Carl von Ossietzky Universität Oldenburg

Fachbachelor Physik, Technik und Medizin

BACHELORARBEIT

The Effects of Independent Automatic Gain Control on Spatial Hearing of Individual Bilateral Cochlear Implant Users: Computational Modeling and Psychoacoustic Experiments

vorgelegt von

Helen Kasim

Betreuender Gutachter: Prof. Dr. Mathias Dietz

Zweite Gutachterin: Dr. Hongmei Hu

Oldenburg, den 29.06.2023

Contents

Acknowledgments

Special thanks to the participants who took part in the long measurements and made this project possible. I am also thankful to my supervisors, Hongmei, Mathias, and Rebecca, who supported me along the way, and of course, many thanks to the rest of the group for always being welcoming and giving advice when needed. Lastly, I'd like to mention Mira, Swantje, and Malte, who supported me during my writing process with their comments.

List of Figures

List of Tables

Abbreviations and Units

Abbreviations

AGC	Automatic gain control
AN	Auditory nerve
BTE	Behind-the-ear
CI	Cochlear implant
CIS	Continuous interleaved sampling
CSSS	Channel-specific sampling sequence
DR	Dynamic range
EI	Excitatory-inhibitory
FSx	Fine-structure processing with x channels
GWN	Broadband Gaussian white noise
HRIR	Head-related impulse response
ILD	Interaural level difference
ITD	Interaural time difference
LGF	Loudness growth function
LSO	Lateral superior olive
MCL	Most comfortable level
NH	Normal-hearing
OLSA	Oldenburger sentence test
PK	Broadband pink noise
PKHP	Highpass-filtered pink noise
RMSE	Root-mean-square error
THR	Threshold

Units

CL	Current level
cu	Current units
dB	Decibel
dB SPL	Sound pressure level in decibel
Hz	Hertz
pps	Pulses per second
qu	Charge units
µA	Micro ampere
µs	Micro seconds

1. Introduction

Cochlear implants (CIs) are devices that can restore hearing in people with severe to profound hearing loss meaning that the sound transmission in the inner ear is impaired, and therefore the brain does not receive sufficient input to perceive sound. A reason for this can be, e.g., missing hair cells, which regulate the transmission and amplification of sound. A CI, therefore, needs to replace all previous steps of sound perception (e.g., Bacon et al., 2004).

CIs are now devices utilized by more and more people and are now used by more than one million users (Zeng, 2022), compared to approximately 736,000 in the year 2019 (U.S. Department of Health and Human Service, 2021). Since the number of CI users has increased significantly, there is a higher need for solving current problems that affect different aspects of their hearing performance, such as speech understanding or spatial hearing.

A typical normal-hearing (NH) listener uses binaural cues such as interaural level differences (ILDs) and interaural time differences (ITDs) for spatial hearing. The ILDs are more relevant for localizing high-frequency sounds as the sound level is attenuated more on the contralateral side because of the head-shadow effect (e.g., Shaw, 1974) and ITDs are too short. Compared to that, low-frequency signals can be localized better using the ITDs.

So far, CI users can mainly access the ILD cues and, to some extent, envelope ITDs (e.g., Grantham et al., 2007; van Hoesel and Tyler, 2003), which describe the ITD obtained from the sound envelope. It is, therefore, even more important that these cues are preserved correctly. Still, bilateral CI users can localize well with a root-mean-square error (RMSE) of about 30° compared to unilateral CI users with an RMSE of around 70° but still worse than NH listeners with an RMSE of about 6° (e.g., Dorman et al., 2016; Grantham et al., 2007).

One of the reasons why localization is more difficult for bilateral CI users compared to NH listeners is the distortion of the ILD cues due to the independent processing of automatic gain control (AGC) at each ear. The AGC is responsible for elevating softer signals and compressing louder signals above a given threshold, consequently lowering the level. Unlike the AGCs in hearing aids, the AGC in CIs are broad-band and therefore applied over the whole frequency spectrum. Using independent AGC compressors for each ear thus distorts those ILD cues at higher

levels because the amount of compression differs on both sides in each frequency band. This can lead to a different perception of localization (e.g., Archer-Boyd and Carlyon, 2019).

Recently, an open-source computer model framework for simulating the spatial hearing abilities of bilateral CI listeners has been developed by Hu et al. (2023) based on the models of Kelvasa and Dietz (2015), Klug et al. (2020), and Hu et al. (2022). It includes a binaural signal generation stage, a CI processing stage with a wide range of speech-coding strategies, and a human model, which includes an electrode-neuron interface stage, an auditory nerve (AN) model stage, an excitation-inhibition binaural interaction model stage, and a decision stage based on the normalized hemispheric rate differences. Overall, the model framework shows persuasive results in simulating the average performance of CI users in different tasks, including left-right discrimination, lateralization, and localization performance of CI users for single- or multiple-electrode direct stimulation and free-field listening experiments. However, not all features of the average CI user's performance could be captured. For example, when simulating the effect of two independent AGCs, the model presents a localization perceived from the opposite direction for broad-band stimuli at high input levels. This means that a broad-band signal presented, e.g., from the left is perceived as from the right side.

As reverse localization does not usually appear in previous localization studies (e.g., Dorman et al., 2014; Seeber and Fastl, 2008) and only in very sparse data for CI users with devices from the manufacturer Advanced Bionics (e.g., Brown, 2018; Pastore et al., 2021), it is now interesting to examine what would lead to reverse localization.

Another aspect of AGCs the model shows is inverted ILD curves in lower frequency channels. Dorman et al. (2014) stated that there are three cases of how the AGC does or does not influence the end results for localization: (a) the AGC is inactivated on both sides, leading to no distortion of ILD cues; (b) the AGC is activated on one side resulting in changed and smaller ILDs, which are sometimes even inverted, especially in combination with more head-shadow making it harder to localize correctly; (c) AGC compression is activated on both sides possibly reducing the ILDs to zero, and localization would not be possible to some extent. Inverted ILD cues in the low-frequency channels have also been shown in previous literature at levels above the AGC compression threshold but still with correct localization (e.g., Dorman et al., 2014; Gray et al., 2021). It could be possible that these inverted ILDs are perceived in another way, e.g., as non-fused auditory objects. This means that one stimulus would be perceived as two separate auditory objects.

A study by Suneel et al. (2017) stated that CI users typically have trouble fusing two different sounds presented to each ear. They suggested a possible correlation between the RMSE for localization and the amount of binaural fusion. In that study, a CI user's perception was simulated through similar CI processing steps, also known as vocoder stimuli, and presented to NH listeners. In general, they did not state that the two different sounds in each ear might result from independent AGCs. To my knowledge, such a study including actual CI users has not been conducted yet.
This thesis investigates how AGC impacts spatial hearing in various scenarios by utilizing psychoacoustic experiments and model simulations. Binaural fusion and localization in the horizontal plane are the main foci of this thesis.
Two bilateral CI users were recruited to achieve this, and a series of psychoacoustic tests based on the assumption of possible perception distortions due to two independent processors were conducted. These include unbalanced loudness, decreased localization performance, and the amount of binaural fusion. In order to examine how AGC impacts the processing pathway, each stage of the model was analyzed.
From the literature and information above, it is now hypothesized that reverse localization does not occur in the experiments. Full binaural fusion should happen if the CI users also present good localization results. The reason for the inverted localization results predicted by the model might be due to a parameter or a combination of the default parameters that do not correctly represent the average CI user. Modifying the parameters in regards to an individual CI user's settings could show which parameter can be improved and also give a better prediction of the data obtained in the experimental part.

2. Cochlear Implants and Spatial Hearing

2.1. Cochlear Implants

2.1.1. General Structure

A CI system consists of two components: an external processor with a transmission coil and an implanted part with an electrode array and receiver coil.

Typically behind-the-ear (BTE) processors are worn with a hook keeping it in its position and connected to the implanted part with a magnet (e.g., Bacon et al., 2004). The electrode array inserted into the cochlea stimulates the AN fibers.
The stimulation is realized by sending electric pulses to specific electrodes. CIs typically include 12 to 22 electrodes depending on the manufacturer, and each is assigned a different frequency band (e.g., Swanson, 2008). Pulses, which make up the final output of the processors, are typically sent biphasically for most manufacturers with different polarities so that the tissue's overall charge results in zero (e.g., Swanson, 2008).
By sending pulses to the frequency-specific electrodes and thus also sending electric fields that stimulate the AN, a signal is transmitted to the brain. The electrodes do not target the frequencies on the basilar membrane typically stimulated by acoustic hearing. Rather, since the electrode array usually does not fully reach into the apical turn of the cochlea, the entire frequency range needed for everyday life is assigned to a more compressed and shifted part of the basilar membrane (e.g., Swanson, 2008).
When stimulating the AN through electric fields, current spread appears, meaning that neighboring frequency regions are also stimulated. This phenomenon is also known as the spread of excitation (Swanson, 2008).

2.1.2. Processing Stages

Different stages are required for the CI to replace the acoustic way of hearing. First, a microphone on the processor receives acoustic input, which will be further processed. A pre-emphasis filter with a cutoff frequency typically at 1200 Hz is applied to increase the energy of the signal in the high-frequency channels, similar to the functionality of the middle ear (Swanson, 2008).
The next step is the AGC, which elevates lower levels and compresses higher levels. It is also partially responsible for compressing the large dynamic range (DR) of around 120 decibels (dB) for NH listeners in the acoustic domain into a small dynamic range of approximately 20 dB accessible to CI users (e.g., Bacon et al., 2004). Usually, the gain is linear until the so-called kneepoint, which indicates the threshold at which higher levels are compressed. The compression depends on different parameters: the compression ratio and time constants such as attack and release time.
First, the knee-point is given in decibels for sound pressure level (dB SPL) and compared to the signal level. For each manufacturer, it is possible to vary the sensitivity range of the kneepoint so that it fits an individual CI user. For the manufacturer MED-EL, the kneepoint can be set by choosing a sensitivity value,

usually at 75%, which equals a value of 52.7 dB SPL. Overall, it can be chosen between 48 dB SPL and 67 dB SPL (Vaerenberg et al., 2014). Once compression is activated, it is applied the same to all frequency bands.
The next parameter is the compression ratio. This differs between different manufacturers and defines the amount of compression and therefore, the reduction of level. One example is a ratio of 3:1 like in MED-EL devices, meaning that the level difference between the signal and the threshold is decreased three times as much. A higher compression ratio such as 12:1 in Advanced Bionics devices would thus lead to more compression (Giannoulis et al., 2012).
The time constants are known as attack and release time. The attack time determines how long it takes to decrease the gain to the level, which depends on the compression defined by the ratio. The release time is responsible for the time needed to increase the gain, if the signal has been reduced to even below the kneepoint level (Giannoulis et al., 2012).
Modern devices often feature dual-loop AGC, which includes fast and slow time constants. The fast constants are used for high-intensity transients, such as slamming doors, and the slow time constants are used for everything else (e.g., Dhanasingh and Hochmair, 2021). The fast constants are only used in extreme situations as they distort speech understanding (Boyle et al., 2009). Slow time constants on the other hand distort localization cues (e.g., Archer-Boyd and Carlyon, 2019).

So far, all these steps are usually classified as front-end processing (Swanson, 2008).
The next stage is the division into different frequency bands by band-pass filtering the signal. The bandwidth is determined by the number of electrodes. For MED-EL, the manufacturer of the devices tested in this thesis, the number is set to 12 and the filter bands are usually log-spaced. If the number of activated channels is lower for different reasons, such as a high impedance, the bandwidth is set accordingly to still cover the whole frequency range. Fast-Fourier-Transform or finite-impulse-response-based filterbanks such as gammatone filterbank are used (e.g., Swanson, 2008).
Once the frequency channels are set, the signal envelope is extracted, rectified, and smoothed. This is needed as the amplitudes of pulse sequences are set by the envelopes (Swanson, 2008).

In the next step, a coding strategy is applied. The most common strategy is the so-called Continuous-Interleaved-Sampling (CIS) strategy, first introduced by Wilson et al. (1991). As the name suggests, the signal in each frequency band is sampled (usually at a fixed rate) and pulses are continuously sent. Here, pulses

are not delivered simultaneously to the specific electrode but interleaved starting from the most basal channel. In MED-EL devices the channel number starts from the most apical channel meaning the lower frequencies.
There are many variants of the CIS strategy nowadays. One of them by MED-EL is called FSx, which stands for fine structure processing in the two or four most apical channels. The idea is that the temporal fine-structure ITDs can still be exploited at lower frequencies and consequently be used for a better perception. The time of pulses is then determined by the zero-crossings of the signal at which pulse packets are consequently sent. The number of pulses delivered to the electrodes in these apical channels is therefore dependent on the signal, also called channel-specific sampling sequences (CSSS). This coding strategy is thus very different from the original CIS strategy in these apical channels and for all other channels remains the same (e.g., Dhanasingh and Hochmair, 2021).
After that, a biphasic pulse train is multiplied with the extracted signal envelope, and the frequency of the pulses is also usually given in pulses per second (pps). The output is then converted into electric current units in micro Ampere (μA). To simulate a similar loudness perception as for NH listeners, a loudness growth function (LGF) is applied. These are usually logarithmic and map the acoustic to the electric dynamic range. The parameters that regulate the acoustic range are the so-called base level and the saturation level. The exponent defining the shape of the function is usually known as α_c or map-law parameter (e.g., Swanson, 2008). Parameters determining the electric dynamic range are given in manufacturer-specific units and are set for the threshold (THR) level and the most comfortable level (MCL). The manufacturers Cochlear Ltd. and Advanced Bionics use specific units such as Current level (CL), and MED-EL's unit is presented either in charge units (qu) describing the charge delivered to each electrode in nano Coulomb (nC) or in current units (cu) that are given in µA.

The output after all these processing stages is called an electrodogram (see Figure (2.1)) and holds all the information for when a pulse is sent to which electrode at which strength that consequently stimulates the AN.

2.2. Spatial Hearing

For NH listeners, spatial hearing is determined by binaural cues that stem from signal differences between the ears. ILDs and ITDs are used for localization in the horizontal plane. ITDs describe the time difference at which the sound arrives at each ear and is usually important for differentiating low-frequency signals. Usually, NH listeners rely more on ITDs for localization which is processed pre-

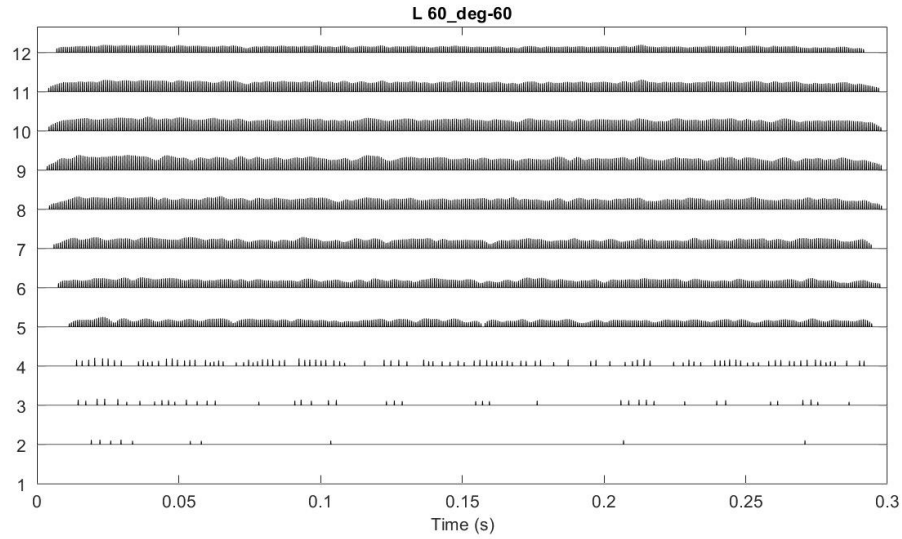

Figure 2.1.: Example electrodogram that includes all the information, including amplitude and time sent to each electrode for broadband Gaussian white noise at 60 dB SPL. It shows the output for the left ear with the stimulus presented at -60°.

dominantly in the medial superior olive in the brain (e.g., Blauert, 1996).
ILDs, on the other hand, are caused by the diffraction of sound waves at the head (e.g., Litovsky et al., 2021), which can be viewed as an object obstructing and dampening the signal at the contralateral side. ILDs are also dependent on the angle. This means that at angles around e.g. ±70°, the ILDs at higher frequencies rise to their maxima since, in that position, the head attenuates the sound the most. This is also known as the head-shadow effect (Shaw, 1974). Lower frequencies are not affected as much because of an obstacle and higher frequencies are attenuated stronger. As a result, ILDs are more effective for higher frequencies. ILDs are processed during later stages in the brain inside the lateral superior olive (LSO) as a simple subtraction between the ipsi- and contralateral side (e.g., Dietz, 2016).
Compared to NH listeners, CI users have a higher difficulty accessing those binaural cues resulting in poorer localization performance on average. They usually cannot utilize ITDs, at most only the envelope ITDs, and the ILDs are distorted due to AGC (e.g., Seeber and Fastl, 2008). Additionally, for higher levels, the ILDs are inverted for lower-frequency channels. The inability of CI users to access ITD cues is caused by factors like unsynchronized hardware and processing (e.g., Dorman et al., 2014).

3. Methods and Materials

3.1. Participants

Two CI users were measured in three appointments: a pilot session and two formal appointments, which were about seven months apart. The first CI user (CI1) participated in the formal appointments with two different processors and the second CI user (CI2) also participated as the pilot subject.
CI1 was a 62-year-old male MED-EL user who was measured in the first appointment with his own old processors (OPUS 2) as well as new ones (SONNET 2) in the second appointment. CI1 had worn the old processors for about seven years before the first measurement. He had around six weeks to get used to the new processors before conducting the experiment again. Before implanted, he was wearing hearing aids, and the cause of his deafness is not known. The participant reported that for the new processor, the old settings were first imported and then fitted.
The other participant, CI2, was a 22-year-old female MED-EL bilateral CI user with two SONNET processors. She was implanted 21 years ago and has used these processors for about seven years. The etiology is not known.
Additional information on the settings in all processors for both participants was read out using the clinical software MAESTRO 9.0 by MED-EL (MED-EL Elektromedizinische Geräte GmbH, 2020). The parameters used in the model can be seen for both participants in table 3.1.
The MED-EL website (MED-EL Elektromedizinische Geräte GmbH, 2022) provides default settings for the different processor types, including automatic sound management 1.0 which is enabled in all of them. This is the first adaptive sound system introduced by MED-EL for all of its devices. It ensures that in all tested processors a dual-loop AGC compressor was included. The newer SONNET processors also include directional microphones (Dhanasingh and Hochmair, 2021). Thus, the participants were asked to use the omnidirectional program for all experiments.

The experiments were approved by the Ethics committee of the University of Oldenburg. Each appointment had an average duration of approximately 90 min and the subjects could take unlimited breaks.

3.2. Test Environments

All experiments were conducted in the Virtual Reality Lab of Oldenburg University, an anechoic chamber covered with melamine wedges on every side of the room to avoid sound reflections. The stimuli were presented from an array of 25 loudspeakers (Genelec 8030 loudspeakers) in the frontal hemisphere (see Figure 3.1), which were placed at an approximate distance of 2.5 to 3 m from the position of the ears. Every loudspeaker was labeled with a number visible to the participants and was chosen on a monitor after the stimulus presentation, even though less than half of them were activated. The actual range was between -75° and +75° with an increment of 15° for most of the measurements. The time needed for each session was therefore reduced and according to a study by Santala and Pulkki (2011), at least for NH subjects, a higher amount of possible options should not result in significantly different outcomes.

To ensure that the CI users were facing the front and that there was no head movement, the head was kept straight to the front either with an infrared head-tracker system using the Qualisys Head Tracker software or during the second appointment by facing the monitor under the loudspeaker labeled with the number thirteen. The participants were allowed to move their heads after the stimulus was finished playing to identify the perceived direction.

The experiments were started in a neighboring room using an alternative-force-choice toolbox implemented in *MATLAB* (Ewert, 2013).

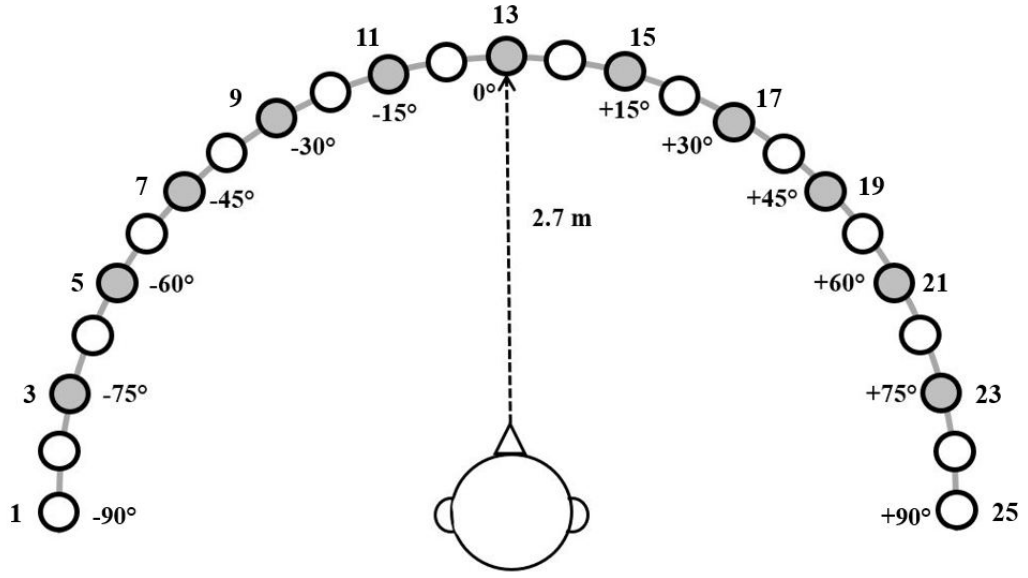

Figure 3.1.: Setup used in all experiments. The loudspeakers were numbered from one to 25. Only those with filled circles were activated and used during the experiments. CI users were still able to choose one out of all 25 loudspeakers.

3.3. Stimuli

For the pilot study, four types of stimuli were used: Gaussian white noise (GWN), pink noise (broad-band (PK) and highpass filtered (PKHP)), and the word "Schuhe" from the Oldenburger sentence test (OLSA) (Wagener et al., 1999). For all other sessions only the stimuli PK and PKHP were tested. The stimuli duration was 300 ms except for the word which was roughly 500 ms. The GWN was band-pass filtered between 150 and 10000 Hz. PK with a bandwidth of 125 to 8000 Hz and PKHP with 2000 to 8000 Hz were tested.
The stimuli were calibrated to 40, 60, and 80 dB SPL using a sound level meter (Brüel & Kjaer type 2250 and pre-polarized free-field microphone type 4189). The microphone was held at the position of the head, and levels were measured using longer versions of the stimuli, which had a duration of about 30 s. There was a pool of 30 pre-generated versions for all noise stimuli. During presentation, the stimuli were selected randomly.

3.4. Psychoacoustic Experiments

3.4.1. Loudness

At the beginning of a test appointment, a binaural loudness scaling was conducted to check whether the stimuli were too loud or too quiet. PK was presented to the five loudspeakers located at ±75°, ±45°, and 0° (see Figure 3.1), at which the amount of attenuation because of the head-shadow effect was different (e.g., Shaw, 1974). The measurement started from the left (-75°, loudspeaker number three) and then changed to the next, clockwise. For each direction, three levels were tested in the order 40, 60, and 80 dB SPL. It was repeated twice for each condition. The scale consisted of eleven categories: from zero ("cannot hear") to ten ("too loud") (see Figure 3.2).
Since unbalanced loudness perception with two independent CIs could affect the localization performance, the same loudness scaling experiment as for the binaural perception was conducted monaurally for each side to check possible asymmetric hearing with the participant's speech processors.
The monaural part was left out during the last appointment as the previous results showed that this information could be obtained from other data, such as the processor settings.

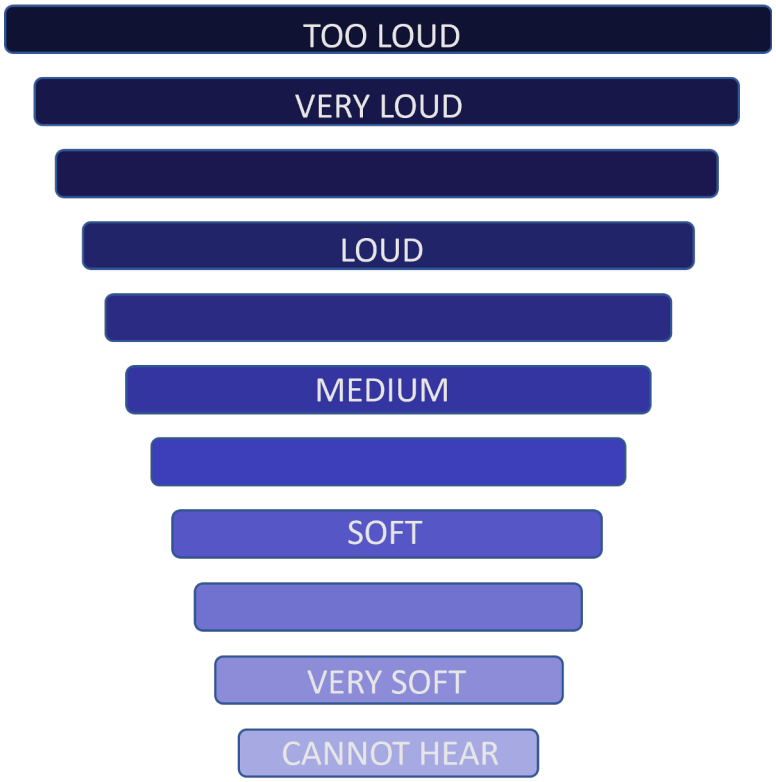

Figure 3.2.: Loudness scale ranging from ”cannot hear” to ”too loud”.

3.4.2. Localization

To check the localization performance with independently activated AGCs, all four types of stimuli were used for subject CI2. The experiment was separated into three runs. In each run, the test order was PK, PKHP, the word “Schuhe”, and GWN. For each stimuli type, the sound was presented randomly from one of the eleven loudspeakers marked in gray (see Figure 3.1). During the experiment, the participants were instructed to face the front until the stimulus was presented. After that, the CI users could look around and choose a number assigned to every loudspeaker. There were two repetitions per condition, resulting in a total of six repetitions per stimulus and angle. To test the effect of levels, the procedure described above was conducted at three levels in the order of 40, 60, and 80 dB SPL. The levels were expected to represent the three different cases concerning AGC compression: 40 dB SPL as below the threshold, 60 dB SPL as slightly above and possibly not activating compression on the contralateral side, and 80 dB SPL as the level exceeding the AGC knee-point on both sides.
After the pilot measurement, the azimuth increment was increased to 30° and the range spanned from -60° to 60° for the first formal appointment. For the second formal appointment, the range was increased again to all marked loudspeakers (see Figure 3.1), and the experiment was only conducted using PK and PKHP.

The reduction of conditions was intended to reduce the time needed for each session.

3.4.3. Non-Fused Images

During the final appointment, the experiment analyzing the amount of binaural fusion after AGC compression was combined with the localization task to test this condition for all azimuth and level combinations. When choosing the perceived direction on the monitor, a text box was also presented asking for the perceived binaural fusion. In case a subject would not perceive the presentation as one auditory object, they were asked to describe it in terms of loudness and frequency. Due to a lack of time, binaural fusion was not tested during the first formal appointment with both CI users.

3.5. Model Framework

Figure 3.3 shows selected stages from the open-source model framework of Hu et al. (2023). During the first stage, all signals from the experiments were loaded and convolved with head-related impulse responses (HRIRs) from Denk et al. (2018). The default HRIRs from the Kayser et al. (2009) database were replaced, which is explained in the appendix.

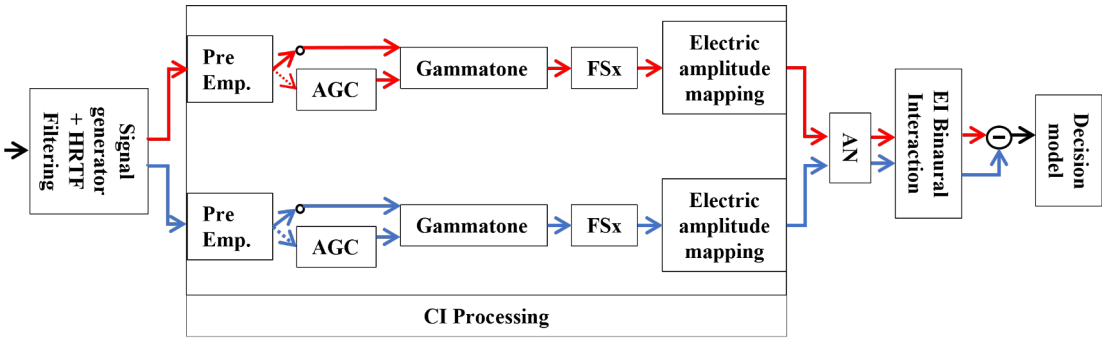

Figure 3.3.: Selected stages in the model simulations by Hu et al. (2023) predicting horizontal localization for bilateral CI users. AN – Auditory Nerve, EI – Excitatory and Inhibitory

The model itself consists of two parts: the CI processing stage and the binaural hearing model. All parameters that were examined in this thesis were changed in the CI processing stage.

The CI processing stage includes a pre-emphasis filter, an AGC (on or off), and a gammatone filterbank-based FSx coding strategy, where all default parameters were later changed to fit the recruited CI users. This is similar to the general processing stages in typical CI processors as described in chapter 2.1.

The FSx strategy implemented here differs slightly from the manufacturer's default strategy. Unlike before, the pulses, not pulse packets, are sent at the local maxima of the input signal's envelope. X=4 refers to the number of apical channels for which it is implemented and can also be changed manually to fit an individual CI user's settings. The other channels are computed similarly to the CIS strategy.
The pre-emphasis filter consisted of a first-order highpass Butterworth filter with a cutoff frequency of 1200 Hz. The broadband AGC was based on the manual by Giannoulis et al. (2012) with manufacturer-specific settings (see Table 3.1).
The gammatone filterbank splits the original signal into the specified number of frequency channels and fits the logarithmic axis accordingly. The signal envelope is then extracted for each channel for further processing.
Here, the loudness growth function was taken from Swanson (2008) and describes logarithmic loudness growth using the following equation:

$$y = \begin{cases} \frac{\log(1+\alpha(\frac{\nu-B}{M-B}))}{\log(1+\alpha)} & B \leq \nu \leq M \\ 0 & \nu \leq B \\ 1 & \nu \geq M \end{cases} \tag{3.1}$$

ν equals the envelope extracted in previous stages, B the base level, and M the saturation level given by a manufacturer (default value in this case from Cochlear Ltd.). The level range was from 25 to 65 dB SPL. Again, the steepness of the curve is described by α, often defined as the map-law value in MED-EL devices. The output y refers to values inside the electrical dynamic range.
After this, the binaural hearing model for electric hearing starts. The next processing stage, the electrode nerve interface, enabled the experiment-specific setting, such as the spread of excitation. The AN model was based on the implementation by Fredelake and Hohmann (2012) and simulated electrically stimulated AN fibers.
For the binaural neuron interaction model, a population of excitatory-inhibitory (EI) neurons along the tonotopic array were used (Klug et al., 2020). Each EI neuron receives input from twenty excitatory and eight inhibitory AN fibers, which are then implemented as inputs for the coincidence counting model by Ashida et al. (2016). The individual AN fibers have different properties regarding the spiking rate.
In the decision model, the normalized hemispheric rate difference was then taken and linearly mapped to azimuthal sound source localization. The hemispheric rate difference describes the difference in spike rates between the left and right sides inside the LSO (Klug et al., 2020).

There were no free parameters that were changed in any stage that was part of the binaural hearing model.

Returning to the initial problem of the model, the model predicted reverse localization for Gaussian white noise (see Figure 3.4). Here, the Kayser et al. (2009) database was used.

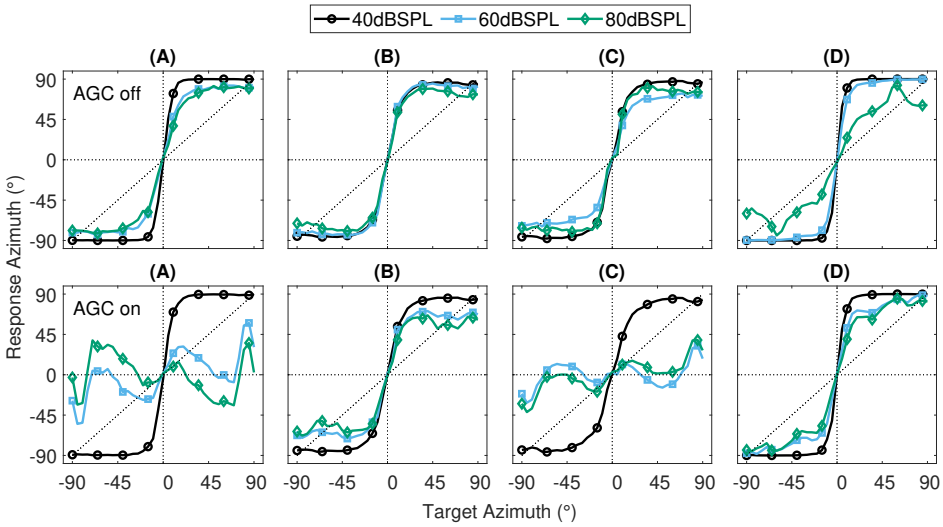

Figure 3.4.: Default model localization outputs without AGC (top row) and with AGC (bottom row) for different stimuli such as (A) Gaussian white noise, (B) the word ”Schuhe” from the Oldenburger sentence test and pink noise ((C) broadband and (D) highpass-filtered).

The model was now individualized for the two participants by implementing their processors' settings into the CI processing stage to analyze their localization performance after AGC compression. However, it's important to note that the binaural hearing model was not changed and therefore a deviation from the experimental results had to be taken into account. Still, the results provided an insight into the changed results in each processing stage after AGC and the influence of each parameter.

3.6. Model and Subject Parameters

A total of thirteen parameters were changed in the model. The following table (see Table 3.1) shows those parameters that were replaced for each subject as well as the default values.
In the model, default parameters by the manufacturer Cochlear Ltd. were used.

		CI1 - old				CI1 - new			
		L		**R**		**L**		**R**	
Number of electrodes		9		12		9		12	
Number of FSx channels		2		3		2		3	
map-law value α_c		500				500			
Channel border freq. in Hz. (log-spaced)	Lower	100 - 5728		100 - 6323		100 - 5728		100 - 6323	
	Upper	237 - 8500		198 - 8500		237 - 8500		198 - 8500	
AGC	Knee-point	52.7 dB SPL				52.7 dB SPL			
	Comp. ratio	3:1				3:1			
	Attack time	100 ms				100 ms			
	Release time	400 ms				400 ms			
Channel number		**1-2**	**3-9**	**1-3**	**4-12**	**1-2**	**3-9**	**1-3**	**4-12**
Pulse rate in pps		1415		1325		2454	1227	1227	
Pulse duration in μs		67.1		59.5		88		55.4	
Default sensitivity		0.75		0.85		0.85		0.85	
Level values in CL	THR	103.3	118.1	4.26	75.1	101.1	98.2	4.26	77.2
	MCL	167.7	217.3	207.4	217.9	229.2	206.6	219.8	218.7

		CI2				Model default	
		L		**R**		**L**	**R**
Number of electrodes		12		12		12	12
Number of FSx channels		4		4		4	4
map-law value α_c		500				415.96	
Channel border freq. in Hz. (log-spaced)	Lower	70 - 6321		100 - 6323		100 - 7000	100 - 7000
	Upper	197.9 - 8500		204.2 - 8500		250 - 8000	250 - 8000
AGC	Knee-point	52.7 dB SPL				50 dB SPL	
	Comp. ratio	3:1				2:1	
	Attack time	100 ms				50 ms	
	Release time	400 ms				200 ms	
Channel number		**1-4**	**5-12**	**1-4**	**5-12**	**1-12**	**1-12**
Pulse rate in pps		6173	1235	7258	1210	900	900
Pulse duration in μs		25.4		25.4		25	25
Default sensitivity		1		0.85		1	1
Level values in CL	THR	9.897	68.2	51.87	77.39	100	100
	MCL	202.9	209.4	205.5	211.5	200	200

Table 3.1.: Default model parameters and individualized parameters used in the model for CI1 and CI2. The channel frequency borders are shown from lowest channel number to highest (all logarithmically spaced). THR – threshold level, MCL – most comfortable level, CL – current level

This means that the level parameters from each subject had to be re-scaled before being implemented into the model using an equation by Di Lella et al. (2020):

$$\text{Current level [CL]} = \log_{10}\left(\frac{\text{Electric current } [\mu A]}{17.5}\right) * \frac{255}{2} \tag{3.2}$$

3.7. Model Simulations

For simplicity, asymmetric hearing was excluded from the current simulations. Thus the center frequency, the threshold (THR), and the most comfortable level (MCL) values were set symmetric for both ears. Because of this, each parameter was tested individually first to test the influence of that parameter and then in

combination with all other parameters for one CI user to simulate their individual localization performance. All are simulated with azimuths between -90° and +90° with an increment of 5°. Each simulation was repeated five times and averaged at the end for a more general prediction.
The influence of every parameter listed in Table 3.1 was then examined at different processing stages, such as ILDs after electric output compression, spike rates for AN and EI neuron outputs, and predicted localization after the decision model. For comparison, the results of all parameter combinations were analyzed with activated and inactivated AGC.

4. Results

4.1. Psychoacoustic Results

4.1.1. Loudness

Informal binaural loudness scaling confirmed that the levels presented were not perceived as too loud or too quiet and were used throughout all experiments. After the first session, participant CI2 complained after the last measurements with the GWN that for this stimulus 80 dB SPL was too loud.
A slight bias to the right appeared in the monaural results for CI1. Here, all scaling values were higher than the opposite side at all angles by approximately one scaling unit. For CI2, a bias occurred for the monaural left results. This bias did not appear consistently in the binaural results. A mean and standard deviation were not computed as each condition had only two repetitions. The table containing all scaling values is included in the appendix.

4.1.2. Non-Fused Images

During the last session, the non-fused perception of objects was not reported by either subject at any level for any stimulus, with one exception. Subject CI1 reported non-fusion in one case but only when the loudspeaker presented a delay within one stimulus, which was not purposefully included.
The subject CI2 reported that non-fusion appeared in general scenarios such as conversations but not during the experiment.

4.1.3. Localization

The localization results by both CI users can be seen in Figure 4.1 and 4.2 and Table 4.1. The figures show the data as response azimuth as a function of target azimuth. Correct results were plotted on the diagonal dotted line. CI1's results can generally be described through a linear fit and CI2's instead as sigmoid functions. The lowest RMSE was almost always present at 60 dB SPL. The results for PKHP and the additional stimuli measured with CI2 can be seen in the appendix.

Participant	CI1						CI2					
Stimulus	PK			PKHP			PK			PKHP		
Level in dB SPL	40	60	80	40	60	80	40	60	80	40	60	80
exp. appointment one	24.2°	12.2°	14.1°	26.1°	23.7°	15.9°	24.5°	18.6°	32.7°	24.9°	15.2°	33.9°
exp. appointment two	24.3°	13.5°	18.7°	22.5°	20.4°	17.5°	20.4°	19.4°	21.9°	20.7°	17.1°	17.8°
def. model	38.8°	15.7°	21.2°	38.9°	34.6°	31.9°	38.8°	15.7°	21.2°	38.9°	34.6°	31.9°
indv. model - L old	36.7°	18.9°	30.4°	35.8°	26.9°	23.5°	39.4°	30°	30.1°	38.6°	32.3°	26.2°
indv. model - R old	38.8°	33.1°	30.3°	37.7°	33.6°	27.6°	38.9°	12.7°	26.6°	37.8°	29.3°	21.9°
indv. model - L new	38.5°	23.3°	34.8°	38.8°	31.9°	27.5°	-	-	-	-	-	-
indv. model - R new	38.8°	31.9°	31.5°	37.8°	34.1°	28.6°	-	-	-	-	-	-

Table 4.1.: Average root-mean-square errors for localization of all levels and stimuli. The results are shown for experiment data, the default (def.) model, and individualized (indv.) model results for each side (L - left, R - right).

Starting with the individual results for each stimulus for CI1, the PK was localized better in comparison to the PKHP. Especially at 60 dB SPL (see Figure 4.1, E), the average RMSE reached a value of 12.2° with the old processors and a value of 13.5° with the new processors. The stimuli presented at the highest level were also localized better than at the lowest level for both the PKHP and PK. Though the results for both stimuli can be almost described as linear, the PKHP data (see Figure A.1) was shifted more toward the right side. This is also slightly the case with both stimuli not being accurately localized at 0°.

Overall, the results with the new processors were in the same range as the old processors' and did not show a consistent improvement, at most for PKHP. Here, the center was localized correctly on average at 40 and 80 dB SPL. Comparing both processors showed a significant difference for a few data points for the PK when using a paired two-tailed t-test but did not indicate consistent behavior for improved or worse results. At 40 dB SPL, the results at 0° and +30° were significantly different with p-values of 0.0065 and 0.043, and at 80 dB SPL, a significant difference appeared at -30° with a p-value of 0.0011.

As reported before, the localization performance by CI2 can generally be described as sigmoid curves for all stimuli, including the GWN (see Figure A.4) and the OLSA word (see Figure A.3). They were also overall perceived better at

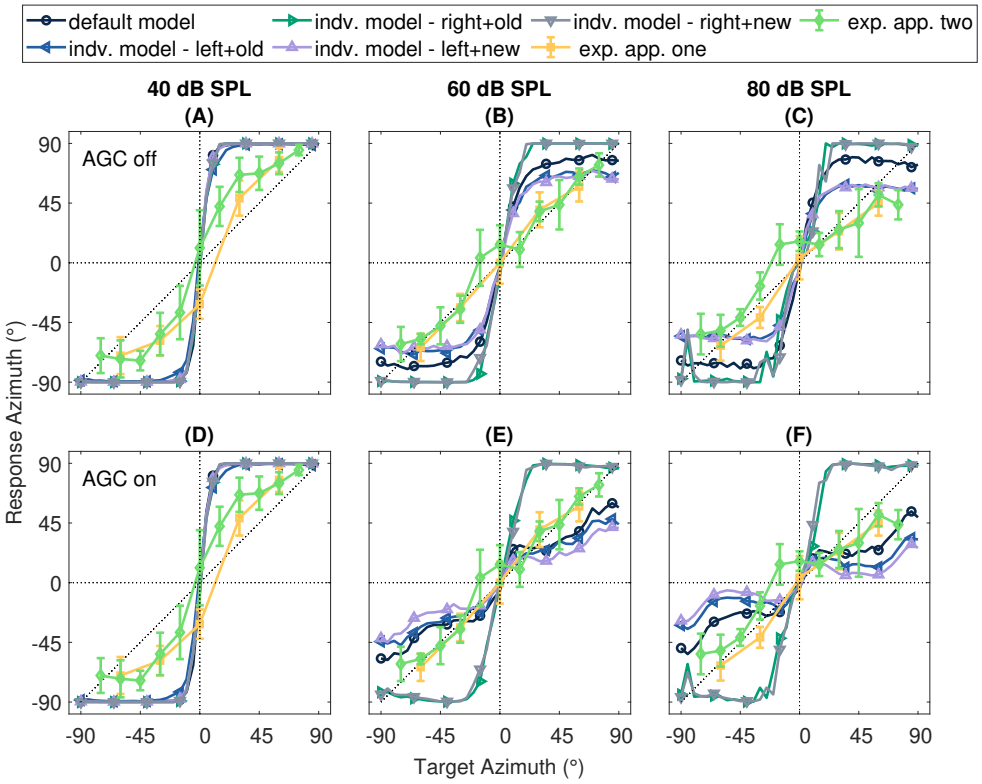

Figure 4.1.: Response azimuth as a function of target azimuth shown for experimental and model results (AGC: off (top) and on (bottom)) in comparison to the default model output for pink noise presented at 40, 60 and 80 dB SPL. The mean is shown for experimental data with the standard deviation. All parameters from CI1 from Table 3.1 were used. Correct answers are shown on the diagonal dotted line.

60 dB SPL. The RMSE, in this case, unlike for the other participant, improved more compared to the first measurements for both retested stimuli. Significant differences appeared at all levels: at 40 dB SPL at -30° and +45° with p-values of 0.0071 and 0.0422 for the PK; at 60 dB SPL at +15° and +45° for the PKHP with p-values of 0.0108 and 0.0034; at 80 dB SPL for the PK at azimuths from -30° to +15° with p-values ranging from 0.0062 to 0.0292 and for the PKHP for angles from -45° to +30° with p-values between 0.000045 and 0.0172.
Similar to CI1, the subject showed a small bias but to the left instead of the right side. In the last measuring appointment, this bias disappeared more, leading to an improvement at 80 dB SPL of more than 10°. For the pilot measurements, the results could be described as sigmoids that become flatter as the level increases. Reverse localization was only reported by CI2 when the OLSA word was presented (see Figure A.3). Here, two presentations, one on each side at ±45° at 60 dB SPL, were perceived as from the opposite side. The PKHP (see Figure A.2) was overall localized best when comparing the RMSE values.

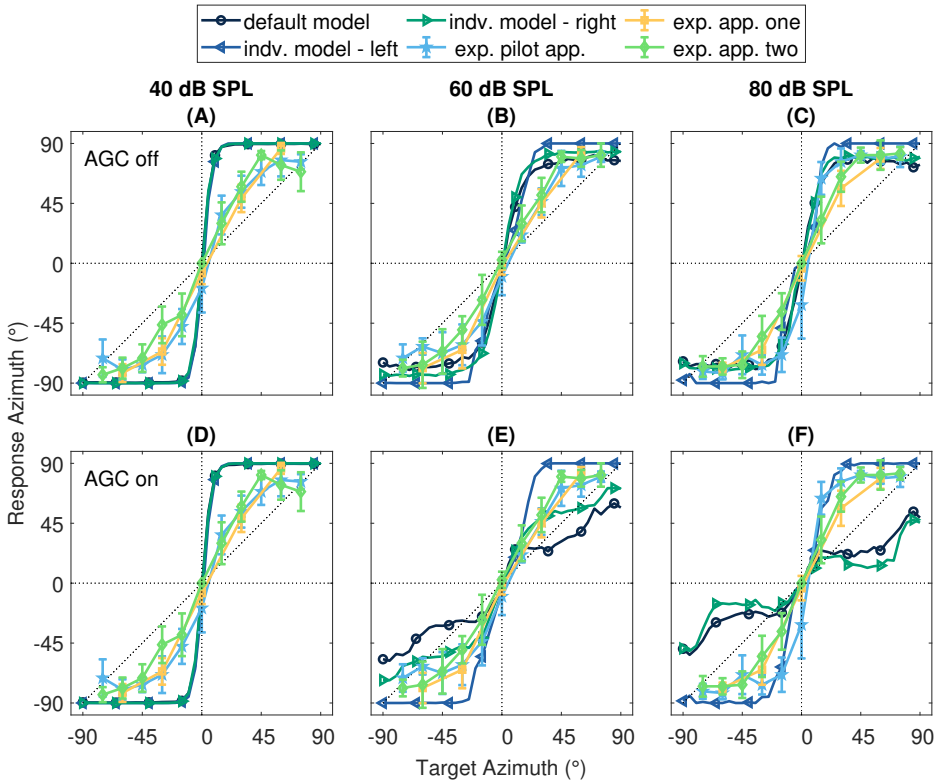

Figure 4.2.: Localization results for CI2 with PK. Description as in Figure 4.1.

4.2. Model Simulations

4.2.1. Loudness

Unbalanced loudness scaling was not tested in the current model with a fitting back-end stage. Still, when using even slightly different level parameters, e.g. MCL or THR values, a directional bias is also predicted by the model using the localization decision stage. This unbalance appeared when using left- and right-sided level values, e.g., from CI2. Applying these to the corresponding side leads to a heavy bias to the left side, with almost all data points shifted.

4.2.2. Localization

Overall, the individualized model did not show a consistent improvement compared to the default results. Looking at the RMSE for localization performance (see Table 4.2), the default model generally showed a lower value meaning better localization. When comparing the model simulations to the experimental results, the RMSE was not consistently better meaning that neither the default nor individualized simulations were better fitted to the data overall. Overall, the model only described the experimental data at best with an RMSE of approximately

15°. The model results at 40 dB SPL show outputs shifted to ±90°.

Participant	CI1						CI2					
Stimulus	PK			PKHP			PK			PKHP		
Level in dB SPL	40	60	80	40	60	80	40	60	80	40	60	80
exp. app. one & def. model	27.3°	19.5°	20.5°	27.8°	47°	40°	26.1°	28°	46.3°	24.4°	30.1°	15.1°
exp. app. one & indv. model - L old	29.2°	20.8°	28°	27.8°	39°	30.7°	25.8°	22°	14.2°	24.6°	29.2°	18.3°
exp. app. one & indv. model - R old	27.2°	35°	39.1°	27.1°	45.6°	38.9°	26.4°	17.1°	53.7°	23.8°	24.5°	17.5°
exp. app. two & def. model	27.3°	21°	20.8°	36.4°	50.1°	44.6°	27.5°	29.3°	38.8°	27°	26.9°	22.2°
exp. app. two & indv. model - L old	-	-	-	-	-	-	29.1°	20.3°	16.7°	27.1°	25.8°	19.5°
exp. app. two & indv. model - R old	-	-	-	-	-	-	27.8°	18.2°	44.8°	26.7°	22.2°	15.4°
exp. app. two & indv. model - L new	27.4°	26.8°	29.7°	36.6°	47°	37.7°	-	-	-	-	-	-
exp. app. two & indv. model - R new	28.2°	39°	49.4°	35.8°	48.4°	41.6°	-	-	-	-	-	-

Table 4.2.: Average root-mean-square errors between the simulated and experimental results. The results are shown for experiment data for each appointment (app.), the default (def.) model, and individualized (indv.) model results for each side (L - left, R - right).

The default model for CI1's data showed plausible results, especially for the PK. When comparing the simulations using the right and left-sided parameter settings for this participant, the left side fit more. This holds for both the old and new psychoacoustic results. Overall though, it appears that the experimental results are within the range of the simulated results on both sides (see Figure 4.1). Comparing the statistical results for both sides using a paired t-test again showed that for all processors almost all results were significantly different at all levels in all of those cases with p-values below 0.05.
The pattern of each simulation for CI1, including the old and new processor's settings, showed that the simulations with left-sided processors predict a localization more towards the center, and the right side shows an extreme, where most results are shifted to ±90 (see Figure 4.1). The latter has the shape of a sigmoid function for the PK, almost like a z-curve, whereas the left side can rather be described as two combined sigmoid functions.
The PKHP generally for all its simulations showed a dip around ±60° to ±70° (see Figure A.1).

The simulations for CI2 showed a similar outcome, but the opposite for each side. For the left side now, the sigmoid curve became flat towards the edges. Comparing all RMSE for CI2, the model generally provided a better prediction for the individualized results. For the PKHP (see Figure A.2), both sides showed similar RMSE. The PK, however, presented a bigger difference between both sides. The PKHP, similar to the results for CI1, showed a dip in the curve at ±60° to ±70°. For the other stimuli, the GWN and OLSA word could also be described as sigmoid curves becoming flatter as the level becomes higher. The results for the OLSA word also included a dip at the levels described before. Testing for significant differences showed that again most of the results presented

a difference to experimental data with p-values lower than 0.05. In general, there were less significant differences when comparing the psychoacoustic data to the simulations for the right processor at 60 dB SPL.

4.3. Influence of Different Parameters

The parameters listed in Table 3.1 show a difference in the effect of the influence that they have on the results. The outcome of the model, meaning the localization stage, is shown for AGC-related parameters (see Figure 4.3) and level-related parameters (see Figure 4.4) with broadband noise, GWN in particular. The stimulus for which reversed localization was predicted at the beginning was chosen. Here, the following parameters were changed one at a time inside the model: the knee-point, the compression ratio, and the attack and release time of the AGC, THR, and MCL values with the other kept at default, an in- or decreased dynamic range (DR) and a shifted DR, where the THR and MCL values were reduced by a constant factor. Other changed parameters are examined in the appendix. These include the number of activated channels, the phase duration of each pulse, the stimulation rate in pps, the number of channels with fine-structure processing, the coding strategy from FSx=4 to CIS, and the map-law value α_c. In addition to the parameters shown in table 3.1 other parameters, like fast-time constants or a higher compression ratio for the AGC were tested.

Each parameter influencing the AGC compression was examined. In general, the results at 40 dB SPL never changed with different AGC parameters (see Figure 4.3, A, D, G, J).
When changing the knee-point only by a few dB, the results remained the same. When the knee-point was increased to even higher values, such as the minimal sensitivity corresponding to a knee-point for MED-EL devices of 67 dB SPL, the results started showing greater similarity to the outcomes without AGC compression, at least for the lower levels (see Figure 4.3, B). As the levels were increased, the outcome was similar to the default result at 60 dB SPL (see Figure 4.3, B-C).
The compression ratio had a larger effect on the outcome. The model default had the lowest ratio and also more correct localization results. As the levels and the compression ratios were increased the results became more reversed (see Figure 4.3, E-F).
The attack and release time constants were increased to at least twice the duration but still showed no big difference in localization (see Figure 4.3, G-L). The influence of these was expressed better through ILDs after the CI processing stage. The default model uses time parameters which are usually for the slow

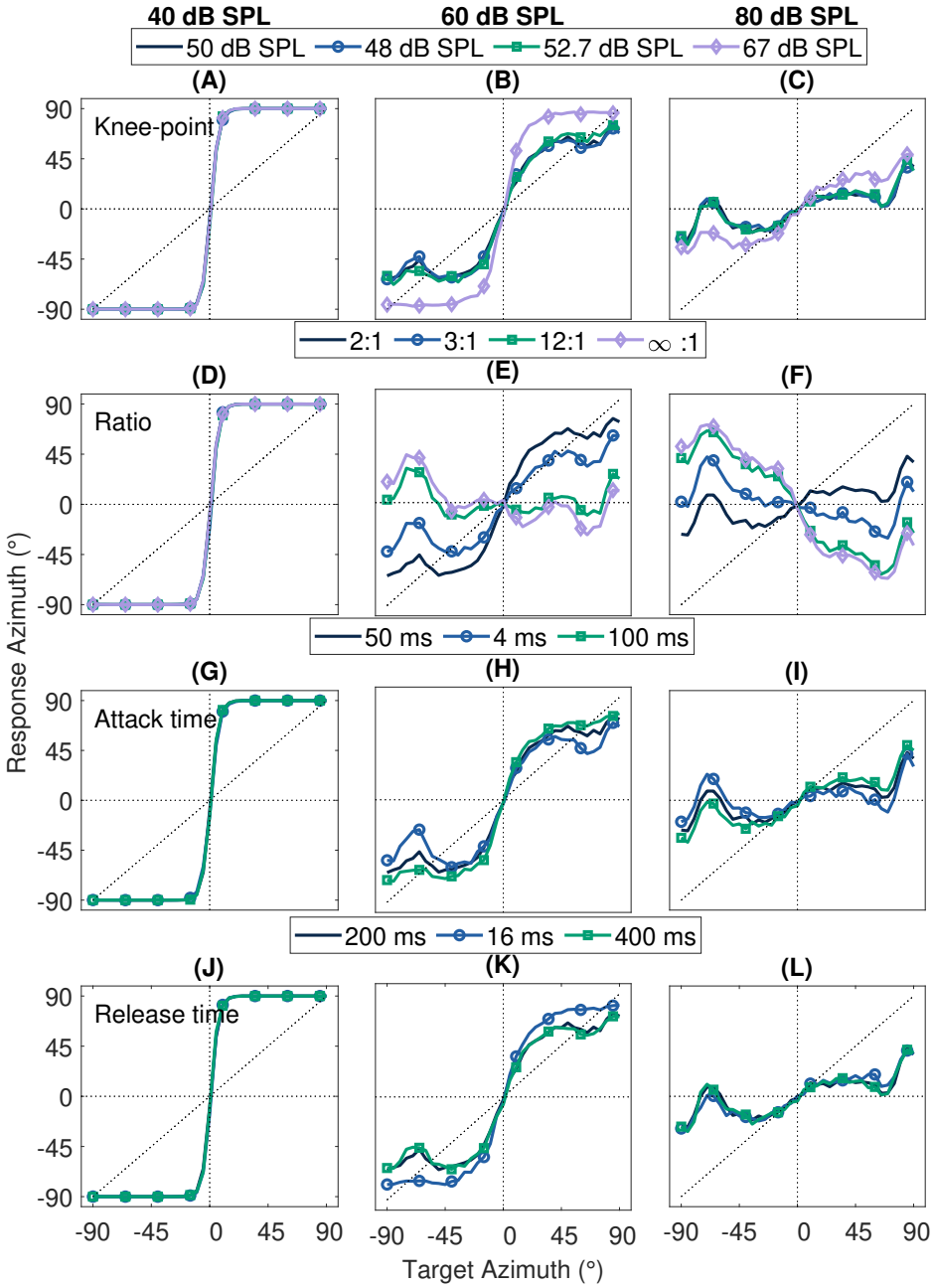

Figure 4.3.: Localization results simulated changed automatic gain control parameters in the CI processing stage for Gaussian white noise. The results are presented at different levels for changed (A-C) knee-point values, (D-F) compression ratios, (G-I) attack time, and (J-L) release time constants. The black functions show the default results.

compression of the AGC, which is also where the reversed ILDs occurred. When switching to constants used for the fast compressors, the low-frequency ILDs were not inverted anymore but the localization results stayed the same. Next, the level-related parameters such as THR and MCL values were changed including the DR

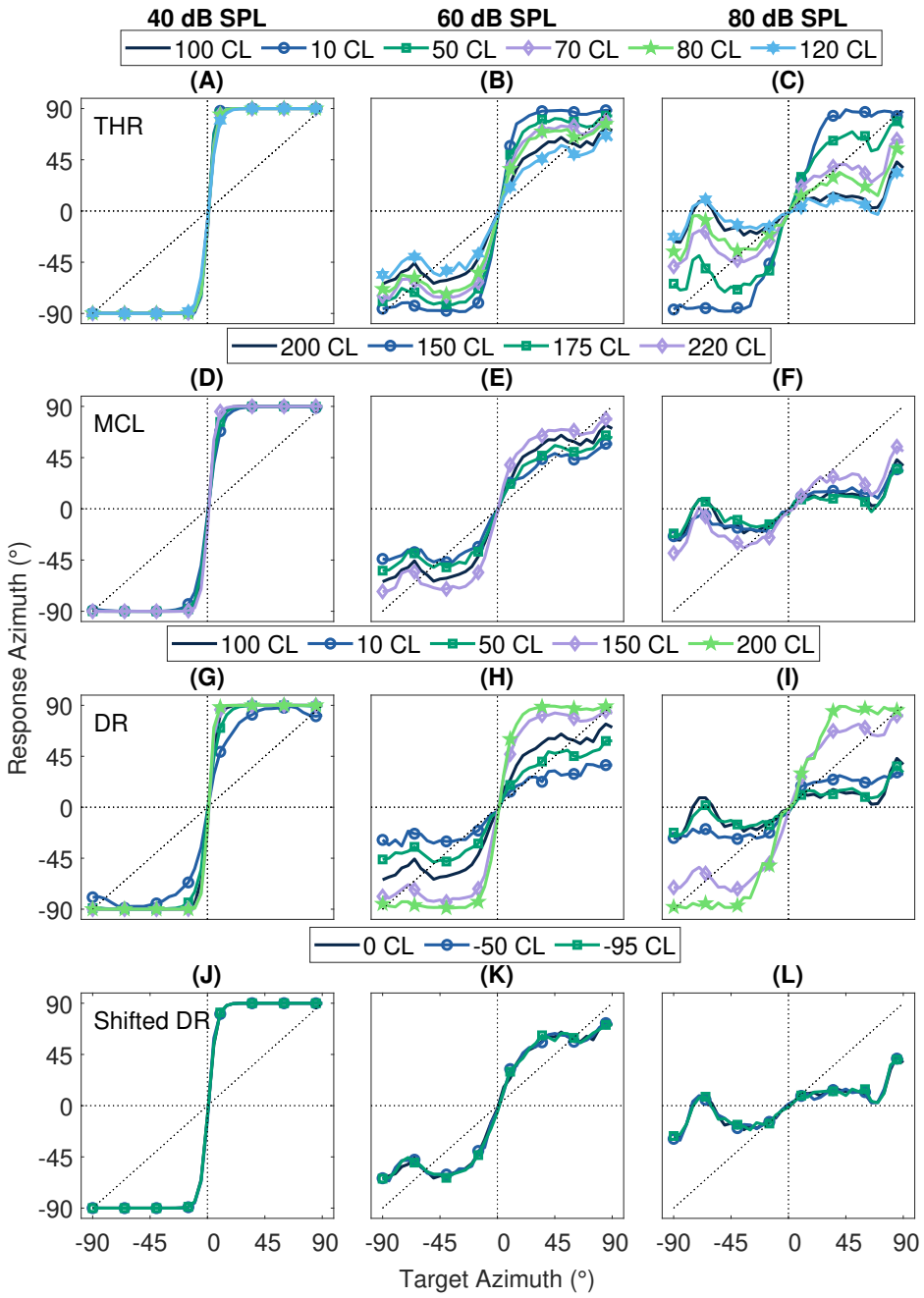

Figure 4.4.: Localization results simulated changed level-related parameters in the CI processing stage for Gaussian white noise. The results are presented at different levels for changed (A-C) THR, and (D-F) MCL values, (G-I) in- or decreased dynamic range, and (J-L) a shifted dynamic range. The black functions show the default results.

in general.

What had the most considerable effect on the results, were the THR values (see Figure 4.4, B-C) or rather the increase in DR range (see Figure 4.4, H-I). There were differences in the parameters and the corresponding results between the old

and new processors in CI1 and likewise compared to CI2 (see Figure 4.5). CI2 had the biggest DR reaching up to a maximum of 200 CL (almost covering the whole range) and, therefore, the lowest THR values at least on the left side. The other side had a DR more similar to the default but the THR values are much lower in comparison at least for the more basal and middle channels. The values for CI1 on the right side were almost alike to CI2's left side. For the other processors, the THR values were even higher, and the DR was even smaller. In general, the stimulation limits spanned a larger DR than the default range of 100 CL, at least for most processors tested. Usually, the MCL values were around 200 CL as before, but the THR values differed greatly. For some channels in CI1 and CI2, the THR values almost reached 0 CL, especially for the channels with activated fine-structure processing. Testing those parameters individually showed that a change in THR values or rather the DR was necessary to show localization results more similar to the subjects. The results for these large DRs lay almost entirely on the edges of the horizontal plane at almost all levels. Having a DR of roughly 150 CL e.g. between 50 and 200 CL or bigger made a big difference in the localization results (see Figure 4.4, H-I). Modifying the THR values or the DR did not always result in different outcomes. For THR values higher than 100 CL (see Figure 4.4, B-C) or DR smaller than 100 CL (see Figure 4.4, H-I) nothing remarkably changed, and the localization was also not predicted to be worse than before.

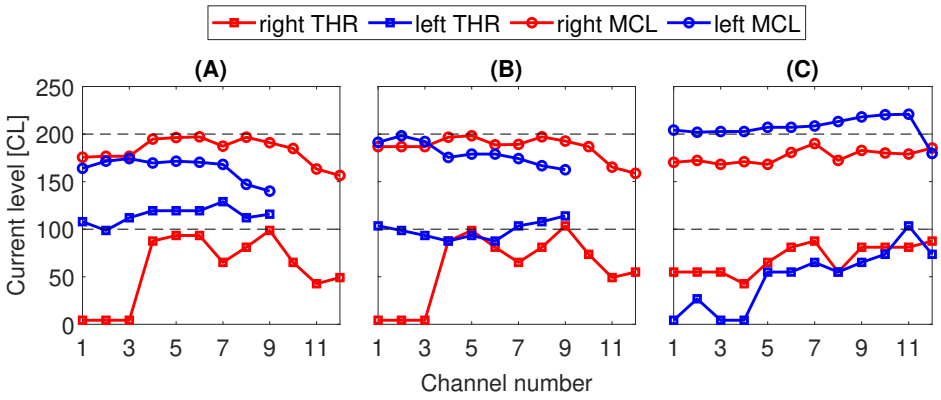

Figure 4.5.: Level values (MCL and THR) for (A) CI1 with old processors, (B) CI1 with new processors, and (C) CI2. The dotted lines represent the model default.

Another interesting occurrence happened when increasing all THR values with a constant factor for CI2 on the left processor or the right for CI1. The idea was to see if some channels having THR values similar to the default would lead to an inverted or almost inverted perception again. Surprisingly, the results stayed unchanged compared to before, and the localization was still not reversed for

broadband noise. This outcome indicated a channel-dependent behavior within the model. To check this, all channels were kept at the default values and always only one channel was changed. It was done for all channels and tested for values of [10:10:90] CL. As initially thought, the results after AGC compression were different for each channel (see Figure 4.6). For some channels, only a change of at least 20 CL was needed to achieve a localization prediction with no reversals. For others even reducing the values to 10 CL still led to reversed localization. Overall, especially the middle channels led to a reduction of reversals, then the apical channels after a higher decrease of the THR values, and for most basal channels the reversal was not removed. Additionally, a change in the basal channels sometimes led to the overall localization being flat for almost all target azimuths. Also, e.g., reducing the THR value in channel eight to 70 or 80 CL in the default model predicted results more similar to the tested CI users.

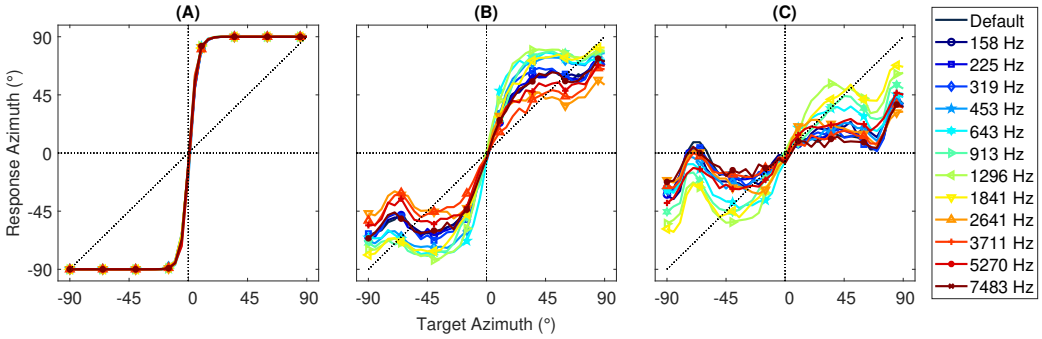

Figure 4.6.: Change in each channel with a THR value of 80 CL after AGC compression at (A) 40, (B) 60, and (C) 80 dB SPL.

Next, the direct position of the DR was investigated. For this, the default range by 50 CL and 95 CL respectively. The THR values were now more similar to the values depicted in Figure 4.5. Still, the results stayed the same compared to the default (see Figure 4.4, J-L).
For the results at 40 dB SPL in all conditions, only a small or no difference was seen compared to the default outcome.

4.4. Influence of AGC on Different Stages

The influence of the AGC compressor can be seen to a different extent in all stages of the model (see Figure 4.7). Here, the default model parameters were used first.
Starting from the ILD plots after the CI processing stage, the differences between non-activated (see Figure 4.7, A) and activated AGC (see Figure 4.7, B) lay in the lower frequency channels. After AGC compression, these were inverted and

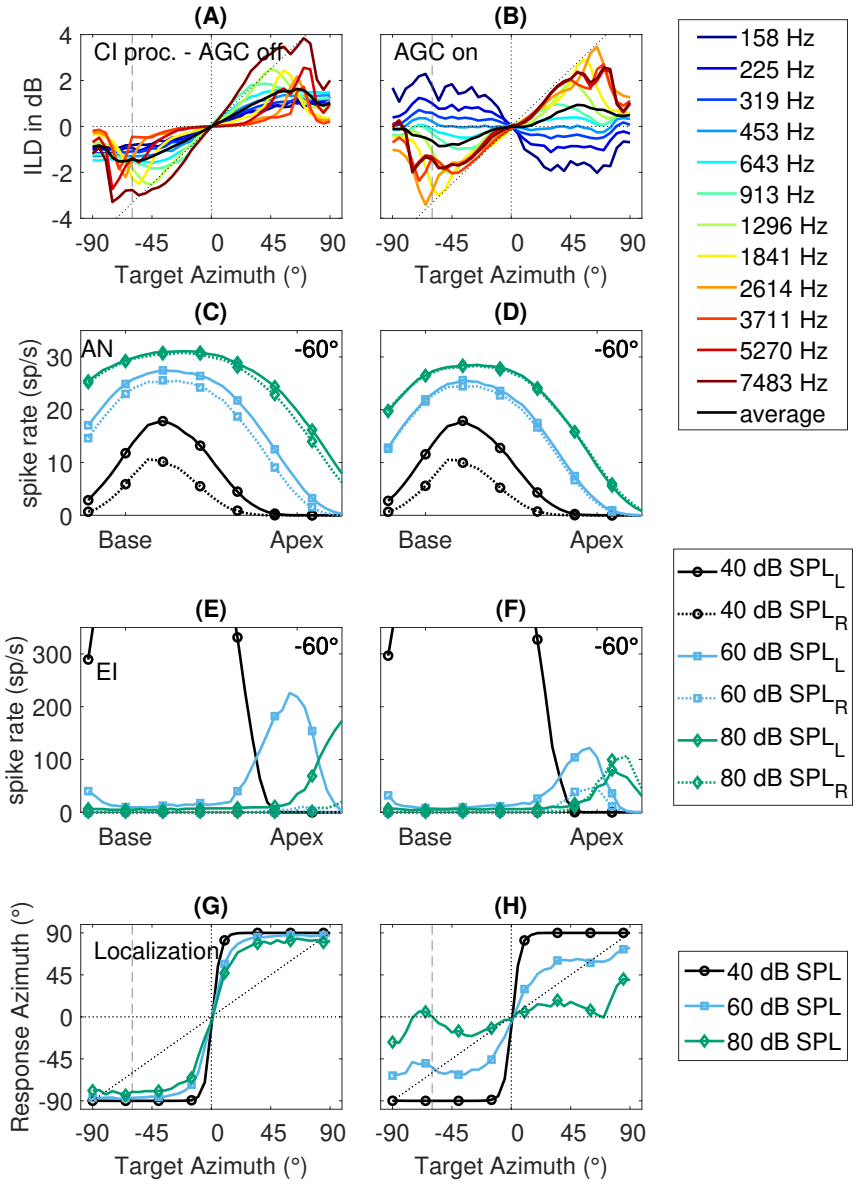

Figure 4.7.: Different stages of the model using the model's default parameters. The left column shows the results without and the right with activated AGC. The ILD results are shown at 80 dB SPL. To see the difference in spike rates for the middle stages, the y-axis was reduced.

also showed a higher absolute value in comparison to inactivated AGC. Only the four most apical channels were reversed in sign.

From looking at the AN spike rates with activated AGC (see Figure 4.7, D), the higher levels resulted in a higher spike count for the contralateral side towards the lower frequencies and apex. As the levels were increased, the spike rates also rose. For the levels that were not being distorted by the AGC or at least not as much, the spike rates showed no differences. The inverted results in the fre-

quencies below 500 Hz were not as clear as in other stages. Another aspect that can be observed from the AN spikes is that a bigger shift to the apex leads to an increased difference between the ipsi- and contralateral sides. The maximum spike rate for all levels was around higher frequencies.
In the EI stage, the spike rates rose when the levels were increased and the reverse in the lower frequencies below 500 Hz was seen better (see Figure 4.7, E-F). The highest spike rates appeared for the lower levels, and the difference was shown more clearly between the left and right sides. For 40 and 60 dB SPL, the contralateral spike count is mostly around zero except for the lower frequencies when the AGC was activated. The maxima were shifted toward the apical frequencies as the level was increased, with and without AGC compression. At 80 dB SPL, where unusual localization was predicted, most spike rates were around zero for all frequency bands with a center frequency above 500 Hz. Here, the contralateral side showed a higher spike rate again. The overall spike rate after AGC compression was decreased for the ipsilateral side and increased for the contralateral side. Lastly, the target azimuth as a function of response azimuth representation shows the results mentioned before (see Figure 4.7, G-H). Here the predicted localization around 0° can be seen for higher levels after the AGC is activated.

When implementing the CI users' settings, the ILDs mostly did not change. The biggest differences appeared in the later stages of the model. As shown before, the localization was shifted toward ±90°, similar to the localization at 40 dB SPL in the default simulation. The AN and EI stages did not show an increased spike rate for the contralateral side now.

5. Discussion

In the following section, the influence of AGC on the spatial hearing abilities of bilateral CI users is discussed based on the experimental and simulated findings. It is examined how the AGC affects the distortion of ILD cues, binaural fusion and the initially predicted inverted localization. Additionally, the results of the previous chapter (see Chapter 4) are compared, either between the participants, the different processors used by one subject, or between the experiments and the model. Lastly, the influence of the dynamic range on the localization for CI users is discussed as well as the limitations and future directions.

5.1. Influence of AGC on Spatial Hearing

5.1.1. ILD Distortions

This thesis aimed to analyze the effect of independent AGC on spatial hearing for individual CI users. As seen in the literature, AGC affects binaural perception (e.g., Archer-Boyd and Carlyon, 2019; Dorman et al., 2014). It was shown to distort localization cues by reversing the sign of some ILDs in low-frequency channels. In extreme cases, it could result in flat localization curves meaning that the results are centered around the mid-line. As described in the introduction (see Chapter 1), Dorman et al. (2014) stated three cases in which the AGC could influence spatial hearing depending on which and how many sides the compression is activated. In the second case described, where the AGC is active only on one side, the compression together with the head-shadow effect works as an extended low-pass filter explaining why the ILDs become reversed. Since most of the signal energy after pre-emphasis filtering is in the higher frequency parts, the inverted parts do not impact the results as much. Dorman et al. (2014) showed a direct correlation between the ILD cues and the resulting localization. They also hypothesized that the inverted ILDs might not be relevant for localization and that the high-frequency components play a dominant role in deciding the direction.

The three scenarios can also be represented through the default simulations shown before. Here, the different levels also tested in the experiments can be used to describe each case. The lowest level at 40 dB SPL is directly related to the first case as the level is below the knee-point. The result, therefore, does not change within the model after the AGC is activated. The second case can be connected to the next higher level at 60 dB SPL, where small compression takes place. At 80 dB SPL the model shows the third case where, for both sides, the average ILD was decreased so much that the default simulation shows localization centered around the middle.

Localization for the three cases by Dorman et al. (2014) is also stimulus-dependent. Though the CI users are both relatively consistent with their results, the model shows a difference e.g. between white and pink noise. For pink noise, the energy of the signal is higher in the lower frequency bands. During pre-emphasis as the first stage in CI processing the energy of these frequencies is attenuated. The threshold for which AGC compression is activated in the first place then starts at higher levels. According to Dwyer et al. (2021), the knee-point for GWN would be around 56 dB SPL and for the PK more around 62 dB SPL, which can also be seen in the model results. The PK at 60 dB SPL is not affected as much. For

the pink noise in general, the reverse ILDs in the lower frequency channels might also not have a large influence since the overall energy in these frequency bands is reduced because of the pre-emphasis filter. This holds for broadband AGC compression, which compares the level of the whole signal to the thresholds.

5.1.2. Binaural Fusion

Regarding the binaural fusion of auditory objects connected to AGC compression, there is no directly connected literature to my knowledge. In general, it could be compared to binaural pitch fusion because there are two peaks on opposite sides for high and low-frequency components perceived as one. Creating a stimulus by combining two sounds with different directions inside the model could be tested and the outcome examined.
Binaural pitch fusion has been investigated in previous studies (e.g., Reiss et al., 2018) and found to happen in up to three to four octaves. This could be transferred to the results here, but the reversed ILDs had a higher absolute value than the ILDs which had a distance of three to four octaves. It was also not directly possible to evaluate the perceived pitch as it was only assessed if a second auditory object was perceived by the participants, which was never the case. Consequently, pitch fusion was not the focus of this study.
This theory can also be connected to one in previous literature, which states that the dominant or average ILD over all channels is responsible for the overall direction of perception (e.g., Archer-Boyd and Carlyon, 2019; Dorman et al., 2014).
Another approach is the comparison to a study done by Suneel et al. (2017) where they tested the correlation of binaural fusion and RMSE in localization for vocoded NH listeners when combining two different sounds received from each ear. Here, they proposed that the binaural fusion, as we define it, might be hard to distinguish for CI users and tested it for NH therefore first. They found RMSE of under 30° correlating with full binaural fusion. Comparing this idea now to the individual CI user's results, it also aligns as no non-fusion was reported during the experiments, and the RMSE was mostly around 20° on average. The experiments will therefore need to be also tested on CI users with worse localization abilities.
Overall, if possible, it could also be helpful to read out ILDs passed through the individual processors since the ILDs in the model are quite big compared to other studies. Dorman et al. (2014) showed relatively low absolute values for the inverted channels and Gray et al. (2021) reported even less reversed ILDs in general.

The CI user CI2 reported non-fusion in general scenarios such as conversations but still reported full binaural fusion during the experiments. This result could

be due to the nature of the task or due to the subject being slightly familiar with the setup.

5.1.3. Reverse Localization

As discussed in the introduction, the default model parameters gave a prediction of a reverse localization for GWN with a frequency range of 150 and 10000 Hz. This result was not observed in either of the two CI users measured for this thesis. In general, though, both localized better than average when comparing their RMSE of roughly 20.8° to the RMSE in Dorman et al. (2016) of 29°. Using their processors in the model did also not result in reversed localization.
To my knowledge, it has only been reported in a few studies (e.g., Brown, 2018; Pastore et al., 2021), and even then, it might also be explained as almost a full bias towards one side. For these studies, the results can be linked to higher compression because both conducted experiments with CI users using Advanced Bionics devices. These include a higher compression ratio of 12:1, which in return results in higher distortion of the binaural cues and could lead to worse localization (see Figure 4.3, E-F). This idea has also been hypothesized, at least for children, when comparing localization performance between different manufacturers (Killan et al., 2019). In other studies, results similar to sigmoid curves or at higher levels a bit more towards the mid-line have been presented (e.g., Dorman et al., 2014; Seeber and Fastl, 2008; van Hoesel and Tyler, 2003).

The path inside the model (see Figure 4.7) until reverse localization occurs or to which extent it is represented in each stage of the model can also be followed. Starting with the ILDs after electric output compression, as described before, the more apical channels are inverted.
Keeping the spread of excitation in mind, the maximum spike rate appears for high-frequency ANs. The highest rate appears in higher frequencies as the signal level for those channels is generally higher and more channels are around the corresponding AN fibers. The current spread is stronger there before being reduced exponentially. The reverse here can be seen for ANs corresponding to up to 500 Hz. This is approximately how deep the electrode array reaches in the cochlea in the model, and the nerve fibers are only stimulated through spread of excitation. Since these are further away from the electrodes with positive ILDs, the effect of the reverse ILD channels or those of middle frequencies with a bigger amplitude in each electrode is the biggest. Consequently, the contralateral side shows a higher spike rate at these positions.
Then, for each side, the ipsi- and contralateral sides are counted with a stronger

weighting on the inhibition to compute the EI neurons' spike rates. The stronger inhibition stems from the spiking properties in the inhibitory fibers as they are higher than for the excitatory fibers on average. Since the number of inhibitory fibers is also lower, additional weighting is applied but overall the spike rate for the inhibition is stronger.
Continuing with the example at -60° (see Figure 4.7, F) at higher levels, most high-frequency neurons show a similar spike rate for both sides. Still, the ipsilateral side has a slightly higher spike rate. This explains why the results here are mostly around zero, but the ipsilateral side still is slightly higher. Towards the apex, the contralateral side with higher inhibition leads to a spike rate that is not as balanced as before, meaning that this combination leads to a stronger reduction in spikes. Even with stronger weighting the ipsilateral side does not seem to have as big of an effect in reducing the spike rate. The amount of compression applied at each level might also explain the shift of maxima as the level increases. For the lowest level, the AGC is still inactivated and the ILD cues are not distorted. The contralateral side has a lower spike rate because of the head-shadow effect and with the weighting for computing EI neuron spike rates, the ipsilateral side has such a high spike rate in comparison.
Finally, the decision model for horizontal localization computes each azimuth by taking the hemispheric rate difference. This means that first the mean of each side for all EI neurons is taken. Then, the difference between both sides is calculated. The results are normalized and interpolated to an azimuth range from -90° to +90° with an increment of 5°. Since for the highest levels, the maxima at the EI stage are at the low-frequency components, the results are inverted. Since the rate difference is not big, the results are also centered around 0°. For the lower levels, it is clear which side has a higher spike rate and, thus, towards which direction the localization is predicted. This also explains why the outcomes at 40 dB SPL are pushed to ±90°.

So far, to my knowledge, there is no literature showing AN and LSO recording confirming the middle stages. It is clear from other studies (e.g., Dorman et al., 2014; Dwyer et al., 2021; Gray et al., 2021) that reversed ILDs appear at higher levels with AGC compression, as discussed before. For the localization results, high levels such as 80 dB SPL are usually not tested as it becomes an uncomfortably loud perception but rather at levels slightly lower and still above the AGC knee-point. Still, reverse localization is not a phenomenon commonly reported. The results for the lower levels can be represented by averaged data of previous studies at least to some degree (e.g., Dorman et al., 2014; van Hoesel and Tyler, 2003).

Since reversed localization did not appear, the AGC might need to be modified to predict an average or an individual CI user's performance better. One possibility could be that, especially for the very high levels that exceed the AGC knee-point by far, a fast-acting AGC processor could be needed inside the model in addition to the slow-acting AGC. As most CI processors use dual-loop AGCs using both fast and slow-acting time constants, it might be necessary to also include that inside the model. Using fast time constants showed no inverted ILDs after the first processing stage for any frequency band and combining the results with the slow constants, where a reversal in sign appears for apical channels, might lead to absolute ILDs in the inverted channels that are slightly lower than the correct ILDs at high-frequency channels. The slow AGC compressor is also known to distort the ILD cues stronger (e.g., Archer-Boyd and Carlyon, 2019; Pastore et al., 2021).
Additionally, the distribution of AN fibers with different spike rate properties might need to be analyzed as so far the inhibition is stronger overall and is one reason for the inverted localization predicted in the model. More balanced properties for both excitatory and inhibitory fibers could result in a more "natural" localization at high levels.
The last reason for reverse localization could lie within the decision stage as the hemispheric rate difference has not been directly confirmed as a stage within the auditory pathway. It is possible that this decision stage is currently too simple to correctly predict localization. In Kelvasa and Dietz (2015), where this stage of the model was first introduced, the LSO rate difference model was recommended as a stage that fits best for most applications of the model. For a more complex decision stage, e.g., an artificial observer using feedback as another input into the decision stage could be implemented (Dietz et al., 2018).

5.2. Differences between Psychoacoustic and Model Results

Comparing the experimental and simulated results, it can be seen that the results for each subject are fitted well to at least one side of the processors or in between both. It is not entirely clear which parameter is responsible the most or determines which side dominates the other. One theory by Anderson et al. (2023) is that the side with the bigger dynamic range dominates at least for speech perception. Linking that to the present study could describe the localization results for both CI users. For CI1 a bias towards the right side and CI2 a bias towards the left side can be seen (see Figure 5.1). For both, the corresponding side has a

bigger dynamic range, therefore agreeing with the idea by Anderson et al. (2023).

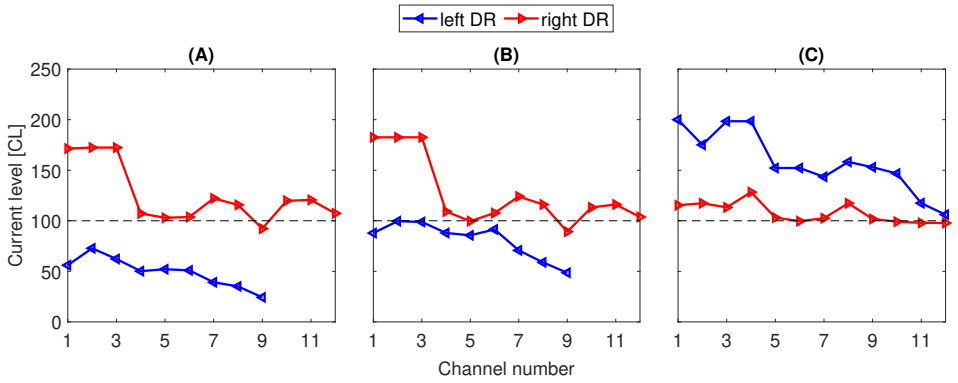

Figure 5.1.: Dynamic ranges for (A) CI1 with old processors, (B) CI1 with new processors, and (C) CI2. The dotted lines represent the model default.

In general, it can also be related to the number of activated channels in both ears. Since there is a mismatch in CI1 on the left side, the spread of excitation or simply the stimulation of the AN might not be as strong as the other side with a full range of activated channels. Though, for the left side with fewer activated electrodes, the phase duration was also increased, possibly compensating for the loss of electrodes. To my knowledge, this has not been shown in previous literature so far. It would be interesting to test in the future how one side having fewer electrodes available would behave inside the model.
Asymmetric modeling for this case should be possible within the current model if some parts are changed. So far, after the CI processing stage, the electrodogram is given as the next input as one variable. Simply saving both sides with different names should solve the problem.

The processor of each CI user that had similar THR values to the default settings follows a similar pattern because most of the other parameters already matched with the default. The processors with very low THR values show a more extreme localization to the sides. A low THR value will result in a smaller output amplitude in the electrodogram. It could explain that the results with lower THR values are more similar to the results of lower levels in general. The next stages therefore can be related to the patterns at 40 dB SPL (see Figure 4.7, H). From analyzing the influence of levels on the output at different stages, a lower level causes a shift of the maximum spike rate to higher frequencies and away from the frequencies at which the contralateral side previously showed a higher spike rate (see Figure 4.7, F).

Following the idea now that one side has a bigger influence on the localization but not being the only influence, the overall localization could be the result of both sides with a higher weighting on the dominant side (Dorman et al., 2014). The question is then how much weight each side has and how to determine that. Since general asymmetric modeling including both individual processors is not included in the model so far, the difference is hard to determine.

Inside the model framework, only parameters in the CI processing stage were changed, and not in the human model. This implies that the same CI user was simulated in the model but stimulated with different processors. Comparing the model to the experimental results (see Table 4.2) shows that the predicted results still have an average RMSE of around at least 20°. For better results, the model would also need to be individualized but this was not possible at the current stage as the participants' properties within the central processing stages are not known.

In general, for all levels, the AGC was activated for both CI users and could not be manually switched off during the experiments. For CI2, the results are mostly localized near ±90° so that the predicted results without AGC processing seem to be a better fit overall. It is difficult to determine why the psychoacoustic results tend to align with the default simulations more in some cases for participant CI1. If one would consider an average result between both sides with a weighting added to it, the results with activated AGC might represent the experimental data better.
Comparing the results now across stimuli for each CI user shows that for almost all stimuli the model results fit better for CI2. Again, this might be because those are already mostly at ±90°. The outcome for PKHP is the most similar at least for CI2. Here even the default model predicted localization to the sides of the horizontal plane already.

5.3. Inter- and Intra-Subject Comparison

Comparing the test and re-test results using a paired t-test showed almost no significant differences. Overall, for CI1 the results became slightly worse, and for CI2 better. For the old and new processors, the settings were first imported and then fitted again, explaining the similarity in the results. Most parameters remained the same. It is difficult to compare the results by only considering the processor types as the dynamic range seems to play the most important role. One difference in the hardware is the microphone placement as in the OPUS

2 processors it is placed almost directly at the front and in the SONNET 2 a little behind the pinna (MED-EL Elektromedizinische Geräte GmbH, 2022). This change was also tested by applying different HRIRs to the model but that also showed no considerable difference. One reason not related to the hardware might also be that the CI user was not accustomed to the new processors and would need more time to reach the previous localization results.
For CI2 the test and re-test results also changed significantly for some cases and improved mostly. This could be due to the CI user being familiar with the experimental setup. CI2's localization curve can be described almost as a sigmoid curve which is similar to other localization literature (e.g., Dorman et al., 2014; Seeber and Fastl, 2008) whereas CI1's almost as a linear function. Comparing that to the findings of the model might explain why the results of CI1 were better even though CI2's dynamic range almost covered the entire range. The increase in dynamic range in general could result in steeper sigmoid curves with almost extreme bias to one side for all angles at higher levels. The dynamic range was bigger in general for more channels and especially those that had a higher impact on the results. This might be the reason for the extent of how well a CI user can correctly localize. Increasing the THR values again for some channels might lead to better localization results but still impact other aspects of their hearing performance such as speech understanding or hearing in quiet environments.
The result for the OLSA word and PK were similar as the spectra were also comparable.

5.4. Dynamic Range Dependency for Localization

Taking a closer look at exclusively the change of dynamic range over different stages showed an insight into how it has the biggest impact on localization compared to other parameters tested here. As can be seen before (see Figure 4.6), increasing the dynamic range in one channel already removed the reversal in localized perception and made it similar to the literature (e.g., Dorman et al., 2014; van Hoesel and Tyler, 2003). Because of this, the next steps were only looked at for one channel.
First at the stage of CI processing after electric output computation, the ILDs showed no difference as the ratio between the ipsi- and contra-lateral side did not change. In the case that the THR value for channel eight was reduced to 80 CL, the overall level at each side was only reduced by a few dB at high levels, where the distortion occurs. This was then expressed as an increase in spike rate in the AN fibers but also as a decrease for the contralateral side in the EI neurons. The peaks and drops were still mostly around the same area.

To analyze the effect of changing the THR or MCL values, one would have to take a look at the loudness growth function. Having a lower threshold means that mostly the lower outputs change as the higher ones approach a limit due to the logarithmic behavior. Theoretically, since the levels were only checked at higher levels, most of the inputs would also be at a higher level. The LGF in the model is currently also fitted to the manufacturer Cochlear Ltd. which ranges from 25 to 65 dB SPL. This means that most of the changes in THR values or dynamic range for the levels tested are mostly not big. Comparing the electrodograms shows that the decrease in THR value results in a difference that appears regarding the transmission of the signal to the next stage.
Having a default dynamic range from e.g. 50 to 200 CL (see Figure 4.4, H-I) results in a predicted localization more similar to previous literature (e.g., Dorman et al., 2014; Seeber and Fastl, 2008; van Hoesel and Tyler, 2003). Comparing that to level values set in other literature shows that the dynamic ranges are usually not as big as suggested (see Figure 5.2). Gray et al. (2021) even stated specifically, that a range from 50 to 200 CL is unusually big. It is therefore unclear which values most accurately describe the behavior of an average CI user as the level values THR and MCL are fitted differently for every individual. Only the THR values might need to be changed as the default MCL values and other literature are typically around 200 CL. Another study by Gajecki and Nogueira (2021) directly presented a table including all channel values for the THR and MCL parameters. All subjects there were using Cochlear Ltd. devices and even had a dynamic range smaller than 100 CL.

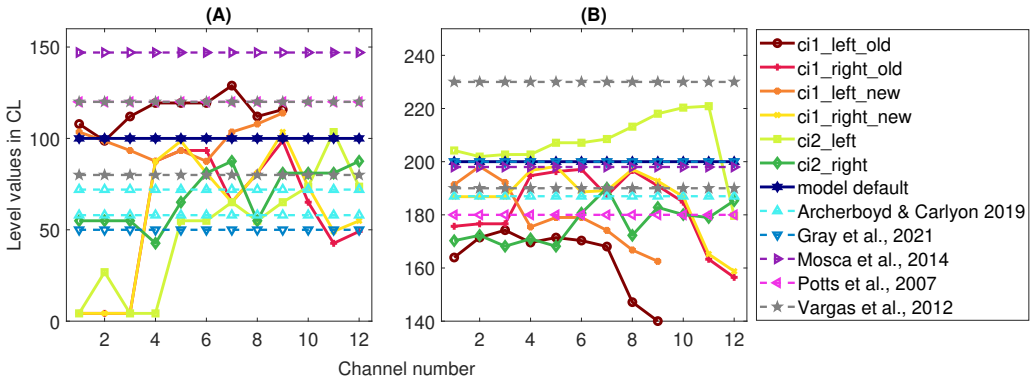

Figure 5.2.: (A) THR values and (B) MCL values compared for the CI participants and literature.

Comparing the values from each study in Figure 5.2, there might be a manufacturer-dependent difference as the lowest values for subjects were reported for studies done by Archer-Boyd and Carlyon (2019) and Vargas et al. (2012). Both used MED-EL devices and the lower values can also be related to the values read out

for this thesis. The study by Mosca et al. (2014) had the highest THR levels similar to the data shown by Gajecki and Nogueira (2021), both tested with Cochlear Ltd. devices.
Overall, according to the THR values tested by decreasing it systematically by 10 CL each time, it might be necessary for the model as a temporary solution to avoid reverse predicted localization to decrease it to around 60 to 70 CL in all or just in a single channel out of the lower or especially middle-frequency channels.

5.5. Limitations and Future Directions

There were several limitations at the current stage which would need to be changed for a more solid understanding of all experiments. On the one hand, the number of participants was too low to find more general conclusions and on the other hand, the model is still kept simple as asymmetric hearing has not been included yet. Also having asymmetric modeling at least for the case of CI1 would help understand the effect of mismatch in the number of channels on localization. It would also help in understanding which parameters determine the directional bias.
Another part of the model could also include a dual-loop AGC usually found in most devices nowadays. The current model only contains the option of changing attack and release time in one type of compressor and it could only be changed from a slow to a fast compressor or vice versa. If the absolute value of ILDs in inverted channels could lead to localization being more towards the contralateral side, the AGC could also have a different effect when including a dual-loop processor. When both slow and fast time constants were tested, the results showed a reverse in low-frequency channels as shown in multiple studies (e.g., Dorman et al., 2014; Dwyer et al., 2021; Gray et al., 2021) for a slow processor and only all with the correct sign for the fast processor. Therefore the absolute value of the inverted ILDs might be reduced. Here, it could also be helpful to compare the model results to ILDs directly read out from real processors to compare. This was done in the study done by Gray et al. (2021) but only for processors from Cochlear Ltd. and not for processors by MED-EL as used here. Since that study showed almost no reverse ILDs, it would also be interesting to test the same conditions for different coding strategies, such as the n-of-m strategy.
It would also be interesting to analyze the correlation between the THR values in each frequency channel and the resulting localization. Using that information would also clarify where to set the default value to predict the average CI user's localization. Additonally, changing different spiking properties for the excitatory and inhibitory fibers in the binaural interaction stage could improve the model's

default output to fit better to an average CI user and might describe an individual CI user's psychoacoustic results better.
Regarding the individual channel behavior, it might be interesting to test this directly in CI users doing direct stimulation experiments. Here, the dynamic range could be reduced or the THR elevated to such an extent that accurate localization would not be possible anymore. Then, if possible, the intensity of stimulation in one channel would be changed each time and then observed.

6. Conclusion and Outlook

The effect of AGC on the spatial perception of bilateral CI users was analyzed in this thesis. For this, a series of experiments were conducted testing the amount of binaural fusion and localization after AGC compression and compared to a bilateral CI model.
First, the psychoacoustic data neither showed reverse localization nor non-fusion in any condition. Compared to the average RMSE in other studies, both CI users localized well. Connecting this to the study done by Suneel et al. (2017) indicates that binaural fusion should always occur for these CI users even though one subject reported non-fusion in general scenarios such as having a conversation. To test whether non-fusion could happen after AGC compression, CI users would need to be tested that do not perform well on the localization task. The number of participants, in general, was too low to confirm the hypotheses. At least for the localization tasks, it was possible to determine the similarities compared to other CI users.
The model gave a deeper insight into the processing of AGC compression at higher levels and the resulting localization. The model first showed the inverted ILDs in low-frequency channels predicted in other studies. Then, AN and LSO spike rates were analyzed for both ears, and the contralateral side showed a higher spike rate than the ipsilateral side for the lower frequencies up to 500 Hz. The default simulations predicted localization centered around 0°, and in the initial results, reverse localization which compared to other studies does not usually occur. It was found that the reversed predicted perception can be traced back to every stage of the model. Since there are no AN and LSO recordings for humans after AGC processing, it is hard to determine which parameter directly caused these results. Possible changes within the model could be different weighting factors for the binaural interaction stage or a change in the decision model.

Additionally, the model indicated other factors that affect spatial hearing. The dynamic range was shown to have the largest influence on the range of localization. Increasing it resulted in better localization performance. Another aspect is that the psychoacoustic results were found to be typically between the predicted outcomes using one side of the processors. Following a study by Anderson et al. (2023) that claimed the dynamic range to indicate a directional bias, the psychoacoustic results could be described as combining both sides with a different amount of influence each.
Another finding in the model was the channel-dependent behavior influencing localization. Here, the results changed differently for each channel when the dynamic range was altered.

The findings of this thesis suggest some modifications for the model to give a better prediction of both average and individual CI users. These changes include a dual-loop AGC system and more balanced excitation and inhibition during the binaural interaction stage. The results also show that more psychoacoustic data is needed especially for the binaural fusion task. For this, the same setup as described in this thesis can be used, and an analysis similar to Suneel et al. (2017) can be conducted to find a correlation between the amount of binaural fusion and localization performance.

In conclusion, the effect of AGC on the individual CI user's spatial hearing can be summed up as follows:

1. Activated AGCs distort the ILD cues and result in distorted localization to some degree but should not result in reversed localization.

2. CI users that perform well in localization tasks could present full binaural fusion in such experiments.

3. The effect of AGC compression can be seen directly in every stage of the model.

4. The dynamic range affects the range of localization the most.

5. For each frequency channel, the amount of influence on localization results is different and the middle channels have the biggest influence.

Bibliography

Anderson, S. R., Gallun, F. J., & Litovsky, R. Y. (2023). Interaural asymmetry of dynamic range: Abnormal fusion, bilateral interference, and shifts in attention. *Frontiers in Neuroscience*, *16*. https://doi.org/10.3389/fnins.2022.1018190

Archer-Boyd, A. W., & Carlyon, R. P. (2019). Simulations of the effect of unlinked cochlear-implant automatic gain control and head movement on interaural level differences. *The Journal of the Acoustical Society of America*, *145*(3). https://doi.org/10.1121/1.5093623

Ashida, G., Kretzberg, J., & Tollin, D. J. (2016). Roles for coincidence detection in coding amplitude-modulated sounds. *PLOS Computational Biology*, *12*(6), e1004997. https://doi.org/10.1371/journal.pcbi.1004997

Bacon, S. P., Fay, R. R., & Popper, A. N. (2004). Compression: From cochlea to cochlear implants. https://doi.org/https://doi.org/10.1007/b97241

Blauert, J. (1996). *Spatial hearing : The psychophysics of human sound localization*. The MIT Press. https://doi.org/https://doi.org/10.7551/mitpress/6391.001.0001

Boyle, P. J., Büchner, A., Stone, M. A., Lenarz, T., & Moore, B. C. J. (2009). Comparison of dual-time-constant and fast-acting automatic gain control (agc) systems in cochlear implants. *International Journal of Audiology*, *48*(4), 211–221. https://doi.org/10.1080/14992020802581982

Brown, C. A. (2018). Corrective binaural processing for bilateral cochlear implant patients. *PLoS One*, *13*(1). https://doi.org/https://doi.org/10.1371/journal.pone.0187965

Denk, F., Ernst, S. M. A., Ewert, S. D., & Kollmeier, B. (2018). Adapting hearing devices to the individual ear acoustics: Database and target response correction functions for various device styles. *Trends in Hearing*, *22*. https://doi.org/10.1177/2331216518779313

Dhanasingh, A., & Hochmair, I. (2021). Signal processing & audio processors. *Acta Oto-Laryngologica*, *141*(sup1), 106–134. https://doi.org/10.1080/00016489.2021.1888504

Di Lella, F. A., Parreño, M., Fernandez, F., Boccio, C. M., & Ausili, S. A. (2020). Measuring the electrical status of the bionic ear. re-thinking the impedance

in cochlear implants. *Frontiers in Bioengineering and Biotechnology*, *8*. https://doi.org/10.3389/fbioe.2020.568690

Dietz, M. (2016). Models of the electrically stimulated binaural system: A review. *Network: Computation in Neural Systems*, *27*(2-3), 186–211. https://doi.org/10.1080/0954898X.2016.1219411

Dietz, M., Lestang, J.-H., Majdak, P., Stern, R. M., Marquardt, T., Ewert, S. D., Hartmann, W. M., & Goodman, D. F. (2018). A framework for testing and comparing binaural models [Computational models of the auditory system]. *Hearing Research*, *360*, 92–106. https://doi.org/https://doi.org/10.1016/j.heares.2017.11.010

Dorman, M. F., Loiselle, L., Stohl, J., Yost, W. A., Spahr, A., Brown, C., & Cook, S. (2014). Interaural level differences and sound source localization for bilateral cochlear implant patients. *Ear and Hearing*, *35*(6), 633–40. https://doi.org/10.1097/AUD.0000000000000057

Dorman, M. F., Loiselle, L. H., Cook, S. J., Yost, W. A., & Gifford, R. H. (2016). Sound source localization by normal-hearing listeners, hearing-impaired listeners and cochlear implant listeners. *Audiology and Neurotology*, *21*(3), 127–31. https://doi.org/10.1159/000444740

Dwyer, R. T., Chen, C., Hehrmann, P., Dwyer, N. C., & Gifford, R. H. (2021). Synchronized automatic gain control in bilateral cochlear implant recipients yields significant benefit in static and dynamic listening conditions. *Trends in Hearing*, *25*, 23312165211014139. https://doi.org/10.1177/23312165211014139

Ewert, S. (2013). Afc - a modular framework for running psychoacoustic experiments and computational perception models. *Proceedings of the International Conference on Acoustics AIA-DAGA 2013*, 1326–1329. http://medi.uni-oldenburg.de/afc/manual.htm

Fredelake, S., & Hohmann, V. (2012). Factors affecting predicted speech intelligibility with cochlear implants in an auditory model for electrical stimulation. *Hearing Research*, *287*(1-2), 76–90. https://doi.org/https://doi.org/10.1016/j.heares.2012.03.005

Gajecki, T., & Nogueira, W. (2021). Enhancement of interaural level differences for bilateral cochlear implant users. *Hearing Research*, *409*, 108313. https://doi.org/https://doi.org/10.1016/j.heares.2021.108313

Gardner, W. G., & Martin, K. D. (1995). HRTF measurements of a KEMAR. *The Journal of the Acoustical Society of America*, *97*(6), 3907–3908. https://doi.org/10.1121/1.412407

Giannoulis, D., Massberg, M., & Reiss, J. D. (2012). Digital dynamic range compressor design—a tutorial and analysis. *Journal of The Audio Engineering*

Society, *60*, 399–408. https://www.eecs.qmul.ac.uk/~josh/documents/2012/GiannoulisMassbergReiss-dynamicrangecompression-JAES2012.pdf

Grantham, D. W., Ashmead, D. H., Ricketts, T. A., Labadie, R. F., & Haynes, D. S. (2007). Horizontal-plane localization of noise and speech signals by postlingually deafened adults fitted with bilateral cochlear implants. *Ear and Hearing*, *28*(4), 524–41. https://doi.org/10.1097/AUD.0b013e31806dc21a

Gray, W. O., Mayo, P. G., Goupell, M. J., & Brown, A. D. (2021). Transmission of binaural cues by bilateral cochlear implants: Examining the impacts of bilaterally independent spectral peak-picking, pulse timing, and compression. *Trends in Hearing*, *25*, 23312165211030411. https://doi.org/10.1177/23312165211030411

Hu, H., Ausili, S. A., Williges, B., Klug, J., Felsheim, R. C., Vickers, D., & Dietz, M. (2023). Matlab code and results for the manuscript submitted to acta acustica 'a model framework for simulating spatial hearing of bilateral cochlear implants users' (version 1.0). https://doi.org/https://doi.org/10.5281/zenodo.7471961

Hu, H., Klug, J., & Dietz, M. (2022). Simulation of itd-dependent single-neuron responses under electrical stimulation and with amplitude-modulated acoustic stimuli. *Journal of the Association for Research in Otolaryngology*, *23*(4), 535–550. https://doi.org/10.1007/s10162-021-00823-1

Kayser, H., Ewert, S. D., Anemüller, J., Rohdenburg, T., Hohmann, V., & Kollmeier, B. (2009). Database of multichannel in-ear and behind-the-ear head-related and binaural room impulse responses. *EURASIP Journal on Advances in Signal Processing*, *2009*(1), 298605. https://doi.org/10.1155/2009/298605

Kelvasa, D., & Dietz, M. (2015). Auditory model-based sound direction estimation with bilateral cochlear implants. *Trends in Hearing*, *19*. https://doi.org/10.1177/2331216515616378

Killan, C., Scally, A., Killan, E., Totten, C., & Raine, C. (2019). Factors affecting sound-source localization in children with simultaneous or sequential bilateral cochlear implants. *Ear and Hearing*, *40*(4), 870–877. https://doi.org/10.1097/aud.0000000000000666

Klug, J., Schmors, L., Ashida, G., & Dietz, M. (2020). Neural rate difference model can account for lateralization of high-frequency stimuli. *The Journal of the Acoustical Society of America*, *148*(2), 678–691. https://doi.org/10.1121/10.0001602

Litovsky, R. Y., Goupell, M. J., Fay, R. R., & Popper, A. N. (2021). Binaural hearing : With 93 illustrations. https://doi.org/https://doi.org/10.1007/978-3-030-57100-9

MED-EL Elektromedizinische Geräte GmbH. (2020). *Maestro system software 9.0*.

MED-EL Elektromedizinische Geräte GmbH. (2022). *Audioprozessoren im vergleich* [Accessed: 15.05.2023]. https://www.medel.com/de/hearing-solutions/cochlear-implants/audio-processor-comparison

Mosca, F., Grassia, R., & Leone, C. A. (2014). Longitudinal variations in fitting parameters for adult cochlear implant recipients. *Acta otorhinolaryngologica Italica*, *34*(2), 111–116. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4025185/

Pastore, M. T., Pulling, K. R., Chen, C., Yost, W. A., & Dorman, M. F. (2021). Effects of bilateral automatic gain control synchronization in cochlear implants with and without head movements: Sound source localization in the frontal hemifield. *Journal of Speech, Language, and Hearing Research*, *64*(7), 2811–2824. https://doi.org/10.1044/2021_jslhr-20-00493

Reiss, L. A. J., Fowler, J. R., Hartling, C. L., & Oh, Y. (2018). Binaural pitch fusion in bilateral cochlear implant users. *Ear and Hearing*, *39*(2), 390–397. https://doi.org/10.1097/aud.0000000000000497

Santala, O., & Pulkki, V. (2011). Directional perception of distributed sound sources. *The Journal of the Acoustical Society of America*, *129*(3), 1522–1530. https://doi.org/10.1121/1.3533727

Seeber, B. U., & Fastl, H. (2008). Localization cues with bilateral cochlear implants. *The Journal of the Acoustical Society of America*, *123*(2), 1030–42. https://doi.org/10.1121/1.2821965

Shaw, E. A. (1974). Transformation of sound pressure level from the free field to the eardrum in the horizontal plane. *The Journal of the Acoustical Society of America*, *56*(6), 1848–61. https://doi.org/10.1121/1.1903522

Suneel, D., Staisloff, H., Shayman, C. S., Stelmach, J., & Aronoff, J. M. (2017). Localization performance correlates with binaural fusion for interaurally mismatched vocoded speech. *The Journal of the Acoustical Society of America*, *142*(3). https://doi.org/10.1121/1.5001903

Swanson, B. A. (2008). *Pitch perception with cochlear implants* (Thesis) [PhD thesis, Faculty of Medicine, Dentistry & Health Sciences, Otolaryngology Eye and Ear Hospital, The University of Melbourne.]. http://hdl.handle.net/11343/39587

U.S. Department of Health and Human Service. (2021). *Cochlear implants* [Accessed: 08.06.2023]. https://www.nidcd.nih.gov/health/cochlear-implants

Vaerenberg, B., Govaerts, P. J., Stainsby, T., Nopp, P., Gault, A., & Gnansia, D. (2014). A uniform graphical representation of intensity coding in current-

generation cochlear implant systems. *Ear and hearing*, *35*(5), 533–543. https://doi.org/10.1097/AUD.0000000000000039

van Hoesel, R. J., & Tyler, R. S. (2003). Speech perception, localization, and lateralization with bilateral cochlear implants. *The Journal of the Acoustical Society of America*, *113*(3), 1617–30. https://doi.org/10.1121/1.1539520

Vargas, J. L., Sainz, M., Roldan, C., Alvarez, I., & de la Torre, A. (2012). Long-term evolution of the electrical stimulation levels for cochlear implant patients. *Clinical and Experimental Otorhinolaryngology*, *5*(4), 194–200. https://doi.org/10.3342/ceo.2012.5.4.194

Wagener, K., Brand, T., & Kollmeier, B. (1999). Entwicklung und evaluation eines satztests für die deutsche sprache i: Design des oldenburger satztests. *Zeitschrift für Audiologie*, *38*(1-3), 4–15. https://www.uzh.ch/orl/dga-ev/publikationen/zfaudiologie/archiv/ZfA_1999_38-1_004-015_Original.pdf

Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K., & Rabinowitz, W. M. (1991). Better speech recognition with cochlear implants. *Nature*, *352*(6332), 236–238. https://doi.org/10.1038/352236a0

Zeng, F.-G. (2022). Celebrating the one millionth cochlear implant. *JASA Express Letters*, *2*(7). https://doi.org/10.1121/10.0012825

A. Appendix

A.1. Loudness scaling

Participant		CI1						CI2					
Monaural side		Left			Right			Left			Right		
Level in dB SPL		40	60	80	40	60	80	40	60	80	40	60	80
Azimuth in degree	-75	1	2	5	0	3	5	1	7	8	1	5	7
		1	3	6	1	3	7	2	7	8	1	5	7
	-45	1	2	5	1	3	5	1	5	8	1	5	7
		1	3	5	2	3	6	1	6	9	1	5	7
	0	1	2	4	1	3	5	1	5	8	1	5	7
		1	3	4	1	4	5	1	5	8	1	5	7
	+45	1	3	5	1	3	5	1	4	7	1	5	7
		1	3	5	2	4	5	1	5	8	1	5	7
	+75	1	3	6	1	3	6	1	4	7	1	5	7
		1	3	6	1	4	7	1	5	7	3	6	7

Table A.1.: Loudness scaling values for the monaural measurements. The range starts from zero ("not audible") to ten ("too loud").

A.2. Localization Results

The rest of the localization results are shown for both CI users. For both, the PKHP in comparison to the model is shown (see Figure A.1 and A.2) and the OLSA word (see Figure A.3) and GWN (see Figure A.4) were also plotted.

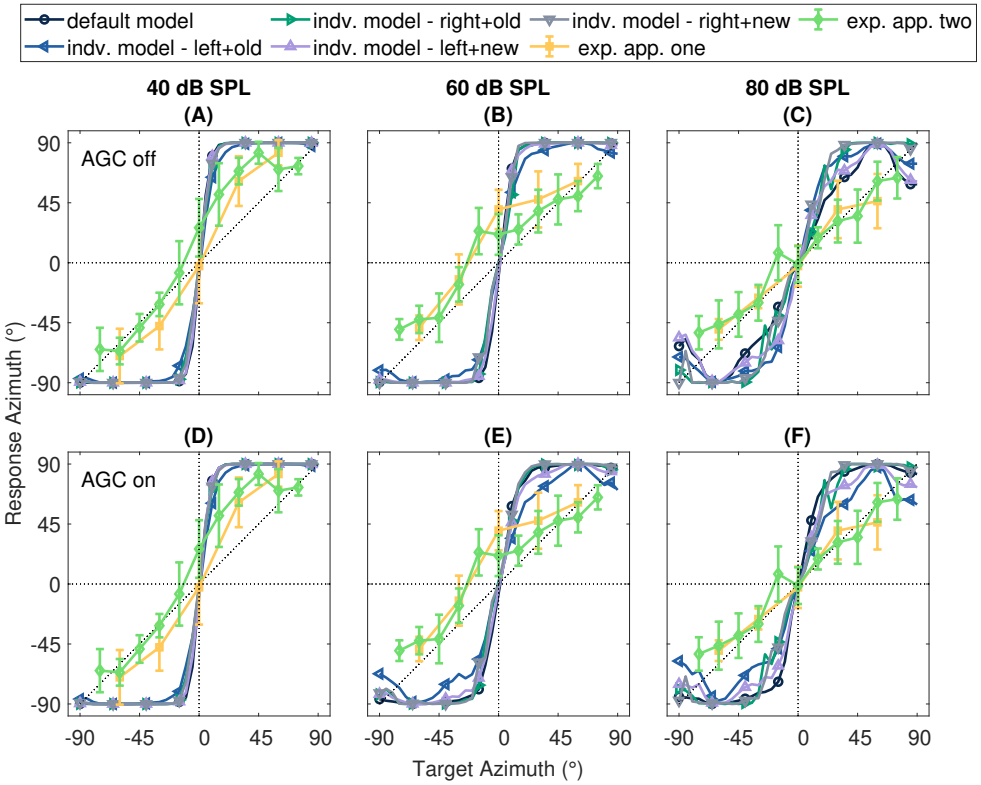

Figure A.1.: Response azimuth as a function of target azimuth shown for experimental and model results compared to the default model output for highpass filtered pink noise presented at 40, 60, and 80 dB SPL. The mean is shown for experimental data with the standard deviation. All parameters from CI1 from Table 3.1 were used. The diagonal dotted line represents the correct answers.

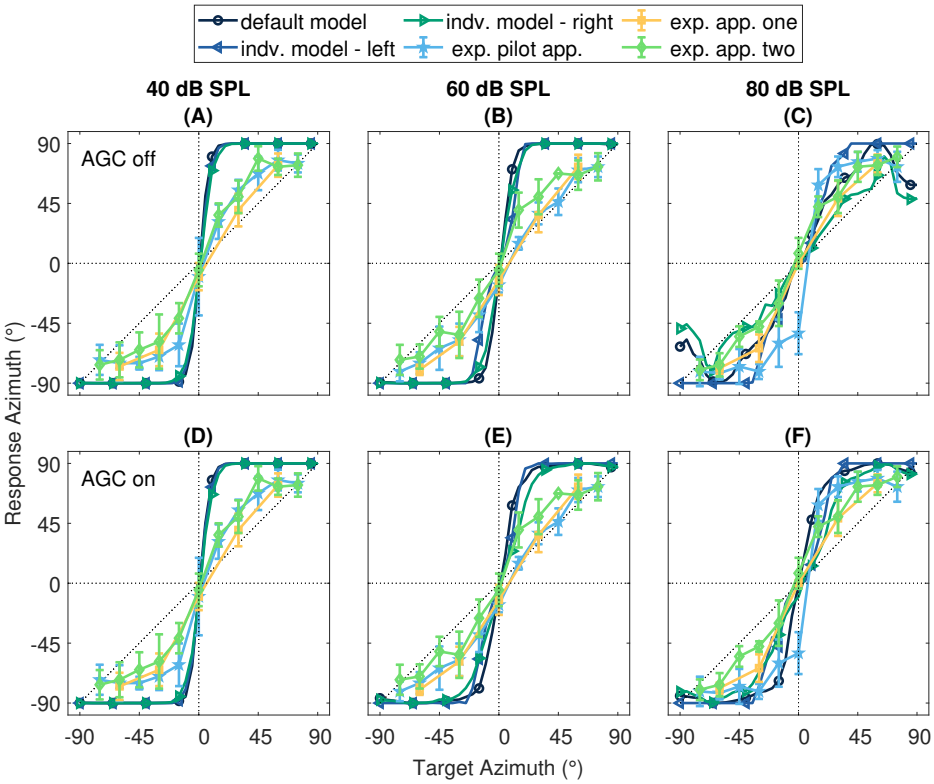

Figure A.2.: Localization results for CI2 with HP.

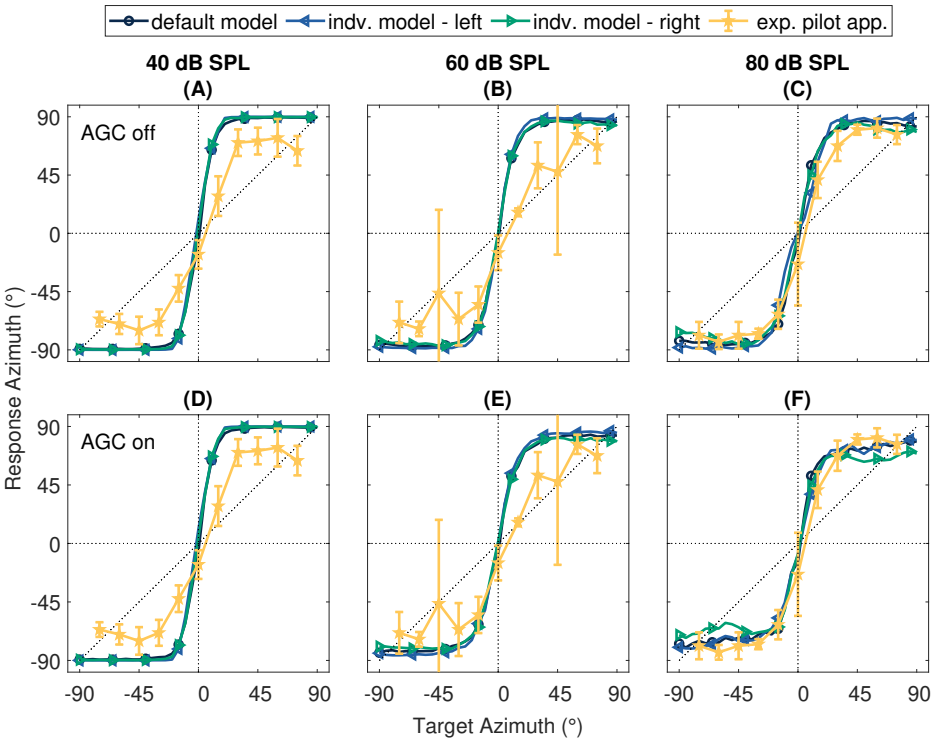

Figure A.3.: Localization results for CI2 with OLSA word.

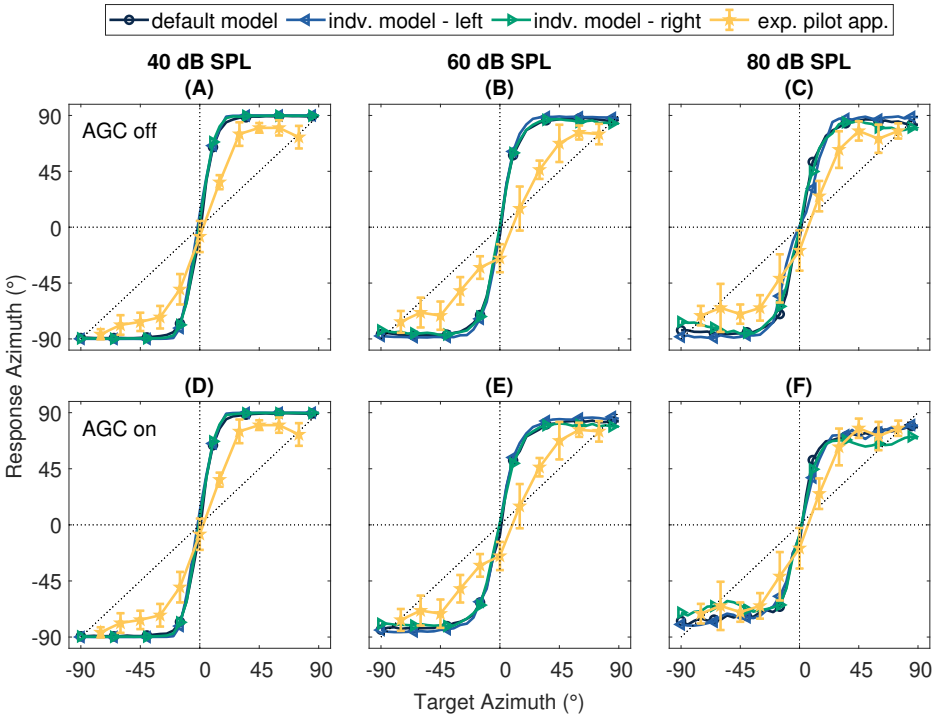

Figure A.4.: Localization results for CI2 with GWN.

A.3. Influence of Additional Parameters on the Model

Additional parameters apart from the ones in chapter 4.3 were analyzed in the model. These contained the frequency range for all channels, the phase duration of each pulse, the number of samples in pps, the total number of channels on each side, the number of channels with fine-structure processing, the coding strategy from FSx=4 to CIS, and the map-law value α_c (see Figure A.5).
Starting from the same order as the parameters were changed, a change in the number of channels and therefore the frequency range did not lead to different outcomes (see Figure A.5, A-C). The MED-EL devices had a range of up to 8500 Hz compared to the model with a range of up to 8000 Hz. When plotting the frequency filters logarithmically, the channel borders stayed around the same range. For CI1, whose processor only had nine active channels on the left side, the range was adapted and read into the model. This resulted in a small shift away from the center (see Figure A.5, C).
Following that, a change in phase duration or rate also did not change the default outcome (see Figure A.5, D-I). Regarding the stimulation rate for each channel, there was a large discrepancy in values. The general default parameters stayed the lowest at 900 pps compared to the individual CI users' around 1300 pps.
The same could be seen for the changed map-law value (see Figure A.5, J-L).
Comparing the influence of the number of channels with fine-structure processing, again, resulted in the same outcome. Changing the coding strategy to CIS showed results similar to higher compression and therefore reverse localization (see Figure A.5, M-O).

Since the first CI user changed from OPUS 2 to SONNET 2 processors, the location of microphones also slightly changed. Because of this, the related HRIRs were also replaced from the database. The one by Denk et al. (2018) included HRIRs for three positions on a BTE device. Changing the microphone position from front to middle to back on the behind-the-ear processor data only resulted in higher compression, with the maximum ILD shifted to different angles. The same results were predicted at the localization stage.
ttest significance

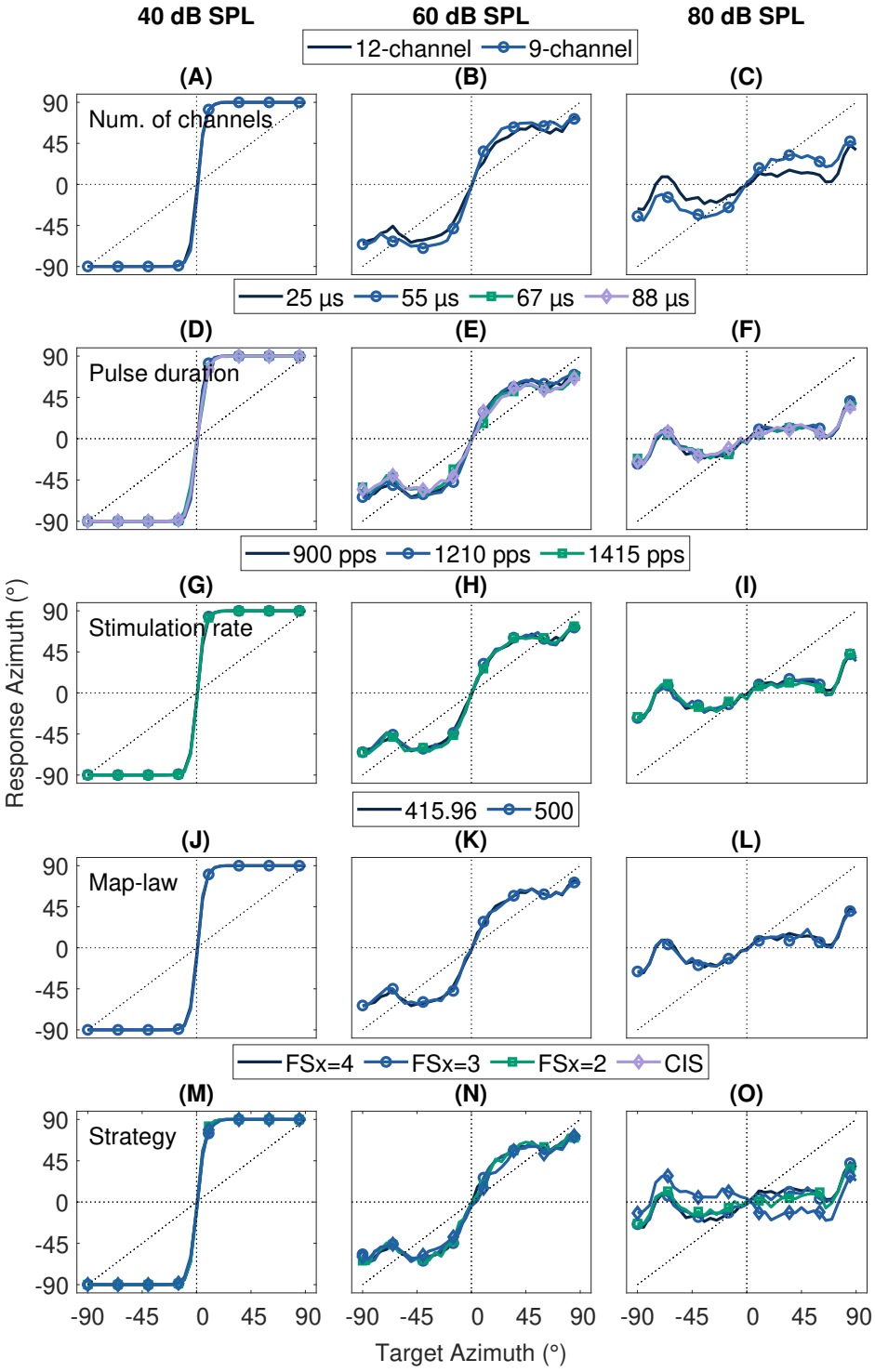

Figure A.5.: Localization results simulated with parameters taken from table 3.1. The results were shown at different levels for changed (A-C) numbers of channels, (D-F) pulse duration, (G-I) stimulation rates, (J-L) map-law constants and (M-O) coding strategies.

A.4. AN and EI spike rate differences

Here, the AN and EI spike rates are shown for the left and right sides for all levels:

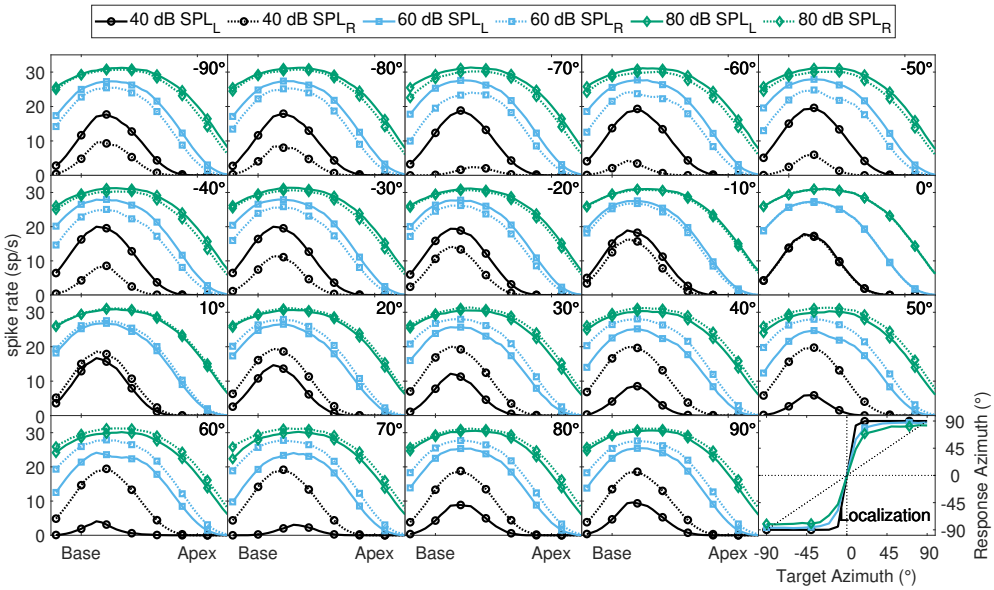

Figure A.6.: Auditory nerve spike rate with inactivated automatic gain control compression at different azimuth and levels. The model output for the corresponding localization is shown in the bottom right corner.

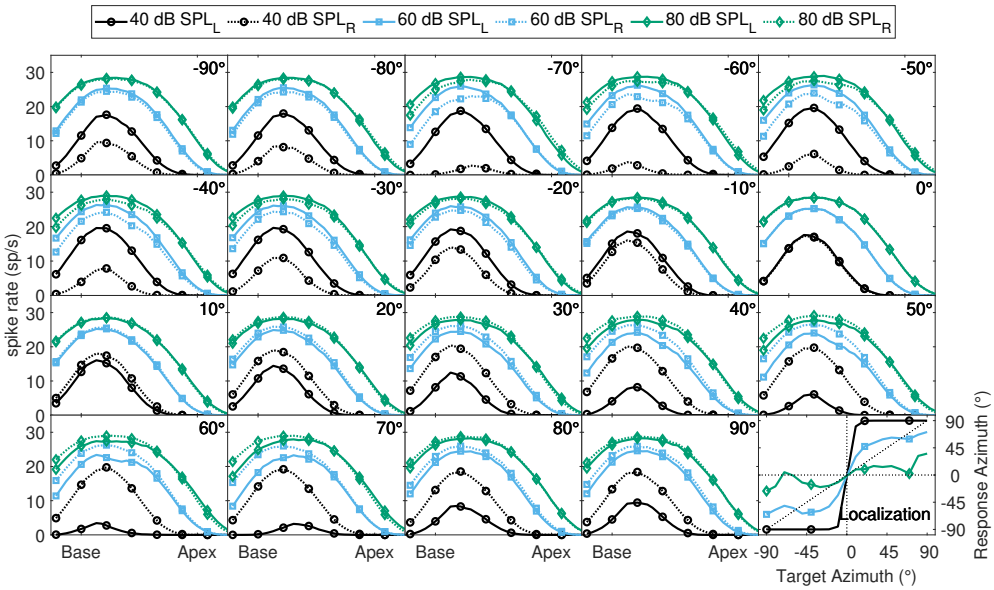

Figure A.7.: Auditory nerve spike rate with activated automatic gain control. The results are depicted similarly to Figure A.6.

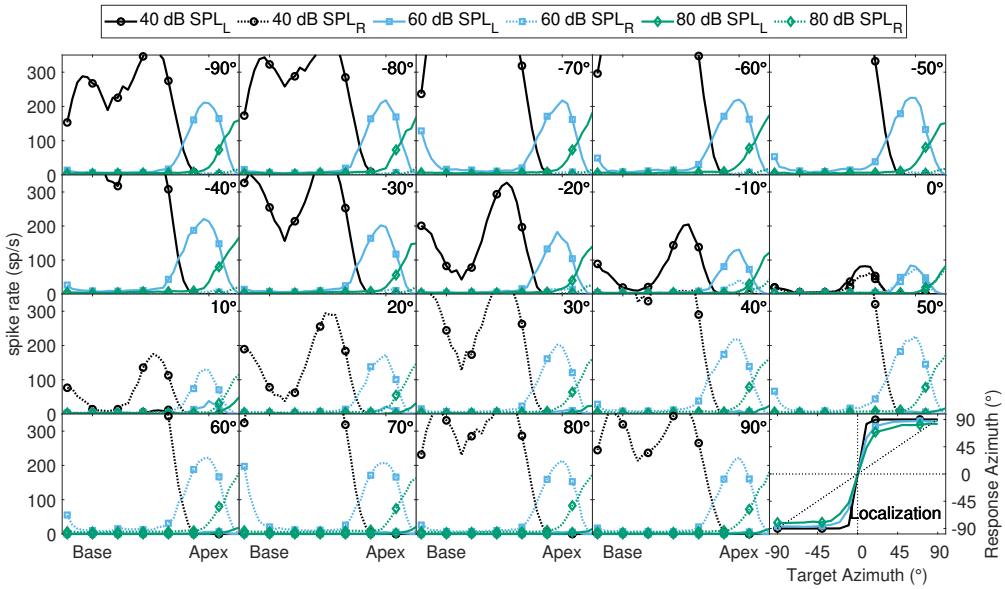

Figure A.8.: Lateral superior olive spike rate with inactivated automatic gain control. The results are depicted similarly to Figure A.6.

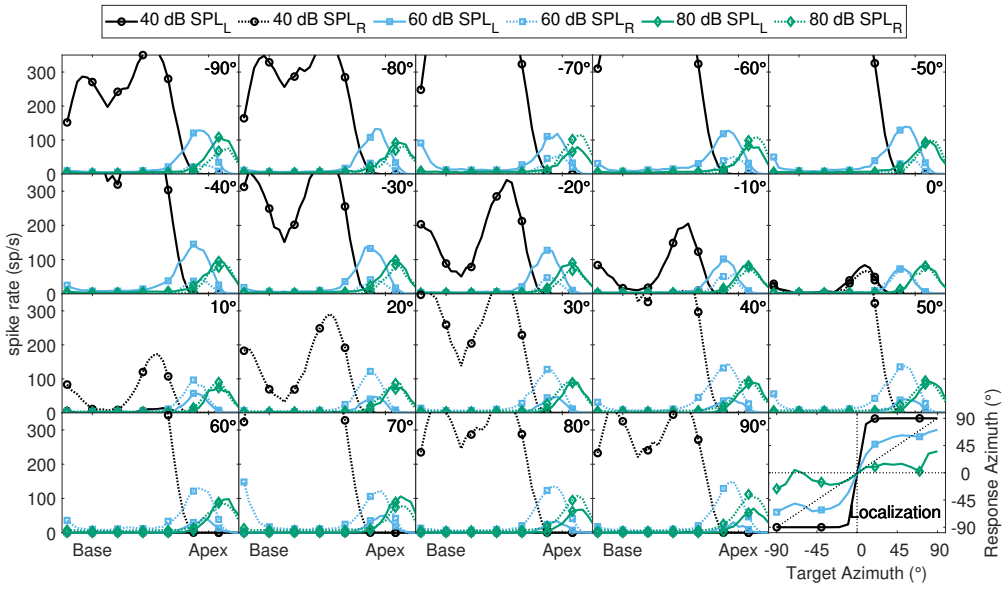

Figure A.9.: Lateral superior olive spike rate with activated automatic gain control. The results are depicted similarly to Figure A.6.

A.5. HRIR changes

In this section, different HRIR databases are compared to explain the reasons for replacing the default one inside the model. It is compared through results from previous literature as well as comparing selected ones directly through different stages of the model. The change was motivated since different literature (e.g. Dorman et al., 2014; Dwyer et al., 2021; Gray et al., 2021) showed different ILD amplitudes for different databases and the default on used in the model by Kayser et al. (2009) also applied in Dwyer et al. (2021) showed comparably large inverted ILDs even ranging up to middle-frequency channels. Since it was also proposed in Dorman et al. (2014) that the dominant ILDs might influence the direction of the localization, a reduction in reversed ILDs could lead to more similar results as in the literature.

The default database by Kayser et al. (2009) was compared to a different one from Oldenburg, by Denk et al. (2018), and also an older database by Gardner and Martin (1995). The latter was used in simulations by Dorman et al. (2014). When comparing all hardware differences taken from the documentation, one striking difference is the distance to the recording device. The oldest database was measured only at half the distance of the others (1.4 m). All of the databases measured HRIRs using a KEMAR, a manikin on which,e.g., hearing aids can be placed, instead of real participants. For both databases from Oldenburg, participants' data was also included but for comparison and having a more general HRIR the KEMAR was used.

In the oldest database, only one microphone for the BTE device was used instead of three in the rest. More microphone placements enabled change to compare if that also has an influence on the results since the different processors used by the CI users here also had different microphone placements.

All of them had their own way of processing the measured data but when comparing the ILDs across frequencies the biggest differences were mostly in the higher frequencies (see fig.A.10).

The most important difference though was the ILDs produced after processing all inside the model (see Figure A.11). The HRIRs resulting from Gardner and Martin (1995) provided ILDs after electric output compression which could not be computed at some azimuths but the overall results showed similar findings as in the literature. Reversed ILDs appeared in low-frequency channels at higher levels, though it seems here that the absolute amplitude of the reversed channels is higher in comparison to the rest. For the Kayser et al. (2009) database there is less of a difference in amplitude and a bigger in the newest.

As the end results showed that reverse localization was predicted for all but each

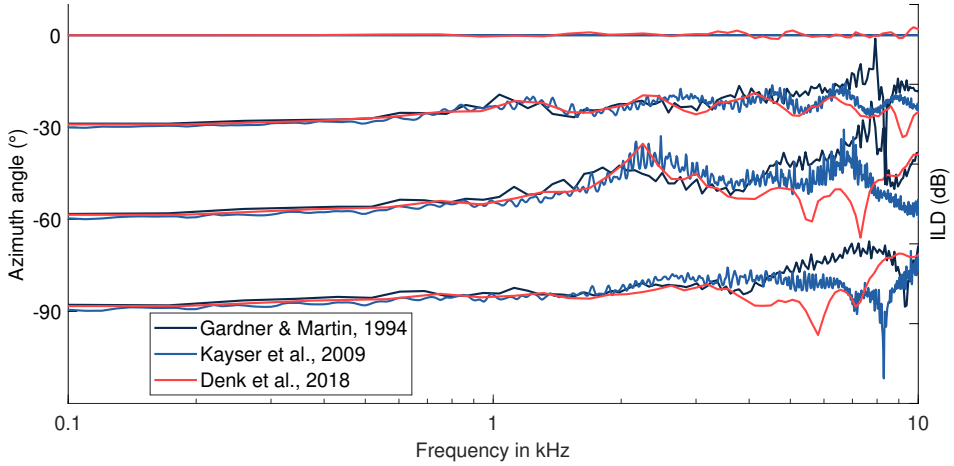

Figure A.10.: ILD comparison over frequency for different HRIRs.

to a different extent. Comparing that to the different stages before like in previous analyses in this thesis, the AN and EI spike rates show a bigger reverse when using the one by Kayser et al. (2009) and for Gardner and Martin (1995) the reverse is almost not even visible at even 100 dB SPL. The results here for the other levels are extremely compressed to the outward azimuth and also do not provide a wider range in microphone placements.
Since both older databases show different weaknesses in using them, the newest one was chosen to be used in the model. This was also because it showed lower reversed ILDs such as in other literature (e.g. Dorman et al., 2014; Gray et al., 2021).

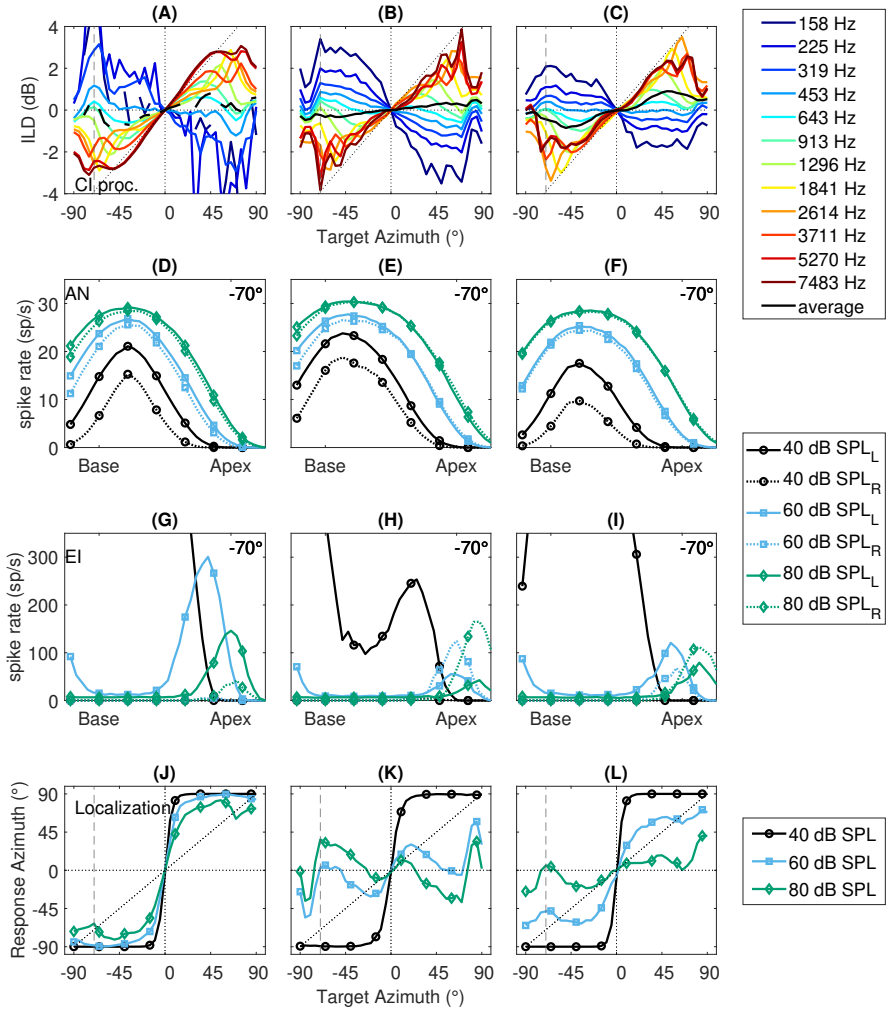

Figure A.11.: HRIR database comparison showing different stages of the model for Gaussian white noise. The first column (A, D, G, J) describes the one by Gardner and Martin (1995), the second (B, E, H, K) by Kayser et al. (2009) and the last (C, F, I, L) by Denk et al. (2018). It shows the model stages in the order: ILDs, AN spike rates, EI spike rates, and localization output.

Hiermit versichere ich an Eides statt, dass ich diese Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Außerdem versichere ich, dass ich die allgemeinen Prinzipien wissenschaftlicher Arbeit und Veröffentlichung, wie sie in den Leitlinien guter wissenschaftlicher Praxis der Carl von Ossietzky Universität Oldenburg festgelegt sind, befolgt habe.

Helen Kasim Oldenburg, den 29.06.2023