listen to one single speaker in a noisy crowd.

Our hearing is an important tool to us, and thus it is vital to protect it from harm. Hearing loss is often permanent, as the cells involved do not regenerate in the same way as our skin or muscles do. Because of this, prevention is the only way to protect the ears in noisy environments; once the damage has occurred, it is too late.

As such, we have developed a wide range of ways to protect our ears. Solutions range from simply covering them with our hands, or stuffing cotton in them to sleep better, to complex electronic hearing protectors who actively measure and analyze the acoustic environment, and generate noise-canceling sound waves for protection.

In a lot of professions, people are subjected to harmful noise levels. It may be constant loud exposure, as in a factory or at an heliport, or it may be the potential risk of gunfire or explosions that military personnel often is exposed to. For the latter case in particular, it is important that the hearing protection employed interferes as little as possible with the user's normal senses.

In this study, one such electronic hearing protection system is examined: the QuietPro Intelligent Hearing System, produced by Nacre AS, Norway. It is a combined communications terminal and hearing protector, and consists of two small earplugs, and a separate digital signal processing (DSP) unit. The earplugs have an outer microphone and inner speaker, and the DSP unit analyzes the recorded sounds. The system is designed to let all sound pass through unaltered, unless designated harmful; e.g. when the impulse

or averaged intensity exceeds set limits. It also features an additional inner microphone, which is used to pick up the user's voice from within the ear canal, for radio communication in noisy environments.

In particular, the focus in this study was on localization of sound sources, and how the users' quality of experience is affected, and potentially reduced, while wearing this system in a low-noise environment. This was explored through a set of listening experiments in an audio lab at the Norwegian University of Science and Technology, NTNU, comparing localization performance with the QuietPro system to a normal hearing condition.

The accurate localization of sounds is again of typical interest in military applications, for which the QuietPro system was originally developed. The fast and accurate detection of, say, a branch snapping in the forest somewhere close to a soldier may be the difference between life and death.

The rest of this report has the following structure: Chapter 2 offers a theoretical background and overview of the auditory system, explaining the choices behind the experiment design based on findings in the literature. In Chapter 3, the experiment design is described in detail, including the lab setup and test protocols. Chapter 4 presents the results, both for individual subjects, and comparing the averaged results with and without the QuietPro system. Chapter 5 is dedicated to discussion of these results, exploring the most likely causes of localization error, and suggesting ideas for future research. Chapter 6 sums up the central findings, and various plots for the individual subjects are added as appendices.

2 Background

This chapter gives a short overview of the ear in general, and localization of sound in particular. The basics of neural signal processing are introduced, and some of the neural mechanisms involved in sound localization are explored. Throughout the chapter, references to relevant papers are given, and the last section will explore a few papers of particular relevance to the current study.

2.1 The ear

The ears are responsible for translating the mechanical vibrations of the air into auditory events which the mind can interpret. This process involves acoustic, mechanical and neural mechanisms, at the center of which is the mechanical Fourier-like transform performed in the cochlea, where vibrations are translated to neural impulses.

Throughout this chapter, and the rest of the report, a spherical coordinate system will be used to indicate direction, where 0° azimuth and 0° elevation is directly ahead, 90° azimuth is to the right, and 90° elevation is straight up. See also figure 1.

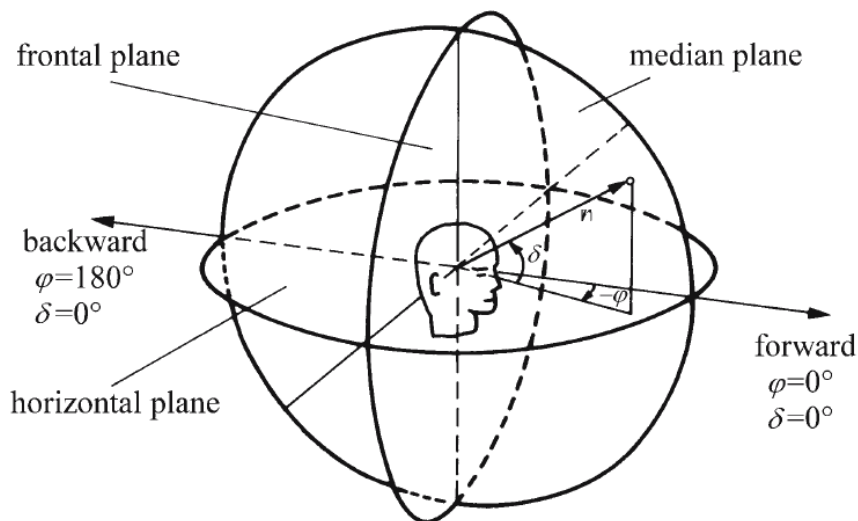


Figure 1: Spherical, head-related coordinate system, adapted from Blauert [1, figure 1.4].

2.1.1 The outer ear

The outer ear (figure 2a) is the only visible part of the human auditory system, and consists of the pinna and the ear canal, which terminates at the tympanic membrane, or ear drum. The pinna functions as a linear filter, depending on the direction and distance of a sound source, caused by several different physical phenomena, including reflection and resonance.

The effects of the concha in particular are of special interest to this study, since it is this part of the outer ear that is occluded when the QuietPro earplugs are being used. One example of the spectral filtering are the natural resonances that occur in the concha at 3 kHz and 5 kHz [2], who can be turned into corresponding dips by filling the concha with putty. This is however just an example of the many filtering effects caused by the pinnae.

One should also consider the effects of reflections from the head, shoulders and torso when describing the filtering effect of the outer ear, as these also contribute to the directional information regarding a sound source. The combined filtering effect—from the periphery to the ear canal—is known as a head-related transfer function, or HRTF [3, 4], and is slightly different for each person.

2.1.2 The middle ear

The middle ear (figure 2b) is composed of the tympanic membrane, the oval window, and the ossicles connecting the two, transmitting and amplifying the vibrations of the air in the outer ear to the fluid-filled inner ear.

The ossicles are three tiny bones—the malleus, incus and stapes—also known as the hammer, the anvil and the stirrup. The malleus is connected to the tympanic membrane, picking up the minute vibrations and transmitting them through the incus to the stapes, which in turn is connected to the oval window and the cochlear complex. The pressure transfer function of the middle ear amplifies the signal with about 15-20 dB, making it possible to transfer the vibrations of the air to the fluid in the inner ear. This is caused by the relative difference in size of the two membranes—the oval window area is $\frac{1}{16}$ th of the tympanic membrane—as well as the structure of the ossicles, who function as a lever. For a more exhaustive look into the functionality of the middle ear, there's an excellent chapter by Hudde in [3, Chapter 3].

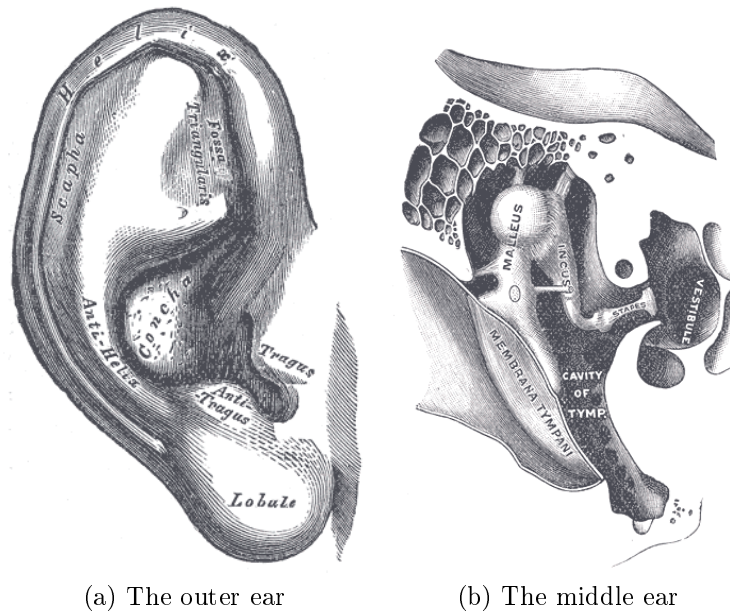


Figure 2: Figures taken from Gray's Anatomy.

2.1.3 The cochlea

The cochlea is the spiral-shaped, fluid-filled part of the inner ear responsible for translating the acoustic waves from the outer and middle ear into neural signals for the brain. The cochlea cross-section is composed of three chambers—the scala tympani, scala vestibuli and scala media—separated by membranes. The scala tympani and vestibuli are linked together at the apex of the cochlea, whereas the scala media is a separate chamber, with a different ion concentration. This difference in concentration generates a voltage potential across the membranes, which is necessary for the transmission of neural signals, covered in the next subsection. Background info can be found in Purves et. al. [5, Chapter 12].

The basilar membrane and the organ of Corti attached to it, are the central components of the cochlea. The basilar membrane varies in width and stiffness from the base to the apex of the cochlea, and thus it is frequency-selective, resonating near the base for high-frequent sounds, and towards the apex for lower frequencies. This allows the cochlea to function as a kind of mechanical Fourier transform or filter bank, detecting the different frequency components of a complex sound.

As can be seen from figure 3, the organ of Corti contains two sets of cells known as the outer and inner hair cells. These cells have thin hairs, or stereocilia, who are connected to the tectorial membrane, a stiff, hinged plate in the scala media. The hinges of the tectorial and basilar membranes are offset, causing a shear force on the stereocilia when the basilar membrane is displaced by a sound wave. The outer and inner hair cells do however have different functions, and will be reviewed separately.

The inner hair cells are the main signal transducers of the ear. When the aforementioned shear force is applied to their stereocilia, small ion gates are opened, letting ions flow into the cell and depolarizing it. This causes the cell to release neurotransmitters, which is picked up by the auditory neurons, triggering the transmission of neural signals. When the shear force is applied in the opposite direction, the ion gates close, hyperpolarizing the cell and stopping signal transmission. This allows the auditory system to code the sound frequency by transmitting neural pulses phase-locked to the wavefronts of the sound. See section 2.2.1 for an in-depth explanation of how neural signal transmission works.

The outer hair cells do not transmit sound information to the brain like the inner hair cells do, but rather function as a pre-amplifier, stiffening or vibrating in reaction to the same shear forces. This causes the resonance of the basilar membrane to become narrower, by dampening the response on the sides of the peak, as well as amplifying it at the peak itself by resonating at the same frequency. This is particularly useful for weaker sounds, that we otherwise might not be able to detect.

2.1.4 Hearing damage

The ear, and in particular the cochlea, is a highly sensitive structure, but this also makes it vulnerable to very loud noises. The hair cells in the organ of Corti do not regenerate, so if some are destroyed, that frequency band is attenuated or lost for ever.

It is useful to divide potentially damaging noise into two categories; one consists of high intensity, impulse-shaped noises, like a gunshot; and the other of continuous noise over longer periods of time, e.g. working in a noisy factory. The two categories are harmful in different ways.

The impulse noise damage is known as an acoustic trauma, and can physically sever the thin hairs linking the hair cells to the basilar membrane. The hair cells sensitive to high frequencies are most vulnerable to this, as they are

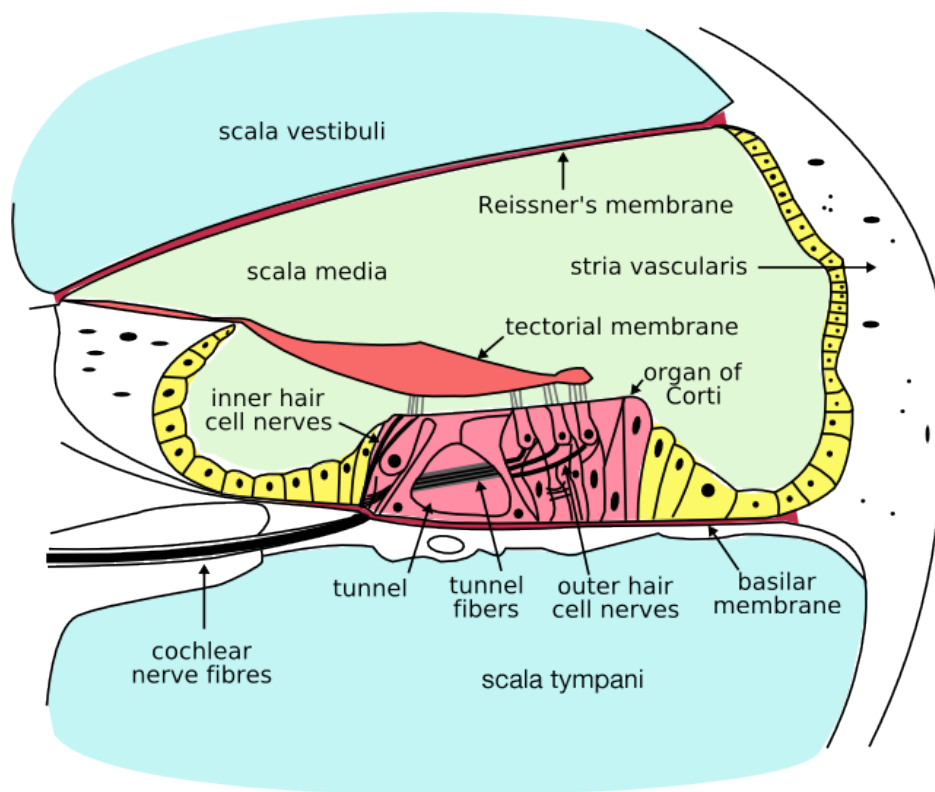


Figure 3: A cross-section of the cochlea, showing the scalae, the basilar membrane and the organ of Corti.

located just inside the oval window, where all wave propagation along the basilar membrane begins.

Noise-induced hearing loss (NIHL) over time is somewhat more complex. When the ear is exposed to continuous loud noise, the outer hair cells work hard to fine-tune the basilar membrane, and the ion pumps in the stria vascularis (see figure 3) are constantly active to restore the ion balance. Both these activities generate reactive oxygen species (ROS), commonly known as free radicals. Normally, mechanisms in the cochlea reabsorb these chemicals, but a continuous strain on the system causes an excess of ROS, which can cause cell damage [6].

Based on this, international standards for noise exposure at work—like Directive 2003/10/EC in the European Union [7]—have been established, defining an average noise exposure level during an 8-hour workday at 85 dBA SPL as the threshold at which hearing damage may occur, and protective measures should be taken. Following the same scale, the exposure limit at 88 dBA

SPL is 4 hours, and only 15 minutes at 100 dBA SPL—which is a typical sound level at a rock concert.

2.2 The neural system

The auditory neural system is a complex, massively parallel structure, that science is only beginning to understand the workings of. This section does not attempt to cover all of its intricacies, but will give a short overview of some of the neural centers relevant to feature extraction and localization, as well as a short introduction to neural signal transmission in general. See again Purves et. al. [5, Unit I, and Chapter 12] for further details.

2.2.1 Neural signal transmission

Nerve cells, or neurons, communicate by means of electrical pulses, known as action potentials. A typical neuron consists of a cell body; one or more dendrites, who can be considered inputs; and an axon, or output. Dendrites and axons of separate neurons are linked together by synapses, which are nerve ends that communicate by transmitting chemicals across a short open span, known as the synaptic cleft. Groups of neurons are linked together in complex functional structures known as neural networks.

The action potential is the basic signal pulse in neural communication. Selective ion permeability in the cell membrane of a neuron generates an electrical difference across the membrane, known as the resting potential. If the membrane is depolarized, mechanisms within the neuron causes a powerful electrical pulse—an action potential—to propagate along the length of the axon.

The depolarization of a neuron can be caused by a number of factors. Examples include synaptic transmission from other neurons, or—as is the case of the inner hair cells, although they are not strictly classified as neurons—direct mechanical manipulation of ion gates in the cell membrane. In either case, the distribution of ions on the inside and outside of the neuron is altered. Conversely, the electrical difference can also be increased, causing a hyper-polarization of the membrane.

Most neurons have some degree of idle activity, where they generate some action potentials without any external input. The frequency of this activity depends on the size and type of the neuron. As mentioned, depolarizing the

neuron will generate an action potential, or in most cases, an increased firing rate. Hyper-polarization on the other hand will stop even this idle activity.

Complex signals are coded as a function of action potential frequency and amplitude, e.g. the phase-locked pulses generated in the cochlea for low-frequency sounds, or more continuous, tonic pulse trains for high-frequency components.

The propagation speed of the action potentials depends on the type of axon, and is faster if the axon is isolated, i.e. wrapped in a white substance known as myelin. In the inner ear, afferent¹ axons from the inner hair cells towards the brain are myelinated, whereas the efferent axons terminating at the outer hair cells are not. The thickness of this myelin sheet is one way for the brain to vary the signaling speed, but at a trade-off for nerve density.

2.2.2 Central auditory pathway

The central auditory pathway is a complex, parallel structure, consisting of a number of nuclei, or nerve clusters, shown schematically in figure 4. These initial stages—located in the brain stem—are mainly responsible for various feature extractions, and preliminary, subconscious analysis.

From the cochlea, the auditory nerve leads to the cochlear nucleus, which is partitioned into a dorsal and ventral region. The ventral cochlear nucleus is tonotopically² organized, and leads to the superior olivary complex, which receives input from both ears, and is responsible for the initial analysis of time and intensity differences.

In the dorsal cochlear nucleus, a more complex neural network is thought to analyze the monaural spectral content of the signal, before transmitting it further up the pathway to the lateral lemniscus and the inferior colliculus.

The inferior colliculus (IC) integrates the input from all the previous nuclei, including the frequency analysis from the cochlear nucleus and the binaural localization analysis from the olivary complex, allowing more complex analysis. In the barn owl, an auditory space map has been demonstrated here, with specific neurons responding to a preferred elevation and azimuth. Such an auditory space map has not yet been identified in humans or other mammals, but it is likely that a similar organization exists.

Further on, the auditory nerve passes through the auditory thalamus, or

¹Afferent: towards the central nerve system

²Tonotopic: anatomically organized as a function of frequency

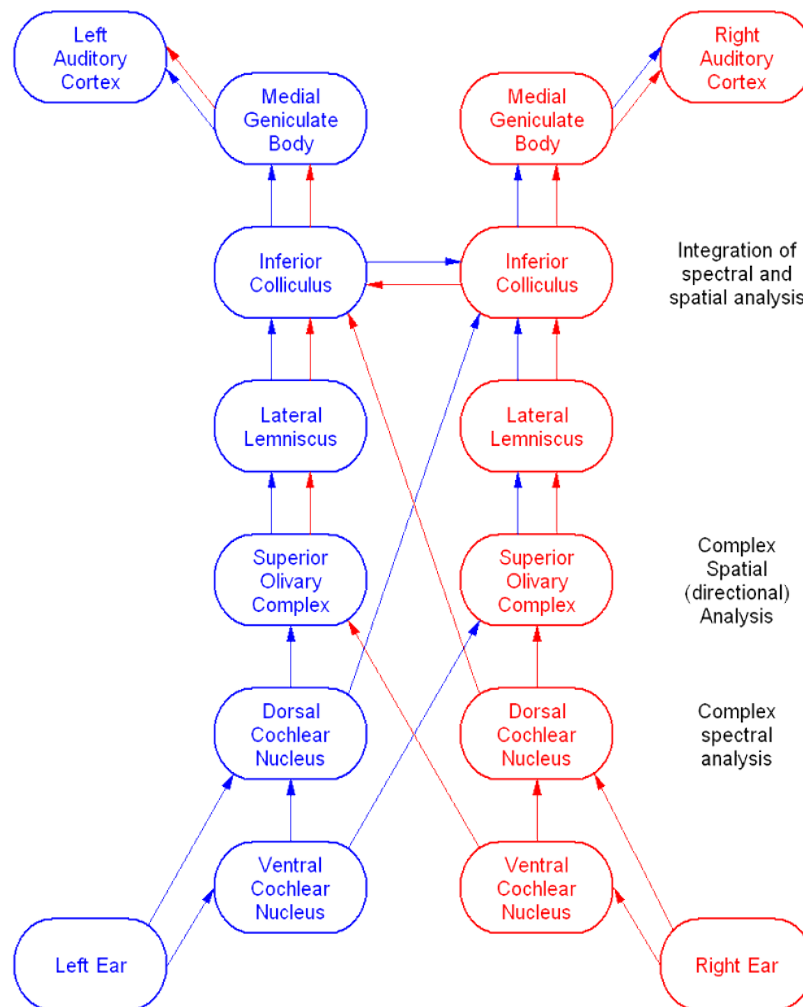


Figure 4: A schematic view of the brain stem, and the nuclei involved in analysis of sound.

medial geniculate body, an area in the midbrain responsible for relaying information from the different senses. A variety of processing is thought to occur here, with neurons sensitive to binaural differences, high- or low-frequency signals, or complex temporal patterns. Axons from the auditory thalamus primarily project to primary auditory cortex, which is the end station of the auditory nerve.

The primary auditory cortex consists of belts of tonotopically organized neurons, mapping the layout of the cochlea, with orthogonal bands sensitive to binaural properties. The detailed workings of the auditory cortex is not well

known, but the area is bordered by regions responsible for the interpretation of speech, among other things.

One final thing to note is that the central auditory pathway is not a one-way system. There are feedback circuits at a number of levels. An important example, mentioned in section 2.1.3, are the efferent nerves projecting from the superior olivary complex to the outer hair cells, who contribute to the spectral tuning of the basilar membrane, in response to input signals.

2.3 Localization of sound

The localization of sound has traditionally been considered as a two-part system, utilizing the time difference of a wavefront arriving at the two ears for low frequencies, and the difference in intensity caused by shadowing from the head for high frequencies. This is known as the duplex theory, and was first introduced by Lord Rayleigh [8] in 1907. Later experiments have mostly confirmed the basics of this model for sources in the horizontal plane, but have also shown that the shape of the pinna (the outer ear) plays an important role, particularly in resolving the elevation of a sound source, or deciding whether a sound originated in the front or the rear, known as a front-back confusion. This section will attempt to give a short overview of these three mechanisms for sound localization. See Blauert [1, Chapter 2] or Moore [9, Chapter 7] for in-depth reviews.

2.3.1 Interaural Time Differences

The interaural time difference, or ITD, is based on measuring the time difference of a wavefront first reaching one ear, then the other. The maximum detectable ITD is about 690 μs , for a sound source 90° off to one side. ITD works best for low frequencies, where the wavelength is larger than the diameter of the head (approximately 22 cm). At periods shorter than about twice the max ITD, or about 725 Hz, ambiguities start to arise, as a wave can be interpreted to come from either 90° extremity. At frequencies above 1500 Hz, ITDs have very limited use, at least for pure tones.

The first neural processing and analysis of ITDs takes place in an area of the brain stem known as the Medial Superior Olive (MSO). For frequencies below approximately 1 kHz, the wavefronts are coded as single phase-locked action potentials in the cochlear nerve.

Afferent axons from both ears meet in the MSO, where they connect to a

group of neurons functioning as a time-delayed coincidence detector, taking advantage of the propagation speed along the axons to allow micro-second accuracy in detecting time differences, as illustrated in figure 5. The neuron where signals from both ears arrive simultaneously is more strongly activated than the rest, and the population as a whole functions as a simple linear space map of the auditory environment.

This theory, known as the Jeffress model, has been well demonstrated in barn owls, but recent research [10] have questioned the validity of this simple model for humans and other mammals. It is assumed that a system with similar functionality exists in mammals, although the implementation may differ somewhat.

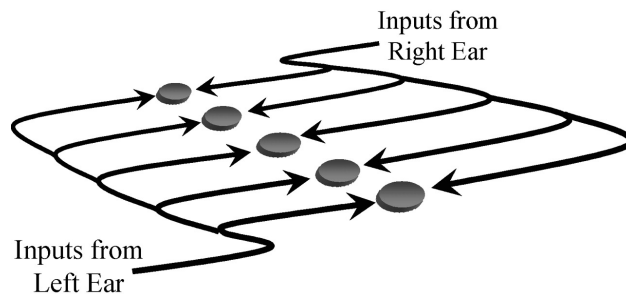


Figure 5: The Jeffress coincidence detection model for interaural time differences

2.3.2 Interaural Level Differences

The interaural intensity or level difference, IID/ILD, is caused by the head itself creating an acoustical shadow on the side furthest from the sound source. This is particularly notable for high frequency sounds, where the size of the head interferes with the wavelength of the sound, generating a difference in intensity of more than 20 dB for the most lateralized cases.

Like the ITD, early ILD neural analysis takes place in the Superior Olive, where action potentials from one ear are passed through a support neuron who 'inverts' the signal; i.e. it is active when no APs arrive at the dendrites, and hyperpolarized if stimulated. The signal from this support neuron is then added to unmodified signals from the other ear. If the sum of these signals is positive, the difference neuron will fire. A number of these difference detector neurons in combination generate another linear auditory space map.

Both the ITD and ILD processing is performed on narrow frequency bands,

and this tonotopic organization continues all the way up to the primary auditory cortex. Findings by Fujiki et. al. [11] also show that the ITD and ILD processing occur at the same time post-stimulus, and primarily in the hemisphere on the opposite side of the stimulus location. This in contrast to monaural spectral cues—covered in the next section—who are processed later, and in different, more specific regions of the cortex.

2.3.3 The Pinna and other factors

The pinna is the third central tool available for sound localization. While the two techniques already presented can solve most ambiguities in the horizontal plane, they are not able to register a difference between two sound sources located in the same angular distance from an imaginary axis passing through the two ears. This is known as the cone of confusion, illustrated in figure 6. Any point on this conic section would give identical results—if only the interaural time and intensity differences caused by the head itself are considered—as the distance to the two ears, and area shadowed by the head, is identical.

The pinna cues are primarily used to detect changes in elevation, as localization in the horizontal plane are less dependent on these cues [12]. They are however also important in clearing up front-back ambiguities in the horizontal plane, which will be of particular interest to the current study.

The neural mechanisms for decoding the pinna cues are not yet known. It is however assumed to be a higher-level process, where the spatial associations are learned based on the direction-dependent filter response of the outer ears, known as head-related transfer functions. One popular theory suggests that the brain compares the complex spectral cues received from the pinnae with a collection of template patterns corresponding to learned locations [13]. It has also been shown that localization performance outside the horizontal plane deteriorates when the stimulus is narrowly bandpass filtered [14], suggesting that the pinna cues have been trained on natural, broadband sounds.

Of particular interest is a study by Hofman and Van Opstal [15], where a group of subjects wore custom-fit molds occluding most of the concha. Over the course of several weeks, their localization performance went from severely disrupted and back to almost normal, indicating that they were able to relearn the pinna cues, without losing their existing ability to locate sounds without the molds. This not only supports the templates theory, but also suggests that the human neural system for localization is highly plastic.

Another interesting study by Langendijk and Bronkhorst [16] used 1-octave wide band-stop filters in combination with HRTFs and broadband noise stimuli. They found that removing the 8-16 kHz band severely increased the number of front-back confusions, and removal of the 6-12 kHz band caused an increase of up-down errors, whereas filtering with only $\frac{1}{2}$ -octave bands did not cause statistically significant interruptions. Their results indicate that a majority of the information needed to solve front-back confusions are located in the 8-16 kHz range.

Other factors that should be mentioned include the familiarity of sounds, and of course vision, which plays a vital role in our daily lives, but whose effects can be controlled in an experimental setting.

Head movements are also very useful in clearing up ambiguities in localization, but not very relevant in this study, as the stimulus length was chosen to be shorter than the time needed to initiate a head turn in response to novel stimuli [17]. Thus, the pinnae and the spectral shape of the sound are the main tools available to the subjects to solve front-back ambiguities.

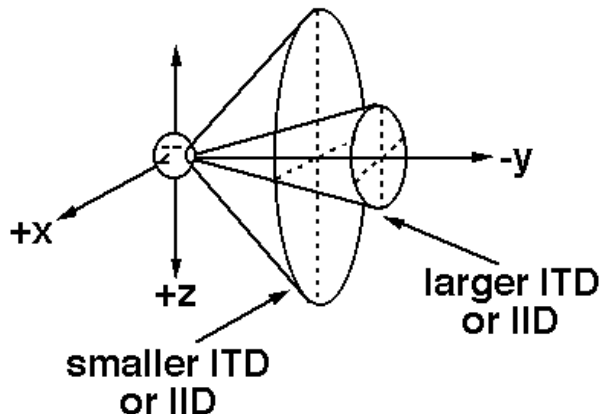


Figure 6: Two cones of confusion, illustrating ambiguous sound source locations

2.4 Similar studies

In addition to the studies already mentioned in the previous section, a study by Abel et. al. [18] has to be included. In their study, a selection of electronic hearing protection systems—including the QuietPro system used in the present study—is tested for localization performance, or more specifi-

cally axis confusions in the horizontal plane. Their setup consisted of eight speakers, positioned pairwise with a 15° offset from the median and interaural axes, and subjects were instructed to detect from which speaker position a signal originated.

Abel found that the QuietPro system in pass-through mode performed better than the active ear muff systems it was compared to, but still with a notable reduction in accuracy compared to normal hearing (71.1% correct responses, compared to 94.1% for the unoccluded condition). See the discussion (section 5.1.3) for further comparisons with the present study.

2.5 Stimuli

As mentioned in the earlier sections, the type of sound stimulus chosen has a significant impact on localization performance. This section examines some of the relevant variables and reasons for the choice of stimuli in the present study.

2.5.1 Spectral contents

One factor is the spectral contents of the stimuli. Wightman and Kistler[19] found that when presenting subjects with conflicting localization cues over head phones, the low-frequency ITD cues were dominant.

A study by middlebrooks [14] using 1/6th octave narrow-band noise centered at 6, 8, 10 and 12 kHz showed that the horizontal localization performance was comparable to a broadband condition, but elevation and front-back reversals increased notably, and varied depending on center frequency. His findings indicated that binaural ILD cues and monaural spectral cues are used independently.

Examples in the literature include various filtered click trains [20], the sound of an M-16 rifle being loaded [21], a telephone ringing signal [22], broadband speech [23], or different variations of high- or low-pass filtered noise.

Overall, broadband (20 Hz – 20 kHz) noise appears to be the stimulus with the best localization accuracy, so it was chosen for this experiment.

2.5.2 Duration

To eliminate the contribution of head motions for resolving ambiguities, the duration of the stimulus must be limited. Makous and Middlebrooks[17] found using head-tracking that in less than 0.3% of their trials, head motion occurred within 150 ms of stimulus onset. Additionally, these registered motions were initiated before stimulus onset, so they concluded that the human neural system needs more than 150 ms to process and respond to novel sound locations.

This is supported by findings by Fujiki et. al. [11] who found that binaural time and intensity cues are processed neurally at around 100–150 ms, and monaural spectral cues later, at 200–250 ms. In addition, acting on these findings and generating a motor response also adds a delay.

Hofman and Van Opstal[24] found that for horizontal location, broadband noise bursts as short as 3 ms gave nearly as good results as a 500 ms control condition, whereas stable localization for elevation needed a burst duration of 80 ms.

Based on these findings, a 150 ms stimulus duration was found to be optimal for this experiment.

2.6 The QuietPro system

The QuietPro Intelligent Hearing System, produced by Nacre AS, Norway, is a combined electronic hearing protection and communications unit. In this study, only the DSP-based hearing protection functionality has been explored; specifically in low-noise conditions, where it is designed to let all sound pass through the plugs as unaltered as possible, yet blocking harmful impulse noises should they occur. Electronic design limitations in the DSP unit does however cause the signal to be low-pass filtered, which limits the total spectrum available to the subject. The cut-off frequency of this filter was observed to be approximately 8 kHz. Additionally, the earplug itself covers most of the concha of the ear when inserted, altering the spectral 'fingerprint' familiar to the brain.

The purpose of this study is to attempt to quantify the effect these limitations on normal hearing have on localization performance in the horizontal plane, and to look for possible causes of reduced localization ability.

3 Methods

The experiment consisted of a sound localization trial, where each subject was asked to identify from which speaker a short noise burst originated from. The subjects completed two trial blocks of 195 trials each; one with normal hearing, and one occluded condition, wearing the QuietPro system. The subjects also underwent a short pure-tone audiometry test and otoscopic examination in advance.

This chapter also describes the lab setup and experiment protocol used, as well as some of the statistical methods employed.

3.1 Subjects

11 subjects were recruited from the students at the university, 7 males and 4 females, between 22 and 29 years old. Each subject participated in a single 1.5-hour session, which included both the examination and the experiment itself. They will be referred to as subjects A to K throughout this report.

Each subject underwent a short otoscopic examination, to ensure that the ear canal was clear and otherwise healthy, in addition to checking the individual shape and size of the ear channel for easier insertion of the earplugs later, as well as selecting the right size of foam plug.

Following the otoscopic examination, a pure-tone audiometry test was run for frequencies between 250Hz and 8kHz, as per the instructions given in [25, Chapter 3.5] This was done to screen for hearing damage—here defined as a pure-tone hearing threshold more than 20 dB above the minimum audibility curve—at either 0.25, 0.5, 1, 2, 3, 4, 6 or 8 kHz. The difference in thresholds between the two ears was also required to be less than 15 dB.

Four subjects (C, D, F and J) exceeded the predefined thresholds (a measured pure-tone hearing threshold of -25 dB at either 6 kHz or 8 kHz). Subject D also had an interaural difference above 15 dB at 6 kHz. All subject were still included in the averaged results, to ensure a broad statistical basis. The audiograms are available in appendix A.

3.2 Lab setup

The listening experiments were performed in the Aura lab at NTNU, which is a test laboratory designed for multimedia and surround sound research. It measures $7.1 \times 5.8 \times 2.7$ meters (length by width by height), and is a semi-reverberant room with some acoustic dampening installed. The subject was placed on a chair in the center of the room, surrounded by an array of loudspeakers.

The loudspeaker setup consisted of 23 individually numbered, two-way active speakers (Dynaudio Acoustics BM6A) placed in a circle with a 2.00 m radius around the subject, with 15 degrees angular distance between each speaker. The subject was positioned so that their head was in the very center of the circle, over a mark on the floor, and with the speakers at ear level. See figure 7 for an illustration.

Speaker position 15 was left empty, as one of the original 24 speakers belonging to the lab could not be located for this experiment, and this position would be one with comparably less impact on the experiment, as the spatial resolution in this direction is poor.

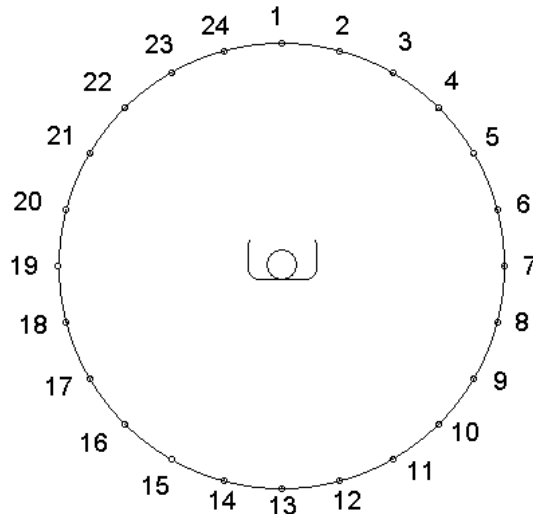


Figure 7: The physical setup of the lab, with numbering and positioning of the speakers. The radius of the circle is 2.00 m.

The speakers were tuned towards equal frequency response, using the 0–4 dB tuning knobs for each of the two channels on the speakers. A $\frac{1}{3}$ -octave frequency band audiometer with a Brüel & Kjær condenser microphone (Type

4165) for measurements was used to measure the speakers' frequency responses, which did reveal some minor per-speaker differences, particularly above 10 kHz, likely caused from different previous uses. These differences were however deemed to be within acceptable limits for the experiment.

The audiometer was first calibrated with a Brüel & Kjær 1 kHz sound level calibrator (Type 4231). The average background noise in the listening room was measured to be 29.1 dBA, with no peaks above 30dB in the 20Hz-20kHz spectrum. The full spectrum can be seen in appendix E.

The speakers were addressed individually using a 24 channel Hammerfall sound card with a software-based equalizer and patching matrix. The signals were then sent via ADAT to three separate RME 8-channel DA converters, and then via a 24-channel analog patching matrix to the separate speakers. See figure 8 for an illustration.

The subject responses were collected using a 17 inch LG touch screen displaying a MatLab user interface with 24 numbered buttons representing the speakers, and a confirm/next trial button in the center, as illustrated in figure 9. The touchscreen was placed either in the subject's lap or on a chair slightly to the right in front of them, depending on individual preferences.

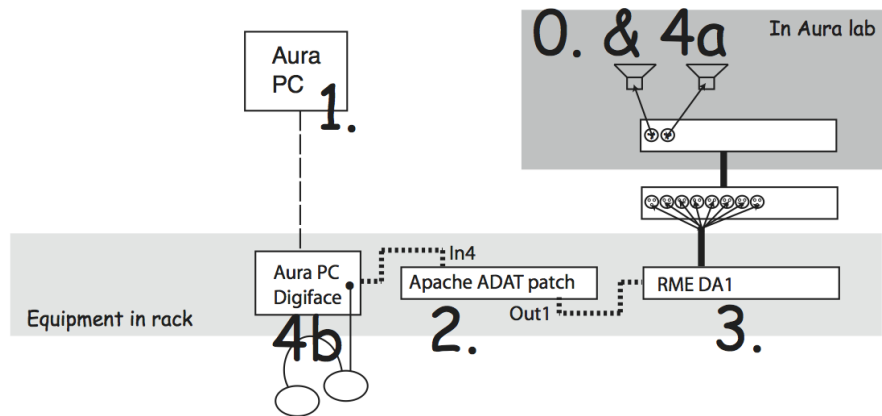


Figure 8: Sketch of the audio path in the Aura lab, for two speakers. It was extended to 24 by adding two more RME 8-channel DA converters in parallel, and connecting the rest of the speakers.

The QuietPro unit used in the experiment was running software versions PIC v.1.38.3938, and ASIC v.2.33/0232. It was equipped with new disposable Lithium batteries before the experiments began. Disposable earplugs of sizes small and medium were used, depending on subject preference and optimum fit.

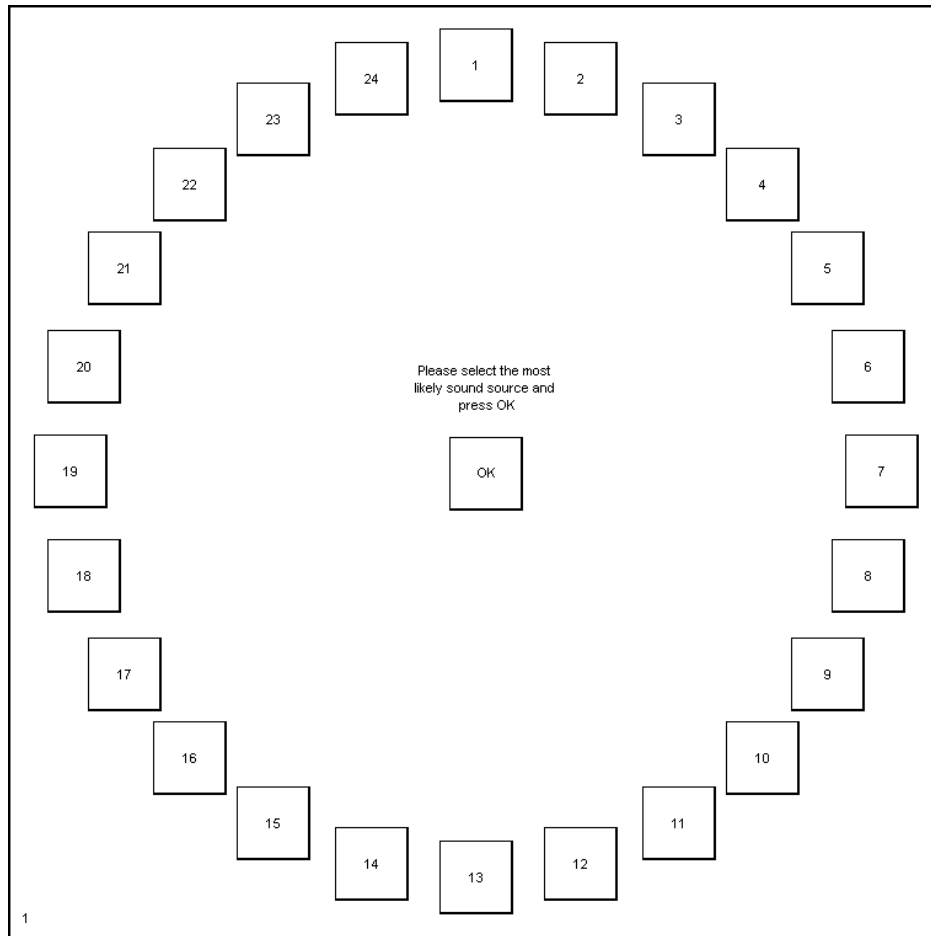


Figure 9: The graphical user interface used in the experiment. After selection, the text changed to "Speaker [X] selected. Press OK to confirm". The number in the lower left corner is a trial counter.

3.3 Experimental protocol

The stimulus used in the experiment was 150 ms bursts of white noise generated in MatLab, and low-pass filtered at 20 kHz to get a smooth (5ms) onset curve. The resulting spectrum of the stimulus recorded at the position of the subject's head (2.00 m from the speakers) was primarily shaped by the system response. The SNR was >10 dB for the [50 Hz – 20 kHz] interval, and >25 dB for the [250 Hz – 16 kHz] interval. See appendix E for detailed data.

The stimuli was then pseudo-randomly presented to one of the 24 speakers,

with an equal number of trials per speaker. Each trial block consisted of 195 trials, giving an average of 8.5 trials per speaker per trial block. The sound level at the listeners position—2.00 m from the speakers—was measured to 66 dB SPL A-weighted.

The subjects responded via a touch-screen display located next to the chair. The response application consisted of 24 numbered buttons arranged in a circle (figure 9), representing the individual speakers, as well as an OK button in the center, which confirmed the selection and triggered the next sound burst. Thus, the subjects controlled the speed of the experiment, taking the time they found necessary to decide on a source location before triggering playback of the next stimulus.

6 of the subjects ran the test with the occluded condition first and normal hearing afterwards, and the other 5 with the normal hearing condition first.

All subjects were instructed to focus their gaze at a visual target marked on speaker 1—straight in front of them, at 0° azimuth and elevation—between each trial, to ensure a neutral initial position. The subjects were however allowed to turn their head between trials, to check the number on a given speaker outside their field of vision, or otherwise clear up uncertainties.

The instructions were not written down, so small differences between subjects may have occurred. Explanations of how to operate the response panel, the procedure and estimated duration of the test, and what the subject was supposed to do, was included each time.

For the trial block with the occluded condition, the subjects were first instructed how to mount the earplugs, then allowed to insert them by themselves. The mounting was then controlled by the researcher, before switching on the system itself, which initiated the automatic acoustic leakage tests, to confirm a proper fit. The system’s sound amplification was reset to default (zero gain), to be comparable to normal hearing.

The source speaker, the subject’s guess, and subject response time was stored for each trial, and each trial block was saved as a timestamped MATLAB data file for later analysis.

3.4 Data analysis

To see if there was a statistically significant difference between the two hearing conditions for a given direction in the horizontal plane, a two-tailed t -test was used. However, since the variance in answers when wearing the ear plugs

appeared to be larger than in the unoccluded condition, Welch's t -test was used, which is an adaption of the Student's t -test for two samples with potentially unequal variances [26]:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}} \quad (1)$$

where \bar{X}_i , s_i^2 and N_i are the i^{th} sample mean, sample variance and sample size, respectively.

The degrees of freedom are calculated as follows:

$$\nu = \frac{\left(\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}\right)^2}{\frac{s_1^4}{N_1^2 \cdot \nu_1} + \frac{s_2^4}{N_2^2 \cdot \nu_2}} = \frac{\left(\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}\right)^2}{\frac{s_1^4}{N_1^2 \cdot (N_1-1)} + \frac{s_2^4}{N_2^2 \cdot (N_2-1)}} \quad (2)$$

It's common to use p-values to calculate the probability of the null hypothesis, i.e. that the two sampled populations have the same mean. The alternative hypothesis used here is that the means are different. This is known as a two-tailed t -test. The p-values are found from the Student- t cumulative distribution function,

$$p = 2 \cdot tcdf(x|\nu) = \int_{-\infty}^x \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} \frac{1}{\sqrt{\nu\pi}} \frac{1}{\left(\frac{1+t^2}{\nu}\right)^{\frac{\nu+1}{2}}} dt \quad (3)$$

where $x = -|t|$ for a given direction/speaker, ν are degrees of freedom, and $\Gamma(\cdot)$ is the gamma function

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt. \quad (4)$$

4 Analysis and Results

This chapter will give an overview of the main findings in this experiment. At first some individual results will be presented, to illustrate the large differences between subjects. Next, the average localization performance for the subjects in the normal hearing condition is presented, followed by results for the occluded condition, and testing for statistically significant differences. Unless otherwise specified, a 95% confidence interval ($\alpha = 0.05$) will be used where applicable.

4.1 Preliminary analysis

The collected data consisted of source-response pairs for each trial, grouped by subject and test condition (normal | occluded). A total of 11 subjects participated in a two-trial session each, with an average of 195 responses per condition.

Before any statistical analysis was performed, all the collected data was controlled manually. This was done to remove outliers caused by interruptions during the experiment, which occurred for technical reasons with a few subjects. These interruptions were easily recognized by a response time of >15 seconds, rather than the 1-3 seconds average, in addition to a random subject response.

The test results from each trial block was pairwise sorted by source speaker, before calculating the difference between actual source position and subject response in degrees. When calculating the error in localization accuracy per speaker, only the absolute distance was considered, and so the maximum possible error is 180° . Differences larger than this, caused by wraparound errors due to indexing—e.g. the error of responding 15° to a stimuli originating at 345° would give a 330° difference—were corrected to $newdiff = 360 - olddiff$. The resolution of the data at this point was 15° , which is the minimum possible error distance for a single trial pair.

This difference measure, representing the basic error in localization, was the basis for further analysis.

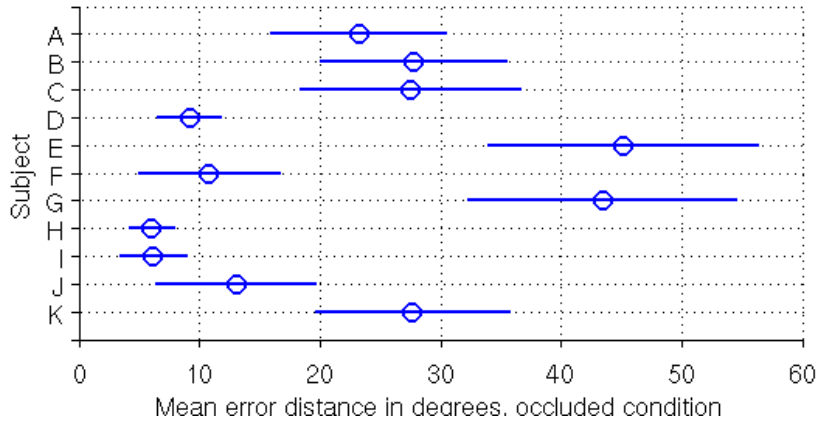


Figure 10: Mean error distance per subject, averaged across all speaker positions. The width of the bars indicate the variance. Note the division into three rough groups. Subjects DFHIJ (left) will be referred to as group 1, ABCK (center) as group 2, and EG (right) as group 3.

4.2 Individual results

The first main observation—notable already when running the experiments—was the large inter-subject variance in the occluded condition, whereas the results in the normal hearing condition were quite similar across subjects.

By averaging each subject’s accuracy across all speaker positions, this inter-subject variance is illustrated (figure 10). The subjects as a whole may be grouped into three separate groups—who are significantly different from each other according to a basic ANOVA test, but that assumes equal variance—a low-error set, denoted as group 1, and consisting of subjects D, F, H, I and J; group 2, consisting of subjects A, B, C and K, with a larger spread of errors; and finally group 3, with subjects E and G. This is not a rigid classification, but quite useful to observe the different types of errors made.

In group 1, there is little significant difference between the normal hearing and occluded conditions, except for some back-to-front reversals in the median plane (a source at -180° perceived as 0°). All group members except subject J ran the experiment with the normal hearing condition first.

Group 2 is more varied. There are reversals in both directions, both in and near the median plane. The overall spread in responses, and thus variance, is also larger than for the other groups. All group members, except subject C, ran the experiment with the occluded condition first.

Group 3 represents the subjects who experienced a complete back-to-front reversal, placing all sources in the front hemisphere. Their localization error is however low, although mirrored across the interaural axis. Both subjects ran the experiment with the occluded condition first.

The scatter plots in figures 11, 12 and 13 illustrate the spread of responses for group 1, 2 and 3, respectively. See appendix B for an explanation of scatter plots, as well as separate plots for each subject.

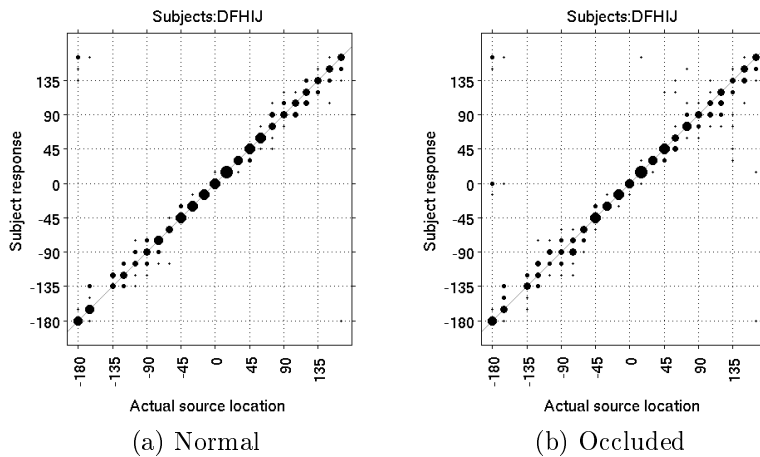


Figure 11: Scatter plots for group 1, subjects [DFHIJ]. This subset had little difference between the normal and occluded conditions, and low variance overall.

Looking at the audiometric results, the subjects with hearing thresholds registered below the predefined 20 dB limit did not perform any worse than those with normal hearing. In fact, three of the four (D, F and J) were classified in group 1, with a low error spread. It should be noted that this is merely an observation by the author, and has not been analyzed statistically. It did however permit the inclusion of all subjects for further averaging and grouped analysis, which is covered in the following section.

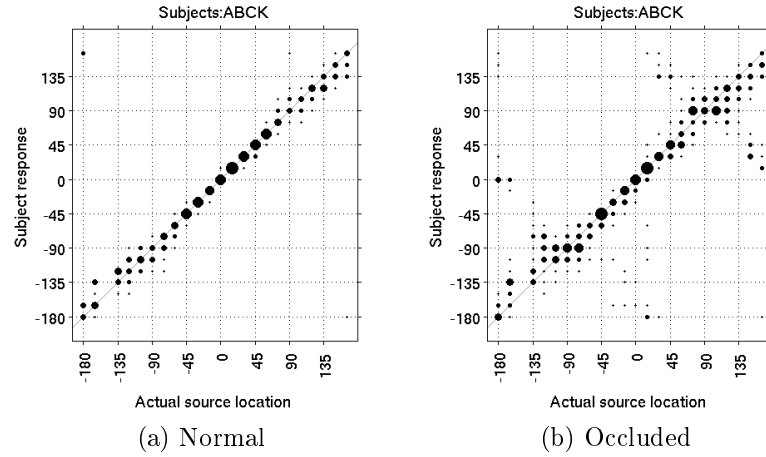


Figure 12: Scatter plots for group 2, subjects [ABCK]. The subset had some reversals, particularly close to the median plane, and a larger variance in the occluded condition.

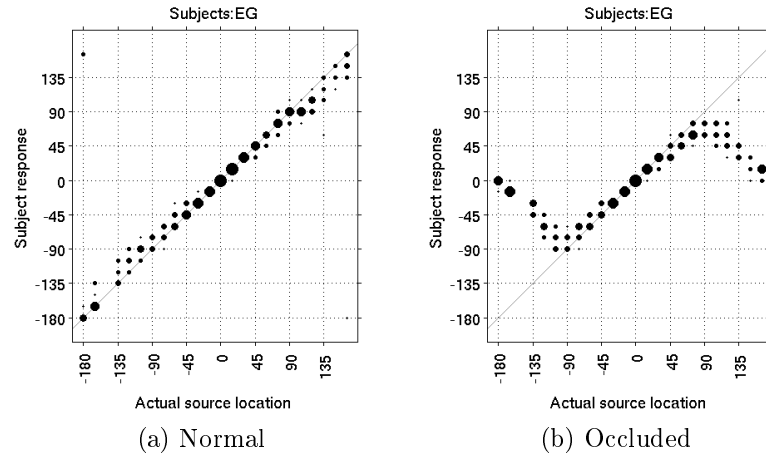


Figure 13: Scatter plots for group 3, subjects [EG]. The subset had complete back-front reversals, but accurate (low variance) localization, even for the mirrored sources.

4.3 Averaged results

To be able to compare the results in this study to previous findings, a common base level is needed. By averaging the results of all subjects in the normal hearing condition, a baseline performance of the lab setup can be inferred, assuming that the subjects are representative of the general population.

Figure 14 shows the averaged localization error for the normal hearing condition for each speaker position. The resolution directly in front is approximately $1^\circ \pm 3^\circ$, with accuracy falling towards the sides and rear, where it reaches $10^\circ \pm 10^\circ$, with somewhat better results at -180° again. It should also be noted that there was little inter-subject variance in the normal hearing condition. No errors larger than 75° were registered in the normal hearing condition.

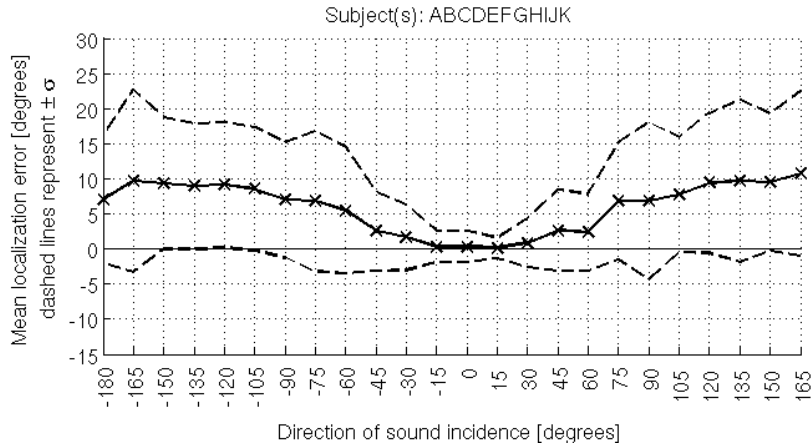


Figure 14: The mean error in localization for the normal hearing condition, averaged across all subjects. The dashed lines are one standard deviation. The data for direction -150° azimuth is interpolated from the two nearest neighbours for readability, due to a missing speaker.

The averaged localization error for the occluded condition is plotted in figure 15. Notable features include the clear increase in errors for directions $> \pm 90^\circ$, due to the source reversals a number of the subjects experienced. This is also the cause of the large variance for the positions close to 0° . Towards the sides the variance is quite low, as any front-back reversal here is indistinguishable from general inaccuracy. See appendix D for individual plots.

The central hypothesis for this experiment was to see if there was any statistically significant difference in localization performance between the normal hearing and occluded conditions, for a given direction of sound incidence.

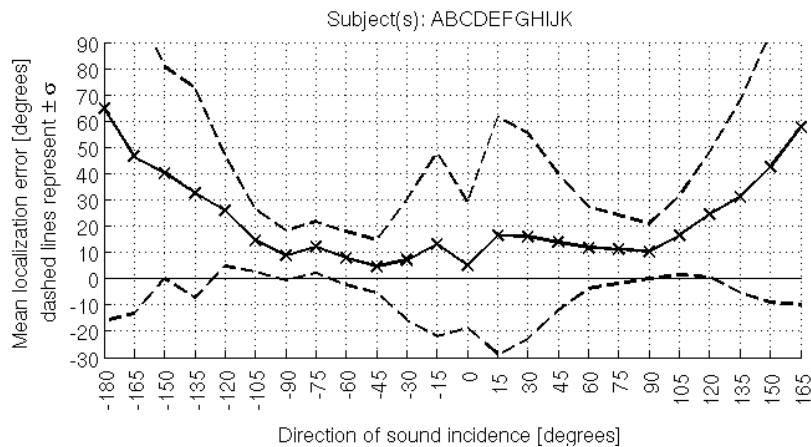


Figure 15: The mean error in localization for the occluded condition, averaged across all subjects. The dashed lines are one standard deviance. The data for direction -150° azimuth is interpolated from the two nearest neighbours for readability, due to a missing speaker. Note the larger scale on the y axis compared to figure 14, for increased readability.

This hypothesis was tested with Welch’s t -test (see section 3.4), with the resulting p-values in table 1 showing that for all but four directions, there is a significant increase ($p < 0.05$) in localization error for the occluded condition. See also figure 16 for a 95% confidence interval plot of the difference means, where the increase in errors at positions above $\pm 90^\circ$ caused by the back-front confusions is notable.

The least significant difference is observed close to the interaural axis. This is because the accuracy in both conditions was equally poor, and no large source reversals add to the variance, as no large left–right confusions were observed.

Direction	p-value	Direction	p-value
-180°	0.0000000...	0°	0.0827284...
-165°	0.0000000...	15°	0.0002789...
-150°	no data	30°	0.0006302...
-135°	0.0000005...	45°	0.0000336...
-120°	0.0000000...	60°	0.0000001...
-105°	0.0002840...	75°	0.0057814...
-90°	0.1913189...	90°	0.0273683...
-75°	0.0002827...	105°	0.0000016...
-60°	0.1338226...	120°	0.0000000...
-45°	0.0576777...	135°	0.0000012...
-30°	0.0319871...	150°	0.0000000...
-15°	0.0005954...	165°	0.0000000...

Table 1: Examining the difference in localization accuracy per speaker between the two test conditions (occluded | normal hearing), averaged across all subjects. The p-values indicate the probability of the two conditions having the same mean, and is significant ($p < 0.05$) for all but four (-90° , -60° , -45° and 0°) of the speaker positions, all located in the frontal hemisphere. See also figure 16 for a 95% confidence interval plot.

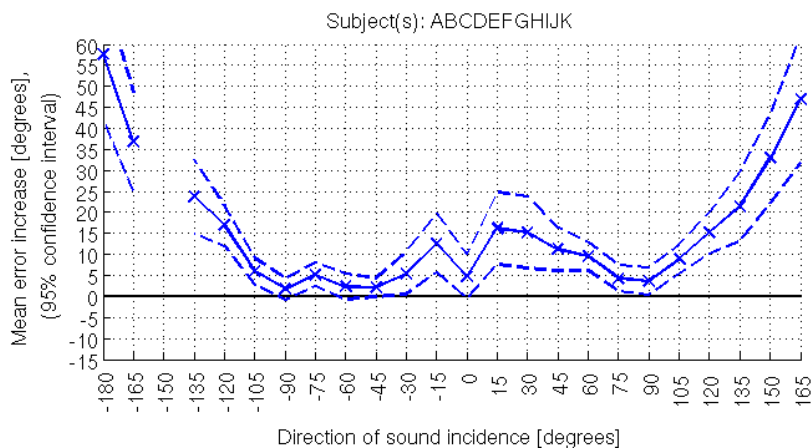


Figure 16: Difference between the sample means of the occluded and normal conditions averaged across all subjects, with a 95% confidence interval, from the t -test. The errors near 0° are caused by partial front-to-back reversals in some subjects, and the large errors in the rear hemisphere (towards the sides) is mainly caused by the complete back-front reversal in subjects E and G, although other subjects were also less accurate in this region.

5 Discussion

The primary purpose of this chapter is to examine and attempt to explain the results of the experiment, comparing with previous findings in the literature. Design errors uncovered in the experiment are described and their consequences analyzed. Finally, suggestions for further studies are given, based upon the findings in this current study.

5.1 Localization errors

5.1.1 General accuracy

The measurements of averaged localization accuracy for the normal hearing condition (see figure 14) are similar to previous findings reported in the literature [1, 9, 17]. This indicates that the audio lab setup and the general design of the experiment is valid, and that the other results in this study may be compared to previous experiments in the literature.

The reasons for both inter-subject differences and the increased uncertainties in the occluded condition are numerous. The rest of this section will attempt to highlight a few important factors.

One contributing factor that is notoriously hard to quantify is the concentration level of the subject; whether they are stressed, tired, bored or focused is likely to have a significant influence on their individual performance. Also, running just a single session with each subject does not give the opportunity to eliminate day-to-day differences.

The novel experience of wearing electronic pass-through earplugs for the first time may have been a contributing factor to the slightly increased inaccuracies in the occluded condition (not counting the source reversals). The subjects also reported a low, buzzing background noise when the system was on, and set to zero gain (default), but did not find it distracting when presented with the sound stimuli.

Visual input is also a factor to be considered. The subjects were sitting in a well-lit room, and could see all speakers in the frontal hemisphere from the default position, and the rest by turning their head. It was not possible to visually identify which speaker was active during a given stimulus, but the subjects were still aware of the limited number of potential sound sources. How much of a contribution this had is not known, but the effect is assumed

to be the same in both conditions, and as such it would be eliminated by the difference tests.

Overall, the result for both conditions appear to be quite similar in accuracy, with the notable exception of the source reversals in the occluded condition, which is the central cause of the large variance in the averaged data. The next section will examine these specifically.

5.1.2 Source reversals

The most interesting type of errors in this experiment were the source reversals in the occluded condition. All subjects except D and H experienced at least some source reversals in the median plane, yet no such reversals occurred in the normal hearing condition. The subjects in group 2 had more severe source reversals, mirroring locations in a wider range, either partially or completely.

The most extreme, and unexpected, example was observed with subjects E and G, who both perceived every single sound stimuli as originating in the frontal hemisphere, yet locating them quite accurately in (mirrored) position. After the experiment, both subjects reported that they had been "waiting for stimuli from behind", or "wondering if there was some trick", indicating that they had correctly understood the instructions given.

One common factor for most of the subjects (A, B, E, G and K) with severe reversals was that they all ran the experiment with the occluded condition first. This is likely to have been one contributing factor, as errors may have been caused by unfamiliarity to the lab setup as a whole.

An exception to this rule in one end was subject C, who tested the normal hearing condition first, yet had several source reversals. This was however also the first subject tested, so interruptions and unclear instructions might have contributed to the inaccuracies.

The other exception was subject J, who ran the experiment with the occluded condition first—just before subjects E and G, in fact—yet experienced very few source reversals, apart from in the median plane. This shows that the complete reversal observed in group 3 was not caused by any technical errors in the lab on one test day.

5.1.3 Comparison to other studies

An unusual feature of the source reversals is the clear frontal preference of the QuietPro system, with the frequency of back-to-front reversals dominating over the front-to-back reversals, and frontal localization generally being quite accurate. This is supported by the findings by Abel et. al. [18], who also found that the QuietPro system had a frontal preference, as opposed to the other active and passive earmuffs they tested, who made more errors in the front hemisphere than in the rear.

Their setup was however somewhat simpler than in the present study, primarily designed to identify source reversals across the median and interaural axes, and measuring the percentage of correct localizations for each unit tested. To compare data, a correct localization in the present study was defined to be a 15° error or less, as the minimum distance between two speakers in Abel et. al. is 30° . Averaged across speakers and subjects, we get 95.0% correct in the normal hearing condition, and 78.6% in the occluded condition, compared to 94.1% and 71.1%, respectively, in [18].

Since the directly comparable results of these two studies are this similar, it is also interesting to note their results for the passive Peltor H10A earmuffs (46.1%), and the Racal Slimgard II ear muff with talk-through functionality (69.2%).

5.2 Causes of localization errors

Based on these findings, and what is known about the human ears' ability to accurately localize sound, the observed reduction is likely caused by two main factors; the physical occlusion of the concha, and subsequent altering of the spectral signature; and the reduced available bandwidth, caused by limitations in the signal processing unit.

Gardner and Gardner [27] found that sound localization in the median plane decreased with increasing occlusion of the pinnae, causing a larger number of source reversals. A more recent study by Hofman and Van Opstal [15] showed that while occluding the concha with custom molds caused large errors in localization, continued use of the molds reduced this error, returning almost to unoccluded levels after 4–6 weeks.

King and Oldfield [28] found, using personalized HRTFs, that in order to get accurate sound localization in both azimuth and elevation, a broadband signal up to at least 13 kHz is needed, as front-back reversals was noted in a

12 kHz low-pass signal. The experiment was only run on 3 subjects, but the findings are still interesting. Similarly, Langendijk and Bronkhorst [16] found that the 8-16 kHz band is vital to resolve front-back reversals. The same was noted by Best et. al. [23] for speech stimuli, but they also found that even a weak contribution from this band (40 dB attenuation) provided a benefit for localization. This is comparable to the passive noise reduction reported by Nacre (42 dB at 8 kHz) for the QuietPro system, which means that users do receive some high-frequent information, but strongly attenuated.

Brungart et. al. [29] offers an engineering angle, comparing the trade-offs between microphone positions on a custom-made completely-in-the-canal (CIC) earplug—which is optimal for conserving the spectral reflections of the pinnae—and the limited bandwidth available for processing due to size restrictions. They too found bandwidth to be the dominating factor, and that microphone position did not have much of an effect when the signal was limited to 6 kHz. Tests with a commercial CIC hearing aid, with a reported cutoff at 7.2 kHz, gave somewhat better results.

To sum up, the errors caused by occlusion of the pinna may potentially be compensated for by extensive training and use of the system. The limited bandwidth of the system appears to be a larger problem, yet it would appear that the limited attenuation provided by the ear plugs may be a small benefit to localization, compared to studies using low-pass filtered stimuli. Increasing the available bandwidth is however an engineering challenge and a trade-off between cost, size, and complexity.

5.3 Sources of error in current experiment design

Due to the limited scope of this study, extended pilot tests were not completed before collecting the research data. As such, some issues with the experiment design were discovered during the data collection. Some could—or had to—be corrected as they were discovered; whereas others were left untouched, as they would either skew the collected data if fixed, or were simply not amendable. This section attempts to estimate the consequences of these design errors.

The most obvious error was the one missing speaker. Rearranging the 23 existing ones to get equal spacing was not considered, as this would disrupt the symmetry of the other positions. Instead, a known low-resolution angle (-150° azimuth) was chosen to be left empty, while keeping the spatial density near the median and inter-aural axes high. This was of course not

optimal, and should be amended in any future studies.

Another variable that was not quantified was the reflection times and acoustical properties of the room. Asymmetries here might generate a bias in the recorded data. However, the differences in frequency response between the individual speakers may be as large a contribution to an eventual bias. Ideally, the experiment should be conducted in an anechoic chamber to eliminate the contribution of reflections altogether, although it might be claimed that the naturalness of a reverberant room is useful for realism. In any case, the central question in this study was to look for a difference between the occluded and normal hearing conditions, and the contributions from both room reflections and speaker differences is the same across these conditions.

On the software side, a memory leak in the external library used to address the individual speakers (`pawavplaya.dll`) caused an overflow after 195-197 repeated calls. 195 trials was however deemed enough trials per condition for further statistical analysis, so attempting to debug and recompile the external library was avoided. One could also have split the trial block into two shorter recording sessions if more data was required.

The result of the aforementioned memory leak was simply that Matlab had to be restarted between trial blocks. This had the unintended side effect of initializing the internal random generator to a default seed, which scrambled the order of the speakers the same way for each subject. This bug was discovered and corrected after a few subjects, but caused a slight overall bias in the number of trials per speaker, which was originally intended to be balanced across positions.

5.4 Suggestions for further studies

The findings in this study have also raised a number of new, interesting questions, which could be explored in future experiments.

5.4.1 Spherical localization

One obvious expansion of the present paradigm would be to add a second dimension, to see how localization above and below the horizontal plane is affected by wearing electronic earplugs. The literature offers several alternative lab designs for spherical localization of sounds.

The first alternative, located at an US Air Force test facility [29, 30], is a

rigid geodesic sphere placed in an anechoic chamber, with speakers placed at every node, spaced 15° apart.

A more minimalist alternative, employed by several research groups [24, 31], is a single speaker mounted on a dual-jointed arm, allowing the positioning of a sound source anywhere on an imagined sphere around the subject. This was also done in an anechoic chamber for both groups.

A third alternative used by [17] should also be mentioned, which consisted of a semicircular hoop with an array of speakers arranged along the length, which could be rotated around the subject.

Common to all these designs is the challenge of simply and accurately collecting the subject responses. Techniques range from head- or eye-tracking to touch-sensitive orbs and spoken coordinate responses. Finding the right solution for a future study would depend on the hardware setup chosen, the experiment design, as well as the tools available to the researchers.

5.4.2 Training

Another very interesting expansion of the current study would be to explore the effect of training and adaptation on sound localization.

As shown in [15], the brain can relearn monaural localization cues from pinnae filled with putty, given 4-6 weeks of training. This would be one way to explore the relative contributions of physical shape and limited bandwidth towards reduced localization performance for the QuietPro system. This assumes that these are the two main factors causing the observed reduction in accuracy from normal hearing, particularly the front-back reversals. After a period of training, the errors caused by the altered shape of the pinna should be minimized, leaving only the limited bandwidth as a central factor.

Performing this experiment in the horizontal plane is possible with the existing lab setup, it's only a matter of recruiting subjects and defining a long-term test schedule. It would also be interesting to see the effect of training on elevation errors, but this would require a more complex lab setup, as described in section 5.4.1.

It would also have been interesting to just run the original experiment with the original subjects once more, to see how large a factor the newness of the lab setup was. The results did show that the subject who tested with the occluded condition first made larger errors than the ones who first ran the experiment with unoccluded, normal hearing. This could also be considered

a form of training, but focused on the experimental setup only, rather than long-term adaptation to new spectral ear signatures.

6 Conclusions

The purpose of this study was to explore and quantify the effects the Quiet-Pro system has on sound localization accuracy, compared to normal, unoccluded hearing.

The results show a small but significant increase in error rate for nearly all directions tested in the horizontal plane, in addition to a notable increase in source reversals for the majority of the subjects tested. The amount of source reversals varied highly between subjects, ranging from zero to a complete back-to-front reversal of a full trial block.

The likely causes of these source reversals are a loss of high-frequency sound information, due to both a physical occlusion of the ear, and electronic low-pass filtering of the sound.

This study has only examined sound localization accuracy in the horizontal plane. Other studies have shown that the same high-frequency spectral information necessary to resolve source reversals, is needed to accurately detect the elevation of a sound source. This indicates that the sound localization error for the QuietPro system may also be significant outside the horizontal plane, but this was not possible to test with the currently available lab equipment.

It should however be mentioned that while the increase in errors compared to normal hearing is notable, the QuietPro system has been shown to perform as good as or better than other comparable hearing protections in earlier studies.

This study does not consider the effect of short- or long-term training, as none of the subjects were familiar with sound localization tests or the QuietPro system before the experiments. This is likely to be a significant factor, which should be explored in more detail.

A Audiograms

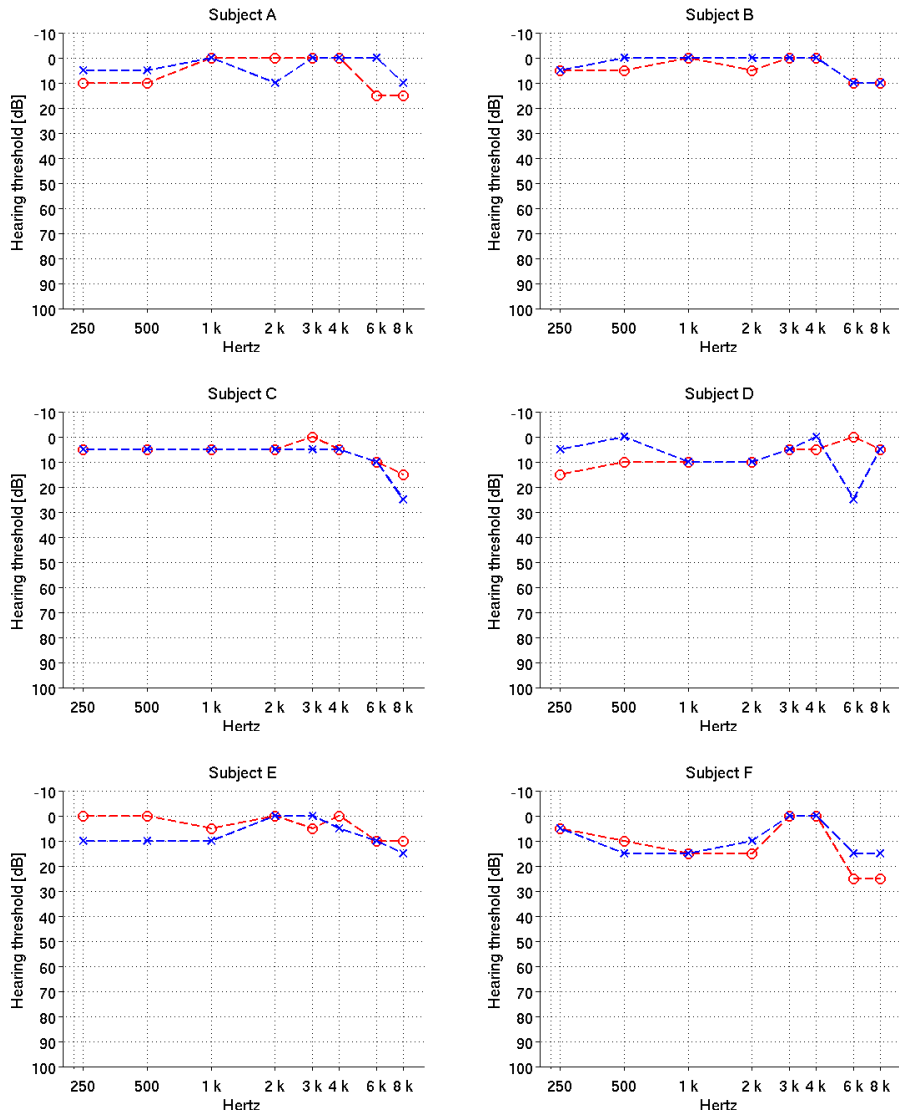


Figure 17: Pure-tone audiograms for subjects A-F. The right ear is represented by red \circ symbols, the left by blue \times symbols.

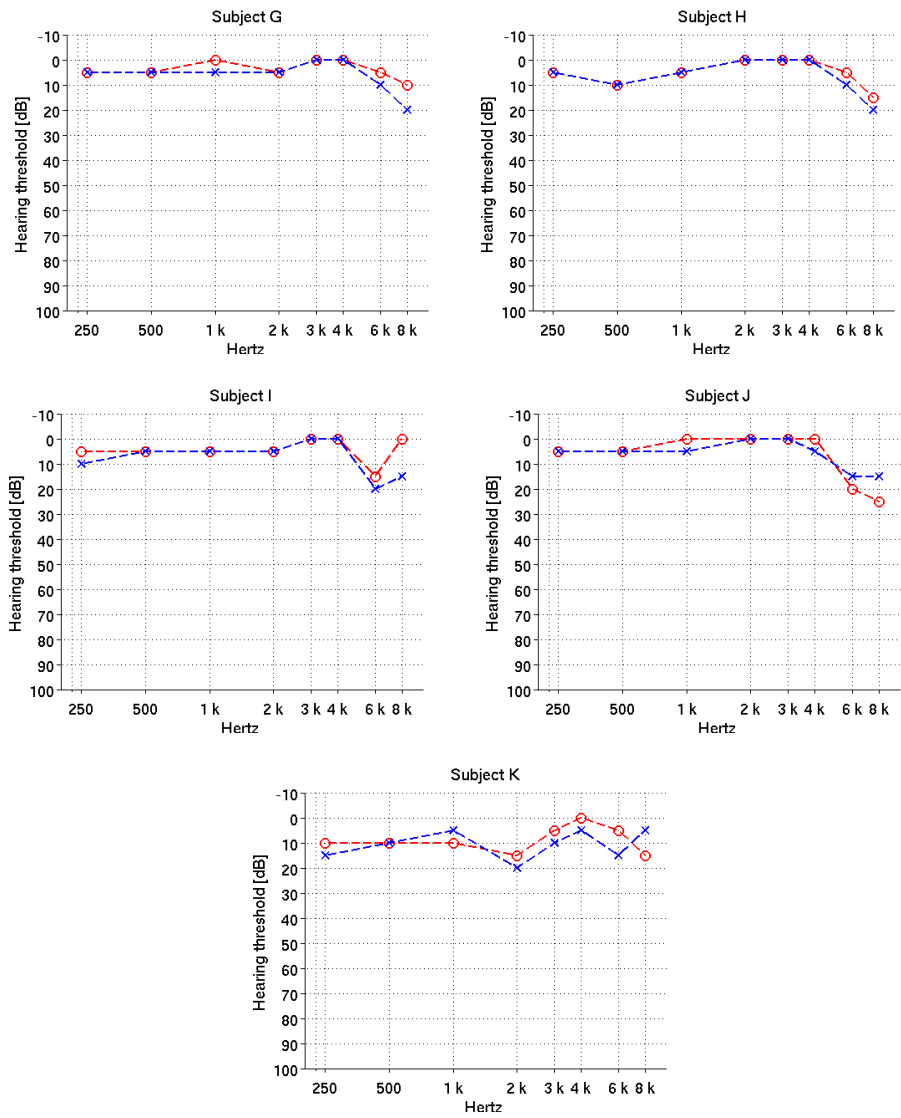
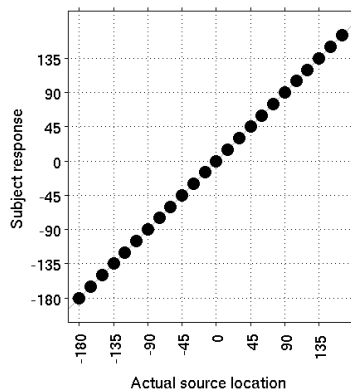


Figure 17: Pure-tone audiograms for subjects G-K. The right ear is represented by red \circ symbols, the left by blue \times symbols.

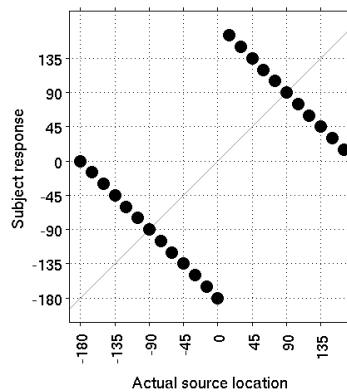
B Subject-specific scatter plots

The scatter plots display every response given by a subject (or several). The diameter of the dots represent the number of responses, where larger dots are more frequent. Each column represents the spread in responses to one specific source position.

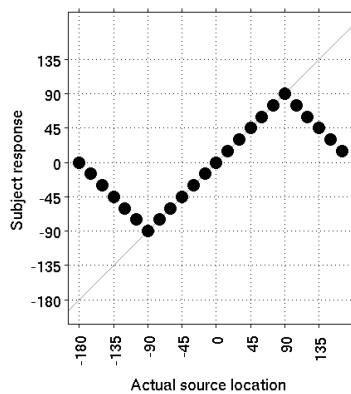
Figure 18 illustrate some stereotyped patterns that reflect responses given by the subjects. The actual plots on the following pages have similar patterns, and combinations thereof, in addition to the scaled dots representing response frequencies.



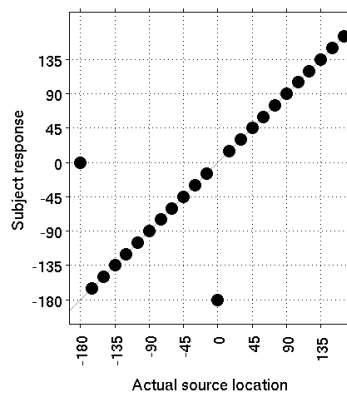
(a) "Correct" responses



(b) Complete front-back reversal



(c) All sources perceived in the front



(d) Median plane reversal only

Figure 18: Examples of scatter plots with generic data, illustrating four typical patterns. Actual recordings may have components of several patterns, as well as a natural variance.

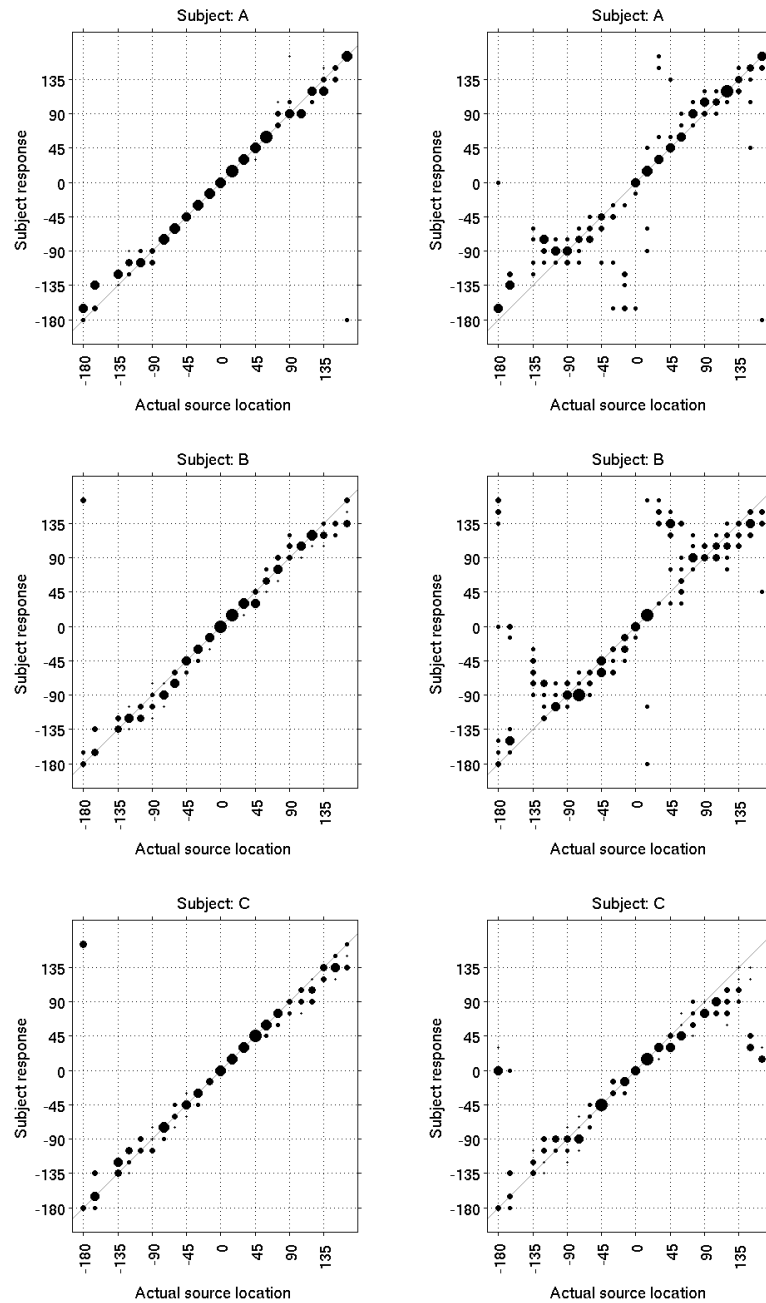


Figure 19: Scatter plots for subjects A-C. Normal hearing in the left column, occluded on the right.

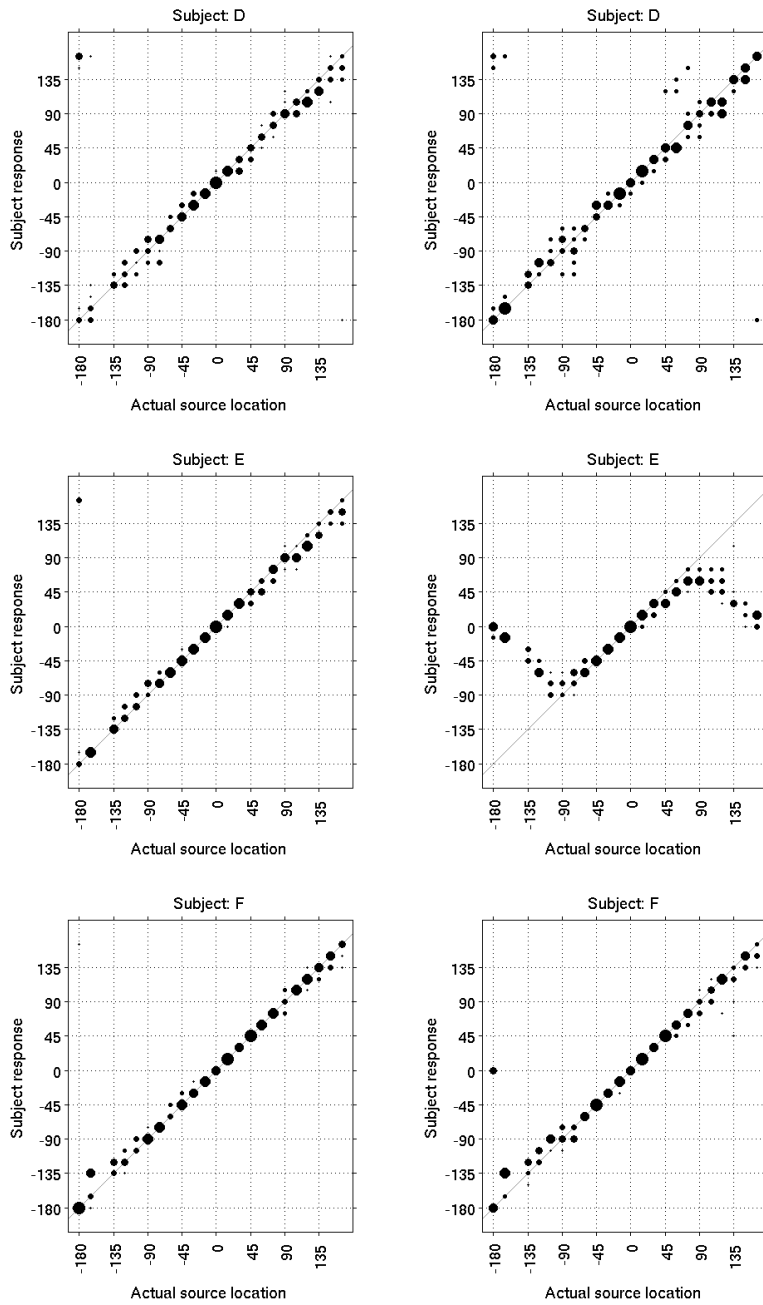


Figure 19: Scatter plots for subjects D-F. Normal hearing in the left column, occluded on the right.

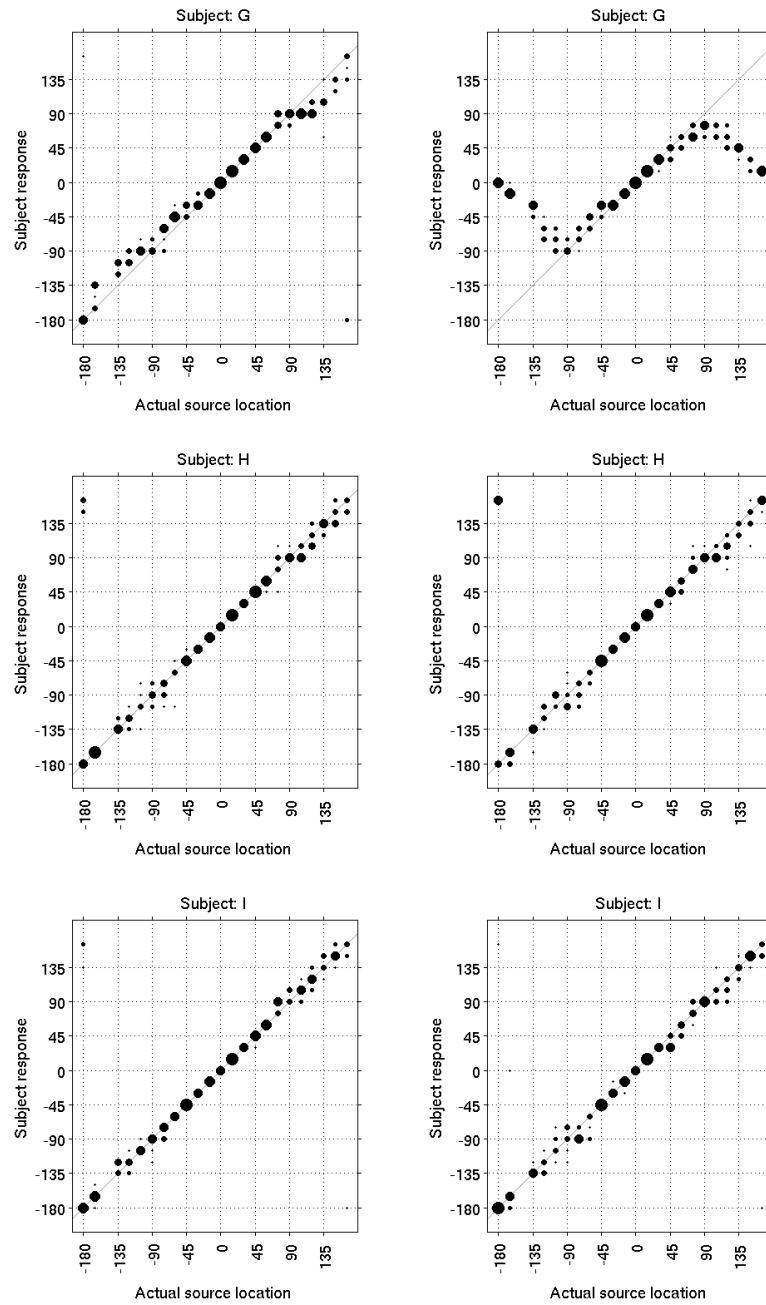


Figure 19: Scatter plots for subjects G-I. Normal hearing in the left column, occluded on the right.

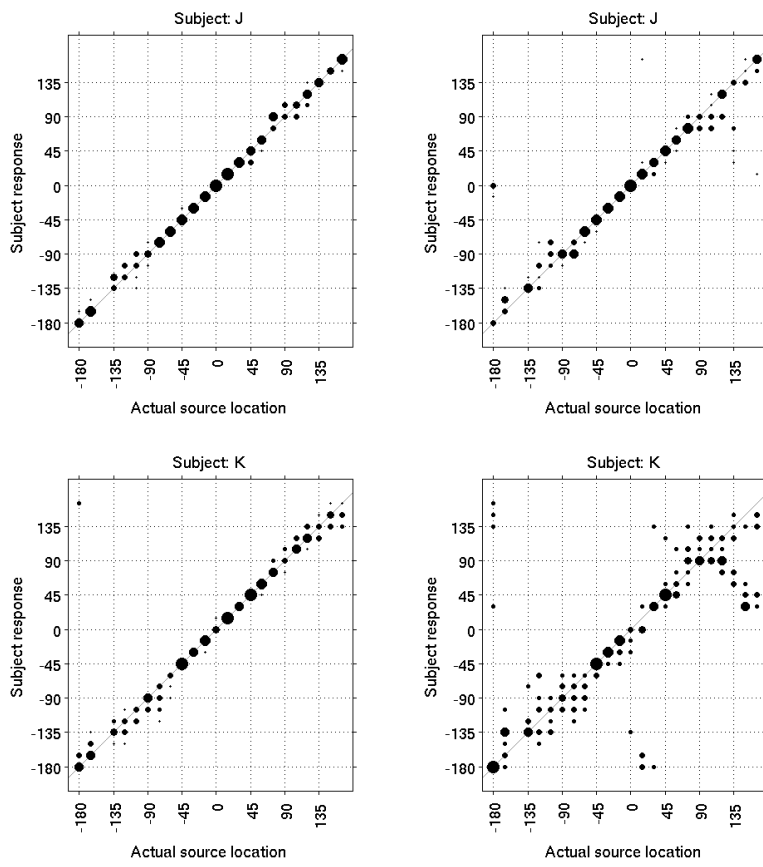
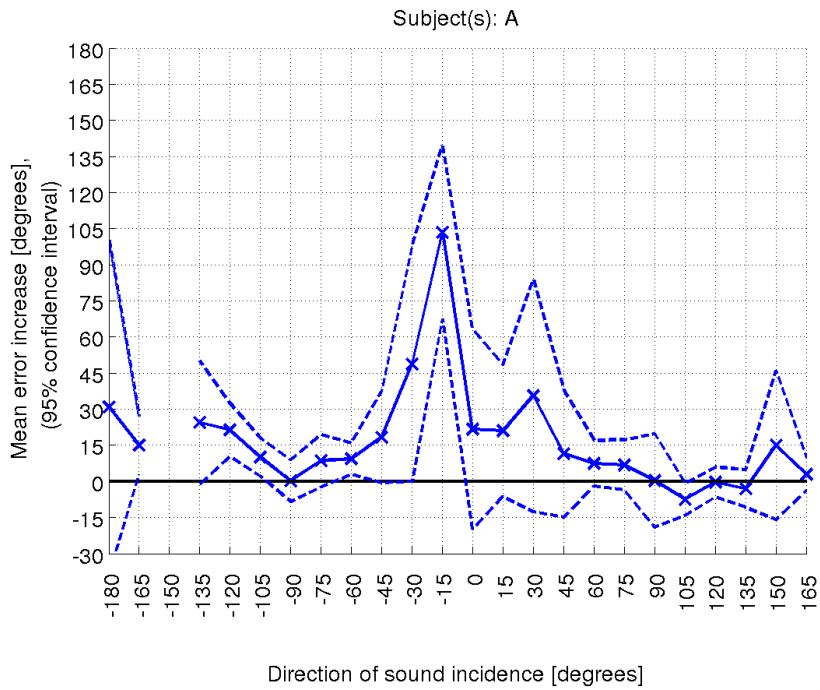


Figure 19: Scatter plots for subjects J-K. Normal hearing in the left column, occluded on the right.

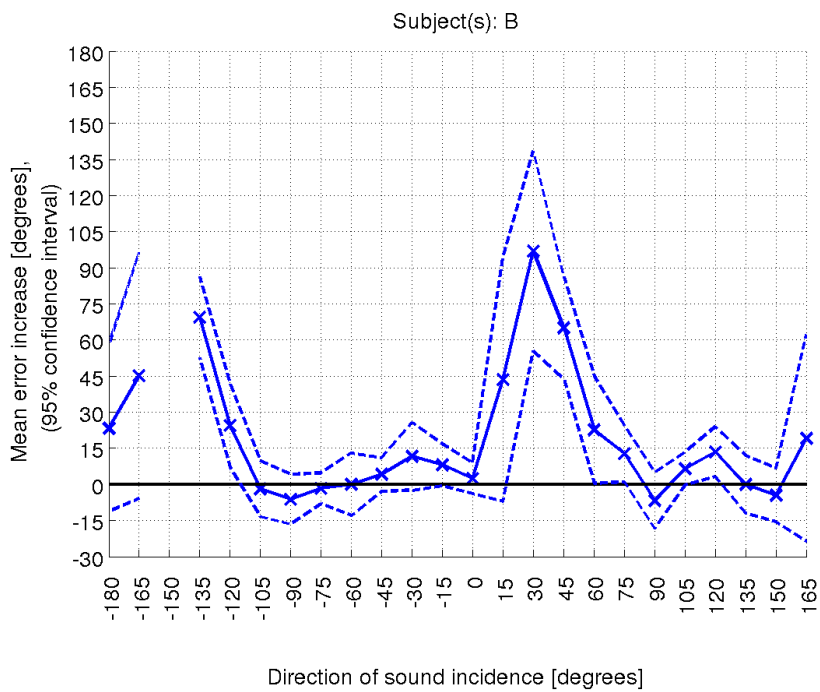
C Subject-specific *t*-test results

These graphs show the increase in mean localization error for the occluded condition compared to the normal hearing condition. The tests were run separately for each speaker position, and is plotted with a 95% confidence interval to indicate the variance in responses.

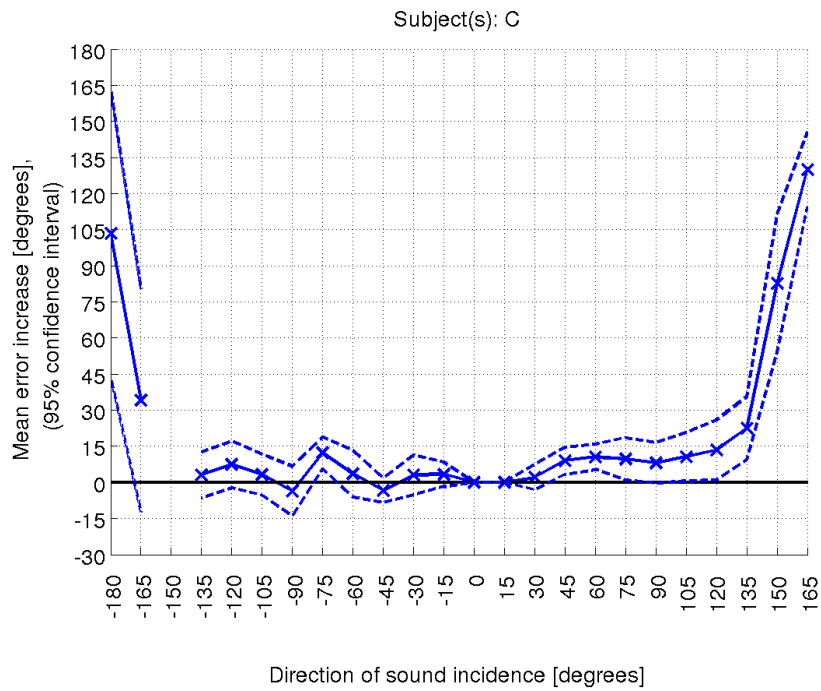
See appendix D for the variance in data per condition.



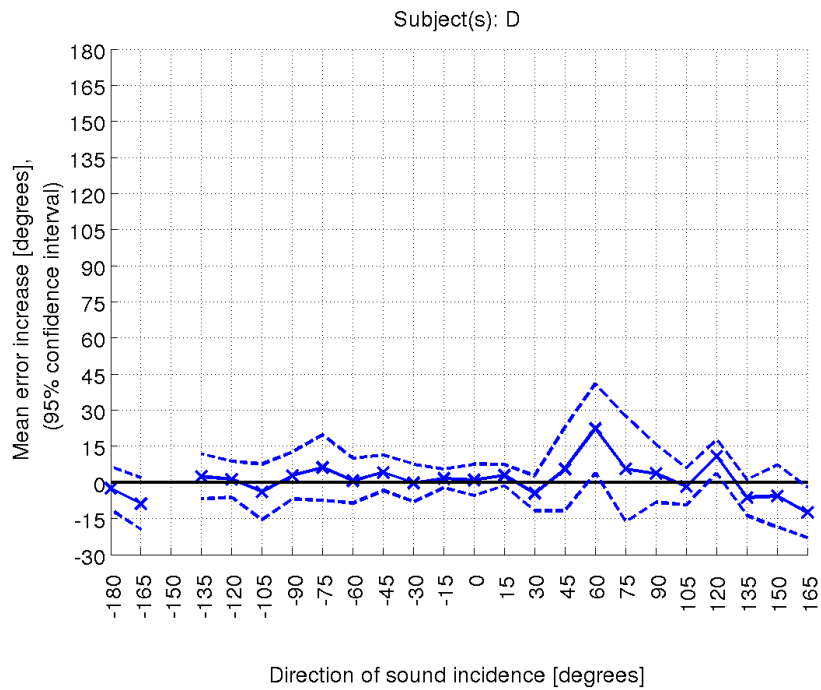
(a) This subject shows a partial front-to-back reversal on the left side.



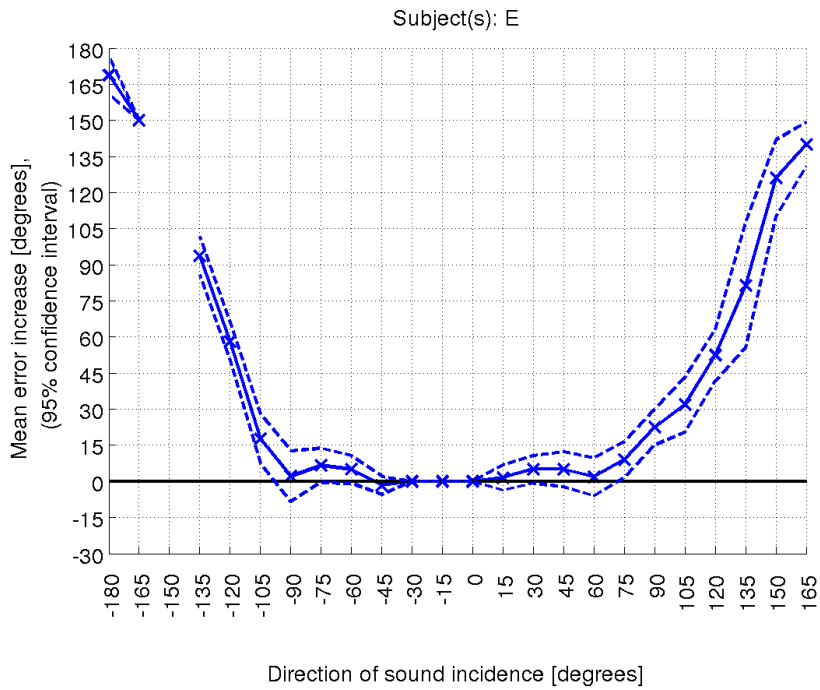
(b) This subject shows a partial front-to-back reversal on the right side.



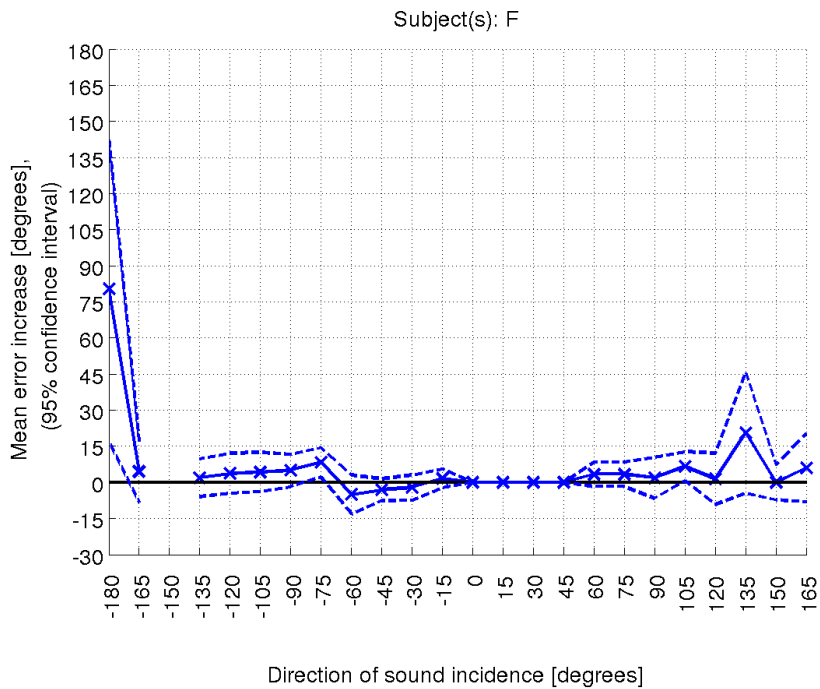
(c) This subject shows a back-to-front reversal close to the median plane.



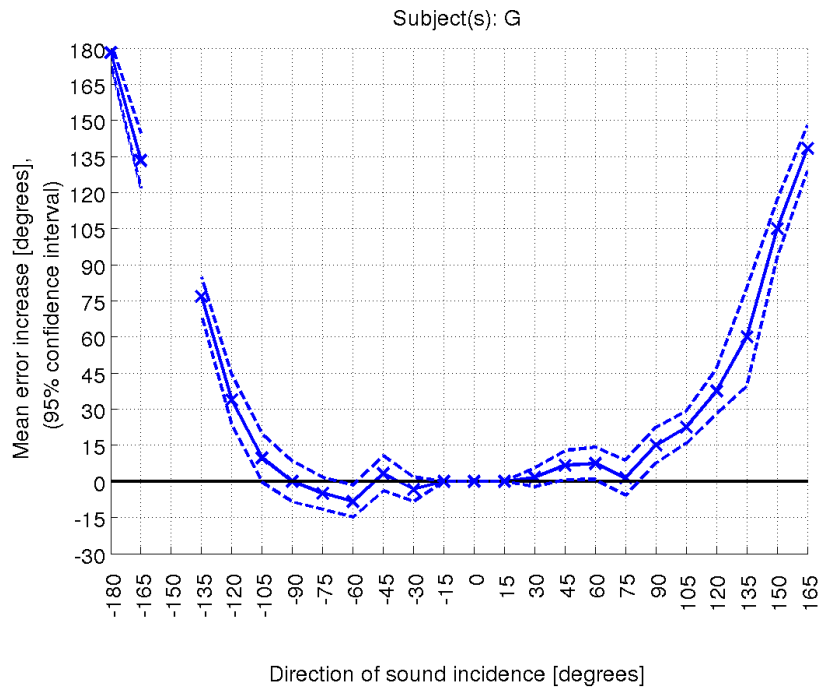
(d) This subject had similar performance in both conditions. Only barely significant reductions at 60° and 120°, and an actual improvement at 165°.



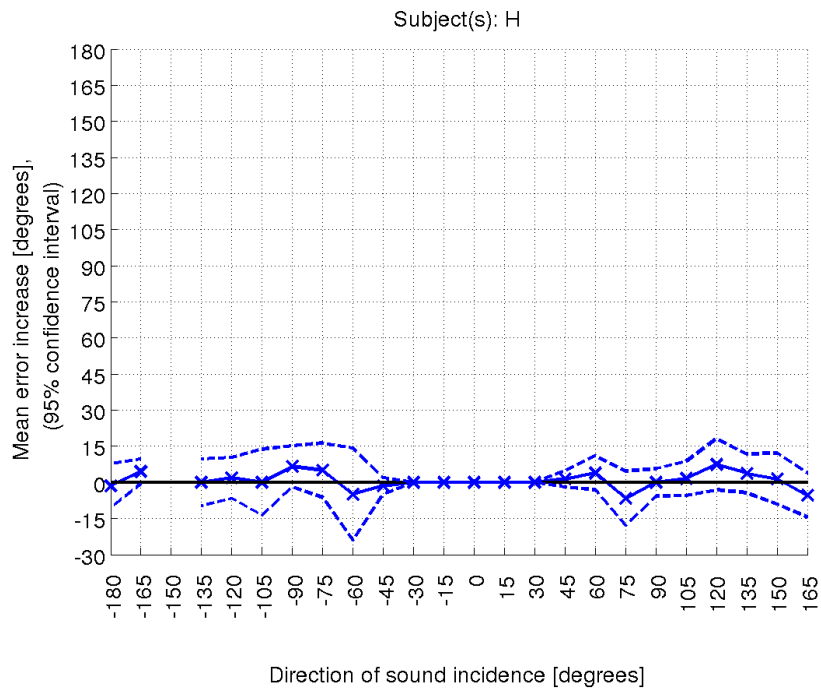
(e) This subject experienced a 100% back-to-front reversal, interpreting all stimuli as originating in the frontal hemisphere. Note the low variance, also for the reversed stimuli.



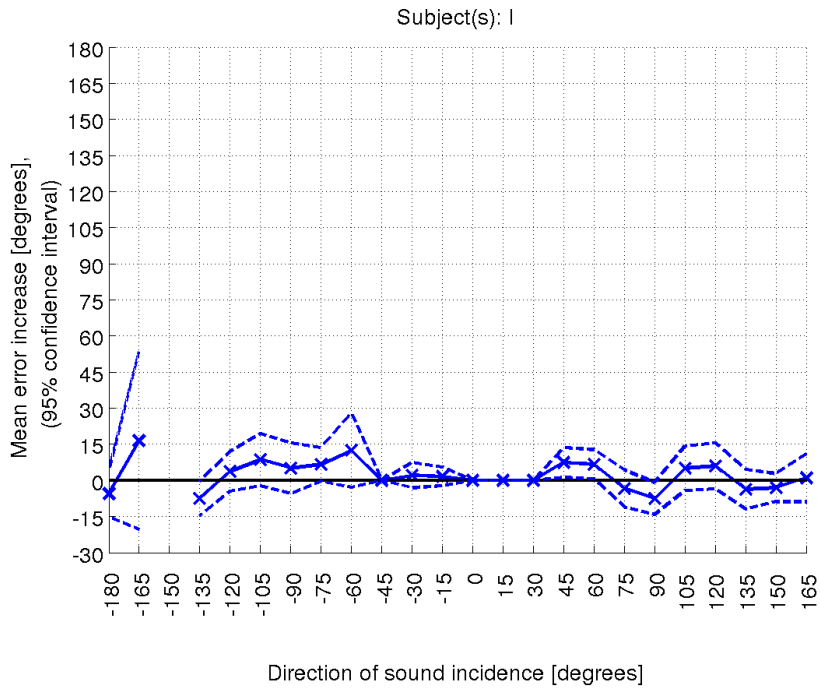
(f) This subject had back-to-front reversals in the median plane only.



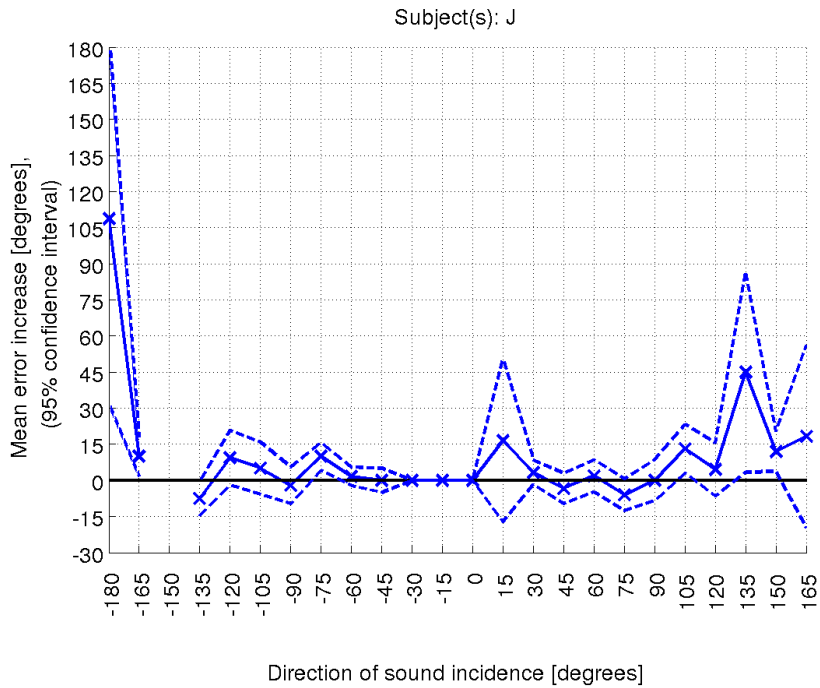
(g) This subject experienced a 100% back-to-front reversal, interpreting all stimuli as originating in the frontal hemisphere. Note the low variance, also for the reversed stimuli.



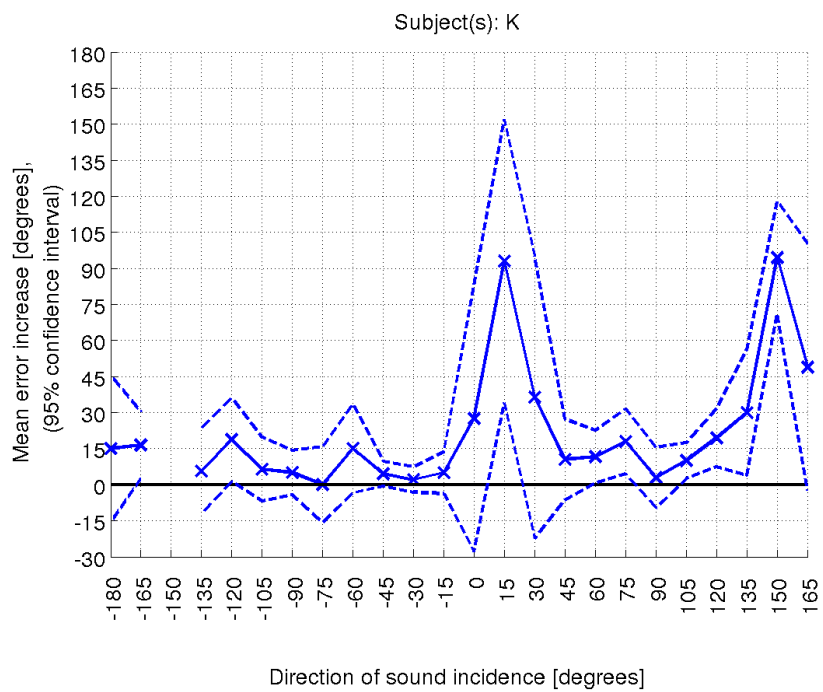
(h) This subject had similar performance in the occluded and normal hearing conditions, no significant differences.



(i) This subject had similar performance in the occluded and normal hearing conditions.



(j) This subject had back-to-front reversals in the median plane only.

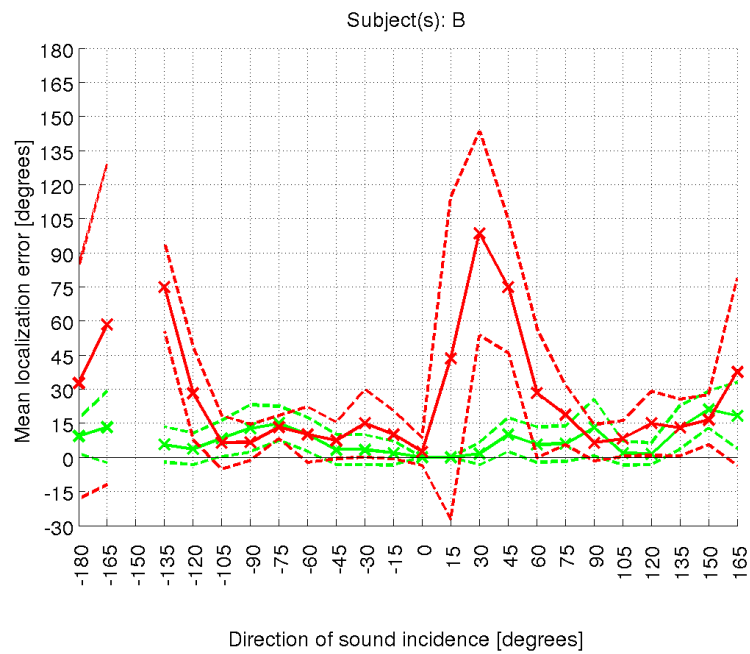
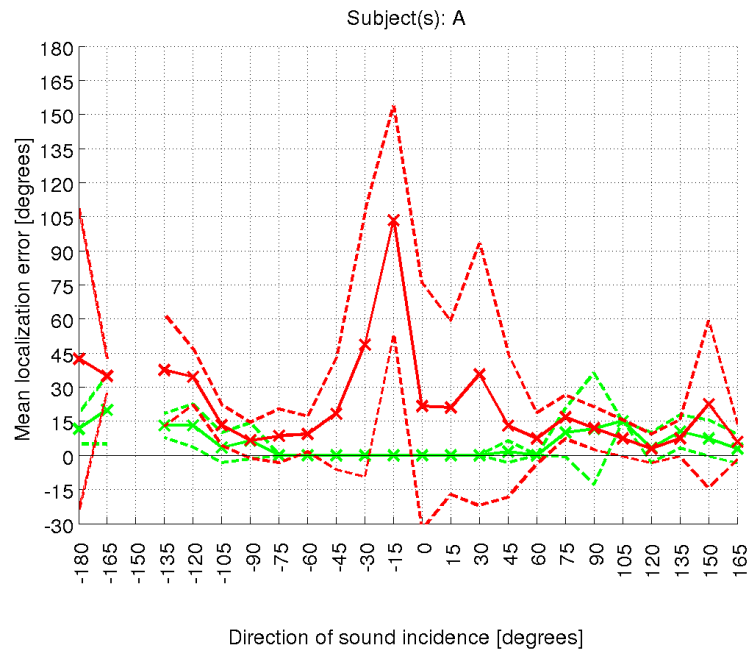


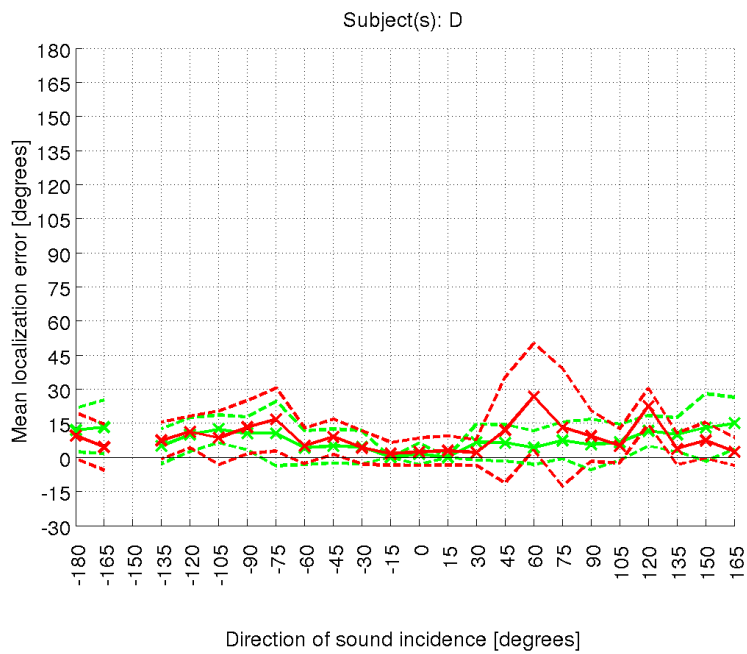
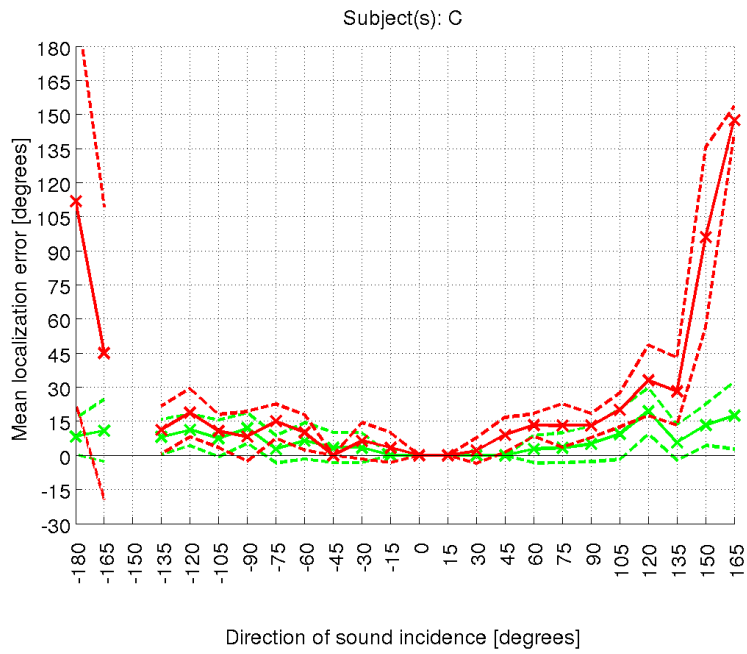
(k) This subject had several off-center reversals, and a large variance in general.

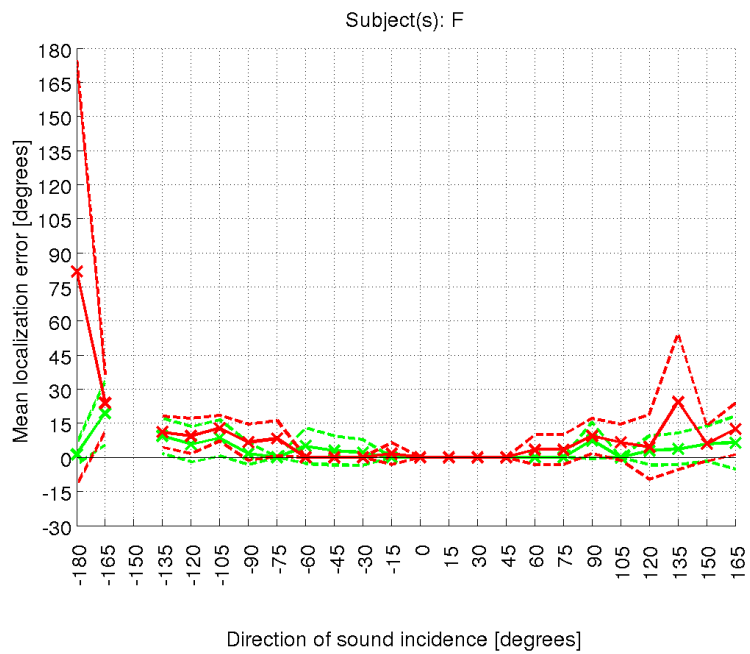
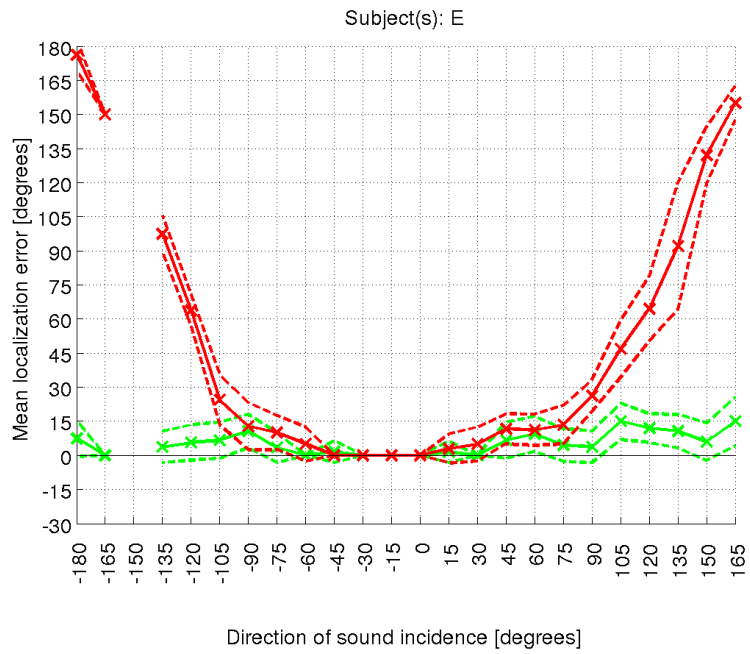
D Subject-specific averaged results

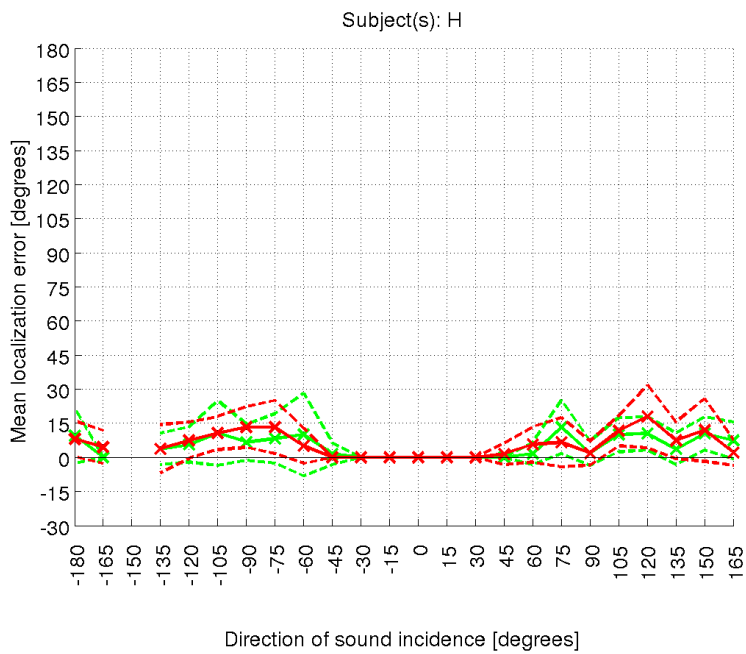
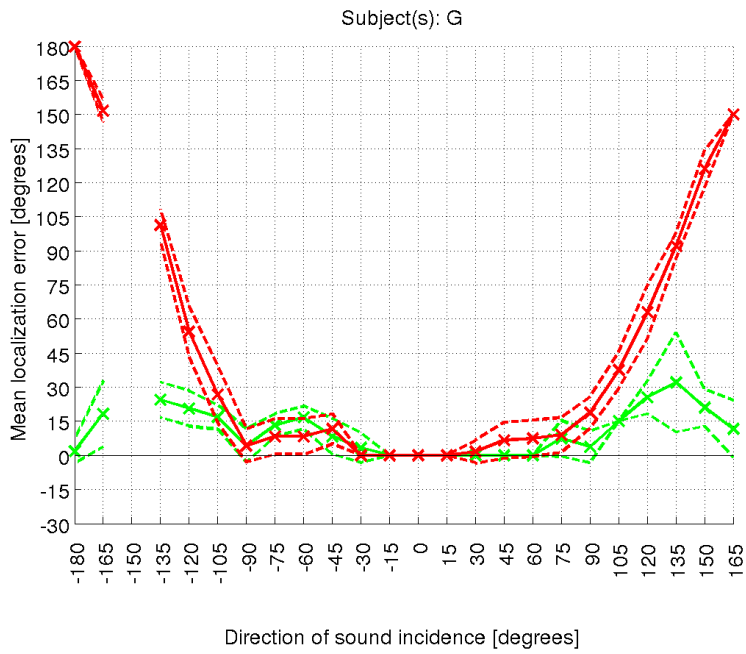
All plots show the per-speaker means for the normal(green|gray) and occluded(red|black) conditions. The corresponding dotted lines represent one standard deviation.

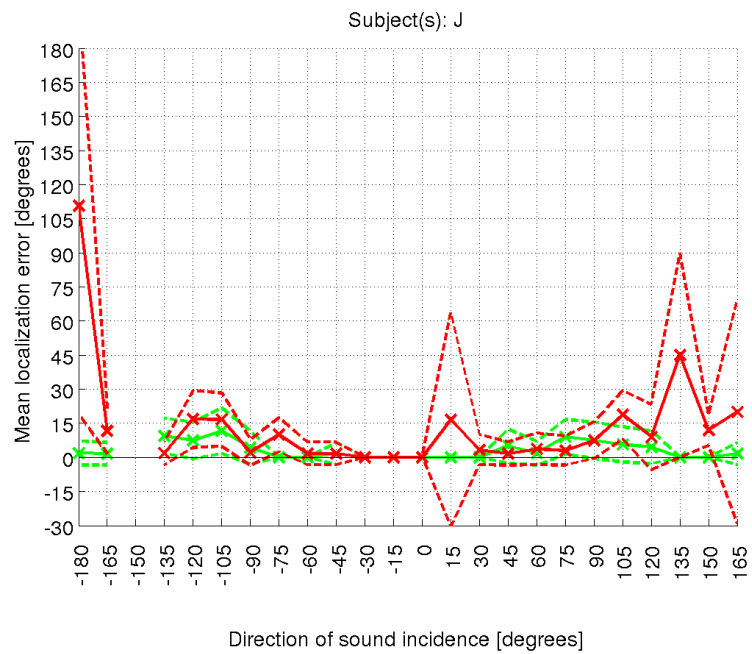
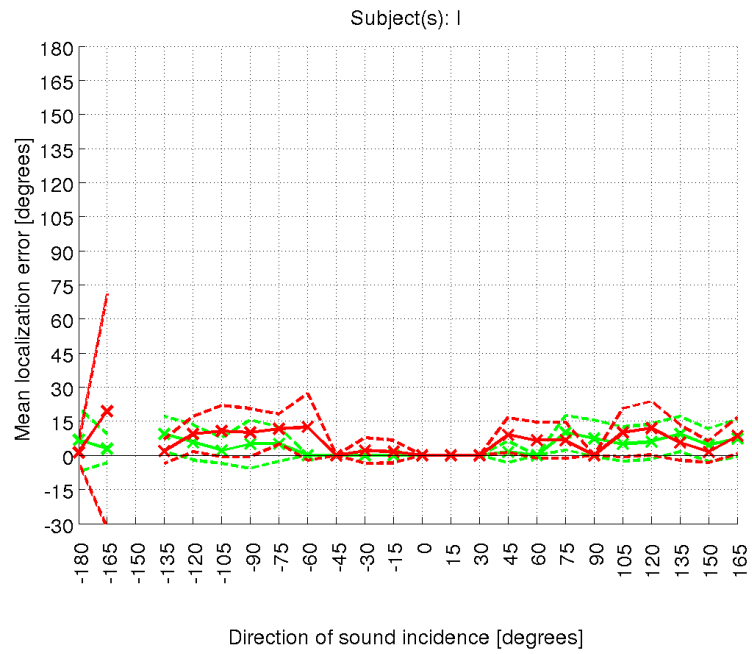
This is useful to see the subjects' performance in either condition separately, not just the absolute difference between the two.

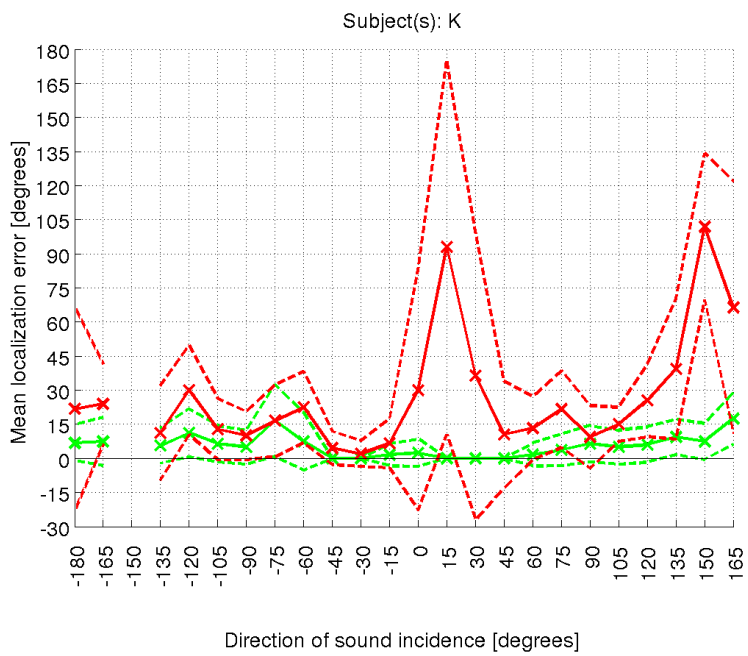




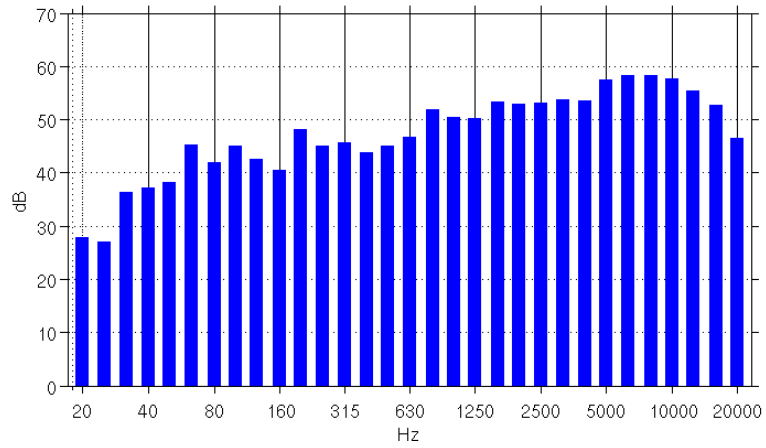




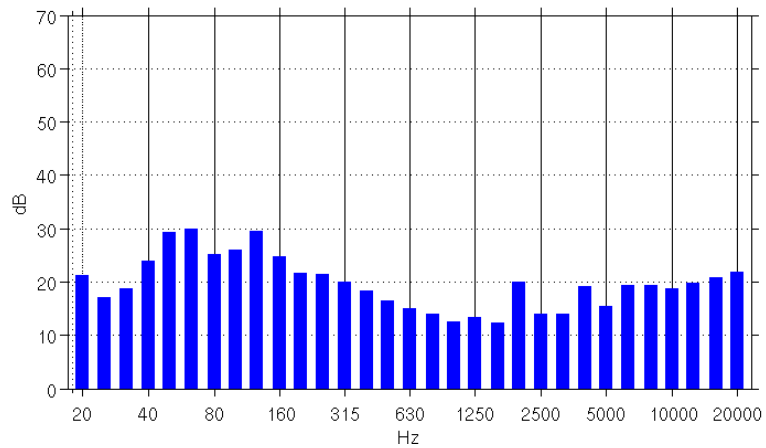




E Sound spectrum measurements



(a) Stimulus (65.7 dBA SPL)



(b) Background noise (29.1 dBA SPL)

Figure 10: Spectra of the noise burst stimulus used, and of the background noise in the listening room.

References

- [1] J. Blauert, *Spatial hearing : The psychophysics of human sound localization*. Cambridge, Mass.: MIT Press, rev. ed., 1997.
- [2] E. A. G. Shaw, "Transformation of sound pressure level from the free field to the eardrum in the horizontal plane," *The Journal of the Acoustical Society of America*, vol. 56, no. 6, pp. 1848–1861, 1974.
- [3] J. Blauert, *Communication acoustics*. Berlin: Springer, 2005. Jens Blauert (editor) ill.
- [4] F. L. Wightman and D. J. Kistler, "Headphone simulation of free-field listening. I: Stimulus synthesis," *The Journal of the Acoustical Society of America*, vol. 85, no. 2, pp. 858–867, 1989.
- [5] D. Purves, *Neuroscience*. Sunderland, Mass.: Sinauer Associates, 3rd ed., 2004.
- [6] C. G. Le Prell, D. Yamashita, S. B. Minami, T. Yamasoba, and J. M. Miller, "Mechanisms of noise-induced hearing loss indicate multiple methods of prevention," *Hearing Research*, vol. 226, no. 1-2, pp. 22–43, 2007.
- [7] The European Parliament and the Council of the European Union, *Directive 2003/10/EC on the minimum health and safety requirements regarding the exposure of workers to the risks arising from physical agents (noise)*, 2003. Retrieved from <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32003L0010:EN:HTML>.
- [8] L. Rayleigh, "On our perception of sound direction," *Phil. Mag.*, vol. 13, pp. 214–232, 1907.
- [9] B. C. J. Moore, *An Introduction to the Psychology of Hearing*. London: Academic Press, 5. ed., 2003.
- [10] D. McAlpine, "Creating a sense of auditory space," *The Journal of Physiology*, vol. 566, no. 1, pp. 21–28, 2005.
- [11] N. Fujiki, K. A. J. Riederer, V. Jousmaki, J. P. Makela, and R. Hari, "Human cortical representation of virtual auditory space: differences between sound azimuth and elevation," *European Journal of Neuroscience*, vol. 16, no. 1, pp. 2207–2213, 2002.

- [12] S. R. Oldfield and S. Parker, “Acuity of sound localisation: a topography of auditory space. II. pinna cues absent,” *Perception*, vol. 13, no. 5, pp. 601–617, 1984.
- [13] F. L. Wightman and D. J. Kistler, “Of vulcan ears, human ears and ‘earprints’,” *Nature Neuroscience*, vol. 1, no. 5, pp. 337–339, 1998.
- [14] J. C. Middlebrooks, “Narrow-band sound localization related to external ear acoustics,” *The Journal of the Acoustical Society of America*, vol. 92, no. 5, pp. 2607–2624, 1992.
- [15] P. M. Hofman, J. G. A. Van Riswick, and A. J. Van Opstal, “Relearning sound localization with new ears,” *Nat Neurosci*, vol. 1, no. 5, pp. 417–421, 1998. 10.1038/1633.
- [16] E. H. A. Langendijk and A. W. Bronkhorst, “Contribution of spectral cues to human sound localization,” *The Journal of the Acoustical Society of America*, vol. 112, no. 4, pp. 1583–1596, 2002.
- [17] J. C. Makous and J. C. Middlebrooks, “Two-dimensional sound localization by human listeners,” *The Journal of the Acoustical Society of America*, vol. 87, no. 5, pp. 2188–2200, 1990.
- [18] S. Abel, S. Tsang, and S. Boyne, “Sound localization with communications headsets: Comparison of passive and active systems,” *Noise and Health*, vol. 9, no. 37, pp. 101–107, 2007.
- [19] F. L. Wightman and D. J. Kistler, “The dominant role of low-frequency interaural time differences in sound localization,” *The Journal of the Acoustical Society of America*, vol. 91, no. 3, pp. 1648–1661, 1992.
- [20] C. Lorenzi, “Sound localization in noise in normal-hearing listeners,” *The Journal of the Acoustical Society of America*, vol. 105, no. 3, pp. 1810–1820, 1999.
- [21] N. L. Vause and D. W. Grantham, “Effects of earplugs and protective headgear on auditory localization ability in the horizontal plane,” *Human Factors*, vol. 41, no. 2, pp. 282–294, 1999.
- [22] T. Van den Bogaert, T. J. Klasen, M. Moonen, L. Van Deun, and J. Wouters, “Horizontal localization with bilateral hearing aids: Without is better than with,” *The Journal of the Acoustical Society of America*, vol. 119, no. 1, pp. 515–525, 2006.

- [23] V. Best, S. Carlile, C. Jin, and A. van Schaik, “The role of high frequencies in speech localization,” *The Journal of the Acoustical Society of America*, vol. 118, no. 1, pp. 353–363, 2005.
- [24] P. M. Hofman and A. J. Van Opstal, “Spectro-temporal factors in two-dimensional human sound localization,” *The Journal of the Acoustical Society of America*, vol. 103, no. 5, pp. 2634–2648, 1998.
- [25] U.S. Centers for Disease Control and Prevention, *Audiometry Procedures Manual*, 2005. Retrieved from http://www.cdc.gov/nchs/data/nhanes/nhanes_05_06/AU.pdf.
- [26] R. E. Walpole, R. H. Myers, S. L. Myers, and K. Ye, *Probability and Statistics for Engineers and Scientists (7th Edition)*. Prentice Hall, 2002.
- [27] M. B. Gardner and R. S. Gardner, “Problem of localization in the median plane: effect of pinnae cavity occlusion,” *The Journal of the Acoustical Society of America*, vol. 53, no. 2, pp. 400–408, 1973.
- [28] R. B. King and S. R. Oldfield, “The impact of signal bandwidth on auditory location: Implications for the design of three-dimensional audio displays,” *Human Factors*, vol. 39, no. 2, pp. 287–296, 1997.
- [29] D. Brungart, A. Kordik, C. Eades, and B. Simpson, “The effect of microphone placement on localization accuracy with electronic pass-through earplugs,” in *Applications of Signal Processing to Audio and Acoustics, 2003 IEEE Workshop on.*, pp. 149–152, 2003.
- [30] R. S. Bolia, W. R. D’Angelo, P. J. Mishler, and L. J. Morris, “Effects of hearing protectors on auditory localization in azimuth and elevation,” *Human Factors*, vol. 43, no. 1, pp. 122–128, 2001.
- [31] S. Carlile, P. Leong, and S. Hyams, “The nature and distribution of errors in sound localization by human listeners,” *Hearing Research*, vol. 114, no. 1-2, pp. 179–196, 1997.