Evaluation of a New Speech Perception Test Based on Ling Sounds in German Language for Cochlear Implant Patients

FAKULTÄT FÜR ELEKTROTECHNIK UND INFORMATIONSTECHNIK
Faculty of Electrical Engineering and Information Technology

DIPLOMA THESIS

submitted in the partial fulfillment of the requirements for the degree of

Diplom-Ingenieur

in

Biomedical Engineering

by

Michael Kalcher, BSc.
Registration Number 01231475

to the Faculty of Electrical Engineering and Information Technology
at the Technical University Vienna

in cooperation with MED-EL G.m.b.H.

Advisors:
Dr. Martin Billinger & Dr. Mathias Kals (MED-EL G.m.b.H.)
Ao.Univ.Prof. Dipl.-Ing. Dr.techn. Eugenijus Kaniusas (Technical University Vienna)

Vienna, 1st June, 2020

Michael Kalcher Eugenijus Kaniusas

Conflict of Interest

The author of this work is employed at MED-EL G.m.b.H. and worked on this project together with the research and development department in Innsbruck as well as the ENT department of the Medical University of Vienna.

It shall also be noted that the author's employment at MED-EL does not cause a conflict of interest.

Acknowledgements

At first, I want to thank Prof. Eugenijus Kaniusas for the possibility of writing this thesis in cooperation with my employer MED-EL. Working a full-time job while trying to finish such a work can be extremely exhausting, which is why I really appreciate the uncomplicated communication and that there was always a quick response if questions occurred throughout the organization or writing process.
I also would like to thank my colleagues Martin Billinger and Mathias Kals, for the constant support when technical questions occurred and the interesting discussions.
Furthermore, I would like to thank Rudolfs Liepins from the Medical University of Vienna for his constant support in the study design and for the countless fruitful discussions in his office or over the phone.
Big thanks also to my boss Ewald Thurner and MED-EL's Director of Research - Signal Processing Peter Nopp for giving me the opportunity to write this thesis for MED-EL.
Special thanks to my friends Michi and Mo, for the constant encouragement and for motivating me when I really needed it.
Finally, I am grateful for my family. For the support and encouragement, they always gave me and for keeping me motivated to finish this work even in difficult and stressful times.
Many thanks also to my significant other Nina, for everything she has done to encourage and motivate me during writing this work and for always being there for me.

Kurzfassung

Die Einstellung der Audioprozessoren von Cochleaimplantaten spielt eine wesentliche Rolle in der Maximierung des Hörerfolgs des Patienten. Die Person, die diese Einstellungen vornimmt, ist sowohl auf das subjektive Feedback des Patienten als auch auf objektive Messverfahren zur Ermittlung des Hörerfolgs angewiesen. Es gibt Sprachtests, um das Verstehen von Wörtern oder Sätzen zu evaluieren und frequenzspezifische Tests, welche die Hörschwellen für unterschiedliche Sinustöne bestimmen, jedoch gibt es keinen standardisierten Test, der frequenzspezifische Sprachsignale als Teststimuli verwendet. Der Ling-6 Sound Test bietet einen Test, der diese beiden Welten vereint indem er als Teststimuli sechs Phoneme verwendet, die den Großteil ihrer Energie in bestimmten charakteristischen Frequenzbereichen haben. Zurzeit gibt es keine standardisierte Version des Tests in deutscher Sprache, bei dem diese Phoneme explizit identifiziert werden müssen.

Es wurde eine Studie an der Medizinischen Universität Wien gestartet, bei der untersucht wurde, ob ein neu erstellter Ling-6 Sound Test sensitiv auf Änderungen in den Map-Einstellungen von CI-Patienten in hohen, mittleren und tiefen Frequenzen reagiert. Weiters wurden drei, in der klinischen Routine übliche audiometrische Tests auf eine solche Sensitivität untersucht. Hierbei handelt es sich um die Freifeldaudiometire, den Freiburger Einsilbertest und den Oldenburger Satztest. Diese Arbeit präsentiert die vorläufigen Ergebnisse der gestarteten Studie. Bei den verwendeten Teststimuli handelt es sich um synthetisch generierte Ling-6 Sounds, welche mittels Klatt-Synthese erzeugt wurden. Die Synthese erfolgte basierend auf Sprachaufnahmen der Ling-6 Sounds in deutscher Sprache, gesprochen von einem männlichen Sprecher. Die synthetisch erzeugten Stimuli wurden spektral analysiert und mit Normalhörenden Probanden vorevaluiert, um ausreichende Klangqualität und Verständlichkeit zu gewährleisten.

Die ersten Ergebnisse zeigten, dass die drei in der klinischen Routine üblichen Tests kaum sensitiv auf große Änderungen in den Einstellungen der Audioprozessoren sind. Die Ergebnisse für den neu erstellten Ling-6 Sound Test allerdings zeigen, dass, wenn man die charakteristischen Vertauschungen der einzelnen Phoneme betrachtet, sich bestimmte Muster erkennen lassen. Diese Muster könnten in Zukunft zur weiteren Optimierung der Einstellung der Audioprozessoren verwendet werden.

Abstract

The fitting of cochlear implants is a crucial step in maximizing the hearing performance in cochlear implant recipients. For the person performing the fitting it is necessary to not only receive the subjective feedback by the patients but also to have objective audiometric measures to assess their hearing performance. While there are dedicated tests for speech signals and pure tones, there is no standardized test that combines the two domains by having a frequency specific speech signal as the test stimulus.

The Ling-6 sound test offers a solution to combine these two domains by using six phonemes that have most of their spectral energy in characteristic frequency regions, however, there is no standardized version of this test. A study was conducted at the Medical University of Vienna to investigate whether a new speech test based on the Ling-6 sounds is sensitive to large changes in the fitting map of cochlear implant patients. Three of the most commonly used audiometric tests were additionally tested on their sensitivity to the mentioned map manipulations. The three audiometric tests included the Freiburg monosyllable test, the Oldenburg sentence test, and the assessment of aided sound field thresholds. This thesis presents the preliminary results of the described study.

The Ling-6 sounds used in this study were synthesized using Klatt synthesis and are based on recordings of each sound spoken by a male German speaker. The synthesized sounds were spectrally analyzed and underwent a pre-evaluation with normal hearing subjects to ensure sufficient sound quality and understandability.

The results of the study suggest that the three most commonly used audiometric tests are not very sensitive to large changes in fitting maps. Results for the new Ling-6 sound test showed, however, that manipulations in certain frequency ranges cause characteristic confusions of the Ling-6 sounds. This information might be used in the future to optimize the fitting maps of patients.

Table of Contents

Abbreviations and Variables

CI	Cochlear Implant
dB HL	Hearing Level
EAS	Electric Acoustic Stimulation
GPI	Ground Path Impedance
IFT	Impedance Field Telemetry
MCL	Maximum Comfortable Level
OLSA	Oldenburger sentence test
PTA	Pure Tone Average
RS	Recognition Score
SFT	Sound Field Thresholds
SRT	Speech Reception Threshold
THR	Threshold
WRS	Word Recognition Score
cu	current units
dur	phase duration of one stimulation pulse
qu	charge units

1 Introduction

1.1 Motivation

Cochlear implantation is a widely spread and well accepted treatment for patients suffering from profound hearing loss or even complete deafness. While a successfully performed surgery is important for a good long-term hearing outcome, other factors such as patient history, rehabilitation as well as the patients' motivation to train their hearing with the cochlear implant and especially the fitting are just as important. Patient history mainly includes type and duration of hearing loss, whereas rehabilitation covers not only the time of recovery after surgery but also the time of learning and understanding the new hearing with the cochlear implant (CI).

Fitting is the process of adjusting the externally worn audio processor, and is usually performed by an audiologist, doctor or medical engineer. The corresponding adjustment which will be stored in the audio processor is called a fitting map. In the fitting process the minimum and maximum stimulation charges are set for each channel along with other stimulation parameters based on subjective feedback by the patient. Since loudness is perceived differently by each patient, additional objective audiometric measurements are performed to verify if proper hearing can be assumed. The most commonly used audiometric tests are aided pure tone audiometry, monosyllable tests and sentence tests. These tests are either based on frequency specific sounds like sine or warble tones without any correlation to natural speech signals, or on broadband speech signals like words or sentences lacking information about the exact frequency range that is heard. So far, there is no standardized audiometric test that assesses speech understanding with frequency specific information.

There is one informal test, however, targeting both speech and frequency information, which is the Ling-6 sound test. This test uses a set of six vowels and consonants that are spoken by an audiologist and the patient must repeat the heard sound. Since the sounds have their main information in certain frequency ranges, and especially the vowels in their so-called formants, the test assesses not only if speech signals (i.e. phonemes in this case) are correctly identified but also gives information on the frequency range that might be lacking amplification to make these sounds audible.

In a clinical study conducted at the Vienna General Hospital, the sensitivity of different speech tests to changes in fitting maps for cochlear implants was investigated as well as the influence of changes in the pure tone audiogram. This thesis describes the process of creating and evaluating a new speech test based on the Ling-6 sounds in German language and comparing its performance to other standardized speech tests. It includes the generation and evaluation of sound samples, the finalizing of the test procedure, the test setup as well as showing and discussing the preliminary results.

1.2 Cochlear implant fitting

In this chapter the author will provide basic information about the CI fitting procedure and the most important parameters. This is necessary to fully understand the relationship between fitting parameters, how they affect electrical hearing and, in further consequence, the performance in audiometric tests. The information provided is based on the fitting procedure and parameters implemented in CI systems by MED-EL and might differ from other manufacturers.

1.2.1 Working principle of cochlear implants

A CI system is an implantable hearing system for patients who suffer from a severe to profound hearing loss or even complete deafness. The system can be separated in two parts, an external and an internal part. The external part is the audio processor and consists of the microphone, the power supply (either by battery or rechargeable batteries), the audio processor’s CPU and the transmission coil. The internal part is the implant consisting of the receiver coil, stimulator/demodulator and the electrode array. Depending on the manufacturer, the reference electrode can either be built on the housing or as a separate electrode to be positioned outside the cochlea.

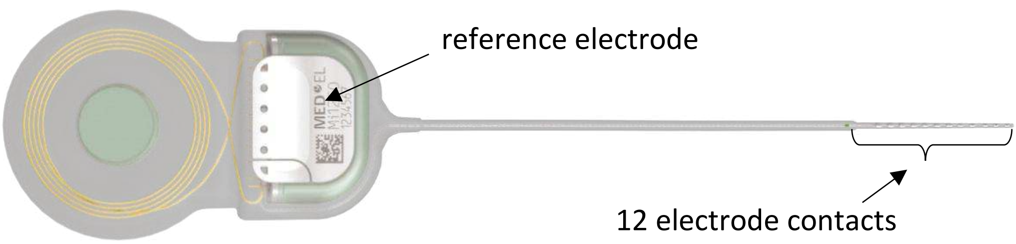

Figure 1.1: Synchrony 2 cochlear implant by MED-EL. Taken from: MED-EL

Figure 1.1 shows the latest CI by MED-EL with the 12 electrode contacts and the reference electrode. When sound reaches the microphone of the audio processor, it is spectrally analyzed and processed to create a stimulation pattern of short bipolar current pulses. This coding information is transmitted through the skin, where the stimulator creates the according pulses and sends them to the individual contacts on the electrode lead. The inner ear has a tonotopic frequency arrangement, where high frequencies are located in the basal regions of the basilar membrane closer to the oval window and low frequencies in the apical regions. This arrangement also applies when stimulating with a CI. The deeper one electrode reaches towards the apical regions of the basilar membrane in the inner ear, the lower are the frequencies that will be heard in the end.

1.2.2 Adjustment of basic fitting parameters in cochlear implants

The necessary fitting equipment consists of the patient's audio processor, a programming cable, the MAX programming interface with USB cable and a computer with the fitting software MAESTRO installed. Figure 1.2 shows a picture of this setup.

The audio processor is connected to the interface via the programming cable. The first task during a fitting session is always to perform an Impedance Field Telemetry (IFT) measurement. In this measurement the impedance of each of the 12 electrodes is determined as well as the ground path impedance (GPI) of the reference electrode located on the implant body.

Figure 1.2: MED-EL cochlear implant fitting setup

The individual IFT is necessary to both functionally check the individual contacts on the electrode array and to ensure accurate stimulation charges.

The acoustic dynamic range of a normal hearing ear ranges from 0 up to about 120 dB SPL, whereas 120 dB SPL corresponds to the acoustic threshold of pain. The main idea behind CI fitting is to map a large acoustic dynamic range from about 30 dB SPL up to 100 dB SPL to the electrical dynamic range of the CI patient. The difference in the dynamic ranges is caused by technical limitations as for example the microphone noise and internal noise floor as well as by the fact that electrical stimulation of the cochlea has a much lower dynamic range than acoustic stimulation.

Figure 1.3 depicts a simplified signal path of an acoustical sound being mapped to electrical stimulation. When sound reaches the microphone of the audio processor, it first undergoes a series of front-end proessing steps including the Automatic Gain Control (AGC). The function of the AGC is to compress the input dynamic range so that it can be easier processed by the CI patient. The AGC is followed by a filter bank and envelope detection (not in the Figure), then by a logarithmic maplaw function to model the electric loudness growth and finally the psychophysic parameters Maximum Comfortable Level (MCL) and Threshold (THR).

While AGC and maplaw are parameters that affect all 12 electrodes globally, MCL and THR can be set individually for each electrode. The two most important of all these parameters which are defining the electrical dynamic range, are MCL and THR.

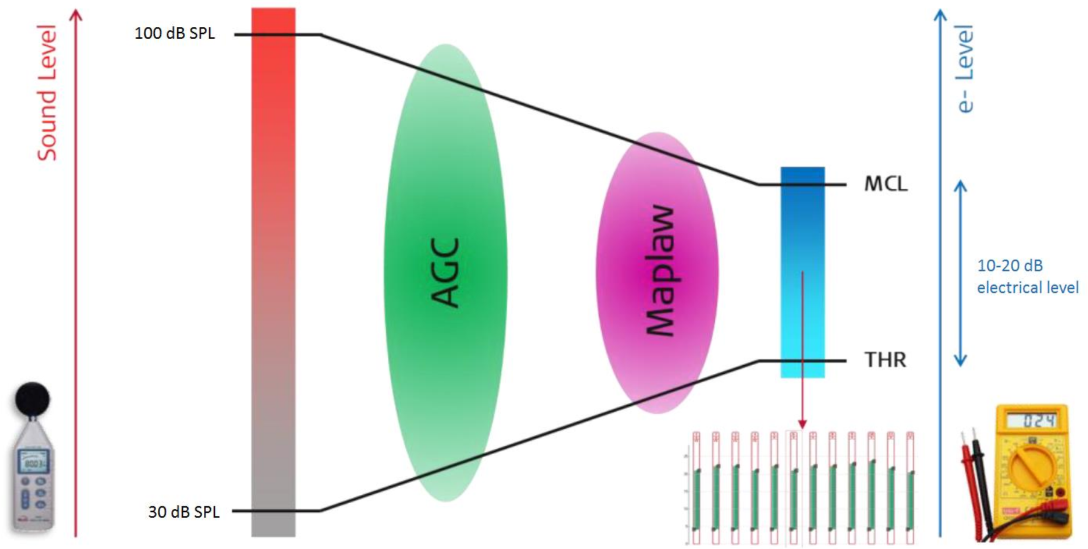

Figure 1.3: Simplified signal path of sound being mapped to electrical stimulation. Taken from MED-EL.

The MCL parameter corresponds to the maximum stimulation charge for one specific electrode contact that will only be presented at an input of about 105 dB SPL and louder. It defines the loudest sound perception on one specific stimulation channel that is tolerated by the CI patient.

In patients with reliable subjective feedback, the MCLs are usually assessed by presenting a stimulation burst and increasing the stimulation charge until the patient reports a loudness level of “loud but comfortable”. It is important to note the difference between the MCL used here and the “Most Comfortable Level”, which is mostly used in hearing aid fitting and represents the sound level that is overall most comfortable.

The THR parameter on the other hand represents the lowest stimulation charge that is applied for inputs below about 25-30 dB SPL. It is assessed by first presenting stimulation bursts above the true THR value and by instructing the patient to give feedback on if the sound was heard. If it was heard, the stimulation charge is decreased until the patient is not able to hear the sound anymore. The THR is the stimulation current that is just inaudible to the patient.

Figure 1.4 on the left shows the relationship between loudness and the stimulation parameters MCL and THR. Figure 1.4 on the right shows one channel strip out of the MAESTRO fitting software, the arrows pointing to MCL and THR respectively.

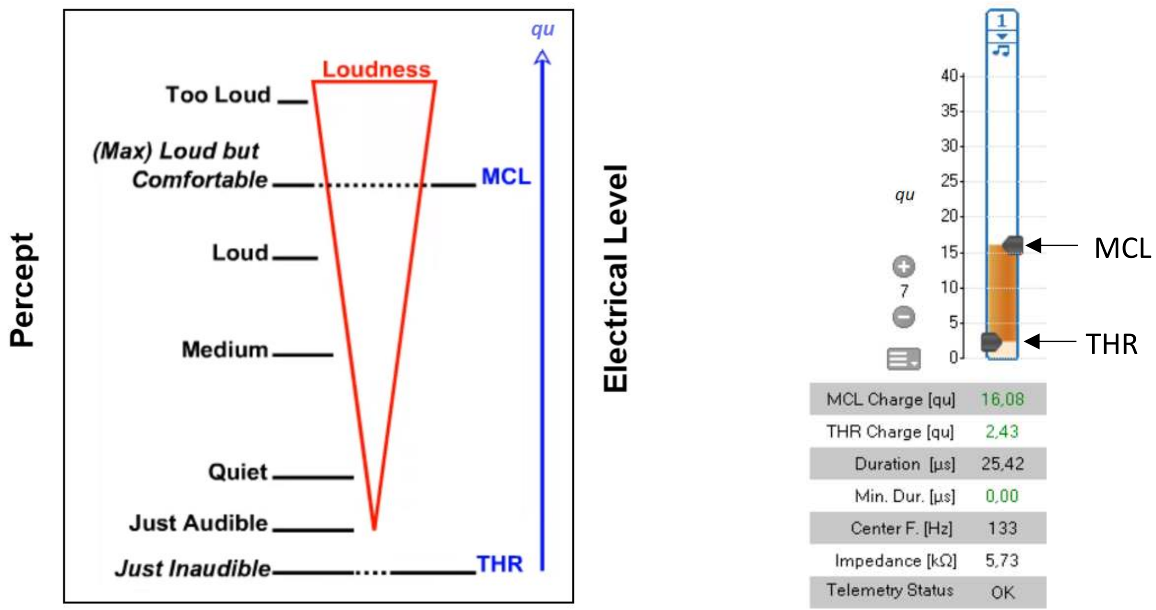

Figure 1.4: Mapping of loudness percept to electrical level. The left side shows the relationship between perceived loudness, the right side shows a single channel strip of the most apical channel in MAESTRO. Taken from MED-EL.

For simplification in clinical practice, MAESTRO uses the unit abbreviations cu (current unit) and qu (charge unit). Current units represent the stimulation current applied whereas 1 cu equals 1 µA of nominal current. Charge units represent the charge that is applied to one electrode in the stimulation process whereas 1 qu equals a nominal charge of 1 nC.

Other fitting parameters that influence the hearing of a CI patient are the steepness of the maplaw function, the used coding strategy, the overall frequency range and the center frequency for each individual electrode channel. These will not be further discussed because they are not relevant for the scope of this thesis.

In the beginning phase after the initial activation of the audio processor, MCL levels are increased gradually in subsequent fitting sessions. This time period of roughly one or two months is dedicated to accustoming the CI patient to the new form of stimulation and hearing that the cochlear implant provides. At the beginning, even low stimulation currents are already perceived as very loud and unpleasant, but the patients tend to adapt quickly so that the very first fitting map seems significantly quieter after just a few hours. This process of habituation is usually performed using a fitting concept that is called progressive maps.

A fitting map is a set of adjustments to the fitting parameters that will be stored in the audio processor. An audio processor can store several fitting maps. Therefore, the patient receives four maps with each having a significant increase in stimulation charge compared to the previous one, the first map being the quietest and the fourth map being the loudest. Then, the patient is instructed to use the first map for about two or three days. After that, he is advised to switch to the next map which will be louder. Once he reaches the loudest map, a follow up fitting with the audiologist or engineer will be performed, which is usually within two weeks. After this initial phase, MCL levels are stabilizing.

1.3 Audiometric tests

Audiometric testing plays a vital role in CI fitting. Sufficient understanding of the test language is a crucial factor when it comes to creating reliable results in audiometric tests. Therefore, every country uses their language specific tests. In this chapter the author will focus on the German speech tests primarily used in clinical routine as they are the ones that were used in the clinical study and will address critical considerations about the individual tests found in literature. In addition, the testing of one's hearing thresholds is briefly described.

1.3.1 Pure Tone Audiogram

Pure tone audiometry describes the procedure of finding the hearing threshold, i.e. the just audible hearing level of the subject. In normal hearing subjects, this test is usually performed with calibrated head phones or in-ear phones in a quiet environment with an ambient noise of maximum 35 dB (A) (*Recommended procedure: Pure-tone air-conduction and bone-conduction threshold audiometry with and without masking*, 2011, p. 7). To test the aided condition of subjects with hearing loss, be it a conventional hearing aid or a CI system, the use of headphones or in-ear phones is not always applicable. Consequently, audiologists usually perform the test with calibrated loudspeakers when assessing the hearing thresholds for patients with hearing aids or hearing implants. When the stimuli are presented with loudspeakers, the test in usually called sound field audiometry, where sound field thresholds (SFTs) for different frequencies are measured.

Presenting pure tones in a test room with parallel walls without sufficient acoustic treatment can lead to standing waves distorting the sound level in the listening position due to interferences. To minimize these standing waves in the test room, warble tones instead of sine tones can be used. Warble tones are pure tones whose frequencies vary periodically several times per second over a small range. This variation in frequency ensures that no standing waves are generated or that they are at least not pronounced. These tones have been proven to be a suitable alternative to pure tones in sound field audiometry (Dockum & Robinson, 1975).

The results for SFTs are given in decibels in hearing level (dB HL), whereas 0 dB HL corresponds to the hearing threshold in dB SPL for the average normal hearing listener. Therefore, if a patient can detect a frequency (e.g. 1 kHz) at 40 dB HL, it was 40 dB SPL louder than the hearing threshold for an average normal hearing person. For example, at 1 kHz the hearing threshold for an average normal hearing person is 0 dB SPL or 20 µPa. The test procedure for an aided sound field audiogram typically includes frequencies between 125 Hz and 8 kHz. The individual frequencies are presented at different levels and the subject indicates if the sound was heard by e.g. pressing a button. During the assessment the tester first ensures that the subject is familiar with the task by presenting a stimulus, usually 1 kHz, at a level that is clearly audible for the subject. If the familiarization leads to a satisfactory positive response by the subject, the actual testing begins by decreasing the level of the stimulus until no further response occurs, i.e. until the stimulus is not heard by the subject anymore. Then the stimulus level is increased in a stepwise manner until a response re-occurs. To verify the measured thresholds the steps of increasing and decreasing the stimulus level will be repeated until the threshold level has been clearly verified.

This procedure is performed for all audiologic test frequencies. The number of test frequencies depends on the time available and on the necessity for detailed information, however, at least the frequencies at 0.5, 1, 2 and 4 kHz should be tested. Typically, 0.25 kHz and 8 kHz are also tested. Averaging the hearing thresholds at 0.5, 1, 2 and 4 kHz results in the so called PTA4 (Pure Tone Average 4) and including 0.25 kHz and 8 kHz yields the PTA6. These values give a rough estimate of a subject's hearing level in each ear. In the study, eleven test frequencies from 0.125 kHz up to 8 kHz will be used.

1.3.2 Sound field audiogram in cochlear implant fitting

Sound field audiometry in cochlear implant fitting helps to assess if the patient can hear soft sounds with the CI at tested frequencies. The goal is hearing thresholds in the range of 20-30 dB HL. SFT can be manipulated by adjusting THR and/or MCL charge levels. Since THR defines the lower end of the electric dynamic range their influence is high, whereas MCL defines the upper end of the electric dynamic range and influences SFT only minorly. This means, elevated SFTs can indicate too low THR levels or to a certain extent too low MCL levels.

1.3.3 Freiburg speech test

The Freiburg speech test is predominantly used in German speaking countries for diagnostics and hearing aid fitting as well as CI fitting. It assesses speech perception with multisyllabic numbers (Freiburg numbers test) and with monosyllabic nouns (Freiburg monosyllable test). Conductive hearing loss correlates with a hearing loss in speech but not necessarily in a discrimination loss. Combined or sensorineural hearing loss correlates with a loss of discrimination in speech, meaning that louder sound levels do not result in better understanding (Hahlbrock, 1970). The Freiburg speech test is standardized in the German norms DIN 45621-1 and DIN 45626-1 (Winkler & Holube, 2014).

The test consists of two parts. In the Freiburg numbers test, multisyllabic numbers are presented. Ten lists of ten numbers larger than 13 and with either two or four syllables are available. The hearing loss in speech is then defined by the 50 % detection threshold for numbers, i.e. by the sound pressure level at which 50 % of the numbers were correctly understood. In the Freiburg monosyllable test, the monosyllabic nouns are tested, where 20 test lists with 20 nouns each are available. Discrimination loss is the difference to 100% in word recognition at the tested presentation level.

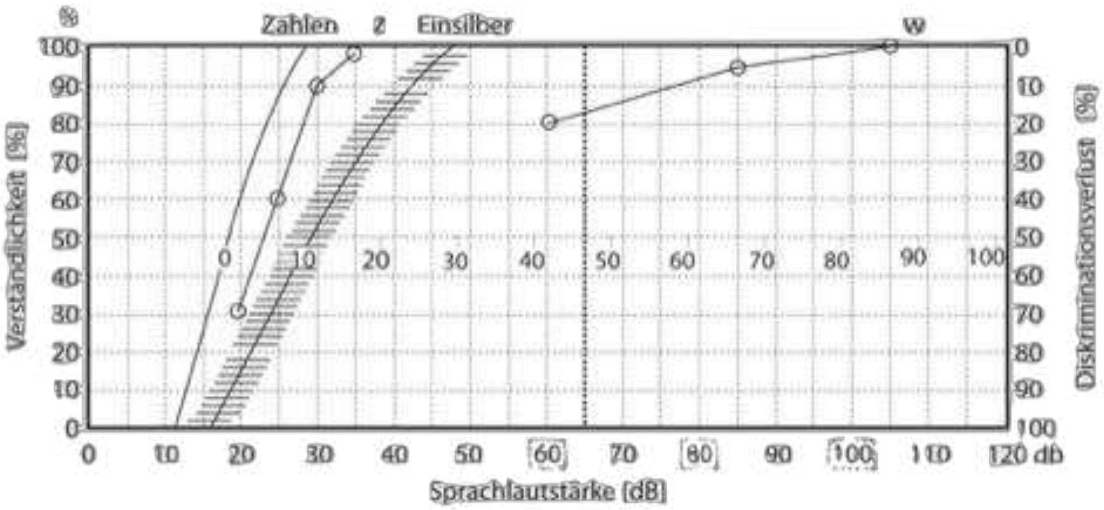

Figure 1.5: Example of a Freiburger speech test evaluation diagram. The word recognition score is printed on the left and the discrimination loss on the right. The presentation level of the speech signals is plotted on the bottom and the scale in the middle represents the hearing loss for monosyllabic numbers in dB. The plotted curves represent the norm curve for multisyllabic numbers, a measured curve for multisyllabic numbers, the norm curve for monosyllabic words and a measured curve for monosyllabic words (from left to right). Taken from (Wilhelm, Lachenmayr, & Hamann, 2015).

The Freiburg speech test with its two parts can assess (together with other clinical assessments) if a conductive hearing loss or a sensorineural hearing loss is present in the patient. The Freiburg monosyllable test is one of the most commonly used speech tests in German speaking countries. It is well-known, widely available and used in everyday clinical practice. The frequent usage of the test and the large number of publications where the test is included leads to a large confidence when interpreting the results. Additionally, for many years, it has played a central role in the indication criteria for hearing aids as well as the evaluation of the treatment outcome. In Germany, the indication criteria for hearing aids states that when the patient only understands a maximum of 80% of monosyllables at 65 dB SPL hearing aids are indicated. The improvement after treatment with hearing aids should be at least 20 % of gain in speech understanding (Steffens, 2015). Another reason is that the Freiburg monosyllable test is standardized in the DIN 45621-1 (Sukowski, Brand, Wagener, & Kollmeier, 2008).

Despite the frequent use and its standardization, there are critical aspects with respect to the reliability of the Freiburg monosyllable test. Firstly, it shows considerably high test-retest differences due to the small number of only 20 test words (items) per test list. Therefore, only a fairly large difference of at least 15-30% can be stated as significantly relevant (Steffens, 2006). Another issue is its large variation in word understandability, which results in different difficulties depending on the test list. For example, Mallinger (2011) showed that list 15 is significantly easier and list 1 significantly more difficult to understand than the other test lists. This difference is caused by fluctuation in sound level of the test words, unnatural over-articulation and the use of outdated vocabulary which is very uncommon in modern everyday German language. Almost 50% of monosyllables used in the Freiburg test were shown to be out of use in the daily life of the average German speaker (Steffens, 2016).

1.3.4 Oldenburg sentence test

Hearing loss is usually most noticeable when trying to communicate in noisy environments. To obtain a realistic measure of the specific hearing loss, hearing diagnostics and rehabilitation use speech tests in noise to simulate these noisy environments. The Oldenburg sentence test (OLSA) is an audiometric test used to assess the speech reception threshold (SRT) in quiet and in noise. The SRT in quiet is the acoustic level at which the patient correctly understands 50 % of the presented speech signals, while the SRT in noise is the signal-to-noise ratio that leads to 50 % speech understanding (*Oldenburger Satztest: Handbuch und Hintergrundwissen*, 2000).

The speech material of the OLSA consists of 40 test lists with each 30 sentences. The sentences all have the form: *name verb number adjective object* with random combinations out of a collection of 10 words per category in total. Therefore, the sentences do not necessarily make sense, making it impossible to memorize the phrases, thus, it is possible to repeat the measurements without interfering learning effects. Another benefit of randomizing the words is that patients cannot make logical assumptions or guesses when performing the test. The interfering noise signal is created by averaging numerous time-shifted target speech sentences and therefore its long term spectrum matches the sentence material (Wagener, Kuehnel, & Kollmeier, 1999).

The SRT in noise is assessed using an adaptive approach. Usually, the noise signal is kept at a fixed level while the speech signal changes adaptively after each presented sentence according to the response of the subject. This way the SRT can be assessed in a more efficient way than by adjusting the levels manually.

The OLSA enjoys a lot of benefits concerning its reliability. Firstly, the possibility to perform the OLSA with noise gives a more natural representation of the hearing loss in everyday life. Secondly, the steep discrimination function provides high accuracy in the test results. Also, it has a large number of repeatable test lists which highly benefits its use in clinical studies where a larger number of speech comprehension test have to performed in different listening conditions. The test lists are phonetically balanced making it a realistic representation of everyday speech and increasing the test-retest reliability, and lastly it was recorded in average talking speed which is a clear advantage over other sentence tests like the Göttingen sentence test and therefore making it more suitable for hearing aid and CI patients.

The main disadvantage of the OLSA is that its procedure can be too difficult and exhausting for elderly patients or patients who generally have low performance with their CI even after a training session. Also, just like the Freiburg monosyllable test, the OLSA gives no insight about the frequency characteristics of the misunderstood words or phonemes. Therefore, the audiologist still does not gain any frequency specific information for the fitting process (Wagener et al., 1999). The pronunciation of the test material has strong characteristics of a dialect that is mainly spoken in northern Germany, which might make it more difficult for patients coming from areas with a completely different dialect. One last disadvantage might be that the names used in the test are not as common in Austria and might already be slightly outdated even in Germany.

1.3.5 Ling-6 sound test

The speech tests described above as well as other available speech tests do not give any insight into a patient's frequency-specific hearing characteristics behind the misunderstood words or phonemes. The Ling-6 sound test is an audiological test method developed by Daniel Ling (Ling, 2002) to determine a patient’s ability to detect or identify speech phonemes in the frequency range of speech. It is generally used for informally verifying the effectiveness of hearing aid and CI fitting. In this test, six sounds are orally presented in a random order at ‘conversational level’ from 1 meter and 3 meters, and the patient either must verify that he or she heard a sound (detection form) or should repeat it back to the audiologist (identification form). A third form would be the discrimination form where patients must discriminate two different sounds without identifying them individually.

Fundamental Frequencies (Hz)		/a/	/i/	/u/
F0	M	124	136	141
	W	212	235	231
Formant Frequencies (Hz)				
F1	M	730	270	300
	W	850	310	370
F2	M	1090	2290	870
	W	1220	2790	950
F3	M	2440	3010	2240
	W	2810	3310	2670

Table 1.1: Averages of fundamental and formant frequencies of vowels by 76 speakers. The speakers were divided into male (M) and female (W). Adapted from (Peterson & Barney, 1952).

The included sounds are /**ʃ**/ as in *fish,* /**s**/ as in *us,* /**i**/ as in *she*, /**u**/ as in *two*, /**a**/ as in *car* and /**m**/ as in *me.* Identifying these sounds correctly is considered to be important because these sounds have concentration of energy (in vowels called formants) across the speech frequency range from 250 to 8500 Hz (*Australian Hearing Manual of Speech Perception*, 2001). The exact formant frequency varies for each speaker so Table 1.1 shows the averages of fundamental and formant frequencies of vowels by 76 speakers as investigated by Peterson and Barney (1952).

Even though the Ling-6 sound test is quick and easily administered, its current oral form of presentation makes standardization difficult. Therefore, it seems logical to seek for a version with sound recordings at a certain sound pressure level. Scollie et al. (2012) presented and evaluated a calibrated version of the Ling-6 sound test for assessment of aided detection thresholds. The current thesis on the other hand aims to create another version that detects weaknesses in specific frequency ranges and eventually gives more detailed information on the quality of a CI patient's fitting map.

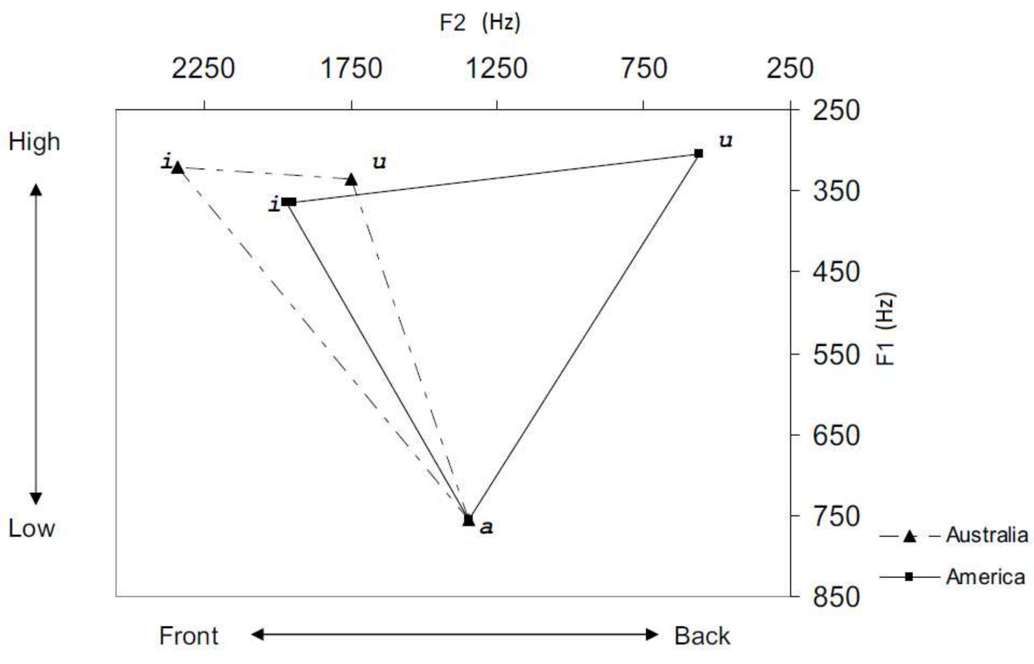

Figure 1.6: Australian versus American English vowel formants. F1 represents the first formant of the respective vowels and F2 the second formant. The higher a vowel is according to this graph (left scale), the closer to the palate is the body of the tongue during its articulation. The lower the vowel, the lower is the body of the tongue. Front and back again refer to areas in the mouth where the respective vowel is articulated. Taken from (Agung, Purdy, & Kitamura, 2005, p. 5)

1.3.6 Language specific differences in the Ling-6 sound test

Speech tests are always language specific and this also applies to the Ling-6 sound test. Not only the spoken language itself affects the test but also a possible dialect of this language. For example, Agung et al. (2005) showed that there are clear differences in the sound /**u**/ when comparing American English to Australian English. In American English the formants F1 and F2 lie at about 300 Hz and 500 Hz, respectively, whereas in Australian English the F2 is at about 1750 Hz (see Figure 1.6).

In the orally performed detection task, the influence of these differences will not have a large impact especially since the test should be assessed by a native speaking audiologist or by a parent. These differences become more critical when creating a speech test based on the Ling-6 sounds in the form of an identification task. The main goal of this test is to receive frequency specific information about the fitting map of patients and therefore the exact analysis of spectra and formant information is of high importance. Therefore, the current thesis will conduct said analysis for synthesized Ling sounds in German language to give more information on the differences to the ones in English and also to have exact frequency information for further discussion of the study results.

2 Methods

2.1 Synthetic Ling-6 sounds

2.1.1 Pilot study with recorded Ling-6 sounds

In a previous project, the author investigated the usability of recorded Ling-6 sound samples in German language (Kalcher, 2019). The test items were recorded by non-professional speakers and underwent post-processing so that the final sound recording for each phoneme was one second long, as stationary as possible, and showed no traces of natural articulation. The reason for this form of post-processing was that the sound sample should be isolated to only contain the frequency information that defines the formants of the sound without its articulation characteristics.

The evaluation of the created sound samples with normal hearing subjects showed that the quality of the samples was not sufficient to be used in the final test. The criterion for usability of the sound files was that they had to be identified correctly by all normal hearing subjects. The criterion was not fulfilled for most of the recorded samples, therefore either the sound samples had to be recorded again with professional speakers or an alternative for recorded samples had to be found. Synthetic Ling-6 sounds created with a formant synthesizer were proposed as an alternative because of the extended possibilities to control all frequency parameters and therefore the resulting versatility when including the samples in the final test.

2.1.2 Creation of synthetic Ling-6 sounds

Before starting the synthesis of new Ling-6 sound samples, the existing recorded samples were analyzed with regard to their formant structure using Praat phonetics software (Praat 6.1.09, January 2020). The voiced Ling-6 sounds **/a/**, **/u/**, **/i/** and **/m/** were then created using a Klatt synthesizer implementation in Python. Here, the previously performed formant analysis was used to determine the center frequency and bandwidths of the formants, which were then extracted into the Klatt synthesizer and empirically modified to achieve a sound sample that was subjectively considered most natural. The fricatives **/s/** and **/ ʃ /** were created by randomizing the phase of their respective audio files in the frequency domain to balance temporal variations and remove any fluctuations caused by the natural articulation.

After synthesis, the sound samples were normalized according to the "Recommendation TU-R BS.1770-4" by the International Telecommunication Union from 2015. During the first three test sessions within the clinical study, it was noticed that the sound /**a**/ was always correctly identified by every subject and they reported that it was the easiest to identify because it had a larger volume compared to the other sounds. When one sound is identified correctly 100% of the time there is no information gain when using this phoneme and the use of this particular sound might be obsolete. To address this issue, a normalization according to subjective loudness percept was performed. The six sounds were presented to ten subjects who were instructed to adjust all sounds to the same subjective loudness. The averaged amplification and attenuation values were then applied to the individual sound and another evaluation with five subjects was performed to verify the resulting corrections. If, after this subjective normalization, the sound /**a**/ is still identified easily, then it will be due to the sound's characteristics in frequency and not due to a difference in volume.

Finally, the synthetic Ling-6 sound samples were analyzed by plotting the power spectrum in 1/3 octave for comparison with the English Ling sounds recorded from a female speaker used in the study by Scollie et al. (2012). Due to missing detailed explanation of the normalizing process in the paper by Scollie et al. (2012) a power spectrum and a different scale were used but since the desired information lies within the spectral peaks this analysis should be sufficient.

2.1.3 Evaluation of synthetic Ling-6 sounds

Before starting the main study in the clinical setting, an evaluation with normal hearing subjects was performed to ensure that the newly created synthesized sound samples are clearly distinguishable. This was a necessary step before starting the clinical study to prevent that the hearing-impaired subjects misunderstood the sound solely because of poor quality which would falsify the results.

A laptop computer was used for the test procedure and the sound samples were presented via over-ear closed headphones (SENNHEISER GSP 300). Overall, the subjects were presented ten repetitions of each sound, yielding 60 stimuli in total. The volume was adjusted to a subjective "quiet conversational level". The possible answers were chosen to be a closed set of the Ling-6 sounds and four additional answers. The additional answers were F, E, O and Not Applicable (N.A.), whereas F represents an alternative for the fricatives and E and O an alternative for the vowel sounds. The subjects should select N.A. if they thought none of the shown options suited the sound they heard. For all five participants in this evaluation, the overall recognition score (RS) of correct answers in percent, as well as a confusion matrix based on the six possible sounds was obtained. The results of this evaluation will be described in section 3.2.

2.2 Study design and ethics committee

In cooperation with the Medical University of Vienna, a clinical study was conducted to evaluate the new speech test for CI patients based on Ling-6 sounds in German language. All participants were informed precisely about the test procedures and signed an informed consent form. All tests and investigations are performed according to the ethical standards of the Declaration of Helsinki 1964 and were approved by the ethics committee of the Medical University of Vienna (1070/2019).

2.2.1 Objectives

The primary objective of the study was to investigate whether a speech test based on Ling-6 sounds is sensitive to changes in a fitting map in high, mid and low frequencies. Surprisingly, no literature was found investigating the sensitivity of other speech tests to said changes in fitting maps, therefore the secondary objectives were to investigate three more audiometric assessment methods which are typically used in the assessment of the hearing performance of CI users in German speaking countries. These methods included the Freiburg monosyllable test, the Oldenburg sentence test in noise, and the assessment of aided hearing thresholds. The goal was to test 20 subjects either unilaterally or bilaterally implanted and also including patients suffering from single sided deafness.

2.2.2 Inclusion criteria

The inclusion criteria for the study were defined as follows:

- Age between 18 and 90 years
- Signed and dated informed consent before start of any study specific procedure
- At least one year of experience with CI
- Fluency in German language

- MED-EL SonataTI100, ConcertoMI1000 or SynchronyMI1200 implant
- OPUS 2, RONDO, SONNET or RONDO 2 audio processor

The reason for the limitation in audio processors was that all fitting parameters had to be transferred to an audio processor solely dedicated to testing in order not to manipulate the patient's personal audio processor. Additionally, not all audio processors share the same specifications for microphone directionality and this way the factor of understanding differently due to also changing the microphone directionality by switching audio processor model could be eliminated.

2.2.3 Exclusion criteria

The exclusion criteria for the study were defined as following:

- Any physical, psychological or emotional disorder that would interfere with the ability to perform the test
- Patients using electric-acoustic stimulation (EAS)
- Pregnancy or breastfeeding

2.2.4 Equipment and listening configuration

Speech perception was assessed in an acoustically treated sound-isolated chamber. The subject was seated in front of a level-calibrated audiometric loudspeaker at a distance of 1 meter. Both the target and the noise signal were presented from the loudspeaker. If necessary, the contralateral side was covered with earplugs and sound isolating earmuffs in order to exclusively test hearing through the CI.

The used equipment consisted of:

- **MED-EL MAESTRO 7:** Fitting Software for programming the CI audio processor
- **MAX Interface Box:** Computer interface for programming the CI audio processor
- **PC and high-end audio interface (RME Fireface UC)** to control the test procedures
- **WESTRA LAB 501 audiometric loudspeaker** to present the test stimuli
- **OPUS 2, RONDO, SONNET and RONDO 2** audio processors + programming cables

2.2.5 Test conditions

To evaluate the sensitivity of the individual speech tests to changes in fitting maps, the test audio processors were programmed with four different maps. The changes in the fitting maps are shown in Table 2.1:

clinical	high	mid	low
preferred clinical map	clinical map, MCLs reduced by 30 % on channels with center frequency above 2600 Hz	clinical map, MCLs reduced by 30 % on channels with center frequency in the range of 750 - 2600 Hz	clinical map, MCLs reduced by 30 % on channels with center frequency below 750 Hz

Table 2.1: Map modifications for testing with each speech test.

2.2.6 Testing

Pure tone audiometry was performed with all four maps (including the clinical map) to evaluate how the hearing thresholds change with the maps. The results will be shown as difference in dB HL compared to the clinical map as well as differences in PTA in the frequency ranges corresponding to the test conditions.

The Freiburg monosyllable test was used to determine the word recognition score (WRS) in percent. In this study two lists were presented at 65 dB SPL for each test condition. The results will be shown as WRS in percent as well as the improvement compared to the clinical map.

The Oldenburg sentence test was used to determine the speech reception thresholds in dB. The OLSA target material was presented at a constant level of 65 dB SPL in adaptive masking noise via loudspeakers. The masker level was adjusted adaptively according to a staircase procedure to determine the SRT, which is defined as the SNR that yields 50% of word recognition on average. The procedure started at an SNR of +10 dB and was adapted depending on the subjects' performance in each trial. Before the test, a training list of 20 sentences was presented to familiarize the subject with the testing procedure. Subsequently, 2 lists of 20 sentences were presented for each test condition. The results will be shown as SRT in dB as well as the improvement compared to the clinical map.

The Ling-6 sounds were calibrated to 65 dB SPL with broadband calibration noise prior to testing. After calibration the subject underwent a training session of two presentations of the Ling-6 sounds to familiarize with the test procedure. Following the training, six sounds were each presented 10 times in a randomized order for each test condition. The subjects entered the respective answers based on their judgment. The results will be shown as recognition score in percent and the improvement compared to the clinical map, and in form of confusion matrices to evaluate if characteristic confusions occurred. The order of testing was randomized between the subjects.

2.2.7 Data analysis

All data was collected in case report forms (CRFs) for each subject and marked with the individual subject ID for anonymization. The data analysis and creation of plots was performed with Microsoft Excel and Matlab.

The improvements between the different test conditions was statistically analyzed using a Kruskal-Wallis test and consequently a one-sample Wilcoxon signed rank test. The non-parametric tests were chosen because a normal distribution could not be assumed due to the small sample size of n=7.

2.3 Test script in Python

One basic requirement for the new Ling-6 sounds test procedure was straightforward usability in order to be used by the clinical personnel at the test center. Another requirement was that the answers and results are stored automatically. Without having the subject answer orally and an examiner writing down the answers, the test can be carried out notably faster, thus helping the patient to remain focused during the whole test run. This can help produce more reliable results because the test can be assumed to be mentally less fatiguing. Therefore, a test procedure was programmed in Python with a simple graphical user interface (GUI) to guide the subject through the test procedure.

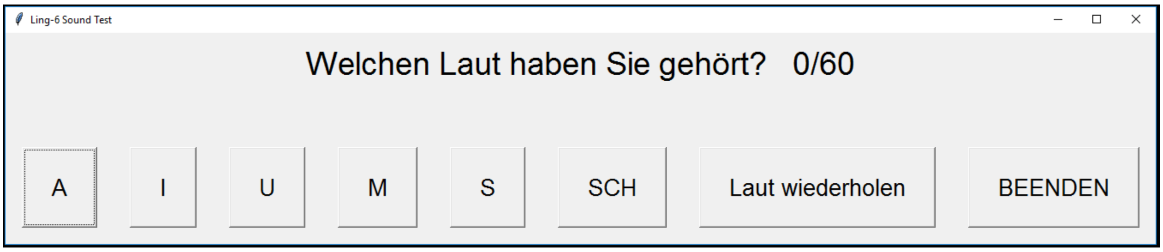

Figure 2.1: Main testing window of the Ling-6 identification test

Figure 2.1 shows the main testing window and the possible available response options. The study participants first are presented with an introduction window. After the introduction window, the supervisor has to enter the subjects' individual study ID needed for documentation. In the next window the task is explained for to the subjects and the possible answers responses presented. Once the subject clicks on START the first sound sample is presented immediately. In the main testing window (figure 2.1), Subjects are instructed to click on the sound they believe to have heard and after giving response via the corresponding GUI button, the next sample is played automatically. Subjects are also given the possibility to repeat the sound to account for temporal lack of focus, since the sample length of one second is rather short.

The number of repetitions for each sound was chosen to be ten, resulting in a sum of 60 sound sample presentations. This way the whole test procedure could be performed in approximately three to four minutes depending on the individual answering speed with enough repetitions to obtain a representative number of answers for the evaluation. The possible answers responses were reduced to a closed set of only the Ling-6 sounds instead of additional sounds that were not tested. This is in contrast to the evaluation described in section 2.1.3 and the reason for this change is, that this study shall first only focus on confusions within the Ling-6 sounds. The set of possible answers should be kept as small as possible to facilitate the analysis of the data and to make the test not too difficult for CI users.

3 Results

3.1 Spectral analysis of Ling-6 sounds

In this section a comparison shall be made between the Ling-6 sounds used by Scollie et al. (2012) and the synthetic Ling-6 sounds created for this study. Figure 3.1 shows the English Ling-6 sound recordings by Scollie et al. (2012). Figures 3.2 and 3.3 show spectral analyses for the synthetic Ling-6 sounds created in the current thesis.

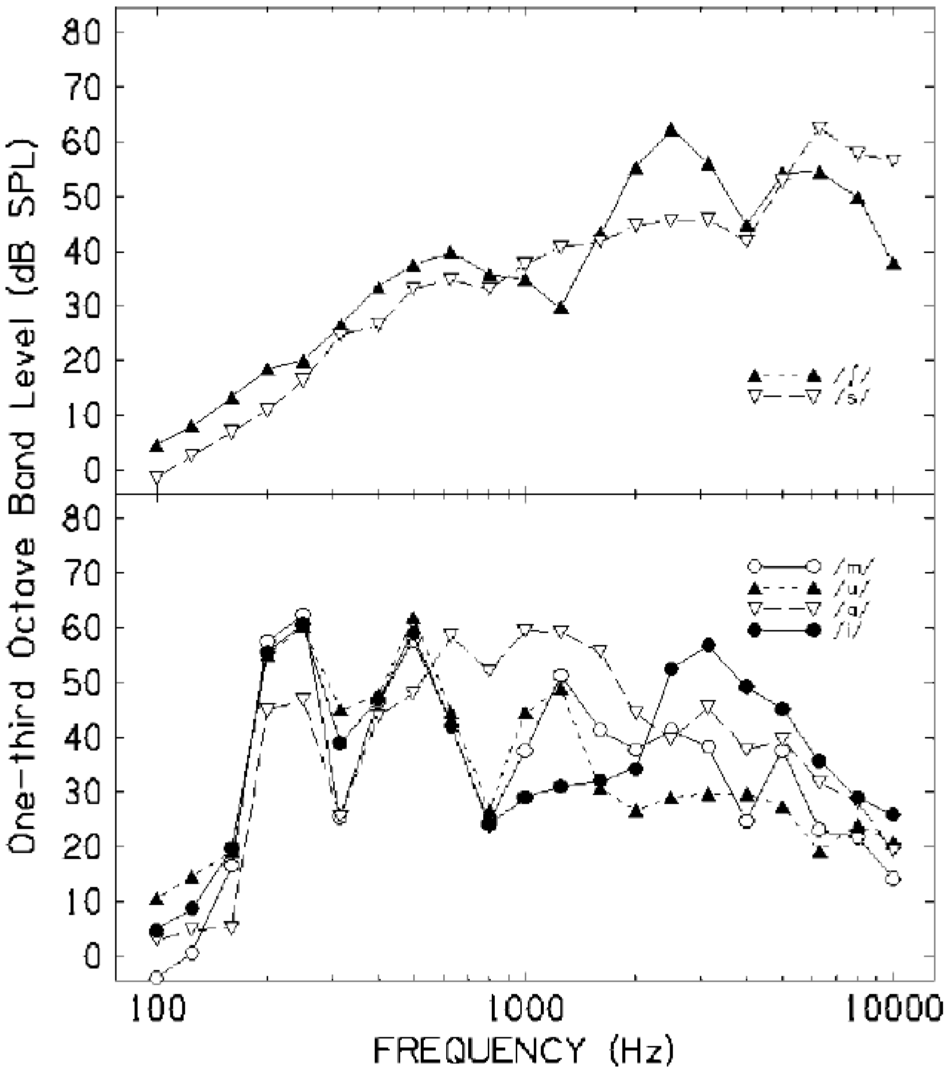

Figure 3.1: One-third octave spectra of the Ling-6 stimuli of a female speaker used by Scollie et al. in 2012. The overall level of each stimulus has been normalized to 65 dB SPL. Taken from (Scollie et al., 2012)

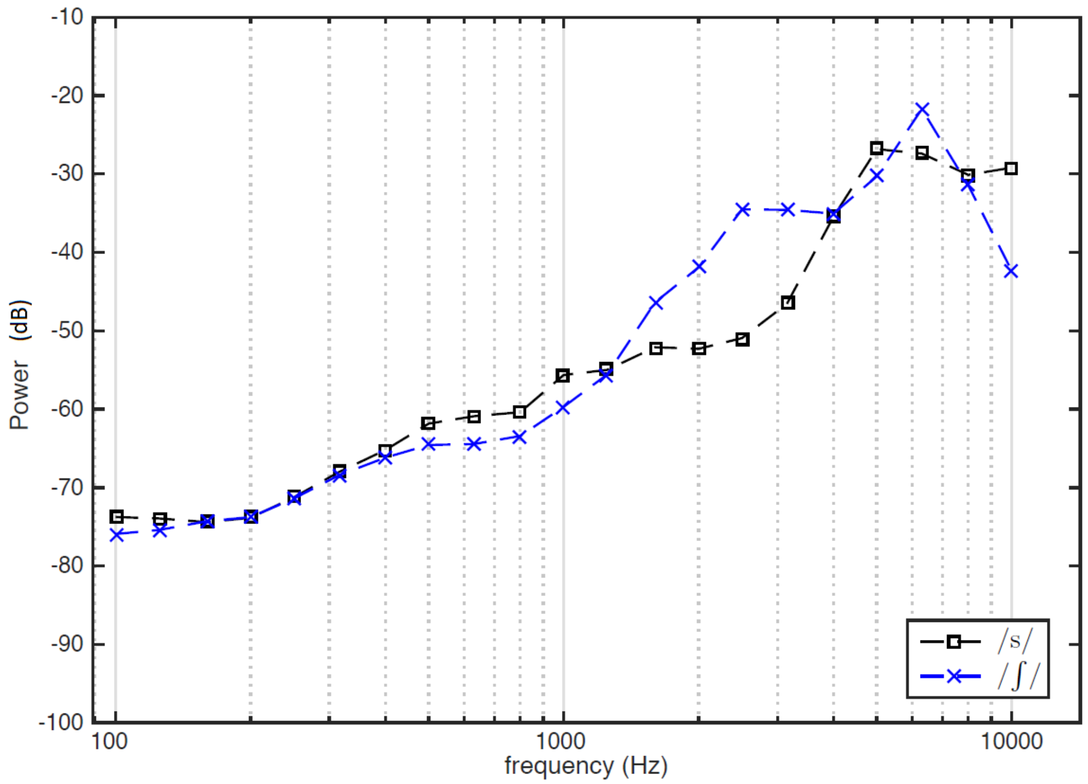

Figure 3.2: One-third octave spectra of synthesized Ling-6 sounds /**s**/ (black) and / **ʃ** / (blue) for a male speaker.

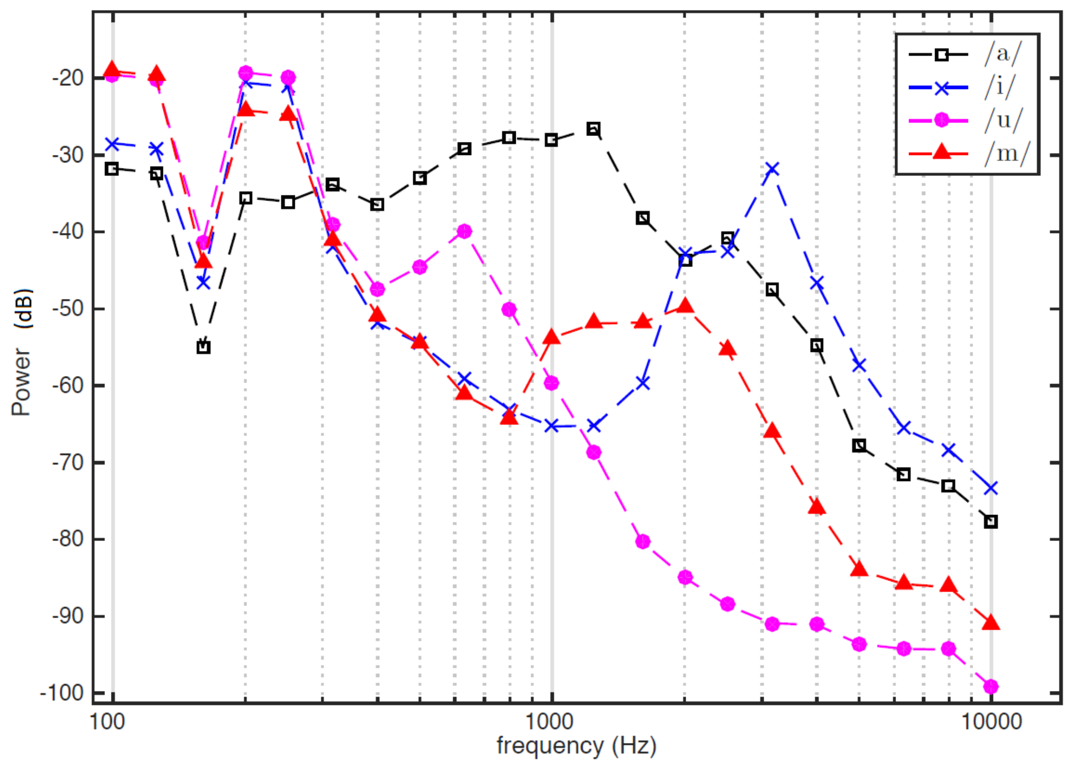

Figure 3.3: One-third octave spectra of synthesized Ling-6 sounds /**a**/ (black), / **i** / (blue), /**u**/ (pink) and /**m**/ (red) for a male speaker.

When comparing the sounds /**s**/ and / **ʃ** / in figure 3.1, the spectra show high similarity up to about 1 kHz. The sound / **ʃ** / shows an almost 20 dB SPL higher level in the frequency range between 2 kHz and 4 kHz, while the sound /**s**/ on the other hand shows a significantly larger level in the range above 5 kHz. With /**s**/ being the "sharper" of the two sounds, these spectra were expected to be of this form. Comparing the vowel sounds from low frequencies to high frequencies shows that the sounds /**u**/ and /**i**/ have similar spectra between 100 Hz and 800 Hz. The differences of the two sounds start to occur at 800 Hz with /**u**/ showing a peak almost 20 dB SPL larger in the frequency range between 800 Hz and 2 kHz compared to /**i**/. Above 2 kHz on the other hand, the sound /**i**/ shows a significantly higher level peaking in a level difference of almost 30 dB SPL at about 3 kHz.
As expected, the sound /**a**/ differs significantly from the rest of the vowels, especially in the mid-frequency range of about 700 Hz to about 2 kHz. In this frequency range, /**a**/ shows the highest energy and an overall spectrum with the least pronounced peaks compared to the other vowels. Finally, the sound /**m**/ shows high similarity with /**u**/ and /**i**/ between 200 Hz and 800 Hz, except for a 10dB notch at 300 Hz compared to /**i**/. At about 1.3 kHz it shares a peak with /u/ and above 1.6 kHz it shows a fairly flat behavior with another notch at 4 kHz.
For the synthetic Ling-6 sounds, figure 3.2 shows the spectral analysis of the fricatives /**s**/ and / **ʃ** / and figure 3.3 the spectral analysis of the voiced sounds /**a**/, /**i**/, /**m**/ and /**u**/ in form of a power spectrum, whereas 0 dB corresponds to the digital full scale. Comparing the Ling-6 sounds by Scollie et al. (2012) and the synthesized Ling-6 sounds used in this study, it can be noted that the overall shapes of the spectra are similar. There are two main differences between the two versions. Firstly, the synthesized sounds show more energy in the low frequency range between 100 Hz and 200 Hz. While the recorded Ling-6 sounds by Scollie et al. (2012) show hardly any energy in this frequency range, the synthesized voiced sounds show large peaks at 100-130 Hz followed by a 20-dB notch and another peak at 200-230 Hz. Secondly, the synthesized sounds show a slight overall shift of the formant peaks towards lower frequencies. The only sound that did not show comparable peaks when comparing the two versions was the sound / **ʃ** /. Above 4kHz, the recorded sound showed a decrease in energy, while the synthesized sound increased in energy and even peaked at 6.6 kHz with more energy than the sound /**s**/.

3.2 Evaluation of synthetic Ling-6 sounds

For the evaluation of the synthetic Ling-6 sounds, the methods were used as described in section 2.1.3. After testing, the individual responses as well as the recognition score were stored in text files for each subject. The data was then imported into Microsoft Excel for further analysis and preparation. Table 3.1 shows the recognition score for each subject and the mean recognition score of all subjects combined. Three of the five subjects identified all the sounds correctly and the mean recognition score was 97%.

Subject	S1	S2	S3	S4	S5	Mean
RS in %	88.33	100	96,67	100	100	97

Table 3.1: Recognition score (RS) for each subject as well as mean recognition score

The confusion matrix for the first evaluation of the synthetic sounds is shown in Table 3.2. The green boxes indicate the correctly identified sounds and the orange boxes the confusions with other possible answers. The sounds /**a**/, /**m**/ and / **ʃ** / were identified correctly 100% of the time. The sound /**i**/ was confused with **E** five times (RS = 90%), the sound /**u**/ with **O** three times (RS = 94%) and the sound /**s**/ with **F** only once (RS = 98%). No confusions within the Ling-6 sounds occurred.

		Played sound					
		/a/	/i/	/u/	/m/	/s/	/ ʃ /
Heard Sound	/a/	50	0	0	0	0	0
	/i/	0	45	0	0	0	0
	/u/	0	0	47	0	0	0
	/m/	0	0	0	50	0	0
	/s/	0	0	0	0	49	0
	/ ʃ /	0	0	0	0	0	50
	E	0	5	0	0	0	0
	O	0	0	3	0	0	0
	F	0	0	0	0	1	0
	N.A.	0	0	0	0	0	0

Table 3.2: Confusion matrix for evaluation of synthetic Ling-6 sounds by normal hearing subjects. E, O and F are not written in phonetic transcription because they only represent possible answers and not actual sound samples.

3.3 Study results

The initial goal of the study was to test 20 subjects. Due to the COVID-19 health crisis in spring 2020, only ten patients could be tested up to the point of writing this thesis. Therefore, this thesis presents the study results of these ten subjects and discusses certain trends that can be observed in these results. The first three patients cannot be included in the results of the speech tests because after they were tested, a re-normalization according to subjective loudness perception had to be performed as described in section 2.1.2. Therefore, the results of the speech tests will only include the remaining seven patients. In this section the term improvement will be used to describe the differences between the test conditions, whereas a negative improvement corresponds to a deterioration in performance.

3.3.1 Pure tone audiometry

Due to a protocol deviation, the first two subjects were not tested according to the test conditions defined in section 2.2.5. With these subjects the MCLs in their fitting maps were reduced by 45% instead of 30%. Therefore, the results for the first two patients will be treated separately. First, the results for the remaining eight patients will be shown.

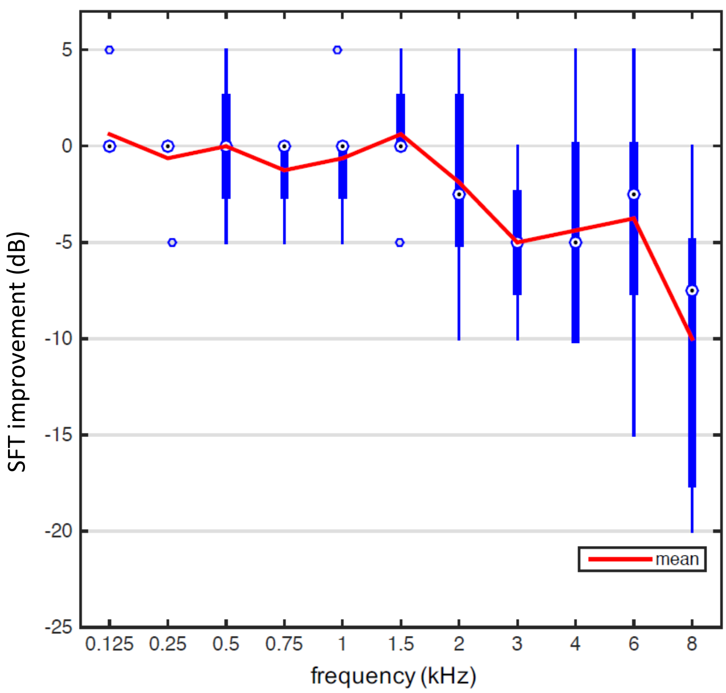

Figure 3.4: Improvement in sound field threshold between the fitting maps '**clinical**' and '**high**'. The number of subjects was eight and the respective decrease in MCLs was 30%. The red line represents the mean difference of all eight subjects for the respective frequency and the dotted circles are the median values.

Figure 3.4 shows the improvement in sound field threshold between the clinical fitting map and the map where MCLs were reduced by 30 % in the frequencies above 2600 Hz. When looking at the mean values in Figure 3.4, from 125 Hz up to 1.5 kHz the mean difference is almost zero, while the median values are exactly zero. Starting at 2 kHz the mean values decrease. From 3 kHz up to 6 kHz they are about -5 dB and reach their maximum difference at 8 kHz with -10 dB. The median values from 2 kHz up to 8 kHz are exactly -5 dB.

Figure 3.5 shows the improvement in sound field threshold between the clinical fitting map and the map where MCLs were reduced by 30% in the frequencies between 750 Hz and 2600 Hz. The mean values indicate a decrease of SFTs between 500 Hz and 3 kHz, especially at 2 kHz, but the results are not as pronounced as in Figure 3.4.

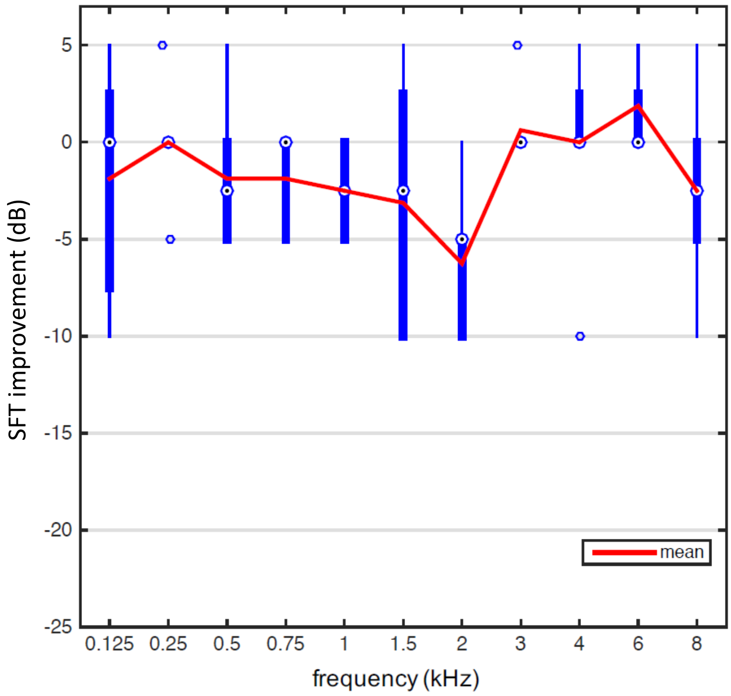

Figure 3.5: Improvement in sound field threshold between the fitting maps '**clinical**' and '**mid**'. The number of subjects was eight and the respective decrease in MCLs was 30%. The red line represents the mean difference of all eight subjects for the respective frequency and the dotted circles are the median values.

Figure 3.6 finally shows the improvement in sound field threshold between the clinical fitting map and the map with MCLs reduced by 30% in the frequency range below 750 Hz. Between 125 Hz and 500 Hz a decrease in SFT of about -5 dB can be seen. At 1 kHz and above, the mean difference in SFT is approximately 0 dB. Only at 8 kHz another drop of about -4 dB can be observed but with a fairly large variation in the results. The median values at 500 Hz and below represent exactly one discrete step of -5 dB in the audiometric measurement. All the median values from 750 Hz and above are 0 dB, again except for the one at 8 kHz being at -2.5 dB.

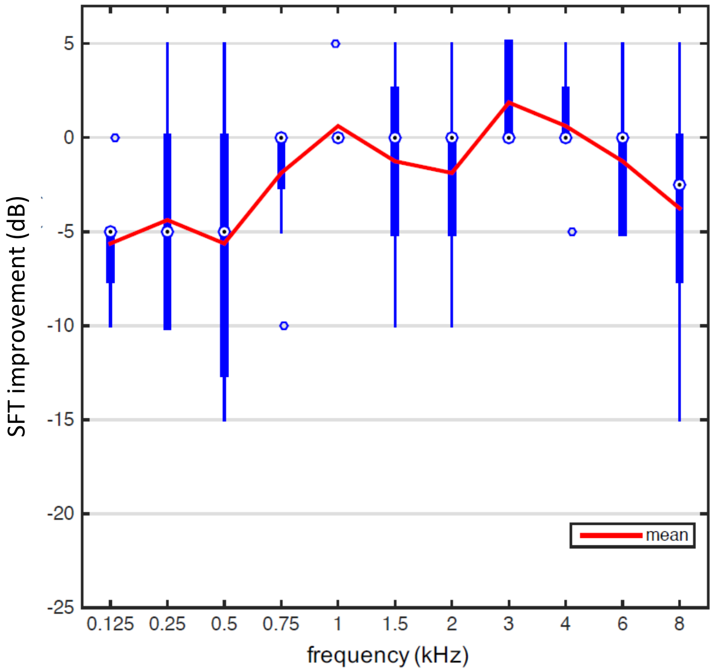

Figure 3.6: Improvement in sound field threshold between the fitting maps '**clinical**' and '**low**'. The number of subjects was eight and the respective decrease in MCLs was 30%. The red line represents the mean difference of all eight subjects for the respective frequency and the dotted circles are the median values.

The PTA for all test frequencies within the frequency range of each test condition was calculated and the difference in PTA compared to test condition 'clinical' was computed. Figure 3.7 shows the results for the PTA differences compared to the clinical test condition. The different PTA values include test frequencies as follows:

- **PTA_high**: 3 kHz, 4 kHz, 6 kHz, 8 kHz
- **PTA_mid**: 0.75 kHz, 1 kHz, 1.5 kHz, 2 kHz
- **PTA_low**: 125 Hz, 250 Hz, 500 Hz

A Kruskal-Wallis test indicated a significant difference within PTA_high ($p = 0.0368$). Consequently, a two-tailed Wilcoxon signed rank test was performed to further analyze this data set and it showed a significant difference for PTA_high in test condition 'high' ($p = 0.0498$, Bonferroni correction for multiple testing applied).

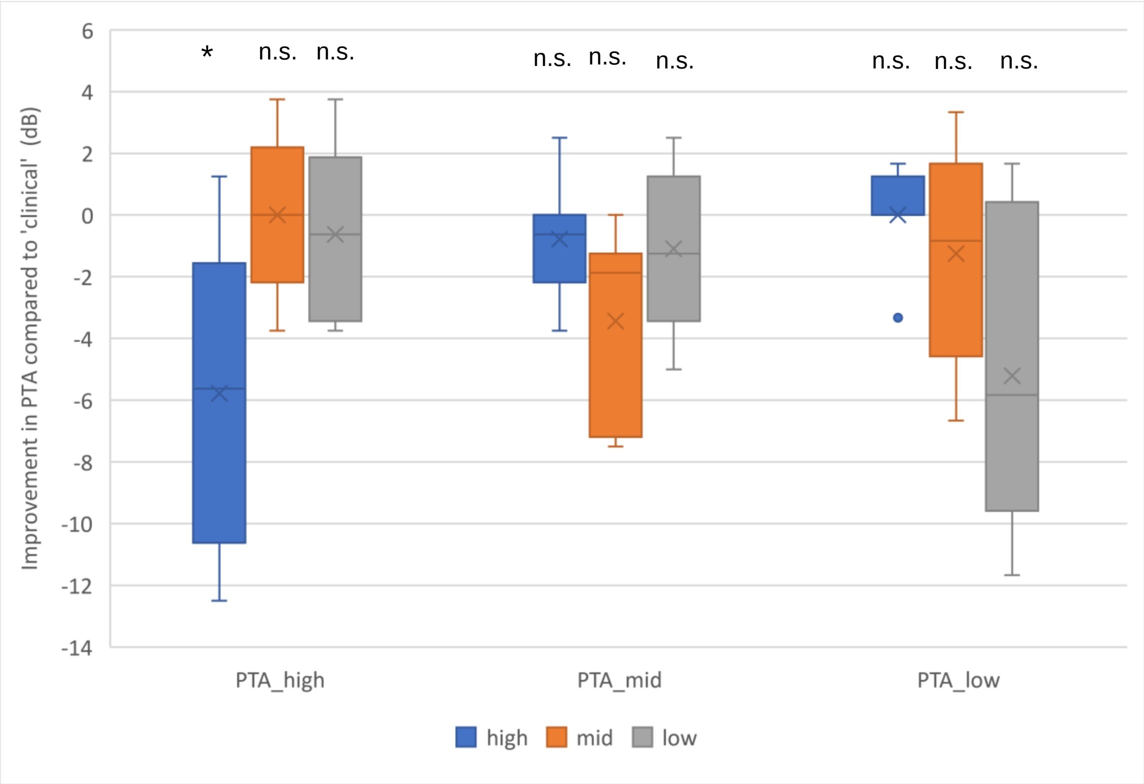

Figure 3.7: Improvement in PTA compared to the clinical test condition for each test condition. X indicates the mean value, the horizontal line within the interquartile range indicates the median value and dots indicate outliers. p-values: * p ≤ 0.05; n.s. (not significant)

Next, the differences in sound field threshold of the two first subjects will be presented. Figure 3.8 shows the differences in SFT when MCLs were reduced by 45% in the frequencies above 2600 Hz. Since this plot only includes two subjects, the mean and median are identical. Figure 3.7 shows a decrease in SFT by -7.5 dB and even a slight increase at 1.5 kHz and 2 kHz. Below 750 Hz, we can again observe a slight decrease in SFT. Figure 3.9 shows the differences in SFT when MCLs were reduced by 45% in the frequencies between 750 Hz and 2600 Hz. In this plot, no clear indications are visible, that the reduction of MCLs by 45 % results in a decrease in SFT in the respective frequency range. Figure 3.10 shows the differences in SFT when MCLs were reduced by 45% in the frequencies below 750 Hz. In the range between 125 Hz and 500 Hz a decrease in SFT between -7.5 dB and -12.5 dB can be seen. Above 500 Hz, no considerable decrease or increase can be observed.

For these two subjects, PTA values were not calculated, and statistical analysis was not performed due to the low number of subjects.

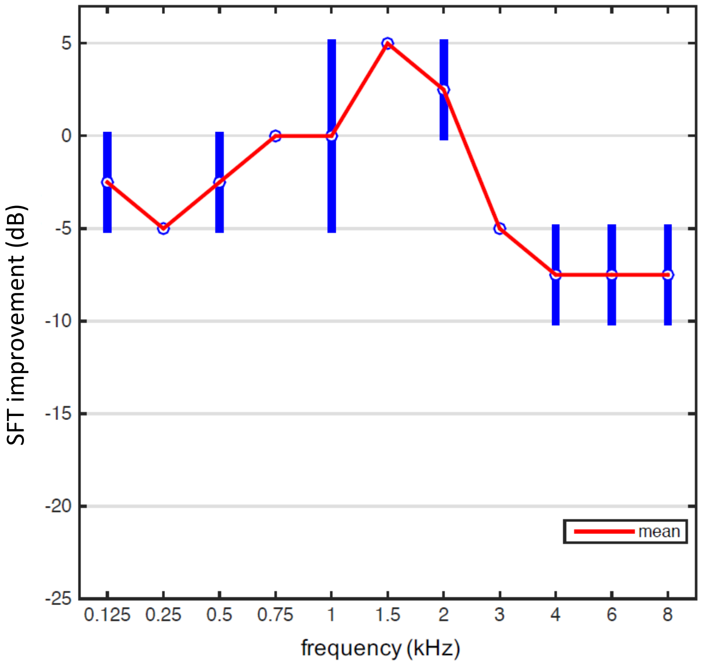

Figure 3.8: Improvement in sound field threshold between the fitting maps '**clinical**' and '**high**'. The number of subjects was two and the respective decrease in MCLs was 45%. Since only two subjects were included in this plot, the mean (red line) and the median (dotted circles) are identical.

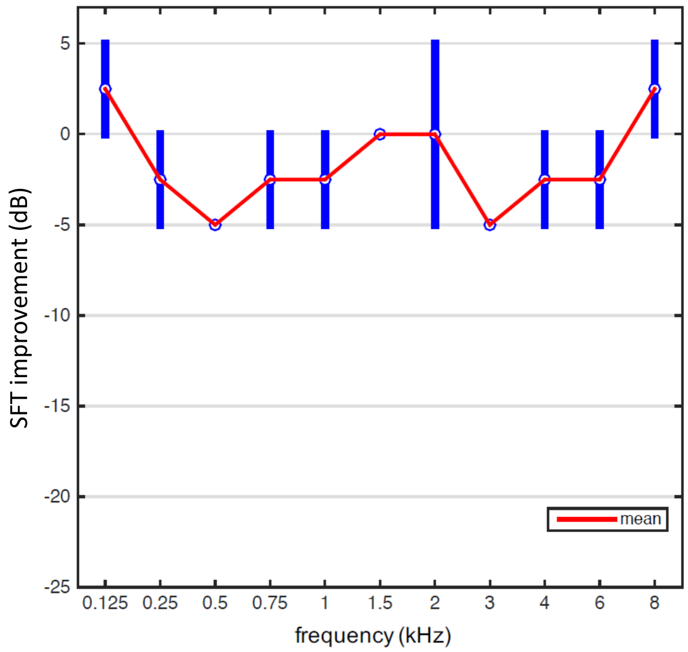

Figure 3.9: Improvement in sound field threshold between the fitting maps "**clinical**" and "**mid**". The number of subjects was two and the respective decrease in MCLs was 45%. Since only two subjects were included in this plot, the mean (red line) and the median (dotted circles) are identical.

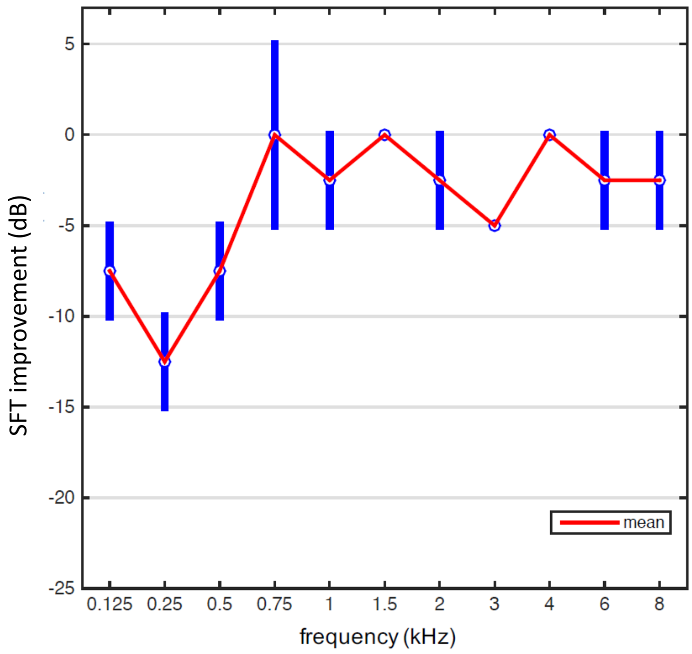

Figure 3.10: Improvement in sound field threshold between the fitting maps '**clinical**' and '**low**'. The number of subjects was two and the respective decrease in MCLs was 45%. Since only two subjects were included in this plot, the mean (red line) and the median (dotted circles) are identical.

3.3.2 Freiburg monosyllable test

Figure 3.11 shows the results for the Freiburg monosyllable test in form of box plots and the median values. Median values as well as the first quartile (Q1) and third quartile (Q3) are also shown in table 3.3.

Test condition	clinical	high	mid	low
Median WRS in %	67.5	65	62.5	57.5
Q3	79.4	75.6	75.0	76.3
Q1	54.4	50.0	40.6	44.4

Table 3.3: Descriptive statistical data corresponding to figure 3.10. The table shows the median word recognition scores (WRS) for each test condition as well as the first (Q1) and third quartile (Q3).

The differences in median WRS compared to the 'clinical' test condition range between 2.5 % and 10 %. The box plots in figure 3.10 show a large variability for each test condition, which is caused by the varying levels of performance of each subject. The improvement in WRS compared to the clinical map is shown in figure 3.12.

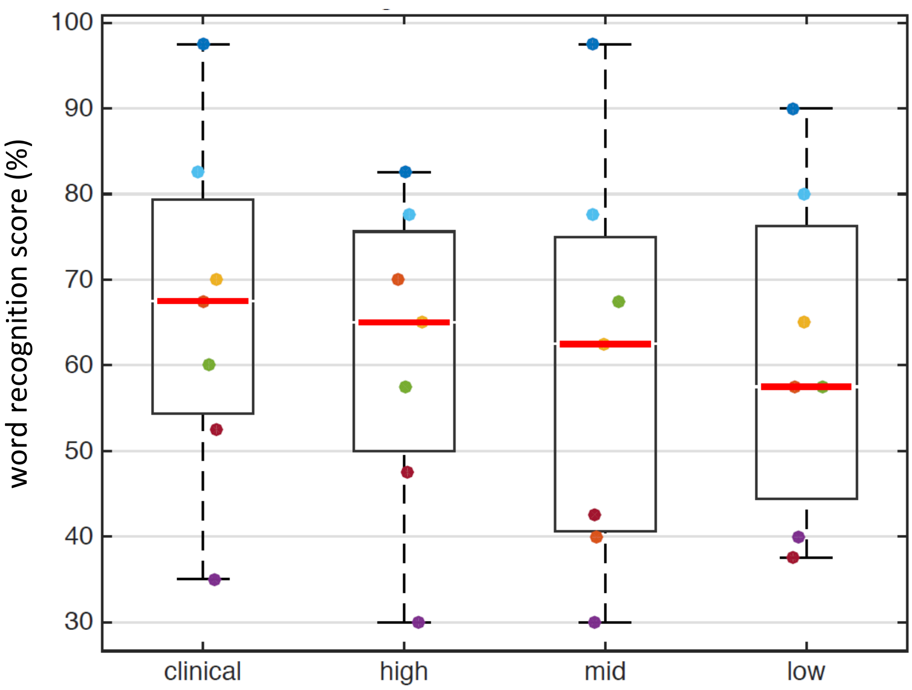

Figure 3.11: Word recognition score for the Freiburger monosyllabic test at 65 dB SPL for each test condition. The colored dots represent the data points for each subject.

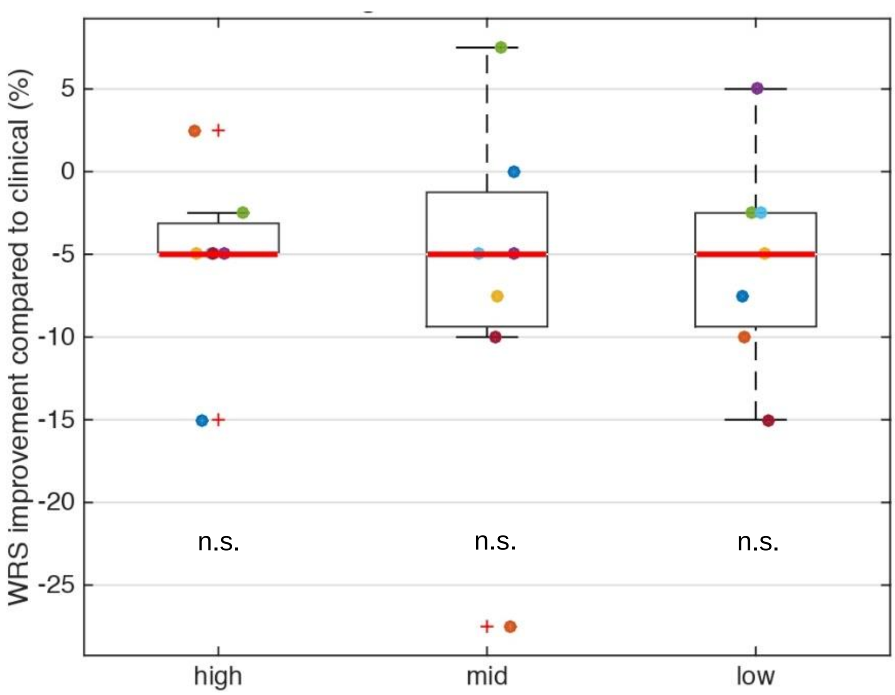

Figure 3.12: Improvement in WRS for the Freiburger monosyllabic test at 65 dB SPL compared to the test condition ‘clinical’. The colored dots represent the data points for each subject. n.s. (not significant)

We can see a median improvement of WRS of -5 % for each test condition. The results for each subject vary strongly. For example, one subject (orange data points) shows a slight improvement in WRS of about 2.5 % in the ‘high’ test condition, a strong decrease of -27.5 % in the ‘mid’ test condition and a less pronounced decrease of -10 % in the ‘low’ test condition. When looking at another subject (blue data points), we can see a completely different response to the map changes. In this subject in ‘high’ test condition, we can see a decrease in WRS of -15 %, in the ‘mid’ condition no improvement, and in the ‘low’ condition again a decrease of -7.5 %. A Kruskal-Wallis test showed no significant differences between the different test conditions (p = 0.105).

3.3.3 Oldenburg sentence test

The Oldenburg sentence test (OLSA) could not be performed in all subjects. One of the seven subjects who are included in the results of the speech tests could not perform the OLSA at all due to not understanding the test procedure and two only could perform one test run due to mental fatigue. Therefore, the results include six subjects, and for two of those only one test run will be included. Table 3.4 shows the median speech reception threshold (SRT) for each test condition as well as the values or the first (Q1) and third (Q3) quartile. The SRT corresponds to the signal to noise ratio of the test stimuli which lead to 50 % speech understanding.

Test condition	clinical	high	mid	low
Median SRT in dB	3.71	5.44	3.43	5.13
Q3	5.40	6.06	4.41	7.50
Q1	3.27	2.17	2.00	4.27

Table 3.4: Descriptive statistical data corresponding to figure 3.12. It shows the median speech reception threshold (SRT) for each test condition as well as first (Q1) and third quartile (Q3).

The best results were achieved in the ‘mid’ test condition (3.43 dB) and the worst in the ‘high’ test condition (5.44 dB). This leads to a maximum difference of 2.01 dB between the test conditions. Figure 3.13 shows the SRT for each test condition for the OLSA. We can again see a fairly large variability, especially with the subject represented by the light blue data points, who performed noticeably better than the others. The SRT improvement compared to the ‘clinical’ test condition is shown in figure 3.14.

Even though the overall performance was lowest in the ‘high’ test condition, the deterioration in performance compared to the clinical map was more pronounced in the ‘low’ test condition. A Kruskal-Wallis test showed no significant differences between the different test conditions (p = 0.088).

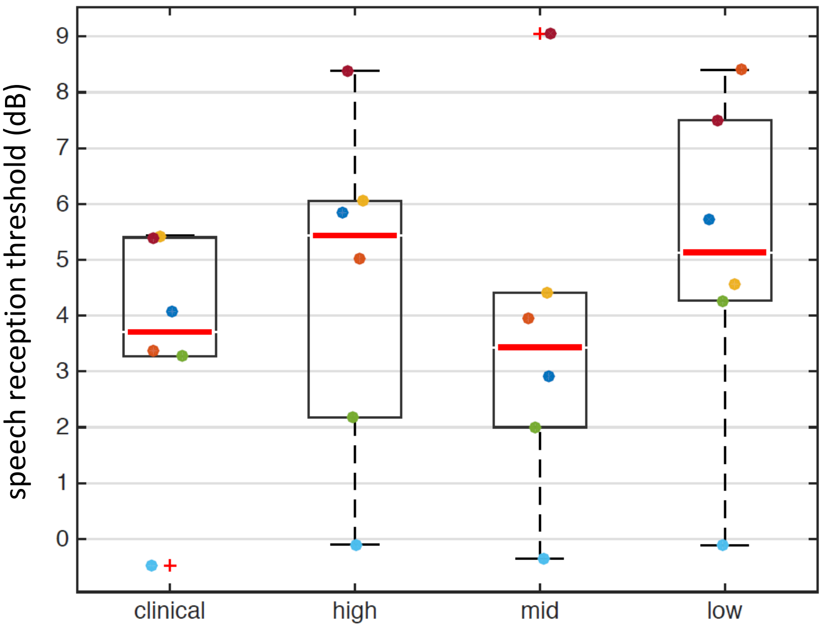

Figure 3.13: Speech reception threshold for the OLSA for each test condition. The colored dots represent the data points for each subject.

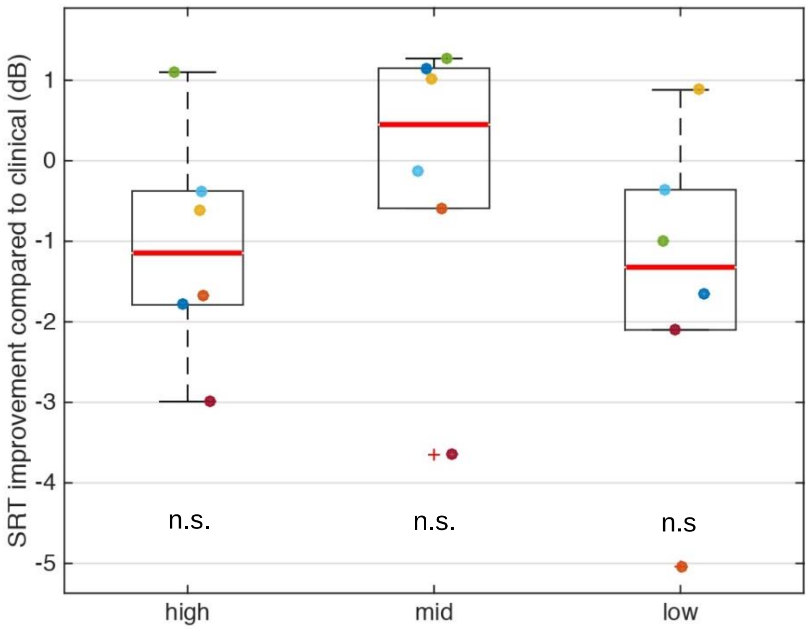

Figure 3.14: Improvement in SRT for the OLSA compared to the test condition ‘clinical’. The colored dots represent the data points for each subject. n.s. (not significant)

3.3.4 Synthetic Ling-6 sounds test

The results for the synthetic Ling-6 sounds test are plotted in figure 3.15. The median percent correct values, as well as first and third quartile are shown in table 3.5.

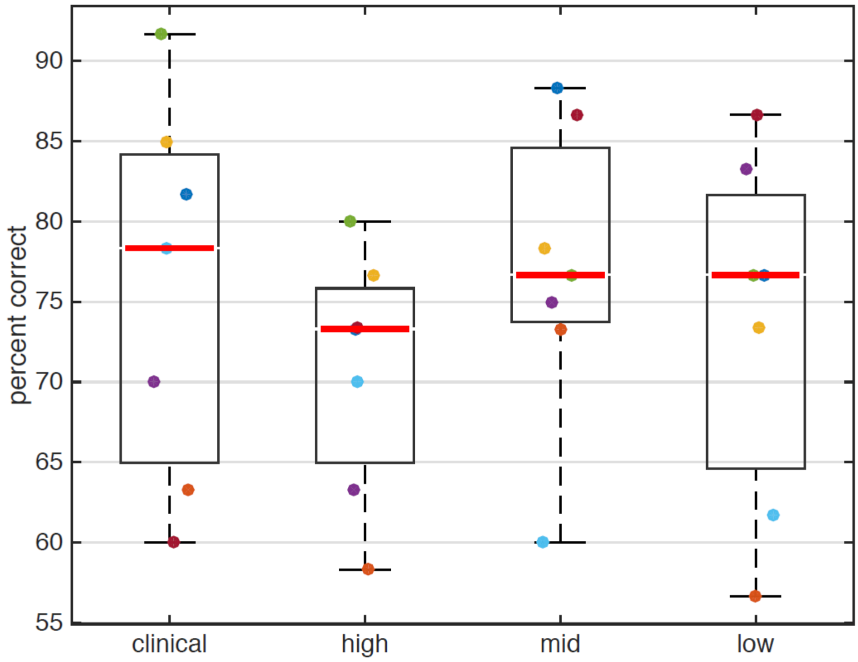

Figure 3.15: Recognition scores for the Ling-6 sound test in percent correct for each test condition.

Test condition	clinical	high	mid	low
Median RS in %	78.3	73.3	76.7	76.7
Q3	84.2	75.8	84.6	81.6
Q1	64.9	64.9	73.7	64.6

Table 3.5: Descriptive statistical data corresponding to figure 3.14. It shows the median recognition score (RS) for the Ling-6 sound test for each test condition as well as first (Q1) and third quartile (Q3).

The median recognition score (RS) was highest in the ‘clinical’ test condition and the overall maximum RS achieved by a subject was also with the clinical map (91.7 %). The overall minimum RS was 56.7 % and was found in the ‘low’ test condition.

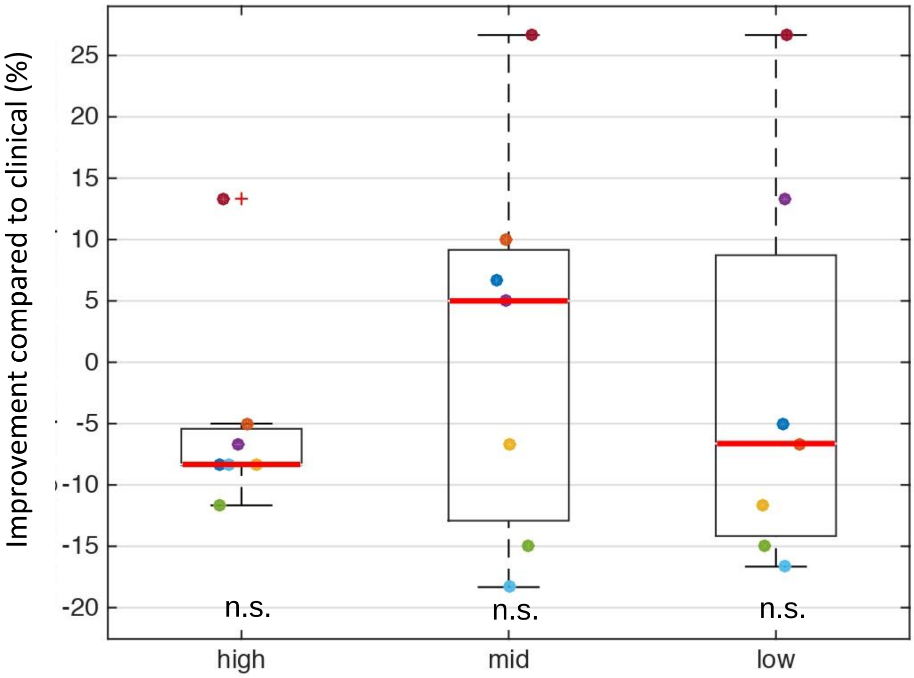

Figure 3.16: Improvement in recognition score for the Ling-6 sound test in percent correct compared to the test condition 'clinical'. n.s. (not significant)

Figure 3.16 shows the improvement in RS for the Ling-6 sound test. Except for one subject, all subjects showed a decrease in RS of at least -5 % in the 'high' test condition. In the 'mid' test condition, more than half of the subjects showed an improvement in RS compared to their clinical map, and in the 'low' test condition, five out of seven subjects again showed a decrease in RS of at least -5 %. It shall be noted that only one subject (wine red data points) showed the lowest performance in the Ling-6 sound test with his clinical map, and improvements with all other maps. A Kruskal-Wallis test did not show significant differences between the different test conditions (p= 0.353).

Figure 3.17 shows the confusion matrices for each test condition. The green boxes indicate the correctly identified sounds.

Figure 3.18 shows the improvement in recognition score for each Ling-6 sound in each test condition and Table 3.6 shows the corresponding descriptive statistical data. The results in figure 3.17 and 3.18 will be described in detail in combination with the discussion in section 4.3.4.

(a)		/a/	/i/	/u/	/m/	/s/	/ʃ/
		Played sound					
Heard Sound	/a/	68	0	0	0	0	0
	/i/	0	42	2	2	8	3
	/u/	2	9	49	28	0	0
	/m/	0	19	19	40	0	0
	/s/	0	0	0	0	60	8
	/ʃ/	0	0	0	0	2	59

(b)		/a/	/i/	/u/	/m/	/s/	/ʃ/
		Played sound					
Heard Sound	/a/	70	0	0	0	0	0
	/i/	0	30	3	6	4	4
	/u/	0	8	54	34	0	0
	/m/	0	32	13	30	0	0
	/s/	0	0	0	0	64	17
	/ʃ/	0	0	0	0	2	49

(c)		/a/	/i/	/u/	/m/	/s/	/ʃ/
		Played sound					
Heard Sound	/a/	70	0	0	0	0	0
	/i/	0	61	2	1	2	0
	/u/	0	2	53	27	0	0
	/m/	0	7	15	42	0	0
	/s/	0	0	0	0	53	26
	/ʃ/	0	0	0	0	15	44

(d)		/a/	/i/	/u/	/m/	/s/	/ʃ/
		Played sound					
Heard Sound	/a/	70	0	0	1	1	0
	/i/	0	53	1	5	0	0
	/u/	0	3	45	29	0	0
	/m/	0	9	24	34	0	0
	/s/	0	4	0	1	55	18
	/ʃ/	0	1	0	0	14	52

Figure 3.17: Confusion matrices for the Ling-6 sound test for each test condition ((a) clinical, (b) high, (c) mid, (d) low). The rows correspond to the sounds that were presented in the test and the lines correspond to the sound that was heard and responded by the subject.

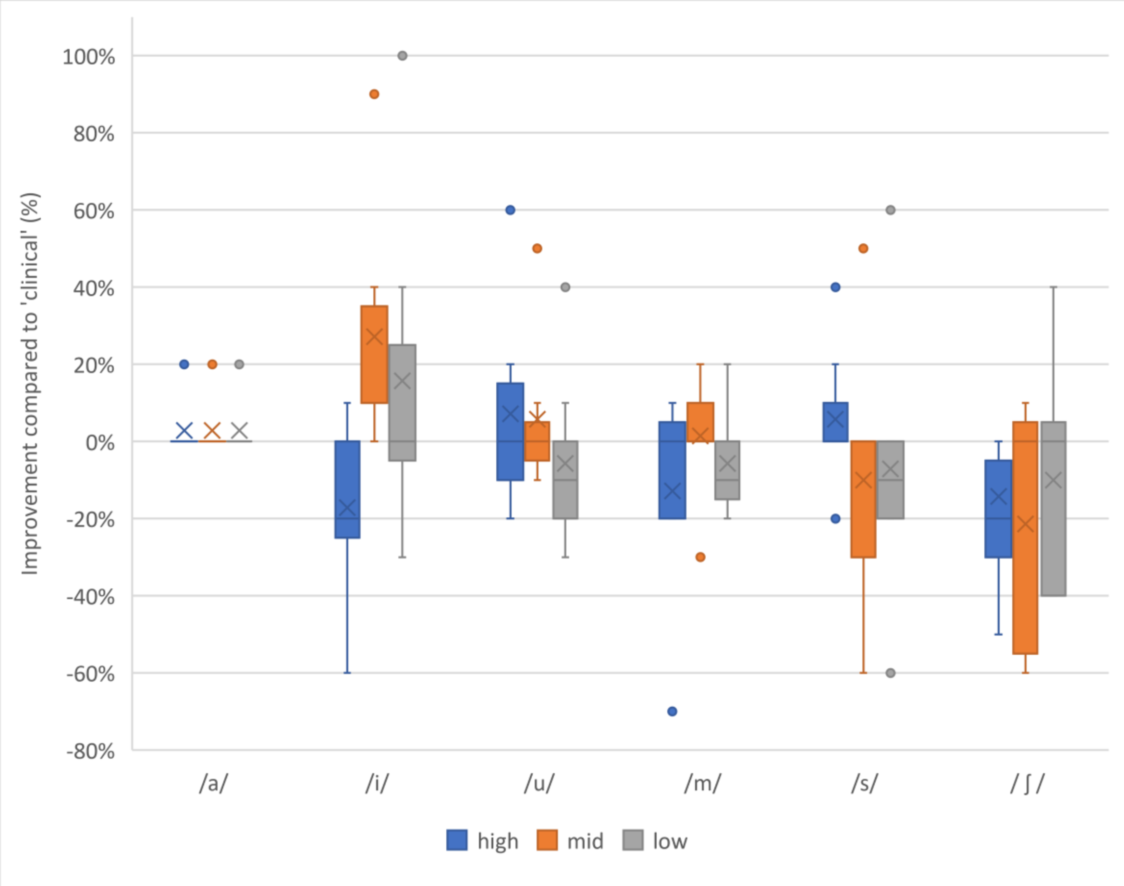

Figure 3.18: Improvement in recognition for each individual Ling-6 sound and each test condition. X indicates the mean value, the horizontal line within the interquartile range indicates the median value and dots indicate outliers.

	/a/			/i/			/u/		
	high	**mid**	**low**	**high**	**mid**	**low**	**high**	**mid**	**low**
Median	0.0%	0.0%	0.0%	-20.0%	10.0%	0.0%	0.0%	0.0%	-10.0%
Q3	0.0%	0.0%	0.0%	0.0%	35.0%	25.0%	15.0%	5.0%	0.0%
Q1	0.0%	0.0%	0.0%	-25.0%	10.0%	-5.0%	-10.0%	-5.0%	-20.0%
Mean	2.9%	2.9%	2.9%	-17.1%	27.1%	15.7%	7.1%	5.7%	-5.7%
σ	7.6%	7.6%	7.6%	23.6%	30.9%	42.8%	27.5%	20.7%	24.4%

	/m/			/s/			/ ʃ /		
	high	**mid**	**low**	**high**	**mid**	**low**	**high**	**mid**	**low**
Median	0.0%	0.0%	-10.0%	0.0%	0.0%	-10.0%	-20.0%	0.0%	0.0%
Q3	5.0%	10.0%	0.0%	10.0%	0.0%	0.0%	-30.0%	-55.0%	-40.0%
Q1	-20.0%	0.0%	-15.0%	0.0%	-30.0%	-20.0%	-5.0%	5.0%	5.0%
Mean	-12.9%	1.4%	-5.7%	5.7%	-10.0%	-7.1%	-14.3%	-21.4%	-10.0%
σ	28.1%	15.7%	14.0%	19.0%	38.7%	35.9%	29.4%	33.4%	31.1%

Table 3.6: Descriptive statistical data for improvement of each Ling-6 sound in each test condition. The median and mean improvement as well as the first quartile Q1, third quartile Q3 and the standard deviation σ are shown.

4 Discussion

4.1 Spectral analysis of Ling-6 sounds

The analysis of the Ling-6 sounds used by Scollie et al. (2012) and the synthetic Ling-6 sounds used in this study (section 3.1) showed that overall the sounds share high similarity apart from a few exceptions. Differences were found to be in the frequency range below 2 kHz, where the synthetic Ling-6 sounds showed more energy than the recorded sounds by Scollie et al. (2012). Additionally, a slight overall shift of the formant peaks was found. The reason for these differences lies in the gender of the respective speaker. While the sounds by Scollie et al. (2012) were recorded from a female speaker, the synthesized sounds were based on a male speaker. As indicated in table 1.1, this difference in F0 accounts for approximately 100 Hz. The lower frequency content of the synthetic Ling-6 sounds was chosen for the study because of the larger frequency range that could be tested.

One more difference was found when comparing the sounds / **ʃ** / of both studies. The synthetic / **ʃ** / showed a peak above 4 kHz at 6.6 kHz, while the levels of the recorded sound decreased above 4 kHz. Whether this difference affects the performance in the Ling-6 sound test will be discussed in section 4.3.4.

4.2 Evaluation of synthetic Ling-6 sounds

The synthesized Ling-6 sound samples had to be evaluated to ensure that they are clearly distinguishable as the sounds that they should represent. The basic requirement for the sounds to be used in the study was a recognition score of nearly 100% by normal hearing subjects. This defined requirement ensured that the sounds will not be misheard due to poor quality of the sound samples.

With a mean recognition score of 97% in five subjects, the requirement was found to be fulfilled (section 3.2). When looking at the confusion matrix in table 3.2 we can even see that no confusions within the Ling-6 sounds occurred but only with sounds that were added as additional possible answers. Since the final test only included the Ling-6 sounds as a closed set of possible answers, the other confusions were considered irrelevant and the synthesized sound samples were considered suitable for the study.

4.3 Study results

MCL is one of the most important parameters in CI fitting and is usually the first parameter that is adjusted or verified if a patient reports problems with their hearing or with speech understanding. In clinical practice, correctly set MCLs form a basis for fitting maps. To investigate this important fitting parameter the following tests were conducted.

4.3.1 Pure tone audiometry

Sound field thresholds (SFTs) for the four test conditions were assessed. The results of section 3.3.1 show the trend that SFTs are affected only if MCLs are lowered in high and low frequencies. SFTs were only affected mildly when MCLs were lowered in the mid-frequency range. Only reducing MCLs in high frequencies showed a significant difference of -5.8 dB in PTA for frequencies between 3 kHz and 8 kHz. Still, the shown decrease of SFTs in high and low frequencies of about -5 dB would not be considered clinically relevant since a 5 dB change corresponds to only one discrete measurement step in sound field audiometry.

When reducing MCLs by 45 % instead of 30 %, the effects in high and low frequencies were more pronounced with a decrease of approximately -7.5 dB (figures 3.8 and 3.10). Again, the mid-frequency range was not significantly affected (figure 3.9). Figure 3.9 shows surprisingly that a substantial decrease of -45 % in MCL has almost no influence in the manipulated frequency range, however, since only two subjects were tested in that condition, this finding is questionable.

After the initial fitting phase of about two months (see section 1.2.2), the adjustments in MCLs typically range between 3-9 % increase or decrease. This is significantly less than in the test conditions of this study. Therefore, it can be assumed that, if large changes in MCLs have little effect on SFTs, more subtle changes as typical in clinical practice will have even less effect. One possible reason for this behavior might be that MCL is the parameter defining the maximum loudness which will be perceived by a CI patient. It is the stimulation charge, that will be applied if loud sounds of about 100-110 dB SPL are picked up by the audio processor microphone. Thus, it hardly affects low level sounds as presented in pure tone audiometry.

The major parameter that does affect low level sounds is THR. It defines the lowest perceivable stimulation charge and therefore should have much more influence on SFTs.

When looking at the literature investigating the effects of THR on SFTs, it was shown that there is little difference in SFTs between accurately setting THRs and setting them to an arbitrary, but still low level (Sainz, de la Torre, Roldän, Ruiz, & Vargas, 2003). The difference was shown to be up to approximately 5 dB when setting THR accurately. Similar results were shown by Boyd (2006), showing no significant difference between accurately set THRs and THRs set to an arbitrary level of 10 % of MCL. Still, both papers showed tendencies towards light improvements in SFTs when accurately measuring THRs especially in patients with reduced dynamic range due to e.g. cochlear fibrosis or otosclerosis.

In conclusion, the results show that pure tone audiometry is not highly sensitive to large changes in fitting maps, at least when manipulating MCLs. When looking at SFTs and fitting parameters in general, there is no fitting parameter affecting SFTs on its own. MCL and THR both have at best only mild effects on pure tone audiometry. Finally, the importance of larger numbers of subjects in measuring SFTs should be pointed out, so that statements about trends could be made more reliably.

4.3.2 Freiburg monosyllable test

The Freiburg monosyllable test is the most frequently used speech recognition test in clinical practice, but it is also often debated (Steffens, 2015). In clinical practice the Freiburg speech test is usually used in combination with pure tone audiometry to determine the performance of CI patients. The preliminary results of this study confirm that even patients who subjectively feel to have proper speech understanding, will not always perform well in speech tests. Looking at figure 3.11, there is a large variability in word recognition scores (WRS) meaning the subjects varied strongly in performance. In 'clinical' test condition they performed slightly better than in the other test conditions. Still, the maximum difference in median WRS between 'clinical' and 'low' is only 10 %. Typically, values smaller than 10 % are not considered clinically relevant. The results in figure 3.12 show that there is a median decrease in WRS of -5 % for each test condition but again, the decrease is too small to be considered clinically relevant. Statistical analysis of the results for the improvements with a Kruskal-Wallis test showed no significant differences.

In conclusion, the results presented here indicate that the Freiburger monosyllabic test is not sensitive to large changes in fitting maps.

4.3.3 Oldenburg sentence test

The Oldenburg sentence test (OLSA) is most commonly used to assess speech understanding in noise. Therefore, not only the speech signal but also the noise signal will be affected by changes in the fitting maps. The results in section 3.3.3 showed the best median SRT to be in the 'mid' test condition. Compared to 'high' test condition, the performance is more than 2 dB better. This difference would usually be considered clinically relevant. The good performance in the 'mid' test condition was rather surprising, mainly because the strong decrease of MCLs in the mid-frequency range was expected to negatively affect the perception of speech, however, the opposite was found. Compared to the 'clinical' test condition there was even a slight improvement of SRT of about 0.5 dB. Statistical analysis of the results for the improvements with a Kruskal-Wallis test showed no significant differences.

In conclusion, the results suggest that the OLSA is minimally sensitive to large changes in fitting maps.

4.3.4 Ling-6 sound test

In this study, a new speech test based on synthetic Ling-6 sounds in detection form was investigated. The study was aimed at showing whether the speech test is sensitive to changes in fitting maps of cochlear implant patients, and whether characteristic confusions between the Ling-6 sounds occurred. The differences in recognition score (figure 3.15) between the test conditions were too small to confidently conclude that the performance was best with the patients preferred clinical map, but the subjects tended to have at least slightly better results with this map. When looking at the improvement of each subject in the different test conditions, it was shown that reducing high frequencies tended to decrease the performance in the Ling-6 test. When reducing MCLs in the mid-frequency range, a slight improvement could be observed and when reducing MCLs in low frequencies a slight deterioration could be shown. Statistical analysis of the results for the improvements with a Kruskal-Wallis test showed no significant differences. The results should therefore be reevaluated once more subjects have been tested.

One of the main interests of this study was the results of the confusion matrices in figure 3.17 and the improvement for each sound in the different test conditions for each subject (figure 3.18). The confusions for each sound will be discussed individually.

The sound /**a**/ showed no confusions except in the test condition 'clinical'. Here, the sound was confused with /**u**/ two times. When looking closer at where these confusions occurred, it was found that these happened in one subject (ID07) where the clinical map was the first map to be tested with the Ling-6 sound test, and the confusions occurred within the first four presentations of sounds. Therefore, we can assume that these confusions occurred during a phase of familiarization with the test procedure, especially since all other /**a**/ sounds were correctly identified afterwards. These findings suggest that using /**a**/ in the detection form of the Ling-6 sound test might be pointless because there is no information gain concerning the audiometric assessment when the sound is correctly identified 100 % of the time. Still, these results support the need of a training or familiarization with the test.

The sound /**i**/ shows a median deterioration of -20 % and noticeably more confusions with the sound /**m**/ in the 'high' test condition compared to the 'clinical' test condition. When looking at the spectra of those two sounds in figure 3.3, this is not surprising. The decrease in MCLs at frequencies above 2.6 kHz leads to a decrease of the F2 formant at 3.3 kHz of the sound /**i**/ thus making the spectra of the two sounds more alike. This should increase the probability for the two sounds to be confused.

In the 'mid' test condition, the sound /**i**/ showed less confusions with other sounds in each subject except for one, in which it stayed the same. When reducing the MCLs for frequencies between 750 Hz and 2.6 kHz, the formant frequencies of /**i**/ are mostly unaffected but the frequencies in between are attenuated. Therefore, the formant peaks are more pronounced possibly making the sound easier to distinguish.

When reducing low frequencies in the fitting map, the sound /**i**/ showed no improvement in the median value, however, 70 % of the data points are in the range of improvement and the mean value indicates an improvement of 15.7%. This again, is not surprising because a reduction of low frequencies results in a reduction of the frequency range characteristic for /**m**/ and /**u**/, making the sound /**i**/ easier to distinguish. In the 'low' test condition, one subject (ID05) also confused /**i**/ with /**s**/ and / **ʃ** /. Since this subject tends to be an outlier throughout all other test conditions in the Ling-6 sound as well, this subject seems to have difficulties with the Ling-6 sound test in general.

The sound /**u**/ shows no difference in confusions for manipulations in frequencies above 750 Hz, most likely because /**u**/ and /**m**/ are similarly affected by the decrease in MCLs. In the ‘low’ test condition, slight differences in confusions occurred compared to the clinical map. A decrease in low frequencies might reduce the F2 formant of /**u**/ at about 600 Hz, therefore making it harder to differentiate from /**m**/. Again, one subject (ID10) reacted differently by having the lowest recognition score of /**u**/ in the ‘clinical’ test condition and thus showing only improvements in the other test conditions. This makes the mean results shift towards an improvement in all test conditions and suggests relying more on the median results.

The sound /**m**/ shows the tendency to be harder to distinguish when high frequencies are reduced in the fitting map. For the ‘high’ test condition, the median improvement compared to the clinical map is 0 %, but the mean recognition improvement shows a decrease by -12.9 %. Furthermore, the sound /**m**/ shows more confusions with /**i**/ in this test condition. The reduction of MCLs in high frequencies might reduce the F2 formant of the sound /**i**/, therefore making the spectrum more similar to /**m**/ and causing more confusions. In the ‘mid’ test condition, no significant improvement could be observed. Reduction of MCLs in low frequencies led to a median decrease in recognition by -10 % and one confusion with /**a**/ and one with /**s**/. Since these confusions are rather untypical the results were investigated in more detail and it was observed that these confusions occurred within the first two presentations in the respective test run. This suggests that these confusions happened either due to lack of focus or during familiarization with the test and were therefore considered irrelevant.

The sound /**s**/ showed the second-best recognition score in the clinical test condition after the sound /**a**/. It showed mainly confusions with /**i**/, which all occurred in one subject (ID05). This subject should be considered carefully due to inconsistent responding throughout the testing. The ‘high’ test condition shows no median improvement but a slight mean improvement of 5.7 %. This improvement is caused by subject ID05 because he or she showed less confusions with the sound /**i**/ in this condition. This is rather unexpected, because a decrease in high frequencies would be expected to shift the spectrum of /**s**/ more towards /**i**/ and therefore making them more difficult to differentiate from one another. A decrease of MCLs in the mid-frequency range resulted in more confusions with the sound / **ʃ** /, which is not surprising because this manipulation causes the spectrum of / **ʃ** / to be more similar to /**s**/ by reducing its mid-frequency content. The results of the 'low' test condition again show more confusions with / **ʃ** / compared to the clinical map and one confusion with /**a**/ in subject ID05. This atypical confusion in subject ID05 occurred at the end of the Ling-6 sound testing and might be caused by lack of focus or by fatigue. Therefore, this result will be disregarded when drawing further conclusions.

The sound / **ʃ** / shows a median decrease in recognition of -20 % and more confusions with /**s**/ when high frequencies are reduced. This manipulation seems to shift the spectra of the two sounds closer together, therefore making them harder to differentiate from one another. Subject ID05 again showed atypical confusions with the sound /**i**/.
The 'mid' test condition shows the most confusions with /**s**/ due to the reduction of the characteristic mid-frequency energy of / **ʃ** / compared to /**s**/, therefore making the two sounds harder to differentiate.
Finally, reducing MCLs in low frequencies results in more confusions with /**s**/ compared to the clinical map but less than in the 'mid' test condition. This was surprising because a reduction in low frequencies should affect / **ʃ** / and /**s**/ in the same way and therefore not result in more confusions than in the clinical test condition. We do currently not have an explanation for this result.

One more difference was found when comparing the synthetic / **ʃ** / sound and the recorded one by Scollie et al. (2012) of both studies. The synthetic / **ʃ** / showed a peak above 4 kHz at 6.6 kHz, while the energy of the recorded sound decreased above 4 kHz. The recognition scores for / **ʃ** / and /**s**/ were relatively high in the clinical map. This suggests that the found differences compared to the Ling-6 sounds by Scollie et al. (2012) do not strongly affect the overall recognition of the two sounds. For future studies and investigations in the field of Ling-6 sounds, it might be an option to analyze a larger number of / **ʃ** / sounds and to adjust the synthesized sounds so that the spectra match the naturally spoken phonemes.

4.4 Conclusion and Outlook

The goal of this thesis was to give insight in the prevailing audiometric tests for CI users and to investigate a newly created speech test based on synthesized Ling-6 sounds. A study was conducted at the Medical University of Vienna in cooperation with the hearing implant manufacturer MED-EL to examine the sensitivity of the most commonly used audiometric assessments to changes in fitting maps of CI patients. The primary objective was to investigate whether a new speech test based on Ling-6 sounds is sensitive to changes in a fitting map in high, mid and low frequencies. As the secondary objective, the Freiburg monosyllable test, the Oldenburg sentence test (OLSA) and the assessment of aided pure tone thresholds were also tested for their sensitivity to these changes. Unfortunately, this thesis could only fully cover the results for seven subjects due to the COVID-19 health crisis in spring 2020.

The results for the secondary objective showed that there is only a subtle effect on aided pure tone thresholds, the Freiburg monosyllable test and the OLSA, especially when considering that changes in MCLs of 30 % are quite substantial compared to the 3-9 % changes in clinical routine. The results might only give information about certain trends due to the fairly small number of subjects. To make more confident statements, more subjects might need to be tested.

Just like for the other tests, the results for the Ling-6 sound test showed that mean or median test scores are not very sensitive to large changes in fitting maps. When looking at the results for each sound and each patient individually, we could see that certain map manipulations lead to characteristic confusions of sounds, which is the main information that should be gained out of performing the test. Again, a larger number of tested subjects will be beneficial to make confident statements about these characteristic confusions. Still, there is room for a few improvements. As a first modification of the test, an optimization of the test items should be considered. For example, the sound /**a**/ was found to give no additional information concerning characteristic confusions and could therefore be eliminated from the list of test items. This statement should not question the use of the sound /**a**/ in the Ling-6 sounds in general but the use for this specific type of test in identification form. Other test items could be included to add more possibilities for characteristic confusions. A suggestion would be to add the German vowels /**e**/ and /**o**/, and the consonant /**f**/. The vowel /**e**/ shows high similarity to /**i**/, while the sound /**o**/ is easily confused with /**u**/. The consonant /**f**/ on the other hand adds an item to possibly be confused with /**s**/ or / **ʃ** /, reducing the probability of guessing within the similar sounding fricatives. It was shown in the evaluation of the synthesized Ling-6 sounds that these new items add potential for more confusions and thus map analyses. Furthermore, they increase the possibilities of frequency analysis for the test. For a future study it might also be an option to divide the fitting maps into four frequency regions to again have more possibilities in the analysis of the confusion matrices.

One more challenge was the addressing of systematic errors potentially caused by the limited number of test stimuli. In the current setup the subject might realize that there are only six sound samples, which might influence the decision-making process of which sound was heard. This might cause the subject to either easily identify all the sounds correctly because of the limited number of answers, or to systematically confusing two or more sounds. To address this problem, slight variations in sound level could be introduced while keeping the spectral information the same to give the illusion that there are more different test items than there actually are.

In conclusion, the investigation of speech tests was found to be very challenging, because there is such a large number of factors involved in the process of understanding speech with a cochlear implant. CI technology has tremendously advanced over the last decades and is now a well-established treatment for patients suffering from severe sensorineural hearing loss or even complete deafness. Still, there is great potential to investigate the field of hearing with a CI and how to objectively assess the performance of CI patients for more precise improvements in their fitting maps.

List of Figures

List of Tables

References

Agung, K. B., Purdy, S. C., & Kitamura, C. (2005). The Ling sound test revisited. *Australian and New Zealand Journal of Audiology, The, 27*(1), 33.

Australian Hearing Manual of Speech Perception. (2001). Australian Hearing.

Boyd, P. J. (2006). Effects of programming threshold and maplaw settings on acoustic thresholds and speech discrimination with the MED-EL COMBI 40+ cochlear implant. *Ear and hearing, 27*(6), 608-618.

Dockum, G. D., & Robinson, D. O. (1975). Warble tone as an audiometric stimulus. *Journal of Speech and Hearing Disorders, 40*(3), 351-356.

Hahlbrock, K. H. (1970). *Sprachaudiometrie: Grundlagen und praktische Anwendung einer Sprachaudiometrie für das deutsche Sprachgebiet*: Thieme Georg Verlag.

Kalcher, M. (2019). Generation of a speech understanding test for cochlear implant patients based on Ling-6 sounds.

Ling, D. (2002). *Speech and the hearing-impaired child: Theory and practice*: Alexander Graham Bell Association for the Deaf and Hard of Hearing.

Mallinger, E. (2011). Trainingseffekte und Listenäquivalenz des Freiburger Einsilbertests im Störschall.

Oldenburger Satztest: Handbuch und Hintergrundwissen. (2000). HörTech gGmbH.

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *The Journal of the acoustical society of America, 24*(2), 175-184.

Recommended procedure: Pure-tone air-conduction and bone-conduction threshold audiometry with and without masking. (2011). British Society of Audiology.

Sainz, M., de la Torre, A., Roldän, C., Ruiz, J. M., & Vargas, J. L. (2003). Analysis of programming maps and its application for balancing multichannel cochlear implants: Análisis de los mapas de programación y su aplicación para el balance de implantes cocleares multicanal. *International journal of audiology, 42*(1), 43-51.

Scollie, S., Glista, D., Tenhaaf, J., Dunn, A., Malandrino, A., Keene, K., & Folkeard, P. (2012). Stimuli and normative data for detection of Ling-6 sounds in hearing level. *American Journal of Audiology*.

Steffens, T. (2006). Test-Retest-Differenz der Regensburger Variante des Oldenburger Kinder-Reimtest (OLKI) im sprachsimulierenden Störgeräusch bei Kindern mit Hörgeräten. *Z Audiol, 45*(3), 88-99.

Steffens, T. (2015). Der Freiburger Sprachtest: ein Klassiker mit Stärken und Schwächen. *HNO Nachrichten, 45*(4), 22-25. doi:10.1007/s00060-015-0370-9

Steffens, T. (2016). Verwendungshäufigkeit der Freiburger Einsilber in der Gegenwartssprache. *HNO, 64*(8), 549-556. doi:10.1007/s00106-016-0163-5

Sukowski, H., Brand, T., Wagener, K. C., & Kollmeier, B. (2008). Untersuchung zur Vergleichbarkeit des Freiburger Sprachtests mit dem Göttinger Satztest und dem Einsilber-Reimtest nach von Wallenberg und Kollmeier. *HNO, 57*(3), 239. doi:10.1007/s00106-008-1727-9

Wagener, K., Kuehnel, V., & Kollmeier, B. (1999). Entwicklung und Evaluation eines Satztests für die deutsche Sprache I: Design des Oldenburger Satztests. *Zeitschrift für Audiologie, 38*, 4-15.

Wilhelm, H., Lachenmayr, B., & Hamann, K.-F. (2015). Sinnesorgane *Medizinische Gutachten* (pp. 417-461): Springer.

Winkler, A., & Holube, I. (2014). Was wissen wir über den Freiburger Sprachtest. *Z Audiol, 53*(4), 146-154.

Erklärung zur Verfassung der Arbeit

Michael Kalcher, BSc.

Hiermit erkläre ich, dass die vorliegende Arbeit gemäß dem Code of Conduct, insbesondere ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel, angefertigt wurde. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet.
Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form in anderen Prüfungsverfahren vorgelegt.

Wien, 01. Juni 2020

____

Michael Kalcher