University of Southampton

Facility of Engineering and Physical Sciences

Web and Internet Sciences

Measuring cochlear implant microphone performance: the accuracy of current methods and implications of inaccuracy.

DOI:

by

Stephen Wetherill

ORCID ID: 0000-0003-0246-8822

Thesis for Doctor of Philosophy

May 2021

University of Southampton

Abstract

Facility of Engineering and Physical Sciences

Web and Internet Sciences

Doctor of Philosophy

Measuring cochlear implant microphone performance: the accuracy of current methods and implications of inaccuracy.

By Stephen Wetherill

Cochlear Implant (CI) microphones are frequently replaced in clinic, but how many of these are actually faulty? How many of these are replaced unnecessarily? How accurate and consistent are current methods used by CI clinics for measuring microphone functionality? This project investigates these questions and also investigates alternative methods for measuring CI microphone performance.

First, the accuracy of current methods for assessing CI microphones are investigated using a survey, measuring microphones that have been reported as broken and testing the accuracy of subjective microphone checks. 57% of the microphones that were reported as broken were within ± 3dB of reference values (i.e. actually working). Furthermore, 90% of clinicians surveyed said they had replaced a microphone that they thought was working. The accuracy of subjective microphone checks was also investigated by presenting 10 microphones that were either "*Working*", "*Partially working*" or "*Not working*". Both the control group (n=10) and CI clinicians (n=4) were both able to identify the "*Not working*" microphones with 100% accuracy. Both groups were significantly less accurate when differentiating the other two groups of microphones.

Previous testing showed that 17% of times that "*Partially working*" microphones were presented to clinicians they were classified as working. Impact of microphones with undetected defects on speech perception is analysed. Partial microphone failures were simulated during a speech perception test. This may mean that partial failures go unreported, and that users are struggling to hear, ascribing it to the situation rather than the microphone.

Objective tests on a CI saw which of the user customizable settings affect the output of the Listening Check adaptor that was used extensively during the experiments. While the current CI microphone selection affects the adaptors' output, none of the other settings were shown to have a significant effect on the output. The consistency CI frequency responses could be recorded was also assessed and over 56 recordings there was a maximum deviation of ± 2dB and an average deviation of ± 0.8dB.

We need to develop better software based systems to be integrated into clinical processes around the testing of microphones both to reduce waste, avoid partially working microphones being overlooked, and to increase patient and clinician confidence in their equipment. These testing processes could be embedded into home monitoring systems to enable more regular and effective testing of implant microphones.

Contents

List of Tables

List of Figures

Acronyms

AIS Auditory Implant Service

BGN Background Noise

BKB Bamford-Kowal-Bench

CI Cochlear Implant

DTT Digit Triplet Test

HA Hearing Aid

HINT Hearing in Noise Test

NHS National Health Service

PTA Pure Tone Audiometry

PTS Pure Tone Sweep

RCT Randomised Controlled Trial

SNR Signal to Noise Ratio

SPL Sound Pressure Level

Academic Thesis: Declaration of Authorship

Print name:	Stephen Wetherill

Title of thesis:	Measuring cochlear implant microphone performance: the accuracy of current methods and implications of inaccuracy.

I declare that this thesis and the work presented in it are my own and has been generated by me as a result of my own original research.

I confirm that:

1. This work was done wholly while in candidature for a research degree at this University;
2. No part of this thesis has been submitted for a qualification an another University;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Parts of this work have been submitted for publication as:

Wetherill, S., Weal, M.J. and Cullington. H. (2021) Is that microphone really faulty? Investigating microphone check accuracy to improve process efficiency in cochlear implant microphone replacement. Cochlear Implant International. Under Review.

Signature:		Date:	

Acknowledgements

Thank you to my supervisors Mark Weal and Helen Cullington who have given of their time and expertise guiding this project. They have both given me the space to get on and work when I need it but also supported me when needed.

David Simpton and David Holland from the University of Southampton Institute of Sound and Vibration Research have both provided very useful insight into my project from an acoustics perspective. This help has been invaluable.

Also a thank you to my dance partner (Emma Nottingham) for your advice, support, encouragement and de-stressing throughout this project.

During this project, my family (especially parents and older sister) have provided support in many forms including emotional and proofreading hundreds of pages between the different versions of this document; this is greatly appreciated and has been instrumental in getting my thesis to this stage.

Chapter 1

Introduction

An estimated 466 million people including 34 million children worldwide have hearing loss (World Health Organisation 2018) with an estimated 11 million people with hearing loss in the United Kingdom[1]. Furthermore, there is an estimated $600,000$ patients worldwide with Cochlear Implants (CIs) with approximately $45,000$ CI processors sold each year[2] and 1553 CIs implanted in the United Kingdom between 1st April 2018 and 31st March 2019 (Hanvey 2020).

1.1 The problem

People who use CI have their processor(s) on the side of their head for nearly every waking hour means that their microphone's get considerable wear and tear in addition to being exposed to wind, rain, dust, debris and more. This often results in the microphone's having reduced performance or breaking entirely. For many people who use CIs, sudden drops in microphone performance will be noticed with relative ease. Devices like the Advanced Bionics Listening Check or Cochlear Kanso 2 enable other people such as parents or carers to subjectively evaluate the microphone. Gradual reductions in performance can be harder to detect. This is because a microphone's performance may only go down by such a small amount as to be unnoticeable. When these changes happen day after day they mount up to a significant reduction in microphone performance. Currently microphone failures are detected one of two ways; firstly, they are noticed by the person who uses the CI or indirectly by a parent or carer using the listening check. This is then reported to the Auditory Implant Service (AIS) who then arrange for a replacement microphone. Secondly, when a person who uses a CI attends a check up appointment at their CI centre and has hearing and speech perception testing. If a test has below expected results then the CI centre will try to locate the problem, replacing the microphone if they believe it is necessary. Further-

[1] www.actiononhearingloss.org.uk facts and figures page
[2] www.earfoundation.org.uk CI information sheet

more, if a working microphone is replaced at clinic due to poor speech perception test results, this could delay diagnosis of the real problem.

While the majority of CIs has some method of listening to the microphone, these all rely on people doing subjective microphone checks and detecting reductions in microphone's performance. While some people who use CI, or their parents/carers, will likely notice failures, they are less likely to notice gradual reductions in performance. While one gradual reduction is unlikely to have a significant effect, if these gradual reductions mount up over time and the cumulative effect will be significant. This can result in people who use CI having a microphone with reduced functionality for a prolonged period of time until their next appointment which could be a year between, or longer with COVID and the increased uptake in telemedicine systems. At check-up appointments doing a speech in noise test is common: this will test the entire CI signal pathway all the way from when the signal enters the CI to the brain's understanding of a sound. There are ways of measuring the functionality of different parts such as connecting a processor to a computer and running diagnostics or measuring the impedance of the electrodes in the array. However, there is no current way of objectively measuring CI microphone performance in isolation. Currently microphones that are reported as broken by either AIS staff or people who use CIs are sent to the manufacturer if they are less than a year old; otherwise they are thrown away and replaced. The main problem is that we do not know the accuracy of the current methods for identifying broken CI microphones.

1.2 Research questions

RQ1 How accurate and repeatable are the current methods for identifying CI microphone failures?

RQ2 What impact do partially failed CI microphones have on speech perception?

RQ3 What processor settings affect the audio output of the Listening Check?

1.3 Hearing system and Cochlear Implants

The hearing system is a transducer that converts sound waves into electrical impulses that are sent along a nerve and interpreted by the brain as sound. When someone's hearing system gets damaged there are a number of options to help including Hearing Aids (HAs); however, these systems try to compensate for the damaged hearing system so have their limits. A CI by contrast works by skipping the outer and middle ear

(from: www.nidcd.nih.gov)

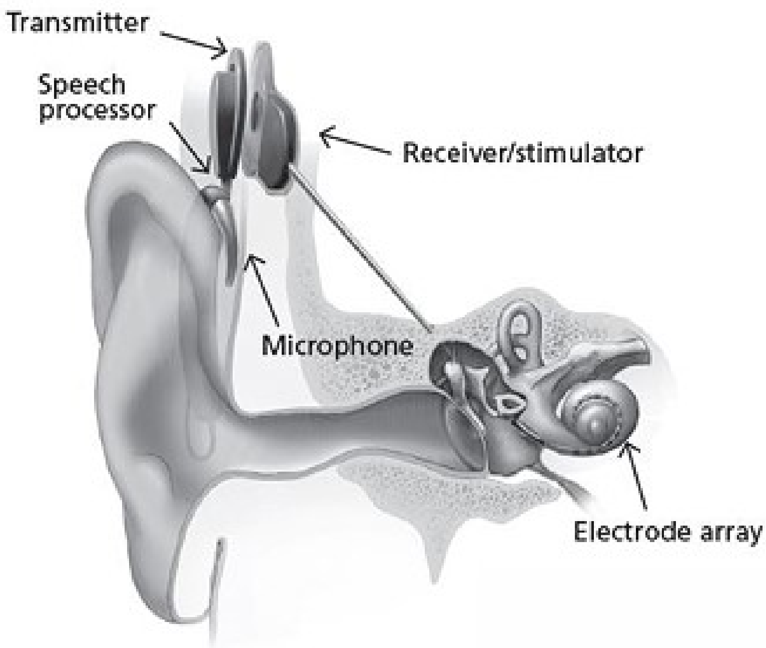


Figure 1.1: Diagram showing the internal and external part of a cochlear implant

entirely and uses an implanted electrode array to directly stimulate the cochlear nerve, in essence replacing the natural transducer. A CI is made up of two parts: the sound processor and the internal implant as shown by Figure 1.1. The sound processor contains the battery, microphones and hardware required for the audio processing. The sound processor takes the signal from the integrated microphones and processes the sound. The internal part of a CI located behind the ear has a cable that extends down through the temporal bone to an electrode array that is inserted into the cochlear. When triggered the electrode array releases an electrical charge that stimulates the nerve cells.

CIs have a number of user specific settings such as microphone selection and directionality; the exact settings that can be customised and how many setting states can be saved varies between CI manufacturer. For example, Naída processors have five programmes that change by a button on the top of the processor. For each programme the user may be able to choose between three microphones, which are shown in Figure 1.2. The T-Mic, the headpiece microphone (both omnidirectional) or the processor microphones which are combined to give directionality options.

Different CI manufacturers have a number of different options to check microphone functionality. For example, Advanced Bionics have a device called the Listening Check™ that connects to a processor and provides a 3.5mm headphone output. This can then be used to subjectively check if a microphone is working. Figure 1.3 shows pictures of an

(from: www.cochlearimplanthelp.com)

Figure 1.2: Picture showing the location of microphones on the Advanced Bionics Naída Q70 CI processor

Advanced Bionics Naída Q70 processor and Listening Check in various configurations. Figure 1.3a shows how it will be generally used by people who use CI with a processor located behind the ear. The coil is not shown in any of these pictures which is connected via a cable to the back top of the CI processor (top right of picture). Figure 1.3b shows how a listening check device can be connected between the CI processor and its battery, this provides a 3.5mm headphone output. Figure 1.3 shows an exploded view of the components: this also shows how the T-Mics can be removed from the CI processors and are available in varying lengths to fit different peoples ears.

Advanced Bionics are not the only CI manufacture, the Kanso© 2 CI sound processor by Cochlear© has a "*Sound Check*" system[3]. Optionally the recordings can then be shared with someone else or a CI centre which could be useful diagnostically but, like the Advanced Bionics Listening Check relies of people subjectively evaluating if a microphone is working.

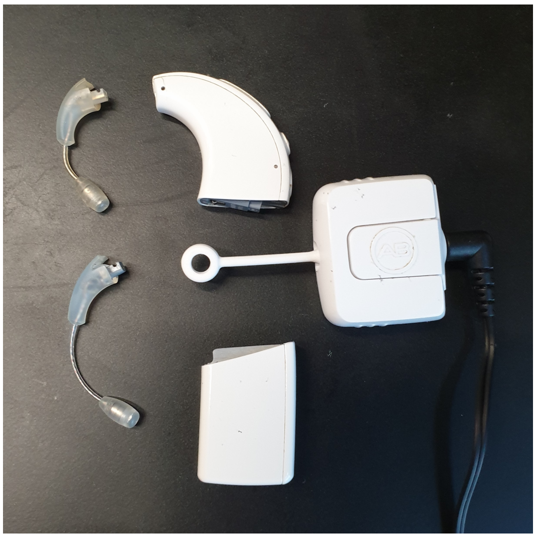
1.4 Existing solutions

As previously discussed, both the Advanced Bionics Listening Check and Cochlear Kanso 2 enable the CIs microphones to be listened to and this can be done by a parent, family member, CI clinic staff or someone else. However, these devices just enable people to listen to the microphones and do not provide any way for microphone performance to be objectively evaluated. The accuracy and repeatability of subjective microphone checks has not previously been investigated, to our knowledge. However,

[3] https://advancedbionics.com/content/dam/advancedbionics/Documents/Global/en_ce/Products/Naida/Naida-CI-System-Check.pdf

(a) Normal use

(b) With listening check



(c) Exploded view

Figure 1.3: Pictures of Advanced Bionics Naída Q70 CI processor with and without the Listening Check attached.

despite the specialist equipment and sound-treated room required for pure tone audiometry, results are only accurate to within ±10dB HL (Frank & Dreisbach 1991, Schmuziger et al. 2004). It is therefore doubtful that subjective microphone checks will detect changes in microphone functionality unless they exceed ±10dB HL. The threshold where a reduction in microphone performance will consistently be detected is likely higher than ±10dB HL as the background noise during a subjective microphone check is likely higher than in a sound-treated room.

In CI clinics speech in noise tests are common place and while there are a multitude of different speech in noise tests such as the HINT test (Molander et al. 2013) and the Bamford-Kowal-Bench (BKB) sentence test (Bench et al. 1979); the most relevant test to this project is the Digit Triplet Test (DTT). The DTT was originally developed in the Netherlands by Smits et al. (2004) as a method for doing a screening hearing test over the telephone; this was then translated into English by Lutman et al. (2006). The tests consists of three numbers between zero and nine[4] being played in quick succession with a varying level of Background Noise (BGN). Subsequently, Cullington & Aidi (2017) investigated if the DTT can be used as part of a remote care system for people who use CIs. 16 people who use CIs took part who did both the DTT and the BKB sentence test. The paper concluded that the DTT correlated with the results from the BKB sentence test showing that it could be used in the long term as a speech in noise test and could then be used as part of a healthcare package. While the DTT is currently being rolled out as part of a telemedicine system by the AIS this measures the hearing of the entire CI signal pathway from microphone to brain interpreting the sound and offers no way of localising a problem. This means that if a speech in noise test has a sub-average result it could be from a wide array of different sources including the CIs microphone, CI processor, coil or electrode array.

1.5 Potential Challenges

CIs also contain a considerable amount of proprietary equipment and software which will make measuring microphone performance harder. For example, while using an Advanced Bionics Listening Check, it is known that the current microphone choice affects the Listening Checks output, but the effect any other user specific settings have is not known. This means that if a processor is set to use the T-Mic microphone then this is what the Listening Check will hear, but if settings such as BGN reduction systems and map levels affect the Listening Check is not known. This becomes especially problematic if aiming to do objective tests through the Listening Check as these other settings could significantly affect results if they affect the Listening Check.

Telemedicine systems are becoming more frequently used for patients both in Audiology and in the wider hospital (See Sections 2.5 and 2.6 for more information). However, the vast majority of these telemedicine systems require access to the internet. A survey of 158 people who use CI found that 60% had access to the internet. Of this 60%, 40% said they may have problems accessing the internet in the future if it requires the use of a computer (Thorén et al. 2013). Therefore, when designing telemedicine systems if it can be used by people who struggle to access the internet should be considered.

The dexterity of people who use CI should also be considered as the microphones are small and many users are unable to change their T-Mic them selves. When aiming to build a device that can be used to objectively test the functionality of CI processor ,

[4] excluding seven as this is two syllables

some aspects of it may require precision such as sliding a CI processor onto a mount and some people who use CI may not have the dexterity for this. Visual acuity should also be considered as the a number of components involved in checking if CI microphones work are small; for example the pin that holds Advanced Bionics T-Mics in is 6mm by 0.7mm which could be hard to see with even a minor visual impairment.

1.6 Report structure

Below is a list of the chapters of this thesis along with a brief description of the subject of each subsequent chapter.

2. **Background**- Discusses literature, areas of research and commercially available products that have relevance to this project.

3. **Methodology**- The subsequent chapters contain a number of discrete experiments. This chapter explains various principles and techniques that are used in a number of the experiments. After this, each experiment is described, also discussing the purpose of each experiment.

4. **The problem with current testing**- This chapter describes a number of experiments investigating how CI microphones are tested, how many microphones are tested, when and why microphones are replaced among other factors. Then microphones that have been reported as broken are tested to investigate the prevalence of false negatives[5]. Also the prevalence of different types of microphone failures are investigated, what frequencies get reduced by what amounts, how many are completely not working.

5. **Implications of failure**- This chapter investigates the implications of the reduced microphone performance on speech recognition. Reductions in microphone performance were simulated during multiple speech tests to determine their effect on speech recognition.

6. **Reference testing**- The previous experiments were all done with the CIs in test settings with all advanced features disabled. While having the CI processor in test settings is good for experimental repeatability, it is not a realistic option for people who use CIs. This chapter conducts a number of experiments to see how CIs processor settings would have affected the results.

7. **Discussion**- This chapter discusses a number of points about the project including: differences between CI manufacturers and microphone type in addition to how future CI telemedicine systems could be introduced.

[5]Fully working microphones marked to be replaced despite being fully working

8. **Conclusion**- This chapter summarises, draws together all the previous chapters and talks about the contributions that this project has made to the subject area.

Chapter 2

Background

The human ear functions as a transducer, converting the analogue sound waves that go into the ear canal into electrical impulses that the brain can understand. Figure 2.1 shows an overview of the human hearing system which can be divided into three chapters. First: the outer ear is made up of two parts, the auricle[1] and the auditory canal. Both the auricle and auditory canal are designed to focus sound towards the ear drum which is between the outer and middle ear. Second: the middle ear includes the ear drum and the ossicles which are in the tympanic cavity. When a sound wave reaches the ear drum, the sound wave causes the ear drum to vibrate, which in turn vibrates the ossicles. Thirdly, the inner ear is comprised of the cochlea and the semi-circular canals. The ossicles transmit sound from the ear drum to the cochlea where it is converted into pressure waves through a liquid. Inside the cochlea there is a series of hair cells that pick up the vibrations in the fluid then convert the changes in pressure into an electrical signal. This electrical signal is sent along the cochlea nerve to the brain where it is interpreted as sound.

The hair cells in the cochlear that are responsible for turning the acoustic sound signal into an electrical signal can die for a multitude of reasons including age and loud noise exposure. These cells dying results in a less sensitive hearing system meaning that people can no longer hear as quiet sounds but the loudest sounds people can hear may remain the unchanged. This results in a significantly reduced usable dynamic range.

2.1 Hearing aids

A common way of improving the quality of life for people who have a hearing loss is though Hearing Aids (HAs) which amplify the quieter sounds so they can be heard. How much sounds are amplified by is customised for each individual according to their

[1] A.K.A. Pinna

(From: www.classichearing.com)

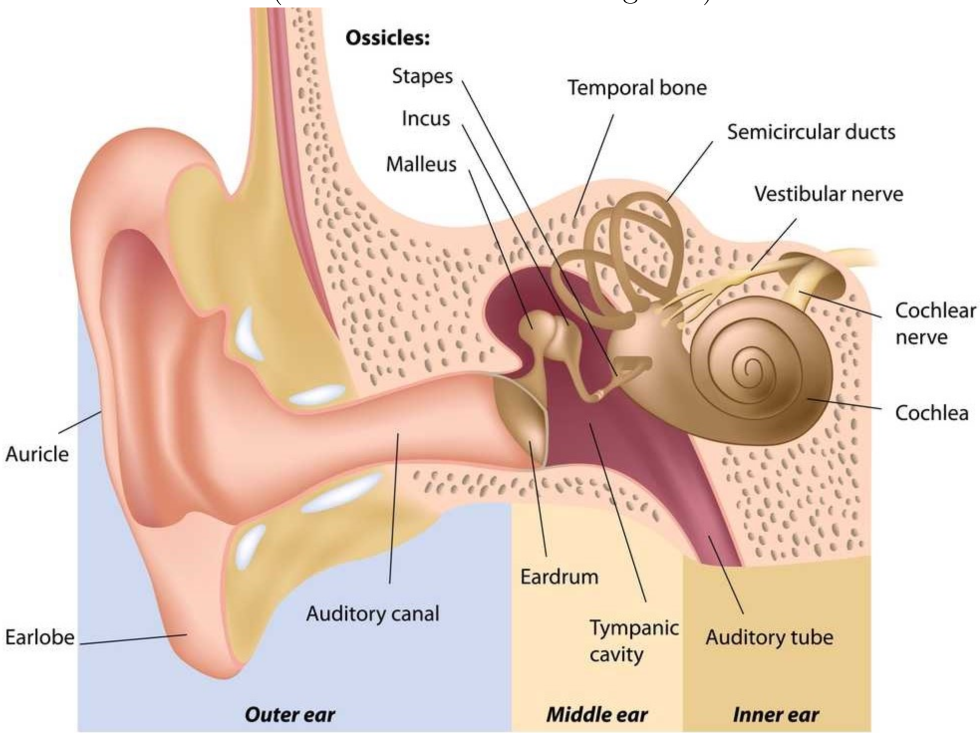


Figure 2.1: Labelled diagram showing the anatomy of the ear

frequency specific hearing loss. Modern HAs utilise a series of advanced features that are designed to improve the listening experience for the HA users. The exact details of how each feature works is not known as they are trade secrets by the various HA companies that use them. However, the underlying principles behind a few of the common HA advanced features are briefly explained.

As previously mentioned, when someone has a hearing loss they will be unable to hear quieter sounds but the loudest sound they can tolerate may remain unchanged. If loud sounds (such as a door slamming) were amplified the same as the quieter sounds this could result in discomfort for the person using the HA. Therefore, modern HAs do not amplify all sound equally but use compressors to reduce how much louder sounds are amplified by to make a more comfortable listening experience.

HAs contain a number of microphones that are more sensitive to sound in-front than behind and other microphones that are equally sensitive to all directions. Adaptive directionality attempts to determine where a persons auditory attention wants to be and adjusts the microphones accordingly. For example, in a noisy environment like a restaurant the adaptive directionality system assumes that the person wants to listen to what is in-front of them so the forward directional microphone is selected. Exactly how a specific direction is prioritised over another varies for each HA manufacturer and the exact details are not known.

A common problem for people who use HAs is Background Noise (BGN) and there is an advanced feature that is designed to help with this (BGN reduction systems). BGN has a significantly lower dynamic range than speech so this advanced feature splits the input signal into a number of frequency bins each containing a part of the frequency spectrum then works out the dynamic range of each frequency bin. The assumption is then made that the frequency bins with lower dynamic range have more noise than speech signal so have the frequency bin specific gain reduced. The opposite is also done so frequency bins with higher dynamic ranges are turned up assuming they have less background noise. The exact method and number of bins used to run this advanced feature are not known.

2.1.1 Test box

HA test boxes are designed to accurately and with repeatability measure the performance of HAs and are common place in audiology clinics. The majority of audiology centres have test boxes[2] and all conform to industry standards (International Electrotechnical Commission 2005, 2012). These devices have a small chamber in them with a speaker built in; also inside the enclosure is a microphone on a small flexible mount and a connector for attaching a HA. A number of the tests that the test boxes are commonly used for require sound be to played at a specific level. However, HAs come in all shapes and sizes so getting an accurate sound level at the different HAs microphones can present a problem. The microphone inside the chamber is designed to combat this issue, the microphone is calibrated and the flexible mount enables this reference microphone to be placed as close to the HA microphone as possible. The test box will then adjust the volume of the speaker to the required Sound Pressure Level (SPL) for each test. This reference microphone is also used to compensate for the enclosure itself altering the presented sound[3], as using an enclosure that did not impact the presented sound at all would be prohibitively expensive and large. Most chambers use a standard 3.5mm headphone connector inside the chamber and each HA manufacture makes adaptors that enable their HAs to be connected. The capabilities of the adaptors varies between manufactures but the vast majority of them do not alter any of the HA settings, meaning the HAs need to be put manually into test settings before testing.

There are a number of different tests that can be done with HAs test boxes but two of the more common tests are frequency response and harmonic distortion. **Frequency**

[2]such as those by Interacoustics and Oticon
[3]such as boosting a specific frequency

response plays a 60dB SPL Pure Tone Sweep (PTS) from 20Hz to 20kHz. This is to assess the full frequency range that the device is capable of reproducing with the devices gain set to a known reference level. **Harmonic distortion** plays 500, 800 and 1600Hz pure tones at 70dB SPL[4] and the output from the test device is analysed. A pure tone stimulus will produce small acoustic artefacts at the harmonics of the stimuli frequency. For example, a HA presented with a 500Hz pure tone signal will produce a small acoustic artefact at 1kHz. The total power output from each of the harmonic frequencies is summed and then presented as a percentage of the stimulus strength and should be below 3%.

A few of the advanced features that are explained in the previous section can interfere with some of the tests. A very common advanced feature in HAs is feedback reduction. As HA feedback in HAs is often a pure tone, when a HA hears a pure tone is introduces a second pure tone of the same frequency but the opposite polarity to cancel out the pure tone. If this advanced feature is left on during a frequency responses which uses a pure tone stimuli, the feedback reduction system can try and cancel the stimuli resulting in anomalous results. Other advanced features can also impact the accuracy of test box measurements. Furthermore, a number of tests require specific gain settings on the HA when testing, such as full on gain and reference test gain. To alleviate this issue, the majority of HAs have a "*test mode*" that they are put in when using a test box: this sets their gain to reference levels and disables all advanced features.

While this shows the array of tests that can be conducted on HAs in clinic however, none of these tests can be done to Cochlear Implants (CIs) to access how well their microphones are working. As discussed in later chapters, there are ways of conducting objectively measuring CI microphone performance in a research environment however, these are not able to be done in CI clinics.

2.2 Cochlear implants

When a hearing loss progresses to the point where HAs can no longer provide sufficient amplification[5] CIs are a good option. The inner ear acts like a transducer by converting an acoustic signal into a series of electrical impulses that the brain then interprets as sound. A CI works by replacing the ears natural transducer with an artificial one. Figure 2.2 shows a diagram of how a CI fits in a human ear. Figure 2.2 item 1 is the CI processor that contains the microphones, batteries and signal processing capabilities.

[4]Except 65dB SPL at 1.6kHz

[5]Hearing could progress for a variety of reasons such as presbycusis, otosclerosis or noise induced hearing loss

A cable from the CI processor extends to Figure 2.2 item 2 which is a called the coil: this is held in place by magnets and sends data to the internal part of the CI using FM radio. The internal part of the CI (Figure 2.2 item 3) has two cables coming off it; one is a grounding cable the other has an electrode array and is inserted into the cochlea.

(from: www.cochlear.com)

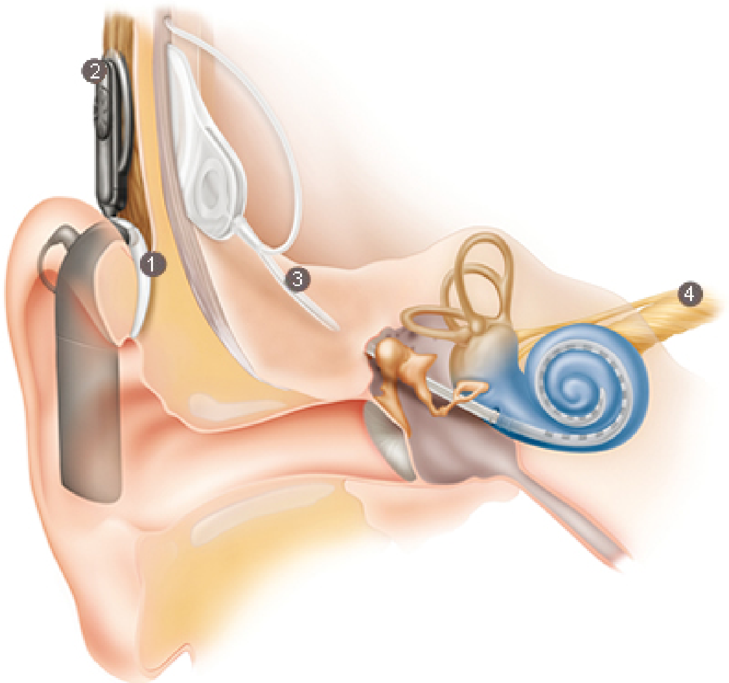


Figure 2.2: Diagram showing the internal and external parts of a cochlear implant processor

1. Cochlear Implant processor
2. Coil
3. Cable connecting the internal part of the cochlear implant processor to the coil in the cochlear
4. Auditory nerve

The inside of a cochlea is filled with a fluid called perilymph with an opening at one end that is coiled approximately 2.5 times. The inner wall of the cochlear is lined with cells that turn the pressure waves that travel down the cochlear into electrical signals. These electrical signals are then sent along the Auditory Nerve (Figure 2.2 item 4) to the brain. The hair cells nearer to the end of the cochlea where the sound enters are all tuned to specifically pick up high frequency sounds and the further through the cochlea a sound gets the lower frequency the cells are tuned to. The electrode array works by releasing a small electrical charge into the cochlea that travels through the perilymph into the cells surrounding it. This means that a single electrode instead of stimulating a specific frequency's nerves in fact stimulates a small range of frequency specific nerves.

As the electrodes array is sending a small electrical current into liquid to stimulate nearby cells a single electrode actually stimulates a small range of frequency sensitive cells. This means that a single electrode could be presenting a single pure tone but it would be perceived as a range of frequencies being presented [6]. This reason is why the number of electrodes that a CIs has been shown to have limited impact on the speech perception for people with CIs (Fishman et al. 1997, Friesen et al. 2001). The electrode array makes it between 1.5 and 2.35 turns into the cochlear so it does not always reach the low frequency nerves, but there does not need to be a direct relationship between a stimulus frequency and the nerves that an electrode is stimulating. This means that while people with CIs may be able to hear frequencies between 100Hz at 10kHz, the CI processor could be sending this frequency information to the brain over the nerves that would normally be used for the frequency range 1kHz to 20kHz for example.

The amplitude of the electrical signal that the electrode array uses to stimulate the cochlear vary depending on a number of factors including the exact electrode positioning and how conductive someone's cochlear is. The amplitude of these signals is altered by clinicians using a "*map*" which sets both the minimum amplitude required for sound to be heard and the maximum tolerable volume for each electrode.

Section 2.1 explains two advanced features that are common to have on HAs, and these features can also be found on CIs. There are a number of other advanced features that are used by either HAs or CIs but explaining every advanced feature that a HA or CI could have is beyond the scope of this project.

To provide additional auditory aid to people who use CIs there are a number of devices that are designed to work in tandem with the CI. One such device is the Phonak Roger pen[7] which connects to a CI processor wirelessly and functions as a wireless microphone. This wireless microphone can then be placed closer to the desired sound source. One such example would be if a person who uses a CI is in a coffee shop with a friend; if using the microphones integrated into the CI processor then it might be still hard to hear the person you are meeting with. However, if the friend puts the Roger pen in their top pocket then this will provide a microphone that the person with the CI can listen to that is significantly closer to the sound source[8] improving the listening experience. Another common use case for this device is in classrooms when a student has a CI because a Roger pen in the teachers top pocket would give significantly better sound signal than the child's CI processor.

[6] like narrow band noise
[7] `https://www.connevans.co.uk/product/6002943/3PRPEN/Phonak-Roger-Pen-transmitter`
[8] than their CI processor microphones

The Advanced Bionics Listening Check[9] is a CI accessory that is used extensively throughout this project. The device connects directly to a CI processor and provides a 3.5mm output for the microphone in the current CI programme. This is designed so that someone with normal hearing close to the person who uses the CI[10] can connect a pair of headphones and listen directly to the microphone to check that it is working. For example, a parent of a person who uses a CI plugs in the Advanced Bionics Listening Check before the child wakes up, as the child's first programme uses the processor microphone exclusively the parent checks this microphone then changes the CI processor to the second programme which uses the T-Mic exclusively. The parent could then speak into the microphone to check that it is working. If the microphone is suspected of being defective then each CI centre will have a way of reporting this and replacing the microphone if necessary.

2.3 Clinical pathways

There are guidelines on patient care pathways for people with hearing loss: however, these focus on HAs because these are more prevalent than CIs (National Institute for Health and Care Excellence 2019, 2020). These documents are only guidance and the exact clinical pathway that a CI candidate takes will vary between countries, but there are a number of common elements which are discussed in the bullet points below.

- Firstly, in order to get a CI there will be an assessment appointment to see if the person meets the audiological criteria for having a CI (which varies between countries).
- Next there is a surgery to implant the internal part of the CI.
- After surgery there is a wait of approximately three weeks to enable the internal part of the CI to settle.
- The external part of the CI is then connected to the internal part and switched on for the first time.
- In the weeks and months following the CI first being turned on there is normally quite a number of appointments to adjust the CI settings as the persons' body gradually gets used to the CI.
- Once someone has acclimatised to their CI, they start off having annual follow ups appointments with their implant centre, reducing over time to one every couple

[9] https://advancedbionics.com/us/en/home/products/accessories/listening-check.html
[10] such as parent, partner or carers

of years or on an as needed basis. During this time people with CIs are given the means of requesting follow-up appointments if they believe it is necessary for any reason.

If the microphone suddenly stops working then in the majority of cases this will be picked up by the person who uses a CI who will then contact the implant centre. Gradual changes are harder to detect and in clinics can be picked up by a sound field audiogram: but this is not a reliable way of measuring the microphones' effectiveness. This is not ideal because you are needing to assess the microphones' functionality at the same time as a people with CIs hearing, making determining the cause of a worse result harder. When a CI T-Mic needs replacing what happens next will vary depending on the implant centre, but the Auditory Implant Service (AIS) in Southampton will send a replacement microphone with a free return envelope for the old microphone. This protocol will be different if the person who uses a CI is not capable of replacing a T-Mic at home: then the processor will either need to be brought or posted to the department where a member of staff could replace the microphone.

Telemedicine systems are being used increasingly frequently by people who use CIs which is discussed in more detail in Section 2.5. This increased uptake of telemedicine systems is resulting in follow up appointments having longer gaps between them. This means that CI centres have fewer opportunities to check how well a CI is working.

2.4 Common clinical tests

Pure Tone Audiometry (PTA) is a common clinical test used for screening and diagnosing of hearing loss (Gelfand 2009). This test consists of playing a series of quiet pure tone beeps with the person who is doing the test pressing a button when they hear a beep. The sound level that the beeps are played back is altered to find the quietest sound that someone can hear at a specific frequency; the test is then repeated on other frequencies. While people with CIs cannot do this test, PTA forms part of the CI assessment criteria that determines if someone is eligible to be implanted or not (National Institute for Health and Care Excellence 2019).

The Bamford-Kowal-Bench (BKB) speech test is a commonly used in CI clinics and plays a specific number of sentences to someone with a background noise present. Each sentence played has five keywords in it and the test is scored on the best Signal to Noise Ratio (SNR) that someone can repeat the required number of keywords correctly (Boothroyd 1968). This is a widely used clinical test that has been researched extensively even showing minimal differences to be classified as significant (Martin

1997, pg. 167).

The Digit Triplet Test (DTT) test was developed by Smits et al. (2004) as a method for screening hearing loss over the phone; this was later translated to English (Lutman et al. 2006). The test plays three numbers between one and nine[11] while a masking noise is being played. During the test the person will type the numbers they hear and the SNR is continually altered until the level is found where the person taking the test gets words correct 50% of the time: this is their threshold. This threshold can then be compared to reference data to get an idea of the state of their hearing. In the case of nationwide hearing screening programmes this would be reference data for normal hearing people. With people who use CIs, the results could be compared to other people who use CIs or to that persons' previous results.

The map of the CI electrode array requires regular adjusting especially when someone's implant is first turned on: in the vast majority of cases this requires a visit to the CI centre where a clinician can adjust the map settings. However, remotely altering the map has been possible for a number of years (Wesarg et al. 2010). In order to remotely alter the CIs map there needs to be a way of adjusting this over the internet and this presents a number of potential security concerns with a variety of potential consequences including: making the device inoperable resulting in deafness or the map could be turned up so far that is causes the user tinnitus or pain (Bodmer & Capkun 2010). It would help mitigate these risks if during remote CI programming appointments someone was physically present with the person who uses the CI so they can help in the event of a technical problems.

Each of these tests have their uses but none of them assess the CIs microphone in isolation. Sound field speech in noise tests such as the BKB do assess the entire CI signal pathway from sound entering the microphones all the way to the brain interpreting the sound. However, because this assesses the entire signal pathway if offers no way of localising any detected problems. For example, if a speech in noise test shows a lower than expected result for someone who uses a CI then this problem could be caused by: a blocked microphone, electrical problem in the CI, a problem with the connection between the CI processor and the internal part of the CI, the electrode array or something else. Being able to localise a problem is essential to knowing how to solve it in a quick and timely manor.

[11] excluding seven as this has two syllables

2.5 Telemedicine for cochlear implants and audiology

The number of audiologists is not expected to meet the rising demand and with an ageing population which will become a greater problem (Hall 2016, Nemes 2002, Windmill & Freeman 2013, Roth et al. 2011). Telemedicine systems have the potential to alleviate some of the growing pressure on CI centres. This would be accomplished by people with CIs using remote care pathways instead of having face to face appointments with clinical AIS staff. This is in line with the National Health Services (NHSs) long term goals[12] (Istepanian 1999, NHS 2019).

There are a number of telemedicine systems that have been investigated and are currently being used for both CIs and HAs. One potential way to alleviate some of this pressure that has been investigated is remotely measuring hearing loss. Hearing loss can be measured in a clinic using specialist equipment operated by trained personnel doing tests such as PTA. PTA relies on the fact that as hearing loss progresses, the quietest sound someone can hear gets louder. PTA presents a series of quiet tones to the person doing the test and they tell the person doing the test when they hear something; the quietest sound someone hears is then used to determine how good someone's hearing is. If this test could do be done accurately remotely, this would significantly help audiology clinics.

Before a test such as PTA can be conducted outside of audiology clinics two problems need to be overcome, controlling the test and the specialist hardware needed. An automated PTA system was trialled on 120 subjects (Garrison & Bochner 2017). This automated system still required specialist hardware but instead of being controlled by a trained audiologist, the test was controlled by a computer and instructions given using a screen. All the participants also did conventional PTA to compare the results to. The results from the automated system were within tolerance of the reference PTA results performed by audiologists 91% of the time. While the automated system still required specialist hardware, this study does show that as long as the hardware is accurate enough hearing tests can be automated and obtain accurate results the majority of the time. The second problem is the specialist hardware required to do the test. A number of studies have investigated if different devices could be used. One such example of this is using an iPad, calibrating its output then using a dedicated application to see if this could conduct accurate PTA (Corry et al. 2017). This set up

[12] www.airedale-trust.nhs.uk

was tested on 26 participants who each did two reference tests[13] and with the iPad[14], the results were then compared. There were significant differences found between the reference results and the iPad results at all tested frequencies except 500Hz. However, the iPad results were consistent with themselves; for example at 1kHz the iPad was consistently 5dB better. A PTA relies on presenting a series of tones at specific dB HL levels (such as 1kHz pure tone at 10dB HL) and the standard way of calibrating an output is to play a pure tone at 60dB SPL[15]. While this way of calibrating will ensure that when the iPad application says it is outputting 60dB SPL, it is outputting that but it tells you nothing about the loudness growth of the device. This only tells you about outputs of 60dB SPL and no other levels, meaning that when the application says it is outputting 30dB HL it could be outputting 27dB HL. As the participants were normal hearing this means that the stimuli would have been presented at less than 20dB HL. This would explain why the results for the application were consistent with themselves and not with the reference testing; this would also be trivial to compensate for by doing level specific calibration. A product is available that uses an iPad with a specialist piece pair of headphones and an application to conduct PTA (ShowBox Audiometry[16]). Two independent studies have found this setup to be with 10dB of PTA which is the industry standard for between test repeatability (Saliba et al. 2017, Thompson et al. 2015). This shows that there are most cost effective methods to get the specialist hardware needed to conduct PTA. The aforementioned studies have shown that PTA can be automated and still get accurate results (Garrison & Bochner 2017) and iPads with calibrated outputs can also be used instead of the specialist hardware required to run PTA (Saliba et al. 2017, Thompson et al. 2015). These two aspects were combined in a study that tested an application that did PTA without the need to be controlled by trained audiologists (Abu-Ghanem et al. 2016). The results were found to be inaccurate at all tested frequencies for the 26 participants. There is still some work to do in combining the application based audiometers and the automated running of the tests.

In order for PTA to work the tones need to be presented at the correct intensities. However, there are a number of other speech tests that can measure hearing loss without needing to be presented at such exact levels. While a number of factors will change between the tests including stimuli and type of masker, all the variations of speech present a speech stimulus with a masking noise[17]; The exact ratio of the speech signal to the background noise (or SNR) is varied during the test to find the SNR where the

[13] Using conventional equipment
[14] during tests tests the iPad was being operated by a trained audiologist
[15] Other levels are also used including 65dB SPL
[16] www.shoebox.md
[17] Could be multi-talker babble or white noise for example

person taking the test gets a specific amount correct[18], the results are then compared to reference data. One such speech test is the Hearing in Noise Test (HINT), this presents a series of ten words presented with a multi talker babble. This was originally tried in Stockholm Central Train Station on 223 participants who also did PTA (Molander et al. 2013). The results from HINT and PTA were compared and they found that when the SNR was > -3.4dB there was a 79% chance of classifying the users hearing correctly. An alternative speech in noise test is the DTT, developed in the Netherlands as a potential method for doing hearing screening over the phone (Smits et al. 2004); this test was then later translated into English Lutman et al. (2006). The DTT plays three numbers for a participant[19] with a multi-talker babble. A study was later conducted investigating if the DTT could be used as part of a CI remote care system (Cullington & Aidi 2017). It is common practice for people who have CIs to do speech test when visiting a CI centre to assess how well their CI is working. Being able to do this remotely would enable CI clinicians to get an accurate idea of how well a persons CI is working. 16 people who use CIs[20] took part who did both the DTT and the BKB sentence test. The paper concluded that the DTT correlated with the results from the BKB sentence test. Furthermore, results from the questionnaire were mostly positive with 15 of the 16 participants agreeing or strongly agreeing that being able to do the DTT remotely would save them time and money. A subsequent paper also found that people with CIs also reported that being able to do the DTT remotely made them feel more empowered to manage their own CI (Cullington et al. 2018).

2020 has seen a significant increase in the using and need for remote care systems as in person appointments now have more risks associated with them because of COVID. Recent studies have shown that remote care for people who use CIs is possible on a wider scale but requires both the staff and the patients to acclimatize to it (Maurrasse et al. 2020, Shapiro et al. 2020). Remote programming of CIs is also happening with encouraging results thus far (Luryi et al. 2020). A recent systematic review of 50 papers on the telemedicine for CIs found only five[21] to be of sufficient quality (Buckman & Fitzharris 2020). Of the total 70 participants between the studies, 69 had no significant difference between their minimum and maximum map levels whether recorded in a clinic or using remote care systems. Furthermore, the studies also reported high satisfaction from the patients, parents[22] and CI staff (Buckman & Fitzharris 2020).

[18]for example, the SNR were the person can repeat half of the presented words accurately

[19]between zero and nine excluding seven as this is two syllables

[20]from the University of Southampton Auditory Implant Service

[21]Wesarg et al. (2010), Hughes et al. (2018), Goehring & Hughes (2017), Goehring et al. (2012), Muñoz et al. (2017)

[22]in the event of the remote care system being used on paediatric patient

Furthermore, a study into the usage of CI specific telemedicine systems found that the 94% of participants[23] said that using the prototyped telemedicine system would save them time and money (Cullington & Aidi 2017). This is relevent because money is always being considered in health care departments and correctly using telemedicine systems could save both the health care department and people who use CIs money.

2.6 Telemedicine in other clinical areas

The effect of telemedicine systems are well documented and one of the best researched examples is for cardiac problems. These are internet connected heart rate monitors that patients with a history of cardiac problems wear and monitor the patients heart rate; if they detect any of the early warning signs of a heart attack they have an alert procedure[24]. Anker et al. (2011) reviewed four different meta-analyses[25] all of which are investigating the effect of these remote heart rate monitors on long term patient mortality. There were only two cohort studies[26] which concluded that telemedicine had a negative or neutral impact on patients. The overwhelming majority of the rest of the papers including more than 12,876 participants between them, found that patients with a history of cardiac problems had a lower mortality rate (however slight) with remote monitoring. Subsequently, a meta-analyses by Wootton (2012) investigated evidence for the effectiveness of telemedicine system covering a variety of applications[27]. 37, 695 patients from 141 Randomised Controlled Trials (RCTs) were included in the analysis and while two studies (1.4%) found that telemedicine had a negative impact, the remaining studies all found that telemedicine had a neutral or positive outcome[28]. However, this study had a mean duration of 6 months which would have ideally been longer as some telemedicine systems may take a significant length of time for patients to become acclimatised to using them. There is significant evidence to support the statement that telemedicine systems have a positive outcome in the majority of cases, however none of these examples are specific to CIs or hearing loss.

An analysis was also carried out on 80 systematic reviews that covered telemedicine systems for asthma, diabetes and heart failure Ekeland et al. (2010). Of the included studies, 26% (n=21) found that the telemedicine systems used had a positive effect, 22.5% (n=18) said that they were promising but could not show anything different; 51% (n=41) of the studies found that there was insufficient data to draw a conclu-

[23] 15 out of 16
[24] For example: direct message to 999 or notification to specific team at hospital
[25] Clark et al. (2007), Klersy et al. (2009), Inglis (2010), Clarke et al. (2011)
[26] Chaudhry et al. (2010), Koehler et al. (2011)
[27] asthma, chronic obstructive pulmonary disease, diabetes, heart failure and hypertension
[28] 31 studies (22%) had a neutral outcome and 108 (76.6%) had a positive outcome

sion. Between all the included studies, none of the telemedicine systems used were shown to have a negative impact. The findings by Ekeland et al. (2010) are backed up by Wootton (2012) who investigated telemedicine systems for: asthma, chronic obstructive pulmonary disease, diabetes, heart failure and hypertension. $37,695$ patients from 141 RCTs were included in the analysis. Of these studies, 108 (76.6%) concluded that telemedicine systems had a positive impact; 2 (1.4%) reported a negative outcome and 31 (22%) had an unclear outcome. A myriad of other studies have shown that when telemedicine systems were used, they either provided a measurable improvement, no measurable improvement or there was not enough data to draw a conclusion (Clark et al. 2007, Koehler et al. 2011, Klersy et al. 2009). While a few studies[29] have shown telemedicine systems to have a negative outcome, the overwhelming majority of studies have shown neutral or positive outcomes. Furthermore, all these studies were conducted before COVID-19 therefore, in a time when face to face appointments are more problematic; the inference can be made that both patients and hospital staff would be more willing to try telemedicine systems.

The financial implications of utilising telemedicine systems have been investigated by a multitude of studies and while some evidence does support the hypothesis that they have financial benefits (Marietta 2001, Bynum et al. 2003, Luryi et al. 2020). Marietta (2001) investigated the cost-effectiveness of telemedicine on patients diagnosed with pre-term labour compared to a control group.[30] and found an average saving of £14,459 per patient. Bynum et al. (2003) investigated potential cost savings to patients of using telemedicine systems and found that 92% of patients saved > $32 in fuel costs, 84% saved > $100 in lost wages[31] and 74% saved > $75 in family expenses[32]. The study (Bynum et al. 2003) was done in Rural Arkansas which limits the real world applicability for the United Kingdom as Arkansas has a significantly lower population density of 56 people per square mile compared to 255[33]. However, a presentation by Helen Cullington shows a map of the patients University of Southampton Auditory implant service[34] stating that "*a lot of them (patients) are travelling for more than two hours*" to get to appointments. In contrast, multiple systematic reviews have concluded that telemedicine does not improve cost efficiency and further research is needed (Roine et al. 2001, Whitten et al. 2002, Mistry 2012, De La Torre-Díez et al. 2015).

[29] Two of the included 80 studies from Ekeland et al. (2010)
[30] treatment n=60, control n=40
[31] as they didn't need to have a day off work to come to an appointment
[32] babysitters, childminders etc.
[33] http://worldpopulationreview.com
[34] 4:20 of https://www.youtube.com/watch?v=kzvocpWUcCY

2.7 The digital divide

The digital divide is defined as "*The gulf between those who have ready access to computers and the internet and those who do not*"[35]. To phrase it a different way, there is a social divide between those who have good technical skills and access to the internet and those who do not. When designing any system in the future for people who use CI efforts should be taken to make it as simple to use as possible.

As CIs themselves are technical devices, the inference can be made that people who use CIs have a certain level of technical competence but this cannot be guaranteed. This means that any future system should be designed so it can be operated by the less digitally literate people who use CIs. A survey of people who use CIs showed that 60% (94 out of 158) had access to a computer and used the internet (Thorén et al. 2013). In the event that a CI microphone test system starts the people that are more technically literate are more likely to be the first ones to volunteer to trial a microphone test system. This study does show how any future system is going to need to be designed so it is inclusive of people who are less digitally literate.

[35] from: `https://www.lexico.com/en/definition/digital_divide`

Chapter 3

Methodology

Rather than looking at the methodology for a single experiment, this chapter will provide an overview as to how the multiple smaller experiments that make up this project work together to meet the overall research objectives. It would not have been possible to do one experiment to meet all the research objectives in Section 1.2.

A number of the experiments described in this chapter required a way of measuring the acoustic performance of a Cochlear Implants (CIs) T-Mic. Section 3.1 describes the various measurement methodologies that were considered, each of their advantages and disadvantages along with why one method was used for the majority of the experiments.

The rest of this chapter provides an overview of the multiple experiments that make up this project. Figure 3.3 is a spider diagram showing each of the experiments and what chapter they are in. The experiments are in blue boxes and the chapters are in orange boxes.

A significant number of the experiments focus on testing Advanced Bionics T-Mics, this was done for practical reasons. The advanced Bionics Listening Check facilitates a direct audio output for the CI microphone making it significantly easier to conduct the experiments. Advanced Bionics processors are in common use at the University of Southampton Auditory Implant Service (AIS) which can supply them for testing at various points in the microphones life cycle. Furthermore, because they detach from the CI processor it is both financially and practically easier to stockpile reportedly broken T-Mics for testing, due to the value of CI processors, it would not be possible to repeatedly test such numbers of processors. Using the T-Mic means there was only two processors needed[1] and all the microphones could be attached to these processors.

[1]a number of the experiments were repeated with both processors to ensure the processor was not affecting the results.

3.1 Measuring T-Mic performance

A number of experiments require accurate frequency responses to be taken of CI microphones and there are three main ways this could be accomplished: anechoic, using a small custom enclosure for the T-Mic or directly connecting to an Advanced Bionics T-Mic. Each of the three options were tried, each with advantages and disadvantages. The three methodologies are described in detail below.

Each person who uses a CI will have a plethora of settings on their CI processor that are specific to them. While it is known that the programme specific microphone setting will affect the listening checks audio output, it was not known which of the other settings would affect the output. Subsequently, for all of these methodologies, when a CI was required it was put into test settings with all the advanced features that could affect the results disabled. For more information about which of the advanced features affect the audio output see Section 6.2.

3.1.1 Anechoic testing

Anechoic testing works by placing a CI processor in an anechoic chamber then connecting it to an Advanced Bionics Listening Check that provides a 3.5mm headphone output. This output was then connected to a computer that plays sound though a speaker[2]. Just after testing, a reference microphone was put in the same location as the CI microphone and another frequency response was recorded. These two frequency responses[3] are then deducted from each other. The CI microphone frequency response will be a combination of a number of factors including: the microphone's frequency response, the response of the speaker and the room acoustics. Deducting the frequency responses from each other should produce a frequency response that is just the T-Mic's performance.

This method of deducting the T-Mic's frequency response from a second frequency response recorded with a reference microphone can be summarized in an equation. A is the other factors that could affect the recording including rooms acoustics and speakers frequency response. B is the frequency response of the microphone that is being measured. So when a frequency response of the test microphone is recorded (B) the other factors are also being recorded (A), this is denoted in Equation 3.1 as (A+B). When a subsequent frequency response is recorded using a reference microphone that has a known flat frequency response, this just records the other factors (A). Therefore,

[2]speaker in anechoic chamber, 1m from T-Mic

[3]Frequency response of the T-Mic and the reference microphone frequency response

deducting the frequency responses from each other results in just the microphones frequency response.

$$(A + B) - (A) = B \tag{3.1}$$

Testing microphones in anechoic chambers is exceedingly common in acoustics. However, testing in an anechoic chamber has limited relevance to CI clinicians as in a clinic they would be testing with a significantly more space efficient set-up. Furthermore, specialist equipment is required for anechoic testing making testing more problematic. Also, if the reference microphone is not placed in the exact location where the T-Mic was then this can cause results to become less accurate. This problem could grow if comparing each microphone to a reference frequency response recorded just after each microphone, but could be mitigated by having one reference microphone recording that all are compared to. Differences between different reference frequency responses could reduce the accuracy of the test and differences are more likely in the probable event that the equipment was needed to be packed down between uses[4].

3.1.2 Custom enclosure

This approach has similarities with the previous method as they both use the output from the Advanced Bionics Listening Check to provide a 3.5mm headphone output. However, this method does the recordings in a 3D printed custom enclosure that was designed to have a headphone attached in one end[5] and a T-Mic inserted in the other end. Figure 3.1 shows a picture of the enclosure and plans are shown in Figure 3.2.

This approach builds on the work by Erson et al. (2020) that built a custom enclosure that could house a small speaker, the Advanced Bionics Listening Check and a CI processor. This then connected using a 4-pole jack to a computer that measured the frequency response of CI microphones. While some frequency responses were recorded, the project lacked repeatability. In order for a test box to be successful it needs to be able to record frequency responses repeatability because you cannot record reductions in CI microphone performance if the results are not consistent. This project did discuss if they should have built a stand alone test box that did not require a computer, this does have significant advantages but it would have significantly complicated the design. The research by Erson et al. (2020) has a number of similarities with this project including the integrated Advanced Bionics Listening Check and conducting testing through a computer. It has emphasised the need for proving a test set-up's

[4]Because anechoic chamber needed by someone else, it is not practical to leave set up for this specific experiment all the time

[5]the headphone was glued in place to increase repeatability.

Figure 3.1: Picture showing a cochlear implant connected to a Listening Check and the custom enclosure

consistency.

The Advanced Bionics Q70 and Q90 technical specifications specify a frequency range of 150Hz to 10,000Hz (Advanced Bionics 2013, 2015*b*). In Equation 3.2, c is speed of sound (in m/s), f is the lowest frequency that the enclosure should affect and L is the length of the enclose in meters. With the maximum distance between the T-mic and the end of the speaker being 30mm[6], equation 3.2 shows that lowest standing wave should be at 5717Hz. This was lower than anticipated but should result in a small increase in level at that frequency.

$$\frac{c}{2L} = f \qquad \frac{343}{2 \times 0.03} = 5717Hz \tag{3.2}$$

To form a point of comparison, frequency responses from a number of known working CI microphones were recorded. The frequency specific average was then calculated so providing an accurate reference frequency response. Similar to the anechoic methodology, each microphone tested had the reference frequency response deducted from it showing each microphones' deviation from the reference frequency response.

[6] as shown by Figure 3.2

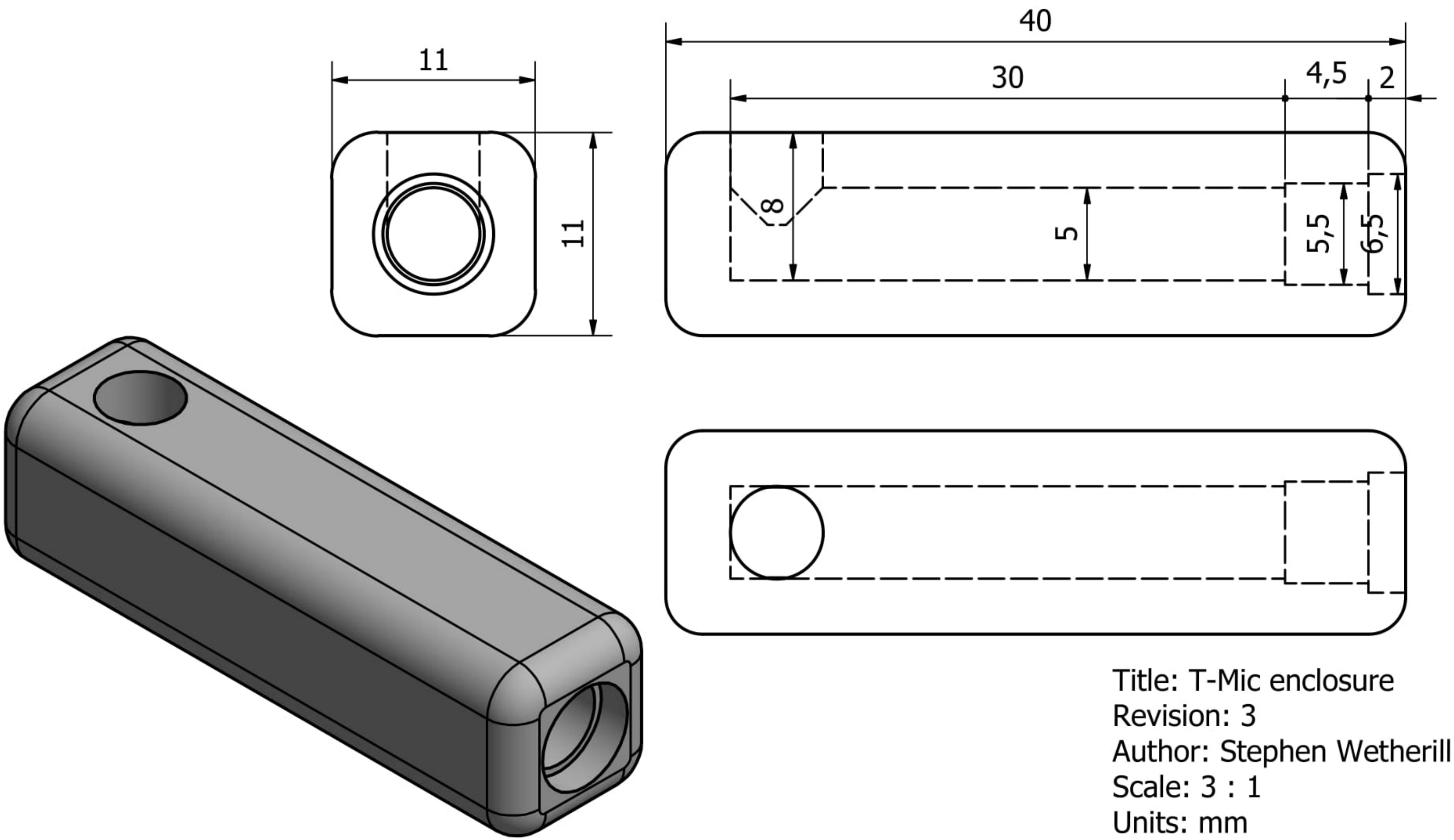


Figure 3.2: Diagram showing the specifications of the T-Mic enclosure

Using a significantly smaller enclosure than the anechoic chamber did increase the risk of the test environment interfering with the results. However, deducting the reference frequency response[7] should cancel this out. But, using a custom enclosure had the advantage where the speaker producing the frequency response stimuli and the microphone being tested were in the same place each time. This was because the custom enclosure was designed to precisely fit an Advanced Bionics T-Mic and hold them in place for the duration of each test. The consistent positioning of the speaker and the T-Mic should increase the test-retest repeatability.

The anechoic methodology has limited relevance to CI clinicians or people who use CIs as it would not be practicable for either group to have dedicated anechoic chambers for microphone testing. In contrast, the custom enclosure methodology would be feasible for a CI clinician or person with a CI to conduct to use. The methodology has numerous similarities with hearing aid test boxes that are widely used in audiology clinics. Both use small enclosures with a small speaker to measure microphone frequency responses.

3.1.3 T-Mic direct

Rather than using the output from the Advanced Bionics Listening Check, this last methodology connected directly to the T-Mic. While the T-Mic is a specialist piece of equipment, at its core it is a microphone. This methodology required knowledge of the electronics required to get the microphone working and the physical connection. Both are explained below.

The first problem was getting something that a T-Mic could be plugged into other than a CI processor. This was done by getting accurate measurements of both the T-Mic connector and what it attached into on the CI processor. From the accurate measurements, a 3D model was made and printed with a supplementary connector. A few iterations were required to get a secure connection. On the CI processor there were two small pins that the T-Mic used; implementing these correctly into the design took a few more iterations but the right measurements and substitute material was found.

Next the T-Mic electrical requirements needed to be determined as in order for the microphone to operate it required the right amount of power supplied in the right way. Furthermore, the system needed to be able to interpret the signals produced by the microphone. This was worked out by connecting a test microphone to a CI processor through the newly 3D printed adaptor, enabling multimeter readings to be taken. Using these measurements a small circuit was designed that both supplied sufficient

[7] combination of knowing working microphones

power and produced a headphone output. The exact specifications of the circuit used are not formally documented here due to copyright concerns, as reverse engineering of CI processor components was required.

It is not practicable for either a CI centre or someone who uses a CI to have an anechoic chamber for microphone testing. The custom enclosure described in Section 3.1.2 was used to measure microphone performance as this set-up would be practicable to use in either a CI clinic or for a telemedicine system. If this was to be used in either environment significant modification would be required to the user interface but it is still significantly more practicable than an anechoic chamber.

A consistent problem with the other methodologies is that the processors are required to be in test settings with the programme using the right microphone in the right ratio[8]. This is problematic as it is not common for people who use CIs to have a dedicated test programme on their processor, making this approach outside of laboratory conditions problematic. As this methodology bypasses the processor entirely, the results are not affected by the processor settings, which are different for every person who uses a CI.

This is an experimental set-up consisting of a 3D printed adaptor using custom electronics and is not manufacturer approved. This could cause significant problems if the manufacturer stops repairing microphones that have been used in this piece of equipment. For this reason the researcher was hesitant to use this device to test a microphone that is currently being used in a CI in case doing so voids the warranty on either the CI processor or the microphone.

3.1.4 Summary of measurement methods

Of the three methodologies examined here, the anechoic methodology was used for two of the experiments to obtain accurate metrics of microphone performance (Sections 6.1 and 6.2). For the remaining experiments, when a frequency response was required the custom enclosure was used for its superior real world relevance and similarity to current hearing aid test box measures. Table 3.1 summarises the advantages and disadvantages of the three different methodologies.

Using an anechoic chamber ensures significant isolation from background noise which is not present in the other two methodologies; but so long as the background noise is of reasonable levels then this should not be a problem when using the custom enclosure methodology. Furthermore, background sound levels were measured during testing and

[8] so not processor microphone 50% and T-Mic 50%

were below 50dB SPL for all testing.

The remainder of this chapter describes the experiments that were conducted and which methodology they used for gathering T-Mic data if required.

3.2 The problem with current testing

The first research objective laid out in Section 1.2 states "*How accurate and repeatable are subjective CI microphone checks*". Chapter 4 uses three experiments to investigate. All the experiments conducted are shown in Figure 3.3.

A survey was sent to clinicians from the University of Southampton AIS. This asked a variety of questions about how often they test microphones, which methods they use for testing CI microphone performance, how often they replace microphones and more. This was to attain information about the current status of testing in CI clinics.

If a CI microphone is suspected of not fully working then it is reported as broken. Section 4.2 is an experiment entitled "*Testing reportedly broken microphones*" that uses microphones from the University of Southampton AIS that were reported as broken. The custom enclosure methodology[9] was used to determine how many of the reportedly broken microphones were, in fact, broken and how many are were false positives. This determined how accurate and repeatable current methods for determining microphones as broken are.

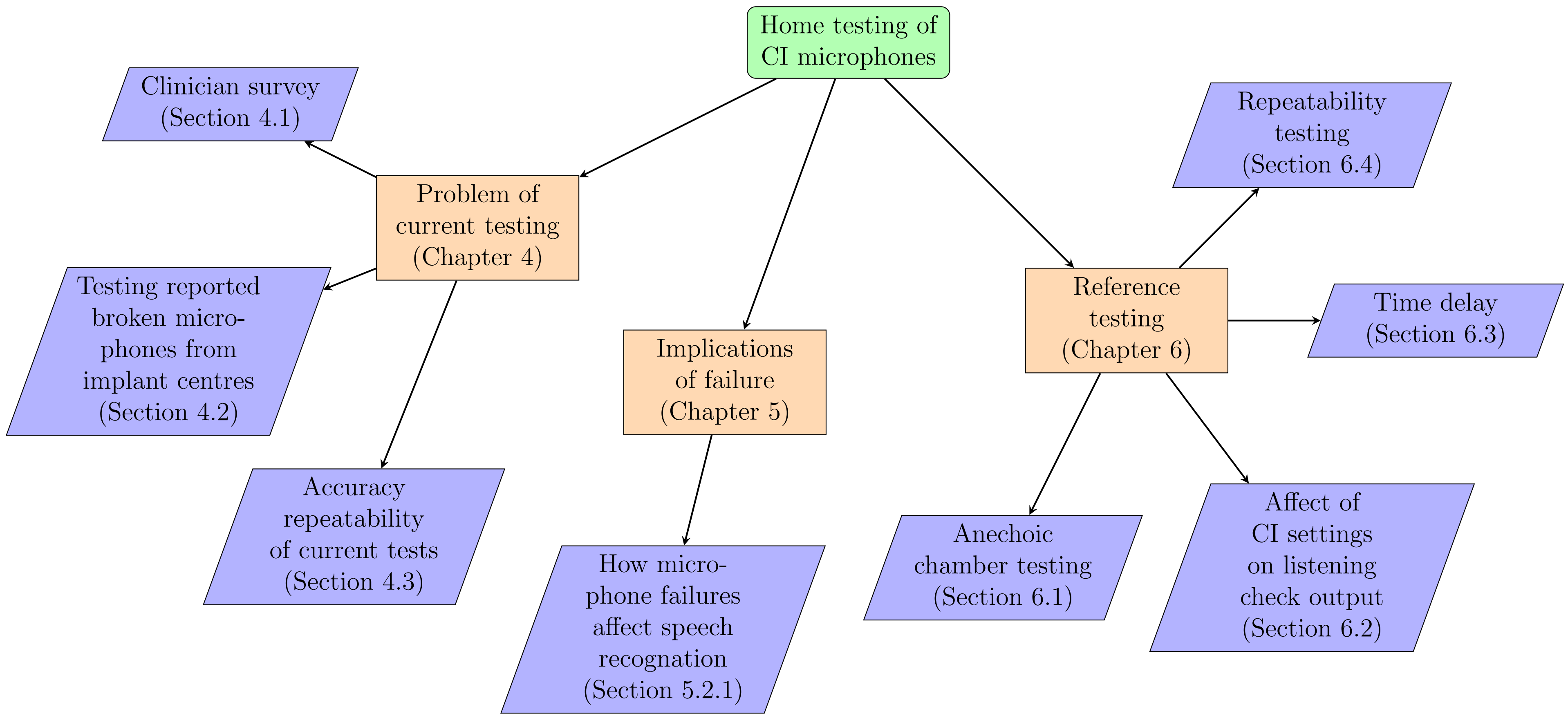
Different CI manufacturers offer different solutions for helping people determine if a CI microphone is working. One option is to use a special adaptor that allows a CI clinician to plug in a pair of headphones[10] and listen to the microphones output. The experiment entitled "*Accuracy and repeatability of current tests*" in Section 4.3 investigated how accurate and repeatable these subjective CI microphone checks were. Ten microphones were placed in front of participants: four fully working, three broken and four that were part working[11]. Participants sorted the microphones into three categories: fully working, part working and not working. This was done with ten normal hearing participants and ten CIs clinicians to see if there was a difference between the groups. The ten normal hearing participants[12] simulated a normal hearing parent or significant other of a person who uses a CIs. The CI clinicians performed subjective microphone checks as they would if they were performing it in their clinic. This ex-

[9] Described in Section 3.1.2
[10] using a standard 3.5mm mono headphone jack
[11] reduced performance at different areas of the frequency spectrum
[12] had a hearing screening with the Digit Triplet Test (DTT)

Table 3.1: Advantages and disadvantages of the different methodologies that could be used to measure microphone performance

Methodology	Advantages	Disadvantages
Anechoic	• Widely used in acoustics • Enclosure least likely to effect the results • Highest backound noise resilience	• Specialist room required • Reduced accuracy if reference microphone placement is not accurate • Hard to position everything in the same place each time, especially when tests not done back to back • Might be influenced by the CI processor settings
Custom Enclosure	• Repeatable T-Mic position • Significant real world relevance	• Smaller enclosure more likely in interfere with results. • Might be influenced by the CI processor settings
T-Mic direct	• Repeatable T-Mic position • Bypasses CI processor settings	• Experimental, required reverse engineering

Figure 3.3: Flow chart showing the experiments conducted and where about in the report structure they are



periment shows the accuracy and repeatability of this specific method of checking CI microphone performance.

3.3 Implications of failure

Chapter 5 addresses the second research objective that is "*What impact do partially failed CI microphones have on speech perception?*" This chapter looks at both the outcomes of false positives and false negatives.

False positives are the microphones that were reported as broken but did not have any detectable defect. Not replacing these microphones unnecessarily presents a significant potential saving to NHS cochlear implant departments. In a private environment where people who use the CI may be expected to pay for each replacement microphone it could be a significant financial expenditure.

False negatives are microphones that do not work but are thought to be fully working. The experiment entitled "*Affect of simulated reductions in microphone performance on people who use CI*" (Section 5.2.1) applied filters simulating the part working microphones to a speech in noise test. This was then used to determine how using different part working microphones would affect speech perception.

3.4 Reference testing

While multiple companies offer devices that can be used to subjectively check the functionality of a CI microphone, the Advanced Bionics Listening Check is designed so a pair of headphones can be plugged in and listen to the microphone output, thus enabling subjective microphone checks. Research objective three is "*What processor settings affect the audio output of the Listening Check?*". Chapter 6 performs various experiments to answer this research objective.

The first experiment (Section 6.1) used the anechoic methodology (Section 3.1.1) to objectively measure the performance of known working CI microphones. This was done to obtain an accurate information on the baseline functionally of a CI microphones.

The next experiment, in Section 6.2, built on the prior experiments results. The anechoic methodology was used to measure the T-Mics output with various advanced features and with the CI processor settings changed to determine which affected the listening checks output. This was important because everyone who uses a CI has different settings and if these settings affect the output then results will not be directly comparable between CI processors belonging to different people. This will significantly affect how results can be compared in the future.

Nothing with technology happens instantaneously, everything takes a finite amount of time to go through the required computers and processing before getting an output, whether it is moving a mouse on a computer or a device processing audio. The third experiment (Section 6.3) used the custom enclosure methodology to investigate how long it takes audio to go from the T-Mic, through the CI processor and to the Listening Check output. This was done so it could be taken into account for any subsequent tests to improve experimental accuracy. This will affect what stimulus should be used when measuring the Listening Check.

The custom enclosure methodology is used for a number of experiments in this project, so this experiment (Section 6.4) investigates how repeatable this methodology is. As this methodology was comparing a known working microphone to another microphone, absolute accuracy was not needed so long as the set-up was providing repeatable results. This was because a drop in performance of a microphone can be assessed by the difference from a known working microphone. Repeatability is measured by taking a multitude frequency responses over a period of time and comparing them to see what the variation between them was.

Chapter 4

The problem with current testing

As Cochlear Implant (CI) microphones are worn on the head they get exposed to a considerable amount of dust, moisture and airborne debris. This, in addition to general wear and tear, causes CI microphones to break. Reductions in microphone performance can be picked up one of two ways; by Auditory Implant Service (AIS) clinicians or by a person who uses a CI.

When people with CIs have check-up appointments at their implant centre[1] they will undergo a series of tests which include a sound field[2] speech in noise test. People with CIs will have a baseline test on record and if the results from the test are significantly lower, then the clinician will do various trouble shooting methods including replacing the microphone. This method for testing has significant limitations; firstly it requires an appointment which can be infrequent[3] and hard to attend for people with CIs. Secondly, this test methodology does not test the CI microphones in isolation but rather a the complete system, from microphone all the way to the speech processing in the brain. Therefore, this test offers no chance for localising a problem and in the occasions of a sub-standard test the microphone is replaced as a precautionary measure leading to significant numbers of false positives.

Currently when a CI microphone is reported as broken to the AIS[4], if it is less than a year old it is sent back to the manufacturer to be replaced, if older it is discarded; this is shown in Figure 4.5. Of the microphones that get reported as broken, it is not known how many of these actually have a fault as a considerable number of them are thrown away. Section 4.2 provides information on an experiment that measures microphones that would have been binned, to see the prevalence of different types of reduction in

[1] the exact duration between check ups will vary depending on a variety of factors including how long they have had their CI for.

[2] Using speakers rather than headphones

[3] sometimes years between

[4] either by a clinician or someone else

microphone performance.

People who use CIs might notice a reductions in microphone performance during their normal day, such as not being able to hear someone speak in a acoustic environment they normally can. This Listening Check device can also be used to do subjective microphone checks by the family, friends or significant other of someone who uses a CIs. Relying on subjective checks either by the person who uses a CI or someone else is not ideal especially when no previous study has investigated the accuracy and repeatability of these subjective microphone checks. This is explored in Section 4.3.

4.1 Clinician survey

When a person who has a CI is seen in a CI clinic there are a lot of things that need to be done in the appointment. The aim of this survey was to find out baseline information about what CI microphones checks are done in a clinic and what affects those decision.

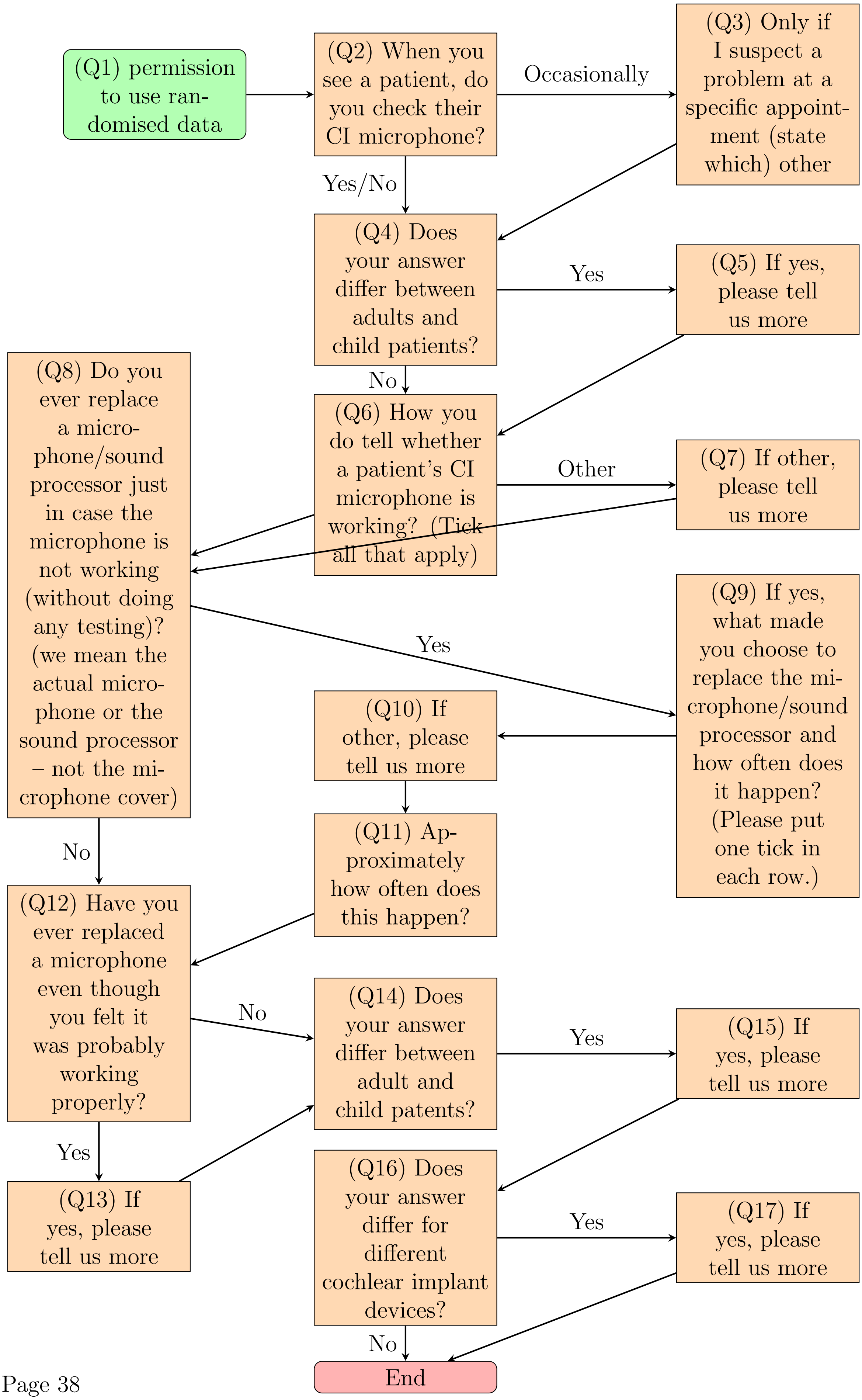
4.1.1 Methodology

The survey was built using Microsoft Forms in compliance with University of Southampton questionnaire guidance and ethics department[5]. The survey was then distributed by email to the University of Southampton AIS clinical staff. The questions were carefully chosen to investigate areas that were suspected to be of importance. Below is the complete list of questions and depending on answers to prior questions not every person saw every question, this is shown in Figure 4.1.

1. I agree that my responses to the survey can be used in this study and I understand that due to this survey being anonymous it is not possible to withdraw my results at a later date.
    - I agree
    - I do not agree
2. When you see a patient, do you check their CI microphone?
    - Yes
    - No
    - Occasionally
3. Only if I suspect a problem at a specific appointment (state which) other

[5]ERGO number: 59833

Figure 4.1: Flowchart showing the survey questions and how they linked together



4. Does your answer differ between adults and child patients
   - Yes
   - No
5. If yes, please tell us more
6. How you do tell whether a patient's CI microphone is working? (Tick all that apply)
   - Patient's report
   - Visual inspection
   - Sound field audiogram
   - Speech perception testing
   - Listening Check
   - Other
7. If other, please tell us more
8. Do you ever replace a microphone/sound processor just in case the microphone is not working (without doing any testing)? (we mean the actual microphone or the sound processor – not the microphone cover)
   - Yes
   - No
9. If yes, what made you choose to replace the microphone/sound processor and how often does it happen? (Please put one tick in each row.)
   - Patient report
   - Visual inspection
   - Sound field audiogram
   - Speech perception testing
   - Unable to verify microphone is working
   - Coming to end of warranty period
   - At a specific appt (state at which appt in the text box below)
   - Annually or at other interval (please specify in the text box below)
   - Other (if other please specify in the text box below)
10. If other, please tell us more

11. Approximately how often does this happen?
    - Multiple times a week
    - Once a week
    - Once every 2 weeks
    - Once a month
    - A few times a year
    - Never/ Less than once a year
12. Have you ever replaced a microphone even though you felt it was probably working properly?
    - Yes (please specify)
    - No
13. If yes, please tell us more
14. Does your answer differ between adult and child patents?
    - Yes (please specify)
    - No
15. If yes, please tell us more
16. Does your answer differ for different cochlear implant devices?
    - Yes (please specify)
    - No
17. If yes, please tell us more

4.1.2 Survey question justification and results

Rather than justifying the questions then later going over the results, each question is justified and then the results are explained. A total of ten clinicians responded to the survey with an average time taken to complete of 6 mins and 22 seconds[6]. The first question is required for the survey to pass the University of Southampton Ethics process.

Question two asked if clinicians check CI microphones, this was to see how often CI microphones are checked at appointments. Of the ten participants four answered no

[6] Minimum 1:56, Maximum of 18:01

and six answered occasionally. This means that none of the clinicians check the microphones for every patient. Those that answered occasionally to question two were prompted for more information in question three. Responses to question three showed that the clinicians who answered occasionally would ascertain from the interview if it was necessary to check the microphone.

Question four asks if the answer to question two differs between children and adults. This was important to determine how often different groups of people who use CIs have their microphones checked. Seven of the clinicians said their answer did not differ between adults and children but three said it did. When asked for more information in question five, they said that due to children not being able to conduct some clinical tests[7] and they are unable to report problems themselves.

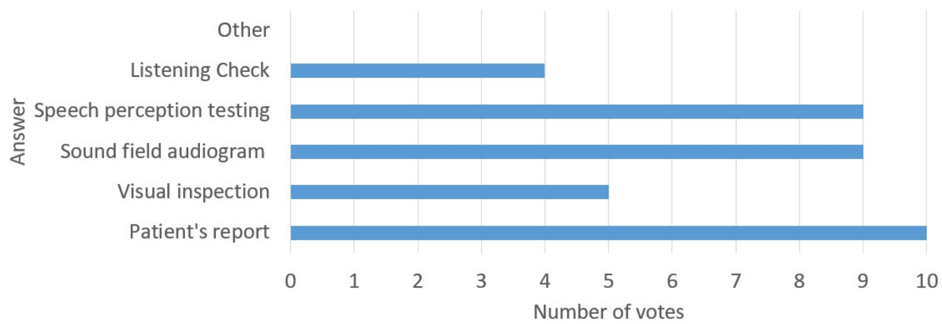


Figure 4.2: Survey results for the question 6: "*How you do tell whether a patient's CI microphone is working?*"

Question six asks how clinicians tell if a microphone is working, there were six options presented (including "other") and they could could tick as many options as applied. Figure 4.2 shows how many votes each of the answers got. This showed that all of the clinicians take the patients report into account and the majority (9 out of 10) also take into account the results from speech perception testing and sound field audiogram. Question seven was a chance for people to write any information they want but some of the answers to question three are also relevant. Between questions three and five, five of the clinicians stated they would check the microphone only if there was a decline in other test results. This is because the sound field testing[8] will test all the CI system from the microphone picking up the signal all the way to the brain interpreting the sound. Therefore, if the results for a sound field test drop then this indicates there is a problem somewhere in the system. In this situation, checking the CI microphone is done to determine if the CI microphone is the cause of the reduction in performance.

[7]specifically Sound field audiogram

[8]Both speech perception testing and sound field audiogram and speech perception testing can be done in a sound field

Question eight asks if the clinicians replace microphones "*just in case*". Six of the clinicians said they had replaced a microphone "*just in case*". An exact binomial confidence interval shows there is a 95% chance the true population mean lies between 26.2% and 87.8%. Question nine asked those those six people that answered yes to question eight what made them replace the microphone. As with question six, as many options as applied could be chosen with and Figure 4.3 shows the results. The results for this question are very similar to those of question six showing how high values placed on patient reports and sound field testing. Question ten was a chance for clinicians to provide more information and there were two responses, both stated if the microphone had not been replaced in the last three months.

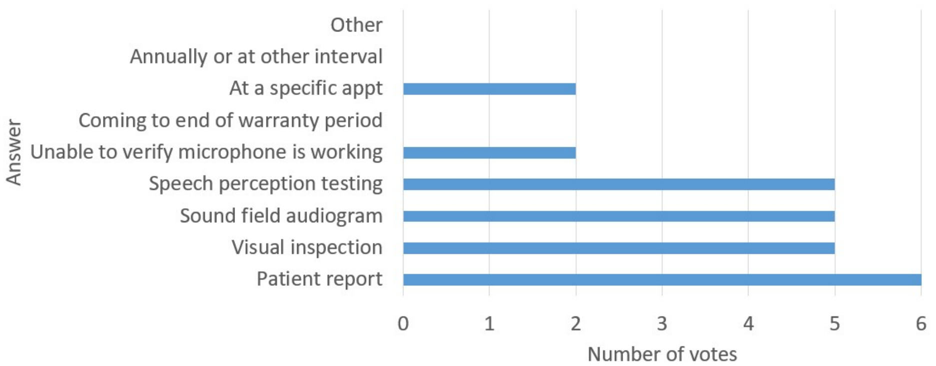


Figure 4.3: Survey results for question nine: "*If yes, what made you choose to replace the microphone/sound processor and how often does it happen?*"

Question 11 was answered by the six people who answered yes to question eight and asked how often this happened. There were six options ranging from "*Multiple times a week*" to "*Never/ Less than once a year*" and the results are shown in Figure 4.4 shows the results. While one clinician did put "*Multiple times a week*". The other five clinicians put less frequent options indicating that on average, clinicians at the AIS are replacing microphones "*just in case*". The reason for the one answer "*Multiple times a week*" is not known but could be due to the number and type of patients that they see.

Question 12 asks all the clinicians if they have replaced a microphone that they thought to be working and nine out of ten clinicians said they had. An exact binomial confidence interval shows there is a 95% chance the true population mean lies between 55.5% and 99.7%. Thus more than half of clinicians have replaced a microphone they thought was working. When asked for more information in Question 13 and two responses were because the clinicians were erring on the side of caution, the other seven responses were variations on ruling out other problems.

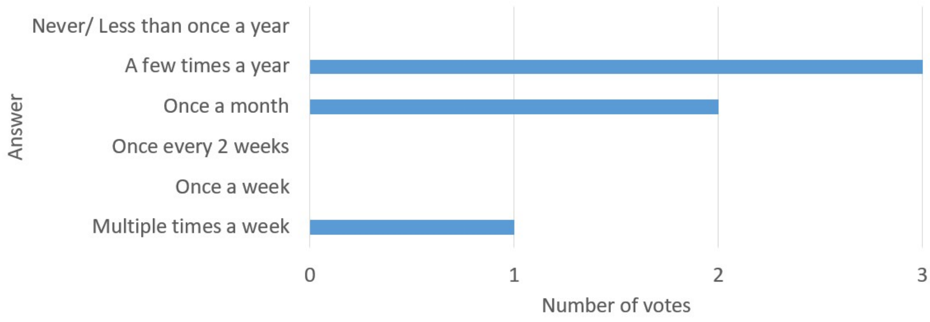


Figure 4.4: Survey results for question eleven: "*Approximately how often does this happen?*"

Question 14 asks if the results differ between adults and children, two clinicians answered yes. When asked for more information in Question 15, one of the clinicians said they were more likely to replace with children as they cannot self report problems themselves as well.

Question 16 asks if the answer differs for different CI manufacturers, five answered yes and five answered no. When asked for more information in next question four of the five that said yes to Question 16 said they were more prone to replace Advanced Bionics T-Mics with one. The remaining clinician[9] said that the CP810 and CP910 microphone covers were more prone to getting blocked.

4.1.3 Summary

Six out of ten clinicians replace microphones "*just in case*" every one or two months and nine out of ten clinicians replace microphones they thought could be fully working. These results show a lack of confidence by clinicians in CI microphone performance. The clinicians cannot be sure if the microphone is fully working and they do not want a broken microphone to negatively impact the auditory life of a person who uses CIs. The number of microphones replaced "*just in case*" likely results in a significant number of microphones that have been reported as broken yet are fully working, the next experiment investigates this (Section 4.2).

Sending this survey out to a larger number of CI centres would have been preferred but when the experiment was started there was not sufficient time to go through the National Health Service ethics process that would have been required.

[9] Of the five that answered yes to Question 16, with the other four saying Advanced Bionics.

4.2 Testing reportedly broken microphones

The previous experiment showed a significant number of microphones are replaced in the CI centre as a precautionary measure. In this experiment Advanced Bionics T-Mics that were reported as broken at the University of Southampton AIS were tested. Their performance was objectively measured to see how many are false positives.

4.2.1 Methodology

As stated at the start of Section 3, this experiment focuses on T-Mics for practical reasons. This experiment took the microphones that were reported as broken and more than a year old, which would have been discarded, as shown by Figure 4.5. Frequency responses for the microphones were collected using the Listening Check device then comparing the frequency response to known working microphones.

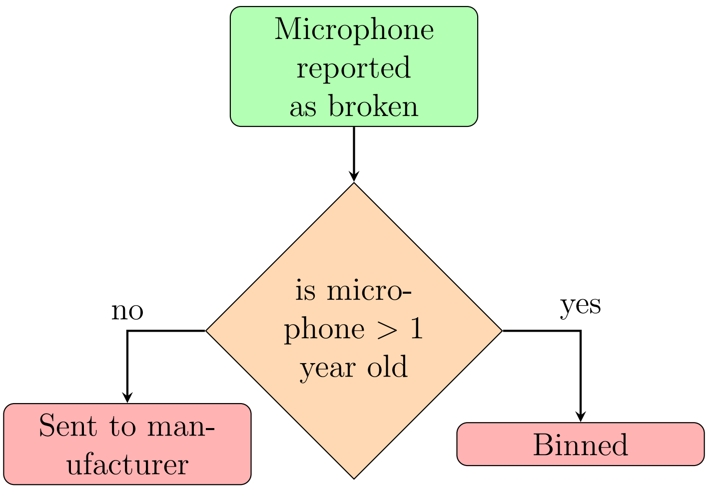


Figure 4.5: Flow chart showing the Auditory Implant Service protocol for broken T-mic

While it is known that the current programme on a CI processor will affect what microphone is being heard from the Listening Check device, it is not currently known what affect changing the processor settings would have on the audio output[10]. Therefore, the CI processors that were used for testing were locked to the T-Mic and has many settings and advanced features disabled.

Measurements for this experiment were done using the custom enclosure methodology as described in Section 3.1.2. White noise was used as an auditory stimulus as this enabled the entire frequency range to be tested simultaneously while not being affected

[10]But this is investigated in Section 6.2

by the time delay that the Listening Check introduces[11]. This meant that white noise could be played and microphone data could be recorded the entire frequency spectrum multiple times a second increasing the number of measurements points. Frequency responses were recorded every 50ms for 30 seconds, then the average frequency response was taken for each microphone.

Originally JavaScript was used to measure the frequency responses as this is a commonly used programming language on the internet. This would likely have made rolling out a test to people with CIs easier and quicker. However, the JavaScript audio interface system is complex with additional stages of digital signal processing that cannot be bypassed making testing less accurate and repeatable. Subsequently, python code was written that both played white noise and recorded the audio output.

The reference frequency response that this experiment compared other experiments to, was an average of the 53 frequency responses from the repeatability testing in Section 6.4, in addition to four further frequency responses from new working T-Mics. Frequency responses of the reportedly broken microphones were compared to the reference frequency response and the results are shown as a frequency differential graph.

4.2.2 Results

Objective frequency responses were recorded for 30 microphones using the methodology described in Section 3.1.2, the frequency responses were then compared to four brand new known to be working microphones to determine functionality. 2 (7%) of the microphones were broken in two so clearly would not have an output. 6 (20%) of the microphones had no detectable output so were completely not working. 2 (7%) microphones only worked some of the time so behaved intermittently. 17 (57%) of the microphones had frequency responses within ±3dB of the reference frequency responses so were fully working. The final 3 microphones (10%) did have an audible output but the frequency response was at least 3dB below the reference frequency responses. These results are summarised in Table 4.1 and the maximum, minimum and average deviation[12] from reference values for each of each microphone is in Table 4.2.

It has not been shown how accurate this set-up is with respect to absolute dB levels but it can repeatability measure microphone performance[13], which is important. As Section 6.4 found the test set-up had an approximate deviation from reference of under ±2dB. It is therefore highly likely that any microphone getting within ±3dB of reference is

[11]See Section 6.3
[12]Measurements between 100Hz and 10kHz, mean averages
[13]as shown in Section 6.4

Table 4.1: Objective performance of reportedly broken microphones

Type	Number (%)	Description
Working	17 (57%)	All values within ±3dB of reference
Partially working	3 (10%)	Some values at least 3dB below target
Intermittent	2 (7%)	Worked some of the time
Not working	6 (20%)	No measurable output
Broken in two	2 (7%)	Not measured. Clearly would not have output

Table 4.2: Microphone difference from reference response

ID	Average	Min	Max		ID	Average	Min	Max
1	-0.21	-2.43	1.66		17	-30.77	-25.67	-39.67
2	-0.26	-2.44	2.27		18	-0.25	-2.27	1.8
3	-0.02	-2.66	1.95		19	-56.36	-69.02	-28.06
4	-0.22	-2.32	1.63		20	-0.06	-1.83	1.82
5	-0.32	-2.94	2.1		21	-30.77	-25.67	-39.67
6	-0.32	-2.34	1.72		22	-0.28	-2.64	1.82
7	-56.37	-69.05	-28.01		23	-0.16	-2.2	1.96
8	-0.28	-2.98	1.69		24	-55.66	-69.71	-27.62
9	-0.26	-3.03	2.39		25	-3.95	-18.16	0.94
10	-0.14	-2.09	1.37		26	-0.16	-2.32	1.6
11	-39.37	-66.57	-28.23		27	-56.37	-69.14	-28.03
12	-0.03	-2.26	1.83		28	-0.48	-3.52	2.06
13	-55.64	-69.13	-27.62		29	-0.06	-2.15	1.51
14	-0.09	-2	1.93		30	-24.21	-60.61	-14.69
15	0.02	-2.05	1.8					
16	-0.03	-1.89	1.7					

still fully working. Approximately 60% of the microphones tested were within ±3dB (peak) of the reference responses and ±0.5dB (average). These were all microphones that were fully working and no reason could be found why they could not be used clinically.

4.2.3 Summary

With 57% of the microphones performing within 3dB of reference values, the current method used clinically for identifying the microphones that need replacing sends a significant number of working microphones to be replaced. The exact confidence interval is 39% to 74% showing that of the microphones that the AIS sent for replacement, between 39% and 74% were working. This was lower than anticipated as clinicians often said the microphones were replaced "*just in case*", especially in paediatrics.

Of the two microphone that were identified by subjective tests as being intermittent, one tested fully working, which shows that objective tests have their limitations but an intermittent microphone fault is likely to be noticed by the person who uses the CI. If a person who uses a CI says there is a microphone problem it would still be advisable to replace the microphone, regardless of the objective test outcome, until an objective test can be shown to identify intermittent microphones consistently.

The impact of false positives cannot be ignored though, with each microphone costing more than £140, departments stand to save significant amounts of money if the number of false positives can be reduced. There is also the impact on the people who use CIs to consider, replacing microphones unnecessarily will cause significant effort for some people who use CI, especially if they are unable to change their own microphone necessitating a need to make a trip to their implant centre that could be hours drive away.

This experiment only tested microphones that had been reported as broken and were more than a year old, as shown in Figure 4.5; microphones that were less than a year old were sent back to the manufacture under warranty. While an exact number of microphones that were sent for replacement was not available; between June 2019 and June 2020 the AIS ordered 267 new microphones and this does not include the T-Mics that are included with new CI processors. If 57% of the 30 microphones tested were fully working, it would seem reasonable that a higher percentage of microphones would be working for the microphones that are less than a year old.

4.3 Accuracy and repeatability of current tests

This experiment was conducted aiming to measure the accuracy and repeatability of getting untrained, normal hearing people to subjectively measure how well CI processor microphones were working. This was necessary because it is one of the few ways of measuring CI microphone performance.

4.3.1 Methodology

Each participant did the Digit Triplet Test (DTT) before starting the study, to check that each participant had normal hearing. In clinic partners or family members do not have any listening checks done before doing any microphone checks. Pure Tone Audiometry (PTA) was a viable alternative screening test but DTT was chosen for practical and time concerns. The DTT is a viable screening test with multiple studies proving the tests accuracy both in clinics and when using a plethora of different hardware (Ozimek et al. 2009, Cullington & Aidi 2017).

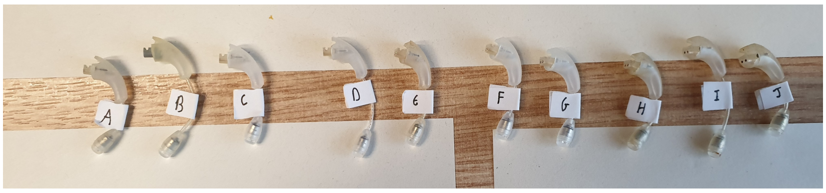


Figure 4.6: Picture of the ten labelled microphones.

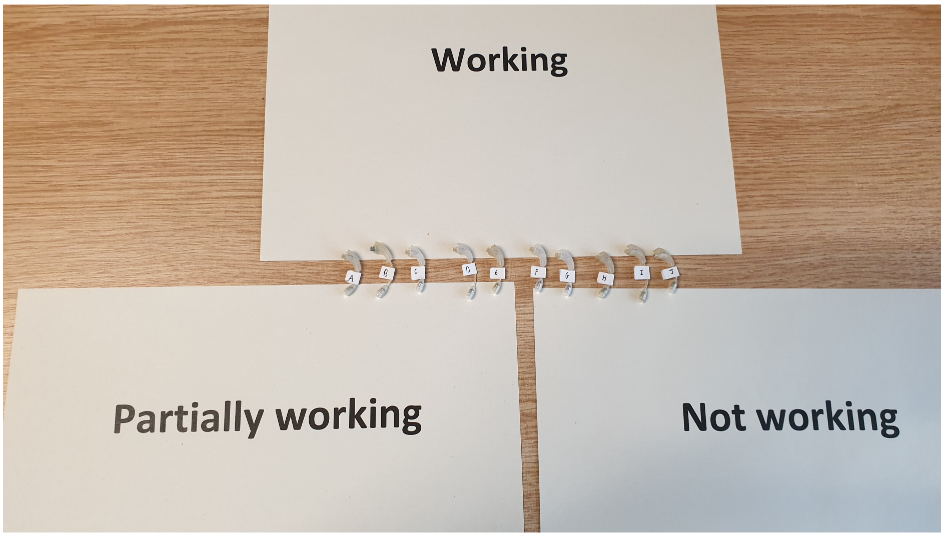


Figure 4.7: Ten labelled microphones as presented to participants with the labelled pieces of paper

Ethics approval was obtained from The University of Southampton (ERGO: 49657) and participants were recruited from around the University of Southampton because this is an abundant source of normal hearing test participants.

Each study participant was then presented with T-Mics which were labelled "A" to "J". Three pieces of paper were placed before each participant labelled "fully working", "completely broken" and "part working". Each of CI microphones had a randomly assigned letter ensuring there was no obvious patten[14] with the labelled microphones shown in Figure 4.6. Figure 4.7 shows the microphones next to the pieces of paper as they were presented to each participant. The microphones were assigned their letters by a third party who did not conduct the test, therefore eliminating any subconscious hints that the researcher could have accidentally given participants. The ten microphones were divided into three categories which are listed below:

- **Fully working:** A microphone that sounds normal at all frequencies

[14]such as the working microphones being D, E and F

- **Part working:** A microphone that may only be working at some frequencies (eg weaker low frequencies or weaker high frequencies) or might only work some of the time.
- **Not working:** A microphone with either no output at any frequency or where only electrical noise is heard (such as humming).

In real world situations clinicians or the significant others of people who use CIs will, if subjectively checking the functionality of a microphone, usually only listen to it for a few seconds before judging if the microphone is fully working, partially working or not working. For this reason test participants took longer than 20 minutes to categorize the ten microphones, they would be asked to stop. While two minutes is longer than someone would likely spend testing an individual microphone, rushing test participants may negativity impact results. 20 minutes was a compromise between giving people a realistic time frame to categorise the microphones while not allowing participants so much time as to reduce the real world applicability of the experiment.

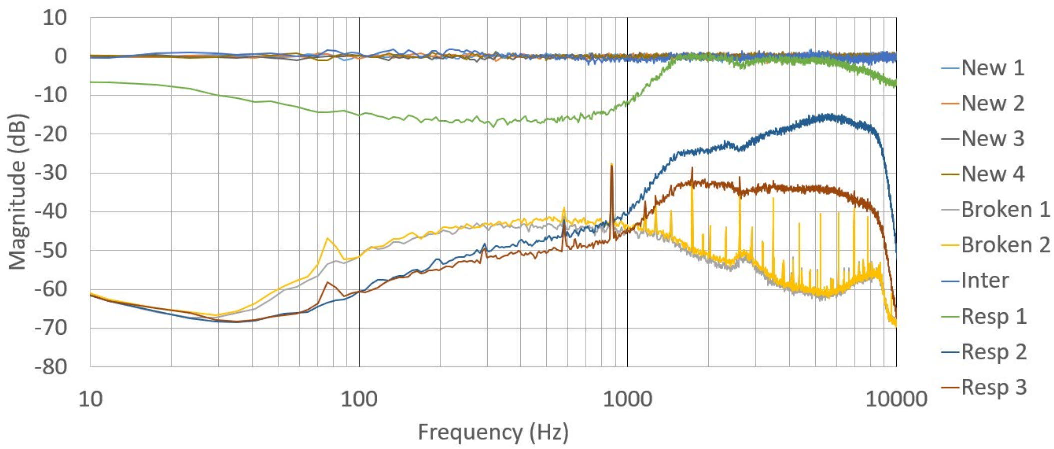


Figure 4.8: Graph showing the frequency responses for the ten microphones that were presented to test participants

The working and partially working microphones were both sourced from the supply of reportedly broken microphones provided by the University of Southampton AIS. Each microphone sourced this way was thoroughly checked for visual defects that could provide a visual indication on the functionality of a microphone[15].

Figure 4.8 shows the frequency responses for all the microphones that were used for testing. Four of the microphones were new and known to be working. The intermittent microphone had a normal response at times, but at other times would not work at all. The ‘Partially working’ microphones had varying frequency responses that were

[15]such as a partially or completely broken cable

chosen to represent a broad range of frequency responses. "Resp 1" have a frequency response close to reference between 1.5kHz and 7kHz. "Resp 2" having reduced performance above 1.5kHz and negligible performance below 1.5kHz. "Resp 3" had a similar response to "Resp 2" only a bit more severe.

Table 4.3: Frequency response deviation from reference recorded prior to testing.

Microphone ID	Average deviation between frequency responses recorded before and after testing (in dB)
Working 1	1.7
Working 2	1.2
Working 3	1.3
Working 4	2.1
Partially 1	1.9
Partially 2	0.9
Partially 3	1.4
Intermittent	30
Not Working 1	1.8
Not Working 2	1.6

Frequency responses for all the microphones were recorded before and after testing. Table 4.3 shows the maximum deviation in frequency response between the measurements for each of the tested microphones. For example, a value of 1.7dB indicates that the maximum deviation from frequency response recorded prior to testing between 100Hz and 10Hz was 1.7dB.

4.3.2 Results

Ten participants were recruited from around the University of Southampton campus and all passed a DTT screening test. Once testing had finished, the ten microphones were retested upon the completion of testing and the results were within ± 1.7dB of their previous results[16] except for the intermittent microphone. This microphone broke completely at some point in during the testing process, therefore, due to it's changing performance it has been excluded from all data analysis. There was also little between subject variability (P=0.9616)[17].

Of the two "*Not working*" microphones presented to the participants, all ten participants identified these microphones correctly with 100% accuracy. This is not surprising

[16] At all measured points between 100Hz and 10kHz
[17] Logistic regression

because all they needed to do was determine if they could hear anything or nothing.

For the three "*Partially working*" microphones, the average accuracy for all ten participants was 40% but Table 4.4 is an error plot showing the total number of microphones in each category presented to the participants. In brackets are the percentage of the presented microphones that were put into each response category. The error plot shows participants were twice as likely to categorise a "*Partially working*" as "*Not working*" than as "*Working*".

Table 4.4: Subjective microphone check error plot.

		Response		
		Working	Partially working	Not working
Presented	Working	30 (75%)	8 (20%)	2 (5%)
	Partially working	5 (15%)	12 (38%)	13 (46%)
	Not working	0 (0%)	0 (0%)	20 (100%)

The "*Working*" T-Mics were identified correctly 75% of the time ($P < 0.0001$)[18]. The fully working category is made up of two different microphones, three new and one good as new. The good as new microphone had a frequency response within 0.5dB of the new microphones yet Table 4.5 shows it was marked as fully working only once out of ten times. The difference between the "*Good as new*" and "*Partially working*" microphones was not statistically significant ($P=0.6724$)[19]. "*Estimated odds ratio for*" "*Good as new*" vs "*Partially working*" = 0.333 meaning the "*Good as new*" microphones less likely to be classed as working, i.e. the wrong way around); 95% confidence interval (0.005, 3.814). The T-Mics come out of the factory white but after being used they tend to get slightly discoloured. This slight visual difference between the "*Working*" and "*Good as new*" could have effected the results. If participants noticed this visual difference either consciously or subconsciously then they might have assumed the microphones that had no discolouration were the "*Working*" microphones. However, this cannot be determined without further testing.

[18]Exact logistic regression, including participant and microphone type as factors.
[19]Exact logistic regression, including participant and microphone type as factors.

Table 4.5: Subjective microphone check error plot with separate good as new

		Response		
		Working	Partially working	Not working
Presented	Working	29 (97%)	1 (3%)	0 (0%)
	Good as new	1 (10%)	7 (70%)	2 (20%)
	Partially working	5 (15%)	12 (38%)	13 (46%)
	Not working	0 (0%)	0 (0%)	20 (100%)

4.3.2.1 Results from clinicians

Once the results from the 10 participants had been analysed, some changes were made to the microphones that were presented and then CI clinicians did the same experiment using the updated 10 microphones. The changes to the presented microphones are outlined in Table 4.6. The number of "*Good as new*" microphones included was increased from one to two as this would enable more accurate comparison between the "*Working*" and "*Good as new*" microphones. There was four "*Partially working*" microphones tested but due to the intermittent microphone failing this was decreased to three included microphones. The number of not working microphones remained unchanged from the previous testing.

Table 4.6: Table showing the ten microphones that were presented to cochlear implant clinicians who participated in the study

Category	**Sub-Category**	**Quantity**	**Notes**
Working	New	3	Same as previous
	Good as new	2	1 more than previous
Partially working	n/a	3	Was previously 4 but intermittent microphone was excluded
Not working	n/a	2	Unchanged

The plan was to test ten CI clinicians from the University of Southampton AIS; however, the COVID-19 lockdown resulted in testing being halted after only four clinicians has completed the test. But for the participants tested there was little between participant variability (P=0.6103)[20].

Table 4.7 shows the error plot of the four clinicians that completed the experiment. As with the previous ten participants "*Not working*" microphones were identified with 100% accuracy. CI clinicians were more accurate at correctly categorising "*Partially working*" microphones with an accuracy or 67% compared to 38% for the other participants. The CI clinicians had a lower accuracy than the other participants and were more likely to mark a "*Working*" microphone as "*Partially working*"[21].

Table 4.7: Subjective microphone check error plot for cochlear implant clinicians

		Response		
		Working	Partially working	Not working
Presented	Working	12 (60%)	7 (35%)	1 (5%)
	Partially working	2 (17%)	8 (67%)	2 (17%)
	Not working	0 (0%)	0 (0%)	8 (100%)

How likely a clinician is to replace a microphone that is fully working can be calculated from these results by dividing the total number of times a "*Working*" microphone was categorised incorrectly by the total number of times that type of microphone was presented. This results in a 40% chance of a "*Working*" microphone being replaced unnecessarily. Also the results show that there is a 16.7% chance of a "*Partially working*" microphone being marked as fully working. The difference between the "*Working*" and "*partially working*" was statistically significant (P=0.0363)[22]. The difference between "*Good as new*" and "*Partially working*" was not statistically significant (P=0.2069)[23].

4.3.3 Summary

Chronologically this was the most recent experiment and the COVID-19 lock-down caused testing to be suspended, because of this only four of the ten planned CI professionals were tested. While testing only part of the participants was not ideal from an experimental prospective but unavoidable.

For this experiment, using new or good as new microphones is a compromise and using

[20]Logistic regression, including participant and microphone type as factors.
[21]38% compared to 20%
[22]Exact logistic regression, including participant and microphone type as factors.
[23]Exact logistic regression, including participant and microphone type as factors.

both has advantages and disadvantages. Good as new microphones are extensively tested and values compared to new known working microphones. There is no way of being 100% certain that these microphones are fully working as they could have reduced performance in a way that a frequency response does not detect. Therefore, having a few new microphones tested alongside good as new microphones ensures that fully functional microphones are categorised. They will also have the same slight discolouration that the other microphones have. However, having some new microphones among those being categorized increases the validity of the experiment.

In many areas CIs are cutting edge technology but the experiments in this section outline this specific area where CIs are significantly behind Hearing Aids (HAs). In audiology clinics there is common to have devices that can objectify measure the functionality of HA microphones. However, CIs clinics lack this functionality and the experiments in this section has shown the inaccuracy of current tests for determining how well CIs microphones are working.

4.4 Chapter summary

The experiments in this chapter were designed to answer the first research question from Section 1.2 that is "*How accurate and repeatable are the current methods for identifying CI microphone failures?*"

The first experiment in Section 4.1 conducted a survey on CI clinician survey from the University of Southampton AIS. This showed that while CI clinicians have a cautious approach to replacing microphones with 60% of clinicians reported replacing a microphone that they thought was working. If they cannot be sure that a microphone is working then they replace it just in case. Would microphones be replaced so often with such a precautionary attitude if the person who uses the CI needed to pay for the microphone? Due to the majority of the AIS patients being from the National Health Service there is no cost to them. However, if they were in a situation where the people who use CIs had to pay for each replacement would the same precautionary approach be taken?

The next experiment in Section 4.2 objectively measured the performance of microphones that had been reported as broken. 59% of the microphones were found to be within ± 3dB of known working microphones. This supports the findings of the first experiments that a significant number of microphones are replaced just in case.

How accurate subjective listening checks are at categorising CI microphones is evaluated in Section 4.3. Both the control group and the CI clinicians were able to correctly identify the "*Not working*" with near 100% accuracy but the ability to differentiate between the "*Partially working*" and "*Working*" microphones was not so clear cut. During testing, CI clinicians marked the "*Working*" microphones as "*Partially working*" 35% of time which is significantly higher than the 20% that the other ten participants did. This supports the findings of the other experiments in the chapter, that clinicians take a cautious approach to replacing microphones; if they cannot be certain that the microphone is fully working then they replace it.

Section 4.3 showed that the four clinicians tested had a 40% chance of categorising a "*Working*" microphone as either "*Partially working*" or "*Not working*". This is lower than the 59% of microphones from the AIS that were reported as broken yet found to be fully working. However, this difference is likely a result of clinicians taking a precautionary approach to microphone replacement. If there is any sign of a microphone not working when talking with a person who uses a CI then they will likely replace the CIs microphone just in case.

Section 4.3 also showed that if a clinician is subjectivity evaluating microphone performance then there is a 16.7% of the microphones being categorised as "*Working*". As shown in Figure 4.8 the "*Partially working*" microphones had extreme reductions in performance of more than 15dB; the value of 16.7% is likely higher when less extreme frequency losses are considered. What are the implications to people who use CIs of using a partially working microphone?

Chapter 5

Implications of failure

As the previous chapter has shown, the methods currently being used in Cochlear Implant (CI) clinics to assess how well CI microphones are functioning are not perfect and can result in microphones that are fully working being marked as not working or microphones that are partially being distributed to people who use CIs. This chapter looks at the implications of both outcomes on the CI centres and those people that use CIs.

5.1 Rejected working microphones (False positives)

These were the microphones that were fully working yet had been reported as broken. The previous chapter has shown that this happens a significant percentage of the time. As the T-mics that were used for the experiments each cost over £140 to replace, and with 60% of the reportedly broken microphones being fully working[1], there is significant expenditure to the CI department that could potentially be avoided.

False positives likely cause significant hassle for people who use CIs and/or their parents/carers. If a microphone is suspected to not be working, then people may need to take time off work to travel to their implant centre to get the microphone replaced. If they have the dexterity to replace their own T-mic they would need to contact their implant centre and get a replacement posted to them, as well as posting their current microphone back. While being less hassle than a dedicated trip to the implant centre, it still takes time for the post, along with the effort and inconvenience and cost of the process. There is also the option of just the processor being posted to the CI centre where they can replace the microphone and post it back. However, this leaves the CI user without a CI processor for a few days, which is not ideal.

[1] Section 4.2

5.2 Unidentified defect (False negatives)

The false negatives were the microphones that were thought to be working but were not and therefore continued to be used by a person who uses a CI. The severity of the impact on a person who uses a CI will depend on the specific frequencies that are reduced by what amount. For example, a microphone that has no response at any frequency will have significantly more impact than a microphone that only has a minimal reduction in performance above 8kHz.

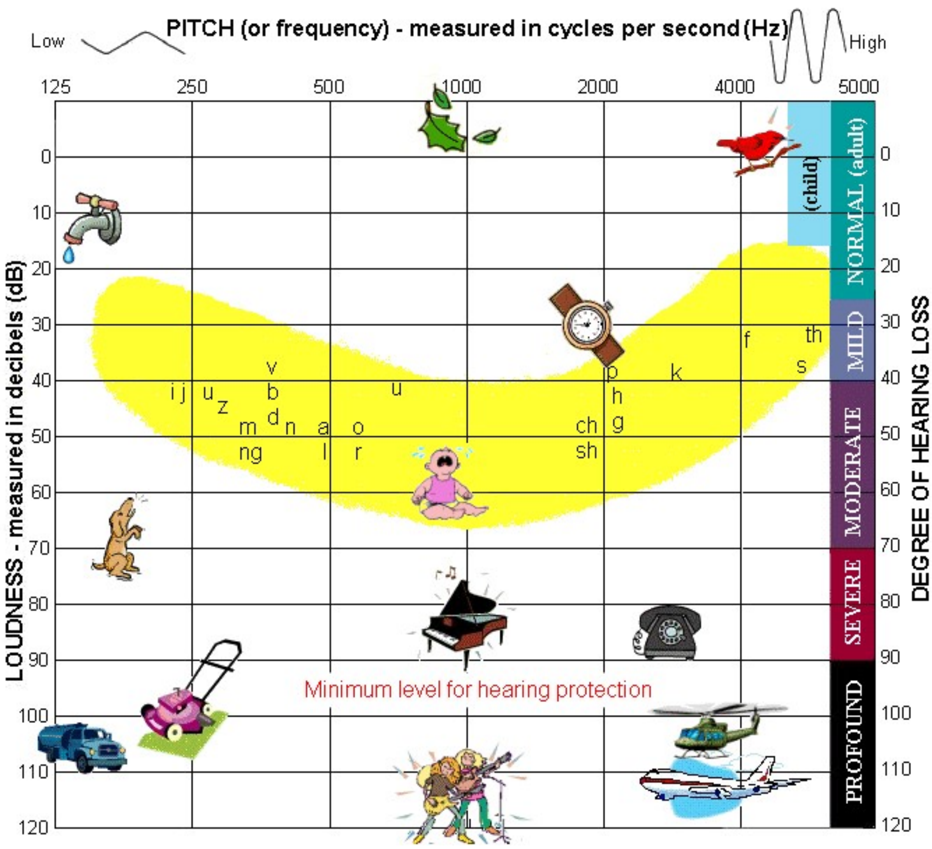


Figure 5.1: Diagram showing the frequency and loudness of a variety of speech sounds and environmental sounds, also known as the Speech banana
(from: ndassistive.org)

Different frequencies are known to have different levels of importance to understanding speech and a common diagram used for this in Audiology is the speech banana which is shown in Figure 5.1. This diagram shows where a variety of sounds such as "*b*", "*sh*" and "*th*" are in relation to both intensity (vertical axis[2]) and frequency (horizontal axis). The yellow shaded area on the diagram encompasses the area required for understanding speech. Several examples of environmental sounds such as birds, watch ticking and dogs barking are also shown on the digram. The speech banana lies

[2] With normal hearing biting at the top of the graph

between 130Hz and 5kHz. However, it is possible to understand speech with a narrower frequency range as evidenced by the fact that telephones still frequently transmit exclusively from 300Hz and 3.4kHz (Bauer et al. 2013).

5.2.1 Effect of simulated reductions in microphone performance on people who use cochlear implants

To determine the effect that different microphone failures have on people who use CIs, an experiment was conducted where degradations in microphone performance were simulated and the effect on perceived speech was identified.

5.2.1.1 Methodology

During this experiment two different speech tests were used. At the start of each experiment each participant had their hearing screened using the Digit Triplet Test (DTT) to ensure that all of the participants hearing was within normal limits. The DTT is well documented and has been shown to work all around the world including Australia (Myles 2017), New Zealand (Greville 1984, Purdy et al. 2000), South Africa (Wilson et al. 1998) and the UK (Westhorp 2009).

To work out the affect of reductions in microphone performance the Arthor Boothroyd speech test was used next which works by presenting lists of ten words in Background Noise (BGN), each word having three phonemes. For each phoneme a participant could repeat correctly they scored a point, meaning that each list of ten words was scored out of 30 points total. For example if the word was "*cat*" and the participant said the word "*Mat*", as the "*a*" and "*t*" was correct they would have scored two points for that word. Ten word lists were used for this experiment these are listed below.

List one: ship, rug, fan, cheek, haze, dice, both, well, jot, move.
List two: fish, duck, path, cheese, rice, hive, bone, wedge, log, tomb.
List three: thug, witch, teak, wrap, vice, jail, gen, shows, food, bomb.
List four: fun, will, vat, shape, wreath, hide, guess, comb, choose, job.
List five: fib, thatch, sum, heel, wide, rake, goes, shop, vet, june.
List six: fill, catch, thumb, heap, wise, rave, got, shown, bed, juice.
List seven: badge, hutch, kill, thighs, wave, reap, roam, goose, not, shed.
List eight: bath, hub, dig, five, wave, reach, joke, noose, put, shell.
List nine: hush, gas, thin, fake, chime, weave, jet, rob, dope, loose.
List ten: jug, latch, wick, faith, sign, beep, herm, rod, vote, shoes.

Five different reductions in microphone were simulated which are shown in Table 5.1.

The first filter serves as a reference setting with the words' frequency response limited to the same frequency response that a T-mic has (Section 6.1). The second filter's frequency response was limited to 100Hz to 3.4kHz to simulate the reduced frequency range of telephones. The third filter removed frequencies below 1kHz as a number of the reportedly broken CI microphones from Section 4.2 presented frequency responses like this. Figure 5.2 is a graph that shows visual representations of the 5 filters; the graph shows some difference between the filters at 0dB however, this variation was added to the graph to better visually differentiate between the filters. A number of the microphones from the experiment in Section 4.2 had deteriorating high frequency performance so filters four and five remove frequencies above 5 and 8kHz respectively.

Table 5.1: Acoustic filter settings for phoneme recognition testing

Number	Signal to Noise Ratio (SNR)	High pass filter (Hz)	Low pass filter (kHz)
1	-6 -12	100	9.8
2	-6 -12	1k	9.8
3	-6 -12	100	3.4
4	-6 -12	100	5
5	-6 -12	100	8

A background noise was played at the same time as the words and testing was done at two different SNRs in order to ensure the results were able to show the differences in performance that the simulated microphone failures could induce. If a SNR was used that produced a baseline response[3] of 15% for example, this would show increases in performance, but it would not show if any of the filters decreased performance. Testing at two SNRs mitigates this issue as a more difficult SNR with a lower baseline has a considerable range to show increases in performance and an easier SNR can be used to show reductions in performance. These exact SNRs were worked out by reviewing previous research and some initial trials of just the reference lists at both SNRs (Boothroyd & Nittrouer 1988). Another option would be to do a test before that varies each participants SNR to find a level where they get approximately 50% correct; however, this would increase test duration and mean testing each participant at different SNRs.

[3] from filter one

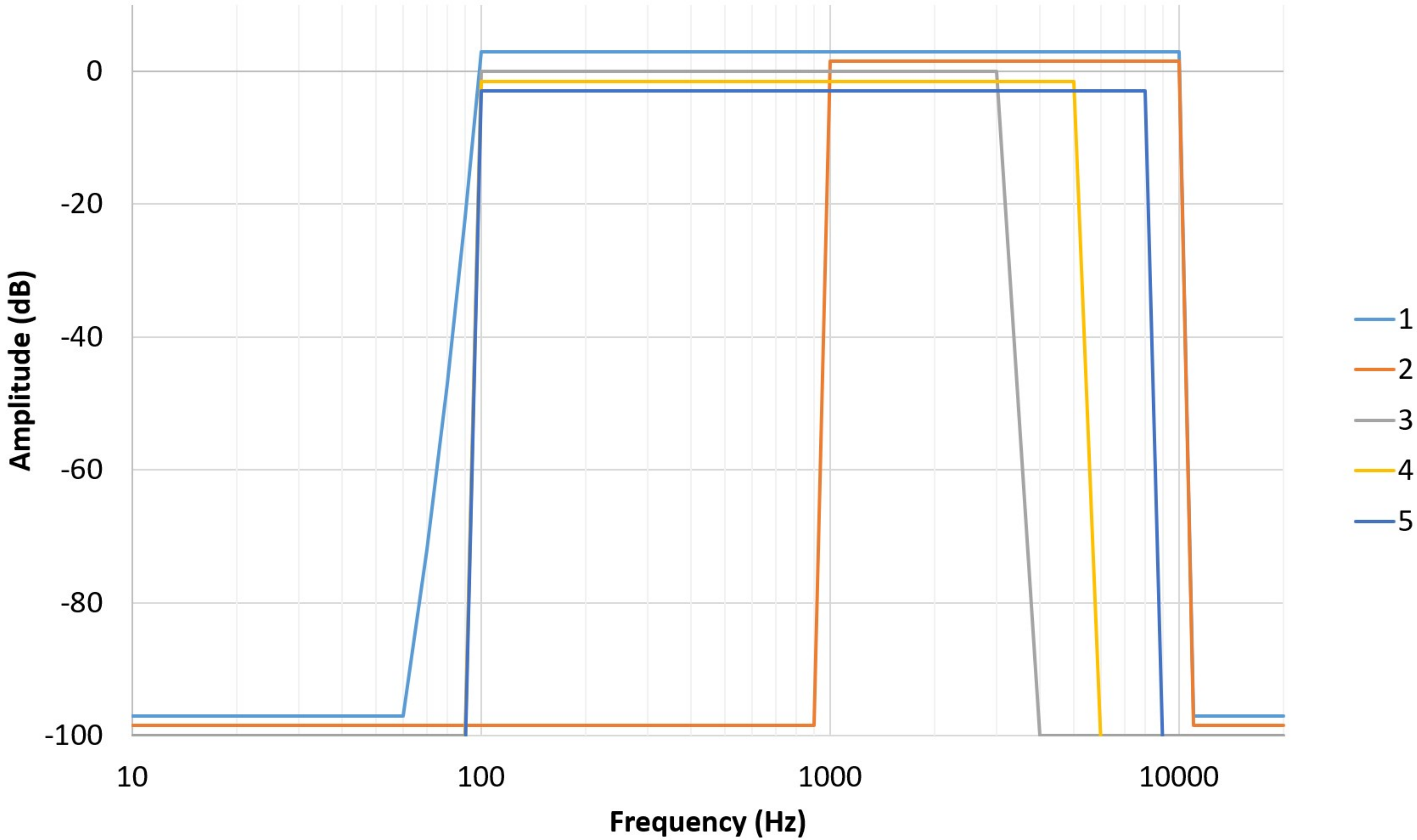


Figure 5.2: Visual representation of the five simulated reductions in microphone performance Variation at 0dB added to aid in visual differentiation between filters.

Upon hearing a word each participant was given the option to either say the word out loud so the tester would enter the word into the computer, or the participant typed the word in themselves. All of the participants chose to type the words themselves. The complete list of data that was recorded is listed below with examples in brackets. The system did not automatically mark how many phonemes were correct but after the testing the presented word and response word were compared and marked.

- Overall order that the words were presented in. (23)
- The word list that was being presented. (List 7)
- Which filter was applied to the word list. (Reference)
- SNR. (-6)
- The word that was being presented. (Cat)
- The stimulus word position in word list. (3)
- Participants response. (Mat)
- points (2) [This was done manually after testing]

Testing was done from a laptop using Audio-Technica® ATH-M40x headphones which had been calibrated to output the speech stimuli at 60dBA SPL[4] with the background noise playing between -6 and -12 dB[5] depending on the SNR that was being tested. A python script randomised the order that the word lists were presented in addition to hiding this order from both the tester and the participant, making this a double blind.

Initial trials were run with a limited number of participants to check that the reference settings for both SNRs was resulting in a phoneme accuracy between 10% and 90% so the filters changes could be detected. These initial trials showed that the SNRs had the required baseline accuracy. Also during these trials the other filter settings were tested as well, which gave some strange results. They found no different at all between all of the filters: research showed that the filters were not being applied accurately or to the extent that was desired. In order to mitigate this issue, each of the filter settings were applied directly to the word list, resulting in consistent and highly accurate filters. While this did mean that during this experiment each filter was only tested on one list of words, half way through the testing which filter was applied to which word list was re-randomised to minimize any issue this may have had on the results.

[4]Peak

[5]Meaning that the background noise was either 6 or 12 dB louder than the words

The number of participants required was calculated before the experiment by assuming a mean difference of 8 and a standard deviation of 10. A paired t-test with alpha=0.05 and 95% power would require a sample size[6] of 20. Ethics approval was obtained from The University of Southampton Ethics and Research Governance (ERGO 46903).

5.2.1.2 Results

Figure 5.3 shows the phoneme accuracy ± 1 standard error for each of the five filters at both of the testing SNRs. The Y axis is the phoneme accuracy for the word list and the X axis being the different filters. The two tested SNRs are represented by different shades. The reference filter (1) had an average accuracy of 56% and 25% for SNR -6 and -12 respectively. The second filter only presented sounds above 1kHz and with average accuracies of 60% and 27% for SNR -6 and -12 respectively, this is no significant change from the reference filter. As Figure 5.1 shows, about half of the speech sounds are above 1kHz and it is possible that the brain is managing to work out the sounds it cannot hear by guessing based on context.

The third filter simulates the same frequency response that phones have used for decades 100Hz to 3.4kHz. With average accuracies of 66% and 35% for SNR -6 and -12 respectively, this is better than the reference filter but not by a statistically significant margin. This filter not worsening phoneme recognition is not surprising as the fact that people have been understanding speech on phone calls for years is evidence that the brain can understand speech even with the narrower presented frequency range (Jax & Vary 2000). As the background noise was put through the same filter it is probable that the slightly increased phoneme accuracy can be attributed to the frequencies above 3.4kHz having little speech information in them but abundance of background noise in them. Furthermore, as the change in average phoneme accuracy is not statistically significant it is possible that the difference is caused by the test re-test variability.

Filters four and five presented sounds below 5kHz and 8kHz and, as with the third filter, there was some deviation from the reference results but none of these deviations were statistically significant. As shown in Figure 5.1, there are negligible speech sounds above 5kHz, so in hindsight, these two filters not having a statistically significant effect on the average phoneme recognition is not surprising.

[6]Sample calculator used (Kane 2019)

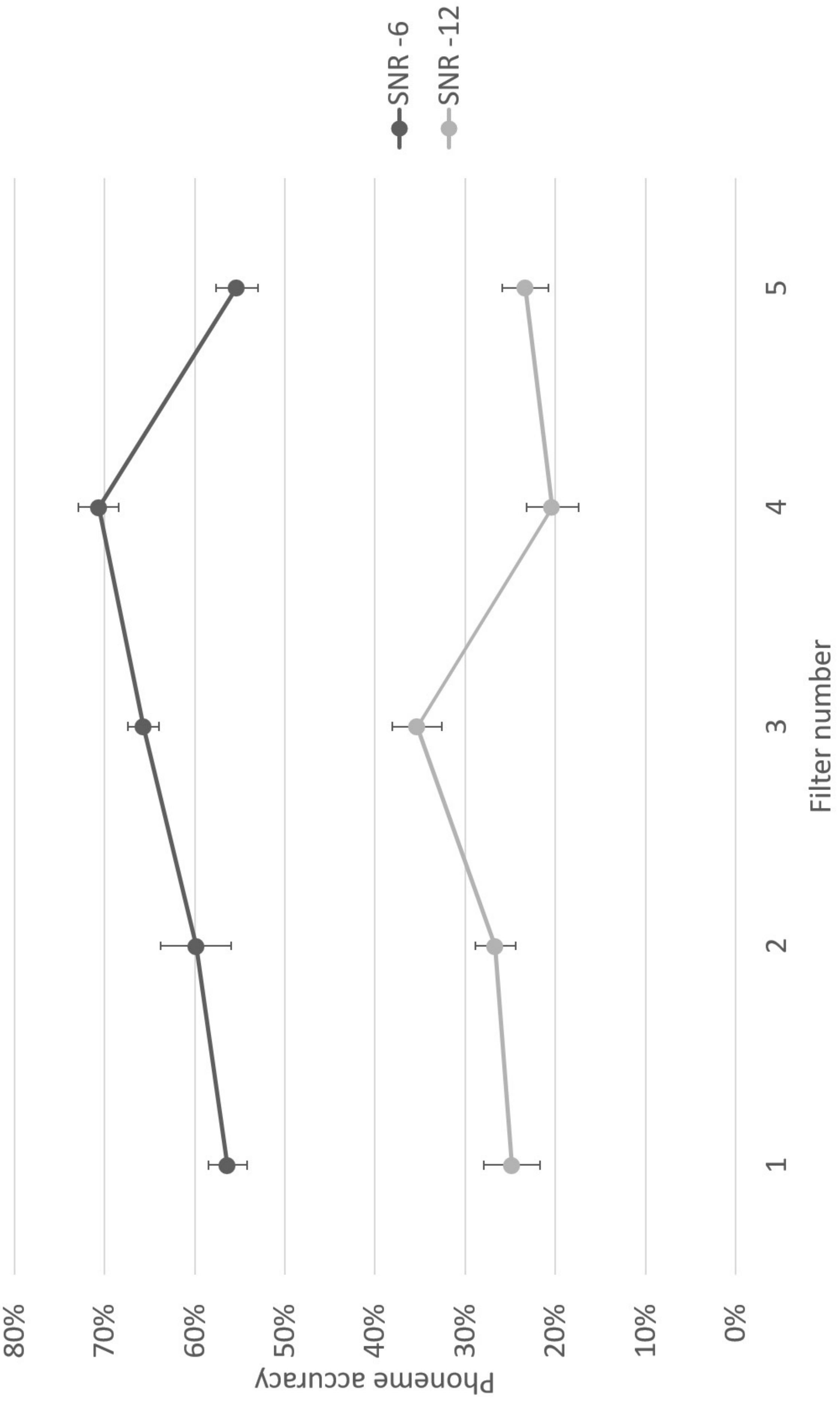


Figure 5.3: Graph showing the average phoneme recognition for the different filter settings as described in Table 5.1 $\pm$ 1 standard error

5.3 Discussion

An alternative type of speech test that could have been used is a sentence based speech test. Speech tests such as the Bamford-Kowal-Bench (BKB) speech test (Bench et al. 1979) or the Hearing in Noise Test (HINT) (Nilsson et al. 1994) are similar to the word test but instead of presenting individual words, they use entire sentences as stimuli. It has been shown that having context for speech stimuli can affect the results, as the sentence structure can be used to work out hard to hear words (Duffy & Giolas 1974, Giolas et al. 1970, Kalikow et al. 1977). Using a word based speech test removes these contextual clues that can affect the results (Boothroyd 1968, Boothroyd & Nittrouer 1988).

5.4 Summary

Research question two states: "*What impact do partially failed CI microphones have on speech perception?*". This was investigated through an experiment that presented a speech stimulus in BGN that had been run through different filters. These filters were chosen to test the effect that reductions in CI microphone performance in different frequency ranges would have on speech recognition. This experiment showed that a reduction in microphone performance in certain frequency ranges does not always lead to reduced speech recognition. The results might have been different if a CI simulator was used on each of the participants, but the participants would not have been used to listening to speech through a CI simulator so this could have affected the results in unpredictable ways. This experiment also only looked at speech sounds and, as shown in Figure 5.1, there are a plethora of other environmental sounds that could affect the quality of lives of people who use CIs.

Chapter 6

Reference testing

This chapter takes a detailed look at the Advanced Bionics listening check by doing a series of objective acoustic tests, firstly keeping the Cochlear Implant (CI) processor in test settings while frequency responses were recorded in an anechoic chamber. Next, remaining in the anechoic chamber, a series of settings on the CI processors were changed to see which affected the listening checks audio output. Then the repeatability of frequency response recorded though the Advanced Bionics listening check was assessed. Finally the delay between putting an input into the CI processor and getting an output from the listening check was measured. These experiments were done to attain baseline information on the acoustic performance of the CI microphones that would inform other experiments and to answer the third research question: "*What processor settings affect the audio output of the Listening Check?*"

6.1 Anechoic chamber testing

Baseline acoustic performance of Advanced Bionics Naída Q70 CI processor when connected to a listening check was ascertained. This was to identify the acoustic characteristics which will be useful to know for future experiments. CI microphones do have a limited frequency response and knowing these accurately will also help enable future experiments to focus on the relevant frequency range.

6.1.1 Methodology

Advanced Bionics Naída Q70 processors were used for testing in conjunction with an Advanced Bionics Listening Check which provides a mono 3.5mm audio output. Three distinct settings of the CI processors were tested, which are shown in Table 6.1: test settings consisted of all advanced features off[1] and linear dynamic processing.

[1] clear voice, wind block, sound reflex and echo block all off

Programmes two and three used these test settings with the only exception that programme two had clear voice set of medium, which is Advanced Bionics noise reduction system.

Table 6.1: CI programme list used for anechoic testing.

Programme number	Microphone	Notes
1	T-mic	Test settings
2	T-mic	Clear voice on
3	Processor microphone	Test settings

The anechoic methodology was used as described in Section 3.1.1. A pure tone sweep was played using a reference speaker from 20Hz to 20kHz[2] with the response first being recorded using a reference microphone in an anechoic chamber. The reference microphone was then replaced by the first CI processor recording frequency responses in all three programmes.

The reference frequency response (B) was deducted from the frequency response for each programme (A) and converted to a dB scale using Equation 6.1. This is because the reference microphone will record the stimulus that had been affected by the speaker and computers audio interface; taking this response from the CI processor response negated these effects.

$$(A - B) - 20\log_{10}(20 \times 10^{-6}) \tag{6.1}$$

6.1.2 Results

Figure 6.1 shows the frequency responses for the three CI processor programmes. Below 50Hz and above 9.5kHz the microphone failed to get any response, which is approximately the stated input frequency range of a CI, and therefore not surprising[3]. The three tested programmes had similar responses below 2kHz but above this frequency there were differences of up to 11dB between the frequency responses.

6.1.3 Summary

Testing using a reference microphone was an efficient way of measuring an accurate frequency response but the accuracy of this measurement system is dependent on getting the microphones of the CI precisely where the reference microphones diaphragm is. Using a simultaneous referencing system where the CI processor and reference microphone

[2] 96 frequencies per octave at 48,000 per second
[3] Advanced Bionics Naída Q70 Technical specifications

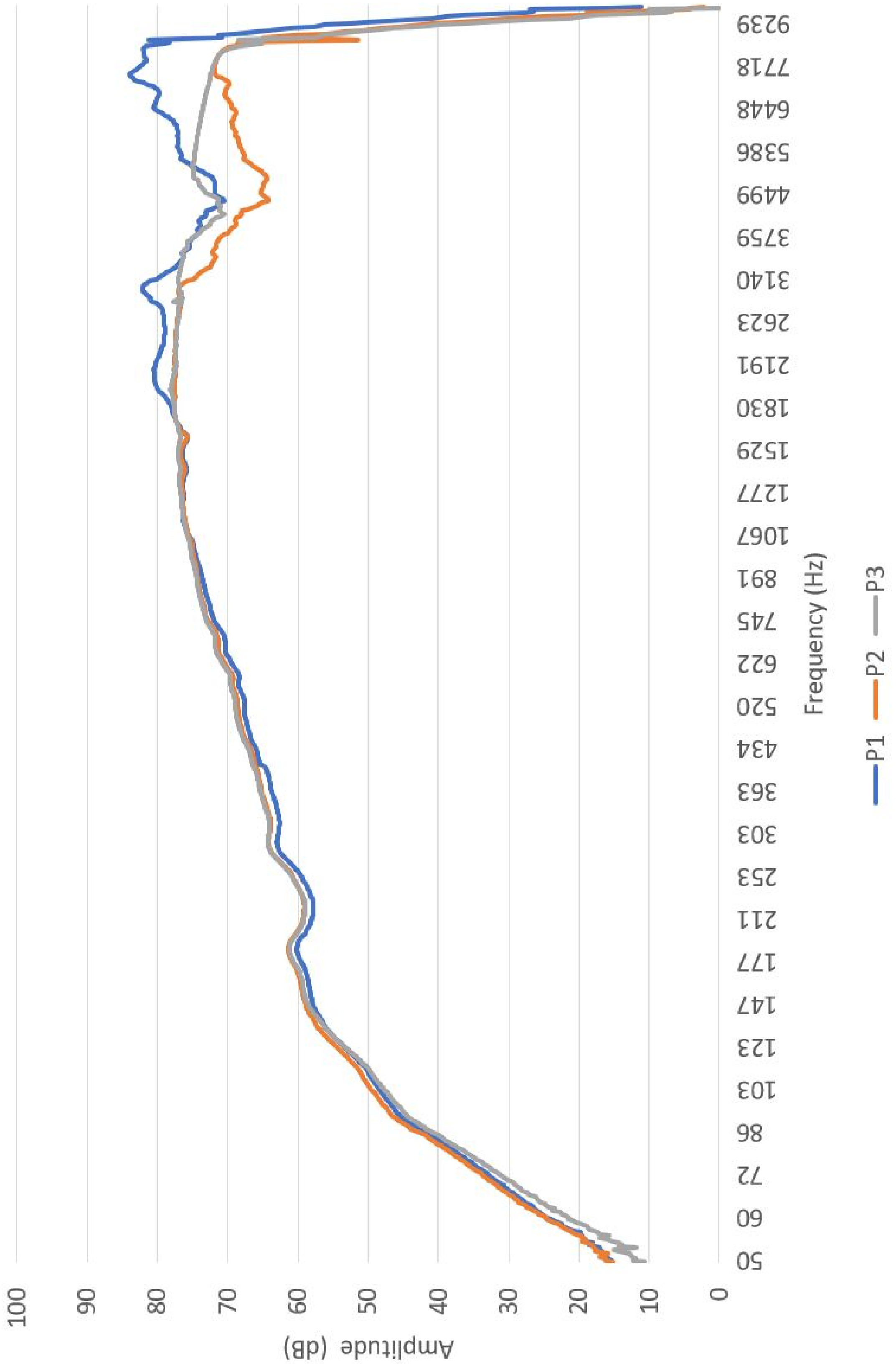


Figure 6.1: Frequency responses of a CI processor with three different programmes.

are placed as close as possible during testing, recording both response simultaneously, might have marginally increased test accuracy.

All the measurements were taken from a single CI processor which is a single factor that could affect all the results. Ideally all the experiments would have been repeated on a second processor which would have both tested repeatability and eliminated any effect that a single CI processor could have on the results. Subsequent experiments will be repeated on at least one other CI processor which should increase the comparability of CI processors that are commonly used.

The Advanced Bionics listening check adaptor is affected by the current programmes microphone selection; this is known and stated in the product documentation. However, it can be concluded from the testing done that the Advanced Bionics listening check adaptor is also affected by some of the CIs advanced feature settings. How much these affect frequency responses measured through the Listening Check would require further investigation.

6.2 Effect of cochlear implant settings on listening check output

The listening check is a device that provides a headphone output to an Advanced Bionic CI. However, each person who uses a CI has settings that are specific to them, and it is not known what effect these settings have on the listening checks output. This is important to ascertain as it will significantly affect the future usability and real world use cases of the listening check. If none of the settings affect the listening checks output then results should be more consistent between people who use CIs. Meaning that regardless of what settings a person says on their CI processor the Listening Check output will be consistent. On the other hand, if various settings affect the listening checks output then this will limit the real world usability of the device for objective testing. Settings affecting the output of the Listening Check would mean that there will be limited comparability between CI processors with differing settings[4].

With each of the CI processors having multiple microphones (Shown in Figure 1.2) it is known that the Listening Check will use the current CI processors programmes microphone ratio. In other words, if a CI has two programmes; one with 100% T-mic and two with 100% processor microphone, then if the CI processor is on programme two through a Listening Check you will hear only the processor microphone.

[4]Including a CI processor on different programmes.

There are a variety of settings that change for each individual person who uses each CI. The following paragraphs will explain each of these filters before explaining an experiment that will determine which of these settings affect the Listening Checks headphone output.

6.2.1 Map Levels

The first of these settings is the map[5] which defines the dynamic range, which is significantly lower than that of someone with normal hearing (Waltzman et al. 2000, pg. 50). To compensate for this multiple dynamic processors (such as compressors) are used to reduce the dynamic range of the signal. The settings of the dynamic processors will be affected by the map as the map forms the required dynamic range for the output.

6.2.2 Adaptive noise reduction

Hearing speech in Background Noise (BGN) is a problem for the hearing impaired and there are various adaptive systems for reducing BGN. Modulation detection and synchrony detection are two of the methods and these are briefly explained: however, each CI company keep the exact methods used confidential.

Modulation detection works on the assumption that the BGN has a lower dynamic range than the speech stimulus. Therefore, the microphones signal is split into a number of discrete frequency bins, with a lower dynamic range being reduced in amplitude (Dillon 2008). An alternative is synchrony detection which identifies specific frequency ranges where there is speech present, which can help focus and inform other noise reduction systems. Synchrony detection relies on the fact that natural speech formants have an inherent amount of discrete frequency modulation in them[6] which is not present in background noise; this modulation is normally synchronous between multiple speech formants (Dillon 2008); these frequencies are then emphasised. A study by Pittman (2011) involving 80 children[7] found that digital noise reduction systems had no significant impact on the children's ability to participate in auditory tasks. However, Nordrum et al. (2006) found that 50% of participants[8] had improved performance in speech in noise tests. Zakis et al. (2009)[9] also found that while no clinically significant difference was found between having noise reduction on or off, 90% of the participants preferred having the noise reduction on. Of these studies only Nordrum et al. (2006)

[5] AKA. M-levels
[6] usually between 3 and 6Hz
[7] between 8 and 12 years old, 50 normal hearing and 30 with a hearing loss
[8] 16 experienced hearing aid users
[9] 10 hearing impaired participants

found a benefit to participants and when compared to the studies by Zakis et al. (2009) and Pittman (2011) they have 5.6 times as many participants involved.

Advanced bionics noise reduction system is called Clear Voice^{TM} with their website[10] saying "*70% improvement in speech understanding in noise (Buechner 2013). Studies have shown up to 6.5dB SNR benefit when the Phonak UltraZoom feature is used together with the Clear Voice by unilateral recipients.*" Research did not reveal a conference recording or journal article, but a paper by Hehrmann et al. (2012) investigated the effect of noise reduction algorithms on 12 unilateral people with CIs and speech ineligibility in noise. They found that the people with CIs speech reception threshold was improved by Advanced Bionics microphone adaptive directionality system (adaptive beam former) by 5.2dB and Clear Voice improved it by a further 0.9dB; providing a total improvement of 6.2dB compared to adaptive beam former and clear voice off. Hehrmann et al. (2012) does account for the majority of the quoted 6.5dB improvement and without the original conference paper it is not possible to know where the extra 0.4dB improvement comes from. How the 6.5dB reduction in speech reception threshold equates to a 70% improvement in speech understanding in noise is unknown. There is conflicting evidence with Dingemanse & Goedegebure (2015) finding that Clear Voice had no significant impact on speech intelligibility for people who use CIs. This was followed up by a further study investigating if altering m-levels altered the effectiveness of Clear Voice (Dingemanse & Goedegebure 2018). Clear Voice was found to give minimal (but statistically significant) improvement to intelligibility and increasing the m-levels did further increase the benefit of Clear Voice. Needing to increase the m-levels to get full benefit from clear voice was not surprising as it will be turning down portions of the frequency spectrum; increasing the m-levels would compensate for Clear Voice lowering some frequency bands. Considering studies have only found a slight improvement from Clear Voice, it is possible that the studies that did not find benefit lacked the required accuracy to pick up the improvement. While exact information about noise reduction methodologies are proprietary information, the inference can be made that adaptive noise reduction systems would be less likely to filter out a multi talker babble than a white noise stimuli. Verifying this would be complex but the experiments could be designed to use a multi talker babble as a background noise reducing the risk of adaptive noise reduction systems influencing the results. When tested, Clear Voice was set to its highest setting on the principle that if this was shown to have a significant effect than more detailed experiments could be done at a later date to see the affect of lower Clear Voice settings.

[10] `https://advancedbionics.com/com/en/home/products/sound-processing/clearvoice.html` as of Thursday 4th October, 2018

6.2.3 Other advanced features

Manufacturers have various advanced features, which in the case of the Advanced Bionics Naída series are: Echo Block, Wind block and Sound Relax. These are all configurable on a programme by programme basis. Information on these advanced features is limited but what is known will be outlined here.

Echo Block is designed to remove early reflections from reverberant sound signals like speech in places of worship. It is unclear what effect this will have when aiming to test microphone functionality, because if using a broadband stimulus that averages over time (like white noise) then any changes this makes should be averaged out: however, other stimuli like pure tone sweeps may more be susceptible. When wind blows into a CI processors microphone it can be loud broadband noise and Wind Block is a feature that is designed to remove this sound. Broadband stimuli will likely be susceptible to the effects of this advanced feature. The Naída CI Q70 User Guide (Advanced Bionics 2015*a*) states that "*Sound relax is designed to soften sudden loud sounds, such as slamming doors or clanking dishes.*" Considering documentation on this feature is limited, the description sounds like a dynamic processing unit, likely made up of compressors and a limiter. When trying to get a consistent level output from a microphone, as would be required for assessing microphone functionality, having dynamic processors affecting the output level in unpredictable ways may be problematic.

6.2.4 Methodology

The anechoic methodology[11] was used for this experiment using the same pure tone sweep as used for the prior experiment. The first frequency responses were recorded with the CI processor in test settings with all advanced features disabled. Each subsequent frequency response only had one advanced feature enabled.

Table 6.2: Cochlear implant programme list for testing map levels

Programme	Map level
1	Low
2	Medium
3	High

6.2.5 Results

6.2.5.1 Map Levels

Three different maps were tested (shown in Table 6.2), all set to minimum, all set to maximum and a half way setting. For each of the settings tested, white noise was

[11] Anechoic methodology described in Section 3.1.1

played for 10 minutes recording frequency responses every 50ms and averaging the results together. All this was done in a room with BGN less than 40dBA during the test. The map levels are a 0 to 500 scale but their specific relationship to the output levels of the CIs electrode array is not made clear by the manufacturer. Like the other systems being analysed, the exact effect of the tested system on the output will be treated as a black box where the effect of the box is being measured rather than trying to determine the exact functionality of the box.

Table 6.3: Map deviation from the reference response in dB

Map level	Average	Peak
Low	-0.44	-2.12
Medium	0.46	2.13
High	0.29	1.6

The frequency responses gathered were compared to the average response from the repeatability testing (Section 6.4); differences between each of the programmes and the reference frequency response is shown in Figure 6.2. The average and maximum variations from the average for each of the map levels are shown in Table 6.3. While there are peaks above the ±2dB tolerance, these are only slightly over at 2.12dB and 2.13dB. Considering the average deviation from the reference responses are all within 0.5dB of the reference response, it can be concluded that the map levels do not affect the audio output of the listening check.

6.2.5.2 Adaptive noise reduction

In order to test the effect of active noise reduction, two programmes were used which were identical except one programme had no advanced features with the second only having clear voice on; both programmes used exclusively the T-Mic. These programmes are shown in Table 6.4. A white noise stimulus was played for a total of 10 minutes taking averages every 50ms, with the volume set to 16% as with the previous tests.

Table 6.4: Cochlear implant programme list for testing adaptive nose reduction

Programme	Features
1	None
2	Clear Voice only

Figure 6.3 shows the difference between the reference response, a programme in test settings and an identical programme but with Clear Voice enabled. Having a maximum deviation from the average of 1.59dB and an average deviation of 0.03dB is not

statistically significant.

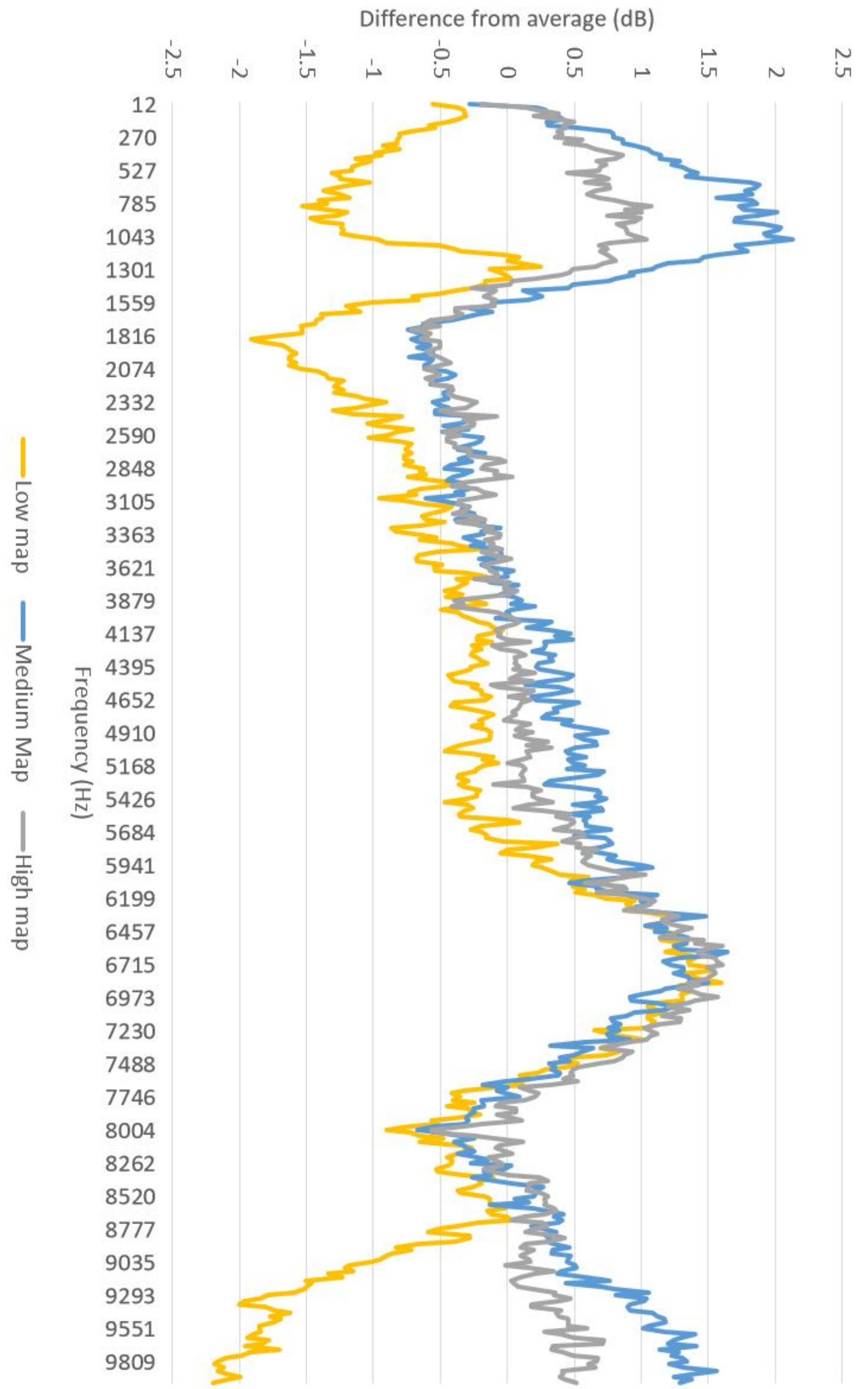


Figure 6.2: Graph showing the effect of map levels on audio adaptor output

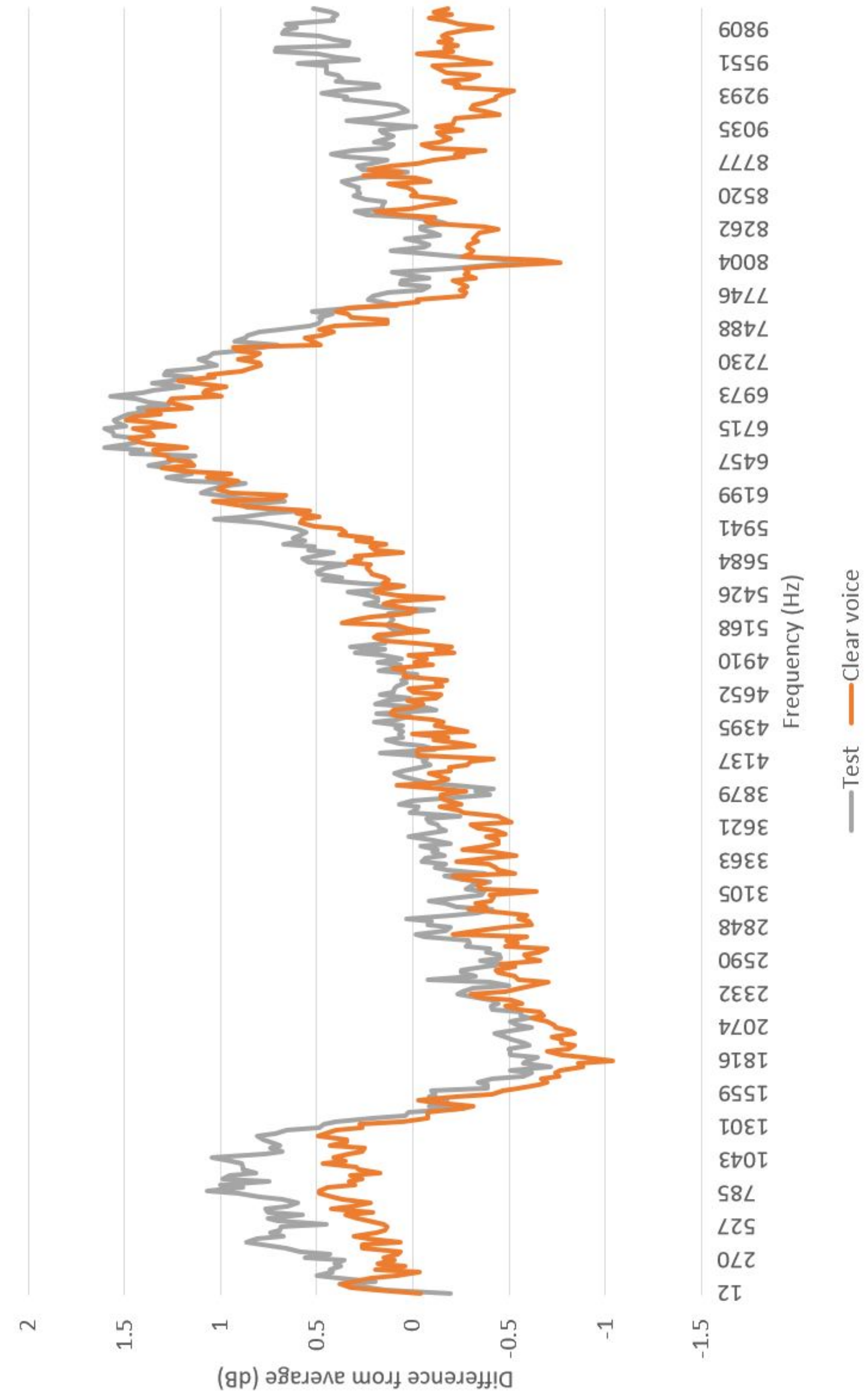


Figure 6.3: Graph showing the effect of the clear voice setting on audio adaptor output

6.2.5.3 Other advanced features

There were three different advanced features that needed to be evaluated, which were evaluated by comparing the output of the three distinct programmes shown in Table 6.5. All the programmes have the same settings with the only exception being which advanced features are turned on.

Table 6.5: Cochlear implant programme list for testing advanced features

Programme	Features
1	Sound relax
2	Echo block
3	Wind block

The difference between the three programmes and the reference response are shown in Figure 6.4 in which a significant dip at 8kHz is visible. Due to the dip the results were repeated a total of four times, twice on each processor and the "*Repeatability*" column in Table 6.6 shows the average difference between the repeats across the frequency spectrum. The repeatability values being below 0.1dB show that the dip is consistent between processors and microphones.

Table 6.6: Advanced feature deviation from the reference response in dB

Advanced feature	Average	Peak	Repeatability
Sound relax	-1.58	-7.14	0.09
Echo block	-1.86	-7.48	-0.02
Wind block	-1.61	-7.28	0.09

6.2.6 Summary of effect of processor settings

In summary, the settings on each CI processor that are specific to each person who uses a CI, map levels, active noise reduction, sound relax, echo block and wind block, were all shown to have no statistically significant effect on the listening checks output, which is encouraging for future implementations into remote care systems. However, the microphone choice and balance still affects the output so this will need to be taken into account.

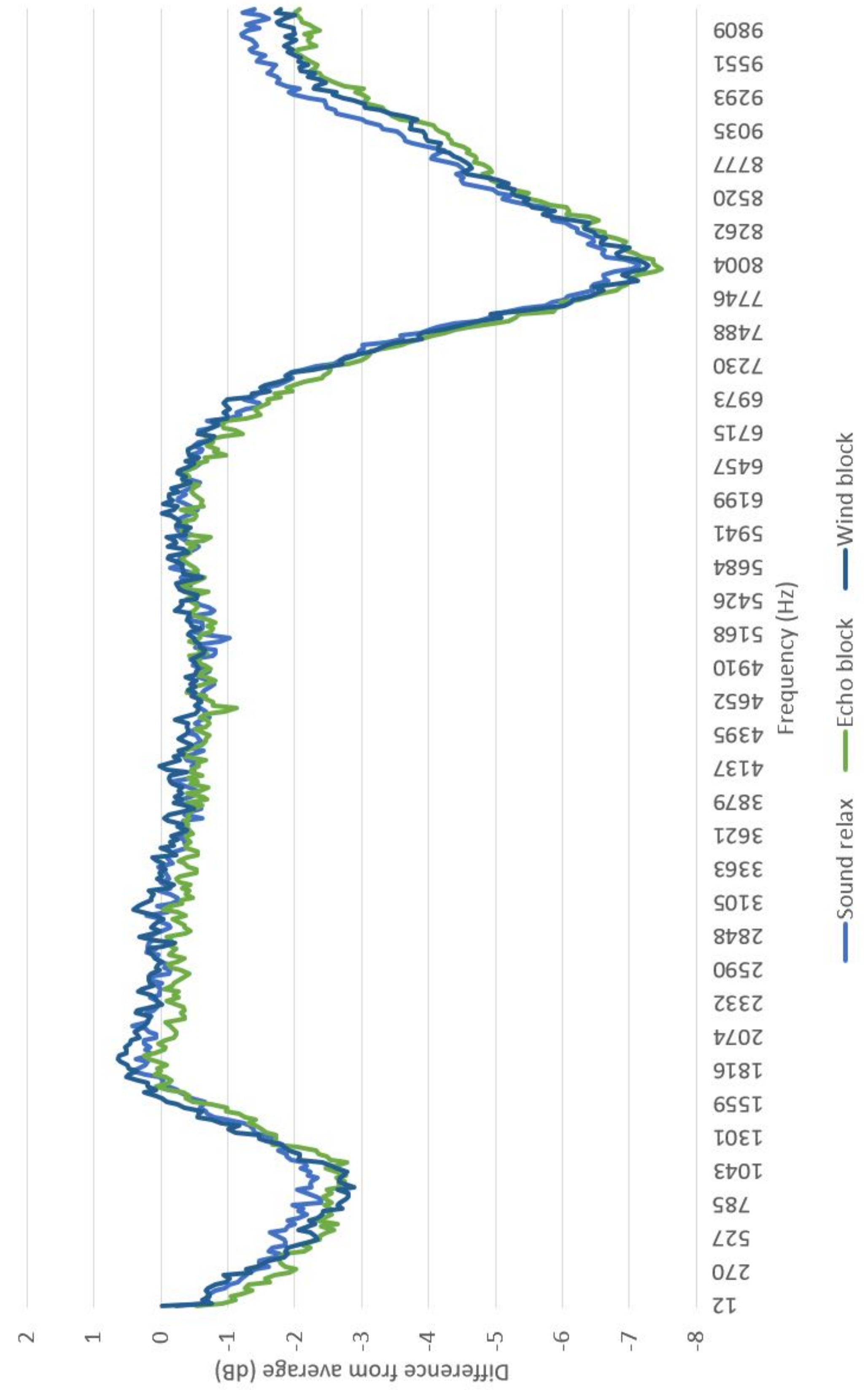


Figure 6.4: Graph showing the effect of advanced features on the audio adaptor output

6.3 Time delay

Audio systems have an inherent delay in them and when an Advanced Bionics CI processor is plugged into a Listening check device it is unknown what time delay this system has. Any time delay is important to know as this will affect how soon any device can start taking measurements. For example, if there is a 500ms delay for a sound signal to get through the CI processor and the listening check device, then results, could not be taken till after this time has expired.

6.3.1 Methodology

For this experiment the custom enclosure methodology was used[12]; a series of 55 pure tones were played from lowest to highest frequencies, with the complete list of frequencies in Appendix A. These frequencies were chosen to be distributed between the known frequency range of the CI processor T-mics as shown in Section 6.1. Each of the 55 pure tones were played for 500ms with frequency responses recorded every 25ms, then averaged over the stimulus duration.

Each frequency response was then analysed and put into one of three categories: previous, none and correct. **Previous** was when an auditory stimulus was being presented but the previous stimulus is still being outputted by the CI listening check[13]. **None** was when there was no clear peak in the frequency response. **Correct** was when a stimulus was being presented and the same stimulus was being detected. For example, if it took 300ms for a sound signal to make it through the CI processor and listening check, then frequency responses recorded 250ms into a 1207Hz stimulus will be detecting a 1090Hz stimuli[14].

CI processors have various advanced features in them that are designed to perform various functions, such as reducing background noise and dynamically altering the microphones directionality. Precisely how each of these advanced features work are industrial secrets so it is unknown how they interact with different auditory stimulus. A pure tone stimulus that was used in this test is not a sound that people who use CIs will encounter on a regular basis, so the advanced features may interact with this type of artificial test stimulus in unexpected ways. This would have also been a factor in the anechoic testing done in Section 6.1 but those used a continuous pure tone sweep rather than the discrete frequency system used here. As many of the advanced features as possible were disabled during this test but some cannot be disabled.

[12] described in Section 3.1.2
[13] Such as 1207Hz being played and 1090Hz being outputted
[14] Which is the previous stimulus frequency as listed in Appendix **??**

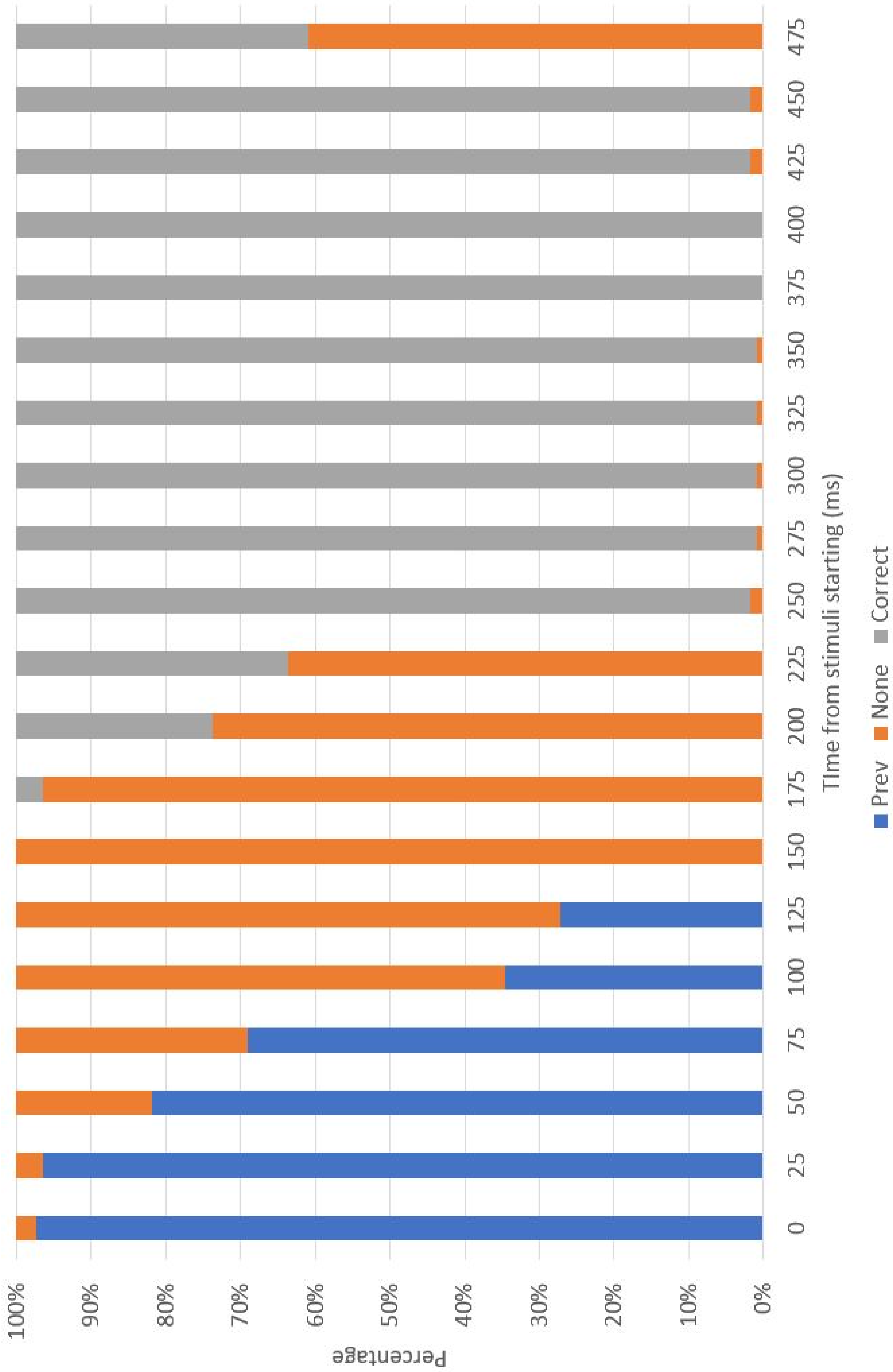


Figure 6.5: Graph showing the time taken for a stimulus to go though the Advanced Bionics Listening Check

6.3.2 Results and Summary

Figure 6.5 shows each of the 25ms increments for the duration that each stimulus was presented. Each increment is colour coded depending on what frequency is most prominent. If a 1kHz tone was being played and a 1kHz tone was easily discernible in the frequency response, then the colour would be grey. If there is no discernible single frequency peak then it is colour coded orange. Finally, if a tone is being played but the previous tone is still present on the frequency response and louder than the stimulus frequency, then this is coloured blue. Figure 6.5 shows the average from all 55 pure tone stimuli frequencies. This graph shows that it takes approximately 250ms for a stimulus to go through the CI processor and make it to the listening check output. Generally when people are using their CI a signal will take significantly less time to stimulate the CIs electrode array; however it taking this long for the listening check is not surprising as there is no urgent need to have such an immediate stimulation as speech requires.

An example frequency response is shown in Table 6.6 which is taken 400ms into 3054Hz stimulus frequency. On this the peak from the stimulus frequency is clearly visible but there is some pure tone artefacts around this frequency. There could be any number of possibilities for this but they are likely an artefact caused by part of the CI processors digital signal processing.

6.4 Repeatability testing

The custom enclosure methodology is used for a number of experiments in this project. This chapter investigates the repeatability of this methodology by preforming a number of frequency responses over approximately two months. A white noise stimulus was used with frequency responses taken every 50ms; the frequency responses were then averaged over a 30 second stimulus duration. Four different new known to be working T-Mics were used for this experiment. Between each recorded frequency response the equipment was packed down and set up again; a variety of different rooms were also used. This was done to get real world relevant results because if such a system was being used by people who used CIs, it would most likely be packed away between uses.

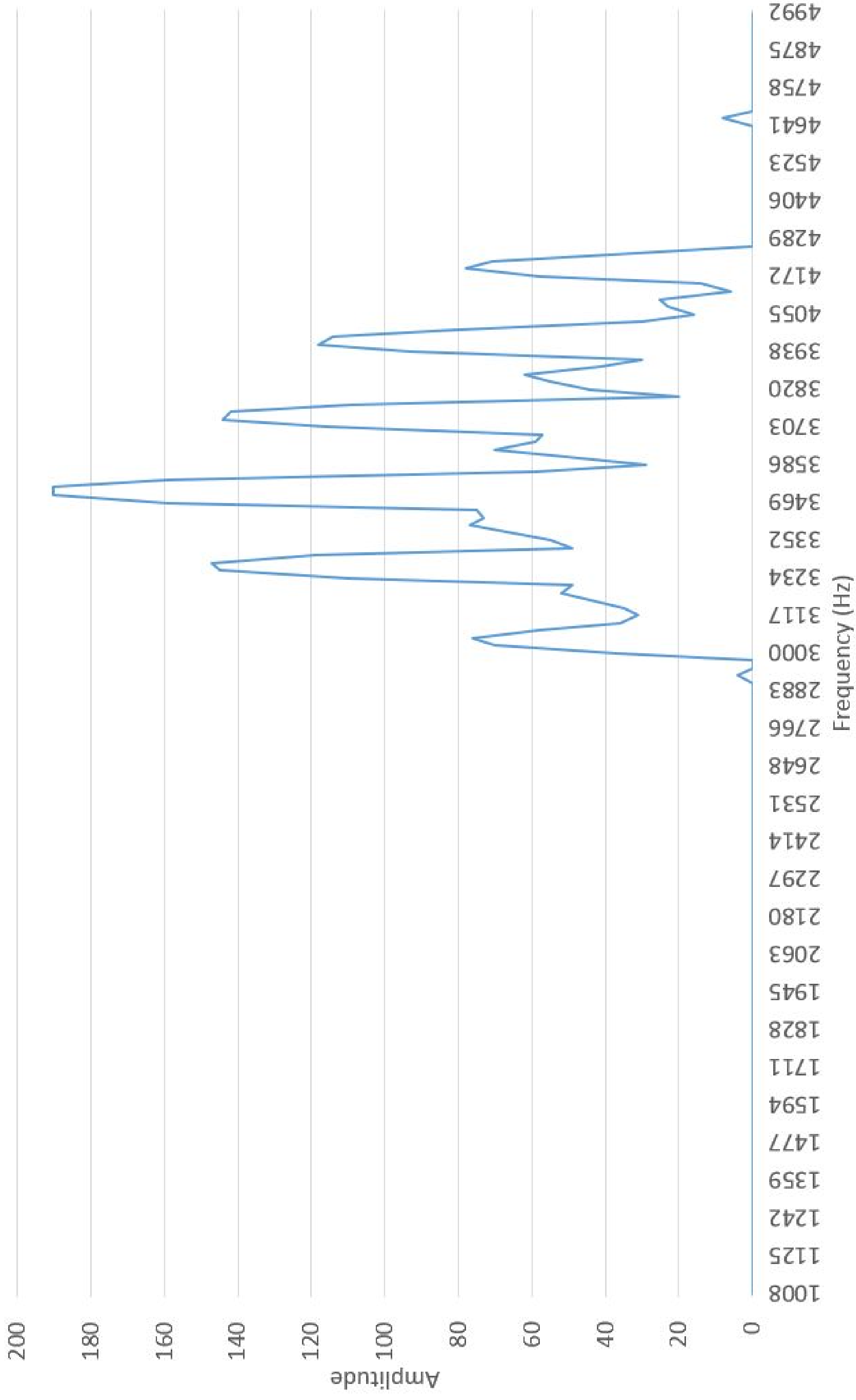


Figure 6.6: Graph showing the frequency response 400ms into 3504Hz

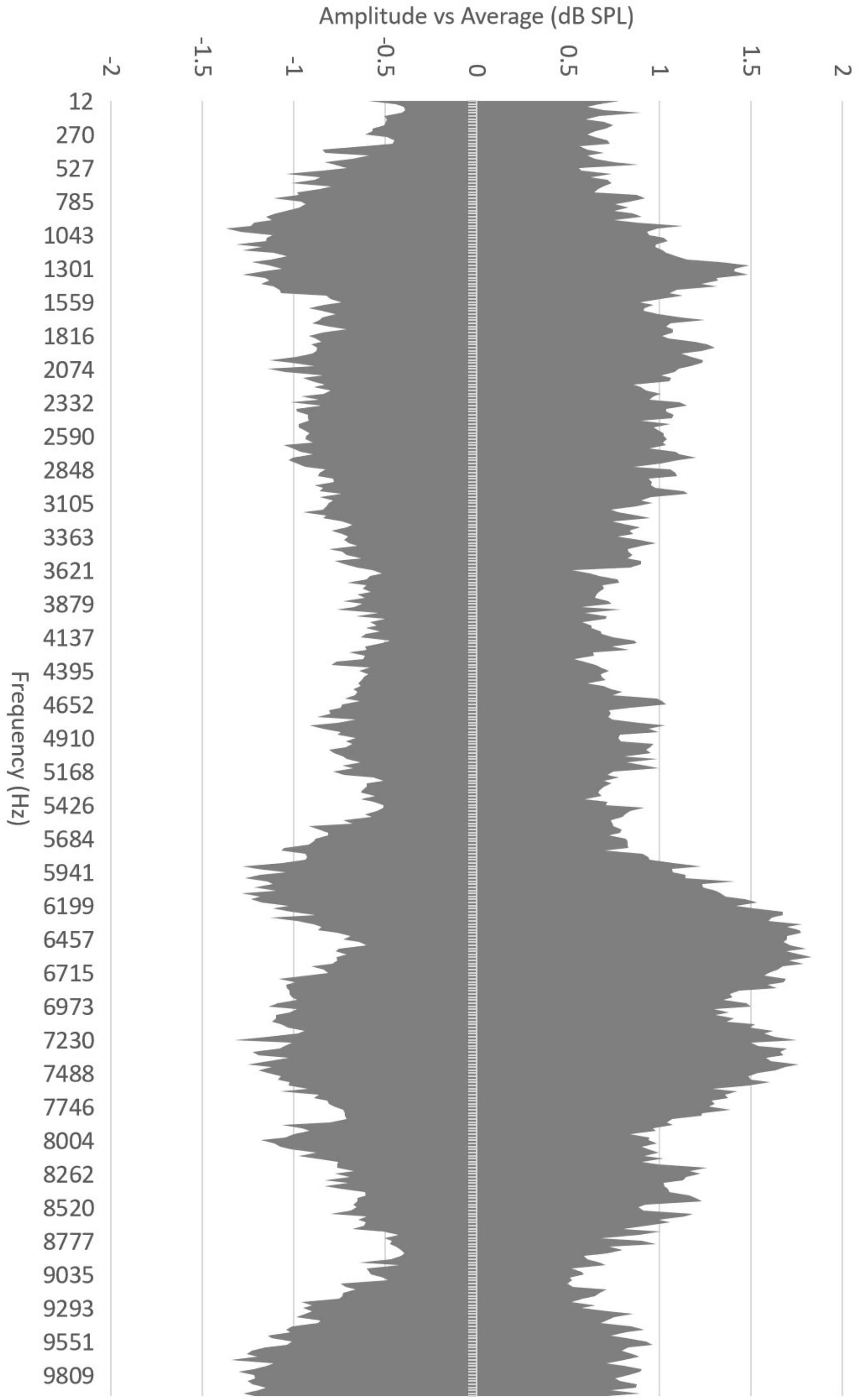


Figure 6.7: Microphone response maximum deviation from average

6.4.1 Results

At every tested frequency, the average of all 53 recordings was taken and Figure 6.7 shows the maximum positive and negative deviation from the average for all of the recordings. The difference between the highest deviation at any frequency deducted from the lowest deviation provides a maximum deviation of 3.13dB[15]. The highest and lowest value from each of the responses was deducted from each other providing a frequency specific range; the average was 1.6dB providing a range of ±0.8dB. A number of the frequency responses deviated from the average at 6.3kHz but this was consistent between tests and only up to 1.7dB from the average. In future tests the average between the responses could be used with a range of ±2dB which is sufficiently above the maximum measured range.

Section 3.1.2 explains the standing waves of the custom enclosure that was used for this experiment. The calculations were done using the maximum distance the T-Mic could be from the speaker, if the distance in the formula is reduced from 30mm to 27mm (the approximate centre of the T-Mic) this then gives a fundamental of 6351Hz which corresponds to a peak in levels on Figure 6.7.

6.5 Summary

While some CI settings were shown to produce minor fluctuations in the Advanced Bionics Listening Check output, these were not bigger than the fluctuation shown in the repeatability testing (Figure 6.7). While there is less importance in regulating CI processor settings for experiments, the programme specific microphone choice still remains a factor and needs to be taken into account when testing. This does increase the ease of conducting checks on Advanced Bionics CI processors using the listening check device, as the specific map levels and advanced features will not affect the results significantly. It should be possible to compensate for individual specific microphone programme settings the majority of the time: however, there are specific use cases where problems could arise, such as if someone has only one programme that is 50/50 split between the T-Mic and the processor microphone. Hopefully, if someone is wanting to assess the functionality of a T-Mic they will have a programme on their CI processor that uses just that microphone[16].

The experiment into the time taken to get sound through the Advanced Bionics listening check reveals a number of important factors that should be taken into account for

[15] Range of ±1.56dB

[16] without either the processor or coil microphone

subsequent testing (experiment in Section 6.3). Sweeping pure tone stimuli are best to be avoided when testing a processor, as compensating for this delay to the listening check output would be difficult to compensate for as it would be hard to know if you were measuring the frequency being presented or the frequency that was presented 0.25 to 0.45 seconds ago. Also, future experiments should allow a short duration of at least 300ms to allow the stimulus to reach the output of the listening check before commencing any measurements or recordings. This is to allow the signal to get to the output so avoiding measuring the BGN before the stimulus was presented rather than the stimulus itself.

Outside of a clinic environment, people do not need to measure microphone performance with any accuracy as long as it is consistent. This is because a system could be rolled out that compares the current state of a microphone to a recording of when it was known to be working, such as during an appointment. Therefore consistency is required and not accuracy, and this is why the results from the repeatability testing in Section 6.4 are so encouraging. Frequency responses used a white noise stimulus averaged over 30 seconds[17] repeated using a laptop connected to a 3D printed enclosure that had a headphone earpiece glued in one end and a slot for a T-Mic to be inserted in the other, shown in Figure 3.2. This set-up was then assembled, recorded a frequency response and then packed up again 56 times over a month in various different rooms and environments to simulate the kind of conditions that it could be used in. This resulted in a maximum deviation of $\pm$ 2dB and an average deviation of $\pm$ 0.8dB[18]. With further research and development of the enclosure and testing process, it should be possible to significantly reduce this variability. Even with the current variability, this is sufficient to notice the majority of the reductions in performance that occur, as shown in Section 4.2.

[17] as per Section 6.3, there was 0.5 second gap after the white noise started playing before frequency responses started to be recorded.

[18] Averaged between 100Hz and 10kHz

Chapter 7

Discussion

This chapter discusses a variety of topics that are related to this project including: how Cochlear Implant (CI) manufacture and microphone type affects someone's microphone replacement options; how having direct support from a CI manufacture would have impacted this project and different ways that microphone performance could be measured in future projects.

7.1 Impact of manufacture and microphone type

Advanced Bionics T-Mics have been used extensively during this project but these only represent a fraction of CI microphones that are used. As shown in Section 4.1, there is a perception by some clinicians that the Advanced Bionics T-Mics are less durable. Considering these are connected with a small cable to the main processor body it is understandable how this could be susceptible to damage by wear and tear. The processor microphone is integrated into the processor body which theoretically makes it more durable, but also significantly more expensive to repair or replace if something goes wrong with the microphones.

How comparable are the results for Advanced Bionics T-Mics to other CI microphones? As discussed above, there are potential differences in reliability between a T-Mic and a processor microphone. Despite this, T-Mics are much easier to replace as they do not need reprogramming. Furthermore, there is a significant financial difference between replacing a T-Mic for more than £140 and replacing a microphone with an integrated processor for significantly more. In a perfect world the cost would not impact the care of people who use CIs, but are CI centres less likely to replace a microphones if it costs so much more? The research in this project still gives a valuable insight into CI microphone performance but the potential difference between T-Mics and processor microphones should be considered.

7.2 Support from manufacturers

This project extensively used Advanced Bionics T-Mics and support from the manufacturer would have made elements of this project easier. The Advanced Bionics Listening Check was the main method used for testing the T-Mics performance but this has disadvantages. As investigated in Section 6.2, CI settings do have an effect on the Listening Checks audio output. Additionally, the Listening Checks output is affected by the CI processors programmes choice of microphone.

It would have been helpful if work could have been done with one of the manufacturers to make a device that bypasses all settings on the CI and provides simultaneous audio outputs for all the microphones. Being able to bypass the CI settings and set simultaneous signals for the microphones would have been preferred. This would have enabled a device to be built that CI clinicians could have put a CI processor into, similar to a Hearing Aid (HA) test box. The test box would bypass the settings that could interfere with the results, play a stimulus and then test all the CIs microphones at once. Such a device would enable clinicians to quickly and accurately find out if the CIs microphones are working. However, building such a device without help from manufacturers would be problematic.

7.3 Physical device

One option that could be used is to build a physical device to test the CIs microphones. Whether the physical device is intended for clinicians or people who use CI will affect the design. There are a multitude of test boxes available for HAs and the different HA manufacturers make adaptors that enable their HAs to be connected to HA test boxes. CIs companies could make adaptors similar to the Advanced Bionics Listening Check but that bypasses processor settings and maybe use a switch on the adaptor to swap between the different CI microphones.

If a test device is wanting to be used for home testing then the device will need to be made portable, cheaper and have a way of reporting the results simply while ideally sending full data back to the CI centre to be reviewed. This would be significantly easier to do with support from a CI manufacture but not impossible without.

7.3.1 Portable CI test box

While testing using the Advanced Bionics Listening Check has provided a wealth of data thus far during this project, this methodology does have disadvantages. Despite experiments showing that the majority of user specific settings do not affect the Listening Checks output, the current programmes microphone selection still affects the output.

As a continuation of the custom enclosure methodology described in Section 3.1.2, a mount and circuit was designed that allowed a T-Mic to be directly connected bypassing the CI processor and Listening Check; this enabled greater control of the microphones digital signal processing. Greater control of the signal chain should result in more accurate and consistent tests, while functioning independently of any user specific CI settings. Figure 7.1 shows the device with a T-Mic fitted and the integrated speaker directly below the T-Mic mount. The system was designed with a 4 pole 3.5mm audio connector that could be used to play stimuli from the integrated speaker and record the output from the T-Mic. However, it would be feasible to embed the required components for the test box to test microphones without a connected computer. The details of how the circuit works are not included here as it involved the reverse engineering of proprietary systems and publishing the findings from these investigations could risk legal repercussions.

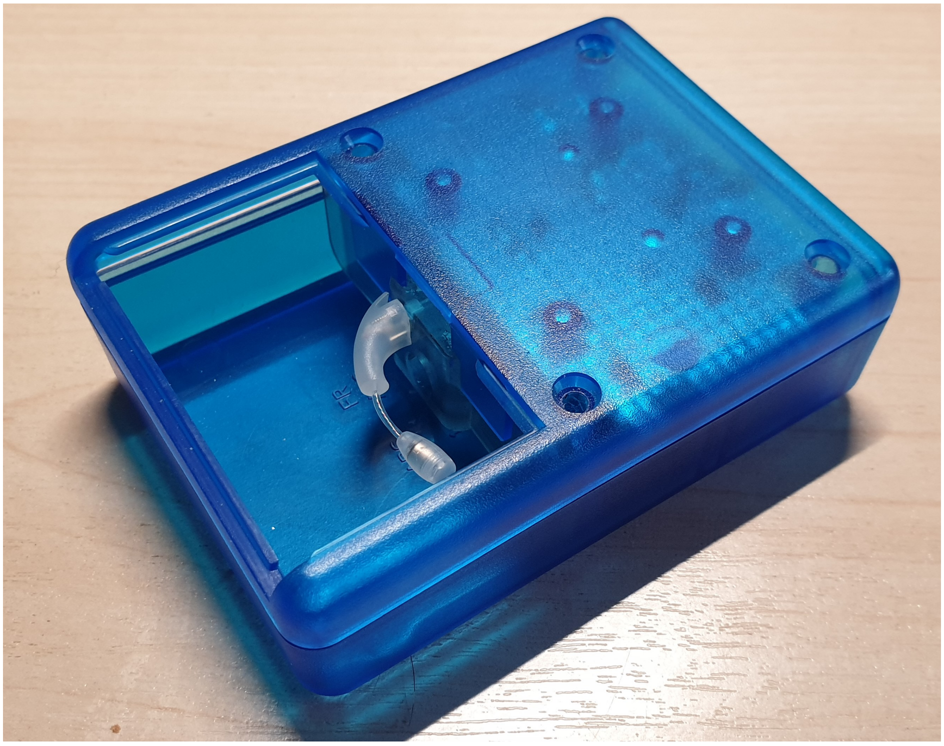


Figure 7.1: Picture showing the prototype portable cochlear implant test box

This project was discontinued as there were doubts that there would be sufficient willingness to use the system in a clinic environment. This is also a reverse engineered device that is not validated by the manufacturer therefore, the effect on the manufactures warranty could not be determined. Furthermore, the clinician survey (Section 4.1) showed that microphones were being replaced because the clinician could not be confident that they were working, as they were replacing microphones "just in case". This means that even if this device did get used extensively in clinic, clinicians are unlikely to be confident in its results and replace any tested microphones "just in case".

This device, or a device similar to this, would be significantly easier with the direct support of a manufacturer. This prototype required the T-Mic to be removed from the CI processor and the test device was only capable of testing this specific microphone. It would provide a significantly better user experience if the test device could connect directly to the CI processor in much the same way the Listening Check does. The device could even then connect via bluetooth to an application[1] that would send results to a CI centre. Such a device and method for measuring CI microphone performance would save CI centres time, effort and money. There would be a myriad of benefits for people who use CI as well, including reduced number to trips to CI centres to check microphones, reducing the numbers of microphones with unidentified defects being used and helping empower them to manage their own CI. However, such a device is not possible without extensive co-orporation from CI manufacturers.

7.4 Passively monitoring microphone input levels

One potential method for detecting microphone problems would be to passively monitor the input level to the CIs microphones. This would work by the CI itself taking measurements of the sound level into the different microphones at sporadic intervals and monitoring them. While the sounds that everyone is exposed to will vary depending on their environment, there is the potential that an AI might be able to notice gradual reductions in microphone performance[2] or sudden changes[3]. These patterns would be hard to spot (hence the need for AI) but it might be possible.

If this approach was to be researched then it would require extensive co-operation from a CI manufacture. This approach is not really possible to research as such a system would need to be integrated into the CI itself. Ease of implementation would depend on the systems currently used in a CI. Some devices may already log the data as part of the data logging process, therefore only the data analysis part of the process would need

[1] or integrated into an existing application
[2] due to a microphone becoming blocked
[3] due to a microphone breaking

to be developed. CI devices are designed to be as small as possible so they might not have the excess processing capacity to do the analysis, although a significant number of CI now have Bluetooth capabilities. This means that if needed the processing could be offloaded onto a phone running an application. The artificial intelligence would likely find it easier if there was also location data as well.

Measuring the long term average microphone input level may prove a useful way of passively measuring CI microphone performance without the need for AI, but there would need to be a way of accounting for both periodic events and infrequent events. Weekly or monthly events, or birthday parties might have significantly higher average sound levels so this would need to be taken into account and compensated for when analysing results. This could be compensated for by analysing longer term averages such as weekly instead of daily.

It was outside the scope of this research to investigate this but it does negate the need for either clinicians or people who use CI to specifically test the microphone (as would be required if using a physical test device like in Section 7.3).

7.5 Power of the mind

Having a method enabling clinicians or people at home to test CI microphones would be great but there needs to be confidence in the results. If the clinicians do not have confidence that the results are accurate and the microphones are actually working, then they are likely to replace them anyway. This is supported by the results from the clinician survey in Section 4.1 where 90% of the clinicians said they had replaced a microphone that they thought was properly working.

This means that even if a new system is developed that enables very accurate measurements to be taken of CI microphones, effort will need to be put into checking that the clinicians have confidence in the results. If the clinicians are ignoring the results and replacing microphones anyway then nothing has been achieved.

Chapter 8

Conclusion

The first research question outlined in Section 1.2 was: "*How accurate and repeatable are the current methods for identifying Cochlear Implant (CI) microphone failures?*". Chapter 4 investigated this research question with a number of experiments. This chapter first found that of the clinicians surveyed, 90% had at some point replaced a microphone they suspected was working. The next experiment tested CI microphones that had been reported as broken from the Auditory Implant Service (AIS). This found that 59% of the microphones that were reported as broken were found to be within ±3dB of known working values. The last experiment in the chapter investigated the accuracy and reliability of subjective microphone checks done by both a control group and CI clinicians. These experiments showed while both groups were able to correctly identify microphones that were not working, they were only able to accurately categorise the working microphones 40% of the time. Together these experiments show that the current methods used for measuring CI microphone performance are not accurate enough, with 60% of the microphones that are being sent for replacement by the AIS fully working and being unnecessarily replaced. Furthermore, testing also showed that 16.7% of the partially working microphones that were presented to CI clinicians were categorised as fully working.

Chapter 7 showed that with the current methods for testing CI microphones, it is possible for them to be tested by clinicians and thought to be fully working when they are only partially working. Chapter 5 investigated the second research question of: "*What impact does partially failed CI microphones have on speech perception?*" An experiment was conducted that measured how accurately participants could repeat words[1] with a reference setting and various simulated reductions in microphone performance. This experiment showed that reductions in microphone performance do not affect speech recognition, but the effect of reductions on microphone performance on environmental

[1]specifically, phoneme accuracy was access

sounds and quality of life are not yet known.

There are devices such as the Advanced Bionics Listening Check that provide a headphone output for a CI microphone. Chapter 6 investigated the third research question which was: "*What processor settings affect the audio output of the Listening Check?*" While the CIs microphone current selection does affect the listening checks output, the other settings do not affect the output. This makes objective tests more practical. The experiments in this chapter also showed that there is a delay between a sound and the output of the listening check of approximately 300ms. This needs to be considered when designing objective tests, especially when using stimulus that changes such as a pure tone sweep. 56 frequency responses of CI microphones were recorded to measure the consistency of the signal chain. The recordings had a maximum deviation of ± 2dB and an average deviation of ± 0.8dB[2]. These experiments show that CI microphone performance can be objectively measured using the Advanced Bionics Listening Check.

Face to face appointments have become more infrequent, especially with the increased risks now associated with them due to COVID-19. The pressure for more remote assessments may persist, especially when patients have to travel long distances. The increased duration between appointments means there will be longer durations between CI microphones being checked. This increases the need for a system to be implemented that ensures that people who use CIs have working microphones.

Objective ways of measuring CI microphone performance would enable clinics to reduce waste, avoid partially working microphones being overlooked and increase patient confidence in their equipment. These testing processes could also be embedded into telemedicine systems enabling regular home testing of CI microphones by the patient.

8.1 Advantages and disadvantages of implementing a remote care system

Implementing a system that enables CI clinicians and/ or people who use CIs to test if CI microphones are working would be a complex job. There are a number of advantages and disadvantages to implementing such a system; a summarised list of these is below.

Advantages

- Previous implementation of telemedicine systems has resulted in people who use CIs feeling more empowered (Chapter 2.5). It is likely that enabling people to

[2] Averaged between 100Hz and 10kHz

test how well their CIs microphones are working would further empower them.

- Chapter 4.2 showed that 59% of the microphones from the AIS that were marked as broken were fully working. If a remote care system was implemented that enabled people with CIs to measure microphone performance, there are a number of potential benefits:

    1. In the event that someone who uses a CI believes their microphone to be broken they will need to arrange a replacement. If they are capable of replacing the microphone themselves then a new microphone can be posted from the CI centre. If they are not able to replace the microphone themselves then this may well require a special trip to their CI centre. Reducing these unnecessary trips to the CI centre would benefit both the CI centre in saved staff time and the people who use the service in saved time and reduced trips to the CI centre.
    2. Reducing the false negatives does have significant cost saving potential for the CI department.

Disadvantages

- Research, prototyping and trials would all be needed prior to implementing any system and this would be expensive. There are possible methods of reducing these costs such as implementing it into a pre-existing system such as CHOICE from the auditory implant service[3].

- It will take time to implement any system; having the full support of a CI manufacturer would reduce this time but it would still not be able to be instantaneously rolled out.

- There are a great many people that would benefit from implementing such a telemedicine system. However, in order to use the system a certain level of technological ability would be required. This is discussed more in Section 2.7.

8.2 Publications and Contributions

The clinician survey (Section 4.1), objective testing of reportedly broken CI microphones (Section 4.2) and repeatability of subjective microphone checks (Section 4.3) were combined into a research paper that being submitted to a journal with the aim of getting it published.

[3] https://ais.southampton.ac.uk/choice/

This project showed that the current methods for identifying CI microphones that need replacing are inaccurate and result in 57% of the microphones that are being sent for replacement being within ±3dB of reference values. This identified and measured the extent of a problem that prior research had not done.

8.3 Future research

Hearing Aid (HA) test boxes are common place in audiology clinics yet from personal experience working in clinics as well as talking to numerous clinicians, they are seldom used on HAs and cannot be used with CIs. While research into the usage of HA test boxes is only tangentially related to CI microphone functionality, knowing how often test boxes are used could give a valuable insight into how CI microphone test boxes could be used in clinic. It would also be valuable to know why audiologists use or avoid using the test boxes, as this will help inform the design of any similar system for CIs.

As previously mentioned in Section 7.4, it might be possible to build a system that can detect problems with CI microphones. If input levels are taken at regular intervals then analysed, long term patterns established, gradual reductions in performance might be detectable. This theory is given credence by recent studies that have been able to distinguish between Covid-19, Bronchitis and Pertussis by analysing the sound of someone's cough (Imran et al. 2020). However, the survey in Section 4.1 showed that clinicians tend to replace microphones they cannot be certain are working, this means that if this passive microphone monitor was to be used to reduce the number of unnecessary CI microphone replacements, clinicians would need to have confidence that it works. Otherwise they would replace the microphone "just in case".

The discontinued project outlined in Section 7.3.1 prototyping a portable CI test box that objectively measures microphone performance also warrants further investigation, but as stated would really require the direct cooperation of a CI manufacture.

8.4 Summary

The first research question asked: "*How accurate and repeatable are the current methods for identifying CI microphone failures?*" This was investigated in Chapter 4 which through a number of discrete experiments showed that CI clinicians are primarily use a combination of speech perception testing, sound field testing and patient reports to identify reduction in microphone performance. Furthermore, nine out of ten clinicians said they replaced microphones "*just in case*" with this happening at varying frequencies from multiple times a week to a few times a year. This section also showed that of the microphones that were reported as broken at the University of Southampton AIS, 57% were still fully working. The only way of testing CI microphones in isolation is by doing subjective listening checks; while CI clinicians were able to identify the microphones that were "*Not working*" with 100% accuracy, they were less than 70% accurate at differentiating between the "*Partially Working*" and "*Working*" microphones. Together these experiments show that the current methods for identifying microphone performance are inaccurate and result in a plethora of microphones that are still fully working being unnecessarily sent for replacement. This will negatively affect the departments budget, produce unnecessary waste and cause unnecessary work for both the CI clinicians and the people who use the CIs.

The second research question asked: "*What impact do partially failed CI microphones have on speech perception?*" While none of the simulated reductions in microphone performance tested in Chapter 5 had a statistically significant effect on phoneme recognition, only a few different reductions in microphone performance were tested. It is likely that some of these reductions in microphone performance would negatively affect the quality of life of someone who uses a CI.

The third research question asked: "*What processor settings affect the audio output of the Listening Check?*" This was investigated in Chapter 6 that performed a variety of tests on CIs processors. Some of the CI processor settings did affect the output of the Advanced Bionics Listening Check; these fluctuations were generally low frequency and of comparable size to the rest re-test repeatability of the test set-up. However, the programme specific microphone choice still affects the output of the Listening Check. It would be possible to implement a test that objectively measures CI microphone performance using the Advanced Bionics Listening Check: however, this would remain problematic without manufacture support.

References

Abu-Ghanem, S., Handzel, O., Ness, L., Ben-Artzi-Blima, M., Fait-Ghelbendorf, K. & Himmelfarb, M. (2016), 'Smartphone-based audiometric test for screening hearing loss in the elderly', *European Archives of Oto-Rhino-Laryngology* **273**(2), 333–339.

Advanced Bionics (2013), 'Naida CI Q70 Technical Specifications'. Available at `https://advancedbionics.com/content/dam/advancedbionics/Documents/Regional/DK/Datasheets/028-M717-02-International%20Naida%20CI%20Q90%20EAS%20Tech%20Specs.pdf`. Accessed 30th May 2019.

Advanced Bionics (2015*a*), 'Naida ci q70 user guide'. Available at `https://advancedbionics.com/content/dam/advancedbionics/Documents/Regional/DE/Produkte-DE/Na%C3%ADda/NaidaCI-Q70/029-M504-87_Rev_A_Naida%20CI%20Q70%20User%20Guide%20(new)_preview.pdf`. Accessed 4th October 2018.

Advanced Bionics (2015*b*), 'Naida CI Q90 Technical Specifications'. Available at `https://advancedbionics.com/content/dam/advancedbionics/Documents/Regional/DK/Datasheets/028-M530-02-Na%C3%ADda%20CI%20Q90%20Tech%20Specs%20Sheet.pdf`. Accessed 30th May 2019.

Anker, S. D., Koehler, F. & Abraham, W. T. (2011), 'Telemedicine and remote management of patients with heart failure', *The Lancet* **378**(9792), 731–739.

Bauer, P., Jones, J. & Fingscheidt, T. (2013), Impact of hearing impairment on fricative intelligibility for artificially bandwidth-extended telephone speech in noise, *in* '2013 IEEE International Conference on Acoustics, Speech and Signal Processing', IEEE, pp. 7039–7043.

Bench, J., Kowal, Å. & Bamford, J. (1979), 'The bkb (bamford-kowal-bench) sentence lists for partially-hearing children', *British Journal of Audiology* **13**(3), 108–112.

Bodmer, D. & Capkun, S. (2010), 'On the security and privacy risks in cochlear implants', *Dept. Comput. Sci., ETH, Zürich, Switzerland* .

Boothroyd, A. (1968), 'Statistical theory of the speech discrimination score', *The Journal of the Acoustical Society of America* **43**(2), 362–367.

Boothroyd, A. & Nittrouer, S. (1988), 'Mathematical treatment of context effects in phoneme and word recognition', *The Journal of the Acoustical Society of America* **84**(1), 101–114.

Buckman, M. & Fitzharris, K. (2020), 'Teleaudiology and cochlear implant appointments', *The Hearing Journal* **73**(5), 30–31.

Buechner, A. (2013), 'Utilization of advanced pre-processing strategies in cochlear implant systems: The new naída ci q70'. Presented at the 11th European Symposium on Cochlear Implants in Istanbul, Turkey.

Bynum, A. B., Irwin, C. A., Cranford, C. O. & Denny, G. S. (2003), 'The impact of telemedicine on patients' cost savings: some preliminary findings', *Telemedicine Journal and e-Health* **9**(4), 361–367.

Chaudhry, S. I., Mattera, J. A., Curtis, J. P., Spertus, J. A., Herrin, J., Lin, Z., Phillips, C. O., Hodshon, B. V., Cooper, L. S. & Krumholz, H. M. (2010), 'Telemonitoring in patients with heart failure', *New England Journal of Medicine* **363**(24), 2301–2309.

Clark, R. A., Inglis, S. C., McAlister, F. A., Cleland, J. G. & Stewart, S. (2007), 'Telemonitoring or structured telephone support programmes for patients with chronic heart failure: systematic review and meta-analysis', *Bmj* **334**(7600), 942.

Clarke, M., Shah, A. & Sharma, U. (2011), 'Systematic review of studies on telemonitoring of patients with congestive heart failure: a meta-analysis', *Journal of Telemedicine and Telecare* **17**(1), 7–14.

Corry, M., Sanders, M. & Searchfield, G. D. (2017), 'The accuracy and reliability of an app-based audiometer using consumer headphones: pure tone audiometry in a normal hearing group', *International Journal of Audiology* pp. 1–5.

Cullington, H. & Aidi, T. (2017), 'Is the digit triplet test an effective and acceptable way to assess speech recognition in adults using cochlear implants in a home environment?', *Cochlear Implants International* **18**(2), 97–105.

Cullington, H., Kitterick, P., Weal, M. & Margol-Gromada, M. (2018), 'Feasibility of personalised remote long-term follow-up of people with cochlear implants: a randomised controlled trial', *BMJ open* **8**(4).

De La Torre-Díez, I., López-Coronado, M., Vaca, C., Aguado, J. S. & de Castro, C. (2015), 'Cost-utility and cost-effectiveness studies of telemedicine, electronic, and mobile health systems in the literature: a systematic review', *Telemedicine and e-Health* **21**(2), 81–85.

Dillon, H. (2008), *Hearing aids*, second edn, Hodder Arnold. Section 8.1, page 227-233.

Dingemanse, J. G. & Goedegebure, A. (2015), 'Application of noise reduction algorithm clearvoice in cochlear implant processing: effects on noise tolerance and speech intelligibility in noise in relation to spectral resolution', *Ear and Hearing* **36**(3), 357–367.

Dingemanse, J. G. & Goedegebure, A. (2018), 'Optimising the effect of noise reduction algorithm clearvoice in cochlear implant users by increasing the maximum comfort levels', *International Journal of Audiology* **57**(3), 230–235.

Duffy, J. R. & Giolas, T. G. (1974), 'Sentence intelligibility as a function of key word selection', *Journal of Speech, Language, and Hearing Research* **17**(4), 631–637.

Ekeland, A. G., Bowes, A. & Flottorp, S. (2010), 'Effectiveness of telemedicine: a systematic review of reviews', *International Journal of Medical Informatics* **79**(11), 736–771.

Erson, S., Rusinov, D., Bata, R. & Barber, T. (2020), 'In home device for remote testing of cochlear implant microphones'. Available at `https://ais.southampton.ac.uk/remote-microphone-systems/`. Accessed 14th January 2021.

Fishman, K. E., Shannon, R. V. & Slattery, W. H. (1997), 'Speech recognition as a function of the number of electrodes used in the speak cochlear implant speech processor', *Journal of Speech, Language, and Hearing Research* **40**(5), 1201–1215.

Frank, T. & Dreisbach, L. E. (1991), 'Repeatability of high-frequency thresholds.', *Ear and Hearing* **12**(4), 294–295.

Friesen, L. M., Shannon, R. V., Baskent, D. & Wang, X. (2001), 'Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants', *The Journal of the Acoustical Society of America* **110**(2), 1150–1163.

Garrison, W. M. & Bochner, J. H. (2017), 'An application for screening gradual-onset age-related hearing loss', *Health* **9**(04), 715.

Gelfand, S. A. (2009), *Essentials of audiology*, Thieme New York, NY.

Giolas, T. G., Cooker, H. S. & Duffy, J. R. (1970), 'The predictability of words in sentences', *Journal of Auditory Research* .

Goehring, J. L. & Hughes, M. L. (2017), 'Measuring sound-processor threshold levels for paediatric cochlear implant recipients using conditioned play audiometry via telepractice', *Journal of Speech, Language, and Hearing Research* **60**(3), 732–740.

Goehring, J. L., Hughes, M. L., Baudhuin, J. L., Valente, D. L., McCreery, R. W., Diaz, G. R., Sanford, T. & Harpster, R. (2012), 'The effect of technology and testing environment on speech perception using telehealth with cochlear implant recipients', *Journal of Speech, Language, and Hearing Research* .

Greville, K. (1984), *Standardisation of NAC-AB speech discrimination materials*, Department of Health, Audiology Centre.

Hall, J. W. (2016), 'Education of audiologists in underserved regions', *The Hearing Journal* **39**.

Hanvey, K. (2020), British cochlear implant group annual data collection financial year 2018-2019, CII.

Hehrmann, P., Fredelake, S., Hamacher, V., Dyballa, K.-H. & Buechner, A. (2012), 'Improved speech intelligibility with cochlear implants using state-of-the-art noise reduction algorithms', pp. 1–3.

Hughes, M. L., Goehring, J. L., Sevier, J. D. & Choi, S. (2018), 'Measuring sound-processor thresholds for paediatric cochlear implant recipients using visual reinforcement audiometry via telepractice', *Journal of Speech, Language, and Hearing Research* **61**(8), 2115–2125.

Imran, A., Posokhova, I., Qureshi, H. N., Masood, U., Riaz, M. S., Ali, K., John, C. N., Hussain, M. I. & Nabeel, M. (2020), 'Ai4covid-19: Ai enabled preliminary diagnosis for covid-19 from cough samples via an app', *Informatics in Medicine Unlocked* **20**, 100378.
URL: *http://www.sciencedirect.com/science/article/pii/S2352914820303026*

Inglis, S. (2010), 'Structured telephone support or telemonitoring programmes for patients with chronic heart failure', *Journal of Evidence-Based Medicine* **3**(4), 228–228.

International Electrotechnical Commission (2005), 'Electroacoustics - hearing aids - part 7: Measurement of the performance characteristics of hearing aids for production, supply and delivery quality assurance purposes'.

International Electrotechnical Commission (2012), ‘Electroacoustics - hearing aids - part 15: Methods for characterising signal processing in hearing aids with a speech-like signal’.

Istepanian, R. S. H. (1999), ‘Telemedicine in the United Kingdom: current status and future prospects’, *IEEE Transactions on Information Technology in Biomedicine* **3**(2), 158–159.

Jax, P. & Vary, P. (2000), Wideband extension of telephone speech using a hidden markov model, *in* ‘2000 IEEE Workshop on Speech Coding. Proceedings. Meeting the Challenges of the New Millennium (Cat. No. 00EX421)’, IEEE, pp. 133–135.

Kalikow, D. N., Stevens, K. N. & Elliott, L. L. (1977), ‘Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability’, *The Journal of the Acoustical Society of America* **61**(5), 1337–1351.

Kane (2019), ‘Sample size calculator’. Available at `https://clincalc.com/stats/samplesize.aspx`. Accessed 4th October 2019.

Klersy, C., De Silvestri, A., Gabutti, G., Regoli, F. & Auricchio, A. (2009), ‘A meta-analysis of remote monitoring of heart failure patients’, *Journal of the American College of Cardiology* **54**(18), 1683–1694.

Koehler, F., Winkler, S., Schieber, M., Sechtem, U., Stangl, K., Böhm, M., Boll, H., Baumann, G., Honold, M., Koehler, K. et al. (2011), ‘Impact of remote telemedical management on mortality and hospitalizations in ambulatory patients with chronic heart failure’, *Circulation* pp. CIRCULATIONAHA–111.

Luryi, A. L., Tower, J. I., Preston, J., Burkland, A., Trueheart, C. E. & Hildrew, D. M. (2020), ‘Cochlear implant mapping through telemedicine—a feasibility study’, *Otology & Neurotology* **41**(3), e330–e333.

Lutman, M., Hall, S. & Athalye, S. (2006), Development of a telephone hearing test.

Marietta, G. (2001), ‘Telemedicine: cost-effective management of high-risk pregnancy’, *Managed Care* .

Martin, M. (1997), *Speech Audiometry*, second edn, Whurr Publishers Ltd.

Maurrasse, S. E., Rastatter, J. C., Hoff, S. R., Billings, K. R. & Valika, T. S. (2020), ‘Telemedicine during the covid-19 pandemic: A paediatric otolaryngology perspective’, *Otolaryngology–Head and Neck Surgery* p. 0194599820931827.

Mistry, H. (2012), 'Systematic review of studies of the cost-effectiveness of telemedicine and telecare. Changes in the economic evidence over twenty years', *Journal of Telemedicine and Telecare* **18**(1), 1–6.

Molander, P., Nordqvist, P., Öberg, M., Lunner, T., Lyxell, B. & Andersson, G. (2013), 'Internet-based hearing screening using speech-in-noise: Validation and comparisons of self-reported hearing problems, quality of life and phonological representation', *BMJ open* **3**(9), e003223.

Muñoz, K., Kibbe, K., Preston, E., Caballero, A., Nelson, L., White, K. & Twohig, M. (2017), 'Paediatric hearing aid management: A demonstration project for using virtual visits to enhance parent support', *International Journal of Audiology* **56**(2), 77–84.

Myles, A. J. (2017), 'The clinical use of Arthur Boothroyd (AB) word lists in Australia: exploring evidence-based practice', *International Journal of Audiology* **56**(11), 870–875.

National Institute for Health and Care Excellence (2019), 'Cochlear implants for children and adults with severe to profound deafness'.

National Institute for Health and Care Excellence (2020), 'Hearing loss overview'. Available at `https://pathways.nice.org.uk/pathways/hearing-loss/hearing-loss-overview.pdf`. Accessed 14th January 2021.

Nemes, J. (2002), 'Fearing shortage of audiologists, profession seeks ways to entice students', *The Hearing Journal* **55**(6), 23–26.

NHS (2019), 'The nhs long term plan'. Available at `www.longtermplan.nhs.uk/publication/nhs-long-term-plan/`. Accessed 10th January 2019.

Nilsson, M., Soli, S. D. & Sullivan, J. A. (1994), 'Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise', *The Journal of the Acoustical Society of America* **95**(2), 1085–1099.

Nordrum, S., Erler, S., Garstecki, D. & Dhar, S. (2006), 'Comparison of performance on the hearing in noise test using directional microphones and digital noise reduction algorithms', *American Journal of Audiology* **15**(1), 81–91.

Ozimek, E., Kutzner, D., Sek, A. & Wicher, A. (2009), 'Development and evaluation of Polish digit triplet test for auditory screening', *Speech Communication* **51**(4), 307–316.

Pittman, A. (2011), 'Children's performance in complex listening conditions: Effects of hearing loss and digital noise reduction', *Journal of Speech, Language, and Hearing Research* **54**(4), 1224–1239.

Purdy, S., Arlington, B. & Johnstone, C. (2000), 'Normative data for the New Zealand recording of the CVC (Revised AB) word lists', *New Zealand Audiological Society Bulletin* **10**(2), 20–29.

Roine, R., Ohinmaa, A. & Hailey, D. (2001), 'Assessing telemedicine: a systematic review of the literature', *Canadian Medical Association Journal* **165**(6), 765–771.

Roth, T. N., Hanebuth, D. & Probst, R. (2011), 'Prevalence of age-related hearing loss in Europe: a review', *European Archives of Oto-Rhino-Laryngology* **268**(8), 1101–1107.

Saliba, J., Al-Reefi, M., Carriere, J. S., Verma, N., Provencal, C. & Rappaport, J. M. (2017), 'Accuracy of mobile-based audiometry in the evaluation of hearing loss in quiet and noisy environments', *Otolaryngology–Head and Neck Surgery* **156**(4), 706–711.

Schmuziger, N., Probst, R. & Smurzynski, J. (2004), 'Test-retest reliability of pure-tone thresholds from 0.5 to 16 kHz using Sennheiser HDA 200 and Etymotic Research ER-2 earphones', *Ear and hearing* **25**(2), 127–132.

Shapiro, S. B., Lipschitz, N., Kemper, N., Zuccarello, M., Breen, J. T., Pensak, M. L. & Samy, R. N. (2020), 'Early experience with telemedicine in patients undergoing otologic/neurotologic procedures', *Otology & Neurotology* .

Smits, C., Houtgast, T. & Kapteyn, T. S. (2004), 'Development and validation of an automatic speech-in-noise screening test by telephone', *International Journal of Audiology* **43**(1), 15–28.

Thompson, G. P., Sladen, D. P., Borst, B. J. H. & Still, O. L. (2015), 'Accuracy of a tablet audiometer for measuring behavioral hearing thresholds in a clinical population', *Otolaryngology–Head and Neck Surgery* **153**(5), 838–842.

Thorén, E. S., Öberg, M., Wänström, G., Andersson, G. & Lunner, T. (2013), 'Internet access and use in adults with hearing loss', *Journal of Medical Internet Research* **15**(5).

Waltzman, S. B., Roland, J. T. et al. (2000), *Cochlear implants*, third edn, Thieme New York.

Wesarg, T., Wasowski, A., Skarzynski, H., Ramos, A., Falcon Gonzalez, J. C., Kyriafinis, G., Junge, F., Novakovich, A., Mauch, H. & Laszig, R. (2010), 'Remote fitting in nucleus cochlear implant recipients', *Acta oto-laryngologica* **130**(12), 1379–1388.

Westhorp, S. (2009), 'Speech recognition tests', *The British Association of Teachers of the Deaf* .

Whitten, P. S., Mair, F. S., Haycox, A., May, C. R., Williams, T. L. & Hellmich, S. (2002), 'Systematic review of cost effectiveness studies of telemedicine interventions', *Bmj* **324**(7351), 1434–1437.

Wilson, W. J., Jones, B. & Fridjhon, P. (1998), 'Use of the NAL-AB wordlists as a South African English speech discrimination test', *The South African Journal of Communication Disorders= Die Suid-Afrikaanse tydskrif vir Kommunikasieafwykings* **45**, 77–86.

Windmill, I. M. & Freeman, B. A. (2013), 'Demand for audiology services: 30-yr projections and impact on academic programs', *Journal of the American Academy of Audiology* **24**(5), 407–416.

Wootton, R. (2012), 'Twenty years of telemedicine in chronic disease management–an evidence synthesis', *Journal of Telemedicine and Telecare* **18**(4), 211–220.

World Health Organisation (2018), 'Deafness and hearing loss'. Available at `www.who.int/en/news-room/fact-sheets/detail/deafness-and-hearing-loss`. Accessed 10th January 2019.

Zakis, J. A., Hau, J. & Blamey, P. J. (2009), 'Environmental noise reduction configuration: Effects on preferences, satisfaction, and speech understanding'.

Appendix A

Pure tone stimuli frequencies

- 105
- 199
- 293
- 410
- 527
- 715
- 902
- 1090
- 1207
- 1301
- 1418
- 1512
- 1605
- 1746
- 1863
- 2027
- 2168
- 2332
- 2496
- 2660
- 2707
- 2801
- 2918
- 2988
- 3105
- 3246
- 3316
- 2504
- 2398
- 3715
- 3855
- 3996
- 4254
- 4488
- 4676
- 4793
- 5004
- 5215
- 5449
- 5754
- 5988
- 6199
- 6480
- 6715
- 6996
- 7207
- 7488
- 7723
- 8004
- 8238
- 8730
- 8988
- 9223
- 9363
- 9504