Improving the New Zealand Digit Triplet Test Using Antiphasic Stimuli

A thesis submitted in partial fulfilment of the requirements for the Degree

of Master of Audiology

in the University of Canterbury

by Rosalie Joan Hadley

University of Canterbury

2022

Abstract

Hearing impairment is a condition which affects many New Zealanders. The World Health Organization estimates that by the year 2050, 1 in every 4 people worldwide will suffer from a hearing impairment of some form (World Health Organization, 2021). As hearing impairment has been found to have negative consequences for both individuals and society at large, having widely available hearing screening tools is of growing importance.

While traditional audiological testing is the gold standard for detecting a hearing impairment, this is not always available – particularly in rural or third world countries. As well as this, the recent outbreak of COVID-19 has highlighted the need of alternative testing methods which can be done in a socially-distanced manner (De Sousa, Smits, et al., 2020).

This thesis looks to improve one such method of contactless testing available in New Zealand known as the New Zealand Hearing Screening Test, which uses a New Zealand English digit triplet test. Traditionally this and other digits-in-noise tests have used diotic stimuli (where identical stimuli are presented to both ears) and have been scored by triplet (where all three numbers in the triplet must be correctly entered). The present study investigated the use of antiphasic stimuli (where the polarity of the speech is reversed in one ear) which was shown by De Sousa, Swanepoel, et al. (2020) to improve the sensitivity and specificity of digits-in-noise testing. This study also investigated whether scoring by individual digit rather than digit triplet could also improve the ability of the test to distinguish between people with normal hearing and those with hearing impairment.

Our results found that similar to De Sousa, Swanepoel, et al. (2020), the use of antiphasic stimuli increased the sensitivity and specificity of the New Zealand English digit triplet test from 94% and 88% to 95% and 90% respectively. However, scoring by digit rather than triplet was not found to improve test performance.

Acknowledgements

This research has only been possible due to a great many academic and emotional supporters. I wish first and foremost to thank my thesis supervisors Greg O'Beirne and Rebecca Kelly and my clinical educators Jonny Grady, Nicole Borland, and Kate Jepsen at the University of Canterbury, without whom I could not have completed this research. Your expertise and guidance was of immense help.

I would also like to thank my fellow classmates who provided a shoulder to lean on and worked patiently around me with testing times and spaces.

I wish to thank my friends Peter and Emma, my brother Craig, my sister Suzanne, and my in-laws Shannon, Laura, Gail, Eric, Veronica and Joe for lending their ears numerous times. Your patience and support was indispensable to the completion of this project.

I want to thank and acknowledge my husband Grant. Without your love, encouragement, and support it would not have been possible for me to undertake and complete this programme.

And finally I would like to dedicate this research to my late parents Lindsay and Kathleen Hosking.

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1. The Importance of Hearing

In the suburbs of Christchurch there is a garden. A silver tabby cat playfully brushes against the leg of an outside chair, its soft purr intermingling with vibrant bird song and children's laughter. An elderly couple sit together, one soaking in the bustling chorus of suburbia, while the other silently sips their tea, oblivious to the surrounding symphony.

"Isn't the tui song magnificent!" one exclaims.

"What?" remarks the other, with a look of confusion.

"I think it's time to get you a hearing aid" the first replies with a sigh.

"Time to get a what?"

Situations such as illustrated above are unfortunately not uncommon, with hearing impairment affecting many people both residing within New Zealand and abroad. The 2021 World Report on Hearing produced by the World Health Organization (WHO) found that over 1.5 billion people worldwide are currently experiencing some form of hearing impairment, and this number is expected to grow to a staggering 2.5 billion – equating to 1 in every 4 people – by the year 2050 (World Health Organization, 2021). The WHO estimate that at least US$1 trillion is lost each year due to a failure to address hearing impairment (World Health Organization, 2021).

With regard to New Zealand, a 2017 study found 18.9% of New Zealanders suffer from some form of hearing impairment, resulting in an estimated cost to the public health system of $131.8 million (National Foundation for the Deaf & Deloitte Access Economics, 2017). Exeter (2015) estimated that 14.87% of those with an impairment in New Zealand were between the ages of 60-69, and 22.23% over the age of 70 (Exeter, 2015). With New Zealand's aging population increasing, they estimate that the number of people aged over 14 with some form of hearing impairment will double by 2061.

While there are significant economic consequences, hearing impairments also have a psychosocial impact on individuals and families living with them. Research has found that hearing impairment limits an individual's ability to communicate and engage with others, leading to increased social isolation, increased problems with friends and family, higher levels of anxiety and depression, and issues in the workplace (Strawbridge et al., 2000; Veiga et al., 2015). While many studies mainly focus on outcomes for older persons, young people are also greatly impacted. A study by Butcher et al. (2019) found that young people with hearing impairment had an increased risk of depression, and were more likely to self-harm and experience peer victimisation (Butcher et al., 2019).

Due to the issues highlighted above, it is of utmost importance from both the viewpoint of individuals and society at large that hearing impairments are identified, and that people with hearing impairments are given help where possible to improve their quality of life and prevent unnecessary burdens on the health care system. But before examining how hearing impairments can be identified through audiologic testing, we will quickly survey how we hear and the forms that a hearing impairment can take.

1.2. The Peripheral Auditory System

The ear can be divided into three different sections – the outer ear, middle ear, and inner ear (see Figure 1). The outer portion of the ear consists of the pinna (ear lobe) and external auditory meatus (ear canal). The middle section contains the tympanic membrane (ear drum) and ossicles (middle ear bones). Lastly, the inner portion houses the cochlea and semi-circular canals, which are the organs of hearing and balance.

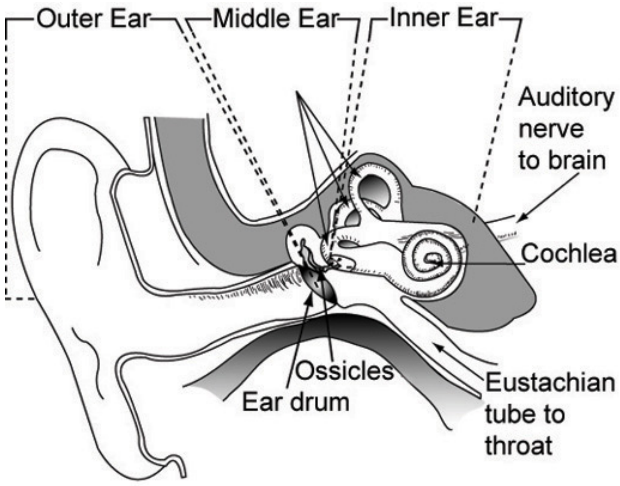


Figure 1: Cross Anatomy of the Right Ear (Coronal Section) (National Aeronautics and Space Administration (NASA), 2004)

When all parts are functioning as they should, sound waves travel through the external auditory meatus where they are intercepted by the tympanic membrane. The tympanic membrane is vibrated by these waves, and this vibration causes sound to pass through the ossicular chain and on towards the cochlea. One of the main purposes of the middle ear is to overcome the impedance mismatch which is created by sound moving from an air medium – the acoustic sound waves that enter from the external ear canal – to a fluid medium when carried on to the inner ear (Kramer & Brown, 2018). A common analogy to illustrate this impedance mismatch is to imagine yourself trying to talk to a person who is underwater while you yourself stand above it. Roughly 99.9% of the sound energy from your voice is actually reflected off the surface of the water, with a mere 0.1% making it through the water and onto the person you with whom are trying to speak (Kramer & Brown, 2018). In terms of decibels, this can be expressed as the equivalent of a 30 dB loss in sound energy (Kramer & Brown, 2018). The ossicular chain overcomes this loss by using a lever system which boosts the sound energy entering the cochlea by around 33 dB (Kramer & Brown, 2018). Once through the ossicles, the sound energy is applied to the oval window. The motion caused by the oval window flexing in response to the movement of the ossicles creates a traveling wave that then traverses the basilar membrane (Emanuel et al., 2009).

The basilar membrane spirals the full length of the cochlea, and is thinner and stiffer at its base and wider and more elastic at the apex. Higher frequencies correlate with the thinner, stiffer areas closer to the base of the cochlea and lower frequencies with the wider, looser areas towards its apex (Kramer & Brown, 2018).

The cochlea itself is a small snail-shell shaped organ consisting of approximately 2.5 turns. It has a length of around 3.1-3.3 cm when fully uncurled, and a height of around 0.5 cm (Møller, 2013). It is encased by bone and made up of three fluid filled ducts – scala vestibuli, scala media, and scala tympani. Scala tympani and scala media are separated by the basilar membrane and organ of Corti while scala vestibuli and scala media by separated by Reissner's membrane (Musiek & Baran, 2020). Both scala tympani and scala vestibuli are made up of an ionic fluid known as perilymph. Its composition is similar to extracellular fluid, containing high amounts of sodium and low amounts of potassium (Møller, 2013). Scala media contains a fluid known as endolymph, which is similar to intracellular fluid and – conversely to perilymph – consists of high amounts of potassium and low amounts of sodium (Møller, 2013).

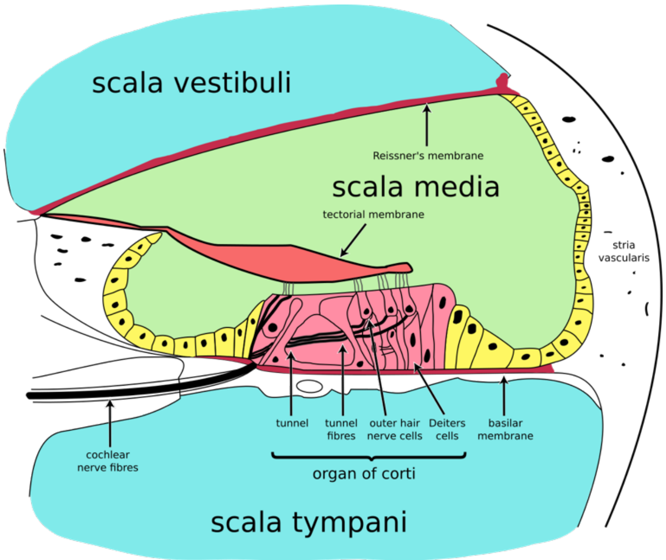


Figure 2: Cross-section of the Cochlea (Oarih Ropshkow, distributed under CC-BY-3.0 license)

The organ of Corti is known more colloquially as the organ of hearing. It contains two types of hair cells which work to convert mechanical vibrations into electrical signals. These are the inner hair cells (IHC), which are responsible for sending electrical signals to the auditory nerve, and the outer hair cells (OHC), which work to amplify sound (Robles & Ruggero, 2001). The organ of Corti contains one row of IHCs which number approximately 3,500, and two to three rows of OHCs that total around 11,000 in number (Ashmore, 2008). The hair cells come in bundles of stereocilia, which vary in length. Movement caused by vibration of the basilar membrane in the direction of the tallest stereocilia of the IHC bundles opens an ion channel which sees potassium and calcium ions enter the cell and generate a transduction current that in turn opens voltage gated calcium channels which allow calcium to enter the cell, triggering the release of neurotransmitter and action potentials in the auditory nerve (Raphael & Altschuler, 2003). The IHCs are responsible for about 95% of the afferent innervation which sees auditory signals sent to the brain, while OHCs are responsible for around 5% (Ashmore, 2008).

The OHCs respond to the transduction current by elongating and contracting, causing increased displacement of the organ of Corti and the amplification of low-level sounds by around 40 to 60 dB (Ashmore, 2008; Murakoshi et al., 2015).

As mentioned, sound moves as a travelling wave through the tonotopically arranged basilar membrane (Musiek & Baran, 2020). Different sections of the basilar membrane correspond to different ranges of frequencies which can be thought of as overlapping band-pass or auditory filters (Moore, 2012). Each specific point on the basilar membrane can be associated with a different auditory filter that centres around a specific frequency known as its characteristic frequency (Moore, 2012).

1.3. Frequency Selectivity

Frequency selectivity is the ability to distinguish different frequency components in complex sound. This is what allows for three tones to be heard simultaneously in a piano chord, or the low purr of a cat to be distinguished from – but heard at the same as – the high pitch song of a Tui. It also helps our auditory system to extract speech from noise.

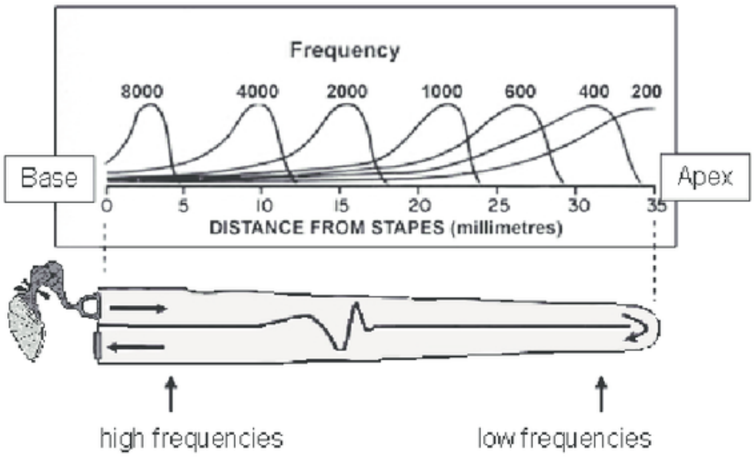


Figure 3: Schematic View of an Uncoiled Cochlea and the Traveling Wave (Emanuel et al., 2009)

As previously discussed, sound vibrations travelling through the middle ear ossicles cause the oval window to flex, which in turn creates a travelling wave which moves from the base of the basilar membrane to the apex. The graded changes in mass and stiffness along the basilar membrane mean that the location where the wave reaches its maximum displacement is dependent on the frequency of the sound wave, with higher frequencies such as 8000 Hz reaching maximum displacement between 0 – 5 mm from the oval window and lower frequencies such as 200 Hz reaching maximum displacement closer to 35 mm from the oval window, as illustrated in Figure 3 (Emanuel et al., 2009).

The nerve fibres located on the basilar membrane have a specific frequency to which they are tuned, referred to as the characteristic frequency or centre frequency. The characteristic frequency is the frequency at which a fibre is excited with the lowest amount of energy or intensity (Emanuel et al., 2009). For example, say we have an auditory nerve fibre whose characteristic frequency is 500 Hz. This fibre can also be stimulated by other tones, such as when a broadband noise is playing. However, to be stimulated by these other tones their intensity level needs to be greater because of the fact that those tones are further away in frequency from the fibre's characteristic frequency. A tone of 100 Hz with enough intensity could stimulate it, but because a much lower amount of energy or intensity is required for it to be stimulated when the tone is at 500 Hz, it is 500 Hz that is considered its characteristic frequency.

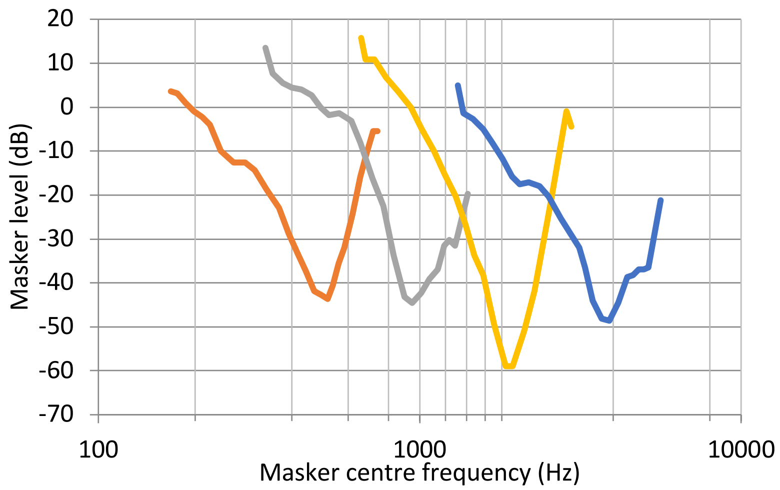


Figure 3: Psychophysical Tuning Curves Illustrate the Sensitivity and Specificity of Different Regions of the Cochlea (image by author)

This phenomenon of frequency selectivity can be measured using psychophysical tuning curves (PTC). Psychophysical tuning curves can be obtained by playing a signal at a specific frequency (usually presented 10 dB above the level at which the listener can just make out the sound) and then playing a masking noise (usually narrowband) at frequencies which are lower than the signal, at the same frequency as the signal, and then higher than it (Plack, 2013). When the masking falls at frequencies lower or higher than the signal, more masking is required to stop the listener hearing it. As the masking nears the signal and is just on top of it, less masking becomes required to stop the listener hearing it. An example can be seen in Figure 4, where PTCs for 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz have been obtained.

This function of the basilar membrane is of importance, as certain forms of hearing impairment can cause diminished frequency selectivity and PTCs that are broader, which in turn leads to a reduced ability to distinguish speech in background noise.

1.4. Forms of Hearing Impairment

There are four main forms of hearing impairment that an individual might develop – sensorineural (inner ear), conductive (outer and/or middle ear), mixed hearing impairment (both sensorineural and conductive), and retrocochlear (involving the auditory nervous system beyond the cochlea). Hearing impairments can also be either unilateral/asymmetric – affecting just one ear – or bilateral – affecting both.

Table 1: Hearing Impairment Classification Based on Goodman Classification (Clark, 1981)

Hearing Thresholds/Level (dB)	**Hearing Impairment Classification**
-10 to 15	Normal Hearing
16 to 25	Slight Hearing Impairment
26 to 40	Mild Hearing Impairment
41 to 55	Moderate Hearing Impairment
56 to 70	Moderate-Severe Hearing Impairment
71 to 90	Severe Hearing Impairment
90 and above	Profound Hearing Impairment

As well as there being different types of impairment, there are also varying degrees. The New Zealand Audiological Society recommends the Goodman classification for hearing impairment, shown in Table 1, for determining the degree of impairment an individual has.

1.4.1. Sensorineural Hearing Impairment

When there is an issue with the cochlea in the inner ear, this is known as a sensorineural hearing loss/impairment (SNHL). This form of hearing impairment is often permanent, and involves damage to the inner and/or outer hair cells that are housed in the cochlea. This damage interferes with the mechanoelectrical transduction process of the IHCs which are responsible for sending electrical signals to the auditory nerve (White et al., 2021).

There are a number of causes of sensorineural impairment, with the most common being presbycusis (age-related hearing impairment), the use of ototoxic medications such as aminoglycoside and macrolide antibiotics, exposure to loud noise over a prolonged period, vestibular schwannomas (auditory nerve tumours), certain infections such as meningitis, and genetic disorders (World Health Organization, 2021). Qi et al. (2019) state that sensorineural impairments are the most common, accounting for around 90% of all impariments (Qi et al., 2019).

The most common form of sensorineural hearing impairment is presbycusis, with over 60% of people expected to have some degree of this by the time they reach 70 years of age (Parthasarathy & Kujawa, 2018). Presbycusis occurs due to a progressive deterioration of IHCs, and typically presents in the higher frequencies.

Researchers also believe that a progressive loss in the synapse between the IHCs and auditory nerve occurs with aging, and can result in reduced frequency selectivity (Parthasarathy & Kujawa, 2018). This type of degeneration does not always present with a noticeable decreasing in hearing thresholds during audiometric testing, but we include it with presbycusis due to the difficulties hearing in background noise due to natural aging that it produces.

Noise induced hearing loss/impairment (NIHL) is also particularly prevalent form of SNHL. This is caused by exposure to excessively loud sounds, which may be for a short or prolonged period of time. This noise exposure can result in either a temporary or permanent shift in hearing thresholds, with permanent changes arising when irreversible damage has occurred to the hair cells (Kurabi et al., 2017). Research has found an arbitrator of this damage is reactive oxygen species (ROS) (Kurabi et al., 2017). These free radicals can chemically interact with cells, and researchers have located free radical species within the cochlea of subjects who have been exposed to loud noise (Kurabi et al., 2017). As well as these free radicals, damage to cochlear neurons has also been found to contribute to noise induced hearing impairment (Kurabi et al., 2017).

Sensorineural hearing impairment is known to result in reduced frequency selectivity, which is the ability of the auditory system to breakdown and distinguish different components in complex sound (De Sousa, Smits, et al., 2020; Moore, 2012). As a result of damage to the cochlea, the filters of the auditory system become broader and are not able to detect differences between frequencies as accurately (Gong et al., 2014). One of the consequences of this is a reduced ability to distinguish speech in a noisy environment, as more noise is able to pass through these broader filters resulting in a reduced signal-to-noise ratio (Moore, 2012).

1.4.2. Conductive Hearing Impairment

When there is an issue with the middle ear, this is known as a conductive hearing loss/impairment (CHL). Conductive hearing impairment can be caused by problems with the middle ear bones or the presence of fluid or masses behind the tympanic membrane, which may impact the ability of sound to move through the middle ear to the cochlea. Conductive hearing impairments can be permanent or temporary depending on their cause. Some common causes are otitis media (ear infections), otosclerosis and other fixations of the ossicles, severe trauma to the head or ear, and malformations of the ossicles (World Health Organization, 2021).

To note, conductive impairments do not result in the same broadening of auditory filters as sensorineural. Although some patients with conductive impairments feel they actually have improved speech discrimination in the presence of background noise, studies have found that subjects with a conductive impairment do in fact have a reduced ability to understand speech in background noise compared to those with normal hearing, which is theorised to be due to long term auditory deprivation (Hsieh et al., 2009). However, while those with a conductive hearing impairment perform worse in background noise to those with normal hearing, research has indicated that they still perform better than those with a sensorineural impairment (De Sousa, Smits, et al., 2020).

1.4.3. Mixed Hearing impairment

A mixed hearing loss/impairment is a combination of a sensorineural and conductive impairments. Some common causes are viral infections such as HIV, measles, and cytomegalovirus, or conditions such as cochlear otosclerosis (otosclerosis that has impacted the cochlea) and superior semicircular canal dehiscence (Abdurehim et al., 2016; Cohen et al., 2014; Katz et al., 2015).

1.4.4. Retrocochlear Hearing Impairment

Retrocochlear hearing loss/impairments are caused by issues with the auditory nervous system itself, with the cochlea and middle ear unaffected. The most common causes are damage to the auditory nerve and central auditory processing system, which send auditory signals from the cochlea to the brain (Shipley & McAfee, 2021). This damage is often due to conditions such as tumours of the eight nerve (vestibular schwannomas), auditory neuropathy spectrum disorder, and intra-axial or extra-axial brainstem disorders (Katz et al., 2015). Retrocochlear hearing impairments often present with fairly normal pure-tone hearing thresholds but unexpectedly poor speech discrimination abilities (Shipley & McAfee, 2021).

2. Audiological Testing Methods

As detailed in the previous sections, identifying and helping those with a hearing impairment has benefits for both society as a whole and the individual affected. This section will detail some of the ways a hearing impairment can currently be identified and discuss their pros and cons.

2.1. Full Diagnostic Puretone Testing

The traditional gold standard method of identifying a hearing impairment is full diagnostic puretone audiometry. This is usually conducted face-to-face in a soundproof room by a qualified audiologist or audiometrist. In this test, an audiometer is used to present different frequencies at different volumes to the listener. To determine how well sound is moving through the ear to the cochlea, calibrated supra-aural headphones or insert earphones are used; a method known as air conduction (AC) testing. An additional bone conduction (BC) test may be done as necessary which investigates how well sound can reach the cochlea via the skull. This test is done by placing a small bone vibrator on the mastoid bone behind the ear, which allows sound from the audiometer to bypass the middle ear and directly stimulate the cochlea. Bone conduction thresholds can be used to distinguish a conductive from a sensorineural hearing impairment. The typical frequencies tested are between 250 – 8000 Hz for air conduction and 500 – 4000 Hz for bone conduction.

As well as this, a full diagnostic session will often include other tests such as speech discrimination (where the listener must repeat back lists of words that take a CVC – consonant-vowel-consonant format), tympanometry to assess the health of the tympanic membrane and middle ear cavity, and acoustic reflex testing to determine if the acoustic pathways are functioning as expected.

While puretone audiometry provides the most complete picture of hearing, it is time and resource heavy, as it must be conducted by a highly trained audiologist or audiometrist in a sound-proof environment with specialised equipment. Such tests can take anywhere from 30 to 60 minutes, and so are not appropriate or cost-effective for mass testing.

2.2. Puretone Hearing Screening

A more basic and quicker version of the face-to-face diagnostic test can be done which obtains air conduction thresholds in a quiet – though not necessarily sound proofed – environment at the main frequencies for speech (500 – 4000 Hz). This screening test is designed to determine if the listener has a hearing impairment in general rather than providing specifics on the type.

The benefit of these screening tests is that they can quickly identify or rule out a hearing impairment and be done in a variety of locations outside the audiology clinic. Those it identifies as having an impairment can be referred to an audiologist to determine the type and scope, while those found to have hearing within normal limits can be discharged.

The downside of such testing is that it does also require a trained specialist and specific equipment to conduct and disadvantages those who live rurally or in locations where audiological testing facilities and practitioners are scarce. As well as this, COVID-19 has seen an increase in the need for physical distancing and has placed restrictions on when and how audiology clinicians and screeners work, making it more difficult for those in need to have their hearing evaluated.

2.3. Automated Screening Methods

As discussed, the gold standard for hearing testing is full diagnostic puretone test, with puretone hearing screenings the next best option for identifying the likely presence or absence of hearing impairment. However, the downsides are that they require trained specialists to conduct them and specialised equipment which must be regularly calibrated. Due to such limitations, research has been conducted around the use of automated, self-administered screening tests which can be conducted from within one's own home.

One of the leading automated tests that has emerged is digit triplet testing (DTT). DTT is a hearing screening method which works by presenting combinations of digits in triplet format (such as '1-3-6') in background noise to the listener either via telephone, mobile device, or computer (Smits et al., 2004).

However, a current drawback of traditional DTTs is that they have a lack of sensitivity for conductive and asymmetrical impairments – which can be missed due to the better-ear compensating in the case of an asymmetrical impairment, and better speech discrimination in background noise for conductive impairments (De Sousa, Smits, et al., 2020; Van den Borre et al., 2021).

However the work of De Sousa, Swanepoel, et al. (2020) showed that the use of antiphasic stimuli could overcome this issue, making such testing a leading contender for large scale hearing screening. In this research project, we seek to replicate the results of De Sousa, Swanepoel, et al. (2020) by investigating if the use of antiphasic stimuli can improve the sensitivity and specificity of the New Zealand English DTT.

3. Digit Triplet Testing

3.1. Introduction

As indicated in above, a promising method of conducting hearing screening tests in a non-evasive, socially distanced, and inexpensive manner is by using digit triplet tests (DTT).

The concept of using digits in background noise as a testing procedure was first explored by Elberling et al. (1989) in the late eighties, who included a digit triplets test as part of their DANTALE speech material set (Elberling et al., 1989). This was later developed into a standalone test by Smits et al. (2004) in the Dutch language which could be conducted over the telephone (Smits et al., 2004). Their test was released nationwide within the Netherlands as a national hearing screening test.

The DTT of Smits et al. (2004) presented triplets in speech-weighted noise monaurally – that is, to each ear individually. Results from uncalibrated home telephones were found to be no different from those measured over a phone in the laboratory at a fixed level of 73 dB A (Smits et al., 2004). A set of 23 digit triplets were chosen at random from a set of 80, with the test lasting around 3 minutes (Smits & Houtgast, 2005). Starting at an initial level of 0 dB SNR (signal-to-noise ratio), a 1-up 1-down adaptive process was utilised where the SNR of the subsequent presentation was increased by 2 dB if the response given was incorrect, and decreased by 2 dB if the participant had gotten the triplet correct (Smits & Houtgast, 2005). Responses were judged correct when all three digits were correctly entered. This enabled them to determine the SNR required for the subject to correctly identify 50% of triplets presented to them. To do the test, listeners would enter the digits they heard back via the telephone keys (Smits & Houtgast, 2005).

Smits & Houtgast recognised at the time that their test was not sensitive to conductive impairments, meaning it was largely restricted to distinguishing between those with normal hearing or a sensorineural impairment. This is because, as discussed, those with a conductive impairment generally have better speech discrimination in background noise than those with a sensorineural impairment due to better frequency selectivity, causing them to perform much better (De Sousa, Smits, et al., 2020; Smits & Houtgast, 2005).

Following the success of the Dutch DTT, Wilson et al. (2006) developed a British-English version. Their initial test presented six sets of digit triplets in background multi-talker speech babble at 14 different presentation levels from 6 to -20 dB with a 2 dB step size over the telephone (Wilson et al., 2006). They later reduced this to three digit triplets presented at levels from 4 to -20 dB with a step size of 4 dB to decrease the test time (Wilson et al., 2006).

A Polish language version was then developed by Ozimek et al. (2009).This used speech-weighted noise created by the superimposition of digit stimuli presented at a level of 70 dB SPL or a level chosen by the participant (Ozimek et al., 2009). They presented 25 trials to each ear separately using a 1-up 1-down adaptive procedure.

A French language DTT was created soon after by Jansen et al. (2010) with 27 trials presented in 65 dB SPL of noise and also scored with a 1-up 1-down procedure and step size of 2 dB (Jansen et al., 2010). Following suit, a New Zealand English version was created by King (2010), with versions in American English by Watson et al. (2012), Finnish by Willberg et al. (2016), South African English by Potgieter et al. (2015), Australian English by Dillion et al. (2016), Persian by Motlagh Zadeh et al. (2020), and Korean by Han et al. (2020).

The majority of these tests only included monosyllabic digits due to fears that digits with more than one syllable would be more easily distinguished from the others, with their length being used as a cue (Van den Borre et al., 2021). However, this was not always possible in all cases, with some languages consisting of largely multi-syllable numbers. Speech-weighted noise is most commonly used for the background noise, usually presented at 65 dB SPL.

Table 2: DTT Publications by target language and type of stimuli presntation and scoring method used

Language	Publication	Diotic/Antiphasic	Scoring
Dutch & Dutch-Flemmish	• Smits & Houtgast (2005) • Smits & Houtgast (2007) • Leensen et al. (2011) • Lyzenga & Smits (2011) • Smits et al. (2013) • Kaandorp et al. (2015) • Koole et al. (2016) • De Graaff et al. (2016) • Smits (2017) • De Graaff et al. (2018) • Koopmans et al. (2018) • Denys et al. (2019)	• Diotic • Diotic • Diotic • Diotic • Diotic • Diotic • Diotic • Diotic • Diotic • Diotic • Diotic • Diotic	• Triplet • Triplet • Triplet • Triplet • Triplet • Triplet • Triplet • Triplet • Triplet • Triplet • Triplet • Digit
British English	• Wilson et al. (2006) • Vlaming et al. (2014) • Cullington & Aidi (2017) • Moore et al. (2019)	• Diotic • Diotic • Diotic • Diotic	• Triplet • Triplet • Triplet • Triplet
Polish	• Ozimek et al. (2009)	• Diotic	• Triplet
French	• Jansen et al. (2010) • Jansen et al. (2012) • Ceccato et al. (2021)	• Diotic • Diotic • Antiphasic	• Triplet • Triplet • Triplet
New Zealand-English	• King (2010) • Bowden (2013) • Spence (2020)	• Diotic • Diotic • Diotic	• Triplet • Triplet • Digit
Māori	• Murray (2012) • Bowden (2013)	• Diotic • Diotic	• Triplet • Triplet
American-English	• Watson et al. (2012) • Williams-Sanchez et al. (2014) • Folmer et al. (2017)	• Diotic • Diotic • Diotic	• Triplet • Triplet • Triplet
Finnish	• Willberg et al. (2016)	• Diotic	• Triplet
South African-English	• Potgieter et al. (2015) • Potgieter et al. (2018) • Brown et al. (2019) • De Sousa, Swanepoel, et al. (2020)	• Diotic • Diotic • Diotic • Antiphasic	• Triplet • Triplet • Triplet • Triplet
Australian-English	• Dillion et al (2016)	• Diotic	• Digit
Persian	• Motlagh Zadeh et al. (2020)	• Diotic	• Triplet
Korean	• Han et al. (2020)	• Diotic	• Triplet

While early forms of the DTT were distributed over the telephone, as computer technology developed and became more powerful and widely available researchers began to develop tests using personal headphones and computers. With regard to mobile phones, Smits and Houtgast (2005) found that landlines produced better results than mobile devices due to better sound quality (Smits & Houtgast, 2005). A later study by Brown et al. (2019) observed better results when using a mobile phone within a soundproof booth, and research by Jansen et al. (2010) found the use of headphones provided slightly better results than testing over the telephone (Brown et al., 2019; Jansen et al., 2010). Other research found that any headphone or even earbuds could be used, as even low-quality devices provided steep psychometric curves and high test reliability (Potgieter et al., 2016). Due to this headphones have become the most commonly used method for administering the test.

3.2. SRT Testing

DTT is considered a form of speech recognition threshold (SRT) testing, with a person's SRT being the signal-to-noise ratio (SNR) required for them to recognise 50% of speech in the presence of background noise (Smits et al., 2009). In the case of DTTs, the speech signal takes the form of numbers or digits. Digits were chosen by early developers because they are extremely common, easily recognisable, less likely to be memorised and remembered by participants when repeated for retest purposes, and enabled testing to be done remotely on telephones, with participants able to use their telephone dial or keypad to enter back the digits they heard (Smits et al., 2009).

To obtain a participants SRT, DTTs typically present sets of three digits (digit triplets) in background noise. They repeat a set number of trials, which on average are between 23 and 30 (Van den Borre et al., 2021).

It is generally accepted that the more trials a test uses the more reliable the results will be, with test reliability decreasing with $\sqrt{n}$ trials (Van den Borre et al., 2021). However, the longer the test the more likely the participant is to give up, meaning there is a trade-off between accuracy and time efficiency.

3.3. PTA and Cut-Off Values

DTTs have an experimentally determined cut-off value which they use to distinguish normal hearing from hearing impaired participants. This cut-off value is determined during test development by comparing the PTA_n of participants with the SNRs obtained during testing and constructing an ROC (receiver operating characteristic) curve (described in Section 3.6.1).

The PTA_n criterion most commonly used is $PTA_{0.5,1,2,4\ kHz}$ which covers the main frequencies for speech discrimination (Van den Borre et al., 2021). However, some versions have adopted different criteria to focus on identifying particular forms of hearing impairment.

For instance, a $PTA_{2,4\ kHz}$ has been found to produce the best correlation between PTA and SRT for noise induced impairments (Leensen et al., 2011). Higher frequency PTAs such as $PTA_{2,3,4,6\ kHz}$ and $PTA_{3,4,6,8\ kHz}$ have also been used in some studies focusing on identifying high frequency hearing impairments such as presbycusis (Jansen et al., 2014; Vlaming et al., 2014). How well the PTA_n correlates to the SRT is therefore dependent on the population that is being tested, as a PTA_n of $PTA_{3,4,6,8\ kHz}$ would not correlate as well as $PTA_{0.5,1,2,4\ kHz}$ for a population consisting of a large number of low frequency hearing impairments (such as those caused by Meniere's disease) but would for one mainly consisting of high frequency impairments.

There can also be a variation in whether the binaural PTA or the better-ear or poorer-ear PTA is adopted. When conducting DTT testing using antiphasic stimuli – discussed in more detail in Section 3.7 – studies have shown SRTs to correlate more strongly to the poorer-ear PTA (De Sousa, Swanepoel, et al., 2020).

3.4. Adaptive Procedure

While various methods (such as fixed-SNR procedures) have been explored, most DTTs use adaptive procedures for scoring (Smits, 2017). Early versions adopted a 1-up 1-down staircase method which produced a SRT corresponding to a score of 50% for identification of triplets as a whole and 79% for individual digits in a triplet (Smits et al., 2009). The adaptive procedure works by either increasing or decreasing the intensity of the SNR based on the result obtained in the previous trial. If the participant does not correctly identify the triplet in one trial, then the SNR intensity is increased in the next trial, and conversely if they do guess it correctly the intensity is decreased. The average of the signal intensity from the trials is used to obtain a SRT for participants, which corresponds to the point where they correctly identify the triplets 50% of the time (Leek, 2001).

Some researchers have also adopted a 1-up 2-down procedure or focused on obtaining a score relating to the correct number of individual digits in a triplet, discussed in more detail in section 3.8 (Van den Borre et al., 2021).

As it is expected that the initial trials will be well above the SRT, it is common for the first several trials to be excluded from the calculation of the average signal intensity, and as such most DTTs calculate a final score from the last 20 SNRs obtained (Van den Borre et al., 2021).

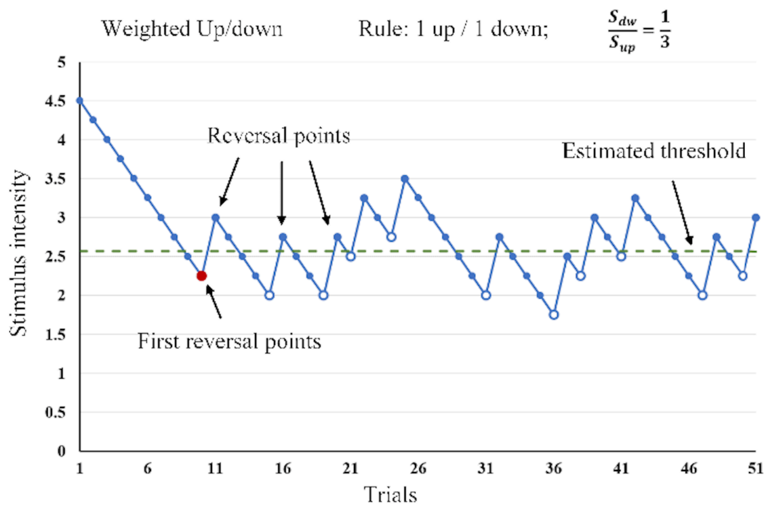


Figure 4: Example of a 1-up 1-down Staircase Adaptive Procedure (Alsaeedi & Wloka, 2021)

3.5. Psychometric functions and Normalisation

Psychometric functions can be used to illustrate a person's ability to identify speech in background noise by plotting speech discrimination in noise as a function of its intensity in dB (MacPherson & Akeroyd, 2014). Researchers use psychometric functions to examine the intensity level required for participants to get 50% of the digit triplets correct and the rate at which their performance improves when the intensity level of the digits is increased, indicated by the slope of the curve (MacPherson & Akeroyd, 2014).

A steep curve indicates that when the SNR is only slightly increased or decreased there is a significant increase in speech intelligibility (MacPherson & Akeroyd, 2014). This is largely dependent on the sensitivity of the test to hearing impairment, with steeper test slopes indicating better tests. A review by Van den Borre et al. (2021) found that psychometric functions associated with DTTs on average produce curves which range between 15%/dB an 20%/dB, although some researchers have achieved higher (Van den Borre et al., 2021).

Psychometric functions are also generally used during the development stage of DTTs as part of a normalisation process to ensure all digits that are used are equally difficult for participants to identify (Van den Borre et al., 2021). Most DTTs have been found to have measurement errors of 0.9 dB to 2.2 dB (Van den Borre et al., 2021).

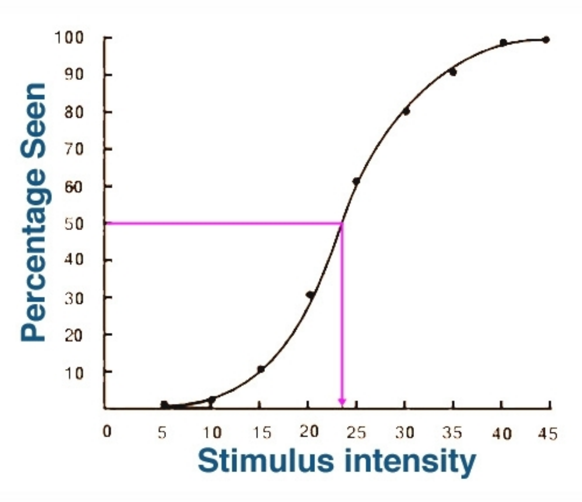


Figure 5: A Psychometric Function (Kalloniatis & Luu, 2005)

3.6. Sensitivity and Specificity

Researchers seek to have tests with a high sensitivity and specificity. In the case of the DTT, sensitivity relates to the proportion of people correctly identified as having a hearing impairment (the true positive rate), while specificity relates to the number correctly identified as having normal hearing (the true positive rate) (Van den Borre et al., 2021).

Sensitivity and specificity for the detection of a condition will obviously depend on how that condition is defined, and can be impacted by factors such as the frequencies tested and what puretone thresholds are used as the cut-off for normal hearing (e.g. > 20 dB HL or > 25 dB HL) (Van den Borre et al., 2021). However, most DTTs have a sensitivity and specificity above 80% even when using stricter definitions such as > 20 dB HL (Van den Borre et al., 2021).

As mentioned, one issue researchers have identified with traditional DTT tests using diotic stimuli is that they are not as sensitive to symmetrical hearing impairments as a participant can use their better hearing ear to obtain a pass. Also, they are considered completely insensitive to conductive hearing impairments. As will be discussed in Section 3.7, the use of antiphasic stimuli has been shown to overcome these issues and improve sensitivity.

3.6.1. ROC and UAC

A plot of sensitivity (the true positive rate) verses 1-specificity (the false positive rate) is known as a receiver operating characteristic (ROC) curve (Cho et al., 2021; Hajian-Tilaki, 2013).

In the case of DTTs, when a test determines a person to have a hearing impairment, if an impairment is confirmed by looking at the puretone average (PTA) from their audiogram this is considered a true positive. If it is found that they do not in fact have a hearing impairment, this is deemed a false positive. When the test indicates a subject has normal hearing, but their PTA indicates they have a hearing impairment, this is considered a false negative. Then in cases where they do in fact have normal hearing, this is called a true negative.

Table 3: NZ-DTT Test Outcomes

Test Result	**Hearing Status**	
	Hearing Impaired (HI)	**Normal Hearing (NH)**
Refer	True Positive (TP)	False Positive (FP)
Pass	False Negative (FN)	True Negative (TN)

ROC curves are used to determine if DTTs are able to distinguish between subjects who have a certain condition – such as normal hearing (NH) or hearing impairment (HI). The point where the sensitivity and specificity are greatest is known as the cut-off point, and this value can be used to classify whether a subject has an impairment or not (Cho et al., 2021). For DTTs, the cut-off point is a signal-to-noise ratio value in decibels (SNR dB) used to classify whether a subject has normal hearing or an impairment.

The area under the curve (AUC) can also be used to indicate the accuracy of the test. The closer the AUC value is to 1, the higher the chance that the model is accurately distinguishing between subjects who have an impairment and those who do not (Hajian-Tilaki, 2013).

The original New Zealand English DTT developed with the diotic presented stimuli, discussed more in Chapter 4, had a sensitivity of 100% and specificity of 85% (King, 2011).

3.7. Diotic verses Antiphasic Stimuli

3.7.1. Introduction

While early DTTs delivered the digit triplet signal in phase binaurally – known as a diotic presentation – there have been recent publications adopting antiphasic presentations.

Antiphasic is the name given when the stimuli are presented 180 degrees out of phase to one ear. An example can be seen in Figure 7. Using an antiphasic presentation has been found to improve the ability of a subject to detect the stimuli when background noise is present (Brown & Musiek, 2013).

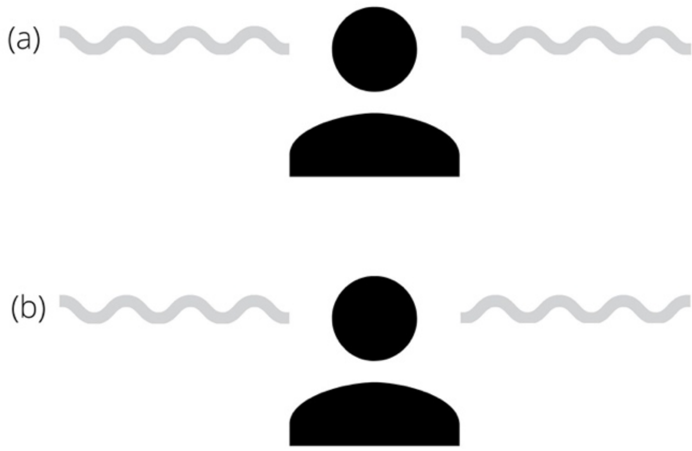


Figure 6: (a) Diotic, Stimuli Presented in Same Phase to Both Ears, (b) Antiphasic, Stimuli Presented Out of Phase 180° Degrees

The difference between detection ability with diotic and antiphasic presented stimuli is known as the Masking Level Difference (MLD) or Binaural Masking Level Difference (BMLD) (Wilson et al., 2003).

Test conditions relating to MLD from Brown & Musiek (2013) are shown in Figure 8 below, where ‘S’ is the signal, ‘N’ is the masking noise, ‘m’ is a monaural presentation, and ‘o’ refers to the signal being in phase or diotic, and ‘π’ refers to the signal being 180 degrees out of phase (Brown & Musiek, 2013).

Diotic stimuli can be defined as stimuli that is presented to both ears in the same relative phase, or the SoNo condition. When the phase of the signal is inverted in one ear but not the other, this is known as antiphasic or the SπNo condition. The BMLD is found by subtracting the SπNo value from the SoNo (Wilson et al., 2003).

An early study by Quaranta and Cervellera (1974) found average MLD values of 8.2 dB for normal hearing subjects, 8.1 dB for those with a symmetrical conductive impairment, 7.9 dB for those with an asymmetric conductive impairment, and 5.7 dB for those with a sensorineural impairment (Quaranta & Cervellera, 1974). Olsen et al. (1976) obtained similar results of 11 dB for normal hearing subjects and 8 dB for those with a conductive impairment, but slightly better results of 8.7 – 10.6 dB for those with a sensorineural impairment (Olsen et al., 1976). A more recent experiment by Wilson et al. (2003) found a MLD of ~13 dB between the SoNo and SπNo condition for 28 normal hearing subjects using a 500 Hz puretone as the signal (Wilson et al., 2003), and Ho et al. (2016) also found subjects with a SNHL were able to detect signals at a lower intensity in the SπNo condition (Ho et al., 2016).

Studies have shown that the frequency of the stimulus can impact the magnitude of the BMLD, with a reduction of up to 2-3 dB being found above 1500 Hz (Hall, 2004; Moore, 2012). It has also been noted that continuous masking noise produces greater BMLDs than burst, as well as using narrowband noise for the masker (Hodgson, 2016).

SmNm	Signal and masker in the same ear (monaural)
SmNo	Signal in one ear and noise in phase at two ears
SmNπ	Signal in one ear and noise 180 degrees out of phase at one ear relative to other
SoNo	Signal and noise in phase at two ears (homophasic)
SoNπ	Signal in phase at two ears and noise 180 degrees out of phase in one ear relative to other (antiphasic)
SπNπ	Signal and noise 180 degrees out of phase at two ears (homophasic)
SπNo	Signal 180 degrees out of phase at two ears and noise in phase at two ears (antiphasic)
SmNm	Signal and masker in the same ear (monaural)
SmNo	Signal in one ear and noise in phase at two ears
SmNπ	Signal in one ear and noise 180 degrees out of phase at one ear relative to other
SoNo	Signal and noise in phase at two ears (homophasic)
SoNπ	Signal in phase at two ears and noise 180 degrees out of phase in one ear relative to other (antiphasic)
SπNπ	Signal and noise 180 degrees out of phase at two ears (homophasic)
SπNo	Signal 180 degrees out of phase at two ears and noise in phase at two ears (antiphasic)

Figure 7: Different Test Conditions for Masking Level Differences from Brown and Musiek (Brown & Musiek, 2013)

While these studies indicate overall better performance with antiphasic stimuli, they also illustrate that the BMLD is not as great in subjects with a hearing impairment, and is most reduced when the impairment is asymmetric and the subjects are older (Brown & Musiek, 2013; Hall, 2004).

To understand why this occurs, it is necessary to take a step back and look at how the auditory system functions. The ability to tell where a sound is coming from – known as sound localisation – is of great importance to survival (Hodgson, 2016). Due to this, mammals have developed the ability to determine where in space a sound is localised, and focus attention on a particular sound amongst a myriad of others. In such situations the signal is attentively honed in on and the background noise attenuated by less focus being placed on it, causing the signal to be at the forefront of the listeners attention (Hodgson, 2016). This is often known as the ‘cocktail party effect’, as it allows a person to focus on just one conversation at a party while ignoring others happening around them (Hodgson, 2016).

The human auditory system receives cues to help in the localisation of sounds, one of these being interaural time difference (ITD) which relates to the difference in time a sound takes to reach each ear (Moore, 2012). For instance, a sound coming from a subject’s right would reach the right ear before the left, helping the listener place the sound on their right. When the sound is sinusoidal, a phase difference can occur between the ears which is known as an interaural phase difference (Moore, 2012). This is frequency dependent, with lower frequency sounds producing a greater difference in the phase of the signal at each ear than higher (Moore, 2012).

It has been found that in the antiphasic condition, ITDs are generated which are similar to those produced when spatial separation of the signal and noise is occurring, meaning the listener is better able to hone in on the signal and ignore the noise, improving the SNR (Gilbert et al., 2015).

However, hearing impaired subjects have been found to produce smaller BMLDs than those with normal hearing due to their having timing irregularities that disrupt the interaural phase difference (De Sousa, Swanepoel, et al., 2020; Wolmarans et al., 2021). Because of this, while the SNR will improve substantially in the antiphasic condition for normal hearing subjects, improvements are less pronounced for those hearing impaired.

3.7.2. Use in DTTs

As discussed, traditional diotic DTTs have struggled to identify asymmetric hearing impairments due to the better-ear being able to compensate (Van den Borre et al., 2021). They can be picked up however when each ear is tested individually (“monaurally”), but this results in the test taking twice as long. As well as this, since the presentation level of the signal is quite loud (often beginning at 65 dB SPL), it is often above the air conduction threshold – that is, at a suprathreshold level – for subjects with a conductive hearing impairment. Since there is no cochlear damage with a conductive impairment, once the signal is loud enough to overcome the attenuation caused by their middle ear pathology the subject will present with SRTs close to those with normal hearing (De Sousa, Swanepoel, et al., 2020).

To help overcome the issue of lack of sensitivity (the proportion correctly identified as having a hearing impairment) for conductive and asymmetric hearing impairments in DTTs, De Sousa et al. (2019) adopted stimuli that was antiphasic, with digit triplets being "phase inverted (antiphasic) between the ears, while leaving the masking noise interaurally in-phase" (De Sousa, Swanepoel, et al., 2020, p. 443). Their research found that antiphasic stimuli could increase sensitivity not only for conductive and symmetric hearing impairments but sensorineural as well, and is discussed more in Chapter 5 (De Sousa, Swanepoel, et al., 2020).

3.8. Triplet verses Digit Scoring

The DTTs mentioned so far rely on the entirety of the digit triplet presented to the listener being identified correctly. However, Denys et al. (2019) have investigated the benefit of scoring each individually recognised digit from within the triplet separately.

In the standard triplet scoring test, there is a 50% probability of recognition for each triplet, as one either correctly identifies the triplet or they do not. This corresponds to a recognition probability of 79% for each individual digit in the triplet, as $(0.7937)^3 = 0.5$. Depending on if the triplet is correctly identified, then the SNR is reduced or increased by a pre-determined step size using an adaptive procedure, as discussed in Section 3.4.

In Denys et al. (2019), the authors compared triplet scoring against digit. They used a standard triplet scoring procedure based around the recognition probability of 50% for each triplet and 79% for each digit with a 1-up 1-down 2 dB step size adaptive procedure that calculated SRT based on the last 21 trials (Denys et al., 2019). They compared this against three different digit scoring procedures which were aimed at targeting recognition probabilities of 79%, 57%, and 35% respectively using adaptive step sizes that depended on the number of digits correctly identified (Denys et al., 2019). In order to provide the same rate of descent as the triplet scored test, they adopted a recognition probability of 50% for the first 6 trials (Denys et al., 2019). They found 79% to have the highest precision, as lower probabilities resulted in more correct guesses (Denys et al., 2019).

They concluded that the digit scoring method provided increased test reliability. However, triplet scoring has been shown to result in a steeper psychometric function and has increased sensitivity (Van den Borre et al., 2021).

4. The New Zealand English DTT

4.1. Test Development

The original New Zealand English DTT was developed by King (2011) in conjunction with Professor Greg O'Beirne of the University of Canterbury, Christchurch, New Zealand. To create the test, they recorded a native New Zealand speaker vocalising digits. These recordings were then used to create 336 triplet combinations, none of which contained any repetition of the same digit (e.g. '2-2-1') (King, 2011).

To create noise that was spectrally identical to the speech signal, they superimposed the recordings collected of the native speaker on each other 10,000 times using an automated process (King, 2011). Due to this, filtering of the test signal resulted in the SNR remaining the same.

The recorded digits were subjected to two normalisation processes to ensure all digits were equally intelligible in noise. They produced psychometric functions for each individual digit and then combined the digits which produced the steepest functions into sets of triplets and evaluated the slopes of the psychometric functions produced (King, 2011). For the individual digits they obtained a mean slope of 16%/dB $\pm$ 5.6%/dB with a measurement error of 2.2 dB, and for the triplet slope they obtained a hypothetical average of 18.7%/dB with an average standard deviation of $\pm$ 1.7 dB (King, 2011). They also obtained the slopes of ten DTT lists to identify if there were any differences between them. To obtain the steepest slope for each triplet list, they identified which digits in each position of the digit (i.e. first digit in the triplet, second, or third) produced the steepest and shallowest slopes. They then used a software programme to construct the lists such that the digit which had the steepest slope in each position of the triplet would occur 75% more frequently than the hypothetical average and conversely the digit with the shallowest slope would occur 75% less frequently. The psychometric functions of the lists resulted a mean slope value of 17.3%/dB with a standard deviation of $\pm$ 3.9% (King, 2011).

To evaluate the DTT itself, 73 participants completed the test. They were presented with 168 digit triplet iterations constituting 27 trials with an average test length of 3 minutes and 30 seconds (King, 2011). Participants were presented the digit triplets separately in each ear and then binaurally, with the order of presentation random.

Trials were delivered by a computer with a graphical interface via the UCAST software and participants entered the digits they heard back via a touch screen (King, 2011). The test used an adaptive 1-up 1-down adaptive procedure with a 2 dB step size. The SNRs of the first 7 trials were disregarded, and a participant's SRT calculated based on the average of the last 20 SNR values obtained (Spence, 2020). Triplets were scored as correct if all three digits were correctly identified.

Participants also underwent a puretone hearing test where their thresholds were obtained. King adopted the best average threshold for each ear across the tested frequencies as the PTA criterion, and found a significant relationship with the DTT results ($r^2 = 0.6539$). King obtained a sensitivity of 100% and specificity of 85% for the diotic triplet test, with the cut-off value for normal hearing set at -8.40 dB SNR and the cut-off for poor hearing at -10.30 dB SNR (King, 2011).

4.2. Modifications

Modifications to improve the original New Zealand English DTT have been attempted by Bowden (2013) and Spence (2020). Bowden (2013) notably improved the digit triplet list by creating a more equal distribution of digits in each position across the triplets. Using software created by O'Beirne (2012), 8 new lists were created with more homogenous digit distribution and triplet slopes. This resulted in each individual digit occurring 26-28 times in each position within a triplet as opposed to 7-62 times per the original version (Bowden, 2013). As a result Bowden was able to improve the sensitivity and specificity of the New Zealand English DTT to 94% and 88% (Bowden, 2013).

In an attempt to improve on the commonly-used 1-up 1-down adaptive procedure, Spence (2020) examined whether implementing the A1 adaptive procedure of Brand & Kollmeier would improve the efficiency of the test. He compared the standard scoring method – which took the average of the SNR data from the last 20 trials – against digit and triplet scoring methods that estimated scores by fitting the test data to the psychometric function using a nonlinear least-squares fit (Spence, 2020).

In the standard triplet scoring method, the scores of 0, 1, 2, and 3 correct digits are 0, 0, 0, and 1 respectively. For digit scoring, the equivalent is 0.00, 0.33, 0.67, and 1.00. Due to the fact that greater sensitivity has been found using triplet scoring, Spence converted the digit scores (0.00, 0.33, 0.67, 1.00) into their triplet equivalents (0.00, 0.00, 0.00, 1.00) after obtaining them using the Brand and Kollmeier A1 approach and fitted a psychometric function to the data to calculate the SRT (Spence, 2020). The A1 approach was found to give a higher AUC than the traditional 1-up 1-down method, and at an earlier point in the adaptive tract, but the ROC curves were not clean enough to enable conclusive determinations of relative test accuracy.

5. Antiphasic DTTs

In this chapter, the development of antiphasic versions of the DTT is explored. Particular focus is given to the work of De Sousa, Swanepoel, et al. (2020) who were the first to develop such a test and whose approach we seek to emulate in the current research project.

5.1. The South African English DTT

The South African English DTT was first developed by Potgieter et al. (2015) for smartphones (Potgieter et al., 2016). They used 20 normal hearing adults between the ages of 18 and 21 to whom they presented 23 digit triplets in speech-weighted background noise, scoring by triplet using an adaptive 1-up 1-down procedure with a 2 dB step size (Potgieter et al., 2016). A further study by Potgieter, Swanepoel, Myburgh, et al. (2018) tested the suitability of the South African English DTT to be used as a large scale hearing screening test, and they reported a sensitivity of 94% and specificity of 77% for a $PTA_{1,2,4,8\ kHz} > 25$ dB HL (Potgieter et al., 2018).

5.1.1. Antiphasic Test Development

The next major development came from De Sousa, Swanepoel, et al. (2020), who conducted a cross-sectional, repeated-measures study that investigated whether the sensitivity of the South African English DTT could be improved by using antiphasic stimuli. They adapted the smartphone DTT application previously developed and created antiphasic triplets by reversing the phase of the original homogenised diotic digits, accomplished by multiplying one of the channels for each sample by -1, thereby reversing the phase of the signal (not the noise) by 180 degrees (De Sousa, Swanepoel, et al., 2020).

They tested 122 participants, 41 who had normal hearing (PTA $\leq$ 25 dB binaurally), 57 with a symmetric sensorineural impairment, 24 with an asymmetric sensorineural impairment, and later included 23 participants with a conductive hearing impairment. The participants first performed a training test using antiphasic stimuli. Then participants performed an antiphasic version of the test, a diotic version, another antiphasic test for retest purposes, and another diotic test for retest purposes.

Each test presented 23 triplets in broadband speech-weighted noise which was interaurally in-phase. They used an adaptive 1-up 1-down procedure with varied step sizes to locate the SNR at which 50% of the triplets were correctly entered. A step size of 4 dB was used for the first 3 steps, which was then reduced to 2 dB for all subsequent steps.

5.1.2. Effect of Test Condition and Hearing Classification

In their statistical analysis, De Sousa, Swanepoel, et al. (2020) used a repeated-measures analysis of variance (ANOVA) with post-hoc comparisons using a Bonferroni adjustment to investigate the effect of test condition (antiphasic or diotic) and hearing classification on SRT. They observed lower SRTs for all four hearing categories in the antiphasic condition, but found them to be particularly lower for those with normal hearing (De Sousa, Swanepoel, et al., 2020).

They then used an analysis of covariance (ANCOVA) to look at the effects of age and English language ability on the SRT in both conditions. They did not find a significant SRT difference between those with high English competence and low competence for either test condition when controlling for poorer ear PTA ($PTA_{0.5,1,2,4\ kHz}$) and age (De Sousa, Swanepoel, et al., 2020).

They used general linear regression to examine whether the slope of the relation between poorer-ear PTA and SRT was different between the antiphasic and diotic tests. They observed a stronger correlation with poorer-ear PTA for the antiphasic condition (De Sousa, Swanepoel, et al., 2020). The slope of the fitted regression was also significantly steeper in the antiphasic condition.

5.1.3. Sensitivity and Specificity

Sensitivity and specificity was obtained by examination of ROC analysis by comparing the ROC curves of both test conditions. They found higher areas under the curve (AUC) for the antiphasic condition compared with the diotic condition for poorer-ear PTA > 25 dB HL (De Sousa, Swanepoel, et al., 2020).

5.1.4. Findings

De Sousa, Swanepoel, et al. (2020) concluded from their research that for any type of hearing classification the antiphasic test was more sensitive and specific. They noted that in the diotic condition asymmetric and conductive hearing impairments produced SRTs that were more similar to normal hearing participants than those with a sensorineural impairment. For asymmetric impaired listeners this was due to the fact that the better-ear could compensate, and for a conductive impairment the conductive component could be overcome once the stimulus was loud enough due to these participants still maintaining good frequency selectivity in background noise (De Sousa, Swanepoel, et al., 2020). When using antiphasic stimuli however, there was a significant decrease in the SRT of normal hearing participants, allowing them to be more easily distinguished from those with a form of hearing impairment.

5.2. French Antiphasic DTT

In a collaboration with some of the authors of the De Sousa, Swanepoel, et al. (2020) study, Ceccato et al. (2021) sought to develop a French version of the antiphasic DTT test.

As mentioned previously, the first French language DTT was developed by Jensen et al. (2010), which consisted of 27 trials and used an adaptive 1-up 1-down procedure with a 2 dB step size. A Dutch-Flemish and French version was later developed by Jensen et al. (2013) which used a similar procedure and for a $PTA_{0.5,1,2,4\ kHz}$ > 50 dB HL had a sensitivity of 100% and a specificity of 92%.

Ceccato et al. (2021) produced a new DTT with mono and bi-syllabic French digits. As with De Sousa, Swanepoel, et al. (2020), they presented the background noise in-phase with the digits being out of phase in one channel. They had 23 trials and used the same 1-up 1-down varied step size method as De Sousa, Swanepoel, et al. (2020), with the first 3 steps being 4 dB and each subsequent step size 2 dB. The average of the last 19 SNRs were used to calculate the SRT.

As with De Sousa, Swanepoel, et al. (2020), they found that poorer-ear PTA was significantly correlated to the SRT for all types of hearing classification in the antiphasic condition. They found that for PTA > 20 dB HL a sensitivity of 92% and specificity of 86% could be obtained at the cut off -12.9 SRT dB SNR (Ceccato et al., 2021). For a PTA > 25 dB HL, this increased to a sensitivity of 96% and specificity of 93% at a cut off value of -11.7 SRT dB SNR (Ceccato et al., 2021).

6. Methods

6.1. Introduction

As indicated in the previous sections, the purpose of the current research was to determine whether an antiphasic version of the New Zealand English DTT could provide improved sensitivity and specificity compared to the diotic version currently in use in New Zealand. This chapter details the methods used to implement and test the antiphasic New Zealand English DTT.

6.2. Participants

A total of 51 participants between the ages of 18 and 86 participated in the research experiment. 10 participants were aged between 18-25, 5 between the ages 26-25, 10 between the ages 36-45, 5 between the ages 46-55, 6 between the ages 56-65, 8 between the ages 66-75, 6 between the ages 76-85, and 1 between the ages 86-95. The median age was 48. Participants were recruited from advertising on social media and through email using details from the University of Canterbury's Speech and Hearing Clinic patient database. All participants took part voluntarily and the project was approved by the Human Ethics Committee.

Table 4: Frequencies of Hearing Classification

Levels	Counts	% of Total	Cumulative %
Normal	25	49.0 %	49.0 %
Symmetric SNHL	14	27.5 %	76.5 %
Asymmetric SNHL	6	11.8 %	88.2 %
CHL	6	11.8 %	100.0 %

Of the participants, 25 presented with normal hearing, 14 with a symmetrical sensorineural hearing impairment, 6 with an asymmetrical sensorineural hearing impairment, and 6 with a conductive hearing impairment.

Normal hearing was defined as a puretone average (PTA) (the average of the hearing thresholds 500, 1000, 2000, and 4000 Hz) of ≤ 25 dB HL (considered within the normal range of hearing) obtained by air conduction (AC) using insert earphones or headphones. Sensorineural hearing impairment was defined as the presentation of a PTA of ≥ 25 dB HL accompanied by bone conduction (BC) thresholds (obtained using a bone conductor) < 20 dB HL of the AC. Conductive hearing impairment was defined as having have a worse-ear PTA of ≥ 25 dB HL accompanied by bone conduction (BC) thresholds ≥ 20 dB HL of the AC. Lastly, an asymmetrical sensorineural hearing impairment was defined as the poorer-ear PTA < 25 dB HL and a difference in PTA of 10 dB between the poorer-ear and better-ear.

6.3. Materials and Apparatus

Audiometry was conducted using GSI Audiostar Pro, GSI, and digital AC40 audiometers. Air conduction thresholds were obtained using E-A-RLINK foam tips and TDH-39 supra-aural headphones. Bone conduction thresholds were obtained using a Radioear B71 bone vibrator.

To perform the New Zealand English DTT, an HP elitebook laptop was utilised with a Soundblaster X-Fi Surround 5.1 Pro USB external sound card and Sennheiser HD 280 Pro headphones. Testing was conducted in soundproof rooms in the University of Canterbury’s Speech and Hearing clinic.

6.4. Procedure

Testing was done in half-hour to one hour sessions in soundproof rooms on the University of Canterbury campus. All participants were given an information sheet and consent form to read and sign. Following this, otoscopy was performed to ascertain the health of the outer ear and tympanic membrane.

Tympanometry was then performed to determine if the tympanic membrane was moving in the expected fashion. Participants without a recent audiogram (obtained within the last year) were then given a puretone hearing test involving air conduction at 250, 500, 1000, 2000, 3000, 4000, 6000, and 8000 Hz and bone conduction at 500, 1000, 2000, and/or 4000 Hz where AC thresholds were equal to or greater than 25 dB HL. Where recent audiograms were available, puretone testing was skipped. The results of the hearing assessment were then explained to the participant and any questions they had regarding the test results answered.

Puretone data was used to obtain right $PTA_{05,1,2\ kHz}$, right $4FA_{05,1,2,4\ kHz}$ (four frequency average), right $ABG_{05,1,2,4\ kHz}$ (air bone gap), left $PTA_{05,1,2\ kHz}$, left $4FA_{05,1,2,4\ kHz}$, left $ABG_{05,1,2,4\ kHz}$, poorer-ear $PTA_{05,1,2\ kHz}$, poorer-ear $4FA_{05,1,2,4\ kHz}$, poorer-ear $AVE_{025,05,1,2,3,4,6,8\ kHz}$ (250-8 kHz average), and better-ear $AVE_{025,05,1,2,3,4,6,8\ kHz}$.

Participants then listened to DTT lists in the UC Adaptive Speech Test Platform (UCAST) software. The lists for each participant were allocated according to a latin-square design. There were a total of 8 lists, each of which contained 27 digit triplets. The noise component of the stimulus was presented at a calibrated level of 65 dB SPL.

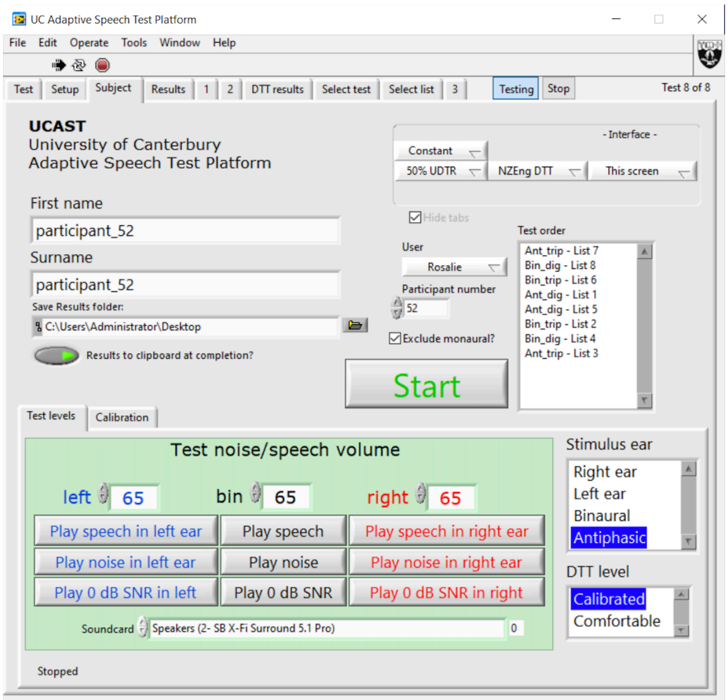


Figure 8: UCAST Platform

Participants completed eight test conditions in an order that was also allocated according to a latin-square design to avoid any order effects. The conditions were: diotic triplet scoring, diotic triplet scoring retest, antiphasic triplet scoring, antiphasic triplet scoring retest, diotic digit scoring, diotic digit scoring retest, antiphasic digit scoring, and antiphasic digit scoring retest. The retest conditions were presented after all the test conditions had been completed.

Digit scoring was performed in a similar manner to that used in Spence (2020), which saw the possible digit scores (0, 0.33, 0.67, 1) converted to their triplet equivalents (0, 0, 0, 1) and fitted to a psychometric function using a nonlinear least-squares fit. The midpoint of this psychometric function is referred to here as the 'FitSRT'.

Triplet scoring was done using the traditional 1-up 1-down adaptive 2 dB step method. The SRT values of the first 7 trials were not used to calculate the final score, but instead the average of the last 20 SNR values, plus the SNR of a "virtual" triplet to account for whether the final triplet was correct or incorrect. Due to digit scoring using a different adaptive procedure, the collected triplet score data was also fitted to a psychometric function, so that its FitSRT could be compared with that from the digit score data.

6.5. Statistical Analysis

Statistical analysis was done using Jamovi 2.0 and Microsoft Excel for Mac (version 16.56). The sample size of 51 provided a large effect size (Cohen's $f = 1$) with 80% statistical power at a 2-tailed significance of level of 0.05.

Intraclass correlation coefficients (ICC) were calculated to investigate the effect of test repetition on SRT. General linear regression was used to analyse whether the slope of the regression line between poorer-ear PTA and SRT dB SNR differed between digit and triplet scoring, and diotic and antiphasic stimuli. Repeated-measures analysis of variance (ANOVA) with post-hoc comparisons using a Bonferroni adjustment were conducted to compare the effect of scoring method and test condition on SRT. An ANOVA was also done to examine the relationship between hearing classifications in each condition. Partial correlations were used to examine the relationship between poorer-ear PTA and SRT. ROC curves were then used to calculate the sensitivity and specificity at different cut off values for the diotic and antiphasic conditions where poorer-ear PTA > 25 dB HL and poorer-ear PTA > 40 dB HL.

7. Results

7.1. Introduction

Table 5: Participant Data

	Normal	Symmetric SNHL	Asymmetric SNHL	CHL
Number	25	14	6	6
Age Range	18 - 75	26 – 95	36 - 75	36 - 85
Mean Poorer-Ear 4FA	8.18	35.83	36.68	53.55
Diotic				
Triplet Scoring Mean SRT	-11.31	-8.76	-10.20	-9.48
Digit Scoring Mean SRT	-11.62	-9.04	-10.17	-9.10
Antiphasic				
Triplet Scoring Mean SRT	-18.43	-13.61	-12.50	-11.04
Digit Scoring Mean SRT	-18.85	-13.60	-12.45	-12.08

As previously mentioned, participants with normal hearing, a symmetric sensorineural impairment (SNHL), an asymmetric sensorineural impairment (SNHL), and conductive hearing impairments (CHL) were used in this research. It was found that normal hearing participants had a better poorer-ear 4 frequency average (4FA) (M = 8.18, SD = 5.44) than the other populations, which is expected. The data showed that participants with a symmetric SNHL (M = 35.82, SD = 10.70) had a similar poorer-ear 4FA to those with an asymmetric SNHL (M = 36.68, SD = 20.93), and those with a conductive hearing impairment had the highest poorer-ear 4FA (M = 53.55, SD = 22.59).

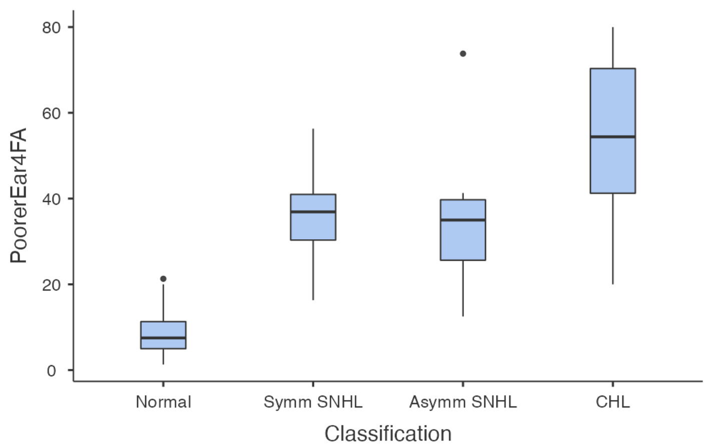


Figure 9: Poorer-Ear 4FA Per Hearing Classification

A box-plot of poorer-ear 4FA and hearing classification identified one outlier for the asymmetric SNHL condition that was not significant. Box-plots of hearing classification in regard to the average test-retest values in the diotic and antiphasic conditions for triplet testing showed one outlier in the diotic condition that was not significant.

7.2. Test verses Retest SRT

Differences between the test and retest SRT for each scoring method and condition were examined by calculating and investigating the intraclass correlation coefficients (ICC). SRT test-retest reliability was found to be high for the diotic digit scoring condition (ICC = 0.76), antiphasic digit scoring condition (ICC = 0.89), diotic triplet scoring condition using the standard scoring method (ICC = 0.79), diotic triplet scoring using FitSRT (ICC = 0.75), antiphasic triplet scoring condition using the standard scoring method (ICC = 0.93), and antiphasic triplet scoring condition using FitSRT (ICC = 0.92), as shown in Tables 6 - 11. Due to these high agreement values, all subsequent analyses were conducted by averaging the test and retest SRT values for each scoring method and condition.

Table 6: Rater Reliability - Diotic Digit Scoring Test-Retest

Intraclass correlation coefficient

	Subjects	Raters	Subject variance	Rater variance	Residual variance	Consistency	Agreement
Value	50	2	2.91	-0.00241	0.913	0.761	0.762

Table 7: Rater Reliability - Antiphasic Digit Scoring Test-Retest

Intraclass correlation coefficient

	Subjects	Raters	Subject variance	Rater variance	Residual variance	Consistency	Agreement
Value	48	2	13.4	0.0273	1.70	0.887	0.886

Table 8: Rater Reliability - Diotic Triplet Scoring Standard Test-Retest

Intraclass correlation coefficient

	Subjects	Raters	Subject variance	Rater variance	Residual variance	Consistency	Agreement
Value	50	2	2.56	0.0969	0.677	0.791	0.768

Table 9: Rater Reliability – Diotic Triplet Scoring FitSRT Test-Retest

Intraclass correlation coefficient

	Subjects	Raters	Subject variance	Rater variance	Residual variance	Consistency	Agreement
Value	50	2	2.59	0.0737	0.682	0.792	0.775

Table 10: Rater Reliability - Antiphasic Triplet Scoring Standard Test-Retest

Intraclass correlation coefficient

	Subjects	Raters	Subject variance	Rater variance	Residual variance	Consistency	Agreement
Value	49	2	12.9	0.0114	1.04	0.926	0.925

Table 11: Rater Reliability - Antiphasic Triplet Scoring FitSRT Test-Retest

Intraclass correlation coefficient

	Subjects	Raters	Subject variance	Rater variance	Residual variance	Consistency	Agreement
Value	49	2	13.6	0.0159	1.22	0.918	0.917

7.3. Digit verses Triplet Scoring

While one of the primary objectives of this research was to determine whether similar results to De Sousa, Swanepoel, et al. (2020) could be found using an antiphasic version of the New Zealand English DTT, a secondary objective was to examine the benefit of digit scoring verses triplet scoring given the results obtained by Denys et al. (2019).

As noted previously, triplet scoring data was converted into a ‘FitSRT’ format which saw the estimation of thresholds by fitting the scoring data to a psychometric function using a nonlinear least-squares fit. This FitSRT data was used to compare digit scoring against triplet scoring.

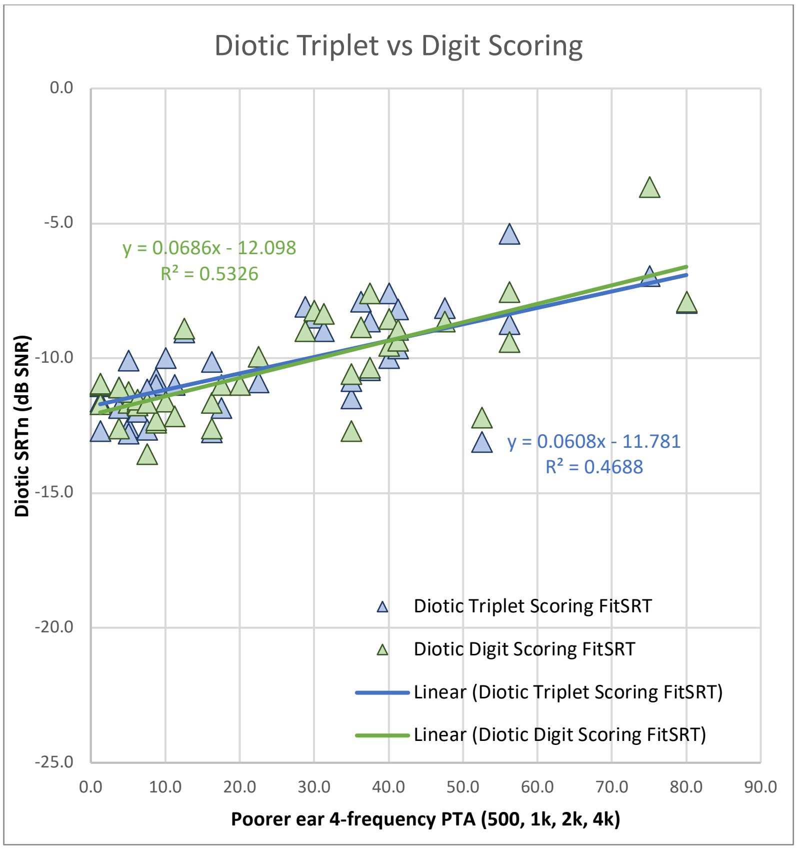


Figure 10: Comparison of the Slope of the Fitted Regression for Digit and Triplet Scoring in the Diotic Condition Against Poorer-ear 4FA

A comparison of the slope of the fitted regression for digit and triplet scoring in the diotic and antiphasic conditions did not show a particular condition produced a steeper line than the other. The regression line slope of the diotic digit condition was 0.069, compared to the 0.061 slope of the diotic triplet condition (Figure 11), while the slope of the antiphasic digit condition 0.16 compared to the 0.15 slope of the antiphasic triplet condition (Figure 12). The test-retest coefficient of determination (R^2) was very similar within each condition, but higher for the antiphasic condition (0.47 and 0.53 for the diotic triplet and digit scoring, and 0.69 and 0.71 for antiphasic triplet and digit scoring).

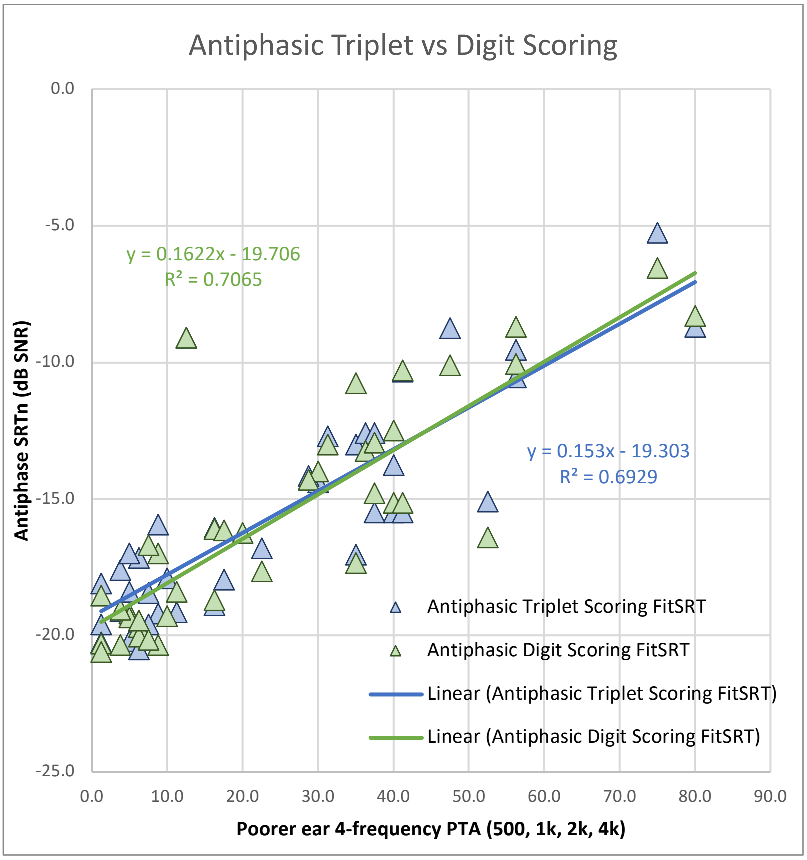


Figure 11: A Comparison of the Slope of the Fitted Regression for Digit and Triplet Scoring in the Antiphasic Condition Against Poorer-ear 4FA

A repeated-measures analysis of variance (ANOVA) was conducted to compare the effect of scoring method on the average FitSRT of the test and retest experiments. Results are shown in Table 12. The observed significant difference between scoring methods ($F(1,48) = 179$, $p = < 0.001$) was driven entirely by the difference in test condition (i.e. diotic or antiphasic), as indicated by post-hoc comparisons using Bonferroni adjustment ($p < 0.001$). It was observed that “diotic digits” verses “triplet diotic’ and “digit antiphasic” verses “triplet antiphasic” did not produce a significant difference ($p = 1.00$).

Table 12: Repeated Measures ANOVA Comparing Effect of Scoring Method on FitSRT

Repeated Measures ANOVA

Within Subjects Effects

	Sum of Squares	df	Mean Square	F	p
Digit vs Triplet	1404	3	468.09	179	< .001
Residual	376	144	2.61		

Note. Type 3 Sums of Squares

Between Subjects Effects

	Sum of Squares	df	Mean Square	F	p
Residual	1346	48	28.0		

Note. Type 3 Sums of Squares

Post Hoc Comparisons - Digit vs Triplet

Comparison							
Digit vs Triplet		Digit vs Triplet	Mean Difference	SE	df	t	$p_{bonferroni}$
Digit Diotic	-	Digit Antiphasic	5.380	0.381	48.0	14.130	< .001
	-	Triplet Diotic	-0.127	0.147	48.0	-0.863	1.000
	-	Triplet Antiphasic	5.196	0.371	48.0	13.995	< .001
Digit Antiphasic	-	Triplet Diotic	-5.506	0.398	48.0	-13.823	< .001
	-	Triplet Antiphasic	-0.184	0.155	48.0	-1.186	1.000
Triplet Diotic	-	Triplet Antiphasic	5.322	0.391	48.0	13.621	< .001

While it was seen that there was no overall difference between the digit and triplet scoring methods, additional analysis was conducted to determine if the digit scoring condition was better able to identify specific forms of hearing impairment. A box-plot analysis was first done comparing digit and triplet scoring in both test conditions and grouping the results by hearing classification. This displayed a difference between test conditions – i.e. whether the test used antiphasic or diotic stimuli – but no significance difference between scoring method.

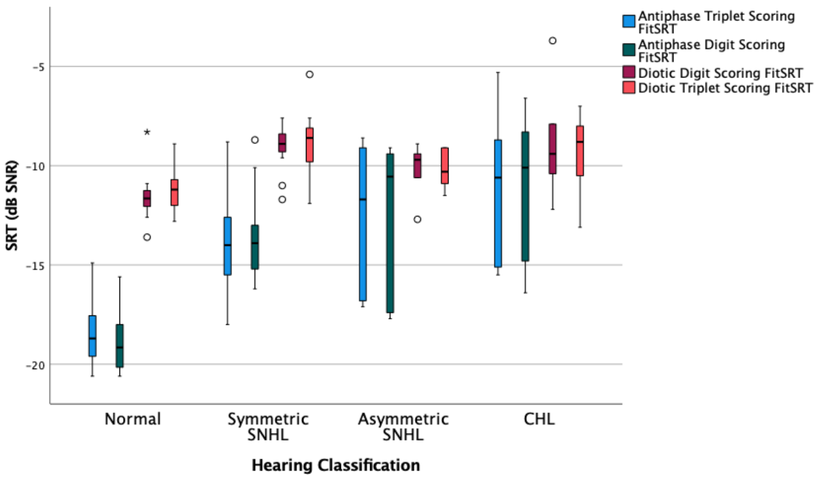


Figure 12: SRT (dB SNR) for each Hearing Classification in the Antiphase Triplet Scoring, Antiphase Digit Scoring, Diotic Triplet Scoring, and Diotic Digit Scoring Conditions

Further analysis was done using a repeated-measures ANOVA first in the diotic condition then the antiphasic, shown in Tables 13 and 14. In the diotic condition the effect of scoring method was not found to be significant (F(1,45) = 0.127, p = 0.72). There was also no significant interaction between scoring method and hearing classification (F(3,45) = 1.59, p = 0.20). With regard to between subject effects, a significant relationship was observed (F(3,45) = 14.1, p < 0.001), but post-hoc Bonferroni testing showed that this was due to the desired differences in SRT between each hearing classification in general, as has been previously established. There was no significant difference in SRT scores (p = 1.00) for normal hearing, symmetric SNHL, asymmetric SNHL, and conductive hearing impaired participants for triplet and digit scoring in the diotic condition.

Table 13: Repeated-measures ANOVA Analysing Effect of Scoring on Hearing Classification in the Diotic Condition

Repeated Measures ANOVA - Diotic Condition

Within Subjects Effects

	Sum of Squares	df	Mean Square	F	p	η^2_p
Scoring (Diotic)	0.0648	1	0.0648	0.127	0.723	0.003
Scoring (Diotic) ✻ Classification	2.4274	3	0.8091	1.593	0.204	0.096
Residual	22.8604	45	0.5080			

Note. Type 3 Sums of Squares

Between Subjects Effects

	Sum of Squares	df	Mean Square	F	p	η^2_p
Classification	137	3	45.64	14.1	< .001	0.485
Residual	145	45	3.23			

Note. Type 3 Sums of Squares

A repeated-measures ANOVA was then conducted for the antiphasic condition. In this condition the effect of scoring method was also not found to be significant ($F(1,45) = 0.127$, $p = 0.52$). There was also no significant interaction between scoring method and hearing classification ($F(3,45) = 1.59$, $p = 0.73$). Post-hoc Bonferroni testing also showed there was no significant difference in SRT scores for normal hearing ($p = 0.76$), symmetric SNHL ($p = 1.00$), asymmetric SNHL ($p = 1.00$), and conductive hearing impaired ($p = 1.00$) participants for triplet and digit scoring in the antiphasic condition.

Table 14: Repeated-measures ANOVA Analysing Effect of Scoring on Hearing Classification in the Antiphasic Condition

Repeated Measures ANOVA – Antiphasic Condition

Within Subjects Effects

	Sum of Squares	df	Mean Square	F	p	η^2_p
Scoring (Antiphasic)	0.250	1	0.250	0.410	0.525	0.009
Scoring (Antiphasic) ✻ Classification	0.788	3	0.263	0.431	0.732	0.028
Residual	27.426	45	0.609			

Note. Type 3 Sums of Squares

Between Subjects Effects

	Sum of Squares	df	Mean Square	F	p	η^2_p
Classification	878	3	292.8	25.9	< .001	0.634
Residual	508	45	11.3			

Note. Type 3 Sums of Squares

7.4. Antiphasic verses Diotic Condition

Due to the finding that digit scoring did not provide significant improvement over triplet scoring, comparisons of antiphasic and diotic triplet data were conducted – unless otherwise stated – using the originally obtained triplet scoring data and not the FitSRT.

A partial correlation on the data from our experiment – controlling for age – found both diotic and antiphasic conditions across all hearing categories were significant correlated with poorer-ear 4FA ($p = 0.003$, $p < 0.001$ respectively). The correlation was found to be stronger in the antiphasic condition ($r = 0.74$) than the diotic ($r = 0.41$), as displayed in Table 15.

Table 15: Partial Correlation of Diotic and Antiphasic Triplet Scoring to Poorer-ear 4FA For All Hearing Classifications

Partial Correlation - All Hearing Classifications

Partial Correlation

		Diotic Triplet Score SRT	Antiphasic Triplet Score SRT	Poorer-Ear 4FA
Diotic Triplet Score SRT	Pearson's r	—		
	p-value	—		
Antiphasic Triplet Score SRT	Pearson's r	0.527	—	
	p-value	< .001	—	
Poorer-Ear 4FA	Pearson's r	0.409	0.744	—
	p-value	0.003	< .001	—

Note. controlling for 'Age'

For SRTs of participants with either normal hearing or CHL, a stronger correlation was found in the antiphasic condition ($r = 0.86$) than the diotic ($r = 0.51$), as per Table 16.

Table 16: Partial Correlation of Diotic and Antiphasic Triplet Scoring to Poorer-ear 4FA for Participants with a CHL or Normal Hearing

Partial Correlation - Normal & Conductive Classification

Partial Correlation

		Diotic Triplet Score SRT	Antiphasic Triplet Score SRT	Poorer-Ear 4FA
Diotic Triplet Score SRT	Pearson's r	—		
	p-value	—		
Antiphasic Triplet Score SRT	Pearson's r	0.675	—	
	p-value	< .001	—	
Poorer-Ear 4FA	Pearson's r	0.515	0.856	—
	p-value	0.004	< .001	—

Note. controlling for 'Age'

Additionally, a significant correlation ($p < 0.001$) between normal hearing or symmetrical SNHL and poorer-ear 4FA was found in our results, with the antiphasic similarly slightly more strongly correlated ($r = 0.69$) than the diotic ($r = 0.52$), as shown in Table 17.

Table 17: Partial Correlation of Diotic and Antiphasic Triplet Scoring to Poorer-ear 4FA for Participants with a Symmetrical SNHL or Normal hearing

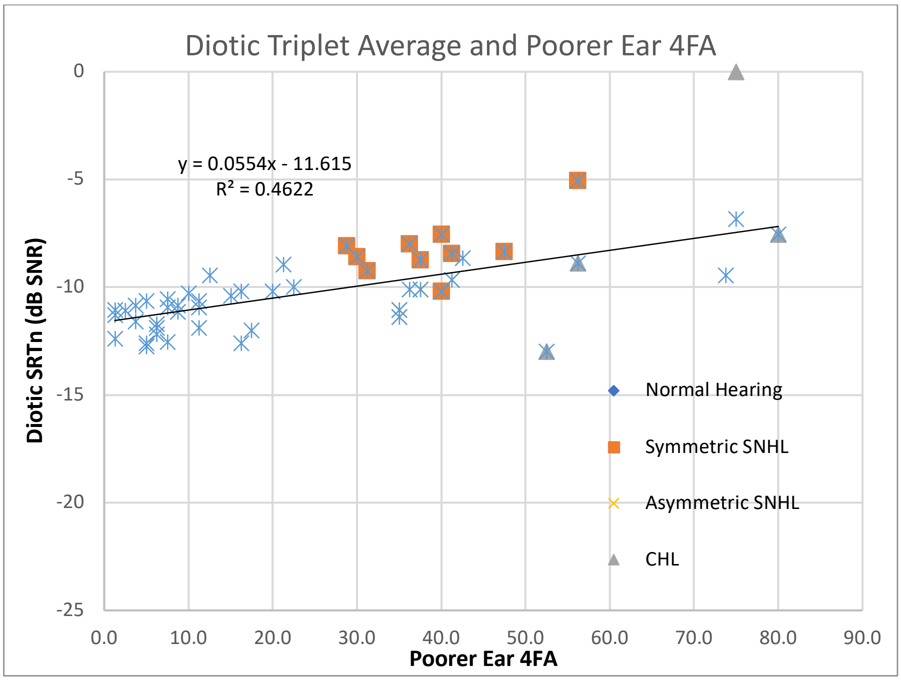
Partial Correlation - Normal & Symmetric SNHL

Partial Correlation

		Diotic Triplet Score SRT	Antiphasic Triplet Score SRT	Poorer-Ear 4FA
Diotic Triplet Score SRT	Pearson's r	—		
	p-value	—		
Antiphasic Triplet Score SRT	Pearson's r	0.521	—	
	p-value	< .001	—	
Poorer-Ear 4FA	Pearson's r	0.562	0.693	—
	p-value	< .001	< .001	—

Note. controlling for 'Age'

No significant correlation was found in our data between each individual hearing classification and poorer-ear 4FA. A significant correlation was found between normal hearing or asymmetric SNHL and poorer-ear 4FA PTA (p = 0.001) in the antiphasic condition (r = 0.55) but no significant correlation (p = 0.59) in the diotic condition (r = 0.10).



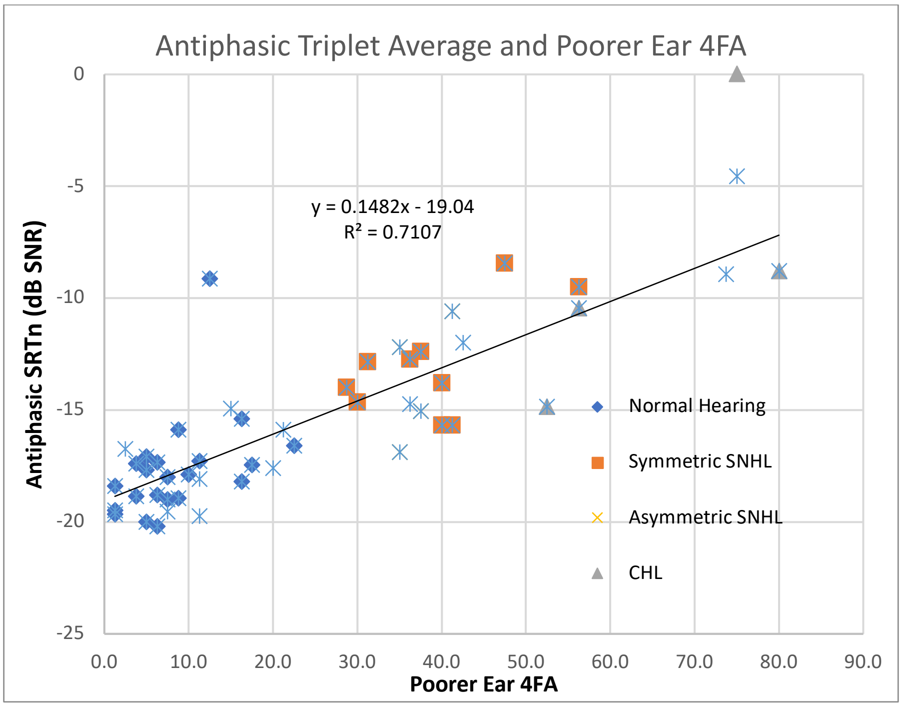


Figure 13: Correlation of Diotic Triplet Scoring (top) and Antiphasic Triplet Scoring (bottom) to Poorer-ear 4FA

The linear regression line for our experiments looking at the correlation between the binaural poorer-ear 4FA and SRT produced regression line values of 0.05 for the diotic condition and 0.15 for the antiphasic (Figure 13).

With regard to each individual hearing classification, the slope was visibly steeper for participants with a symmetrical SNHL in the antiphasic condition (0.18) than the diotic (0.10), and also for an asymmetrical SNHL in the antiphasic condition (0.60) than the diotic (0.25).

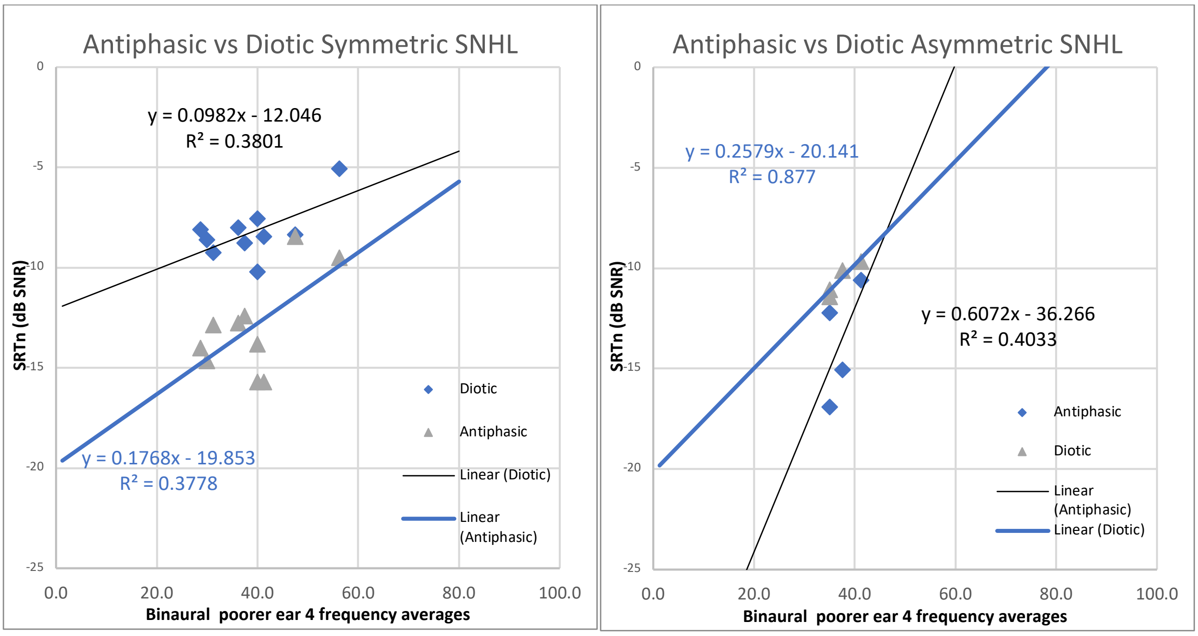


Figure 14: Comparison of Diotic and Antiphasic Triplet Scoring to Poorer-ear 4FA for Participants with a Symmetrical SNHL (left) and Asymmetric SNHL (right)

This was also found with normal hearing participants, who had a steeper slope in the antiphasic condition (0.17) than the diotic (0.05), displayed in Figure 15. While the slope was still steeper in the antiphasic condition (0.32) for participants with a conductive hearing impairment, it was closer to the diotic (0.28) than for the other hearing impairment classifications.

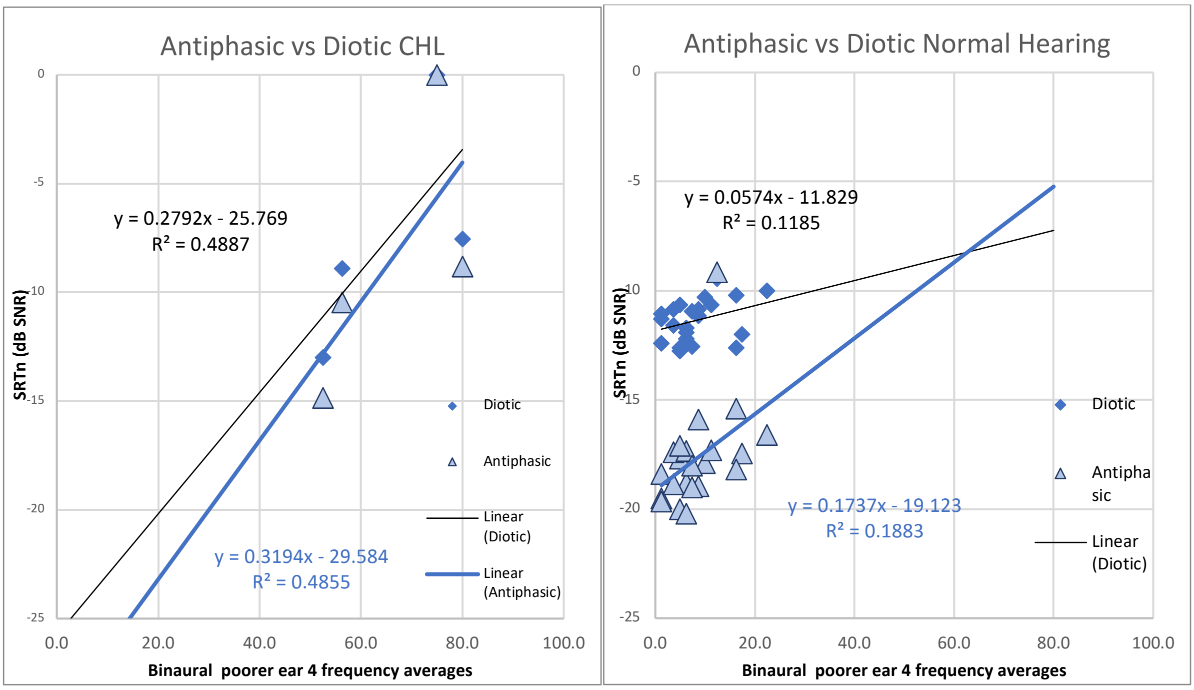


Figure 15: Comparison of Diotic and Antiphasic Triplet Scoring to Poorer-ear 4-frequency PTA for Participants with a CHL (left) and Normal Hearing (right)

7.4.1. Asymmetrical and Conductive Hearing Impairment

We conducted an ANOVA on the diotic triplet SRT to investigate if there was any significant difference between hearing classifications, shown in Table 18. This indicated a significant difference ($F[3,46] = 11.7$, $p < 0.001$). Post-hoc Bonferroni testing illuminated that there was not a significant difference between normal hearing participants and those with an asymmetrical SNHL ($p = 0.42$) in this condition. There was also no significant difference between symmetric SNHL and asymmetric SNHL ($p = 0.23$), symmetric SNHL and CHL ($p = 1.0$), and asymmetric SNHL and CHL ($p = 1.0$).

Table 18: ANOVA Investigating Any Significant Difference Between Hearing Classifications in the Diotic Condition

ANOVA

ANOVA – Diotic Triplet Scoring

	Sum of Squares	df	Mean Square	F	p	η^2p
Classification	60.8	3	20.25	11.7	< .001	0.432
Residuals	79.7	46	1.73			

Post Hoc Tests

Post Hoc Comparisons - Classification

Comparison							
Classification		Classification	Mean Difference	SE	df	t	$p_{bonferroni}$
Normal	-	Symmetric SNHL	-2.479	0.439	46.0	-5.642	< .001
	-	Asymmetric SNHL	-1.108	0.598	46.0	-1.851	0.423
	-	CHL	-2.008	0.645	46.0	-3.113	0.019
Symmetric SNHL	-	Asymmetric SNHL	1.371	0.642	46.0	2.135	0.229
	-	CHL	0.471	0.686	46.0	0.687	1.000
Asymmetric SNHL	-	CHL	-0.900	0.797	46.0	-1.129	1.000

Note. Comparisons are based on estimated marginal means

An ANOVA was then conducted on the antiphasic triplet SRT, shown in Table 19, which also indicated a significant difference (F[3,46] = 22.8, p < 0.001). Post-hoc Bonferroni testing showed that for this condition there was a significant difference between normal hearing participants and those with an asymmetrical SNHL (p < 0.001). There was also a greater difference between normal hearing participants and those with a conductive impairment in the antiphasic data (p < 0.001). There was no significant difference between symmetric SNHL and asymmetric SNHL (p = 1.0), symmetric SNHL and CHL (p = 0.31), and asymmetric SNHL and CHL (p = 1.0).

Table 19: ANOVA Investigating Any Significant Difference Between Hearing Classifications in the Antiphasic Condition

ANOVA

ANOVA – Antiphasic Triplet Scoring Test+Retest

	Sum of Squares	df	Mean Square	F	p	η^2p
Classification	417	3	139.10	22.8	< .001	0.598
Residuals	281	46	6.10			

Post Hoc Tests

Post Hoc Comparisons - Classification

Comparison							
Classification		Classification	Mean Difference	SE	df	t	$p_{bonferroni}$
Normal	-	Symmetric SNHL	-4.82	0.824	46.0	-5.844	< .001
	-	Asymmetric SNHL	-5.93	1.123	46.0	-5.284	< .001
	-	CHL	-7.39	1.210	46.0	-6.110	< .001
Symmetric SNHL	-	Asymmetric SNHL	-1.11	1.205	46.0	-0.925	1.000
	-	CHL	-2.57	1.287	46.0	-2.001	0.308
Asymmetric SNHL	-	CHL	-1.46	1.495	46.0	-0.976	1.000

Note. Comparisons are based on estimated marginal means

7.4.2. ROC Curve Results

As discussed previously, the sensitivity and specificity of a DTT is of high importance to determine the ability of the test to correctly identify people who have hearing impairments from those with normal hearing. To investigate this a plot of sensitivity (the true positive rate) verses 1-specificity (the false positive rate) known as a receiver operating characteristic (ROC) curve was created.

A large area under the curve was obtained in our experiments (AUC = 0.92), with a sensitivity of 95% and specificity of 90% found at a cut-off of -15.9 dB SNR for poorer-ear 4FA > 25 dB HL in the antiphasic condition (displayed in Figure 16).

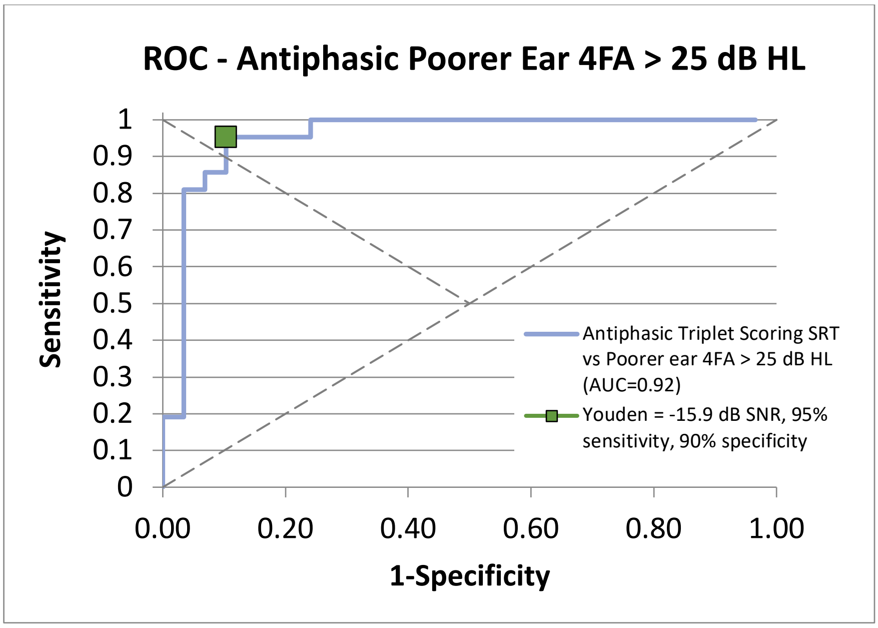


Figure 16: ROC Curve Analysis for Poorer-ear 4FA > 25 dB HL in the Antiphasic Condition

In the diotic condition, the area under the curve was not as substantial (AUC = 0.88), and a sensitivity of 81% and specificity of 90% were obtained at a cut-off of -10 dB SNR for poorer-ear 4FA > 25 dB HL, shown in Figure 17.

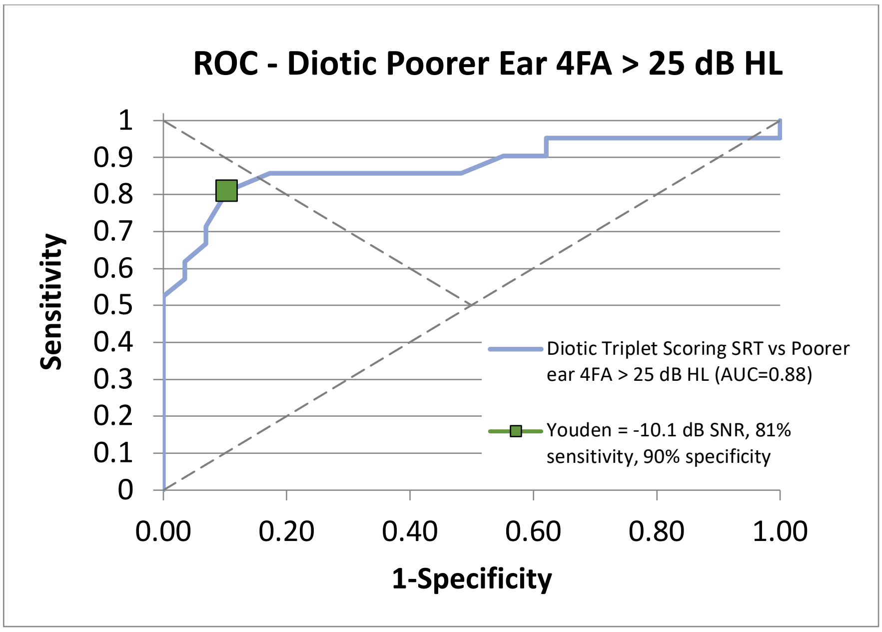


Figure 17: ROC Curve Analysis for Poorer-ear 4FA > 25 dB HL in the Diotic Condition

With regard to poorer-ear 4FA > 40 dB HL, a sensitivity of 100% and specificity of 71% was found at a cut-off of -15.9 dB SNR in the antiphasic condition and a sensitivity of

83% and specificity of 82% for the diotic condition at a cut-off of 10 dB SNR (Figures 17 and 18). The area under the curve was slightly larger in the antiphasic condition (AUC = 0.89) compared to the diotic (AUC = 0.84).

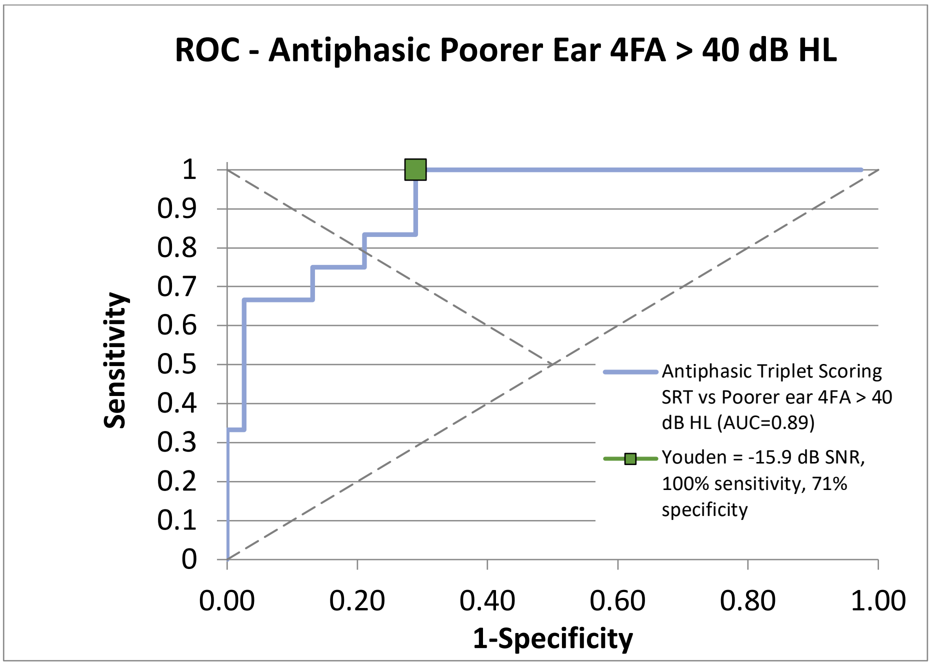


Figure 18: ROC Curve Analysis for Poorer-ear 4FA > 40 dB HL in the Antiphasic Condition

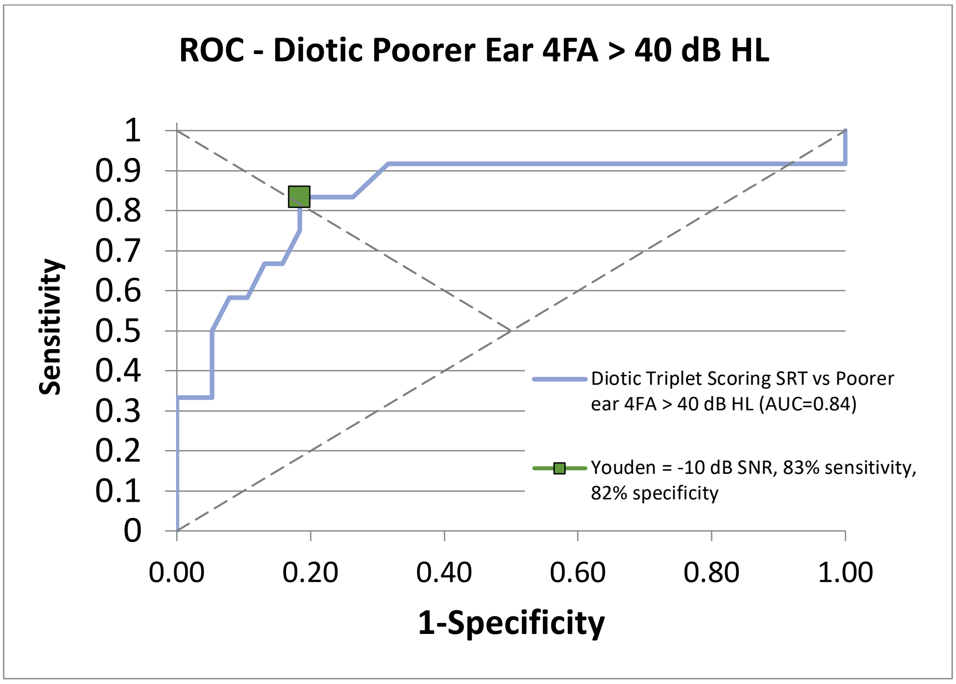


Figure 19: ROC Curve Analysis for Poorer-ear 4FA > 40 dB HL in the Diotic Condition

SRT cut-offs were examined to find that which produced the best trade-off between sensitivity and specificity for each condition at poorer-ear PTA > 25 dB HL and poorer-ear PTA > 40 dB HL, as well as better-ear PTA > 25 dB and better-ear PTA > 40 dB HL, which are shown in Table 20. Best results were found in the antiphasic condition for poorer-ear PTA > 25 dB HL.

Table 20: Poorer and Better-ear Cut-offs in the Antiphasic Condition where PTA > 25 dB HL

	Antiphasic Condition					
	PTA > 25 dB HL			PTA > 40 dB HL		
	SRT dB SNR	Sensitivity %	Specificity %	SRT dB SNR	Sensitivity %	Specificity %
Poorer-Ear	-16.1	95	83	-16.1	100	66
	-15.9	95	90	-15.9	100	71
	-15.7	86	90	-15.7	83	71
Better-Ear	-15.9	100	77	-15.8	100	63
	-15.3	87	80	-15.2	71	65
	-14.8	87	89	-14.9	71	70

Table 21: Poorer and Better-ear cut-offs in the Diotic Condition where PTA > 25 dB HL

	Diotic Condition					
	PTA > 25 dB HL			PTA > 40 dB HL		
	SRT dB SNR	Sensitivity %	Specificity %	SRT dB SNR	Sensitivity %	Specificity %
Poorer-Ear	-10.2	86	83	-10.2	92	68
	-10.1	81	90	-10	83	82
	-9.9	71	93	-9.6	75	82
Better-Ear	-10.1	93	83	-10.2	100	77
	-9.7	87	89	-10.1	93	83
	-9.4	87	97	-9.4	87	97

7.4.3. Effect of Reducing Number of Trials

To investigate the effect reducing the number of triplet trials had on sensitivity, specificity, and the AUC, we conducted a comparison of these elements for poorer-ear 4FA > 25 dB HL and better-ear 4FA > 25 dB HL in both the diotic and antiphasic conditions.

We first analysed digit triplet scoring SRT with regard to poorer-ear 4FA > 25 dB HL for the diotic condition (Table 22). It was observed that the maximal AUC, sensitivity and specificity was found at trial 25 at the Youden cut-off SRT of -10.1, with a AUC of 0.89, sensitivity of 86%, and specificity of 90%.

Table 22: Diotic Triplet Scoring SRT verses Poorer-ear 4FA > 25 dB HL

Diotic Triplet Scoring SRT vs Poorer ear 4FA > 25 dB HL																		
Diotic Trials	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
AUC	0.86	0.86	0.86	0.85	0.87	0.86	0.87	0.87	0.86	0.85	0.88	0.87	0.89	0.89	0.89	0.89	0.89	0.88
Youden Cut-off SRT	-10.0	-9.6	-10.1	-10.0	-10.1	-10.1	-10.0	-10.7	-10.6	-9.4	-10.1	-9.8	-10.1	-10.1	-10.1	-10.1	-10.1	-10.1
Sensitivity	71%	62%	67%	67%	71%	71%	67%	86%	81%	67%	71%	67%	76%	81%	81%	86%	81%	81%
Specificity	86%	97%	90%	93%	90%	90%	93%	79%	79%	97%	93%	97%	93%	93%	93%	90%	93%	90%

Next we analysed digit triplet scoring SRT with regard to better-ear 4FA > 25 dB HL for the diotic condition (Table 23). It was observed that the maximal AUC, sensitivity and specificity was found at trial 23 at the Youden cut-off SRT of -10.1, with a AUC of 0.98, sensitivity of 100%, and specificity of 89%. Note that similar values can be found at trial 21, although here the specificity reduces from 89% to 86%.

Table 23: Diotic Triplet Scoring SRT verses Better-ear 4FA > 25 dB HL

Diotic Triplet Scoring SRT vs Better ear 4FA > 25 dB HL																		
Diotic Trials	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
AUC	0.90	0.91	0.92	0.92	0.95	0.94	0.96	0.96	0.96	0.96	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
Youden Cut-off SRT	-10.0	-9.6	-9.6	-10.0	-10.0	-9.8	-10.0	-9.5	-10.0	-9.4	-10.1	-10.3	-9.4	-10.1	-10.1	-9.4	-9.3	-9.4
Sensitivity	87%	80%	80%	87%	87%	87%	87%	87%	87%	87%	93%	100%	87%	100%	100%	87%	87%	87%
Specificity	83%	94%	91%	91%	91%	91%	91%	94%	91%	94%	91%	86%	97%	89%	89%	97%	97%	97%

We then analysed digit triplet scoring SRT with regard to poorer-ear 4FA > 25 dB HL for the antiphasic condition (Table 24). It was observed that the maximal AUC, sensitivity and specificity was found at trial 25 at the Youden cut-off SRT of -15.9, with a AUC of 0.92, sensitivity of 95%, and specificity of 90%. A similar value with a reduced sensitivity and specificity can be found at trials 12 (cut-off SRT -14.6, AUC 0.94, sensitivity 90%, specificity 90%) and 18 (cut-off SRT -15.2, AUC 0.95, sensitivity 90%, specificity 93%).

Table 24: Antiphasic Triplet Scoring SRT verses Poorer-ear 4FA > 25 dB HL

Antiphasic Triplet Scoring SRT vs Poorer ear 4FA > 25 dB HL																		
Antiphasic Trials	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
AUC	0.94	0.94	0.94	0.89	0.94	0.95	0.94	0.95	0.95	0.95	0.95	0.91	0.91	0.88	0.92	0.92	0.88	0.92
Youden Cut-off SRT	-15.0	-15.2	-14.6	-15.4	-15.2	-15.3	-15.2	-15.4	-15.2	-16.4	-16.5	-16.1	-16.1	-16.0	-16.0	-15.9	-15.8	-15.9
Sensitivity	95%	95%	90%	90%	90%	90%	90%	90%	90%	100%	100%	95%	95%	95%	95%	95%	95%	95%
Specificity	86%	86%	90%	86%	90%	90%	90%	90%	93%	79%	79%	86%	86%	86%	86%	90%	90%	90%

Lastly we analysed digit triplet scoring SRT with regard to better-ear 4FA > 25 dB HL for the antiphasic condition (Table 25). It was observed that the maximal AUC, sensitivity and specificity was found at trial 18 at the Youden cut-off SRT of -15.2, with a AUC of 0.89, sensitivity of 93%, and specificity of 80%.

Table 25: Antiphasic Triplet Scoring SRT verses Better-ear 4FA > 25 dB HL

Antiphasic Triplet Scoring SRT vs Better ear 4FA > 25 dB HL																		
Antiphasic Trials	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
AUC	0.86	0.85	0.87	0.82	0.87	0.87	0.87	0.88	0.89	0.88	0.88	0.86	0.86	0.83	0.87	0.87	0.85	0.87
Youden Cut-off SRT	-15.0	-15.2	-14.6	-15.4	-15.2	-15.3	-15.2	-15.4	-15.2	-16.1	-16.1	-16.1	-16.1	-16.0	-16.0	-15.9	-15.8	-15.9
Sensitivity	93%	93%	93%	93%	93%	93%	93%	93%	93%	100%	100%	100%	100%	100%	100%	100%	100%	100%
Specificity	71%	71%	77%	74%	77%	77%	77%	77%	80%	71%	71%	74%	74%	74%	74%	77%	77%	77%

8. Discussion

The results of our experiment showed that the antiphasic New Zealand English DTT was a more powerful diagnostic tool than the diotic, but scoring by digit did not provide any additional benefit over scoring by triplet.

8.1. Digit verses Triplet Scoring

Denys et al. (2019) found that digit scoring resulted in a higher test-retest reliability than scoring by triplet. However our results showed similar intraclass correlation coefficients for the diotic digit scoring condition (ICC = 0.76), antiphasic digit scoring condition (ICC = 0.89), diotic triplet scoring using FitSRT (ICC = 0.75), and antiphasic triplet scoring condition using FitSRT (ICC = 0.92). The test-retest coefficient of determination (R^2) of the digit and triplet scored regression line slopes were also similar for the diotic digit and triplet scoring conditions (0.47 and 0.53 respectively) and antiphasic digit and triplet scoring conditions (0.69 and 0.71 respectively). This indicates no significant difference in test-retest reliability between digit and triplet scoring.

We also found that with regard to specific hearing classification, digit scoring was no better at distinguishing between normal hearing, symmetric SNHL, asymmetric SNHL or conductive hearing impairment than triplet scoring. Our research found the only differences that presented between digit and triplet scoring occurred when one scoring method was conducted using the diotic condition and the other the antiphasic.

8.2. Diotic verses Antiphasic Stimuli

The results of our experiment are in agreement with those obtained by De Sousa, Swanepoel, et al. (2020). In their research, De Sousa, Swanepoel, et al. (2020) found the slope of the fitted regression was significantly steeper for the antiphasic condition, with their regression line values being 0.07 for the diotic condition and 0.16 for the antiphasic. The linear regression line for our experiments looking at the correlation between the binaural poorer-ear 4FA and SRT also produced a steeper slope for the antiphasic condition (0.15) compared to the diotic (0.05). Similar to De Sousa, Swanepoel, et al. (2020), we conducted a series of repeated-measures ANOVAs with post-hoc comparisons using a Bonferroni adjustment to compare the overall SRT value in each condition as well as for each individual hearing classification. We found that SRT values were lower in the antiphasic condition for all hearing classifications. We also found that in the antiphasic condition normal hearing SRTs were significantly lower than SRTs for symmetrical SNHL, asymmetric SNHL, and conductive hearing impairment, making the antiphasic version better able to distinguish those with normal hearing from those with a hearing impairment.

With regard to overall diagnostic power, De Sousa, Swanepoel, et al. (2020) found larger areas under the curve for the antiphasic condition against poorer-ear PTA. In their results for poorer-ear PTA > 25 dB HL the antiphasic condition produced 90% sensitivity and 80% specificity while the diotic produced 75% sensitivity and 71% specificity (De Sousa, Swanepoel, et al., 2020). For their PTA > 40 dB HL data, the antiphasic condition produced 87% sensitivity and 91% specificity while the diotic resulted in 75% sensitivity and 75% specificity.

Similar improvements with antiphasic stimulation were found in our results. A large area under the curve was obtained in our antiphasic experiments (AUC = 0.92), with a sensitivity of 95% and specificity of 90% found at a cut-off of -15.9 dB SNR for poorer-ear 4FA > 25 dB HL in that condition. The diotic New Zealand English DTT, in comparison, obtained a sensitivity and specificity of 94% and 88% respectively.

In a similar fashion to De Sousa, Swanepoel, et al. (2020), SRT cut-offs were examined to find that which produced the best trade-off between sensitivity and specificity for each condition at poorer-ear PTA > 25 dB HL and poorer-ear PTA > 40 dB HL, as well as better-ear PTA > 25 dB and better-ear PTA > 40 dB HL. We also found that best results were obtained in the antiphasic condition for poorer-ear PTA > 25 dB HL.

One limitation of our research was a relatively low number of participants (*n* = 51). This was roughly one-third of those in De Sousa, Swanepoel, et al. (2020) (*n* = 145) (De Sousa, Swanepoel, et al., 2020). However our results of 95% sensitivity and 90% specificity are comparable with the 90% sensitivity and 80% specificity they obtained. Slightly better results were found by Ceccato et al. (2021), being 96% sensitivity and 93% specificity, who had a higher number of participants (167) than both our research and De Sousa, Swanepoel, et al. (2020) (Ceccato et al., 2021).

Table 26: Sensitivity and Specificity Values for Poorer-ear PTA > 25 dB HL for Diotic and Antiphasic Conditions in the Literature

	Poorer-Ear PTA > 25 dB HL					
	Antiphasic			Diotic		
	SRT dB SNR	Sensitivity %	Specificity %	SRT dB SNR	Sensitivity %	Specificity %
De Sousa et al., 2020	-15.9	90	80	-10.3	75	71
Ceccato et al., 2021	-11.7	96	93	N/A	N/A	N/A
Current Study	-15.9	95	90	-10.1	81	90

Table 27: Sensitivity and Specificity Values for Better-ear PTA > 25 dB HL for Diotic and Antiphasic Conditions in the Literature

	Better-Ear PTA > 25 dB HL					
	Antiphasic			Diotic		
	SRT dB SNR	Sensitivity %	Specificity %	SRT dB SNR	Sensitivity %	Specificity %
De Sousa et al., 2020	-14.0	75	67	-9.8	83	72
Ceccato et al., 2021	N/A	N/A	N/A	N/A	N/A	N/A
Current Study	-15.2	93	80	-10.1	100	89

8.3. Asymmetric and Conductive Hearing Impairment

Traditionally DTTs have struggled to identify those with an asymmetric hearing impairment due to the better-ear compensating during diotic testing.

De Sousa, Swanepoel, et al. (2020) found that participants with a moderate asymmetrical SNHL had SRTs similar to those with normal hearing in the diotic condition, but in the antiphasic condition their SRTs were poorer. They also found that in the diotic condition the majority of their participants with a conductive hearing impairment were classified as having normal hearing, but their SRTs also differed considerably to those with normal hearing in the antiphasic condition.

Our study also found that in the diotic condition the mean SRT of participants with an asymmetrical SNHL did not differ significantly from those with normal hearing, but did for the antiphasic. Our study showed that SRTs for those with a conductive hearing impairment did differ in both conditions from those with normal hearing, but this difference was greater in the antiphasic condition with the difference between the mean SRT in the diotic condition for normal hearing and conductive hearing impairment being 1.82 dB SNR compared with 7.39 dB SNR in the antiphasic. This illustrates that the antiphasic version was better able to identify asymmetric and conductive hearing impairment than the diotic New Zealand English DTT.

This increase in the SRT difference between participants with normal hearing and those with an asymmetric or conductive impairment is believed to be caused by the advantage that participants with normal hearing obtain in the antiphasic condition. As has been previously discussed, the 180 degree interaural phase difference has the effect of making speech better separated from background noise due to interaural timing differences that come into effect. This allows those with normal hearing to distinguish the digits even more clearly in background noise, and as a result they obtain much lower SRT scores. Due to hearing impairment impacting the interaural timing differences of the auditory system, this advantage is not present for those with an impairment, resulting in a larger gap between the SRTs of normal hearing participants and those with a impairment in the antiphasic condition.

8.4. Reduction of Trial Numbers

One of the limitations of current DTTs (the New Zealand English DTT included) is the time required for participants to complete the test. The average test time of the diotic New Zealand English DTT is 3 minutes and 30 seconds. While this may not seem like a long time, entering the digit triplet that is spoken twenty-seven times can be tedious for some, risking abandonment of the test.

The number of trials used in the literature is varied, but most tests utilise between 23 – 30 trials, with focus placed on obtaining the best sensitivity and specificity possible (Van den Borre et al., 2021). Our comparison of the AUC, cut-off SRT, sensitivity and specificity for triplet scoring verses poorer-ear 4FA > 25 dB HL and better-ear 4FA > 25 dB HL in the diotic and antiphasic conditions indicated the highest values were obtained at trials 25, 26, and 27 in the antiphasic condition when comparing triplet scoring SRT with poorer-ear 4FA > 25 dB HL. When compromising sensitivity and specificity values to obtain a lower amount of trials, the best value was found at trial 18 for the antiphasic condition when comparing triplet SRT with poorer-ear 4FA > 25 dB HL (cut-off SRT -15.2, AUC 0.95, sensitivity 90%, specificity 93%).

However, while obtaining the maximal value for AUC, sensitivity, and specificity is promoted in the literature – with the strength of a DTT judged by these values – it may be the case that obtaining values close to but below these may be more beneficial. For instance, our results indicate that if one is ready to sacrifice sensitivity and specificity (meaning less people will be correctly identified as having a hearing impairment) for time reduction, then the number of trials in the antiphasic DTT could be decreased from 27 to 18. However, the maximal value of a AUC of 0.92, sensitivity of 95%, and specificity of 90% found at trial 27 was also achieved at trial 25, indicating that a slight reduction in the number of trials may still result in optimal sensitivity and specificity being obtained. We believe a conservative reduction in trial length to 25 should be implemented for the antiphasic New Zealand English DTT. However more research is recommended investigating the test length which will encourage the greatest completion of the test while still maintaining an adequate AUC, sensitivity, and specificity.

8.5. Implications and Applications

Our results have confirmed that the antiphasic New Zealand English DTT has greater diagnostic power than the diotic version. This is because not only is the antiphasic version better able to distinguish hearing impaired listeners from those with normal hearing, it can do so in a reduced time. This is because of the emphasis on poorer-ear PTA. De Sousa, Swanepoel, et al. (2020) found that after controlling for age, the antiphasic condition was more strongly correlated with the poorer-ear PTA. A similar partial correlation on our data also revealed the antiphasic condition across all hearing categories to be more strongly correlated with poorer-ear 4FA (antiphasic [$r = 0.74$], diotic [$r = 0.4$]). While diotic testing can obtain poorer-ear information and is also correlated with poorer-ear PTA (though less so than antiphasic), it requires the testing of each ear individually. With antiphasic stimuli the poorer-ear can be tested simultaneously with the better, obtaining similar results in half the time.

However, as noted by De Sousa, Smits, et al. (2021), a limitation of antiphasic DTTs is that while they are better at determining if a hearing impairment is present, they do not have the capability to accurately determine with confidence the particular classification of that impairment (i.e. symmetric SNHL, asymmetric SNHL, or conductive) (De Sousa, Smits, et al., 2021). Our results also show this limitation, with there being no significant difference in SRT found between the different hearing impairment classifications.

One suggested solution by De Sousa, Smits, et al. (2021) is to identify a potential asymmetric SNHL or conductive hearing impairment by having participants complete both the diotic and antiphasic versions of the DTT (De Sousa, Smits, et al., 2021). Participants would complete the antiphasic version, determining with more accuracy if they have a hearing impairment, then complete the diotic version. If they received a refer on both the antiphasic and diotic versions, this would suggest they have a symmetric SNHL, as diotic versions have high sensitivity and specificity for identifying this impairment (De Sousa, Smits, et al., 2021). On the other hand, if they receive a refer on the antiphasic and a pass on the diotic, then because of the lack of sensitivity and specificity for asymmetric SNHL and conductive hearing impairment with the diotic this could suggest one of these pathologies (De Sousa, Smits, et al., 2021). However, in their attempts to apply this technique, an asymmetric SNHL and conductive hearing impairment were only able to be identified with a fair degree of accuracy, indicating more research is required before this can be practically applied large scale (De Sousa, Smits, et al., 2021).

In a study investigating puretone audiometry and digits-in-noise testing, De Sousa, Smits, et al. (2020) determined that through a combination of the two symmetric SNHLs could be distinguished from asymmetric SNHL and conductive hearing impairments (De Sousa, Smits, et al., 2020). In De Sousa, Moore, et al. (2021), they propose combining DTTs with calibrated self-test kits which would allow puretone AC audiometry to be done at home on a digital device (De Sousa, Moore, et al., 2021). A limitation however of using the antiphasic DTT for such purposes is that it is particularly sensitive at identifying asymmetric SNHL and conductive hearing impairments. It is often the case that a SNHL can be suspected from pure-tone AC results. As diotic tests are more sensitive at picking up a symmetrical

SNHL, a ‘refer’ would help confirm this diagnosis, while an asymmetric SNHL or conductive hearing impairment would produce a ‘pass’. If an antiphasic DTT was used however, all three hearing impairments would produce a ‘refer’.

An easier and more viable option is to update the New Zealand Hearing Screening Test – currently using the diotic New Zealand English DTT – to the antiphasic version. This test is already available online and in a variety of pharmacies throughout New Zealand. By updating this to the antiphasic version, our results suggest a greater number of listeners with an asymmetric or conductive hearing impairment will be given a ‘refer’. The limitation of this however is that if a ‘refer’ is obtained the listener will still need to visit an audiology clinic to determine the specific type and degree of impairment they have.

We suggest further research be conducted to create a test that can better identify specific types of hearing loss – such as the combined diotic and antiphasic recommended by De Sousa, Smits, et al. (2021) – and which can identify the degree of impairment, which will help participants discern whether a visit to the audiology clinic is required sooner rather than later.

9. Conclusions

In this research we investigated whether the use of antiphasic stimuli over diotic, and digit scoring over triplet, could improve the sensitivity and specificity of the New Zealand English DTT. Our study implemented the current New Zealand English DTT – developed by King (2010) and improved upon by Bowden (2013) and Spence (2020) – and a new antiphasic version which included the option to test by digit or triplet. We tested 51 participants with either normal hearing, a symmetrical SNHL, an asymmetrical SNHL, or a conductive hearing impairment. Lists for each participant were allocated according to a latin-square design (totalling 8 lists each with 27 digit triplets) and presented at a calibrated level of 65 dB SPL.

We followed De Sousa, Swanepoel, et al. (2020) for our statistical analysis, with the goal of replicating their results. We successfully increased the sensitivity and specificity of the New Zealand English DTT from 94% and 88% to 95% and 90% respectively using antiphasic stimuli. These results were comparable with, and in fact slightly better than, those found by De Sousa, Swanepoel, et al. (2020), who obtained a sensitivity and specificity of 90% and 80% respectively.

We also found (similar to De Sousa, Swanepoel, et al. (2020)) that in particular there was greater sensitivity for asymmetric SHNL and conductive hearing impairments with the antiphasic version. However the test was still unable to distinguish between the types of hearing impairment more generally.

With regard to digit scoring verses triplet scoring, we did not find that digit scoring produced improved re-test reliability. Digit scoring was also not found to have more diagnostic power with regards to identifying specific classifications of hearing impairment.

We found that when analysing the sensitivity and specificity of each trial in both conditions the antiphasic test still produced a high sensitivity and specificity when the number of trials was reduced to 25 and 18 indicating that useful results could still be produced by shorter tests with less trials.

Our recommendation is that the New Zealand Hearing Screening test be updated to use the antiphasic New Zealand English DTT given its ability to better detect all forms of hearing impairments, and for the number of trials to be reduced to 25 to decrease test time and still achieve a sensitivity of 90% and sensitivity of 95%.

References

Abdurehim, Y., Lehmann, A., & Zeitouni, A. G. (2016). Stapedotomy vs Cochlear Implantation for Advanced Otosclerosis: Systematic Review and Meta-analysis. *Otolaryngology–Head and Neck Surgery, 155*(Generic), 764-770.

Ashmore, J. (2008). Cochlear Outer Hair Cell Motility. *Physiological Reviews, 88*(1), 173-210. https://doi.org/10.1152/physrev.00044.2006

Bowden, A. T. (2013). *Normalisation, Evaluation and Verification of the New Zealand Hearing Screening Test* University of Canterbury]. https://go.exlibris.link/2w3kV35n

Brown, L., Mahomed-Asmail, F., De Sousa, K. C., & Swanepoel, D. W. (2019). Performance and Reliability of a Smartphone Digits-in-Noise Test in the Sound Field. *American journal of audiology, 28*(3S), 736-741.

Brown, M., & Musiek, F. (2013). Pathways: The Fundamentals of Masking Level Differences for Assessing Auditory Function. *The Hearing Journal, 66*(1), 16,17. https://doi.org/10.1097/01.HJ.0000425772.41884.1d

Butcher, E., Dezateux, C., Cortina-Borja, M., & Knowles, R. (2019). OP55 Mental health in young people with hearing loss: longitudinal analysis of the millennium cohort study. *Journal of epidemiology and community health (1979), 73*(Suppl 1), A27.

Ceccato, J.-C., Duran, M.-J., Swanepoel, D. W., Smits, C., de Sousa, K. C., Gledhill, L., Venail, F., & Puel, J.-L. (2021). French Version of the Antiphasic Digits-in-Noise Test for Smartphone Hearing Screening. *Frontiers in public health, 9*, 725080-725080.

Cho, S., Kim, Y. J., Lee, M., Woo, J. H., & Lee, H. J. (2021, 2021/01/25). Cut-off points between pain intensities of the postoperative pain using receiver operating characteristic (ROC) curves. *BMC Anesthesiology, 21*(1), 29.

Clark, J. G. (1981). Uses and abuses of hearing loss classification. *ASHA (Rockville, Md.), 23*(7), 493.

Cohen, B. E., Durstenfeld, A., & Roehm, P. C. (2014). Viral Causes of Hearing Loss: A Review for Hearing Health Professionals. *Trends in Hearing, 18*.

De Sousa, K. C., Moore, D. R., Smits, C., & Swanepoel, D. W. (2021). Digital technology for remote hearing assessment—current status and future directions for consumers. *Sustainability (Basel, Switzerland), 13*(18), 10124. https://doi.org/10.3390/su131810124

De Sousa, K. C., Smits, C., Moore, D. R., Myburgh, H. C., & Swanepoel, D. W. (2020). Pure-tone audiometry without bone-conduction thresholds: using the digits-in-noise test to detect conductive hearing loss. *International journal of audiology, 59*(10), 801-808.

De Sousa, K. C., Smits, C., Moore, D. R., Myburgh, H. C., & Swanepoel, D. W. (2021). Diotic and Antiphasic Digits-in-noise Testing as a Hearing Screening and Triage Tool to Classify Type of Hearing Loss. *Ear and hearing, Publish Ahead of Print*.

De Sousa, K. C., Swanepoel, D. W., Moore, D. R., Myburgh, H. C., & Smits, C. (2020). Improving Sensitivity of the Digits-In-Noise Test Using Antiphasic Stimuli. *Ear and hearing, 41*(2), 442-450.

Denys, S., Hofmann, M., van Wieringen, A., & Wouters, J. (2019). Improving the efficiency of the digit triplet test using digit scoring with variable adaptive step sizes. *International journal of audiology, 58*(10), 670-677.

Elberling, C., Ludvigsen, C., & Lyregaard, P. E. (1989). Dantale: A New Danish Speech Material. *Scandinavian Audiology, 18*(3), 169-175.

Emanuel, D., Maroonroge, S., & Letowski, T. (2009). Auditory function: Physiology and function of the hearing system. In C.E. Rash, M.B. Russo, T.Letowski, & E. Schmeisser (Eds.), *Helmet-Mounted Displays: Sensory, Perceptual, and Cognitive Issues* (pp. 307-334). USAARL.

Exeter, D. J. (2015). The projected burden of hearing loss in New Zealand (2011-2061) and the implications for the hearing health workforce. *New Zealand medical journal, 128*(1419), 12-21.

Gilbert, H. J., Shackleton, T. M., Krumbholz, K., & Palmer, A. R. (2015). The neural substrate for binaural masking level differences in the auditory cortex. *The Journal of neuroscience, 35*(1), 209-220. https://doi.org/10.1523/JNEUROSCI.1131-14.2015

Gong, Q., Wang, Y., & Xian, M. (2014). An objective assessment method for frequency selectivity of the human auditory system. *Biomedical engineering online, 13*(1), 171-171.

Hajian-Tilaki, K. (2013). Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation. *Caspian journal of internal medicine, 4*(2), 627-635.

Hall, J. W. (2004). Hearing Loss and the Masking-Level Difference. In R. D. Kent (Ed.), *The MIT Encyclopedia of Communication Disorders*. MIT Press.

Ho, C.-Y., Li, P.-C., Young, S.-T., Yeh, W.-Y., Wang, C.-H., & Chu, W.-C. (2016). Preliminary Measurements of Binaural Masking-Level Difference When Using Hearing Aids for Sensorineural Hearing Loss. *Journal of medical and biological engineering, 36*(5), 668-674. https://doi.org/10.1007/s40846-016-0174-4

Hodgson, S. L. (2016). *Release from masking: Behavioral and physiological masking level differences* James Madison University].

Hsieh, D.-L., Lin, K.-N., Ho, J.-H., & Liu, T.-C. (2009). Hearing in Noise Test in Subjects With Conductive Hearing Loss. *Journal of the Formosan Medical Association, 108*(12), 937-942.

Jansen, S., Luts, H., Dejonckere, P., van Wieringen, A., & Wouters, J. (2014). Exploring the sensitivity of speech-in-noise tests for noise-induced hearing loss. *International journal of audiology, 53*(3), 199-205.

Jansen, S., Luts, H., Wagener, K. C., Frachet, B., & Wouters, J. (2010). The French digit triplet test: A hearing screening tool for speech intelligibility in noise. *International journal of audiology, 49*(5), 378-387.

Katz, J., Chasin, M., English, K. M., Hood, L. J., & Tillery, K. L. (2015). *Handbook of clinical audiology* (Seventh ed.). Wolters Kluwer Health.

King, S. M. (2011). *Development and evaluation of a New Zealand Digit Triplet Test for auditory screening*. University of Canterbury.

Kramer, S., & Brown, D. K. (2018). *Audiology*. Plural Publishing, Incorporated.

Kurabi, A., Keithley, E. M., Housley, G. D., Ryan, A. F., & Wong, A. C. Y. (2017). Cellular mechanisms of noise-induced hearing loss. *Hearing research, 349*, 129-137. https://doi.org/10.1016/j.heares.2016.11.013

Leek, M. R. (2001). Adaptive procedures in psychophysical research. *Perception & psychophysics, 63*(8), 1279-1292.

Leensen, M. C. J., de Laat, J. A. P. M., & Dreschler, W. A. (2011). Speech-in-noise screening tests by internet, Part 1: Test evaluation for noise-induced hearing loss identification. *International journal of audiology, 50*(11), 823-834.

MacPherson, A., & Akeroyd, M. A. (2014). Variations in the slope of the psychometric functions for speech intelligibility: a systematic survey. *Trends in Hearing, 18*.

Møller, A. R. (2013). *Hearing: anatomy, physiology, and disorders of the auditory system* (3rd;3; ed.). Plural Pub.

Moore, B. C. J. (2012). *An introduction to the psychology of hearing* (6th ed.). Emerald.

Murakoshi, M., Suzuki, S., & Wada, H. (2015). All Three Rows of Outer Hair Cells Are Required for Cochlear Amplification. *BioMed research international, 2015*, 727434-727412. https://doi.org/10.1155/2015/727434

Musiek, F. E., & Baran, J. A. (2020). *The auditory system: anatomy, physiology, and clinical correlates* (Second ed.). Plural Publishing.

National Aeronautics and Space Administration (NASA). (2004). Human Vestibular System in Space. https://www.nasa.gov/audience/forstudents/9-12/features/F_Human_Vestibular_System_in_Space.html

National Foundation for the Deaf, & Deloitte Access Economics. (2017). *Listen hear! New Zealand: social and economic costs of hearing loss in New Zealand : a report*.

Olsen, W. O., Noffsinger, D., & Carhart, R. (1976). Masking Level Differences Encountered in Clinical Populations. *International journal of audiology, 15*(4), 287-301. https://doi.org/10.3109/00206097609071789

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

Parthasarathy, A., & Kujawa, S. G. (2018). Synaptopathy in the Aging Cochlea: Characterizing Early-Neural Deficits in Auditory Temporal Envelope Processing. *The Journal of neuroscience, 38*(32), 7108-7119. https://doi.org/10.1523/JNEUROSCI.3240-17.2018

Plack, C. J. (2013). *The sense of hearing* (2nd ed.). Psychology Press.

Potgieter, J.-M., Swanepoel, D. W., Myburgh, H. C., Hopper, T. C., & Smits, C. (2016). Development and validation of a smartphone-based digits-in-noise hearing test in South African English. *International journal of audiology, 55*(7), 405-411.

Potgieter, J.-M., Swanepoel, D. W., Myburgh, H. C., & Smits, C. (2018). The South African English Smartphone Digits-in-Noise Hearing Test: Effect of Age, Hearing Loss, and Speaking Competence. *Ear and hearing, 39*(4), 656-663.

Qi, R., Su, L., Zou, L., Yang, J., & Zheng, S. (2019). Altered Gray Matter Volume and White Matter Integrity in Sensorineural Hearing Loss Patients: A VBM and TBSS Study. *Otology & neurotology, 40*(6), e569-e574.

Quaranta, A., & Cervellera, G. (1974). Masking Level Difference in Normal and Pathological Ears. *Audiology, 13*(5), 428-431. https://doi.org/10.3109/00206097409071702

Raphael, Y., & Altschuler, R. A. (2003). Structure and innervation of the cochlea. *60*(Generic), 397-422.

Robles, L., & Ruggero, M. A. (2001). Mechanics of the Mammalian Cochlea. *Physiological Reviews, 81*(3), 1305-1352.

Shipley, K. G., & McAfee, J. G. (2021). *Assessment in speech-language pathology: a resource manual* (Sixth ed.). Plural Publishing, Inc.

Smits, C. (2017). Improving the Efficiency of Speech-In-Noise Hearing Screening Tests. *Ear and hearing, 38*(6), 385-388.

Smits, C., Kapteyn, T. S., & Houtgast, T. (2004). Development and validation of an automatic speech-in-noise screening test by telephone. *International journal of audiology, 43*(1), 15-28.

Smits, C., Kapteyn, T. S., & Houtgast, T. (2009). Development and validation of an automatic speech-in-noise screening test by telephone. *International journal of audiology, 43*(1), 15-28.

Smits, C. H. M., & Houtgast, T. (2005). Results from the Dutch speech-in-noise screening test by telephone. *Ear and hearing, 26*(1), 89-95.

Spence, K. J. (2020). *Improving the performance of the New Zealand Hearing Screening Test* [Thesis, University of Canterbury]. Christchurch, New Zealand. https://go.exlibris.link/njJ8zBpV

Strawbridge, W. J., Wallhagen, M. I., Shema, S. J., & Kaplan, G. A. (2000). Negative Consequences of Hearing Impairment in Old Age. *The Gerontologist, 40*(3), 320-326.

Van den Borre, E., Denys, S., van Wieringen, A., & Wouters, J. (2021). The digit triplet test: a scoping review. *International journal of audiology*, 1-18.

Veiga, S. M., Alexandre, J. D., & Esteves, F. (2015). Living with Hearing Loss: Psychosocial Impact and the Role of Social Support. *Journal of Otolaryngology-ENT Research, 2*(5).

Vlaming, M. S. M. G., MacKinnon, R. C., Jansen, M., & Moore, D. R. (2014). Automated Screening for High-Frequency Hearing Loss. *Ear and hearing, 35*(6), 667-679.

White, H., Helwany, M., & Peterson, D. (2021). Anatomy, Head and Neck, Ear Organ of Corti. Retrieved 22/11/2021, from https://www.statpearls.com/articlelibrary/viewarticle/26313/

Wilson, R. H., Burks, C. A., & Weakley, D. G. (2006). Word Recognition of Digit Triplets and Monosyllabic Words in Multitalker Babble by Listeners with Sensorineural Hearing Loss. *Journal of the American Academy of Audiology, 17*(6), 385-397. https://doi.org/10.3766/jaaa.17.6.2

Wilson, R. H., Moncrieff, D. W., Townsend, E. A., & Pillion, A. L. (2003). Development of a 500-Hz Masking-Level Difference Protocol for Clinic Use. *Journal of the American Academy of Audiology, 14*(1), 1-8. https://doi.org/10.3766/jaaa.14.1.2

Wolmarans, J., De Sousa, K. C., Frisby, C., Mahomed-Asmail, F., Smits, C., Moore, D. R., & Swanepoel, D. W. (2021). Speech Recognition in Noise Using Binaural Diotic and Antiphasic Digits-in-Noise in Children: Maturation and Self-Test Validity. *Journal of the American Academy of Audiology, 32*(5), 315-323. https://doi.org/10.1055/s-0041-1727274

World Health Organization. (2021). World Report on Hearing. https://www.who.int/publications/i/item/world-report-on-hearing

Appendices

Appendix 1: Approval letter from University of Canterbury Human Ethics Committee

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

HUMAN ETHICS COMMITTEE

Secretary, Rebecca Robinson
Telephone: +64 03 369 4588, Extn 94588
Email: human-ethics@canterbury.ac.nz

Ref: HEC 2021/26/LR

1 June 2021

Rosalie Joan Hosking
Psychology, Speech and Hearing
UNIVERSITY OF CANTERBURY

Dear Rosalie

Thank you for submitting your low risk application to the Human Ethics Committee for the research proposal titled "Improving the New Zealand Digit Triplet Test Sensitivity and Specificity using Antiphasic Stimuli".

I am pleased to advise that this application has been reviewed and approved, **subject to the following**:

- *In the Information Sheet, please mention all potential publication possibilities, e.g. academic journals, conference or professional presentations, etc., so that participants can provide prior informed consent for these.*
- *Also in the Information Sheet, in the section regarding the right to withdraw, please add that any decision to withdraw will not affect the participant's relationship with the University, or the provision of any services that they are receiving from the UC clinic, etc.*

With best wishes for your project.

Yours sincerely

Dr Dean Sutherland
Chair, Human Ethics Committee

University of Canterbury Private Bag 4800, Christchurch 8140, New Zealand. www.canterbury.ac.nz

F E S

Appendix 2: Research Participant Consent Form

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

School of Psychology, Speech and Hearing
Email:
rosalie.hosking@pg.canterbury.ac.nz

Improving the New Zealand Digit Triplet Test Sensitivity and Specificity using Antiphasic Stimuli

Consent Form for Participants

- ☐ I have been given a full explanation of this project and have had the opportunity to ask questions.
- ☐ I understand what is required of me if I agree to take part in the research.
- ☐ I understand that participation is voluntary, and I may withdraw at any time without penalty. Withdrawal of participation will also include the withdrawal of any information I have provided should this remain practically achievable.
- ☐ I understand that any information or opinions I provide will be kept confidential to the researcher and research team and that any published or reported results will not identify the participants in any form. I understand that a thesis is a public document and will be available through the UC Library.
- ☐ I understand that all data collected for the study will be kept in locked and secure facilities and/or in password protected electronic form and will be destroyed after five years.
- ☐ I understand the risks associated with taking part and how they will be managed.
- ☐ I understand that I can contact the researcher Rosalie Hosking at rosalie.hosking@pg.canterbury.ac.nz or supervisor Professor Greg O'Beirne at +64 3 369 43139 or gregory.obeirne@canterbury.ac.nz for further information. If I have any complaints, I can contact the Chair of the University of Canterbury Human Ethics Committee, Private Bag 4800, Christchurch (human-ethics@canterbury.ac.nz)
- ☐ I would like a summary of the results of the project.
- ☐ By signing below, I agree to participate in this research project.

Name: ________________________ Signed: ______________________ Date: _________________

Email address *(for report of findings, if applicable:* ____

Appendix 3: Participant Information Sheet

Information Sheet

Rosalie Hosking
Master of Audiology student
Te Kura Mahi ā-Hirikapo | School of Psychology, Speech and Hearing
rosalie.hosking@pg.canterbury.ac.nz

- *The purpose of this information sheet is to inform you about this project so you can decide whether you want to take part.*
- *It is important that you understand this information. Take as much time as you need to decide.*
- *Feel free to talk about your participation in this project with your family or health care providers.*

Improving the New Zealand Digit Triplet Test Sensitivity and Specificity using Antiphasic Stimuli

Our purpose

- The aim of our project is to improve the New Zealand Digit Triplet Test – also known as the New Zealand Hearing Screening Test – to make it function better for New Zealanders.

What is involved?

- If you choose to take part, we will ask you to take a traditional hearing test (approx. 30 mins) and then the New Zealand Digit Triplet Test (approx. 30 mins). Total time = approximately 1 hour.
- This will be done at the University of Canterbury Speech and Hearing Clinic.
- The hearing tests and hearing screenings are free, and you will be put into the draw to win a $150 Westfield Riccarton voucher.

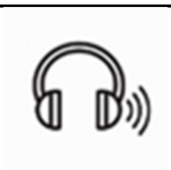
Who is eligible for the project?

- You are eligible to take part if you are aged 18 and above and can understand and repeat the numbers zero to nine

What will I do if I take part?

As mentioned above, there are two main parts to the testing session, and each part takes about 30 minutes.

1. **The hearing test – (approx. 30 minutes).**

	• We will ask you some questions about your general health, your hearing and balance. • The answers to these questions help us to understand what difficulties you may be having. We want to know: o if you have any problems with your hearing, balance, or communication, o when these problems started.
	• We check the outside of your ears. Then we will use a bright light to look at the inside of your ears. This helps us to make sure your ears are healthy inside and out and clear of wax.
	• We will then measure how your ears react to sound and pressure changes. • This test tells us if your eardrum, ear bones, and muscles are moving as they should.
	• We will fit earphones on or in your ears. Then we will play some sounds at different listening levels. Some of the sounds will be quieter or louder than others. We will ask you to tell us when you can hear them by pressing a button. • Sometimes we play background noise at the same time and you may find it hard to tell whether there is a sound or not. This shows us how well you can hear in everyday situations.
	• We only carry out the tests that are needed to detect a hearing loss. We will help you to understand your results and how to find out more information. Please feel free to ask us lots of questions and to bring someone with you for support. • If your results show that you need to see a doctor, we can refer you.

2. **The Digit Triplet Test (approx. 30 minutes).**

	• We will fit headphones on your ears, and we will give you a computer tablet to use. You will hear a woman's voice saying sets of three numbers at different listening levels, and with different levels of background noise. • The numbers are presented in a randomised order, either to both ears at the same time, or to the right and left ears separately. • We will ask you to enter the numbers you hear on the keypad. • As the test progresses you will find it much harder to hear the numbers. But that's OK! Please continue to try - even if you think it is wrong, just have a guess. This helps us to check the accuracy of the screening test.

Data, Confidentiality, and Privacy

- The results of the project will be published but your data will be strictly confidential. Your identity will not be made public. The results may be published in academic journals, and conference or professional presentations.
- To ensure anonymity and confidentiality, your data will be stored in a secure room inside a locked cabinet. In the case of digital data, this will be password protected and stored on password-protected devices. Any backups made will be stored on a secure UC server. Your data will be assigned an identification code so that the data can be de-identified. Participant data will only be accessible to members of the research team.
- You may ask for your raw data to be returned to you or destroyed at any point. If you withdraw, I will remove information relating to you. However, once the analysis of raw data starts in September 2021, it will become increasingly difficult to remove the influence of your data on the results.

Results and dissemination

- The final written submission will be made to the University in February 2022.
- A thesis is a public document and will be available through the UC Library after this date.
- Please indicate to the researcher on the consent form if you would like to receive a copy of the summary of the results of the project. You will have the opportunity to receive and provide verbal feedback on the findings of the study.

Ethics

- This project has been reviewed and approved by the University of Canterbury Human Ethics Committee, and participants should address any complaints to The Chair, Human Ethics Committee, University of Canterbury, Private Bag 4800, Christchurch (human-ethics@canterbury.ac.nz).

Consent

- If you agree to participate in the study, you are asked to complete the attached consent form and return it to Rosalie Hosking (Master of Audiology Student) in person, by post, or by email (see below for details).

Right to withdraw

Taking part is your choice, and you can withdraw at any stage without giving a reason. Should you decide to withdraw, this will not affect your relationship with the University or any services received from the UC Speech and Hearing Clinic.