The role of cognitive functions in the hearing of speech-in-noise and the role of auditory and cognitive training in individuals’ speech-in-noise performance.

SCOTT HAMILTON RICHARDS

Doctor of Philosophy

ASTON UNIVERSITY

June 2018

SUMMARY

Aston University

The most consistent determinant for any variation in speech-in-noise performance between individuals is their own audibility. There is, however, the potential that cognitive functions support listening when situations are challenging, such as when speech is being listened to in background noise. This research explores which cognitive functions support listening when speech is degraded in different manners, with working memory, speed of information processing, response latency and control of inhibition all indicating significant relationships. Based on these findings, both cognitive and auditory training was employed with participants to establish if this could demonstrate any improvement in both on-task performance, such that the task itself improved, and additionally if training could lead to changes in performance of other closely related (near transfer) or more loosely related (far transfer) tasks.

There was no convincing evidence that cognitive training involving inhibition control and processing speed could lead to improved listening performance. Potentially, as cognitive functions play a more minor role in speech-in-noise performance when compared to audibility, training of these cognitive functions may need a more significant training input to demonstrate any significant change in speech-in-noise performance. There was, however, evidence that targeted auditory training may offer some advantages and that far transfer gains appear to be achievable and are most evident when the training material is highly challenging. When the role of a commercial communication training package involving both auditory and cognitive exercises was explored it gave very limited evidence of any significant advantage for first time hearing aid users over normal acclimitisation.

Areas for future research are discussed, particularly with a view to training package development. Further, there is a discussion on how effective rehabilitation may lead to improved communication and the potential social and mental well-being that improved communication may bring.

Key words: Degraded speech, Cognition, Training, Communication, and Transfer.

DEDICATION

I dedicate this thesis to my wife Sandy and my two children, Jordan and Bethany.

ACKNOWLEDGEMENTS

Firstly, I would like to express my sincerest gratitude to my supervisor Professor Carol Holland, for her seemingly endless enthusiasm and motivation throughout this research. Her excellent guidance and expert advice has been invaluable. Her insight and feedback on various written iterations have allowed for the development of this final thesis. My thanks also go to Professor Paul Furlong for his helpful insights and feedback.

I owe an impossibly large debt to my wife Sandy who has played an incredible role in supporting me. Without her, this work would never have been possible and I cannot not thank her enough. Her encouragement when there were set-backs assisted tremendously in overcoming any problems. I would also like to acknowledge the role my two children, Jordan and Bethany played. They were both studying at the same time as this research was taking part. Their support for my own work and general everyday demeanour helped to keep me focussed and motivated.

Many thanks must also go to all the participants that took part in the various pieces of research for this thesis. I greatly appreciate the time they gave and the good will they displayed towards me and my research. Their adherence to the training regimes was of the highest order.

I would like to thank both the Aston Research Centre for Healthy Ageing and the British Society of Audiology for awarding grants to allow for sufficient funds to perform the research in Chapter 6.

Lastly, I would like to thank my colleagues and friends who have been willing volunteers where needed and provided insightful thoughts and guidance. In particular, I would like to thank Amanda Casey as the head of Audiology at Aston for allowing me the opportunity to undertake this program of study. Her support and interest in my progress has been greatly appreciated.

TABLE OF CONTENTS

LIST OF TABLES

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Appendix 1

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Appendix 1

CHAPTER ONE

1. Introduction

This introduction consists of four sections. The first section describes the rationale for the research presented in this thesis, outlining the important mental and social health benefits that effective communication brings. It also summarises the background that provides the context for this work, including a review of the complex process of how speech is perceived. Next it considers cognition, the cognitive functions that are the focus of this thesis, and cognitive changes with age, followed by a brief discussion of the potential relationship between hearing loss and cognitive decline.

Having introduced the cognitive functions that are the focus of this thesis, the second section of the introduction explores the theoretical background through a literature review and gives evidence of the research supporting the link between the ability to hear speech-in-noise and cognitive support mechanisms. Two different types of training, cognitive and auditory, are then reported, with an exploration of the manner in which they may offer improvements for individuals in listening performance in challenging conditions.

After the literature review, the next section of this introduction characterises the magnitude and nature of the problem of hearing loss in the United Kingdom. There is a trend towards an increasing population of older adults, amongst whom there will be a greater number of individuals with raised hearing thresholds. This same section then discusses the effectiveness of hearing aids and the potential that they offer, not just for improved audition but also improved cognitive functioning. The introduction ends with a summary of the aims and hypotheses explored in the different chapters of this thesis.

1.1. Rationale for the research and an introduction to the perception of speech and cognition

People with and without hearing loss often struggle to hear and understand speech in situations where speech quality is degraded or when the speech is competing with background noise (Mattys, Davis, Bradlow, & Scott, 2012). For those with peripheral hearing loss (as distinct from a central processing problem that brings additional difficulties), there are increased challenges in understanding speech (Dubno, Dirks, & Morgan, 1984). Hearing aids may be of benefit for people with hearing loss, but access to and use of these remains problematic (Gianopoulos, Stephens, & Davis, 2002; McCormack & Fortnum, 2013; Mizutari et al., 2013; Smeeth et al., 2002). There is evidence, however, that compared to traditional behind-the-ear hearing aids, the newer, slim-fit, mini behind-the-ear hearing aids offer greater satisfaction and a greater reduction in perceived handicap, as self-reported by participants in the Marke Track VIII hearing aid market survey (Kochkin, 2011).

For others, even if hearing loss itself is not the main presenting problem, hearing in adverse listening conditions can still be a significant challenge (Veneman, Gordon-Salant, Matthews, & Dubno, 2013). Indeed, it has been reported that for older individuals whose hearing levels have not shown a decline, understanding speech in noisy situations is still more challenging than it is for younger listeners with similar hearing ability (Dubno et al., 1984; Pichora-Fuller, Schneider, & Daneman, 1995).

Any decline in an individual's ability to understand speech may have an adverse effect on their communication skill. Communication can be defined as the giving and receiving of information, and for this exchange to take place, the following steps must occur. If the communication is verbal, the words must be heard. It must be possible to discriminate between them, and the topic of the conversation must be preserved and processed in working memory. This enables an appropriate reply to be formulated and given by the receiver of the information, and for the information to be processed and stored for future use. The ability to communicate well is not, therefore, simply a question of audition, but involves other complex processes that need to be understood if a high-standard audiology service is to be delivered to individuals (Kiessling et al., 2003a).

In addition to allowing for the exchange of information, communication also plays a key role in emotional and physical well-being, no less for people with an acquired hearing loss (Gopinath et al., 2012; Hallberg, Hallberg, & Kramer, 2008). For example, in a community environment for older people, Carabellese et al. (1993) found that communication problems associated with hearing loss had a negative effect on emotional well-being. In a study from Japan, Saito et al. (2010) found that depression levels could be predicted by the level of hearing loss in a community-dwelling older population.

Furthermore, in a longitudinal ageing research study in Amsterdam, Pronk, Deeg, and Kramer (2013) report on the isolating effects of hearing loss. A self-report of hearing status and a simple speech-in-noise task were used as measures of hearing loss. Both outcome measures were found to be significantly associated with loneliness, i.e., poorer performance on the speech-in-noise task or a lower self-rated listening performance was associated with greater reported loneliness. This was only, however, for non-hearing-aid users and men. There was no significant association of either of the measures with depression.

It appears, therefore, that if individuals experience hearing loss but do not seek intervention to support their communication ability (such as a hearing aid), they are at increased risk of becoming socially isolated. This emphasises the importance of offering effective support for individuals with hearing loss, including access to hearing aids and high-quality support in their use. However, problems remain with hearing aid uptake, in particular where reports of non-use are associated with the users not perceiving sufficient benefit from them (McCormack & Fortnum, 2013). Therefore, rehabilitation strategies should be explored, including communication training, which may allow for supportive strategies to deliver greater benefits. Similarly, support and understanding of the problems arising from communication difficulties for individuals with ‘normal’ hearing should also be available; for example, working with those who struggle to hear in adverse listening conditions.

The next section deals with the perception of speech and the processes involved in processing and interpreting spoken information. It is introduced here because before focusing on the roles that cognitive functions may play in understanding speech-in-noise, the complicated process of being able to understand speech needs to be understood.

1.1.1. Perception of speech

The perception of speech is not straightforward, and there are many different stages involved. One model to describe these stages (Cutler & Clifton, 1999) focuses on how an auditory input must initially be decoded by the listener, then segmented (both segmentally and suprasegmentally), recognized, and integrated. It must also be borne in mind that one of the first roles of the listener is to separate the signal of interest (speech) from any extra unwanted auditory information, which is achieved by sorting the inputs into different ‘streams’. Speech itself is periodic and constant noise is aperiodic, meaning that speech may be easier to separate from noise. The same is not true when the ‘noise’ also has periodic features, such as background speech babble. For this reason, when researching speech testing in noisy environments, one must consider which test is more appropriate, or indeed whether testing in both types of background noise may be useful for highlighting any differences.

The segments of speech can be broken down into phonemes, the smallest units of speech, which, when put together sequentially, form words. Differentiating between phonemes is one of the keys to understanding words. Having identified the presence of a phoneme, the listener is able to discriminate between words containing this phoneme and reject those in which it is not present. This can lead to activation of the listener’s lexical candidates for words containing this phoneme. Once the word has been recognised, interpretation using syntactic analysis can be performed and the utterance can be integrated by the individual into the communication that is taking place.

The ability to distinguish between two phonemes is a type of ‘bottom-up processing’ in which the auditory system can remove misinterpretations of words by distinguishing the correct phoneme, such that only the appropriate lexical candidates are considered. To distinguish between phonemes, the listener’s auditory system analyses the frequencies present, which differ depending on the sounds. The large-amplitude frequencies in utterances can be seen on speech spectrographs and are termed ‘formants’ (Carlson, Fant, & Granstroem, 1974). Formants are created by resonances in the vocal tract, and their frequency construct allows them to be distinguished from other speech sounds. Although some speech sounds have many formants, it is often the first two or three that give meaning to the sound (Liberman, 1995). The large amplitude activity at certain frequencies that is generated by these speech sounds can also be recorded in the auditory brainstem (Skoe & Kraus, 2010), allowing analysis of how faithfully an individual listener's auditory neural system represents incoming speech sounds.[1]

The basic process of phoneme identification outlined above is complicated by several factors. The first is the speed at which speech is delivered, which is normally on the order of 12 phonemes per second—although individuals can process up to 60 phonemes per second and still understand the meaning (Werker & Tees, 1992). In addition, individual phonemes are affected by the phonemes immediately before and after them. This co-articulation means that the speech sound for a phoneme is variable. This variability may provide some clues as to what has gone before or which speech sound is to come, with the acoustic difference of the phoneme acting as a clue for the listener. One example is the way in which the speech sound 'k' changes in its articulation depending on where it is used in a word; for instance, 'keep' compared to 'peak'. Phoneme identification difficulties also arise from differences between speakers, as different accents and rates of speech have an effect on phonemic understanding. Segmenting of words may also cause problems, as words themselves may not be discrete and tend to run into one another.

A further component of speech perception is that of top-down processing (Davis, Johnsrude, Hervais-Adelman, Taylor, & McGettigan, 2005; Pichora-Fuller, 2008; Sohoglu, Peelle, Carlyon, & Davis, 2012). Top-down processing utilises an individual's prior knowledge and experience of words and semantics to inform their speech perception. In research using degraded speech, Sheldon, Pichora-Fuller, and Schneider (2008) recruited younger and older adults to listen to recordings of sentences in which the signal quality was degraded. Participants were given priming to allow them to become accustomed to the envelope of the degraded signal, as well

---

[1] The potential that these neural representations are linked to speech-in-noise performance was explored in this thesis when recordings were made on a sub-section of participants recruited for the research in Chapter 4. The report is in Appendix 1.

as contextual information about the test sentences. The final target words in the speech task were of either low or high semantic predictability. In general, highly contextualised words were easier for the listeners to understand than unpredictable words. Both the younger and older group benefitted from the support of the priming and context. Older adults generally performed more poorly when the speech was more degraded and benefitted more from the sentence priming and context. The authors suggest that this gives evidence that older adults utilise top-down processing more than younger adults.

In similar fashion, other researchers have demonstrated that older adults use support mechanisms for general auditory processes (Craik, 1986) and more specifically when listening to speech in degraded listening situations (Davis et al., 2005; Pichora-Fuller, 2003b; Wingfield & Tun, 2007). When considering the potential of training to improve communication ability, a key consideration of this thesis, the fact that older adults utilise this top-down support effectively needs to be considered.

Overall, it appears that there is a balance to be achieved between bottom-up (discriminating between phonemes/words) and top-down processing (making use of semantic knowledge) in the effective perception of speech, particularly when the speech is challenging, for example when it is presented against background noise (Davis et al., 2005; Mattys et al., 2012; Pichora-Fuller, 2008; Sheldon et al., 2008).

Having considered the complexities of speech perception above, the next sections introduce cognition in general and the cognitive functions that are explored in this thesis.

1.1.2. Cognition, cognitive functions and executive control

Broadly, cognition refers to the ways in which an individual uses their mental processes and acquires knowledge. There are many different cognitive functions that contribute to this; for this thesis specifically, the speed of information processing, the control of inhibition, and working memory (both complex and simple) are explored. These functions are introduced below, along with their potential roles in linking cognition and communication. It should also be borne in mind that the ‘central executive’ component of working memory has a controlling influence over these different functions (Miyake et al., 2000). The role of the central executive is introduced below.

1.1.2.1. Executive function roles

The central executive is responsible for the allocation of an individual’s cognitive functions. This includes planning, organising, and performing intentional behaviour; updating and checking working memory; selective attention; the inhibition of unwanted stimuli; sustained attention; the planning and monitoring of performance; task maintenance and initiation; and cognitive flexibility (Smith & Jonides, 1999; Stuss & Knight, 2002). These higher-order executive functions appear to offer complex cognitive processing that is composed of a range of different functions. Many of these functions would be expected to have a role in the discrimination of speech-in-noise and are thus worthy of investigation. Additionally, control of these executive mechanisms and their changes over time may also explain variations in the perception of speech in noise.

Miyake et al. (2000) describe an empirical basis for the theory of how executive functions are organised and the roles that they play in complex cognition. The most commonly postulated executive functions are task shifting, updating of working memory, and general inhibition (involving selective attention, ignoring distractors and the inhibition of prepotent responses). Using factor analysis, (Miyake et al., 2000) demonstrate that these three functions are moderately correlated, though separate.

Task-switching, one of the functions, is needed for a variety of cognitively demanding tasks, hearing in adverse conditions amongst them. For example, listening in challenging situations often requires the ability to switch between different target speakers in group environments. This may be further complicated if other group members are also carrying on conversations that the listener needs to inhibit (another executive function). Therefore, if an ageing individual has declining hearing and control of inhibition, this may adversely affect their ability to communicate in difficult environments. Furthermore, this may lead to social withdrawal and the potential for related mental health issues (Livingston et al., 2017). As well as any decline in the cognitive functions themselves, there is the potential that the executive control of such resources also declines with age (Troyer, Graves, & Cullum, 1994).

An understanding of executive control may assist in the development of a model of the interaction between these different factors and potential rehabilitation activities. Planned training programmes focusing on assisting audition, improving cognitive function, or both, when listening in noise, are explored in Chapters 3, 4, and 5. The cognitive functions under executive control that are the focus of this thesis are considered below. As speed of information processing and control of inhibition are the main focuses of this thesis, they are considered first, followed by a description of working memory.

1.1.2.2. Speed of information processing

Information processing speed can be simply defined as the time taken to perform a given task. The task may be straightforward but could involve the processing of some information before responding in the required fashion. In brief, it is the time taken between stimulus reception and response. A considerable amount of research has explored the link between intelligence and the ability to process information quickly (Sheppard & Vernon, 2008). However, there has been limited research on the possible relationship between information processing speed and the discrimination of speech-in-noise. There are very few explorations of this relationship in the comprehensive review of cognitive functions and speech-in-noise performance by Akeroyd (2008). The research reported in the Akeroyd review, as well as further studies in this area, are reported in greater detail in Section 1.2.2.

Speed of information processing can be measured by different methods that explore different aspects of an individual’s performance. These include tasks that involve simple reaction times, general speed of processing, short- and long-term memory processing, and inspection time. Examples of the different tasks employed in this thesis are given below.

Simple reaction time tasks (also known as response latency tasks) involve presentation of a single stimulus, with instructions to the participant to carry out some action as soon as they observe a stimulus (which could be auditory or visual). The reaction will often involve some form of motor activity, such as pressing a button or giving spoken responses. These tasks may also involve choosing which of different stimuli conditions needs to be responded to (a choice reaction time, CRT). Tasks such as the ‘go’ condition of the ‘go/no-go’ task described in Chapter 2 and the congruent condition of the ‘Stroop’ task used in Chapters 2, 3, and 5 are such CRTs. A computer push-button simple reaction task (SRT) is employed in Chapter 5. The methods for each of these tests are explained in greater detail in the relevant chapters.

Inspection time information-processing speed tasks do not require a quick response from an individual. Rather, information is displayed for short periods of time, and the individual must indicate what they have seen. By shortening the display time, the minimum time required to process the information can be calculated for each participant (the information displayed may be simple or complex). This type of assessment is described in Chapter 3 (within the Useful Field of Vision test), alongside an auditorily presented inspection time task, where sentences are time compressed by 50%. The individual must then attempt to indicate what they saw, or repeat what they heard. Again, with this task, there is no speed element to the response; rather, it is a measure of how quickly and/or how accurately the presented information can be processed.

There are also short- and long-term memory speed tasks that test how quickly a participant recognises previously presented information. These speed-of-information-processing tasks, however, are not discussed in this thesis.

As noted above, previous research has demonstrated a significant correlation between intelligence test performance and information processing speed (Deary, Der, & Ford, 2001; Stough, Nettelbeck, & Cooper, 1993). The relationship being that the quicker an individual can process information, the faster they can act on it, allowing for better fluid intelligence. The potential for a link between information processing speed and speech discrimination is that the more quickly an individual is able to process spoken information, the more quickly they will find phonological matches with spoken utterances. If the speech is presented in degraded condition, such as with background noise, this becomes more important, as the information may be incomplete. If the information is incomplete, semantic knowledge from long-term memory will assist an individual in completion of partially heard words, however, this extra processing takes time, and so an individual with faster processing speed may be better able to discriminate speech-in-noise.

For the different tasks outlined above, faster responses or the ability to process information more quickly are indicative of superior performance (Salthouse, 1994, 1996).

1.1.2.3. Control of inhibition

Hasher and Zacks (1979) cite the following three areas as relevant to the control of inhibition:

- Controlling access to attention’s focus
- Deleting irrelevant information from working memory and attention
- The suppressing or restraining of strong but inappropriate responses

The above points are potentially key to the detection of a relevant speech signal in noise, as only the target speech should be attended to once the incoming information has been streamed (meaning that target speech is processed and the unwanted background noise stream is inhibited). Therefore, inhibition tasks are designed to measure how well individuals perform these functions.

The two tests used in this thesis to measure prepotent inhibition control are the Stroop test (Stroop, 1935) and a random number generation task (RNG) (Ginsburg & Karpiuk, 1994). Prepotent inhibition is where an individual must ‘overcome’ or inhibit a previously learned response. In the case of the classic Stroop test, participants are presented with words in either a congruent condition (the meaning of the word and the colour of the word match, such as ‘red’ printed in red ink) or an incongruent condition (the meaning and the colour do not match, such as ‘red’ printed in green ink). For the incongruent condition, an individual’s prepotent (or automatic) response to read the word must be overcome (inhibited) in order to state the colour of the ink. In the most widely used method for calculating inhibition cost, the speed of the responses to the congruent and incongruent conditions are taken as measures, and the average congruent response time is subtracted from the average incongruent response time (Macleod, 1991). The smaller the difference in time between the two measurements, the better the inhibition control. RNG tasks also measure prepotent inhibition, where number patterns and number repetitions must be inhibited. For this task, an individual is instructed to choose numbers between one and nine at random and to say them aloud, avoiding patterns. The RNG gives two measure of prepotent inhibition control and one of control of updating. Both the above tasks are employed in Chapters 2, 3, and 5 of this thesis.

Secondly, inhibition can be exhibited as an individual's ability to attend to target information and disregard irrelevant information in real-time, ensuring that the irrelevant information does not interfere with the processing of the target information. In the case of conversations in adverse conditions, a lack of real-time inhibitory control may result in the inability to attend to the target voice, causing unwanted information to be processed and using up valuable cognitive resources. Selective auditory attention tasks have been employed in Chapters 2 and 3 of this thesis, with Chapter 3 also using an analogous visual selective attention task as part of the UfoV test.

1.1.2.4. Working memory

Working memory essentially concerns how much information an individual can keep 'live' and recall after a short period of time. This could be a simple list of objects or words (heard or seen). The task can be made more complex by asking the participant to undertake a secondary task at the same time. Baddeley and Hitch (1974) proposed a model in which the central executive has a controlling influence over working memory. Baddeley (1986) expands on the role of the central executive, introducing the 'episodic buffer' as a further component (Baddeley, 2000). The episodic buffer offers back-up storage and communicates between the long-term and working memories. Rudner and Rönnberg (2008) hypothesise that the episodic buffer may work alongside the central executive when language processing becomes particularly effortful. This is potentially achieved by the engagement of posterior neural networks alongside the frontal networks activated by the central executive. Below is the model of working memory conceptualised by Baddeley.

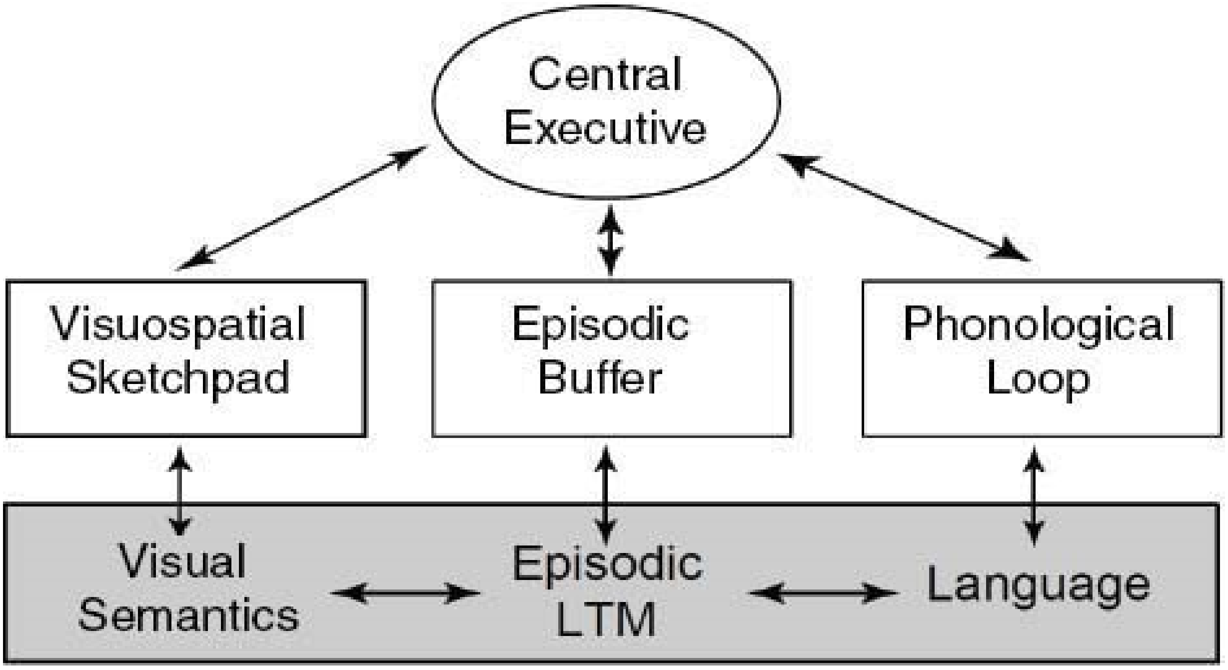


Figure 1. The working memory model (Baddeley (2000).

Explaining the model of Baddeley in more detail, it proposes that the central executive controls two slave systems: the visuospatial sketchpad, dealing with what is seen, and the phonological loop, dealing with what is heard. The phonological loop comprises the phonological store, which holds what is heard in memory (although this is prone to rapid decay), and the articulatory rehearsal component, which is able to refresh the heard information and keep it 'live' in working memory (whilst also allowing for manipulation of the information). How well these processes function will dictate the efficiency of individuals' auditory working memory span. A memory span can be defined as the number of items that can be recalled immediately after their observation or presentation and is an objective measure of an individual's performance. The more items that can be remembered, the better the memory span.

Common working memory span tasks for assessing performance include a listening span (LSPAN), where target information is delivered auditorily, and a reading span (RSPAN), where the information is in written or pictorial format. The tasks can be simple, with an individual recalling information that was previously seen and/or heard. This task can also be made more complex by the addition of a secondary task, such as recalling them in the order they were presented (forward, which is simpler, or in reverse, which is more challenging). Additionally, the imparted information can be made harder to perceive, which will also demand more cognitive support, thus making the task harder and the memory span shorter. There is a thorough discussion in Section 1.2.4.1 of the application of these tasks for researching links with speech-in-noise performance. Complex and simple LSPAN and RSPAN tasks are presented in all chapters of this thesis. Working memory and its support of listening in challenging situations have been well explored (Janse & Jesse, 2014; Ng et al., 2014; Rudner, Ronnberg, & Lunner, 2011), and evidence of a potential link is reported in Section 1.2.4.

For all of the cognitive functions described above, a potential link to speech-in-noise performance is well-conceptualised by the 'Ease of Language Understanding' (ELU) model (Rönnberg et al., 2013; Rönnberg, Rudner, Foo, & Lunner, 2008). The ELU theorises that where stored phonological speech representations are quickly matched, giving a rapid automatic binding of phonology (RAMBPHO), little cognitive support is needed for speech understanding. Where the speech is novel or degraded in some manner, explicit cognitive processing is required to understand the incoming speech. In these degraded listening conditions, those individuals with better cognitive performance may perceive speech-in-noise in a superior fashion. Further, when cognitive support is necessary, there is a reduction of the cognitive resources available for performing other cognitive activities. Additionally, if phonological matches become more automatic (following auditory training for example) this may allow for some of the cognitive support that was needed to assist in speech discrimination to be 'freed' up so it can be allocated to other processes. The ELU model is shown below.

Figure 2. The ease of language understanding (ELU) model (Rönnberg et al. (2013)

Having introduced the cognitive functions above and the link to a model of ease of language understanding, the next section describes how the control of these functions and the functions themselves are prone to change as an individual ages.

1.1.3. Cognitive changes with age

In parallel with declining hearing thresholds (Schmeidt, 2010), cognitive functions may also decline in an ageing adult (Darowski, Helder, Zacks, Hasher, & Hambrick, 2008). Though not necessarily the case for all older adults, declines are most often demonstrated in processing speed (general cognitive slowing), attention, working memory, executive functions and the overall control of these functions. These are all explored in more detail below. In circumstances of normal ageing, although this slowing may affect the performance of some tasks representing these functions, it should not affect the formation of new or well-established memories (Holland & Rabbitt, 1991).

There are several global explanations proposed for changes in cognitive performance due to ageing. One such proposal suggests that changes in performance are mainly due to a global slowing of neural transmission, resulting in reduced information-processing speed (Salthouse, 1996; Salthouse, 1985). Indeed, this slowing can be demonstrated when examining nerve conduction velocities in the auditory system (Goodin, Squires, Henderson, & Starr, 1978), which typically exhibit a reduction in transmission velocity recordings as an adult ages. In relation to cognitive performance, this neural slowing results in slower task execution, leading to information loss from earlier cognitive operations (Kail & Salthouse, 1994; Salthouse, 1994, 1996). If temporal co-ordination is affected by the slowing of some processes, the cognitive function of an individual may become inefficient, resulting in poorer cognitive performance. Slower reactions to sounds have also been demonstrated; for example, Tun, O'Kane, and Wingfield (2002) report that older adults have significantly slower processing speeds than younger adults do. The measures in the Tun et al. study included simple visual and auditory reaction tests and the digit symbol substitution test[2] (DSST). The reaction time tasks were simple sensorimotor tasks, whilst the DSST, a more complex task, assesses how many symbols can be matched with numbers in appropriate pairs within 90 seconds. For all of these measures, older adults perform significantly more slowly than younger adults do.

Craik and Byrd (1982) and Craik (1986) proposed a further global explanation: that ageing results in a reduction in available cognitive resources. This is demonstrated by a reduction in a person’s mental ‘energy’, leading to impaired performance in mentally demanding tasks (such as those involving working memory and divided attention), whilst tasks requiring little or no attention, such as automatic routines or skills, do not show such impairments. Automatic processing occurs when access to the required information is easily retrieved, for example, from recently used words (Logan, 2004). This model could readily account for the ease with which spoken utterances are matched, as predicted by the Ease of Language Understanding model. Controlled processing, conversely, uses cognitive resources to search for phonological matches. As control of where attention needs to be focussed declines with age (Bialystok, Craik, & Luk, 2008), it is possible to envisage how this may affect an older listener attempting to discriminate speech, potentially whilst switching from one speaker to another. This may be more problematic still for older adults if their pool of available cognitive resource is limited.

---

[2] The DSST is a task that is used to assess processing speed. The participant being tested is instructed to use information from a key of digit and symbol pairings to match a new series of letters or symbols with their appropriate pairs. They do this as quickly and accurately as possible.

Neuro-imaging studies provide evidence of compensatory mechanisms that allow older adults to counterpoise poor cognitive performance by recruiting both hemispheres to assist in cognitively challenging tasks (Cabeza, 2002; Reuter-Lorenz & Cappell, 2008). Younger adults are able to perform these tasks with activation focussed on only one hemisphere. There is also evidence of neural plasticity, as even where there is a degree of neural atrophy, there may be a subsequent over-activation of other neurological sites—particularly within the pre-frontal cortex—that allows for the maintenance of cognitive performance by older adults (Greenwood, 2007; Reuter-Lorenz & Cappell, 2008). It is, therefore, possible to theorise that for older adults when they are seeking extra cognitive support to assist with the discrimination of speech-in-noise, that the identified brain areas above may be recruited.

The above section has provided an overview of the theories of cognitive changes with age; the following sub-section briefly explores the relationship between declining hearing and a potential link to cognitive decline/dementia.

1.1.4. The relationship between hearing loss and cognitive decline/dementia

As this thesis does not seek to present new evidence on the relationship between cognitive decline/dementia and hearing loss, this section is purely included to highlight this important health issue. It also serves to promote the importance of offering good rehabilitation services to individuals with hearing loss, as high-quality rehabilitative strategies for individuals with hearing loss will potentially lessen any associated cognitive impact (Kiessling et al., 2003b).

Hearing loss often leads to inferior communication, which can cause an individual to withdraw from social interaction. Lack of social interaction has itself been demonstrated to be a predictor of cognitive decline/dementia in older individuals (Livingston et al., 2017). It remains the subject of some debate whether it is the inability to hear well that leads to cognitive decline, or whether it is individuals isolating themselves and avoiding social interaction because they struggle to hear or are embarrassed by their impairment. A further postulated hypothesis is that declining hearing is an indicator of a less-well-functioning neural network (Harrison Bush, Lister, Lin, Betz, & Edwards, 2015; Humes, Busey, Craig, & Kewley-Port, 2013; Lindenberger & Baltes, 1994). The study by Lindenberger and Baltes (1994) of a cohort from the Berlin Ageing study (BASE)[3] reported that, when vision and hearing acuity are combined, they account for 49.2% of the variance in intellectual functioning and 93.1% of the age-related variance in intelligence. Therefore, sensory functioning may be an accurate predictor of age-related variance in intelligence, and as such, may indicate the physiological integrity of the brain.

Middle-aged adults[4] were recruited for a later study that set out to determine whether declining auditory and visual senses hampered performance in cognitive tasks (Lindenberger, Scherer, & Baltes, 2001). The same tests were administered in the same order as in the older age study above, and participants had their sensory acuity artificially reduced by the use of occlusion filters. Compared to a control group and a placebo group, there was no significant reduction in cognitive performance demonstrated by the single-sense reduction group. This led the authors to argue that sensory impairments did not affect the results of cognitive assessments in the original research.

However, as the authors acknowledge, the artificial reduction of hearing by acoustic filters does not accurately mimic the sensory loss of the older individuals in the original research. Acoustic filters, despite raising thresholds, give a listener a conductive hearing loss; therefore, although the thresholds may be similar, the effects of presbycusis on the cochlea will not be observed, such as loss of sensory hair cells, a reduced endocochlear potential, and a reduction of auditory neurons. These cochlear changes induce much greater speech perception problems than the artificially created conductive loss. Although recruitment issues would pose challenges, the study would have been more persuasive if a group of middle-aged adults with sensorineural hearing loss had been recruited for the assessments.

---

[3] The Lindenberger and Baltes study reported on 156 participants 70–103 years old.

[4] 218 middle-aged adults between 30 and 50 years.

In a study by Lin, Ferrucci, et al. (2011), participants enrolled in the Baltimore Longitudinal Study of Ageing (BLSA) had their pure tone average hearing thresholds measured in the fifth year of this longitudinal study. They also had cognitive assessments in years 5, 8, 10, and 11. When measured at baseline (year 5), those participants who were identified as having normal hearing in their better ear performed significantly better on a global cognitive test battery (the Min-Mental State Exam (MMSE[5])), compared to the group with hearing loss. However, the difference between the groups was not significant for a test assessing executive function (the Digit Symbol Substitution Test, DSST), although the normally hearing group did tend to perform better on this test. When considered across all of the assessment years, the average annual change in scores for both the MMSE and the DSST indicated a significantly increased decline for the hearing-loss group of participants compared to the normally hearing group. Lin et al. (2013) further report that, on average, a decline for normal hearing individuals of five points on the MMSE would occur over 10.9 years, but for individuals with hearing loss, it would be just 7.7 years. A five-point change in the MMSE is considered indicative of cognitive impairment.

In a similar long-term ageing study conducted in France using self-reported hearing status (Amieva et al., 2015), a similar link between hearing loss and associated cognitive decline, as measured by the MMSE was identified. Additionally, Amieva et al. reported that hearing aids may offer positive protective effects against such decline. The Amieva study was somewhat limited in that hearing impairment was self-judged and this was only recorded at baseline so any potential changes in hearing status over the 25 years of the study were not noted. Further, participants simply reported whether they had hearing aids or not and there was no measure of usage. When considering the problems of getting patients to consistently use hearing aids (Smeeth et al., 2002) it does make the reported findings of hearing protection against cognitive decline more questionable.

---

[5] The MMSE is a 30 point questionnaire used to measure cognitive impairment.

In contrast to both the Lin et al. and Amieva et al. studies where hearing status was either measured or self-reported only at baseline, a pilot study from Washington county exploring atherosclerosis risk (Deal et al., 2015) measured hearing loss after cognitive data had been collected over a 20 year period. Reported findings, however, were similar in that the participants with hearing loss were at a greater risk of accelerated cognitive decline. The decline, in similar fashion to Ameiva et al. was also observed to be modifiable, to some extent, if a hearing aid was worn.

Below is a summary of the possible relationship between cognitive decline and hearing loss is from work by Wayne and Johnsrude (2015). The first model deals with cognitive decline driving hearing loss. Models 2 and 3 are the reverse in that hearing loss drives cognitive decline. Model 4 suggests that there is a common factor that is driving declines in both.

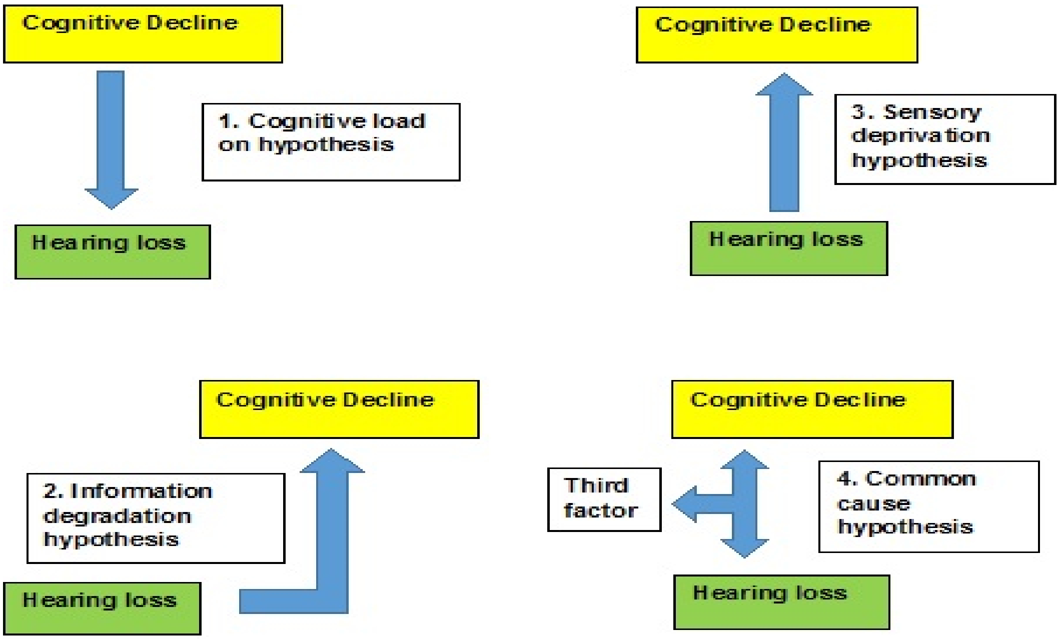


Figure 3. 4 models of the potential relationships between cognitive decline and hearing loss (Wayne & Johnsrude, 2015)

The above sections have given a general introduction to cognition and cognitive functions and the complex process involved in the perceiving of speech. Also covered was the fact that as an individual ages, there are cognitive changes. The literature review that follows explores the research into the role of cognitive support when listening to speech-in-noise. It goes on to review the evidence surrounding cognitive and auditory training.

1.2. Literature review

This section discusses the principal literature that informs this thesis. The first sections (1.2.1-1.2.4.) reviews previous research on the role that cognitive functions may offer in supporting speech-in-noise perception (Akeroyd, 2008; Besser, Koelewijn, Zekveld, Kramer, & Festen, 2013a). Sections 1.2.5.and 1.2.6 respectively explore the evidence surrounding cognitive and auditory training and how this may impact speech-in-noise performance (Au et al., 2015; Brouns, 2011; Henshaw & Ferguson, 2013).

1.2.1. Reviews of cognitive function and the relationship with speech-in-noise discrimination

In a review of 20 papers published since 1989, Akeroyd (2008) explored the relationship between speech reception and cognitive function. The review asks the question of whether individual differences in speech reception are related to differences in cognitive ability. Summarising the findings, Akeroyd reports that previous studies ‘clearly demonstrate some link between cognitive performance and speech reception’ (page S68).

A reliable relationship with speech-perception-in-noise performance was evidenced by tests of working memory (some of which involved executive function control). This was not the case for every study examined in the review, but the weight of the reported evidence supported such a relationship. Information processing speed was found in some studies to have a significant correlation with speech-in-noise tests (see Section 1.2.2), but IQ tests generally failed to demonstrate a significant relationship.

Overall, the evidence reflects that the degree of hearing loss is more predictive of any variation in speech-in-noise performance, rather than the cognitive measures. None of the papers reviewed by Akeroyd comment on the role of inhibition in speech-perception-in-noise performance. The control of inhibition, a key executive function introduced at the start of this section (1.1.2.3.), and its potential role in speech-in-noise perception is another of the key questions to be addressed by this thesis.

A later review by Besser et al. (2013a) looks at how linguistic closure (the ability to identify sentences where information may be incomplete) and working memory relate to speech-in-noise listening performance. The review focuses on research papers produced after the Akeroyd (2008) review. In addition to reporting on speech-in-noise tests, Besser et al. include papers where written text tests were used, specifically the text reception threshold test (TRT) (Zekveld, George, Kramer, Goverts, & Houtgast, 2007). In the TRT, black bars obscure some of the written text; the more black bars there are, the more difficult the text is to distinguish. This provides a measure of linguistic closure and may thus be related to speech-in-noise performance as part(s) of the target information may be missing. The more skilled an individual is at linguistic closure (i.e., deducing what a sentence may be, despite some of the text being hidden), the better they may be at completing a sentence that was only partially heard when presented auditorily. For the TRT, scores are based on how much of the text can be covered before the participant is unable to correctly identify 50% of the sentences. It gives a measure of an individual's ability to utilise top-down context to comprehend the information being delivered. The findings of the Besser et al. review are in line with those of Akeroyd (2008), although the review was not as comprehensive because it limited its focus to the relationship between working memory and speech (or text) in 'noise'. In general, Besser et al. reported that when one working memory span task (an RSPAN[6] and/or LSPAN[7]) was correlated with speech-in-noise performance, then if both were performed in the same piece of research, they both correlated. It was also found that linguistic closure (as measured by the TRT) was associated with speech-in-noise scores, but the relationships with speech-in-noise performance differed from those of RSPAN and LSPAN. The relationships also appeared to depend upon the type of speech-in-noise task undertaken.

For middle-aged and older adults, when the target speech was provided with a background of constant babble noise, both linguistic closure (TRT) and working memory (RSPAN) were

---

[6] This is a complex task where sentences are read whilst target information has to be remembered. The greater the number of pieces of target information remembered, the greater an individual's memory span.

[7] Similar to the reading span task but the sentences are delivered auditorily.

associated with speech-in-noise performance. However, when the speech had a background of fluctuating non-speech noise, linguistic closure (TRT) was the better predictor. The difference between these types of speech-masking noise is that when the background noise is babble, it is often comprised of real speech and thus more difficult to ignore. However, the natural babble of speech does have gaps between words, allowing for short periods of reduced noise when competing speech sounds/words may be ‘glimpsed’. If the background noise is non-speech or speech-shaped it is non-informational and is comprised of the different frequencies found in speech averaged over a period of time. As the noise is constant, it contains no dips, thus affording fewer of these ‘speech glimpses’.

A possible conclusion could be that when less speech can be distinguished from the non-speech noise individuals must rely more on the limited amount of speech they do perceive, so the task becomes more about general audition. Once they have heard any words, (or parts for words) they can then utilise their top-down semantic knowledge to fill in the rest. Alternatively, if sufficient speech is detected as is possible when the noise is background babble instead of steady noise, working memory may lend more support by keeping the heard information ‘live’ before it is lost, so matches to the spoken information can be sought. Sentences or words may then be understood.

Neither of the above reviews was able to fully address the still largely unexplained variation in speech-in-noise performance. Furthermore, where a relationship between speech-in-noise performance and a cognitive function has been demonstrated, the relationships have often been inconsistently defined by different researchers. The reason for the inconsistency has not been fully explored or explained but may be due to speech and cognitive task selection: how challenging a task is may affect which cognitive functions lend support. This will be tested in Chapters 3 and 5 where two speech tests will be employed, one with a steady state noise at a very challenging SNR and one with adaptive background speech babble. The hypothesis to be tested is that when speech is easier to pick up amongst background speech babble, there will be different relationships of different cognitive functions with the performance of the two speech-in-noise tasks.

Some of the cognitive functions described in the two reviews above have demonstrated a relationship with understanding speech-in-noise to a greater or lesser extent, particularly working memory performance. The different individual cognitive functions employed in this thesis that potentially may account for some of the variation in speech-in-noise performance are explored below.

1.2.2. The role of speed of information processing in understanding speech-in-noise

The cognitive function of information processing speed was introduced in Section 1.1.2.2. This current section explores the evidence surrounding its potential to explain individual variation in performance of speech-in-noise tasks.

Two pieces of research exploring the role of the speed of information processing in speech discrimination were included in the review of cognitive functions and speech-in-noise by Akeroyd (2008). One utilised a comprehensive battery of cognitive tests (van Rooij & Plomp, 1989), and the authors report that, after hearing thresholds, only auditory and visual reaction times (simple and choice) and a more complex information-processing (short-term memory) speed task[8] demonstrated significant correlations with speech-in-noise performance. There was no correlation between speech-in-noise performance and auditory or visual short-term memory, which was measured using a simple forward and backward digit span.

Conversely, the second study of speed of information processing in the Akeroyd review, by Jerger, Jerger, and Pirozzolo (1991), found that for a group of 200 older individuals with varying degrees of hearing loss, different cognitive functions (including information processing speed) did not demonstrate a significant relationship with speech-in-noise performance.

---

[8] Measured by a memory scanning task where participants had to decide as quickly as possible if a digit presented was part of an earlier memory set or not. The memory set included between two and four digits.

Desjardins and Doherty (2013) later reported that processing speed on a digit symbol substitution test was significantly correlated with speech perception in background noise in adults, as was working memory. The study recruited three groups of participants (15 young normally hearing adults (YNH), 15 older normally hearing adults (ONH) and 16 older hearing-impaired (OHI) adults, some with and some without hearing aids) to explore performance differences but only reported correlation results for all participants when grouped together. Unfortunately, as the study did not separate the results by age group, the strength of any correlational relationship with age group is unknown. Such an analysis would be useful for determining which age ranges rely more on cognitive factors such as processing speed for speech-in-noise tasks but would require greater numbers of participants recruited in each age range to make result interpretation statistically meaningful. This in turn may also clarify whether it is the general slowing of the older group (Salthouse, 1994) that differentiates older individuals' performance from that of the younger groups.

The Desjardin and Doherty research also demonstrates significant differences in processing speed using a non-auditory-based task between the YNH adults and the two older adult groups. There was also a significant difference in performance between the ONH and OHI groups (*p* =.043), with the ONH group performing faster. This is in line with research indicating that older individuals with hearing loss do not function as well cognitively as those with normal hearing (Deal et al., 2015).

The evidence presented above demonstrates that with age, there is declining performance in reaction/sensorimotor and processing speeds. The relationship between speed of information processing and perception of speech-in-noise was also established, leading to the hypothesis that training to prevent decline or improve performance of this function may result in better communication abilities for individuals. This forms the basis for Chapter 3 of this thesis, which explores processing speed training and seeks to ascertain whether this improves perception of speech-in-noise.

1.2.3. The role of inhibition control in the understanding of speech-in-noise

As noted in Section 1.1.2.3., control of inhibition may play a role in speech-in-noise performance, as the ability to maintain focus on a target whilst inhibiting unwanted interferences or stimuli is a key skill in understanding speech-in-noise. The Stroop test (Stroop, 1935) is a classic measure of prepotent inhibition. A review of Stroop test performance and speech-in-noise performance (Knight & Heinrich, 2017) concludes that there may be several confounding factors affecting the relationship between inhibition and speech-in-noise performance. These variables may include the speech test material and the way in which the Stroop test is performed and the score calculated. These factors are explored further in Chapter 3.

The relationship between inhibition and speech-in-noise performance has been studied using the neighbourhood activation model (NAM) (Luce & Pisoni, 1998; Pisoni, Goldinger, & Luce, 1988). NAM is explained in detail in Section 3.1.1., but in brief, it has been found that a word with many phonetically similar-sounding competitors (i.e., a high neighbourhood density) is more difficult to distinguish than one with fewer. In addition, frequently used words are easier to discriminate than rarely used words. There is an interaction between these two concepts, as the bottom-up sensory input and top-down information occur in decision-making units that select the appropriate word and ignore potential competitors.

Research by Taler, Aaron, Steinmetz, and Pisoni (2010) demonstrated a significant relationship between the Stroop test's interference measure and the effect of neighbourhood density when words are presented in a challenging signal to noise ratio[9] (SNR of -3 dB). This indicates that poorer inhibitory functioning is associated with greater differences in performance when identifying high- and low-neighbourhood-density words in challenging listening conditions. The same relationship is not demonstrated when the listening conditions

---

[9] Signal to noise ratios (SNRs) are defined as the volume ratio of speech to the background noise in which it is presented. For example, an SNR of -3 dB means the speech is 3 dB quieter than the background noise. Conversely, if the SNR is +10 dB, the speech is 10 dB louder than the background noise.

are at a more favourable SNR. Conversely, measuring inhibition with a Stroop task, Desjardins and Doherty (2013) did not identify a significant relationship with speech-in-noise performance. The Taler et al. study offers an important variation from Desjardins and Doherty, as its methods allow a comparison of inhibition in very challenging listening situations with that in easier listening situations, whilst the latter used three adaptive speech-in-noise tasks that may not involve the same listening load and did not tax the listener in the same way.

It appears, therefore, that a significant relationship between performance on speech-in-noise and inhibition tasks may be most evident when listening conditions are highly challenging, as under these conditions, cognitive support is increased as speech is harder to perceive. However, there is limited evidence that control of inhibition affects speech-in-noise performance; therefore, the scenarios in which control of inhibition assists an individual needs confirmation. For this reason, the speech-in-noise task chosen for the exploration of this relationship in Chapter 2 is set at a particularly difficult SNR, and this consistently difficult task with no adaptive element was employed throughout the research for this thesis. For later chapters in this thesis, the adaptive QuickSIN test in speech babble was also employed to determine whether the cognitive functions that support performance of the speech tasks differ under those conditions.

1.2.4. The role of working memory in the understanding of speech-in-noise

The control of working memory also comes under the umbrella of executive functioning. The Baddeley (1986) working memory model was introduced in Section 1.1.2.4., along with description of the tasks that can be employed to measure individual working memory span.

Souza and Arehart (2015) report a particularly robust relationship between working memory and speech-in-noise comprehension. The study involved 94 older participants (50–91 years), including some with hearing loss and some hearing aid users. The working memory task was an RSPAN, and the speech-in-noise task was the QuickSIN (which uses short, low-context sentences at changing SNRs), with the background noise a four-person babble. Listeners with poorer working memory had greater difficulty understanding speech-in-noise even after accounting for age and any hearing loss. Thus, those participants with better working memory spans appear to be potentially more able to hear speech in background noise.

Ng et al. (2014) examined the relationship between working memory and speech-in-noise for first-time fitted hearing-aid patients. Using a visually presented working memory capacity task, a significant relationship was demonstrated at the initial fitting, but when re-tested at three months and six months post-fitting, the relationship was less significant. The authors hypothesise that working memory was initially of importance for participants performing the speech-in-noise task, but as they acclimatised to their hearing aids, matches to newly stored representations of speech became more automatic (Rönnberg et al., 2013), which reduced the need for explicit processing and demand on cognitive resources. It could be anticipated, therefore, that this 'freeing' up of cognitive resources may allow for better processing of information or improved performance on other cognitive tasks. This thesis will investigate this by employing tasks that assess memory span, inhibition, processing speed, and other cognitive functions before and after interventions such as auditory or cognitive training.

Primary, or short-term, working memory is involved in the recall of material immediately after it has been delivered during conversations or speech assessments. For more complete communication, the information must be stored in secondary, or long-term, memory. When listeners are presented with speech in the presence of background noise, the noise interferes with the encoding of speech into secondary memory, which affects recall (Murphy, Craik, Li, & Schneider, 2000). Murphy et al. conducted five different experiments, using lists of paired words presented in either quiet or noisy conditions, with two groups of younger and older adults.[10] They investigated differences in working memory performance and encoding of information into longer-term memory for the two groups. The method involved the presentation of a list of five paired words in different listening conditions, with the listener asked to recall the corresponding word at a varying series of time points when given the other word of the pair. In

---

[10] The older adult group had some mild degree of hearing loss.

the first experiment, with the younger group alone, recall ability was affected by serial word position and the presence of background noise. There was the anticipated ‘recency effect’, with words positioned in fourth and fifth place being significantly easier to recall. Noise also had a significant effect on performance, as when all of the words were presented in quiet, they were significantly easier for participants to recall than when the words were presented in noise.[11] This was thought to be due to the more recently presented words (positions four and five) being held in working memory and thus remaining available to the listener. The deficit in performance for the earlier words was inferred to be due to failure to encode these as easily into secondary memory due to the noise interfering with the process of encoding.

The latter experiments compared the performances of the older and younger groups, again reporting an impaired ability to retrieve from secondary memory the word pairs in positions one, two, and three. This was after the researchers had adjusted the speech input level so it allowed all participants equal auditory access to the words. They proposed that the differences between the groups are due either to a reduction in sensory representation (although this should not have been an issue, as the participants’ access to the speech material was adjusted to make this equal) or to a decline in available cognitive resource. Thus, the diversion of cognitive resources in order to identify hard-to-perceive speech affects the performance of ‘on-line’ cognitive tasks and also appears to have a negative effect on memory recall. This is such that even if the speech in a given task is perceived and understood, the recall of the auditory information is impaired due to the impact of the reduced resources on encoding (McCoy et al., 2005; Rabbitt, 1991).

The concern here for older individuals is that even if they are able to follow a conversation, the information imparted will not be recalled later. The ability to recall previously delivered information is key if deeper communication is to take place as opposed to simple surface perception. For this reason, recognition tasks are performed throughout this thesis to assess

---

[11] Identification of the words at positions 1, 2 and 3 was also significantly more difficult at a harder SNR when compared to an easier SNR,

the relationships between heard information and secondary memory encoding, as well as to identify any effect of a training programme. It could be argued that if the information imparted is not stored by the receiver, then true communication has not taken place.

1.2.4.1. Choice of working memory span task when investigating speech-in-noise performance

Working memory span can be measured using different methods and in different sensory modalities. The classic complex working memory span measure (Daneman & Carpenter, 1980) involves participants being asked to hold information in their memory whilst performing a secondary task. For example, they may be asked to read aloud whilst storing the last word of each sentence in their memory. A simple straightforward working memory span task does not involve a secondary task. Both of these span tasks can be presented auditorily and/or visually and are measured by how many pieces of information a person is able to accurately remember.

There remains some disagreement as to the best tools to assess working memory span when working with people with any degree of hearing loss. If an RSPAN task is used, there will be fewer confounding factors, as there will be no difference in access to speech sounds that could be problematic if the material were presented auditorily. Conversely, if examining working memory in relation to perception of speech-in-noise, working memory should be measured using the same modality because the individual's working memory may be affected by their hearing loss. This may give a more accurate picture for that particular individual of the problems they have with hearing in noisy conditions. The same may apply even when participants in research present with similar hearing, as even small hearing differences may be sufficient to affect auditory access.

Some researchers investigating the relationship between working memory and perception of degraded speech (e.g., Foo, Rudner, Rönnberg, and Lunner (2007); Zekveld et al. (2011)) have used the RSPAN, whilst others have used an LSPAN task (Besser et al., 2013a). Other studies have used both SPAN tests (Baldwin & Ash, 2011; Koelewijn, Zekveld, Festen, Ronnberg, & Kramer, 2012).

A further complication for the question of which working memory span task to use is that a significant relationship has not been consistently demonstrated between the two modalities of reading and listening working memory performance (Besser, Koelewijn, Zekveld, Kramer, & Festen, 2013b; Koelewijn et al., 2012; Zekveld, Festen, & Kramer, 2013). For research incorporating both RSPAN and LSPAN tests, Smith and Pichora-Fuller (2015) recruited 72 participants in three groups[12] and reported that the range of performance was much wider for the LSPAN and that only the older-old group demonstrated a significant relationship between the two span measures.

The difference between the reading and listening modalities may be due to a difference between the transfer of cognitive resources needed for discriminating words for the LSPAN and that needed to read words for the RSPAN, resulting in different levels of performance for the two tasks. For example, if an individual has any degree of hearing loss, they may need cognitive support to hear what is being said for the LSPAN, but if their vision (even after correction) is good, there will be no cognitive support necessary for an RSPAN. This reasoning could also explain why these memory span measures may have different relationships with speech-perception-in-noise tasks, even with the same participants (Baldwin & Ash, 2011). Baldwin and Ash further report that the RSPAN was comparable for younger and older groups but that, again, the LSPAN was poorer for the older adults, as were the speech recognition scores. For the older group, the LSPAN predicted speech-in-noise performance and the RSPAN did not.

This relationship between the LSPAN and speech-in-noise performance for older adults may also be due to a decline in available cognitive resources. As some of the remaining resources are reallocated to support discrimination during the LSPAN, the older adults with better

---

[12] A younger listener group with normal hearing and two older groups, a ‘young-old’ group and an ‘old-old’ group, both with hearing loss.

discrimination reallocate fewer resources for discrimination, thus having more available for the working memory task. This is not the case for the RSPAN, as the task does not require the same amount of reallocation. This hypothesis will be tested in Chapter 4, where both an RSPAN and an LSPAN task are undertaken by participants.

The research presented above demonstrates that there is no clear consensus on which of the span tasks best establishes a relationship with speech-in-noise performance. The RPSAN is more easily controlled, but the LSPAN better reflects an individual's real performance when listening to speech. Where training is involved in this thesis, an LSPAN will be used to determine whether improving discrimination allows for improvements in working memory when presented auditorily. Furthermore, where the LSPAN is used, studies will employ degraded speech (or speech-in-noise) to add extra stress to the working memory task to explore whether this explains variation in speech-in-noise performance. An RSPAN will be employed where a working memory test without any confounding factors is needed.

1.2.4.2. Speech presentation for degraded speech tasks and LSPAN in this thesis.

When participants present themselves as having no hearing problems, all listening tasks will be performed without alteration or amplification of the speech material. Although this may slightly reduce a listener's access to the presented speech, the participants will have established phonological patterns already encoded in their auditory cortex; thus, when seeking phonological matches whilst performing any speech-in-noise tasks, they will be able to use these (Rönnberg et al., 2013). In Chapter 5, where the participants all use hearing aids due to notable hearing loss, all hearing aids will be set to a standard fitting rule to provide reasonably equal access to speech sounds.

1.2.5. The role of cognitive training in understanding speech-in-noise

As cognitive functions have demonstrated a relationship with speech-in-noise performance (Akeroyd, 2008; Besser et al., 2013a; Rudner et al., 2011), it would seem possible that training in these functions may allow for speech-in-noise performance to be improved, with resultant gains in an individual's well-being as they are happier to seek out social engagement.

Before describing the role of cognitive training in the understanding of speech-in-noise, there follows a brief introduction to the general research behind cognitive training for children and adults. The ideal of any cognitive training is the transfer of improvements to other functions and areas of performance. As working memory is a key component of fluid intelligence,[13] training working memory with the goal of demonstrating transfer to improved fluid intelligence would appear to be achievable.

Research by Jaeggi, Buschkuehl, Jonides, and Shah (2011) on junior school children trained using a working memory task found significant on-task improvements for the trained group compared to an active control. However, it was only those children who showed the highest ability after training that demonstrated transfer to a fluid intelligence task. A conclusion of the Jaeggi et al. study is that the question should not be whether cognitive training works, but rather which training regimes and conditions give rise to the best transfer.

Research with young adults (Jaeggi, Buschkuehl, Jonides, & Perrig, 2008) demonstrated a significant improvement in working memory, as measured by span tasks, for a group trained using the dual *n*-back[14] compared to the performance of an active control group. Fluid intelligence also significantly improved after training. In addition, the more training done on the dual *n*-back, the greater the gains in the tests of fluid intelligence.

In a meta-analysis by Au et al. (2015) of 20 studies of adult working memory training and potential gains in fluid intelligence, a small but significant positive effect was found. Although some studies did not find a relationship between improved working memory and fluid

---

[13] Fluid intelligence can be defined as the ability to use logic to solve novel problems in new situations.

[14] An *n*-back task is where a series of stimuli are presented one after another and the participant has to decide if the current stimulus is the same as one presented 'n' presentations before. The higher the 'n', the more difficult the task becomes.

intelligence, such a relationship does appear to be present overall. Those participants who perform poorly on the task prior to training tend to see the greatest improvements from it.

On a note of caution, Au et al. report that some of the sample sizes were small and that there was an influence of outliers. Additionally, some studies used passive controls, and where this was the case, the passive control and the experimental group exhibited larger differences. As taking part in training is known to create an expectation of improvement (Boot, Simons, Stothart, & Stutts, 2013), questions have been raised regarding this meta-analysis as other authors have found no or limited evidence of improved fluid intelligence after working memory training (Conway & Getz, 2010; Morrison & Chein, 2011; Redick et al., 2013). In response to the critique, Au, Buschkuehl, Duncan, and Jaeggi (2016) argued that any differences between treatment/active control and treatment/passive control have no significant bearing on the conclusions drawn.

As working memory appears to play a role in understanding speech-in-noise (Rudner et al., 2011), it could be hypothesised that working memory training may also lead to improvements in speech perception for some, if not all individuals. The same may also hold true for the cognitive functions of information processing speed and control of inhibition. To some extent, the success of any training is dependent on whether any improvements are transferable to other processes; i.e., not simply on-task improvements, but also communication enhancement in real-world situations. If transfer is possible, an assessment package that identifies performance weaknesses may allow for better targeting of areas for improvement, giving the potential for improved speech performance via this targeted training. Further research along these lines would seem advantageous and consequently a proposal is put forward at the end of this these for further investigation surrounding this area.

As an example of a study using cognitive training to improve speech perception in noise, Ferguson and Henshaw (2015a) reported on three studies. One of these involved a cognitive training programme (COGMED) that had a focus on working memory training. Previous research using COGMED has shown improvements for both younger and older adults in untrained attention tasks after a training period (thus demonstrating a degree of transfer) as well as self-reports of improved cognitive functions (Brehmer, Westerberg, & Bäckman, 2012). The participants in Ferguson and Henshaw's research were all hearing aid users, and after a period of visual working memory training, the trained group performed significantly better on auditory working memory tasks (some of which were novel) than an active control group, demonstrating a degree of near transfer. There was, however, no transfer to any improvement in speech-perception-in-noise performance.

It could be theorised that the lack of far transfer to the speech-in-noise task may have been due to the training being insufficiently intense or not challenging enough. Additionally, as the correlation of working memory and speech-in-noise performance is generally fairly weak, in order to demonstrate far transfer, any improvements from working memory training alone would have to be very significant. A fuller exploration of the evidence surrounding the potential for training in different cognitive functions to induce speech-in-noise improvements is given in Chapter 3.

The next section explores the evidence surrounding what, if any, benefit there is from undertaking auditory training as opposed to cognitive training. Access to speech in terms of hearing level is known to be the dominant factor in predicting speech-in-noise performance, as if there is a reduction in hearing level due to any degree of hearing loss, speech is less easily accessed (Akeroyd, 2008; Humes & Dunbo, 2010; Van Rooij & Plomp, 1992b).

1.2.6. The role of auditory training in understanding of speech-in-noise

Cognitive training, as described above, offers potential advantages for listeners when attempting to perceive speech under difficult listening conditions, however, success in evidence of far transfer performance improvement remains elusive (Henshaw & Ferguson, 2014). This section considers how auditory training may support individuals with perceiving speech under degraded conditions.

Auditory training can be defined as ‘a systematic procedure designed to increase the amount of information that a person’s hearing contributes to his total perception’ (Sanders, 1971, p. 205). Auditory training itself is not a new area, and contentions about its effectiveness are longstanding (Rubinstein & Boothroyd, 1987). Several recent reviews have examined the effectiveness of auditory training for patients fitted with hearing aids (Besser et al., 2013a; Brouns, 2011; Henshaw & Ferguson, 2013), and all identified a lack of rigour in at least some of the research, finding flaws in the designs and methods employed. They also highlight future directions for research. Three of the most common problems identified were a lack of a control group (particularly an active control group), a lack of randomisation, and a lack of calculation of the necessary sample size to demonstrate an effect. These points are addressed in the research conducted for this current thesis.

Auditory training has been performed with disparate groups of participants, such as children with hearing loss (Sullivan, Thibodeau, & Assmann, 2013), children with normal hearing (Millward, Hall, Ferguson, & Moore, 2011), children with auditory processing disorder (APD) (Cameron, Glyde, & Dillon, 2012), adult and child cochlear implant users (Fu & Galvin, 2007a, 2007b; Oba, Fu, & Galvin, 2011; Stacey et al., 2010), older hearing-impaired individuals (Barcroft et al., 2011; Humes, Burk, Strauser, & Kinney, 2009), and people using hearing aids (Sabes & Sweetow, 2007; Sweetow & Sabes, 2006a).

Auditory training can also be performed with normally hearing individuals, for the purpose of research. These participants can be of any age who report no problems with understanding speech-in-noise. In these cases, speech is adjusted to make it either novel (vocoding[15]) or difficult to understand (filtering speech or placing it within noise). Normally hearing participants who have undergone programmes of auditory training have shown improvements in speech recognition scores, both in quiet listening situations with degraded speech (Hervais-Adelman, Davis, Johnsrude, & Carlyon, 2008; Hervais-Adelman, Davis, Johnsrude, Taylor, & Carlyon, 2011) and for non-degraded speech against background noise (Burk, Humes, Amos, &

---

[15] The process of vocoding is dealt with in some detail in the upcoming methods section in Chapter 4.

Strauser, 2006; Song, Skoe, Banai, & Kraus, 2012). A further consideration surrounds the length and type of auditory training and the impact this may have on its success (Burk & Humes, 2007; Humes, Kinney, Brown, Kiener, & Quigley, 2014). A greater exploration of this somewhat contentious area is offered in Chapter 4.

Auditory training may be split into two broad pathways—analytic and synthetic—and may encompass a range of activities (Moore & Amitay, 2007). Analytic auditory training requires the trainee to make judgements on the recognition of speech sounds, such as between syllables/phonemes (Ferguson, Henshaw, Clark, & Moore, 2014) or between similar words (Humes et al., 2009). Synthetic training is based on the completion of sentence-type tasks, with participants asked to listen to a complete sentence and make a judgement on what they hear (Besser, Zekveld, Kramer, Ronnberg, & Festen, 2012). Analytic training aims to improve a listener's bottom-up processing by improving the representations of speech sounds in the hearing periphery. Synthetic training is intended to nurture the process of top-down processing, allowing listeners to make use of their syntactic knowledge. The majority of auditory or communication training uses a combination of the two (Sweetow & Sabes, 2006a).

One major drawback of this combined approach is that it is unclear which of the training exercises is effective for each individual. Indeed, it could be argued that if training is not directed where it is needed, the outcome could be detrimental to the trainee (Cameron et al., 2012). The picture is further complicated if training in cognitive tasks is also included.

In addition to the stated reasoning behind analytic and synthetic training, a further potential rationale for the occurrence of far transfer from training and its associated potential gains is that the training undertaken may benefit shared tasks. For example, being cognitively stimulated by auditory training—even if aimed at improving a skill such as perception of speech-in-noise—may also benefit performance of cognitive tasks (Ferguson & Henshaw, 2015c). Henshaw and Ferguson argue that it is the act of being stimulated by the training that may lead to any improvement.

In an analytic auditory training study where 'on-task' near and far transfer gains were explored, older adults with mild hearing loss[16] were trained to identify differences between 11 pairs of phonemes in quiet conditions (Ferguson et al., 2014). Compared to an active control group, the participants demonstrated a significant improvement in the phoneme discrimination task after a training period. This improvement in phonemic identification, however, was not combined with a subsequent improvement in two speech perception-in-noise tasks, thus demonstrating no evidence of far transfer.

An alternative view is that the phonemic training does not lead to improvements in the speech-in-noise assessments because the speech tasks themselves have a certain amount of task redundancy and thus do not allow for differences to be evidenced. In addition, if the phonemes that are trained do not occur regularly enough in the speech material used for the assessment (or in real world conversations), demonstrating a significant improvement will be problematic.

In the same research, following the auditory training, there was a reduction in hearing disability as measured by a self-reported questionnaire, the Glasgow Hearing Aid Benefit Profile (GHAB), which was used despite the participants being non hearing aid users. Of the four questions on the GHAB, the only significant difference between the trained and active control groups after training was in the reduction of disability in group conversations (with the trained group reporting larger improvements). There was, however, no significant difference between the groups for the hearing-in-background-noise question, which constitutes another complex listening situation. The authors propose that the improvement in complex cognitive tasks and the reduction in disability of listening in group situations, both indexing complex cognitive functions, highlights the potential further hidden benefits of auditory training. However, there is a degree of inconsistency in this argument, as although the GHAB improved, there was no significant improvement in the speech-in-noise tasks that may have been anticipated if there had been a real reduction in hearing handicap. Furthermore, this would have been a more convincing argument if the questionnaire provided evidence that there were improvements in

---

[16] The participants were not hearing aid users

all situations that involve complex listening situations. The GHAB's single questions regarding listening situations are also not as sensitive for investigating potential changes compared to some of the more exhaustive questionnaires available. An argument could be made that a more comprehensive questionnaire such as the full 'Speech, Spatial and Qualities Hearing Scale' (Gatehouse & Noble, 2004) or a structured interview may lead to different interpretations of the effects of the training.

Having expressed some caution around the interpretations, the Ferguson et al. research offers the hypothesis that analytic auditory training, whilst not leading to improved hearing, does promote better cognitive function, especially on complex cognitive tasks, and seems to improve self-perceived function. This present thesis looks specifically at whether auditory training can lead to working memory improvements and, conversely, whether cognitive training may lead to improved speech-in-noise performance.

Having discussed in the above literature review the evidence surrounding the role that cognitive functions play in speech-in-noise discrimination and the potential improvement that both cognitive training and auditory training may offer, the next section introduces information surrounding the size of the problem in terms of the incidence of hearing loss at a population level and how it potentially may lead to a poorer quality of life.

1.3. The size and nature of the incidence of hearing loss

The following sub-sections begin by introducing the size of the problem of old-age hearing loss in the United Kingdom and describing how, if hearing loss starts affecting communication, there may be associated quality-of-life problems. Since a common rehabilitation option for presbycusis hearing loss[17] is the fitting of hearing aids, the next sub-section focusses on the success or otherwise of hearing aids and the possibility that any improvement in audition may lead to subsequent cognitive improvements.

---

[17] Presbycusis is described in the upcoming Section 1.3.1

1.3.1. Hearing loss, impact on quality of life, and hearing aids

The older population in the UK is increasing, with the number of people aged 75 and over projected to rise by 89.3% to 9.9 million by mid-2039 (Office of National Statisitics, 2015). The proportion of the population over 70 that will have a noticeable hearing loss is 71% (Action on Hearing Loss, 2015). Ageing is frequently associated with a decline in hearing thresholds (as measured using a pure tone audiogram), particularly in the higher frequencies, termed ‘presbycusis’, which was first described by Zwaardemaker (1899) (cited in Lesperance and Burmeister (2008). There are several reasons why presbycusis hearing loss may develop with age, including the loss of sensory hair cells in the cochlea, declining endocochlear potential due to age-related changes in the stria vascularis, and a reduction of auditory neurons (Schmeidt, 2010; Schuknecht & Gacek, 1993; Yang, Schrepfer, & Schacht, 2015). The pure tone audiogram itself, which measures the quietest sound that an individual can hear at different frequencies and comparing these to a national standard, does not provide accurate information on how well speech can be discriminated (Vermiglio, Soli, Freed, & Fisher, 2012). It also does not predict how well speech is understood in a typical environment with background noise. There is strong evidence that a hearing-impaired listener will perform significantly worse in noisy spaces than in a quiet environment (CHABA, 1988; Dubno et al., 1984; Era, Jokela, Ovarnberg, & Heikkinen, 1986). This poses the question of whether there should be more emphasis placed on a different range of assessment options, including speech-in-noise performance assessment, rather than simply measuring hearing thresholds using a PTA.

Since the most frequent type of acquired hearing loss is presbycusis, the aforementioned elder population growth will have profound implications for the number of people presenting with hearing problems. Extrapolating the figures and with a knowledge of the incidence of deafness in different age groups, by 2035, there could be as many as 15.6 million people in the UK with hearing loss (Action on Hearing Loss, 2015). However, the exact age at which presbycusis will begin to manifest is very difficult to predict, as is the effect that this will have on speech perception performance. In a recent study, an exponential decline was reported in speech-in-noise scores after the age of 50 years (Moore et al., 2014). There is the potential, therefore that the previous hearing loss figures may underplay the real situation.

Hearing loss, however, may not necessarily be apparent in older-age individuals. Where an older individual's hearing thresholds remain normal (or close to normal), research demonstrates that the speech-in-noise performance of older listeners may still be inferior to that of younger listeners (Dubno et al., 1984; Pichora-Fuller et al., 1995; Schoof & Rosen, 2014). Schoof and Rosen note that this inferior performance is not consistent across all types of background noises, as there is a significant difference between older and younger adults when the speech is heard with 'babble' (with older adults performing worse) but not when it is presented with speech-shaped noise. The Schoof and Rosen research indicates that, with babble, older adults are potentially better able to, or more reliant on using cognitive support and 'top-down' processing skills to identify words when speech sounds are glimpsed amongst the noise. The implication of the speech testing results is that speech-in-noise performance is affected by variables beyond simple age and hearing level.

1.3.2. Hearing aids and the potential for cognitive improvement

As outlined in the previous section, hearing loss restricts communication, which may bring social and mental health problems. Therefore, there is a need to effectively manage this issue. Hearing aids, especially for non-reversible hearing conditions such as presbycusis, may offer assistance with communication and provide protection against the cognitive decline reported in hearing aid non-users with hearing loss (Amieva et al., 2015; Deal et al., 2015; Qian et al., 2016).

Although the fitting of hearing aids for individuals with hearing loss is reasonably common, this should not be the only intervention considered (Pryce & Gooberman-Hill, 2012; Pryce, Hall, Laplante-Levesque, & Clark, 2016). Other options include assistive listening devices such as amplified telephones and television listening aids, and counselling and training programmes

(Hawkins, 2005). In addition to hearing technology, communication and cognitive training could also be useful (Sweetow & Sabes, 2006a).

Although hearing aid provision is free of charge in the UK, their use is not as widespread as would be ideal (Liston, Solomon, & Banerjee, 1995; McCormack & Fortnum, 2013). Additionally, Smeeth et al. (2002) state that 40% of the hearing aids issued to patients by the National Health Service (NHS) are no longer in use. There are differing reasons cited for this. McCormack and Fortnum (2013) found that the two most important reasons were a lack of perceived benefit and discomfort. This was similar to the findings of Mizutari et al. (2013), who identified older individuals (over 65) in a community who did not have hearing aids but would benefit from them. Of the 1,414 people assessed, 103 had hearing aids, and a further 103 were identified who did not, but met the research criteria to be offered one. Of these, 68 consented to have a fitting. At subsequent visits of the newly unilaterally fitted 68 patients, five were lost to follow-up and 25 returned their aid after the research. The predominant reasons for returning the hearing aids were not wanting to use the aid or not seeing sufficient benefit. For the Mizutari et al. and McCormack and Fortnum studies, the cost of the hearing aid was not an issue.

The study by Mizutari et al. highlights an issue with hearing aid usage itself, as there was no cost involved and all patients received high-quality support. Although participants in the Smeeth et al. and Mizutari et al. studies had access to support services, neither of these studies explored whether simple acclimatisation to hearing aids was sufficient, or whether more formally structured activities or training (as explored in Chapter 5 of this thesis) or group support (Hawkins, 2005) could have led to improved hearing aid acceptance and use. One of the hypotheses to be tested in Chapter 5 is whether involving structured activities from day 1 after hearing aid fitting confers advantages over simple acclimatisation to the hearing aids without the employment of structured exercises.

There is some contention over whether or not there is a natural acclimitisation period to hearing aids. Many researchers have reported the need for an acclimitisation period (Munro & Lutman, 2003; Yund, Roup, Simon, & Bowman, 2006) whereas in contrast others have not (Arlinger et al., 1996; Humes, Wilson, Barlow, Garner, & Amos, 2002). There is limited research available surrounding acclimitisation to hearing aids that use non-linear processing with Dawes, Munro, Kalluri, and Edwards (2014) reporting the only paper available up until 2014 was the Yund et al. research. Dawes et al. when also reporting on acclimitisation of users to non-linear processing hearing aids, did not report speech recognition improvements over a 12 week period for hearing aid users when compared to unaided performance and a control group, thus indicating no great acclimitisation effect.

It would be of interest to explore if non-linear processing of sounds via hearing aids requires more or less acclimitisation time (if indeed any is needed at all) than previous reports on acclimitisation times for linear amplification aids. There is the possibility that as non-linear amplification attempts to mimic the functioning of the ear system in a more accurate fashion, then acclimitisation would be less of a problem, however, this may vary depending on degree of hearing loss.

Further studies examining the utility of hearing aids have attempted to determine whether their use, in addition to improving hearing levels, may have an impact on cognitive functioning and social well-being. If lack of sensory input affects cognitive functioning due to sensory underload (Sekuler & Blake, 1987), then re-establishing better auditory input could lead to improved cognitive performance. This ‘cascade’ hypothesis would be supported if individuals’ cognitive functions were seen to improve after hearing aid fitting, although any change would not be instantaneous and it would take time to reverse the sensory underload caused by the reduced audition. The rationale is that the sensory underload is compensated for by the hearing aid, thus facilitating cognitive improvement.

An alternative to the ‘cascade’ hypothesis would be the ‘common cause’ hypothesis (Lindenberger & Baltes, 1994). This states that general neural degeneration with age is responsible for concomitant declines in hearing, memory, and cognitive function. These theories have been explored above in Section 1.1.3. For the research reviews in this section, the ‘cascade’ hypothesis is potentially supported where after hearing aid fitting, there has been a significant improvement in cognitive functioning over some months or improved social well-being is self-reported. If, however, there is no significant improvement noted, or there is only a small degree of instantaneous positive change, the common cause hypothesis provides a better explanation. If a common cause is the underlying reason for declines in both hearing and cognitive ability, then when audition is restored, the limited pool of cognitive resources can be reallocated in the short term. It may appear, therefore, that as an individual’s hearing improves, their cognitive function also improves. However, this improvement will not be sustained or increased over time.

It should be further borne in mind that there is a finite amount of cognitive resources available for necessary cognitive functions, and this amount is known to decrease across one’s lifespan (Craik & Byrd, 1982). If resources must be reallocated to support important activities such as listening, less is available for other tasks, such as processing information into secondary memory, thus limiting deeper processing of any received information.

A narrative review of hearing aid fitting and cognitive changes was undertaken by Kalluri and Humes (2012). One of the seven studies reviewed was that of Mulrow et al. (1990), who reported that alongside the expected improvement in hearing performance, the hearing aid group saw significant improvements in emotional and cognitive functioning and a reduction in depressive symptoms when compared to a control group. These improvements were demonstrated using a range of self-report questionnaires and demonstrate the influence of improved hearing on other aspects of individuals’ lives, namely cognitive functioning and emotional well-being.

Studies by Tesch-Römer (1997) and van Hooren et al. (2005) were also included in the review; both failed to demonstrate global benefits attributable to hearing aid use, aside from the expected improved audition. The former study reported that six months post-hearing-aid-fitting, no significant self-reported differences in cognitive function, social activities, social relations, or general well-being were demonstrated compared to control groups. The latter study utilised a battery of cognitive tests rather than relying on self-reporting, and when the hearing-aided group's performance was compared to that of a hearing-matched non-aided group at six months and one year post-fitting, there was no significant improvement in any of the cognitive tasks. These tasks measured selective attention, attention focus, processing speed, verbal learning, and verbal fluency.

Other studies in the narrative review provide mixed evidence of the cognitive benefits of hearing-aid-fitting. Lehrl, Funk, and Seifert (2005) found a significant improvement in working memory capacity for a hearing-aided group compared to a matched unaided group after two months. However, there was no significant improvement in memory span or processing speed. Acar, Yurekli, Babademez, Karabulut, and Karasen (2011), in the same review, offered evidence of short-term cognitive improvement based on the Mini Mental State Exam (MMSE) for a group of older hearing-aided patients after three months of hearing aid use, although this study lacked a control group. The final study reported here, performed by Choi, Shim, Lee, Yoon, and Joo (2011), demonstrated a significant improvement for a hearing-aided group compared with a control group at six months post-fitting for the cognitive functions of recall and recognition, with no significant difference in learning ability or delayed recall.

Short-term gains in working memory performance following hearing-aid-fitting are also reported by Doherty and Desjardins (2015), in research published after the Kalluri and Humes review. Doherty and Desjardin's study involved a middle-aged and a 'younger older' group[18] of newly fitted hearing aid patients. Working memory was demonstrated to have improved at six weeks post-fitting when measured using two auditory working memory tasks. This was the case for both age groups, for whom the auditory working memory span improved in both quiet and background noise conditions. For a second memory task, a more complex *n*-back (where both a *1*-back and *2*-back were run), improvement was only seen in the younger older group, both in quiet and background noise conditions and this was only for the *1*-back. To analyse whether the cognitive function itself had improved, two age-matched control groups without

---

[18] Middle aged group 50-60 years and the younger older group 63-74 years.

hearing aids undertook memory span tasks. No significant difference in performance on the *n*-back task was found between the sessions for the control groups.

A reason the younger-older group improved on the *1*-back, and why the middle aged group did not was reported by the authors to potentially be due to ceiling effects as the middle aged group performed the *1*-back test very well at baseline. For the younger-older group, it would be of interest to explore, if in the longer term the more complex *2*-back also demonstrated improved performance. It may be that in order to demonstrate such improvements on more complex cognitive tasks that more time is needed to derive benefit from improved audition.

Overall, the Kalluri and Humes review and Desjardin and Doherty research indicate that there are cognitive advantages for newly fitted hearing aid users and that these tend to impact cognitive functioning immediately. They further report that there is limited evidence regarding the long-term positive impact of hearing aids on cognition, and that when studies focussed on this area have been performed rigorously, they have not observed a positive effect.

In summary, of the five papers in the Kalluri and Humes review for which outcomes were measured between 2 and 6 months after fitting, four demonstrated a significant cognitive improvement from hearing aid use (Acar et al., 2011; Choi et al., 2011; Lehrl et al., 2005; Mulrow et al., 1990). The one study that did not report a cognitive improvement measured outcomes at the upper end of the 6 month duration (Tesch-Römer, 1997). This would lend support to the common cause hypothesis, where improved audition allows for the reallocation of some cognitive resources, which is why cognitive improvement is seen fairly immediately. When cognitive benefit was measured after 1 or 2 years of hearing aid use, no significant effects were observed (van Hooren et al., 2005). This would not, therefore, seem to support the ‘cascade hypothesis’.

It is possible that for a full evaluation of the ‘cascade’ theory of sensory underload (Sekuler & Blake, 1987), a longer period of auditory input may be needed before cognitive gains from hearing aids are evidenced, as the sensory underload may have been manifest for many years. Thus, longer-term follow-ups for hearing aid patients are indicated, as a timeframe of six months or less may be insufficient, and research over shorter timescales may risk longer-term benefits being missed. If the sensory underload theory of Sekuler and Blake (1987) does hold true, longer-term follow-ups to assess the possibility of reversing the effects of sensory underload need to be better explored. Whilst not addressed in this thesis, a plan to explore these long-term benefits in future studies is offered in the final discussion.

A large-scale study by Dawes et al. (2015) of a subsample of the UK Biobank dataset (n = 164,770) explored cognitive function changes using a speech-in-noise task. The cognitive functions were reaction time, spatial and prospective memory, and reasoning. Hearing performance was measured using a triple-digit test, which was a speech-in-noise task that can be performed online. In the triple-digit test, the listener hears three numbers presented with a background of noise and responds by typing the numbers into a computer or telephone keypad. The background noise becomes more or less intense according to the participant’s responses, and a performance score is achieved. Controlling for age, sex, general health, and socioeconomic status, the hearing and cognitive task performance scores were combined with straightforward rating scales to assess hearing aid use, social isolation, and depression. In conclusion, hearing aid use (compared to non-use) was associated with better cognition.

Dawes et al. argue that cognition is associated with hearing aid use, independent of any association with social isolation and depression. It appears, therefore, that this superior cognitive performance is not simply a result of the general effects of social integration and less depressive symptoms. Dawes et al. suggest that hearing aid use may have two positive effects. The first is ameliorating the effects of auditory deprivation from hearing loss, which potentially reduces cognitive decline. Second, they argue that hearing aids may be sought out by more ‘cognitively able’ individuals, and that it is the self-efficacy of this more active group that influences the aided group’s performance on cognitive tasks. That is, the drive to obtain the hearing aids was due to being more cognitively active to begin with, naturally leading to a stronger cognitive performance.

Dawes et al. further report their surprise at the lack of an association between hearing aid use and a comparative reduction in isolation and depression. They argue that noisy situations may remain challenging even with hearing aid use, because background sounds are amplified by the aid to detrimental levels. Thus, when responding to a simple yes/no question on social isolation, hearing aid users state that social isolation has not improved, as they still wish to avoid these situations. The simple yes/no measures on social isolation may not have captured the actual situations or level of isolation and any potential reasons for the loneliness experienced has to be inferred. Therefore, a measure that provides multiple levels or a continuous variable may be a more useful tool to explain the isolating effects of hearing loss and why hearing aids may not offer any advantage.

Although sampling a large database allows for greater statistical power, the number of participants and the necessity of automation of tasks and questions mean that subtler relationships may be missed. The simple yes/no answer to the social isolation question and the triple-digit test are two such examples. The triple-digit test does not allow for top-down processing or working memory to be adequately explored, as sentences in noise-based tasks would do. Further, authors of such large-scale studies must interpret the data rather than following up with interviews to explore the reasons behind the responses given. Although the number of participants in the research for this thesis is small (albeit appropriately powered), detailed testing and questionnaires enabled exploration of possible relationships that would otherwise need to be interpreted from large-scale study data. That is not to say that these large-scale studies do not play a key role in providing valuable information, as the links they report between hearing, cognition, and cognitive decline have brought this important health issue to the attention of clinicians and the general public alike.

The next section sets out the aims and hypotheses explored in this thesis.

1.4. Present research (aims and hypotheses)

Thus far, this thesis has introduced several areas in which the knowledge base on the perception of speech-in-noise is either weak or contradictory. There is also limited conclusive evidence on the role of auditory and cognitive training in improving speech perception in noise or in potentially releasing cognitive resources. Although the evidence is limited regarding the success of cognitive training (in particular) for improving speech-in-noise performance, the underlying theoretical background suggests that it could work. This thesis aims to address some of the gaps in the knowledge surrounding this area. Further explanation of the variation in performance between older and younger adults is required, as is evidence for the rehabilitative measures that may assist in improving an individual’s communication. This is particularly the case for older adults, as improved communication may facilitate an improved quality of life if their tendency has been to withdraw from social interaction.

Theoretically, it is anticipated that individuals who have better speech perception abilities and/or are able to better employ cognitive support, as hypothesised by the Ease of Language Understanding model (Rönnberg et al., 2013), will be the better speech-in-noise performers. If this is the case, training to improve cognitive function (to enhance cognitive support) will allow for improved speech-in-noise performance. Conversely, improving phonological representations (using auditory training) will necessitate less cognitive support, as speech recognition is more automatic. It is hypothesised that both these training regimes will lead to improved speech-in-noise performance, with the added benefit of releasing cognitive resources for other tasks such as encoding information into secondary memory. It is reasonable to assume that directly targeting the cognitive or auditory discrimination area where individual performance is weakest will allow for more targeted training and consequently better outcomes.

Furthermore, it is anticipated that older listeners with any degree of hearing loss (even those who have not sought help) will allocate more cognitive resources to support speech understanding than will younger listeners in the same situation (Füllgrabe & Rosen, 2016b). It is also hypothesised that the older individuals who have been best able to compensate for any cognitive decline will maintain superior speech-in-noise performance.

The overall research for this thesis is organised according to two main themes. First, this work explores which independent variables demonstrate a significant relationship with speech-in-noise performance and how this may explain any variation in speech perception. Second, where relevant, the effectiveness of any training intervention undertaken by participants is reported. The main objective is to identify which type of training delivers the best outcomes for individuals in terms of speech-in-noise performance and whether this frees up cognitive resources. The aims and hypotheses to be tested in each chapter are provided below.

Chapter 2: Speech discrimination in noise and correlation with control of inhibition and speed of information processing.

A wide age range of adults with no self-reported hearing loss were recruited for the research for this chapter. The purpose was to explore which cognitive functions may offer support when attempting to discriminate speech in very challenging listening conditions, with the speech presented in speech-shaped noise at -8 dB SNR.

Speed of information processing tasks and both pre-potent and ‘real-time’ control of inhibition tasks were employed to ascertain which of these functions may assist with an individual’s speech-in-noise discrimination performance. To minimise the effect of semantic knowledge on speech-in-noise performance, short sentences were used for which there was limited contextual support. The sentences used, such as ‘The glass bowl broke’, were taken from the BKB sentences for partially hearing children (Bench, Kowal, & Bamford, 1979). The challenging SNR afforded fewer glimpses of any heard speech than would a more favourable SNR, and utilising speech-shaped noise as the masking noise instead of babble allowed for limited access to the target speech. It was hypothesised that those individuals who were better able to respond and process information quickly and inhibit both prepotent responses and unwanted extraneous signals would also better perceive speech-in-noise.

The hypotheses to be tested are as follows:

H2.1. If the control of inhibition and speed of information processing play a role in the understanding of speech-in-noise, then the tasks that measure these cognitive functions will correlate significantly with speech-in-noise discrimination performance.

H2.2. If the cognitive functions noted in H2.1 provide an explanation for variation in speech-in-noise discrimination performance, then they will also add substantially to any explanation of performance variation once age, PTA, and working memory have been factored out.

Chapter 3: The potential role of inhibition and processing speed training in improving performance for speech-in-noise perception and other associated cognitive tasks

The primary aim of this chapter was to investigate whether cognitive training offers any advantage for individuals without hearing loss (hearing aid non-users) in perceiving speech-in-noise. It is hypothesised that if the cognitive functions of speed of information processing and control of inhibition contribute to speech-in-noise perception, training of these functions will thus improve speech-in-noise discrimination performance. Participants were randomly assigned to one of three training pathways, one involving a computer-based program to train individuals' visual processing speed and inhibition control, the second involving a computer program to improve individuals' auditory processing speed and inhibition control, and the third an active control pathway that involved completing a cognitive workbook. The use of three groups means that any differences in training gains between the different types of training can be explored. The hypothesis was that training auditorily will provide the greatest advantage for speech-in-noise discrimination performance, as training gains are generally task-specific, and it is anticipated that the same will be true for sensory modality.

As the cognitive function tasks and subsequent training were visual- or auditory-based, this affords the opportunity to analyse at baseline whether visual or auditory cognitive function tasks are better predictors of speech-in-noise discrimination performance. As two types of speech-in-noise task were used, it also afforded the opportunity to explore which cognitive functions were most useful in the different types of background noise tested. Further, there was also the opportunity to explore which of the two types of inhibition, prepotent or more ongoing 'real-time' inhibition, is more closely associated with the ability to hear speech-in-noise.

The hypotheses to be tested were as follows:

H3.1 If training on the cognitive functions of information processing speed and inhibition control results in improved speech perception in noise, then the groups that are cognitively trained via computer will demonstrate significantly better SIN performance than the active control group trained on general cognitive exercises.

H3.2. If training in a sensory modality allows for better performance of tasks in the same modality, auditorily trained participants will perform auditory cognitive-based tasks significantly better after training than will visually or control-trained participants. The reverse will be true for any vision-based tasks.

H3.3. If cognitive functions are most useful when glimpses of speech can be more readily perceived, then cognitive functions will correlate better with a ‘speech-in-babble’ task than a ‘consistently challenging speech in speech-shaped noise’ task presented at a particularly adverse SNR.

H3.4. Auditory tasks will predict performance in speech-in-noise tasks at baseline better than visual tasks will.

H3.5. If real-time inhibition control is the dominant cognitive performance factor influencing speech-in-noise discrimination performance, then assessments utilising this function will correlate to a greater extent with speech-in-noise performance than those utilising the inhibition of prepotent responses will.

H3.6. If ageing and small declines in hearing levels are not the only predictors of speech-in-noise discrimination performance, then once age and hearing levels have been controlled for, cognitive functions will add to the explanation of any variation in speech-in-noise performance.

Chapter 4: Exploring auditory training in degraded speech, near transfer, and release

If auditory training is to be truly successful, evidence is required of transfer to different degraded speech material as well as targeted improvements in the trained speech material. This would indicate that transfer to real-world performance is possible. Furthermore, working memory appears to offer support in understanding speech-in-noise, at least for some individuals; therefore, if listening becomes easier due to training, some cognitive resources previously used to assist with understanding of speech-in-noise will be released. The freed resources may allow for better encoding of received information into the secondary memory for later use.

The primary aim, therefore, is to investigate whether training in one type of degraded speech leads to discrimination improvements in that particular type of degraded speech ('on-task') as well as in other types of degraded speech (near transfer). A further aim is to investigate whether there is a release of cognitive resources when an individual's discrimination performance improves. There is also the theoretical possibility that the simple act of undertaking any training will be cognitively stimulating and that this will allow for improved performance on unrelated tasks.

The hypotheses to be tested are as follows:

H4.1. If working memory plays a positive role in discriminating speech-in-noise, then there will be a significant relationship between participants' performance in the speech-in-noise task and working memory at baseline, and the relationship will be stronger for a listening span (LSPAN) than for a reading span (RSPAN) task.

H4.2. If training in discriminating degraded speech leads to improved phonological representations, then speech discrimination performance for that type of degraded speech will significantly improve, above the improvement of groups offered different training.

H4.3. If training allows for improved discrimination performance in one type of degraded speech, then the process of auditory training may also allow for significantly improved performance in discriminating other types of degraded speech, thus demonstrating a degree of near transfer.

H4.4. If discriminating degraded speech becomes less effortful and more automatic, then any released cognitive resources will allow for significantly improved performance of speech-based working memory tasks in the same type of degraded speech.

H4.5. If the act of undertaking a training regimen itself is cognitively stimulating, then there will be a significant improvement in performance of non-related cognitive tasks following a training period.

H4.6. If discriminating degraded speech becomes less effortful and more automatic, then processing of heard information will become simpler, allowing for better memory of the delivered information.

Chapter 5: The role of a commercial auditory training package (LACE) in the rehabilitation of newly fitted hearing-aided patients

As previous chapters have explored whether cognitive and auditory training offer an individual advantages in improving their speech-in-noise perception, this chapter tested the theory that offering a combination approach of the two types of training to newly fitted hearing aid users may offer success. The computer-based Listening and Communication Enhancement (LACE) training package aims to improve an individual’s hearing ability using such a combination approach. The primary aim of this chapter is to determine not just whether LACE training results in ‘on-task’ improvements for newly fitted hearing aid patients, but also whether any improvements lead to improvements in deeper processing (the encoding of auditory information into secondary memory) and the release of cognitive resources. This will be tested by assessing whether recognition of auditory material and working memory performance improve after LACE training. If improvements in speech perception do release cognitive resources previously used for discrimination purposes, this should allow for improved working memory span and secondary encoding.

An additional aim of this section is to assess which (if any) of three care pathways (LACE training, placebo training and no training), after hearing aid fitting and training, is associated with an increase in participants’ self-ratings on a 12-item questionnaire that assesses different listening skills,. Furthermore, the self-rating judgements can be compared with the individuals’ performance on the different speech-in-noise tasks. It is anticipated that participants on all pathways may self-rate as improving, as undertaking any training or receiving a new intervention (such as hearing aids) may allow for the anticipation of success (Crow et al., 1999). If the gains are real, however, then any change in self-rating performance will correlate strongly with improved task performance post-training.

The final aim is to establish which cognitive functions explain the variation in speech-in-noise performance for hearing aid users, after controlling for age and hearing loss. This is also expected to reveal whether the relationship of cognitive functions with speech-in-noise perception differs from those noted in the portions of this thesis that recruited unaided participants with normal (or near normal) hearing. It is anticipated that older patients with hearing loss will utilise more of their remaining cognitive resources to support their listening, as although hearing aids give patients better levels of audition, the aids are still amplifying these speech sounds to an inner ear that has undergone physiological changes due to age (presbycusis). As a result, although speech may be heard, the broader tuning peaks on the basilar membrane may provide reduced clarity, such that phonological mismatches occur with previously stored information, necessitating explicit cognitive processing to support discrimination. This would be exhibited by a greater range of cognitive functions demonstrating a relationship with speech-in-noise performance compared with the ranges from other chapters where participants' hearing levels were normal (or close to it).

Two types of speech-in-noise tests were employed: adaptive speech in babble and constantly challenging speech in speech-shaped noise. The different cognitive support mechanisms for listening under these different conditions can thus also be explored. The cognitive functions examined are working memory, information processing speed (a response latency task), and control of inhibition. To control as much as possible for individual variation, variables such as type and degree of hearing loss, type of hearing aid, length of time of hearing aid usage, and fitting strategy have all been standardised.

The hypotheses being tested to achieve the above aims were as follows:

H5.1. If auditory and cognitive training, as delivered by LACE, correctly target the areas needed to improve speech-in-noise performance, then the LACE-trained pathway group will perform the speech-in-noise tasks significantly better at the second assessment compared to the other training pathways.

H5.2. If speech discrimination improves following training, then there will be a subsequent release of cognitive resources that allows for significantly improved performance on auditory-based memory span tasks (both simple and complex) and a speech recognition task of previously delivered auditory information.

H5.3. If LACE training offers significant advantages over placebo and no communication training, then there will be a significant difference in self-rating improvements between the pathways, both in general and for specific listening situations such as competing noise, sound quality, spatial awareness, and effort.

H5.4. If improved performance affects perceived hearing efficacy following hearing aid fitting, then the self-rating improvements noted by the participants in each of the pathways will correlate with the participants' speech-in-noise task performance.

H5.5. If cognitive functions assist with discriminating speech-in-noise for newly fitted hearing aid users, then the tasks assessing the different cognitive functions will significantly correlate with speech-in-noise performance.

H5.6. When testing H5.5, different sets of significant relationships will be identified for the two types of speech-in-noise tests, as steady-state speech noise and fluctuating noise may necessitate the support of different cognitive functions.

H5.7. If the speech performance variation is due to cognitive functions themselves and not raised hearing thresholds or the effects of age, then once age and hearing level have been accounted for, the tasks assessing the cognitive functions will significantly add to the explanation of speech performance variation.

Chapter 6: General discussion

One of the theories tested in the course of this thesis was that when listening conditions were particularly challenging, individuals would need to rely on cognitive support mechanisms in order to perceive speech-in-noise. It was theorised that this may involve the functions of information processing speed and inhibition control. It was further theorised that if the above functions assisted with perceiving speech-in-noise, then training those functions may lead to improved performance of speech discrimination tasks.

The thesis additionally theorised that following auditory training, there would be improved speech performance for different types of degraded speech material. It was further theorised that gains in perceiving one type of speech following training would transfer to improved performance for other types of speech material. It was also proposed, however, that this will only become evident when the training is particularly challenging, as the more challenging the task, the greater the development and strengthening of new neural auditory representations. Furthermore, the improvements will only be evident when the outcome measures involve auditory-based tasks.

It also proposed that as listening became more automatic, there would be a subsequent release of cognitive resources to perform other cognitive tasks.

Lastly, a proposed model for a bespoke training package is put forward. If targeted training offers the greatest advantages, training designed to meet an individual’s specific needs will offer greater improvements and is thus worthy of further exploration.

Appendix

The appendix contains a summary report of an extra study performed as part of the background to this thesis that involved a sub-group of participants that were recruited for the research performed in Chapter 4. This sought to explore the potential relationships between speech-in-noise performance and the neural representations evoked by a speech sound. Previous research has demonstrated that both the amplitude and latency of speech neural representations may assist in explaining some of the variation noted in individual performance in understanding speech-in-noise. The hypothesis for this research is that the amplitude of neural responses generated by the speech sound /da/ (namely, the fundamental, first formant, and overall amplitude) will have a significant relationship with speech-in-noise performance.

Additionally, a further hypothesis is that since cognitive functions rely on neural transmission, better auditory neural representation may be indicative of superior cognitive function, which could be demonstrated by a longer working memory span. There is also the possibility of an alternative relationship: if better neural representations allow for better speech-in-noise performance, those individuals may rely less on other cognitive functions, such as working memory. These participants may therefore have lower working memory capacity but good speech perception in noise. The report on these neural representations and their relationships with speech-in-noise performance or cognitive functioning is presented in the appendix because no significant relationships were demonstrated.

CHAPTER TWO

2. Speech discrimination in noise and correlation with control of inhibition, speed of information processing and working memory

2.1. Introduction

If working memory research is excluded, there has been a general under-exploration of the involvement of cognitive functions in speech-in-noise discrimination. This led to the research developed for this chapter. The primary aim of the research was to determine if the cognitive functions involved in speed of information processing (response latency tasks) and inhibition control contributed to speech-in-noise performance. Working memory was also considered, as it has been found to have the most reliable relationship with speech-in-noise performance (Akeroyd, 2008). By including it in this research, and after accounting for any noted relationship between working memory and speech-in-noise performance, any remaining significant relationships can be better established.

Response latency, as a measure of information processing speed, is potentially linked to speech-in-noise discrimination ability in that a faster processing speed would allow for quicker access to speech sounds, such that speech glimpses heard amongst background noise can be acted upon quickly. This would allow, for example, phonological matches to be identified quickly for spoken utterances. Quicker processing speed would also enable speedy entry of information into working memory for any necessary manipulation. In turn, attention could be paid to a particular source, focusing on target speech quickly and being able to speedily switch between speakers. All of these abilities would seem to offer advantages for perceiving speech-in-noise.

Intuitively, it would also seem that control of inhibition should play a role in speech-in-noise performance. The ability to inhibit unwanted stimuli and focus on a target should offer the opportunity to attend to a speaker while ignoring any background noise. This can be measured by the ability to inhibit prepotent responses as well as to inhibit in ‘real time’ by concentrating on a target and inhibiting the intrusion of other unwanted information.

In the general introduction to this thesis in Section 1.1.2., the cognitive functions of processing speed, inhibition control and working memory were introduced. The control of these functions (Hegarty, Shah, & Miyake, 2000; Miyake et al., 2000) and changes that may occur with age (Troyer et al., 1994) were introduced in the same section. The potential relationships of these functions with speech-in-noise performance, alongside the role that executive functioning may play in perceiving speech-in-noise, were discussed in the literature review, Section 1.2. As each of the above functions may play a role in understanding speech-in-noise, the evidence surrounding this is explored in more detail in the subsections that follow here, after a general review of other factors that may have an influence on an individual's ability to perceive speech-in-noise.

It is not necessarily the case that even small hearing losses will lead to individuals having poor speech perception in noise. Amongst groups of individuals of varying ages with similar hearing thresholds, some listeners are able to discriminate speech in background noise much more readily than others (Veneman et al., 2013). It has further been reported that when older adults are matched for their ability to hear speech in quiet conditions, there is still a wide range of variability within the group with regards to their ability to then hear speech in background noise (Duquesnoy, 1983). It appears, therefore, that there are other factors involved that allow some individuals to perform much better than others. The hypothesis for this chapter is that information processing speed and control of inhibition may account for some of these observed differences.

There is evidence that performance on tasks involving inhibition control (Hasher & Zacks, 1988) and processing speed (Salthouse, 1996) declines with age. It can, therefore, also be envisaged that older individuals who are better able to maintain performance of these functions will be best protected against declining speech discrimination ability due to even slight age-associated hearing changes.

It could be anticipated that if the common cause hypothesis of general slowing with age (Salthouse, 1996) holds true, and information processing speed plays a role in speech-in-noise discrimination performance, then individuals with slower processing speed may exhibit poorer speech-in-noise discrimination. In similar fashion, if the control of inhibition (Hasher & Zacks, 1988) plays a role, then this may also influence speech-in-noise discrimination performance.

The theory of declining mental energy with age, as proposed by Craik (1983), suggests that ageing results in a reduction of the amount of cognitive resources available to an older person. If this is accurate, then it will again be older individuals who will show a greater variation in performance on speech-in-noise tasks if they rely on these cognitive functions to assist in their discrimination ability. As speech-in-noise discrimination is a complex task and reduction in a person's mental energy leads to impaired performance of tasks that are mentally demanding (such as working memory and divided attention tasks), it is possible to imagine how ageing would affect speech-in-noise discrimination performance. It is also noteworthy that tasks requiring little or no attention do not show such impairments, such that when listening conditions are optimised, older individuals may show no decrements in speech discrimination performance.

In a comprehensive report on cognitive functions and speech-in-noise discrimination performance, it was concluded that the review had 'clearly demonstrated that there is a link between cognition and speech reception in noise' (Akeroyd, 2008, p. S65). Akeroyd reported on a wide range of hearing losses.[19] The review classified the cognitive measures into a) general scholastic ability, b) standard IQ tests and tests of memory, c) tests of working memory, d) miscellaneous tests, e) simple and choice reaction times and f) visual analogues of speech reception. A limitation of the report, although not a criticism, was the lack of reports on the role of executive function in speech-in-noise performance, as there was no research on that topic identified up to the report date. As can be seen from the above cognitive groupings, working memory and speed of information processing research did form part of the review. These functions, and the way they apply to the research in this chapter, are considered alongside other, more recent research in Sections 2.1.1 and 2.1.3, respectively. Inhibition

---

[19] Some participants used hearing aids and some did not.

control as part of executive functioning and its relationship with speech-in-noise performance is discussed in Section 2.1.2.

Of the other cognitive functions listed above, where measures of intelligence were explored for their link to speech understanding in noise, the intelligence measures tended to be standard crystallised measures and not measures of fluid intelligence. Scores on the Scholastic Aptitude Test (SAT) for verbal and mathematical skill (with normal-hearing young adults) failed to demonstrate a significant correlation with speech-perception-in-noise performance (Kidd, Watson, & Gygi, 2007; Surprenant & Watson, 2001). One paper did report a minor relationship between intelligence quotient (IQ) and speech perception in noise (van Rooij & Plomp, 1990), but only when the IQ score was used as part of a canonical regression model attempting to explain variance in speech-in-noise performance. Being limited to information on the relationship between crystallised intelligence and speech-in-noise perception restricted what the review could report, as fluid intelligence could potentially demonstrate a more powerful relationship. The reason for this possibility is that as fluid intelligence involves the ability to overcome novel problems with the use of logic, applying this skill to develop a strategy to focus on one speaker whilst ignoring others, and then switching targets whilst maintaining a conversation thread, would seem to draw upon fluid intelligence skills.

For the research in this chapter, it was decided to employ a speech-in-noise task set at a consistently challenging -8 dB SNR in speech-shaped noise. In real life, this could be envisaged as going from a quiet environment to having to listen to a speaker in a room that was consistently noisy, such as a room where a party is in progress. Much of the previous research has used adaptive background noise, allowing for a ‘tuning-in’ in to the target speaker. Making the speech task non-adaptive by removing this ability was expected to show that individuals who could quickly adapt by processing information quickly and efficiently inhibiting the intrusion of background noise would discriminate speech best.

The next sub-sections initially explore the roles of processing speed, control of inhibition, and working memory with speech-in-noise discrimination performance. These sub-sections discuss how the previous literature assisted in designing the research carried out in this chapter.

2.1.1. Information processing speed

The role of information processing speed and its relationship with the ability to perceive speech-in-noise was introduced in Section 1.5.3.1. The research in this chapter was also informed by the literature below.

Desjardins and Doherty (2013)[20] tested participants using three different speech-in-noise tasks and a cognitive test battery consisting of a working memory task (RSPAN), a processing speed task (Digit Symbol Substitution test) and a selective attention task that taps inhibition (a Stroop task). The tests of working memory and processing speed correlated significantly with three speech-in-noise tasks[21] when all participants (younger adults, older adults and older adults with hearing loss) were considered together. In similar fashion, Schoof and Rosen (2014) demonstrated that both working memory and processing speed were related to speech reception threshold in noise. The speech test employed for this research was the Institute of Electrical and Electronic Engineers (IEEE) Harvard sentences (IEEE, 1969), presented in either noise or speech babble. The task required participants to identify five key words in sentences in an adaptive noise testing paradigm. This evidence of a relationship was, however, only significant when the speech reception threshold test was performed in noise, not when the speech was embedded in speech babble.

Moore et al. (2014) reported that cognitive measures used in the UK Biobank study of 40- to 69-year-old participants were independently associated with both age and speech-in-noise discrimination performance. The speech test employed was a digit triple test (DTT) in which listeners had to identify three digits amongst changing background noise levels. There were a number of cognitive measures undertaken, one of which was an information processing speed

---

[20] With three groups made up of younger adults, older adults and older adults with a hearing loss

[21] Speech in 2-person babble, in 6-person babble and in speech-shaped noise involving last word in sentence identification in adaptive noise.

task (choice reaction) asking participants to press a button when pairs of shapes matched. Other cognitive tests also demonstrated a relationship with speech-in-noise perception performance, including memory span and fluid intelligence, both of which had previously been identified as being associated with speech-in-noise discrimination ability (Akeroyd, 2008). Participant performances on all of the cognitive tests demonstrated relationships with speech reception thresholds. Across the age ranges, better cognitive functioning accounted for a 0.7 dB reduction in DDT reception threshold, indicating superior performance.

In all of the above research, the speech-in-noise test was adaptive and involved either single digit or word(s) identification. As was introduced at the start of this chapter, the speech test used here was designed to be consistently challenging at an SNR of -8 dB. It was hypothesised that there may be different compensatory cognitive strategies invoked to assist in discriminating in this challenging condition. This type of listening situation, of listening to a speaker in a noisy room with no chance of tuning in to voices, may occur regularly in normal life. Discovering, therefore, which functions may explain some of the variation in speech-in-noise discrimination performance would seem advantageous. The processing speed tasks employed to test for any relationship were a choice reaction time test, the go/no-go, and the congruent condition of the Stroop task.

2.1.2. Control of inhibition

Control of inhibition is a function under executive control. The two main ways that it can be exhibited were described in Section 1.1.3.2: the ability to inhibit prepotent responses (in this study, measured by a Stroop task and a random number generation task) and the ability to focus on one task while ignoring unwanted stimuli. In the research for this chapter, the latter was measured by performance of an auditory selective attention task that involved both inhibition control and task switching, both of which tap executive functions. Before turning to how previous research surrounding inhibition control helped develop the study in this chapter, it is first worth considering in more detail how inhibition changes with age and some of the contention around inhibition theory itself.

A decline in inhibition control has been noted in older healthy adults (Hasher & Zacks, 1988). The Stroop effect (a measure of inhibition) has been demonstrated to show changes across the lifespan (Uttl & Graf, 1997), with older individuals having poorer control of inhibition. There is some contention as to the exact nature of inhibition control. Uttl and Graf state that although the control of inhibition demonstrates a significant direct link with age, this could be due to the influence of processing speed (i.e., colour identification and word reading), rather than a decline in the actual qualitative process of inhibition. Indeed, in a meta-analysis of the Stroop effect, when younger and older adults' performances were compared (Verhaeghen & De Meersman, 1998), it was reported that the Stroop effect appears to be an artefactual effect resulting from general slowing. As both processing speed and control of inhibition are being considered in this chapter, it may be possible to distinguish their unique contributions and respective relationships with speech-in-noise performance. Also, by factoring out age and hearing level in the research for this chapter, any additional explanation of variance in performance in the speech task associated with these two functions can be determined.

Previous research using the neighbourhood activation model (NAM; (Luce & Pisoni, 1998; Pisoni et al., 1988) was introduced in Section 1.2.3. The NAM has proven to be a useful tool in the exploration of the relationship between control of inhibition and speech-in-noise discrimination. The NAM proposes that to perceive a word, the listener must discriminate it from lexically similar neighbours, and any of its competitor words must be inhibited. It, therefore, follows that a word with many phonetically similar-sounding competitors is harder to distinguish than one with fewer acoustically similar words. That is, the more unique a word in terms of sound, the easier it is to discriminate. In addition, more frequently used words are easier to discriminate than rarely used words. There needs to be an interaction between these two concepts as the bottom-up sensory input and top-down information are applied in decision-making units to select the appropriate word and ignore potential competitors.

The research findings of Taler et al. (2010) using NAM have already been considered in Section 1.2.3. In summary, both older and younger adult groups performed less well at discriminating a word when its neighbourhood density was high. This indicates that the greater the number of other likely contender words, the more the contenders need to be inhibited. It could be interpreted, therefore, that an individual with better control of inhibition may perform this task in a superior fashion. The effects of neighbourhood density were stronger for the older adult group compared to the younger group, indicating a decline in older adults' control of inhibition. This difference was significant when listening conditions were challenging (at -3 dB SNR), but not at an easier SNR of +10 dB. Thus, individuals who exhibit better control of inhibition are seemingly able to ignore/inhibit contender words more efficiently, allowing better speech reception performance in noise, when the task is difficult.

Research by Sommers (1996) and Sommers and Danielson (1999) found results similar to those of Taler et al. regarding the effect of declines in the control of inhibition on the recognition of high neighbourhood density words when presented in noise. The former study found performance differences between a younger and older group of participants. This took the shape of a significant interaction between age and lexical difficulty of words presented in noise, thus supporting an age-related decline in inhibitory control. This was the case whether words were presented in isolation or in sentences. However, Sommers argues that this may be due not to inhibition but to other factors, such as a decline in available cognitive resources or in processing speed. In the latter study, also utilising older and younger groups of adults, older adults had greater difficulty in identifying words with many likely contenders, similar to the earlier study. Two additional aspects were reported. Firstly, the older adult group showed greater benefit from the addition of contextual clues than the younger group i.e. older adults appear to be better able to use environmental support and get extra assistance from 'top-down' processing as their sensory and cognitive skills decline (Craik & Byrd, 1982). Secondly, using an auditory Stroop task revealed a significant relationship between the identification of low-frequency, lexically challenging words (when presented in noise) and an auditory Stroop interference measure[22]. This was not the case for the lexically easy words. Although Sommers

---

[22] The audio Stroop task involved a woman or a man saying mother, father or person. Participants responded as quickly and accurately as possible. The auditory Stroop interference measure was the

and Danielson used different masking noises[23] than Taler et al., they both found that it was only under challenging situations that differences between age groups and relationships with inhibition control were observed.

Although the NAM research procedure was not used for the research in this chapter, as the BKB sentences employs words that are simple and in frequent use, the words used here do have a number of similar competitors. In addition, the research reported here made the task difficult by employing a speech-in-noise test in a very challenging SNR. It could be argued that rather than manipulating the speech task by using more ‘unusual’ words, as in the NAM, using a speech task that employs simpler language better reflects real-life listening situations.

Still, the Stroop interference measure and speech perception in noise do not always demonstrate a significant relationship (Knight & Heinrich, 2017). Knight and Heinrich suggest several confounding factors, including the speech test material and the way the Stroop score is calculated. As an example, Helfer and Freyman (2014) examined speech-in-noise discrimination performance and inhibition control as measured by a Stroop task but found no significant relationships for middle and older age participants (45–85 years) with any of the speech tasks. The speech tasks employed used an average score of speech discrimination performance in speech-shaped noise at three set SNRs, -3, 0 and +3 dB, hence not as difficult as the present research level of -8 dB.

However, a correlation was found for a processing speed task when the background noise was either a 1- or 2-talker competing speech masker but not for speech-shaped noise. If as some researchers argue, that inhibition control and speed of information processing are tapping the same function of speed then where there is a relationship with one of the two, it may have been expected that both would demonstrate a relationship. The different relationships between processing speed and the inhibition measure potentially offer some evidence that inhibition

---

time difference between the congruent and incongruent responses, e.g., a woman saying mother and a woman saying father.

[23] Speech-shaped noise and babble, respectively, both of which have been explained in Section 1.4.

and processing speed that whilst similar are indeed different functions rather than as argued by Uttl and Graf (1997).

2.1.3. Working memory

As introduced in Section 1.2.4, the role of working memory has demonstrated the most reliable relationship with speech-perception-in-noise performance (Akeroyd, 2008). Although working memory is not the focus of this chapter, it is worth reviewing briefly here the evidence that guided the choice of assessments used for the research in this chapter.

As a summary, previous researchers have demonstrated a link between working memory and speech perception in noise (Besser et al., 2013a; Foo et al., 2007; Lunner, 2003; Rudner et al., 2011). The hypothesis was that when listening to speech in background noise, the immediate speech information needs to be decoded and matched to stored information (Rönnberg et al., 2013; Rönnberg et al., 2008) before it is lost downstream. For example, Souza and Arehart (2015), testing younger and older participants,[24] reported that the listeners with poorer working memory (measured by two reading span working memory tests) had greater difficulty understanding speech-in-noise even after controlling for age and any hearing loss. They reported on the robust nature of this relationship. However, this relationship is not universally reported (Besser et al., 2013a; Zekveld et al., 2013). Indeed, Füllgrabe and Rosen (2016b) reported that working memory is not a strong predictor of speech-in-noise discrimination performance for younger/middle aged adults with normal hearing.

In addition, Baldwin and Ash (2011) reported that when speech is challenging to listen to, there is a decrease in auditory working memory performance as measured by a listening span task (LSPAN). They reported that it was an individual’s speech recognition ability that significantly accounted for LSPAN performance. This was the case for normally hearing younger adults (18–31), but particularly for normally hearing older adults (60–80 years).

---

[24] The older participants had hearing loss.

The above summary assisted with developing the methods for the research for this chapter. The age range recruited was wide (18–80 years), with all participants self-presenting as without hearing loss.[25] It was, therefore, decided to employ a listening span task to measure working memory, as the evidence above and in Section 1.2.4.1. suggests that this may establish subtle relationships with speech-in-noise discrimination performance where the RSPAN may not. Building a hierarchical regression model of factors explaining variance in performance of a speech-in-noise task was a key aim of this chapter, as by partialling out age, hearing level and general cognitive performance (working memory), any additional explanation of speech-in-noise performance variations accounted for by information processing speed, and/or inhibition control could be determined. The next section sets out the aims and hypotheses that were to be tested in this chapter.

2.1.4. Aims and hypotheses

A wide age range of adults with no self-reported hearing loss were recruited for the research for this chapter. The purpose was to explore which cognitive functions may offer support when attempting to discriminate speech in very challenging listening conditions, with the speech presented in speech-shaped noise at -8 dB SNR.

Speed of information processing tasks and both pre-potent and ‘real-time’ control of inhibition tasks were employed to ascertain which of these functions may assist with an individual’s speech-in-noise discrimination performance. To minimise the effect of semantic knowledge on speech-in-noise performance, short sentences were used for which there was limited contextual support. The sentences used, such as ‘The glass bowl broke’, were taken from the BKB sentences for partially hearing children (Bench et al., 1979). The challenging SNR afforded fewer glimpses of any heard speech than would a more favourable SNR, and utilising speech-shaped noise as the masking noise instead of babble allowed for limited access to the

---

[25] Pure Tone Audiometry (PTA) was performed at the initial assessment to rule out any significant problems.

target speech. It was hypothesised that those individuals who were better able to respond and process information quickly and inhibit both prepotent responses and unwanted extraneous signals would also better perceive speech-in-noise.

The hypotheses to be tested are as follows:

H2.1. If the control of inhibition and speed of information processing play a role in the understanding of speech-in-noise, then the tasks that measure these cognitive functions will correlate significantly with speech-in-noise discrimination performance.

H2.2. If the cognitive functions noted in H2.1 provide an explanation for variation in speech-in-noise discrimination performance, then they will also add substantially to any explanation of performance variation once age, PTA, and working memory have been factored out.

2.2. Methods

2.2.1 Ethics

The Aston University Research Ethics Committee (UREC) approved the proposed research. (Reference number 701).

2.2.2. Participants

To calculate the sample size needed for a statistically significant correlation, previous research in the area of inhibition control and speech-in-noise discrimination performance was explored. Research by Sommers and Danielson (1999) reported a medium-sized effect for the relationship between a Stroop interference measure and speech-in-noise performance. Based on this, G*Power software (Faul, Erdfelder, Lang, & Buchner, 2007) was used to calculate the number of participants necessary using the previous effect size (r = 0.3), an alpha value of 0.05 and a probability of 80%. As previous research had indicated the direction of the expected relationship, a one-tailed test was used for the calculation. The resulting number of participants needed was 67.

Ultimately, 71 adults were recruited (28 men and 43 women, 18–76 years old, mean 45.09 years, SD 21.45) to allow for any attrition. The participants were recruited through an invitation from the Research Centre for Healthy Ageing (ARCHA) or from the psychology undergraduate research participation scheme or were colleagues/friends at university. Participants in the ARCHA group were given £7.50 to cover travel expenses, and students were given course credit. All participants spoke English as their first language, since performing speech-in-noise discrimination tasks when a participant’s first language is other than English leads to poorer scores (Roberts et al., 2006). No participant reported any problems with colour blindness. There were no hearing aid users, and no participants had sought help with regards to their hearing.

2.2.3. General design

A correlational design was initially used, and none of the variables were manipulated. Participants’ performance on the different cognitive tasks was correlated with their BKB speech-in-noise discrimination performance score. Based on previous research and in order to test the hypothesis that speed of information processing and control of inhibition would explain some of the variation in speech-in-noise performance a hierarchical regression model was constructed. As a further check of which functions to include in the model the strength of the correlations between the different cognitive functions and speech-in-noise performance was also examined.

2.2.4. Test material

2.2.4.1. Audiometry

A calibrated pure tone audiometer as part of the aurical+ testing and assessment system (Otometrics) was used to perform the hearing assessment.

2.2.4.2. Discrimination task

Speech-shaped noise was used as the masking noise for the speech-in-noise task affording limited glimpses of speech as there are no dips or troughs in the masking noise. The entire speech-in-noise test was delivered at a demanding non-adaptive SNR of -8dB SNR, meaning the speech was 8dB quieter than the masking noise and being non-adaptive maintained a consistent level of difficulty.

Lists 2, 3 and 4 of the original BKB material (Bench et al., 1979) were employed, so 48 sentences in total (3 lists of 16 sentences) were used. The participants repeated the words that they heard immediately after each presented sentence and were encouraged to guess if they could not discriminate the words clearly. The tester recorded the number of key words correctly repeated by the participant. As an example, 'The **clown** has a **funny face**' is a standard BKB sentence, with the letters in bold being the scoring words. The individual performing the task would listen to the sentence presented and then immediately repeat aloud what they heard to the researcher. If they did not hear the entire sentence, they were encouraged to say any single words they heard. A loose scoring method was employed where a correct mark is given even when only the route of the word is correct (i.e., 'run' would score a mark even if the target word was actually 'running').

2.2.4.3. Listening sentence span task (LSPAN)

A complex listening span task (LSPAN) was also performed with all participants, based on the work of Daneman and Carpenter (1980). This involved listening to a sentence, making a judgment on whether or not the sentence made 'sense', and then being presented with a letter auditorily, which the participant attempted to hold in their working memory. They were then presented with another 2-6 sentences depending how many sentences were in the block, all with a letter following to hold in their memory for the length of that testing block.

For example, a sentence could be **'The children walked across the sky'** which the participant would indicate was false. They would then hear a letter spoken which they had to remember, for example **‘A’**. The next sentence could be **‘Peter got out his camera to take some pictures’** which the participant would indicate was true. They would then hear another letter, for example **‘G’.** The participant now has to hold **A** and **G** in that order in their memory. This would continue for however many sentences and letters were presented in the block. On the collection screen they would choose the letters they had been presented with in the correct order.

The number of sentences for each block was randomised between 3 and 7, so the length of letter sequences was by default the same, 3 to 7. After the complete sequence of sentences and letters in a block, participants were presented with a selection of 16 letters on a computer screen. This selection was the same whether they had been presented 3 sentences and letters or 7 sentences and letters i.e. there were always 16 options. The participant then had to click on the letters they had seen in the order that they had seen them. In total there were 15 collection blocks of between 3 and 7 sentences/letters in length.

Participants were given a training period to familiarise themselves with the LSPAN task. This involves remembering letter sequences, making judgements on sentence veracity and then combining the two tasks. The screen from which to choose the letters they had seen was presented at the end of each test block. If a participant did not make sufficiently accurate judgments of the sense of the sentences (thereby avoiding the complexity of the task) and their correct identification of sentences fell below 85%, their results were excluded. The absolute (LSPAN) score is recorded for each participant, which is the sum of all the sequences of letters in which the participant got all the letters correct. If they make one error on a letter in a sequence, that sequence does not count towards their score. The number of letters from all the correct complete sequences are added together to produce an overall score. The participant only gets a score when all the letters in a sequence are correct.

2.2.4.4. Stroop Test

Control of the ability to inhibit a prepotent response (in this case the reading words) while performing a different task was assessed by a Stroop test (Stroop, 1935). This test has been described in Section 1.1.2.3. and has been established over time as a classical measure of cognitive control. Words were presented to participants on a white background computer screen in either a congruent condition, where the meaning of the word and the hue of the word were the same (i.e., “red” written in red font), or in an incongruent condition, where the word “red” could be written in green font. In the incongruent condition, the over-learned response of word reading must be overcome (inhibited), requiring attention in order to report the colour (Cohen, Dunbar, & McClelland, 1990). The congruent condition was chosen instead of a neutral condition for the Stroop interference measure as it has been argued that by using the same words and colours, it is only congruency that is being adjusted (Sabri, Melara, & Algom, 2001). Participants when seeing a colour written on a computer screen named the colour of the writing (not what he word said) as quickly but as accurately as possible. The time they took was recorded by the inbuilt microphone in the computer. As the Stroop test used microphone voice recordings, all responses were checked for accuracy and all incorrect colour responses were ignored. In addition, errors in the microphone recording itself were also ignored. The test was a screening Stroop task consisting of only 20 presentations. The average congruent recorded response time was subtracted from the average incongruent response time, this being the most widely used way of calculating Stroop scores (Macleod, 1991). This is a measure of how well an individual can inhibit their prepotent response to read the word rather name the hue and is reported as Stroop interference (or Stroop effect). The smaller the difference between the incongruent and congruent response, the better the control of inhibition.

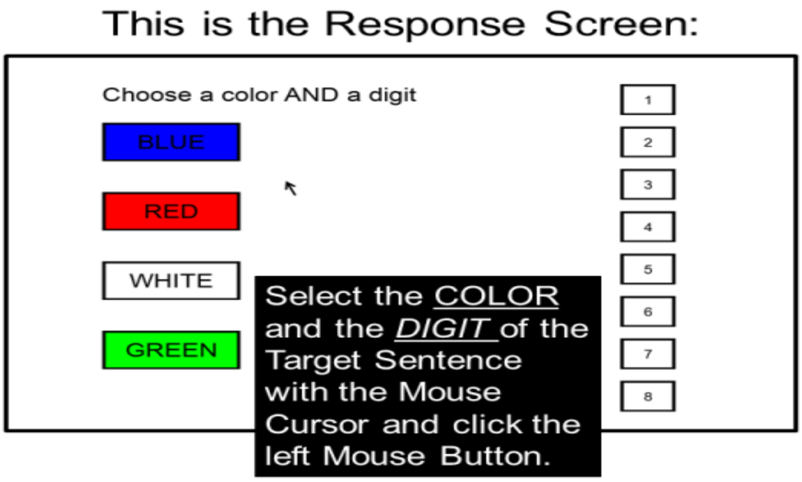
All recorded utterances of the participants were analysed for incorrect or unwanted responses (e.g., saying ‘umm’). Further, any response times below 300msecs were removed for being due to anticipation, and responses greater than 1500msecs were also removed due to task inattention (Macleod, 2005).

2.2.4.5. Selective attention task

A selective attention task (Humes, Lee, & Coughlin, 2006) was used as an auditory measure of inhibition and task switching. In this test, sentences are presented in dichotic fashion by different talkers (a male and female speaker talking simultaneously) through headphones. The sentences were delivered to participants at a calibrated 65dBSPL via the testing apparatus as described in Section 2.2.5. The sentences took the form of 'Ready (call signal) go to (colour) (number) now'. For example, one might be, 'Ready Charlie go to Blue 6 now' The participants were given the call signal in written form on the computer screen ('Charlie' in the example above) before each auditorily presented sentence to instruct them to attend to one of the two talkers. One talker would say 'Charlie' (attend to this talker), while the other talker would say 'Arrow' (so this talker should be ignored). Figure 1 shows how the test was explained to the participants. During the test, they would only see the word Charlie on the screen.

Figure 1. Sample of the call sign for selective attention task.

The call sign itself was not part of the results screen, meaning there was no visual aid during the result collection screen. After each pair of simultaneous/dichotic sentences, a collection screen then appeared with a choice of 4 colours and 8 numbers (but no call sign), and the participant clicked by means of a computer mouse to identify which ones they had heard from the target talker. The collection screen is shown below in Figure 2. Here the correct response would be to click on BLUE and 8 as they are the targets given above.



This is the Response Screen:

Choose a color AND a digit

Select the COLOR and the DIGIT of the Target Sentence with the Mouse Cursor and click the left Mouse Button.

Figure 2. Sample of the response screen for selective attention task.

Participants were given 32 practice pairs of sentences along with feedback on correct or incorrect responses, followed by 32 sentence pairs in the trial itself with no feedback. The score for each participant was the number of colour words and numbers they could correctly identify out of the 32 spoken presentations by the target talker. A maximum score per sentence would therefore be 2 (the colour and digit correctly identified), and the minimum score would be 0 (no colours or digits correctly identified). By averaging the number of correctly identified colours and digits across all 32 sentences a performance score was calculated. The closer the score to 2 the better the performance of the task, the closer to 0 the worse.

2.2.4.6. Random number generation (RNG) Task

A random number generation task (RNG; (Ginsburg & Karpiuk, 1994) was created using Microsoft Office Power Point. Blank slides with an asterisk (*) symbol presented in 44 font appeared every 800msecs. The participant watched a computer screen and said a number out loud between 1 and 9 inclusive each time an asterisk appeared. They were asked to be as random as possible when saying their numbers and to avoid patterns. They were given 13 practice attempts to become accustomed to the timing and then 55 presentations during the main trial. Their responses were written down and entered into RGCalc, which gave the three output measures given below (Towse & Neil, 1998).

1. The RNG index examines sequence frequency, with a lower score demonstrating a better ability to inhibit patterns.

2. The R-score represents repetitions (how often numbers are repeated), with low scores indicating a better updating ability.

3. The A-score, a combined adjacency score, focuses on the numerical sequences, with low A-scores showing a better ability to inhibit patterns that have been learnt.

2.2.4.7. Go/no-go task and response latency measure

The go/no-go task has been used by researchers to examine responses to a ‘go task’ and the ability to inhibit responses to an alternate presentation (no-go). In this study, the task was used to examine if there were changes related to age or more fundamental differences in inhibition.

The test consisted of 250 trials that involved a rectangle appearing on a computer screen either horizontally or vertically. Before the rectangle appeared, a plus sign (+) appeared on the screen for 800ms, followed by a blank white screen for 500ms. A clear rectangle then appeared and changed colour either very quickly or more slowly (100, 200, 300, 400 or 500ms). If the rectangle turned green, the participant had to press the space bar (go response). However, if the rectangle turned blue, the participant was told to take no action (no-go).

The orientation of the rectangle acted as a cue for this task. While in the vertical condition, there was a 4:1 greater likelihood of requiring a ‘go’ response. In the horizontal presentation, the reverse was true, as it was more likely to turn blue (no-go response). The orientation of the rectangle therefore does gives pre-potency of the likelihood of a ‘go’ response. The participant sat in front of a computer screen and had to press the space bar every time the shape they were presented with turned green. They were told to do this as accurately and quickly as they could. If the shape turned blue they were not to press the space bar. For the response latency task, it was the speed of pressing the space bar correctly that was recorded. For the inhibition portion of the go/no-go it was the number of incorrect presses that were recorded.

The performance during the go/no-go task non-inhibition trials was used as a measure of reaction time for the ‘go’ condition (125 presentations). Any incorrect responses were ignored in the reaction time calculation.

2.2.5. Procedure

The pure tone audiometry and all the cognitive tests were performed in a well-lit (320-360 lux) sound proof room at Aston University. Detailed instructions were given to the participants before the start of any of the tasks with clarification about the procedures given where necessary. The order of the tasks performed was as follows: otoscopy, pure tone audiometry, BKB speech-in-noise test, Stroop test, auditory working memory test (LSPAN), auditory selective attention task, go/no-go task, and random number generation task. Breaks were offered to the participants after each of the tests.

Ear examination was performed (BSA, 2016) and if any adverse findings were discovered during this examination, participants were advised as to where to seek further medical assistance. The hearing test was performed to BSA protocol (BSA, 2011) to confirm participants’ hearing thresholds. Although no participants with known hearing loss were purposefully recruited, it had been anticipated that, given the age of the participants, some may have some mild degree of hearing difficulty. The participant’s thresholds for both ears were recorded from 500-4000Hz in octave bands. There was a debrief of the hearing test results with the participants, with the plan that if any degree of hearing loss was found, the participant would be given guidance about how to obtain further information and help. There were no exclusions necessary due to the medical condition of the ears or participants’ hearing. The average hearing levels at each frequency are shown after the summary of descriptive statistics.

The BKB speech-in-noise test was delivered to the participants through a Samsung R350 personal computer and an Edirol UA-25 audio capture box. The output was played through

Sennheiser HD250 headphones and was calibrated to deliver the speech at 65dBSPL (normal conversational voice level) as measured by a Brüel and Kjær 2250 sound level meter.

All of the cognitive assessments were run through a Toshiba Satellite Pro personal computer. The Stroop, selective attention and go/no-go tasks were all performed using Millisecond software (Inquisit). The RNG task was a stand-alone program run on the same personal computer.

2.3. Results

During results collection, one of the participants struggled significantly with both the listening span and random number generation task. As this would have an unacceptable influence on the results, the results of this particular participant were considered an outlier and excluded from the study. This left 70 participants in the study.

As a comment on two of the cognitive function tasks, as only 6 participants had made incorrect responses on the go/no-go inhibition task, indicating problems with ceiling effects, this particular task was excluded from the final analysis. The response latency performance for the go condition was, however, retained. Additionally, the results of the Stroop task presented with a large standard deviation, due in part to it being a screening task and as such, individual's with very strong inhibition could potentially record a negative score (i.e. their incongruent performance was better than their congruent performance).

As a further general comment, the research had set out to make the speech-in-noise task particularly challenging in order to explore which cognitive functions may explain performance under these trying conditions. This particular aim was achieved in that the overall mean performance on the BKB hearing-in-noise speech test was 31.62% (SD 12.57), confirming that the test was indeed difficult. The full set of results are given below in Table 1.

Table 1
Summary of descriptive statistics

Measure	**Mean**	**Standard deviation**
Average PTA	10.10	8.83
BKB Hearing-in-Noise Test %Score at constant -8 dB SNR	31.62	12.57
Listening Span	23.88	15.13
Go/no-go latency (msecs)	409.70	59.67
Stroop interference measure (msecs)	95.06	91.69
Selective Auditory score Attention task	1.67	0.30
Random Number (RNG)	0.25	0.05
Random Number (R score)	1.71	1.11
Random Number Adjacency (A)	21.78	8.91

Figure 3 below presents additional information on the hearing levels of the participants. These are displayed as average thresholds (with standard deviations) for each frequency. No participant had a threshold at any frequency greater than 40 dB HL, and no participant had an average loss (when averaged across 500, 1000, 2000 and 4000 Hz in both ears) greater than 30 dB HL.

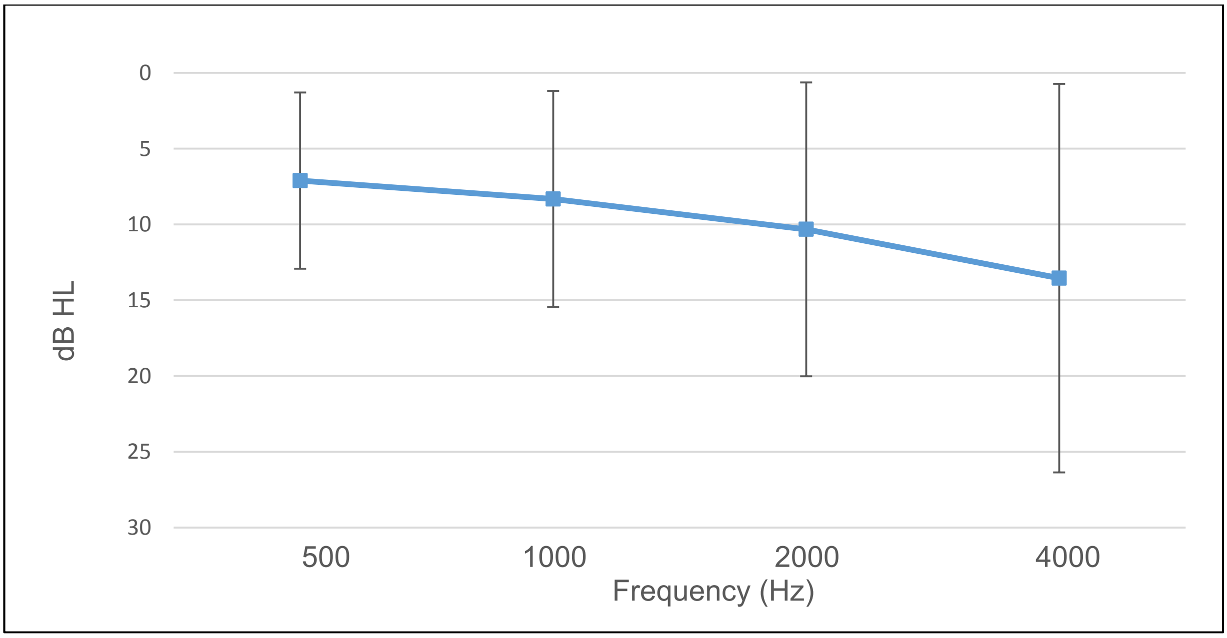


Figure 3. Average pure tone thresholds at each frequency with standard deviations.

Prior to running the correlation analysis, all the data were explored for normality. Three of the cognitive tests, the auditory selective attention task and 2 of the random number generation measures (the RNG and the R scores) were found to deviate from normality based on both visual observation of the Q-Q plots and the Shapiro-Wilk test, $p < 0.05$. Attempts to transform the data to a normal distribution were not successful, so a Spearman's rho test was used to run the correlational tests.

Hypothesis 2.1, predicated that speech-in-noise discrimination performance would be related to the cognitive functions of control of inhibition and speed of information processing. This was tested using the correlation matrix analysis where other significant relationships were also explored between speech-in-noise performance and other variables. Due to large number of comparisons being made (10 in total), a Holm-Bonferroni correction was applied (Holm, 1979). This utilises a changing critical *p* value from alpha/10 to alpha/1 following a stepwise procedure, where alpha was set at 0.05. The criterion for significance, therefore, slightly relaxes. For example, the first level of significance is 0.05/10 = 0.0050, the second is 0.05/9 = 0.0055, the third is 0.05/8 = 0.00625, etc.

The Holm-Bonferroni procedure has the effect of reducing family-wise error rates (type 1 or false positive errors). It does, however, potentially increase the risk of type 2 errors, so a true relationship may be dismissed when one is present. The analysis aimed to determine the strength and direction of any relationships; the full set of results are given below in Table 2.

Table 2

Correlation matrix (N = 70) with Spearman's rank correlation coefficients.

	Age	Average PTA	BKB Score	LSPAN score	Selective attention	Go-No go	RNG RNG score	RNG adjacency	RNG R score	Stroop Interference
Age	—									
Average PTA	.736*	—								
BKB Score	-.104	-.173	—							
LSPAN score	-.480*	-.476*	.326	—						
Selective auditory attention	-.614*	-.446*	.267	.490*	—					
Go/no-go latency	.369*	.393*	-.410*	-.460*	-.455*	—				
RNG RNG score	.123	.028	-.168	-.066	-.217	.107	—			
RNG adjacency	.355*	.218	-.165	-.269	-.284	.256	-.010	—		
RNG R score	-.214	-.195	.012	-.078	-.067	.013	.270	.070	__	
Stroop Interference	-.168	-.258	-.241	-.072	.025	-.029	.075	030	.022	---

Note: Key: Average PTA = average pure audiometry threshold calculated from thresholds at 500, 1000, 2000 and 4000 Hz. BKB = Bamford-Kowal- Bench hearing-in-noise test score as the percentage of correctly identified key words. LSPAN = Listening Span. RNG = Random Number Generation tasks Correlations: *Correlation is significant at the 0.05 level (2-tailed), Holm-Bonferroni correction applied.

2.3.1. Correlational analysis

As seen in Table 2, the only cognitive function that demonstrated a significant correlation with speech-in-noise performance was information processing speed (response latency). This was measured by the ‘go’ section of the go/no-go task, a simple visual reaction time task. The direction was negative; thus, the quicker the reaction time, the better the speech-in-noise discrimination score, $p < 0.001$.

There were no other significant relationships demonstrated for the other cognitive functions after the Holm-Bonferroni correction was applied. However, the selective attention task, the memory span (LSPAN) and the Stroop measure of interference were all significant at the non-corrected $p < 0.05$ level. The three random number generation tasks, two of which (the RNG and Adjacency (A)) measure inhibition, and the R updating measure did not demonstrate a significant relationship either with or without correction.

The relationships of the various cognitive functions with BKB speech-in-noise performance is demonstrated in Figures 4-7. Further, there are also further correlations in Figures 8 and 9, demonstrating the relationship of age with average hearing threshold and age with response latency. The two graphs, Figure 8 and 9 illustrate how cognitive functions show a decline with age.

As Hypothesis 2.1 stated that there would be significant relationships between speech-in-noise discrimination performance and the range of cognitive functions assessed, the hypothesis can only be partially supported by the results, as only speed of information processing was significantly correlated.

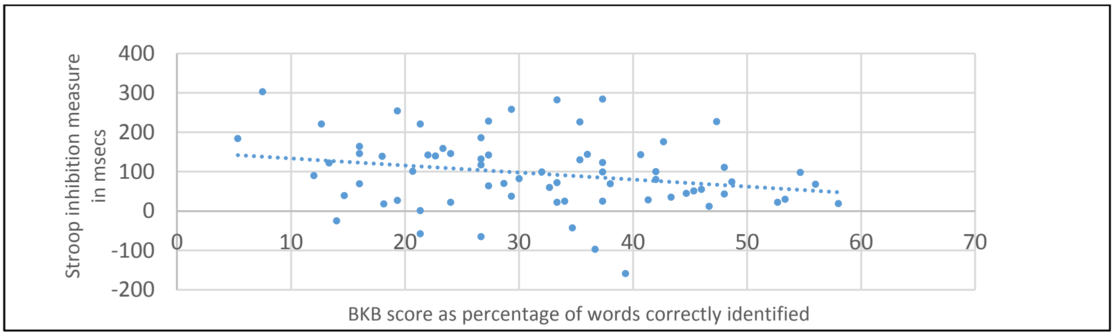


Figure 4. The correlation of BKB performance with the Stroop interference measure. Spearman's rho = -0.241.

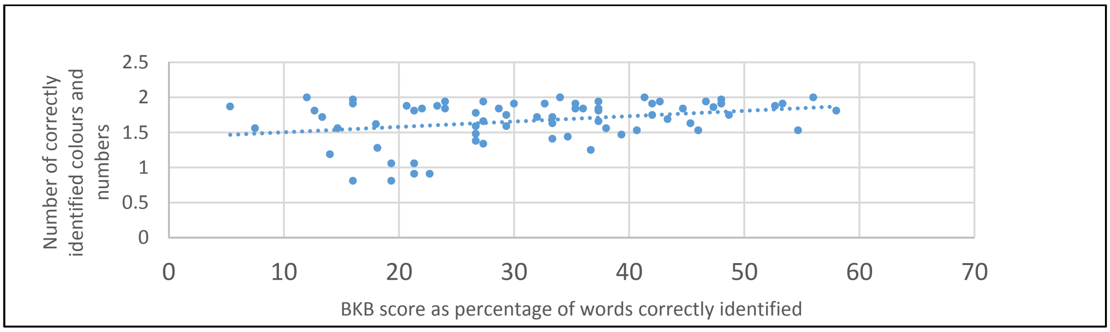


Figure 5. The correlation of BKB performance and selective auditory attention task performance. Spearman's rho = -0.267.

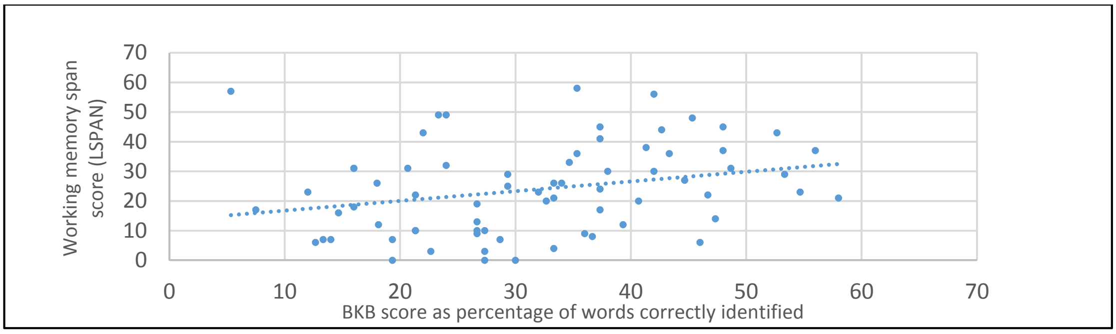


Figure 6. The correlation of BKB performance and working memory (LSPAN) performance. Spearman's rho = 0.326.

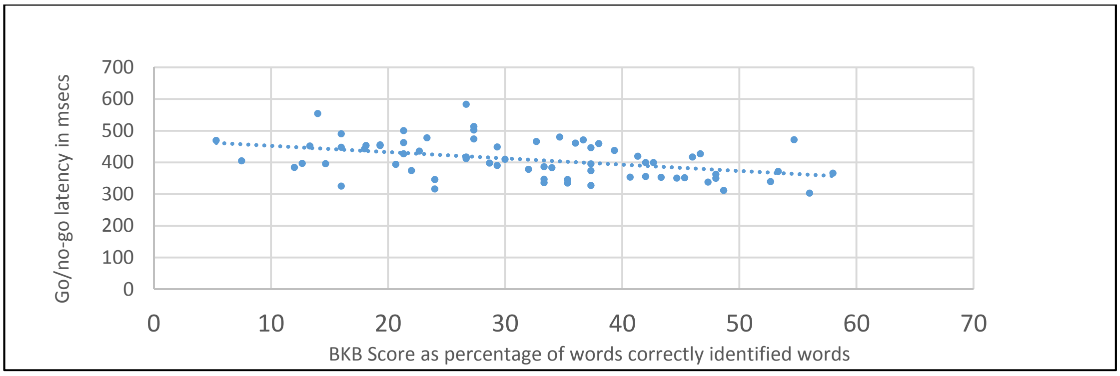


Figure 7. The correlation of BKB performance with response latency performance (go/no-go task). Spearman's rho = -0.410.

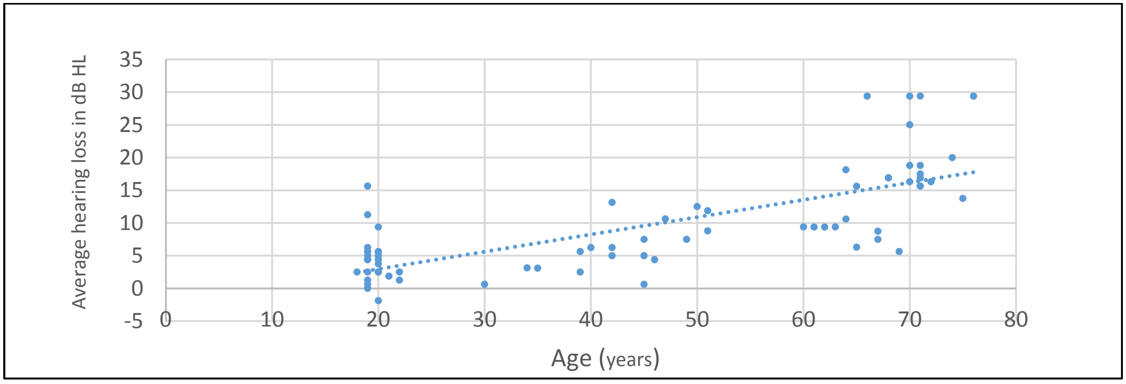


Figure 8. The correlation of age with measured PTA average. Spearman's rho = 0.736.

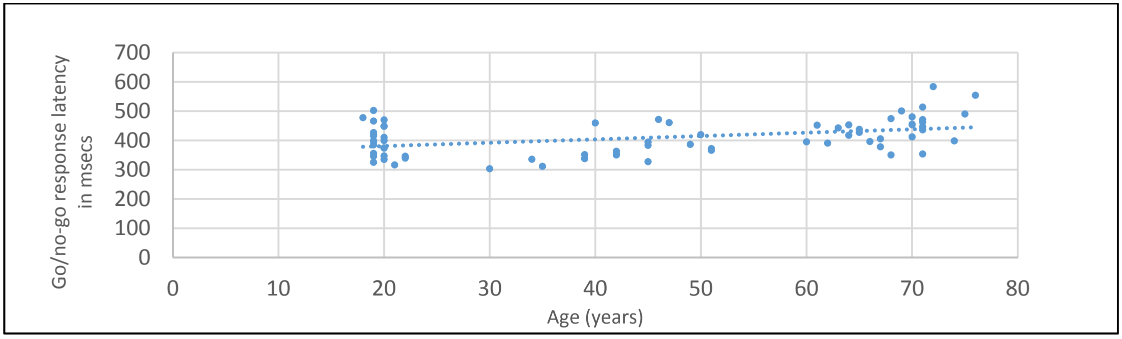


Figure 9. The correlation of age with response latency performance (go/no-go task). Spearman's rho = 0.369.

2.3.2. Hierarchical regression modelling

To evaluate Hypothesis 2.2, hierarchical regression modelling was employed, as it has been found to be the most appropriate method when testing theory (Field, 2016, p. 512). The hypothesis for this chapter was to determine if the cognitive functions of speed of information processing and control of inhibition may offer an explanation of BKB variance without the undue influence of other variables. The regression model was, therefore, designed to partial out other variables that may also be associated with speech-in-noise discrimination performance.

The first model of the hierarchical regression analysis accounted for relevant demographic information by including age and hearing level (PTA). The PTA average threshold was included, as it has been reported in previous research that hearing thresholds are the main predictors of speech-in-noise discrimination performance (Akeroyd, 2008; Humes & Dunbo, 2010; Van Rooij & Plomp, 1992b). Age has also been shown to have an effect on the performance of many cognitive tasks, with performance declining as individuals' age (Darowski et al., 2008; Salthouse, 1996). These accounted for model 1.

Model 2 allowed for any extra variance due to working memory (measured by the LSPAN), which has been shown to be linked to speech-in-noise performance (Pichora-Fuller et al., 1995). Model 3 added speed of information processing, which has been demonstrated as being associated with speech-in-noise performance (Jerger et al., 1991). The measure used for the research in this chapter was the response latency of the go response of the go/no-go test. Model 4 added tasks involving the control of inhibition, as this cognitive function has also demonstrated a previous relationship with speech-in-noise performance (Taler et al., 2010). Additionally, the tasks chosen to represent the different cognitive functions in the regression model had all demonstrated a significant correlation with the BKB speech-in-noise task before or after the Bonferroni correction, $p < 0.05$.

Before the hierarchical analysis was performed, assumptions of linearity, the independence of errors (Durbin-Watson statistic of 1.793) and homoscedasticity were checked and met.

The four successive models were, therefore, used to investigate the amount that each model in turn added to explaining the unique variation of the dependent variable of speech discrimination in noise. The full set of models used for the hierarchical regression is shown below in Table 3.

Table 3

Summary of models used in hierarchical regression and proportion of variance in BKB test scores

Model	**Predictors**	R^2	R^2 **change**	**Significance *F* change (p value)**
Model 1	Variation due to PTA and Age	0.043	0.043	0.231
Model 2	Variation due to working memory (LSPAN)	0.089	0.046	0.071
Model 3	Variation due to information processing speed (response latency of go condition of no/go task)	0.203	0.113	0.003
Model 4	Variation due to Inhibition measures (Stroop and selective attention task)	0.287	0.084	0.030

The overall hierarchical regression model comprising age, PTA, working memory span performance (LSPAN), speed of information processing (reaction time- go condition of go/no-go) and inhibition measures (Stroop interference and auditory selective attention) was statistically significant, $R^2 = 0.287$, $F(2, 63) = 3.710$, $p < 0.05$. This accounted for 29% of performance variance.

The effect of adding each model individually gave the following results. Adding working memory, the LSPAN task (model 2), to model 1 (PTA scores and age) did not result in a statistically significant increase in the prediction (F change) of speech-in-noise discrimination performance, $R^2 = 0.089$, $F(1, 66) = 3.357$, $p = 0.071$. When adding speed of information processing (model 3), this added significantly to the explanation of performance variance ($R^2 = 0.203$, $F(1,65) = 9.246$, $p < 0.01$). Adding the control of inhibition measures, the Stroop interference measure and auditory selective attention performance, to models 1, 2 and 3 also yielded a statistically significant increase in the prediction of the speech-in-noise discrimination scores ($R^2 = 0.287$, $F(2, 63) = 4.219$, $p < 0.05$).

Individually, the only variables that added significantly to predicting the variance in performance on the speech-in-noise task once other variables had been accounted for were reaction time (go condition of go/no-go) and inhibition measure (Stroop interference), both at $p < 0.05$.

In the hierarchical models 2, 3 and 4 (in Tables 3 and 4), some of the cognitive functions do individually demonstrate a significant contribution to the explanation of the variance in speech-in-noise performance, even after age and hearing thresholds have been partialled out. In view of this, Hypothesis 2.2 is supported. This is not true, however for all of the cognitive functions that were assessed.

The full set of results is given below in Table 4.

Table 4

Hierarchical regression predicting BKB test scores from PTA and age (model 1), when working memory (model 2), speed of processing (model 3) and control of inhibition (model 4) are added.

	b	***SE b***	***ß***	***R^2***	***R^2* change**	**Significance of *F* change**
Model 1				.043	.043	.231
PTA	-.316	-.222	-.222			
Age	.015	.091	.025			
Model 2				.089	.046	.071
PTA Scores	-.229	.212.	-.161			
Age	.070	.095	.119			
Listening span score	.122	.114	.147			
Model 3				.203	.113	.003
PTA	-.159	.212	-.112			
Age	.111	.090	.190			
Listening span score	.122	.114	.147			
Go/no-go latency	-.082	.027	**-.389****			
Model 4				.287	.084	.030
PTA scores	-.245	.213	-.172			
Age	.129	.092	.220			
Listening Span Score	.056	.112	.068			
Go/no-go latency	-.072	.027	**-.342***			
Selective Attention Task	6.664	6.079	.161			
Stroop interference	-.037	.016	**.267***			

Note: ** $p < 0.05$, ** $p < 0.005$*

2.4. Discussion

To establish the roles of different cognitive functions in the discrimination of speech from challenging background noise, a wide age range of participants was recruited. They performed a variety of cognitive tasks, and the significant relationships have been reported above. Of the presented hypotheses, H2.1 was only partially supported in that, after a correction for multiple correlations, only information processing speed correlated significantly with the BKB speech-in-noise performance. Hypothesis 2.2 was also supported in that after factoring out age and hearing level (and other cognitive functions) in hierarchical regression modelling, information processing speed and prepotent inhibition control added significantly to the variance in performance on the BKB task.

Information processing speed, as measured by the response latency task of the go condition of the go/no-go task also significantly correlated with the variables of age, PTA, and the cognitive functions measured by performance of an LSPAN, and an auditory selective attention task (‘real-time’ inhibition), the latter two, indicating the broad range of performance tasks associated with the ability to process information quickly.

It appears, therefore, that when performing a speech task that is presented in a non-adaptive fashion at a particularly challenging SNR, it is the ability to process any received information quickly and use this information effectively that is of significance. This might have been anticipated for older participants, as it is known that there is a general slowing of processing speed due to ageing (Salthouse, 1996) but to demonstrate a significant relationship it must also be the case that it is both younger and older adults who have good processing speed who demonstrate better discrimination when the speech is particularly hard to perceive. It may also be the case that for older adults, faster processing is simply an indicator of a better-functioning neural network and of intelligence, and that this is what gives these older participants with better processing speeds the advantage in working out what may have been said.

For the research in this chapter, speed of information was measured by a choice reaction time (CRT), a response latency measure. Also using a CRT task, Jerger et al. (1991) did not give evidence of a relationship with two speech-in-noise tests (using high and low context sentences). The participants had mild/moderate hearing losses but were given an adjusted signal allowing equivalent access across participants to the speech test material. The major difference with this present research being that Jerger et al. presented the sentences at an SNR of +8dB. It could be hypothesised, therefore, that it is only when speech is particularly challenging to listen to that information processing speed demonstrates a significant relationship with speech perception in noise.

The above arguments are made even more convincing because a significant relationship was demonstrated not only on correlations but also when hierarchical regression was performed. Indeed, in the final model, when the other factors of PTA and age had been controlled for in model 1 and other cognitive functions in models 2,3 and 4, it was the contribution of processing speed that accounted for the most significant amount of the variance in performance.

There was no significant correlation of working memory span, as measured by an LSPAN, with speech-in-noise discrimination performance. However, other authors have established such a relationship using an LSPAN (Baldwin & Ash, 2011; Pichora-Fuller et al., 1995). There may be two potential reasons for the differences between those studies and the present work. Firstly, as the speech-in-noise test used here was particularly challenging at a -8 dB SNR, little of the speech is heard, and as the BKB sentences are short, holding information in working memory potentially offers no great advantage. Therefore, under these conditions, working memory does not offer a significant advantage in performing the task. Secondly, a wide age range of participants were recruited for this study. Since for younger listeners better working memory performance does not offer a great advantage when perceiving speech-in-noise (Füllgrabe & Rosen, 2016a), no potential relationship was evidenced, however, one may have been if all participants were older, as better working memory seems to offer an advantage for older individuals or those with a hearing loss when perceiving speech-in-noise (Rudner et al., 2011).

In similar fashion, control of inhibition as measured by the Stroop interference measure, selective auditory attention, and the RNG task also did not show a significant correlation with speech-in-noise performance. This again could be due to the speech test material and task, where although the ability to inhibit unwanted information and focus on a target would intuitively seem to offer an advantage when perceiving speech in noise, if the speech task is too challenging, then the ability to inhibit the unwanted stimulus perhaps does not offer an advantage, as not enough of the target speech has been heard.

However, when performing the hierarchical regression analysis, tasks that assessed inhibition control did add significantly to speech-in-noise performance variance. It appears, therefore, that this effect may only be seen when removing the factors of age, hearing level and other cognitive functions. Individually, on hierarchical regression modelling it was only response latency (the go condition of the go/no-go task) and inhibition control (as measured by the Stroop interference task) that significantly predicted speech-in-noise performance variance when all of the variables had been added. Previous authors have argued that control of inhibition and speed of processing are in fact tapping the same cognitive function (Uttl & Graf, 1997), and that it is simply the quicker access to information that allows for superior control of inhibition. This is also argued by Knight and Heinrich (2017), who offer explanations, including generalised slowing and sensory declines, as to why the Stroop interference measure may not be a simple measure of inhibition. The fact that, individually, the two tasks (representing the two functions) both contribute to the speech-in-noise performance may suggest that while linked, the two tasks do represent different functions. The exact split between the contributions of the two functions, however, is hard to precisely quantify, as the two tasks themselves significantly correlate. If, as argued by Uttl and others, Stroop performance is mainly a processing speed task, then having accounted for response latency in Model 2, there should be no additional explanation of speech-in-noise performance variance when adding the Stroop interference measure. The fact that, when examining the hierarchical regression model, the Stroop results added a further significant explanation of the variance suggests that control of inhibition, whilst involving a processing speed element, is indeed a separate function. This would at least appear to be the case when it comes to perceiving speech-in-noise.

The report from the hierarchical regression analysis and the significant prediction of speech-in-noise performance is in contrast to the research by Helfer and Freyman (2014), which did not indicate a significant relationship with Stroop performance. The Helfer and Freyman research used performance on discrimination of speech in speech-shaped noise, averaged across three SNRs of -3, 0 and +3 dB. The potential reason for the difference in the reported relationships is that the research for this chapter was significantly more challenging, as the SNR was fixed at -8 dB SNR. This suggests that the influence of inhibition control as assessed by the Stroop may only be evident when the speech task is particularly challenging and when the influences of other cognitive functions on speech-in-noise performance have already been factored out.

There could be two potential reasons for this finding. Firstly, in less challenging SNRs/listening conditions or when a listener is able to 'tune' in to the speech (adaptive speech tests), listeners use other cognitive repair mechanisms if they cannot hear clearly, so there is no evidence of a relationship with prepotent inhibition control. Secondly, and pertinent for this research, a screening Stroop task was used that required participants to adapt to the task quickly. This potentially allowed for greater performance differences to be exhibited. The reason for this is that there is a build-up of selective inhibition that affects the two Stroop conditions of congruence and incongruence differently (Ridderinkhof, van den Wildenberg, Wijnen, & Burle, 2004). Recording results before any selective attention has had time to be activated may give a better representation of an individual's control of prepotent inhibition, particularly when compared to speech-in-noise performance where there has also been no adaptation to the task.

There was additionally, a difference in the strength of the relationships of speech-in-noise performance with performance on the Stroop task (the interference measure) and the RNG (the RNG and the 'A' score, the two inhibition measures). Both the Stroop and the RNG tasks involve inhibiting prepotent responses, yet they were not significantly correlated with each other. This could be due to the type of prepotent inhibition control needed. The Stroop does involve some degree of more active 'real-time' inhibition control in that a participant is presented with information and needs to judge what to attend to whilst suppressing the other unwanted stimulus. Although the RNG requires prepotent inhibition to avoid naming numbers in sequences or patterns, there is no auditory or visual stimulus that needs to be inhibited.

The ability to respond quickly to a stimulus, on a fairly basic level may, therefore, be linked to an individual's speech-in-noise perception. A potential reason being that if a swift match can be achieved with stored phonological information there is a better chance of keeping this information 'live' in working memory until the necessary matches for partially heard words has been completed. This potentially illustrates how working memory, another factor associated with speech-in-noise performance, may also be linked to how quickly information enters. It appears, therefore, that individuals rely on factors other than simply audition and the ability to perceive speech to maintain their speech-in-noise performance. Therefore, older individuals who have better cognitive performance are indicated to have superior speech-in-noise performance.

2.5. Limitations

A short screening Stroop task was used in this research, which meant that participants had no time to adjust to the task. This was intended, but it did mean that if participants made errors in colour naming, there was not a full set of results from which to calculate an average response. Some results, therefore, may have differed from those of a Stroop task where task acclimatisation is possible. A shorter Stroop task, however, might offer an opportunity to be run as part of a clinical test battery that may be ultimately used to design a communication training program. This, along with not allowing participants an acclimitisation period, was the reason it was kept as a screening test, as it could then potentially guide rehabilitation advice.

Testing sessions lasted 90 minutes. As not all measures of inhibition showed a relationship with the speech-in-noise task at -8 dB SNR in speech-shaped noise, it would have been a useful extension to this research to have performed speech-in-noise tasks in more than one background noise condition. Speech tests in different background noise conditions, such as babble, would have been a useful addition, but assessment time was already long and involved tests needing consistently high amounts of concentration.

It could further be argued that using nonsensical sentences rather than the grammatically and structurally sound BKB sentences may have resulted in greater difficulty, as basic grammatical rules would not have allowed for support from top-down expectation processes. However, keeping the task transferrable and relatable to real-life situations was preferred to using nonsense words.

2.6. Conclusions

For the participants recruited for this study (all of whom had expressed no hearing concerns, although some had a mild degree of hearing loss), it was only speed of information processing (a choice reaction time (CRT) response latency task) that correlated with speech-in-noise performance. The faster the information processing speed the better the speech-in-noise score. It would appear to be the case that when speech remains consistently challenging, stressing the auditory system, that it is under these conditions that a relationship between information processing speed and speech-in-noise performance becomes evident. This was the case regardless of hearing thresholds, age and hearing threshold level which were partialled out in a hierarchical regression model.

Control of inhibition as measured by a screening Stroop test also explained some of the variance in speech-in-noise performance. This again was after demographic information and other cognitive functions were partialled out. The Stroop test employed was a quick screening test and as yet it is unclear if the screening Stroop is a reliable tool for recording inhibition. Although the results recorded for Stroop performance were suitably varied, recording the interference caused by the Stroop test before any adaptation to the task can occur also seems to give evidence of those individuals that have better instant prepotent control. In turn, this offers insight into how a listener can cope with a speech task when there is no adaptation time given. Previous researchers have suggested that the Stroop is actually based largely on processing speed (Uttl & Graf, 1997) which may be a further reason why a relationship was seen with the speech-in-noise task in this chapter. As an added advantage the screening Stroop, being a quick measure, could be readily incorporated as a clinical tool in a hearing evaluation appointment.

Since the speed of responses plays a role in the majority of cognitive measures (Salthouse, 1996), it would seem key to try and preserve response speed, employing perhaps one of the many computer-based training programs available on the market (COGMED). This may then moderate any decline in performance of many of the complex tasks that feed into hearing speech-in-noise. The general underpinning of functions by speed of information processing and how they affect speech performance is also worthy of further investigation.

As there was evidence that cognitive functions did support (or at least explain some of the variance in speech-in-noise performance) then a training program to improve the performance of these functions may offer support to an individual when attempting to discriminate speech-in-noise. The potential for this to occur is explored in the next chapter.

CHAPTER THREE

3. Does inhibition and processing speed training improve performance of speech-in-noise and other associated cognitive tasks?

3.1. Introduction

This chapter opens with a summary of the previous chapter in this thesis, along with a brief outline of the aims of this chapter and the theory that informed them. Then, a section on the general aims of cognitive training is followed by a further section exploring previous research where cognitive training has been used in an attempt to improve performance in speech-in-noise perception. The remaining part of this chapter focusses on the cognitive training study undertaken for this thesis.

In Chapter 2, it was demonstrated that speed of information processing, as evidenced by a choice reaction time test (the go condition of the go/no-go task), and control of inhibition, as evidenced by the Stroop interference measure, can both add significantly to the explanation of variability in speech-in-noise perception performance after both hearing level and age have been partialled out. However, it remains a fact that as has been previously reported, the most consistent predictor of speech-in-noise perception performance is an individual's level of audition (Humes, 2002; Humes & Dunbo, 2010; van Rooij & Plomp, 1992a).

This chapter, therefore, explores whether training the cognitive functions identified in the previous chapter of processing speed and control of inhibition could lead to improved speech-in-noise performance. By recruiting participants similar to those who volunteered for Chapter 2, it was hypothesised that if the same cognitive functions again demonstrated significant relationships with speech-in-noise performance, then training these same functions may consequently lead to improved speech-in-noise performance.

No previous research has compared cognitive training in two different sensory modalities and the effect that this may have on speech-in-noise performance. The aim of using the two sensory modalities of vision and hearing was to ascertain whether training the cognitive function itself works just as efficiently in either modality or if there are greater improvements when the training is carried out in the sensory modality of the task in which the desired improvement is sought. For example, if seeking to improve speech-in-noise perception by training working memory, should the working memory training be delivered auditorily or visually, or in fact, does it make no difference? If the research in this chapter demonstrates that cognitive training does prove beneficial for speech-in-noise perception, then further information on how such training should be delivered would be an important addition in developing rehabilitative strategies.

Furthermore, the research for this chapter set out to distinguish between the two types of inhibition: prepotent inhibition (as measured by a Stroop test) and the more ‘on line’ inhibition, in which irrelevant information has to be ignored to prevent it from being processed and using up cognitive resources (as measured by an auditory selection type task). By attempting to differentiate between these two types of inhibition control in greater detail, the type of inhibition control most relevant to speech-in-noise perception could be used in future cognitive training for listeners who have problems hearing in noise.

3.1.1. General cognitive training

The concept of cognitive training was explained in the introduction to this thesis in Section 1.2.5. Cognitive training has been able to demonstrate improved performance of tasks associated with the training (on task) as well as closely associated (near transfer) and loosely associated tasks (far transfer; Au et al., 2015). The latter aspect of far transfer is a key principle of cognitive training and is particularly pertinent to the research in this chapter, in that the hypothesis is that training different cognitive functions may allow for improved speech-in-noise performance (far transfer).

Much of the research into cognitive training has focused on improving fluid intelligence (Au et al., 2016; Au et al., 2015), but previous research on such training has not reported universal success, with the opposing view that any improvements are due more to the placebo effects of undertaking the training (Redick et al., 2013). There is currently, therefore, no definitive answer as to whether cognitive training really provides an advantage in terms of fluid intelligence.

Aligned with attempts to improve fluid intelligence, there has also been recent interest in the role of brain training in the prevention of cognitive decline and dementia (Smith, 2016). In a meta-analysis, Smith outlined many of the positive benefits of training, with a plea for further research into health promotion services such as 'brain fitness' that have shown potential for reducing dementia. Gates and Sachdev (2014), in a review that found that cognitive training was effective for mild cognitive impairment and preclinical early Alzheimer's disease, reported on the potential for training to improve cognitive skills, lower the incidence of dementia and also slow the rate of any decline.

As an example of the application of auditory cognitive training to facilitate improved listening skills, Smith et al. (2009) utilised a novel computerised cognitive training program with older, community-dwelling adults and compared them to an active control group.[26] The experiment involved six auditory-based computerised exercises aimed at improving processing speed and accuracy of perception of auditory information. Training was performed on computers with remote access. The active control group performed training of a similar length that engaged learning processes but was not designed to improve auditory system function. The active control group received the same amount of audio-visual and computer input as the auditory training group. The training for this study was intensive, requiring 40 hours of training over 8 weeks.

Smith et al.'s research employed the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS auditory modality), which tests orally presented lists of words, story memory, digit span forward and delayed list recall and recognition. The experimental group performed the RBANS significantly better when compared to the active control after the training period. The RBANS-trained group also demonstrated significantly greater improvement than the active control in other tasks (delayed recall, backwards digits, and letter number sequencing, all at $p < .05$), thus demonstrating a degree of near transfer. There was, however, no evidence of far transfer to narrative memory. The Smith et al. research

---

[26] All 65 years and older with no significant cognitive decline as measured by MMSE.

was not aimed at improving speech-in-noise perception, but they did question whether there may have been improvements that were not measured. It would have been a useful addition if there was potential for communication improvement for these older adults as well. The research in this chapter attempts to demonstrate the possibility of such improvements.

In view of the close association of sensory decline and intelligence (Lindenberger & Baltes, 1994) and the further demonstrated association between hearing loss and cognitive decline/dementia (Lin, Metter, et al., 2011; Lin et al., 2013), remedial strategies that may lead to improved outcomes need exploration. Dementia prevention can be assisted by cognitive training (Ball et al., 2002; Livingston et al., 2017), and there is evidence that hearing loss may be modifiable by intervention (Amieva et al., 2015; Deal et al., 2015; Livingston et al., 2017). Re-introducing better audition has been hypothesised as preventing further cognitive decline (Sekuler & Blake, 1987). The interactions between auditory training, cognitive training and greater access to and continued use of hearing intervention are not well understood and require further exploration. This chapter focusses mainly on the role that cognitive training may offer in terms of any communication improvements.

3.1.2. Working memory training and fluid intelligence

Before considering the potential gains that working memory training may offer in terms of communication support, it is worth considering the role that it may play in improvement of fluid intelligence. The concept that working memory training leads to gains in the performance of tasks in different cognitive domains is appealing, as it may lead to improved work-based performance and its associated societal benefits. The same argument could be made for general cognitive training if it can be seen to offer any improvement to communication and the protection that communication ability may provide against social isolation and other associated mental-health problems (Saito et al., 2010).

Evidence of far transfer from training on working memory to gains in fluid intelligence (Au et al., 2015) has been demonstrated in the form of improved performance on tasks such as the Tower of Hanoi (ToH) and Tower of London (ToL). These tasks involve logical reasoning and the thinking through of novel problems independent of previously acquired knowledge. Although the findings and conclusions of the Au et al. meta-analysis have been disputed (Melby-Lervåg & Hulme, 2016) and a 'theoretical grounding' in identifying the mechanics of any transfer that occurs is lacking (Redick et al., 2013), working memory itself certainly improves significantly after training. There also appears to be sufficient evidence to suggest that some degree of far transfer is a strong possibility. Since working memory span is a known predictor of speech-in-noise perception performance for older adults (Fuellgrabe, Moore, & Stone, 2015), then hypothetically, training to improve it may lead to improved speech-in-noise perception.

3.1.3. Cognitive functions and the perception of speech in noise

Much of the research exploring the relationship between cognitive functions and speech perception in noise has been discussed in Section 1.2.1; therefore, this section simply summarises these main areas.

The relationship between working memory performance and speech perception in noise has been extensively researched (Rönnberg et al., 2008; Rudner et al., 2011). In-depth reviews by both Akeroyd (2008) and Besser et al. (2013a) give further examples of where this relationship has been demonstrated. The general relationship for older adults and those with hearing loss is that individuals with better speech-in-noise discrimination scores have a tendency to have better working memory spans and vice versa. This relationship does not universally hold, as it does appear to be dependent on age and hearing thresholds. Indeed, in a meta-analysis, working memory does not seem to be important for speech-in-noise perception in younger adults (Füllgrabe & Rosen, 2016b). When partialling out hearing loss, the relationship between working memory capacity and speech-in-noise perception performance for younger listeners[27] gave a non-significant correlation ($r = .18$; $p = .162$), but for middle-aged to older adults,[28] the relationship was significant ($r = .44$; $p \leq .011$).

---

[27] 18–39 years

[28] 40–91 years

As listening conditions deteriorate or the incoming speech signal becomes degraded or altered, there are proposed models predicting that the listener will have to invoke the use of cognitive support mechanisms to maintain their listening ability (Rönnberg et al., 2013). Training the supporting cognitive functions that may enable the explicit processing loop proposed by Rönnberg et al. to function better could aid in speech discrimination. Additionally, auditory training would allow better auditory patterns to be developed, such that binding of incoming speech to these phonological representations becomes more automatic with less demand on cognitive processing, freeing cognitive resources for tasks such as secondary encoding of the information.

Alongside the importance of working memory discussed above, Chapter 2 of this thesis identified other cognitive skills that explained some of the variation in speech-in-noise perception performance. The key skills identified in that chapter were speed of information processing (the ability to access information quickly) and control of inhibition. These two cognitive functions added significantly to the hierarchical model of the explanation of variance in speech-in-noise perception ability after factoring out age, hearing level and working memory. There appears, therefore, to be a need for further investigation into the possibility that training in a range of different cognitive functions, outside of purely working memory, may have effects on speech-in-noise discrimination.

Information processing speed, including response latency measures have been demonstrated as factors that predict speech-in-noise discrimination performance (Pronk, Deeg, Festen, et al., 2013). Indeed, when analysing data from the Amsterdam Longitudinal Ageing Study, Pronk et al. stated that only processing speed influenced the changes in speech reception with age, not global functioning or fluid intelligence. The processing speed task used was a coding task that involved matching as many characters as possible from a top row with characters in a lower row in one minute while ignoring other distracting information. The speech reception task was the digit triplet test described in Section 1.3.2. Testing was performed in participants' own homes, so the type of speech-in-noise test that could be used was limited due to calibration issues. A limitation of the delivery of single digits in background noise, as in the triple test, is that it does not allow for adequate exploration of the role of all possible cognitive functions. Listening to three digits in noise and holding them in memory is not particularly taxing, and the background noise is non-informational.

This limitation may explain why global cognitive functioning and fluid intelligence did not also demonstrate a relationship with speech-in-noise performance in the Pronk et al. study. Also, there were no direct measures reported of working memory or inhibition. It would appear advantageous to use sentence-based tasks for speech-in-noise performance where possible, as these will tap into more support mechanisms afforded by cognitive functions as well as exploring any potential reliance on semantic knowledge.

In a review of different types of Stroop tests and their relationship with speech-in-noise tasks, Knight and Heinrich (2017) reported that the control of inhibition does not consistently demonstrate a relationship with speech-in-noise perception. Confounding factors such as the modality of the Stroop test, the type of speech task, and computation of Stroop interference measures all have an effect on any potential relationship. Furthermore, when considering any possible relationship, Knight and Heinrich (2017) stated that hearing sensitivity in conjunction with the Stroop measure has a further influence. For this present research, the participants all presented with normal hearing, thus eliminating some variability. The same standard, single-word, computer-based visual Stroop presentation method that established a significant relationship with speech-in-noise performance in Chapter 2 was used again here although it was a full Stroop and not a screen.

Research into computer-based online training aimed at improving speech-in-noise discrimination performance is discussed in Chapter 5, including where hearing-impaired individuals have shown evidence of improvements.[29] Training to improve processing speed and control of inhibition with the end goal of improving speech-in-noise discrimination performance for individuals with no reported hearing loss has not been researched and hence was one of the reasons for the research performed in this chapter.

---

[29] For a further review see Henshaw and Ferguson (2013).

The specific areas of training of cognitive functions in an attempt to improve speech-in-noise discrimination performance are discussed in the next section. While research exploring the individual roles of cognitive functions in speech-in-noise perception has been carried out, previous research has focused primarily on the role of working memory training. The training of other cognitive functions is employed in general communication training packages such as LACE, but there is no research evidence regarding effects of processing speed and inhibition when these cognitive functions are trained in isolation.

3.1.4. Cognitive function training with the aim of improving the discrimination of speech-in-noise

There is a general paucity of research on the effect of training of cognitive functions on speech perception in noise. Where research has been performed, it has focused on working memory. A report by Ferguson and Henshaw (2015c) has compared three randomised, controlled studies of participants with hearing loss;[30] the improvements observed in two of these studies after auditory training have already been reported in Section 1.2.5. The third study, by Henshaw and Ferguson, involved training on verbal, visuospatial working memory and memory storage tasks. The training utilised an on-line training program (COGMED) that adjusts itself to remain challenging at all times. Although near transfer was demonstrated by the trained group following training, performing significantly better than a control group, there was no evidence of far transfer improvements on untrained tasks such as speech-in-noise perception.

Wayne, Hamilton, Huyck, and Johnsrude (2016), also using COGMED, researched the potential gains in speech-in-noise discrimination performance following working memory training. However, in similar fashion to Henshaw and Ferguson, after training for 25 sessions, on-task improvement of the trained working memory tasks did show a significant improvement, but the far transfer goal of improved speech-in-noise performance did not. Additionally, the

---

[30] Using two studies of auditory training and one of cognitive working memory training.

researchers were not able to demonstrate any near transfer improvements on other working memory tasks.

For the evaluation of training delivered via different methods, the inability to infer whether one form of delivery is superior to another is a restricting factor. To attempt to address this question, although working memory training was not employed for this chapter, the cognitive functions of processing speed and control of inhibition were trained in different sensory modalities so that differences could be observed. One of the working hypotheses was that training in an auditory modality, although tapping the same cognitive function as visual training, will offer greater advantages because the auditory system is stimulated and new or strengthened phonological representations are established. It was also hypothesised that training these functions would still offer an advantage over no training whichever sense the training was delivered in. If this holds true, then training the cognitive functions visually, although potentially not giving the same results as auditory training for auditory-based tasks, would still demonstrate an improvement of tasks demanding the input of these cognitive functions.

Previous research by Schneiders et al. (2012) has explored this potential cross-over of training in one sensory modality to potentially invoke performance improvements in another sensory modality. This research involved training working memory and was conducted in an auditory mode. The researchers wanted to explore whether working memory improvements achieved via auditory-based working memory training could be transferred to a visual task. Training involved performing an auditory *n*-back, which is based on an adaptive working memory task. Following training, the participants' performances were compared to those of a control group that undertook no training. Similar to the other research reported in this section, significant 'on task' improvements were demonstrated, yet there was no cross-over or near transfer to improvement on an analogous visual working memory task. This would seem to indicate that if cognitive training is used, then the sensory modality, at least where working memory training is concerned, does have an important bearing. Theoretically, as the auditory and visual cortices are separate, then although the pre-frontal cortex will be stimulated by cognitive training (strengthening responses), there will also be neural links developed with the individual sensory cortex that is trained. It may, therefore, be reasonable to hypothesise that the most effective way to induce improved cognitive function to assist with a specific task such as speech-in-noise discrimination is to train in the sensory modality of the task. As the methods for the research in this chapter allowed for training in the different sensory modalities of processing speed and control of inhibition, this research may help corroborate if the sense in which the training is undertaken makes a difference.

The above section demonstrates that concrete evidence of far transfer gains from cognitive training remains somewhat elusive. This in turn brings into question whether any training gains may be transferable to improved performance of tasks in the real world, such as discriminating speech in degraded listening conditions.

3.1.5. Aims and hypotheses

The primary aim of this chapter was to investigate whether cognitive training offers any advantage for individuals without hearing loss (hearing aid non-users) in perceiving speech-in-noise. It is hypothesised that if the cognitive functions of speed of information processing and control of inhibition contribute to speech-in-noise perception, training of these functions will thus improve speech-in-noise discrimination performance. Participants were randomly assigned to one of three training pathways, one involving a computer-based program to train individuals' visual processing speed and inhibition control, the second involving a computer program to improve individuals' auditory processing speed and inhibition control, and the third an active control pathway that involved completing a cognitive workbook. The use of three groups means that any differences in training gains between the different types of training can be explored. The hypothesis was that training auditorily will provide the greatest advantage for speech-in-noise discrimination performance, as training gains are generally task-specific, and it is anticipated that the same will be true for sensory modality.

As the cognitive function tasks and subsequent training were visual- or auditory-based, this affords the opportunity to analyse at baseline whether visual or auditory cognitive function tasks are better predictors of speech-in-noise discrimination performance. As two types of speech-in-noise task were used, it also afforded the opportunity to explore which cognitive functions were most useful in the different types of background noise tested. Further, there was also the opportunity to explore which of the two types of inhibition, prepotent or more ongoing ‘real-time’ inhibition, is more closely associated with the ability to hear speech-in-noise.

The hypotheses to be tested were as follows:

H3.1 If training on the cognitive functions of information processing speed and inhibition control results in improved speech perception in noise, then the groups that are cognitively trained via computer will demonstrate significantly better SIN performance than the active control group trained on general cognitive exercises.

H3.2. If training in a sensory modality allows for better performance of tasks in the same modality, auditorily trained participants will perform auditory cognitive-based tasks significantly better after training than will visually or control-trained participants. The reverse will be true for any vision-based tasks.

H3.3. If cognitive functions are most useful when glimpses of speech can be more readily perceived, then cognitive functions will correlate better with a ‘speech-in-babble’ task than a ‘consistently challenging speech in speech-shaped noise’ task presented at a particularly adverse SNR.

H3.4. Auditory tasks will predict performance in speech-in-noise tasks at baseline better than visual tasks will.

H3.5. If real-time inhibition control is the dominant cognitive performance factor influencing speech-in-noise discrimination performance, then assessments utilising this function will correlate to a greater extent with speech-in-noise performance than those utilising the inhibition of prepotent responses will.

H3.6. If ageing and small declines in hearing levels are not the only predictors of speech-in-noise discrimination performance, then once age and hearing levels have been controlled for, cognitive functions will add to the explanation of any variation in speech-in-noise performance.

3.2. Methods

3.2.1. Ethics

This research was approved by the Aston University Research Ethics Committee, reference number 1034.

3.2.2. Participants

There was no previous research with the same design as this present study. The sample size calculation was, therefore, based on research by Schneiders et al. (2012) that had reported an significant effect size of .28 for a three way interaction of task modality (auditory or hearing), training group (trained v untrained) and time. Using G power (Faul et al., 2007), in order to demonstrate a significant interaction of training group and time employing an alpha value of .05 and a probability of .80 the calculated number of participants was given as 36. This was calculated for 3 groups using a mixed methods ANOVA with two measurement points. As the research was not exactly the same design and dropouts were anticipated, 48 participants were recruited to allow 16 to be randomly allocated to each training group.

There were 48 adult participants recruited for the study that had not taken part in any previous studies. Their ages ranged from 18 to 80 years old, with a mean age of 49.67. (SD 22.56) The participants were recruited from three different sources. They were either recruited through an invitation from the University Aston Research Centre for Healthy Ageing (ARCHA), from the psychology undergraduate research participation scheme (SONA) or were colleagues at University or associates of the researcher. Participants in the ARCHA group were given £7.50 per visit to cover travel expenses, and students from SONA were given course credits. All participants were required to have English as their first language. Participants did not report any problems with colour blindness. There were also no hearing aid users, and no one had sought help concerning their hearing. All participants had the ability to access the internet and access to a personal computer/tablet. In consenting to perform the research, each participant agreed to undertake six training sessions as well as to make two visits to Aston University.

Stratified randomisation was used (Suresh, 2011) to allocate participants to groups. This was performed using the computer software program (GraphPad). The training groups are summarised below;

- Group 1 received computerised auditory training on tasks involving an element of inhibition and processing speed.
- Group 2 underwent computerised visual training on tasks requiring inhibition and processing speed.
- Group 3 were issued a cognitive exercise book and served as an active control group.

The participants were age matched as far as possible in each of the groups in order to balance the effects of age on both the speech discrimination in noise performance tasks and the cognitive tasks, both of which have been shown to decline with age.

The demographics of the participants are outlined below in Table 1.

Table 1
Participant demographics

Training Pathway	**Mean Age (years)**	**Age Standard Deviation**	**PTA average 500-4000Hz**	**PTA Standard Deviation**	**Total (N=48)**
Auditory training	51.69 (range 18-80)	22.41	10.97	5.78	16
Visual training	45.75 (range 19-74)	23.55	8.59	5.57	16
Cognitive workbook	51.56 (range 19-80)	22.66	10.97	5.49	16

Key: PTA: pure tone audiogram thresholds
Note: ANOVA testing for age and PTA threshold demonstrated no significant difference between groups, *p* >.05.

3.2.3. General design

A within-between study design was used. Baseline assessments were performed, then, following a two-week intensive training program[31] a further set of the same assessments were undertaken. The between-participants' factor was the type of training undertaken: visual training, auditory training or active control. Participant performance of the two speech-in-noise scores for the BKB test run at -8dBSNR and the QuickSIN tests served as the dependent variables.

The cognitive functions measured were assessments to explore working memory, processing speed and the control of inhibition. The inhibition tests involved both prepotent inhibition and real-time inhibition. These were all analysed before and after the training period. Participants completed all the cognitive assessments in each of the two sensory modalities of auditory and vision.

Although the main study design was between and repeated measures, correlational analysis was performed at baseline to explore for predictors of speech-in-noise performance for both types of speech-in-noise.

3.2.4. Test materials

3.2.4.1. Audiometry

The equipment to carry out the initial screen and perform audiometry was the same as used in Section 2.2.4.1.

3.2.4.2. Discrimination tasks

BKB test

The Bamford-Kowal-Bench (BKB; (Bench et al., 1979) speech test material was used, and the method for conducting the test and the challenging SNR of -8dB was the same as described

---

[31] Each of 6 training sessions lasting 30 minutes so as not to over train which can be detrimental (Wright & Sabin, 2007) and to allow for consolidation between training sessions (Goedert & Miller, 2008)

in Chapter 2 (Section 2.2.4.2). BKB lists 2, 3 and 4 were used for the baseline assessment which meant 48 sentences were listened to. Lists 9, 10 and 11 were used after the training period. A loose scoring method was adopted as described in Section 2.2.5.

QuickSIN test

The QuickSIN test (Killion, Niquette, Gudmundsen, Revit, & Banerjee, 2004), was delivered through the same test apparatus as for the BKB test described above. The talker for the QuickSIN test is a woman speaker of American English against a background babble of four talkers (three women and one man). The sentences do not contain many semantic clues, making the material not highly predictable, for example 'The **beetle droned** in the **hot June sun**' (target words in bold font). The test and re-test reliability has been found to be good (Wilson et al, 2007). The sentences are organised in groups of six, and the SNR starts at the first of the lists at +25dB and then reduces the SNR in 5dB steps to 0dBSNR for the sixth sentence, so in effect getting progressively harder to discriminate the target talker. The participant repeats the sentence immediately after they hear it, and the score for each sentence is recorded. The SNR automatically reduces for the next sentence. The number of correct key scoring words in each six-sentence list is summed and subtracted from 25.5 to calculate the SNR in dB that is needed for an individual to identify 50% of words correctly.

On their first visit, participants were presented list 3 list for practice, then were scored on lists 4 to 7 and an average dBSNR was recorded. At the second visit, list 8 was used as a practice list and performance was measured across lists 9 to 12. Lower dBSNR levels indicate that the individual is better at discerning speech from the background babble, as the speech can still be understood at a level just slightly higher or even below the background babble level. It is possible for individuals to achieve a negative dBSNR score.

3.2.4.3. Reading sentence span (RSPAN)

The working memory task based on the work of Daneman and Carpenter (1980) was performed in identical fashion as described in Section 2.2.4.3. However, for the research in this chapter the task was performed as an RSPAN. So instead of the participants hearing the sentences (to make judgements on) and letters to remember, these were both delivered in written format on a computer screen.

3.2.4.4. Random number generation tasks

The method used for the random number generation task was the same as in Chapter 2 (Section 2.2.4.6) based on the work of Towse and Neil (1998).

3.2.4.5. Useful field of vision test (UfoV)

The Useful Field of Vision test (UfoV; version 6.1.4) The UfoV comprises of three subtests of increasing difficulty in a set order:

1. ***Central vision and processing speed task.*** This task required the participant to identify a central target, a 2.0 cm × 1.5 cm outline of a car or truck on the computer screen, and then, using a computer mouse click on the target that they saw, i.e. whether it was a car or a truck.

2. ***Divided attention task.*** This test required the participant to identify whether the central target was a car or truck but also to localise a peripheral target (a 2.0 cm × 1.5 cm outline of a car) presented simultaneously at one of eight radial locations, approximately 11 cm from the centre of the screen where the central target was displayed. Using a computer mouse, the participant would then click on whether the central target was a car or truck and where the other target was located on one of eight different radial points.

3. ***Selective attention task.*** For this test, the peripheral targets were embedded amongst visual distractors (triangles of the same size as the targets). The task was the same as for the divided attention task, requiring identification of whether the central target was a car or a truck and where on eight radial points the other car was located. The participant responded using a computer mouse in the same way as in Test 2.

For each of the three test conditions described of processing speed, divided attention and selective attention, the participant had to identify what the target was for the first test and additionally where a secondary target may have been for the second two tests. These targets were displayed for between 16.7 and 500.0 msecs. Using a double staircase method[32], the individual's score recorded was the quickest time in which a participant could achieve a 75% correct response rate for each of the three tests. The test took approximately 15 minutes for participants to complete.

3.2.4.6. Auditory selective attention task

The auditory selective attention task was in the same format as described in Section 2.2.4.5. This involved a male and female speaker with the listener being cued as to which one to attend to and then having to identify what the target talker had said. On analysing the previous testing, it was shown that after the first four practice presentations, the majority of listeners had grasped what was necessary for task performance so only four practice presentations were given during the research in this chapter.

3.2.4.7. Stroop task

A screening Stroop test was employed in the research in Chapter 2, which is described in detail in Section 2.2.4.4. A longer Stroop test was used in this chapter as rather than a measure exploring quick adaptation to the task a longer set of responses was used to give a more classical measure of prepotent inhibition control (Stroop 1935; MacLeod, 1991).

Response latency was taken from the Stroop congruent trials and used as a further measurement of participant performance. On seeing the word on the computer screen, the participant was asked to say aloud the hue of the writing. The computer's built-in microphone recorded the elapsed time and the colour named. As the Stroop test used microphone voice

---

[32] Targets and any distractors are displayed for shorter and shorter times until the target was missed and then target display time is increased until a correct response is achieved.

recording, all responses were checked for accuracy and all incorrect colour responses were ignored. In addition, errors in the microphone recording itself were also ignored. The test used four lists of twenty word presentations. The first list was a practice list consisting of 20 presentations in a ratio of 2/3rds incongruent to 1/3 congruent. The subsequent three lists of 20 words made up of the same ratio were used to obtain two measures. Firstly, a processing speed measure of recognition in the congruent condition and secondly a measure of inhibition control where the average congruent recorded response time was taken from the average incongruent response time. This is the most widely used way of calculating Stroop scores (MacLeod, 1991).

All the recorded utterances of the participants were analysed before inclusion in the final results. Some responses had to be discounted as the participant had got the colour wrong or uttered a sound such as ‘umm’ before saying a word, which went against the detailed given instructions. This was due to the microphone having recorded the first uttered sound, leading to errors in recorded results. These were all checked and removed.

3.2.4.8. Time compressed speech test

A list of 20 of the BKB sentences created by Patel (2012) were used for the compressed speech test. These sentences were spoken by a male speaker with no background noise. Speech tempo compression was achieved using Audacity (R), (Audacity, 2017). A time compression of 50% was used, which has the effect of speeding up speech as it is delivered in half its normal time, without its pitch being altered. This percentage of time compression was used as doubling the speed of presentation has been shown to reduce intelligibility of what is spoken to 50% (Versfeld & Dreschler, 2002). This had the effect of making the task consistently challenging.

Participants were given four practice compressed sentences to acclimatise them to the speech material and the task. The test involved listening to 20-sentences and the participant repeated what they had heard immediately after each of the 20 sentences. The same loose scoring method as described for BKB testing in Section 2.2.4.2 was adopted.

The time compressed speech score was calculated as a percentage of the number of key words correctly repeated by the participant out of the total number of key words in the BKB sentences.

3.2.5. Procedure

All the testing was carried out in a well-lit (320-360 lux) soundproof room at Aston University. All the testing apart from pure tone audiometry was performed at baseline and at follow-up after the training pathway. The ear examination and PTA were the same as performed in Section 2.2.5 and there were no exclusions necessary due to the medical condition of the ears or participants' hearing.

The test order after the initial ear examination and hearing test was: 1. BKB and 2. QuickSIN, 3. RSPAN 4. Random number generation task, 5. UfoV 6. Auditory selective attention task 7. Stroop and 8. Time compressed speech. The same order of testing was used before and after training.

The speech discrimination tasks were delivered through a Samsung R350 personal computer and an Edirol UA-25 audio capture box. The output was played through Sennheiser HD250 headphones and was calibrated to deliver the speech at 65dBSPL (normal conversational voice level) as measured by a Brüel and Kjær 2250 sound level meter. The UfoV was also run via the Samsung R350 but without the soundbox as it was a visual task. The rest of the cognitive test battery (apart from the time compressed sentences) were administered via a Toshiba Satellite Pro personal computer. The Stroop, Selective attention and RSPAN were all performed on the Satellite pro using Millisecond software (Inquisit). The RNG task was a stand-alone program run on the same computer. The time compressed sentence test was run through a Lenovo X260 personal computer with the same sound box calibrated at the same level as for the discrimination tasks. Breaks between the assessments were given to participants where requested.

3.2.6. Training

Participants were randomly allocated to training groups and they performed all the above procedures at their first visit as a baseline measure. After training they performed all the tasks again at a second visit. The types of training and exercises undertaken and the training pathways are expanded upon later in this section. The training pathway and the assessment sessions are given below in Figure 1.

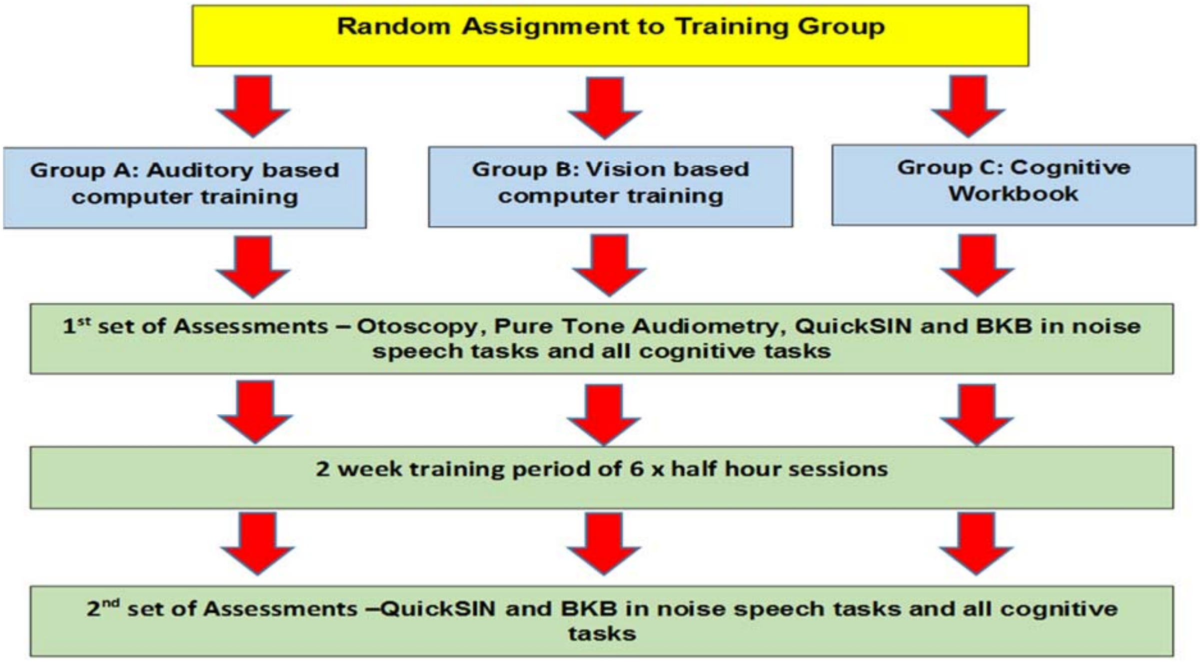


Figure 1. Training pathways for auditory, visual and active control groups.

Participants were randomly assigned to one of the three different groups. Pathway Groups 1 and 2 were set up with training accounts on BrainHQ ("BrainHQ," 2012) by Posit Science. BrainHQ was chosen as previous work using auditory training via Brain Fit had demonstrated that both ‘on-task’ and generalisable improvements as measured by the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) in memory and attention (auditory modality) could be achieved, when compared against an active control group (Smith et al., 2009).

Participants on Pathway 3 were given a cognitive workbook to work through with instructions. The completed workbook was returned to the researcher during the second visit to Aston University. The participants on the computer-based training pathway used their email address as their account name and were given a unique password to access the BrainHQ website. The researcher acted as the administrator, set up each account and monitored training compliance online. If there was evidence that no training was taking place, an email was sent to a participant to enquire if there were any issues.

As administrator, the researcher was able to personalise the training undertaken by the participants and assign them either mainly auditory or visual tasks depending on the pathway to which they were allocated. All the tasks involved an element of processing speed training and/or inhibition control. As with many cognitive training tasks, there was an element of other cognitive skills being trained alongside processing speed and inhibition. The computer-trained participants had their training set for 30-minute sessions and were asked to perform 3 sessions a week for 2 weeks (6 sessions in total).

The training exercises for the visual and auditory speed of processing and inhibition groups were as follows:

Vision training.

Targeting processing speed/inhibition.

Divided attention. The participant had to focus on and react to a particular detail, such as matching colours, shapes, and/or the interior filling of a shape, whilst at the same time dismissing competing information. Participants were shown two shapes on a screen and were asked to press the left arrow key when the two shapes met a certain criteria. For example, they may have had to press the left arrow key when the two shapes were the same colour, but the right arrow key when they were not. As the participant improved when working through the training, the presented images flashed by more quickly, requiring faster processing to attend to the target information whilst ignoring the wrong information.

Double decision. The participants were trained to improve their processing speed by extracting visual information more quickly. The task adjusted such that the participant was working at the edge of their useful field of view. By doing this, it was pushing the field of view outward. This was achieved by having the participant focus their attention on a task in the middle of the computer screen, such as having to identify one of two cars which appeared briefly in the middle of the screen. At the same time, they had to indicate where a Route 66 road sign appeared on the periphery of the computer screen.

Initially the road sign on the periphery was presented on its own, but over time, distractors appeared. As the participant progressed through the training on this task, the road sign moved from the centre of the screen to farther out on the edges. In addition, the central cars became more similar in appearance and the backgrounds grew more complicated, making it harder to spot subtle differences in the cars and road signs.

Targeting processing speed

Visual sweeps. The participant watched two spatial frequency sweeps of movements of bars that made up a shape on a computer screen. They had to determine the direction of sweep of the bars, for example, whether the bars had swept inward or outward. There were four different sweep orientations: vertical, horizontal, and diagonally in two directions. The bars in the sweeps became thicker and thinner in order to ensure the brain responded equally well to all sweeps. As the participant progressed through the training, the sweeps gradually sped up.

Hawk Eye. The participant was tasked with locating a bird that was different from a group of other birds. They were all presented around the periphery of the computer screen. The birds only appeared on screen briefly, and although the target bird was initially distinct from the other birds, the birds appeared reasonably close together and were presented against a simple background. As participants progressed through the training, the pairs of birds became more similar, challenging the participant to make finer distinctions. The birds were also spread farther apart, which pushed the participant to see details farther from the centre of visual attention, and the background became more complex. As the participant became able to process the information accurately at faster speeds, the birds appeared on the screen for less and less time.

Auditory training.

Targeting processing speed.

Sound sweeps. The participant had to listen either to frequency sweeps in sounds that began at a low frequency and then rose or vice versa. For example, if the sound went up in frequency, it would be heard as a ‘weep’, or down in frequency as a ‘whoop’. As the participant progressed through the levels of the training, the sweeps changed in frequency, meaning some were presented at higher frequencies and some lower. Furthermore, there were also changes in the inter-stimulus interval to attempt to have the participant process sounds separately, even when they were presented very close together. This shortening of the interval necessitated the participant to process the sound information faster. This additionally assisted with the phenomenon of ’backward masking’ that occurs when a second sound comes on the heels of the first sound.

Fine tuning. In this training exercise, participants were asked to choose between two auditorily presented syllables that sounded alike, such as ‘*bo’* and ‘*do’*. These sounds were selected as they represent ‘confusable pairs’ in speech. The syllables were computer synthesised in order to allow for maximum manipulation. The brain needed to be able to accurately perceive and to represent them clearly. As the participants moved through the training, the consonant sounds were gradually sped up, necessitating faster processing speeds.

Targeting processing speed/inhibition.

Mixed signals. The challenge for the participant was to focus on certain auditorily presented target information and ignore competing information. This involved listening to a specific number, letter, colour, or other piece of information whilst looking at a set of numbers, symbols, letters, words, or other information. If what the participant heard matched what they saw in a specific way, they responded. If not, they had to resist the temptation to respond, thus demonstrating inhibition control. For example, they might see the letter sequence 'aaiaa' on screen and be told to respond if what they heard matched the middle letter, in this case the letter 'i'. If they heard a different letter, 'a' for example, they were not to respond. The information had to be processed quickly and judgements made in real time as to whether or not to respond.

Targeting processing speed/Working memory.

Auditory Ace. The participant was presented with auditory information about playing cards. Initially, the information was presented one card at a time on the computer screen, then hidden. The participant had to decide if the spoken information about the new card matched the auditory information about the previous one and respond positively if it did and negatively if it did not. The participant was only given a limited amount of time to respond. As the participant progressed through the training, the time allowed to respond was reduced. As the participant progressed further, an element of matching to cards seen earlier in the sequence introduced an element of working memory in addition to the processing speed and inhibition control training.

Cognitive workbook.

A cognitive workbook was used for an active control group to give participants mental exercises to perform for the same amount of time that the computer-based training groups were training. These involved tasks like numerical tracking exercises, identifying shapes within shapes, general word puzzles, 'dingbats', and word searches. They were asked to carry out their training for 6 thirty-minute sessions over 2 weeks. Participants handed in the completed exercise workbook at the end of the training.

3.3 Results

All of the 48 participants recruited to the study attended an initial appointment and a follow-up visit. They completed all of the assessment tasks at both visits and fulfilled their training commitment as dictated by their allotted training pathway. Training compliance was monitored either by the completion of a workbook (the cognitive training group) or by monitoring the on-line training logs available to the researcher on the BrainHQ website.

3.3.1. Descriptive statistics

The general demographic information for the participants in the three training pathways is recorded in Table 1 in Section 3.2.2.

The average age of the group undertaking vision-based inhibition and information processing speed training was slightly younger and had better pure tone audiogram thresholds than the other two groups. ANOVA testing for age and PTA threshold, however, demonstrated no significant difference, $p > .05$.

Below in Figure 2 are the average hearing thresholds for each frequency (with standard deviations) for the participants who took part in the research for this chapter. No participant had a threshold at any frequency greater than 40 dB HL, and participants had to have an average loss (when averaged across 500, 1000, 2000 and 4000 Hz) no greater than 30 dB HL.

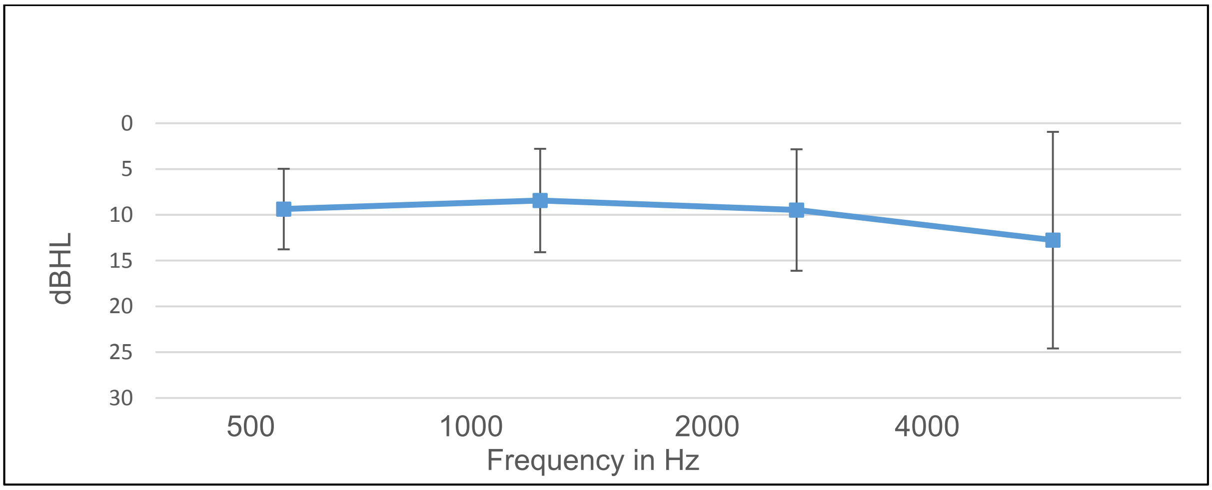


Figure 2. Average pure tone thresholds at each frequency with standard deviations.

A series of ANOVAs was performed to determine whether there were any significant differences between the average thresholds of the three training pathway groups for the individual frequencies recorded for the participants. No significant differences were found. As for Chapter 2, when using hierarchical regression modelling, hearing level was partialled out when exploring for relationships between speech-in-noise discrimination performance and different cognitive functions.

In addition to the ANOVAs performed on the PTA frequencies, further ANOVAs were performed to determine whether there were any significant differences at baseline in performance of the different cognitive tasks between the training groups. None were demonstrated, all at $p > .05$.The mean scores for the measurements at baseline (Assessment 1) and at the subsequent visit (Assessment 2) are given below in Table 2. For some of the tasks, a lower recorded value is indicative of superior performance, such as for the QuickSIN speech test. In the case of QuickSIN, the lower the signal-to-noise ratio task score, the better the performance is deemed to be. The same interpretation, where the lower the recorded value for undertaking the task, the better the performance, also applies to the UfoV, Stroop and RNG tasks. Performance is measured in length of time for the two former tasks and as a score for the different RNG tasks. Table 2 below gives the full set of recorded results.

Table 2

Performance on the different assessments at baseline and after training for the different training groups.

Training group	**Type of Assessment**	**Assessment 1 (baseline)**	**Standard Deviation**	**Assessment 2**	**Standard Deviation**	**Change**
Auditory group	Age	51.69	22.41			
	PTA	10.97	5.78			
	QuickSIN	1.80	1.2	-.021	1.30	-1.811
	BKB	30.58	12.46	33.62	13.26	3.04
	UfoV processing speed	23.59	11.15	18.99	4.68	-4.6
	UfoV divided attention	67.75	106.75	52.76	62.29	-14.99
	UfoV selective attention	152.78 (129.63)*	131.00	132.12 (107.59)*	121.36	-20.66 (-22.04)*
	RSPAN absolute	30.06	17.40	35.75	16.74	5.69
	Compressed speech	45.73	11.43	54.65	13.19	8.92
	Stroop interference	97.08	54.51	117.55	70.21	20.47
	Stroop normalised	1.149	.082	1.182	.10	0.033
	Stroop response latency	671.50	104.23	642.07	107.74	-29.43
	RNG R	1.977	.986	1.850	1.31	-.127
	RNG RNG	.2292	.035	.2395	.053	.0103
	RNG Adjacency	22.13	7.41	18.78	8.48	-3.35
	Auditory selective attention	80.61	11.22	83.10	12.42	2.49
Vision group	Age	45.75	23.54			
	PTA	8.60	5.57			
	QuickSIN	1.31	.69	-.203	.812	-1.513
	BKB	29.33	12.68	33.91	10.76	4.58
	UfoV processing speed	17.53	1.92	17.95	5.00	0.42
	UfoV divided attention	38.53	38.06	19.82	8.36	-18.71
	UfoV selective attention	92.29 (92.29)*	50.98 (50.98)	61.33 (61.33)*	43.74 (50.98)	-30.96 (-30.96)*
	RSPAN absolute	33.44	18.363	39.19	20.9	5.75
	Compressed speech	54.85	13.78	59.25	9.54	4.4
	Stroop Interference	75.8	56.20	70.96	40.51	4.84
	Stroop normalised	1.107	.064	1.108	.067	.001
	Stroop response latency	680.60	107.84	662.19	80.67	-18.41
	RNG R	1.601	1.04	1.554	1.22	-.047
	RNG RNG	.2314	.045	.2427	.055	.0113
	RNG Adjacency	19.82	7.38	20.24	8.31	0.42
	Auditory selective attention	85.67	12.29	87.26	10.52	1.59
Cognitive workbook group	Age	51.56	22.66			
	PTA	10.97	5.49			
	QuickSIN	1.34	.66	.323	.92	-1.017
	BKB	33.96	7.18	36.84	7.89	2.88
	UfoV processing speed	25.88	29.31	29.98	39.67	4.1
	UfoV divided attention	70.65	101.30	45.41	62.35	-25.24
	UfoV selective attention	134.19 (109.81)*	130.80	104.84 (91.82)*	71.26	-29.35 (-17.99)*
	RSPAN absolute	31.50	17.43	27.94	13.43	-3.56
	Compressed speech	47.38	12.63	57.86	12.08	10.48
	Stroop interference	63.17	57.92	71.2	50.50	8.03
	Stroop normalised	1.090	.083	1.099	.07	.009
	Stroop response latency	731.82	104.23	731.44	107.74	-0.38
	RNG R	2.210	1.22	2.132	1.49	-.078
	RNG RNG	.2438	.058	.2812	.046	.0374
	RNG Adjacency	17.73	6.88	15.23	9.39	-2.5
	Auditory selective attention	81.63	13.91	82.13	10.68	0.5

Key: QuickSIN: Quick Speech-in-Noise Test; the lower the score the better the listening performance. BKB: Boothroyd, Kowal and Bench speech-in-noise test at -8 dB SNR. UfoV: Useful Field of View test; for each condition, the lower the time in msec, the better the processing speed performance. RNG: Random number generation; R score examines number repetitions, lower scores indicate better updating ability, and these scores examine the frequency of number sequences (a lower score indicates a better ability to inhibit patterns) and adjacency (gives a score for number sequences with a low score exhibiting a better ability to inhibit learnt patterns). *The numbers in brackets for the UfoV selective attention task denote analysis with Participants 14 and 25 removed, as they recorded no result and were assigned a processing time of 500 msec by the UfoV program.

Analysis plan

As four of the independent variables—the RNG (RNG measure), auditory selective attention task, UfoV selective attention task and Stroop information processing speed measures—were all found to have non-normal distributions on Shapiro-Wilk testing ($p < .05$) and on visual inspection, they also demonstrated non-normal Q-Q plots. This meant that for the correlational part of the analysis that follows later in this results section, a Spearman's rho test was used. For within/between group comparisons, ANOVA testing was used, as it has proven to be a robust statistical test that is reasonably resistant to deviations from normality (Blanca, Alarcón, Arnau, Bono, & Bendayan, 2017; Keselman, Lix, & Keselman, 1996).

The ANOVA testing was performed to explore hypotheses 3.1 and 3.2. These hypotheses posited that there would be significant effects of training group and time following a training period, and that this may be seen both within and between training groups. For H3.1, the difference would be between the speech discrimination tasks, and for H3.2, this would be for the cognitive tasks. For H3.2, it was hypothesised that the auditorily trained group will perform auditory tasks better after training than will visually or control-trained participants. The converse would be true for visually trained participants on visual tasks.

Correlational analysis and hierarchical regression were performed with the baseline data in order to investigate Hypotheses 3.3, 3.4 and 3.5, exploring which cognitive functions will correlate better with a speech-in-babble task rather than a task with consistently challenging speech in speech-shaped-noise at a particularly adverse SNR. As there were 13 variables in the correlation matrix a Holm-Bonferroni correction was applied and this is explained further in Section 3.3.3. A hierarchical regression model was developed to test hypothesis 3.6 based on previous research and with evidence from the correlational matrix. This is explained further in Section 3.3.4.

3.3.2. Two-way mixed-methods ANOVA

As no significant difference was found at baseline for the two speech tasks (or the cognitive tasks) as reported in Section 3.3.1, adding a covariate of starting level performance was not necessary. This had been a concern, as previous work has demonstrated that if starting at a lower performance level for speech testing, greater improved performance may be demonstrated (Sabes & Sweetow, 2007). Levene's test for homogeneity of variance for the pre-group scores was also performed and found to be normal.

ANOVAs were performed for each measure of interest. There were two factors, time and group. Time had two levels and group had three. This was to determine whether the training that was undertaken by each of the three groups brought about any significant changes in each of the dependent variables assessed.

3.3.2.1. QuickSIN performance before and after training

For the two speech tests reported below, Hypothesis 3.1 suggested that cognitive training on processing speed and inhibition control would give an advantage compared to general cognitive exercises (control group).

For the QuickSIN measures, Box's test indicated equality of homogeneity of covariances (*p* = .285). For the mixed methods analysis of variance, the results were as follows:

- There was no overall significant effect of pathway group, *F*(2, 45) = .719, *p* = .493.
- There was, however, a significant effect of time *(F*(1, 45) = 128.02, *p* < .001). Post hoc analysis of within-group performance of the QuickSIN between baseline and follow-up demonstrated statistically significant improvements in performance for the auditory trained group (*t*(15) = 17.49, *p* < .001), the vision trained group (*t*(15) = 7.36, *p* < .001) and the control group (*t*(15) = 4.72, *p* < .001). All training groups, therefore, demonstrated a significant improvement in QuickSIN after training.
- There was also a statistically significant interaction of training pathway*time (*F*(2, 45) = 3.28, *p* = .047). The plot in Figure 3 demonstrates the improvements in

speech-in-noise discrimination performance on the QuickSIN for all groups, with the auditorily trained group demonstrating the most significant improvement.

Although there was a significant interaction of pathway*time, partially supporting Hypothesis 3.1, on follow up post hoc analysis, the between-group QuickSIN performance after the training period did not indicate one training group performing significantly better than another.

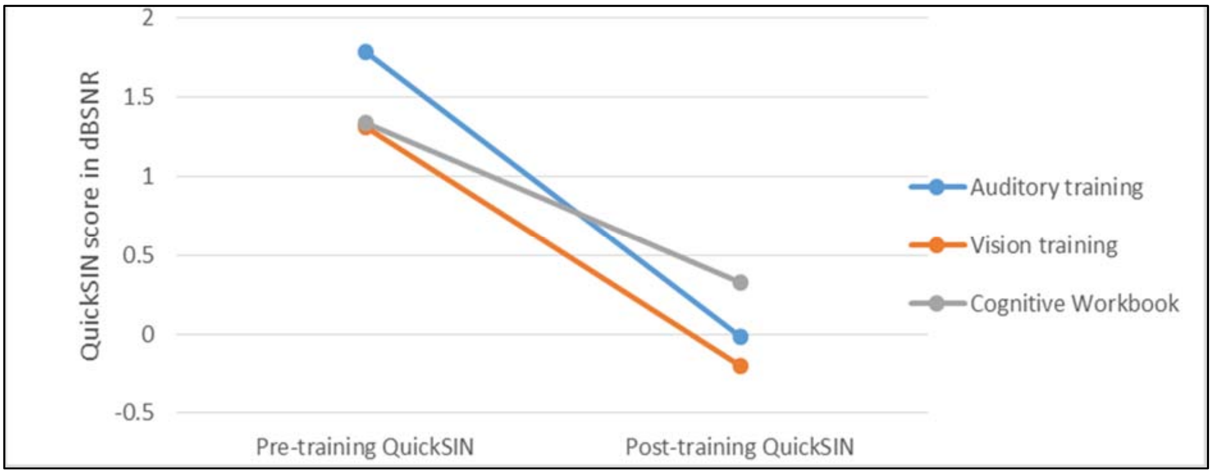


Figure 3. QuickSIN scores before and after the training pathways.

3.3.2.2. BKB performance before and after training

For the BKB measures, Box's test indicated equality of homogeneity of covariances (*p* = .22). For the mixed methods analysis of variance, the results were as follows:

- There was no overall significant effect of pathway group (*F*(2, 45) = .649, *p* = .527).
- There was, however, a significant effect of time *(F*(1, 45) = 9.02, *p* < .005). The post-hoc analysis of within-group performance of the BKB between baseline and follow-up did not, however, demonstrate a statistically significant increase in performance for the auditory trained group (t(15) = -1.78, *p* = .095), the vision trained group (t(15) = -1.99, *p* = .065) or the control group (t(15) = -1.44, *p* =.172).
- There was no statistically significant interaction of training pathway*time (*F*(2, 45) = .217, *p* = .806).

Hypothesis 3.1 was, therefore, not supported for BKB performance either, as there was no significant difference between the groups after training. However, for within-group performance, although none of the training groups demonstrated a significant improvement over time, the two cognitively trained groups did demonstrate a greater improvement than the active control group. This is shown below in Figure 4.

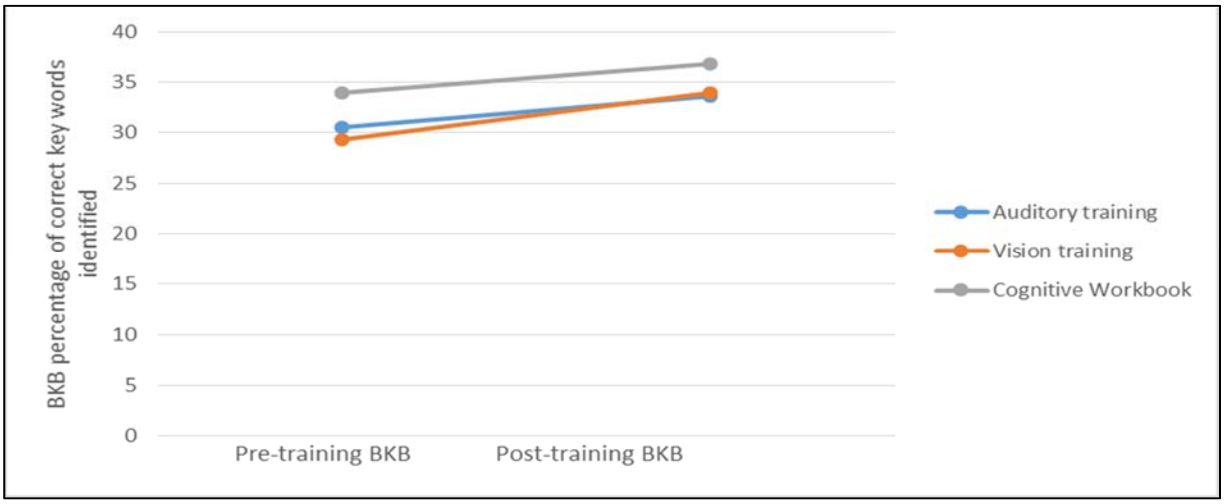


Figure 4. BKB scores before and after the training pathways.

3.3.2.3. Real-time inhibition control performance before and after training.

The tasks used to assess the cognitive functions of on-going inhibition control for auditory and visual tasks were the auditory selective attention and the UfoV selective attention task, respectively.

For the UfoV selective attention task, a mixed methods analysis of variance provided the following results:

- There was no significant effect of pathway group ($F(2, 45) = .869$, $p = .426$).
- There was a significant effect of time ($F(1, 45) = 10.435$, $p = .002$). As there was a significant effect, within-group post-hoc analysis was performed. This analysis demonstrated no statistically significant increase in performance between baseline and follow-up for the auditory-trained group ($t(15) = 1.20$, $p < .066$) or the control group ($t(15) = 1.10$, $p = .291$), but a significant effect for the vision trained group ($t(15) = 2.48$, $p < .05$).
- There was no significant interaction of training pathway*time ($F(2, 45) = .147$, $p = .864$).

Thus, for the UfoV selective attention task, Hypothesis 3.2 could be supported, as the only group that demonstrated a significant within-group improvement after training was the visually trained group.

For the auditory selective attention task, the results were as follows:

- There was no significant effect of pathway group (*F*(2, 45) = .1.980, *p* = .150).
- There was no significant effect of time (*F*(1, 45) = 2.310, *p* = .136).
- There was no significant interaction of training pathway*time on the task (*F*(2, 45) = .331, *p* = .720).

The performances of the three training groups on these tasks are shown below in Figures 5 and 6.

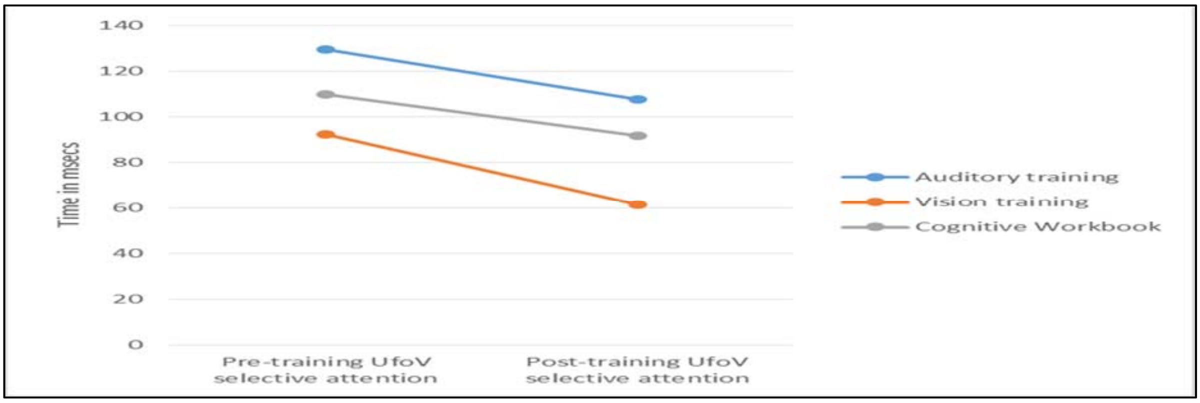


Figure 5. UfoV selective attention scores before and after the training pathways for the different training groups.

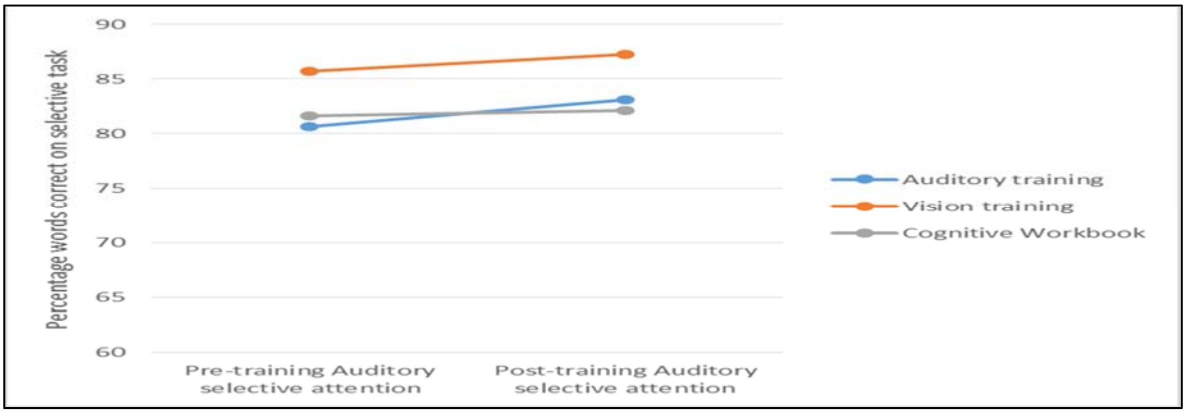


Figure 6. Auditory selective attention scores before and after the training pathways for the different training groups.

3.3.2.4. Auditory information processing speed performance before and after training

For the auditory processing speed task, the identification of 50% time-compressed sentences, there was an increase in performance after each group's respective training. For the mixed methods analysis of variance, the following results were recorded and are displayed in Figure 7 below:

- There was no significant effect of pathway group (*F*(2, 45) = 1.595, *p* = .214).
- There was a significant effect of time (*F*(1, 45) = 27.991, *p* < .001). As there was a significant effect, within-group post-hoc analysis was performed. This analysis demonstrated a statistically significant increase in performance between baseline and follow-up for the auditory trained group (*t*(15) = 10.835, *p* < .005) and the control group (*t*(15) = 18.76, *p* < .005) but not for the vision trained group (*t*(15) = 2.76, *p* < .118).
- There was no significant interaction of training pathway*time (*F*(2, 45) = 1.474, *p* = .240).

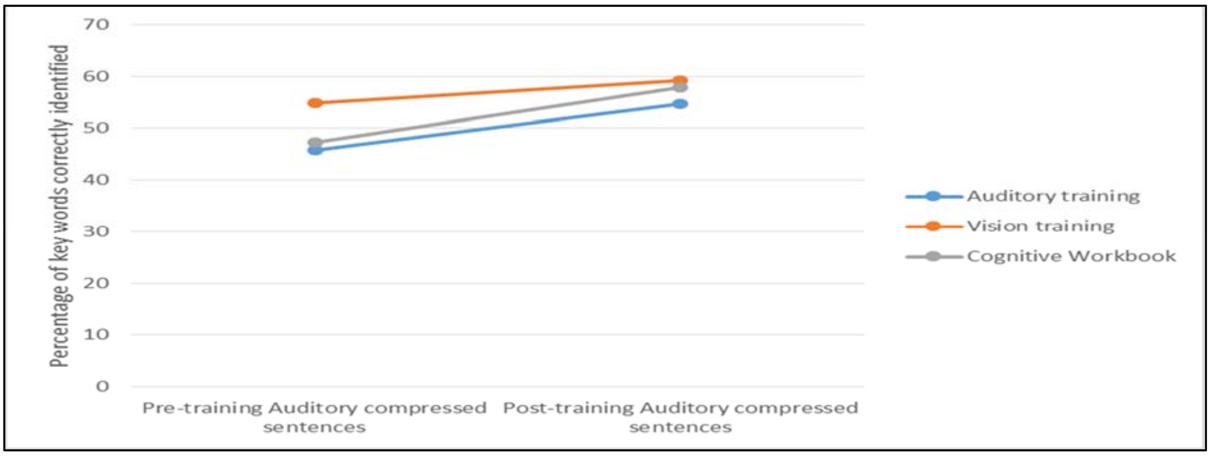


Figure 7. Auditory time compressed speech scores before and after the training pathways.

Of the three training pathways, it was the cognitive workbook group that showed the greatest change within their group after training, with the auditory trained group also demonstrating a significant improvement. Hypothesis 3.2 could not be supported, as although for an auditory-based task, the auditory-trained group demonstrated a significant improvement over time, the control group also demonstrated a significant improvement, which does not, therefore support the hypothesis.

3.3.2.5. Visual prepotent inhibition before and after training

The results for the Stroop interference test are illustrated below in Figure 7. Mixed methods analysis yielded the following results:

- There was no significant effect of pathway group (*F*(2, 45) = 3.020, *p* = .059).
- There was no significant effect of time (*F*(1, 45) = 1.186, *p* = .282).
- There was no significant effect of training pathway*time (*F*(2, 45) = 1.017, *p* = .370).

For the Stroop interference measure, the workbook group and the auditorily trained group demonstrated poorer performance after training, and only the visually trained group demonstrated an improvement. As the interaction was not significant, this does not support Hypothesis 3.2.

Figure 8 below illustrates that the visually trained online exercise group did show an improvement following training; however, there was no significant interaction between time and pathway group.

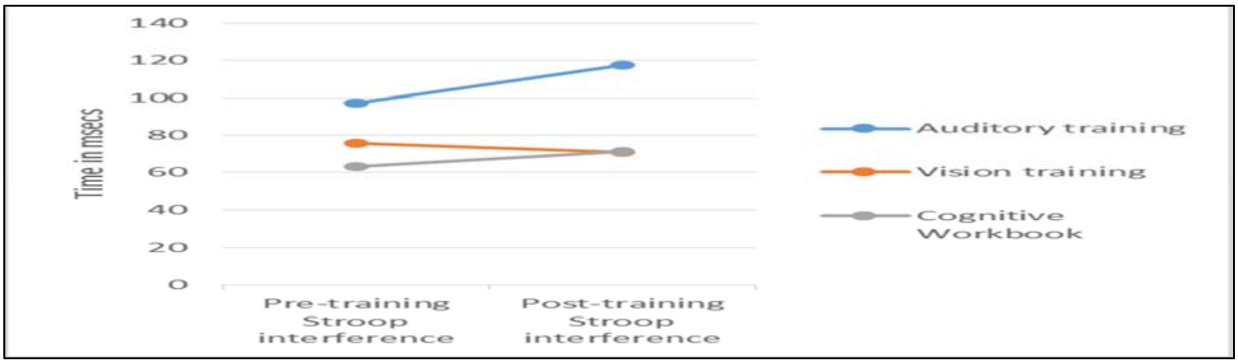


Figure 8. Stroop interference measure before and after the training pathways.

3.3.2.6. Response latency performance (using the congruent condition of the Stroop) before and after training

For the Stroop latency task, the following results were obtained from the mixed methods analysis of variance.

- There was no significant effect of pathway group (*F*(2, 45) = 2.470, p = .096).
- There was no significant effect of time (*F*(1, 45) = 3.196, *p* = .081).
- There was no significant interaction of training pathway*time interaction (*F*(2, 45) = .886, *p* = .419).

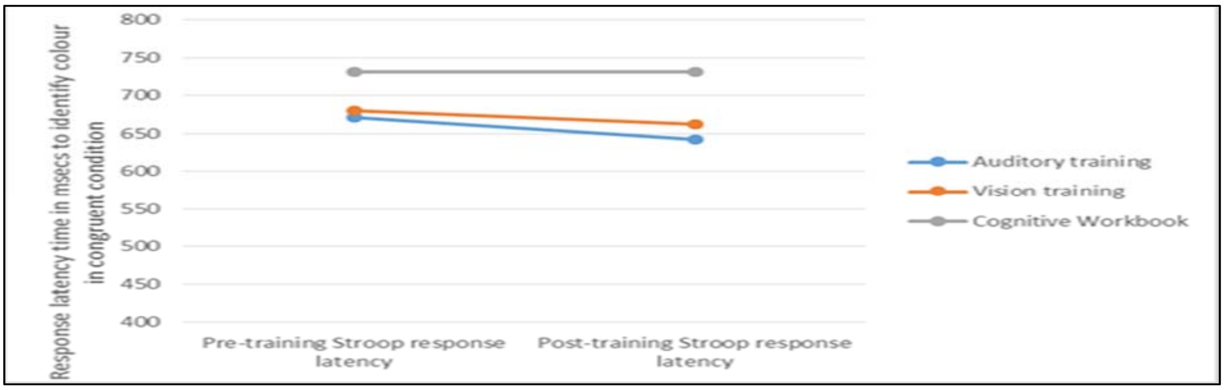


Figure 9. Stroop response latency measures before and after the training pathways.

Figure 9 illustrates that the auditorily and visually trained online computer exercise groups did show improvement following training on the response latency task, as the time taken to respond got shorter. As the improvement was not significant, however, Hypothesis 3.2 cannot be supported for the response latency task.

3.3.3. Correlation analysis (excluding UfoV processing speed and divided attention)

The initial analysis explored potential significant relationships between participants' performance on both of the speech-in-noise tests, demographic information and cognitive function performance data. The relationships were explored using Spearman's rho, as four of the variables were found to be not normally distributed, as reported above. Due to the number of comparisons being made (13), the same Holm-Bonferroni (Holm, 1979) correlation procedure was used as in Chapter 2, with the difference that the critical *p* value changed from alpha/13 to alpha/1 following the step-wise procedure. Alpha was set at .05. The criterion for significance, therefore, slightly relaxes during the steps. For example, the first level of significance is .05/13 = .0038. The second is .05/12 = .0041, and so on.

Table 3:

Correlation matrix (N = 48) with Spearman's rank correlation coefficients

	Age	PTA	Quick SIN	BKB	UfoV selective attention	R span absolute	Comp speech	Stroop IM	Stroop speed	RNG R	RNG RNG	RNG A	Auditory selective attention
Age	-												
Pure tone average	.754*	-											
Quick SIN	.488*	.566*	-										
BKB	.099	-.052	-.186	-									
UfoV selective attention	.675	.563*	.306	-.053	-								
R span absolute	-.151	-.051	-.206	.007	.022	-							
Compressed speech	-.579*	-.659*	-.643*	.336	-.372	.051	-						
Stroop IM	.289	.360	.359	-.176	.043	-.347	-.344	-					
Stroop speed	.523*	.406*	.133	-.036	.476**	.061	-.379	-.098	-				
RNG R	.128	.146	.277	-.112	.025	.030	-.239	.133	.011	-			
RNG RNG	.307	.238	.031	.066	.093	-.225	-.200	-.049	.252	.264	-		
RNG A	.013	.025	-.056	-.174	-.022	-.116	.001	.344	.142	-.120	.186	-	
Auditory selective attention	-.512*	-.653*	-.468*	-.025	-.398	.135	.513*	-.166	-.361	-.019	-.296	-.123	-

Key: PTA = Pure audiometry threshold average calculated from average thresholds at 500, 1000, 2000 and 4000 Hz. QuickSIN Quick Speech-in-Noise test. BKB = Bamford-Kowal- Bench hearing-in-noise test score as percentage of correctly identified key words. RSPAN = Reading Span. RNG = Random Number Generation tasks. Stroop IM = Stroop interference measure. UfoV = Useful field of Vision.

Correlations: *Correlation is significant at the 0.05 level (2-tailed); Holm-Bonferroni correction applied

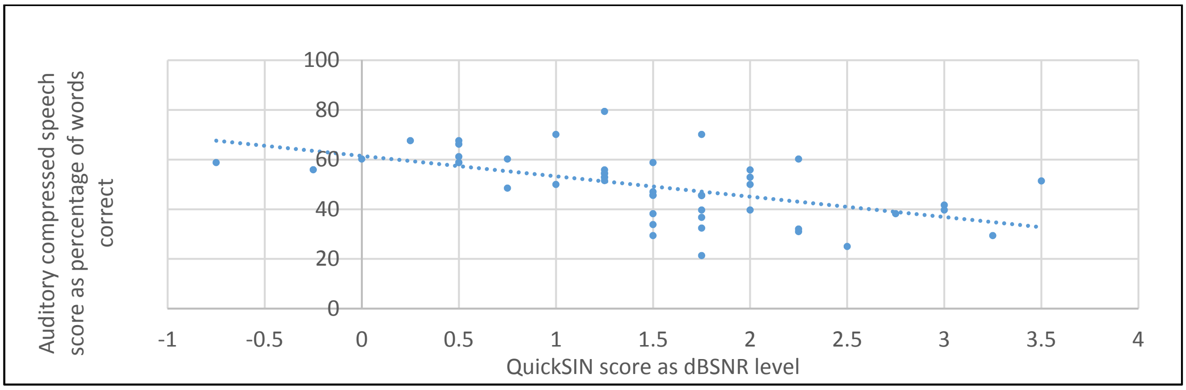


Figure 10. The correlation of QuickSIN performance with auditory compressed speech performance. Spearman's rho = -.643.

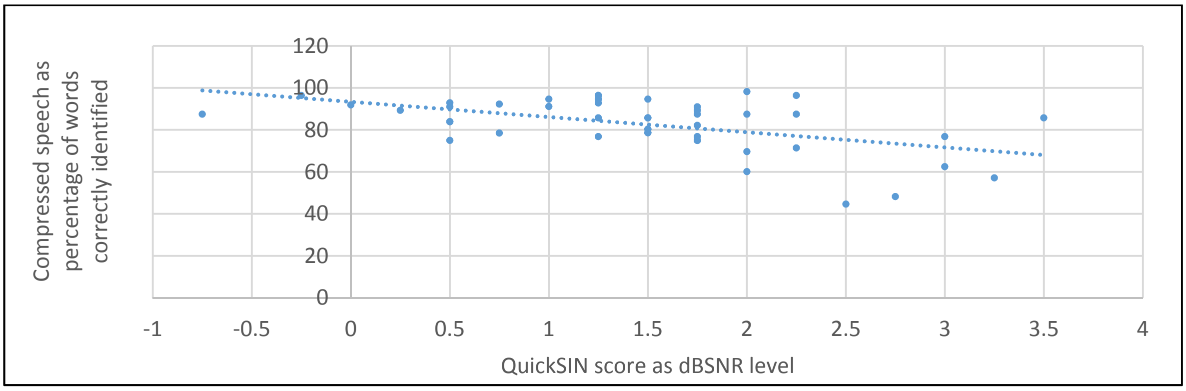


Figure 11. The correlation of QuickSIN performance with auditory selective attention performance. Spearman's rho = -.468.

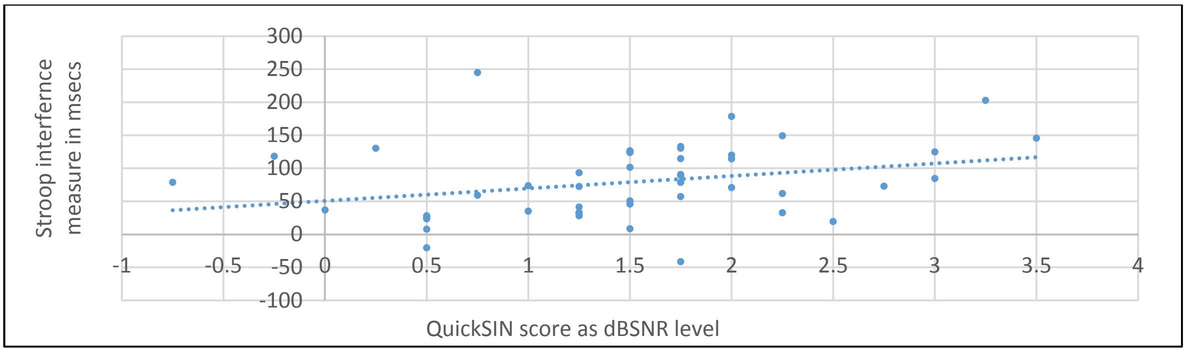


Figure 12. The correlation of QuickSIN performance with the Stroop interference performance. Spearman's rho = .359.

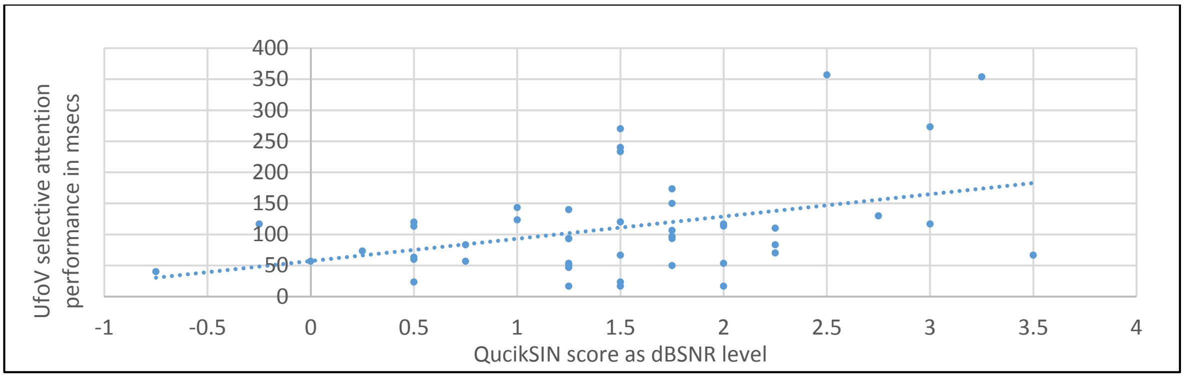


Figure 13. The correlation of QuickSIN performance with the selective useful field of vision performance. Spearman's rho = .306.

The complete set of results above in Table 3 show the significant correlations after the Holm-Bonferroni correction has been applied for each family when exploring for other correlations. The scatterplots above illustrate the significant relationships of the QuickSIN with different variables and cognitive functions (Figures 10, 11, 12 and 13).

Hypothesis H3.3 suggested that glimpses of speech within the masking babble would allow for significantly stronger relationships to be demonstrated between cognitive functions and the QuickSIN test rather than the BKB test. There were no cognitive functions demonstrating a significant relationship with BKB scores. In contrast, for QuickSIN performance, speed of information processing as evidenced by the time compressed speech task (ρ = -.643) demonstrated a significant relationship, as did ‘real-time’ inhibition performance, evidenced by the auditory selective attention task (ρ = -.468). There were no other significant relationships demonstrated after the Holm-Bonferroni correction had been applied.

Hypothesis 3.3 was therefore supported, as the results demonstrate that the auditory based cognitive functions of information processing speed and control of inhibition demonstrated a significant relationship with performance on the QuickSIN (in background babble), but not performance on the BKB in speech-shaped noise.

Hypothesis 3.4 theorised that auditorily performed cognitive tasks would demonstrate significant relationships with speech-in-noise perception task performance whereas visually performed tasks would not. This hypothesis was partially supported in that only the auditory time compressed speech and auditory selective attention tasks demonstrated significant relationships with speech-in-noise perception after the Holm-Bonferroni correction had been applied, but this was only for the QuickSIN speech-in-noise test and not for the BKB. The visual tasks, although tapping the same functions of selective attention (UfoV) and inhibition control (Stroop), were not significant after the application of the correction.

Hypothesis 3.5 was also partially supported, as after the application of the Holm-Bonferroni correction, it was only the task that represented ‘real-time’ inhibition control, namely the auditory selective attention task that demonstrated a significant correlation with QuickSIN speech-in-noise performance. None of the tasks representing prepotent inhibition demonstrated a significant relationship. Again, though, this was only for the QuickSIN and not for the BKB.

The performance of the three RNG tasks, two of which index prepotent control, demonstrated no significant relationships with performance on either of the speech-in-noise tasks.

3.3.3.1. Correlation analysis (including UfoV processing speed and divided attention)

The addition of the UfoV processing speed and divided attentions tasks the correlation matrix did not contribute to the support of Hypotheses 3.4 or 3.5, as neither the UfoV processing speed nor divided attention correlated significantly with the QuickSIN or the BKB, although relationships were stronger with the QuickSIN. As the processing speed and divided attention UfoV did not demonstrate any significant relationships, the correlation matrix involving them has been excluded from this results section.

3.3.4. Hierarchical regression modelling

The same rationale for constructing the hierarchical regression model was used as was employed in Chapter 2 (Field, 2016, p. 512). The model was created using baseline data. The QuickSIN hierarchical regression model was created to address Hypothesis 3.6, which stated that after controlling for age and hearing level, cognitive functions would add to the explanation of speech-in-noise performance variation. Only the results from 46 participants were used to build the regression model as participants 14 and 25 had recorded no result (time) on the UfoV selective attention task.

The model was designed to partial out age and hearing level (Model 1), working memory as assessed by RSPAN performance (Model 2), information processing speed as measured by performance on an auditory compressed speech test (Model 3) and finally control of inhibition, both prepotent and ongoing, as measured by performance on Stroop and Auditory and Vision selective attention tasks, respectively (Model 4). These cognitive functions have all previously demonstrated relationships with speech-in-noise discrimination performance (as reported in the hierarchical regression in Section 2.3.2.). Additionally, these factors had all demonstrated a significant relationship with QuickSIN performance in the correlational analysis, either with or without the Holm-Bonferroni correction. A hierarchical regression model was not created for the BKB speech test as there were no significant correlations of BKB performance with cognitive functioning measured by the different tasks used in the research or with any demographic factors.

As can be seen from Table 4, the addition of processing speed did add significantly to the explanation of variance, ($p < .05$). Once processing speed had also been factored out, the addition of the inhibition measures did not add significantly to the explanation of performance variance. As none of the RNG measures correlated significantly with QuickSIN performance, they were excluded from the hierarchical regression.

Table 4

Summary of models used in hierarchical regression and proportion of variance of QuickSIN performance accounted for by each model

Model	**Predictors**	**R^2**	**R^2 change**	**Significance F change (*p* value)**
Model 1	Variation due to PTA and Age	.283	.283	.001
Model 2	Variation due to working memory (RSPAN)	.316	.034	.158
Model 3	Variation due to auditory compressed speech measure (speed of information processing)	.412	.095	.014
Model 4	Variation due to Auditory and Visual Selective Attention and Stroop interference measures (control of inhibition)	.444	.032	.542

Key: PTA: pure tone average hearing threshold

As auditory processing speed added significantly to the explanation of the variance in QuickSIN performance, Hypothesis 3.6 was partially supported in that after age and hearing had been partialled out, it was the only cognitive function that added extra support as demonstrated in Table 5 below.

Table 5

Hierarchical regression predicting QuickSIN test scores from PTA and age (model 1), working memory performance (model 2), compressed speech performance (model 3) and from control of inhibition (model 4).

	b	*SE b*	*ß*	*R²*	*R²* change	Significance of *F* change
Model 1				.283	.283	.001
PTA	.062	.030	**.391***			
Age	.007	.007	.176			
Model 2				.316	.034	.158
PTA	.066	.029	.416			
Age	.006	.007	.156			
Working memory (RSPAN)	-.009	.007	-.184			
Model 3				.412	.095	.014
PTA Scores	.040	.029	.251			
Age	.002	.007	.051			
Working memory (RSPAN)	-.011	.006	-.221			
Compressed speech	-.028	.011	**-.401***			
Model 4				.444	.032	.542
PTA scores	.008	.037	.053			
Age	.003	.008	.078			
Working memory (RSPAN)	-.010	.007	-.204			
Compressed speech	-.025	.011	**-.358***			
Auditory selective attention	-.018	.013	.260			
Visual selective attention	.000	.002	-.006			
Stroop interference measure	.001	.002	.66			

Note: * $p < .05$.

3.4. Discussion

The main focus of this chapter was to identify whether training in the cognitive functions of processing speed and control of inhibition could improve speech-in-noise discrimination performance. There was no clear evidence from the research undertaken that this was the case. In general, undertaking any of the types of training on offer gave individuals an advantage, as task performances improved over the training period across training groups. The changes, therefore, may be more to do with practice effects rather than any 'real' improvement (McCabe, Langer, Borod, & Bender, 2011). This would be in line with some of the contentions surrounding cognitive training in general, for which it has been argued that there is no convincing evidence that cognitive training (utilising working memory) leads to improved fluid intelligence (Melby-Lervåg & Hulme, 2013; Redick et al., 2013). Indeed, in the present research, there is no clear evidence of 'on-task' improvements beyond gains due to simple practice effects, as for a number of the cognitive functions either all the training pathways demonstrated significant improvements or none did. The one exception would be the UfoV selective attention task at which only the vision trained cognitive gave evidence of significantly improved performance after training, the other training groups did not. The only auditory based cognitive task that demonstrated a significant improvement for the auditory trained group and not the vision group was on auditory compressed speech (an information processing speed function), however, the control group also improved significantly which lends no support to H3.2.

As there was limited evidence of any of the cognitive functions demonstrating significant improvement after training, it was unsurprising that there was also no improvement noted in either of the speech-in-noise perception tasks. This would fit with previous research in this area (Brehmer et al., 2012; Ferguson & Henshaw, 2015c), albeit both Brehmer et al. and Ferguson and Henshaw employed working memory training for their research.

The only significant interaction demonstrated for speech-in-noise improvement following training was on the QuickSIN. This gave evidence of an interaction between time and pathway group, potentially offering some support to Hypothesis 3.1, but as it was a single result in isolation and only for the QuickSIN, not for the BKB task, there is the possibility that this was simply due to chance, so concern over this being a type 1 error needs to be considered. Further, on post-hoc testing, although the computer-trained groups did perform better than the cognitive workbook group, there was no significant difference exhibited in the end point performance on the QuickSIN between any of the training pathways. The fact, however, that there was a significant interaction of pathway and time for the QuickSIN but not for the BKB perhaps gives evidence that the cognitive functions assessed in this chapter are more useful for an individual when glimpses of speech can be more readily perceived as in the QuickSIN.

As an overall summary, it appears that cognitive training did not clearly demonstrate any clear 'on-task' improvements and there was no clear indication of any transfer to a speech-in-noise task. Indeed in general, cognitive functions do seem to have a minor role in supporting speech-in-noise perception (this thesis, Chapter 2), whereas audition is a much more substantial contributor, especially when any degree of hearing loss is present (Humes & Dunbo, 2010). In the hierarchical regression model for this chapter, it was only speed of information processing (the compressed speech task) that individually added significantly to the explanation of performance variation in speech-in-noise discrimination (after audition and age had been accounted for in model 1). It would, therefore, be unlikely that the training of cognitive functions that did (information processing speed) and did not (control of inhibition) contribute significantly to any of the performance variation would be likely to demonstrate a significant improvement of performance of speech-in-noise tasks.

However, it seems reasonable to investigate whether a tailored cognitive training regime would offer greater success. It could be suggested that it is only individuals who have weak cognitive functioning or who have a great reliance on certain cognitive functions who may benefit from training. Identifying the need for training and designing any program is a concept that is put forward in this thesis as an area for future research. A potential reason why auditory cognitive training may be preferred to vision based training tasks is that auditory training may offer added strengthening of auditory neural pathways (and speech representations) as well attempting to improve the cognitive function itself. These neural pathways will in turn become more 'hard-wired' with training. When presented with any tasks that involve making judgements on auditory inputs, these enhanced representations/pathways may allow for quicker and better matches, allowing for superior task performance. For visually trained individuals, although the cognitive function may be improved following training, there is no accompanying improved access to the audiological areas.

With the evidence observed here for the fuller Stroop task in that it did not offer any explanation of speech-in-noise performance variance, it could be argued that using the screening Stroop in the earlier chapter was more akin to assessing processing speed, which is an often put forward argument about what the Stroop test assesses (Uttl & Graf, 1997; Verhaeghen & De Meersman, 1998). The screening Stroop, therefore, demonstrated a relationship because the participants had to adapt to the task quickly, with the ones being able to do so achieving the better scores. For this reason, the short screening Stroop demonstrated a relationship with BKB performance in Chapter 2 when the fuller Stroop task did not in this chapter.

Additionally, as another example of difference in performance on the speech tasks, although participants in this chapter (age 18–80 years) had no self-reported hearing difficulties and their overall hearing thresholds were within the accepted inclusion criteria for this chapter, there were significant correlations of both PTA and age with QuickSIN performance. As older adults, even those with reasonable hearing, still have problems hearing speech-in-noise compared to younger adults with similar hearing (Pichora-Fuller et al., 1995), and where there is any degree of hearing loss the hearing loss will become a main contributor to how well speech is heard in noise (Humes & Dunbo, 2010), this finding could have been expected. Conversely, however, there was no significant relationship observed between PTA or age and BKB performance. It may be the case, therefore, that the BKB task at -8 dB SNR in speech-shaped noise was so challenging that being younger and having slightly better hearing did not aid discrimination, as all participants struggled equally with the task.

The only correlation that approached significance for the BKB task was that with the auditory compressed speech performance, a measure of information processing speed. The strength of the relationship was $\rho = .336$ ($p < .050$); however, this was before the application of the Holm-Bonferroni correction. There is a risk of this correlation being a false positive, but it is considered here as it was the only cognitive or demographic function that potentially demonstrated a relationship with BKB performance. In Chapter 2 of this thesis, it was also only speech of information processing (performance of the go element of the go/no-go task) that correlated significantly with BKB performance (with Holm-Bonferroni correction applied). This does offer the suggestion that processing speed is a key cognitive function when attempting to discriminate speech in particularly adverse listening conditions. The reason that it was not significant in the research for this chapter could be that the processing speed was an auditory rather than a visual, as was used in Chapter 2. The auditory element, therefore, introduced different confounding elements that possibly prevented the relationship from being demonstrated as significant.

The theory linking speed of information processing and BKB performance is that an efficient ability to process auditory information quickly may allow listeners who only catch very brief glimpses of speech to ‘lock onto’ a part of the speech heard in noise. They are then able to match all or some parts of words, and then utilise ‘top-down’ processing to complete words by searching for matches. Using their lexicon, the listener can then also select or guess at other words that may go with this original word to make grammatical sense, thus using their support mechanisms.

One question beyond the scope of this chapter is which of the age groups responded best to training. The design of the research did not allow for a large enough number of trained participants in each age group to explore for differences in performance both by age grouping and by training pathway. This would be a very useful further piece of research to undertake, as if seeking to develop training that is individually tailored, understanding the age groupings that would respond best is of high importance.

3.5. Limitations

The usefulness of the UfoV speed of processing and divided attention tasks was unfortunately limited by ceiling effects, in that too many participants were able to perform these tasks at the fastest speed at which the program could function. It could have been the case that if the test system had allowed a greater differential in performance by not having a ceiling of performance at 16.7 msec (for both the processing speed and the divided attention task), a significant relationship with the two speech tasks may have been demonstrated as information processing speed tasks have demonstrated significant relationships with speech-in-noise performance in both Chapter 2 and this chapter.

It may also have been advantageous to have participants undertake more than 6 training sessions; however, most of the gains achieved in training are evidenced by improved performance over a short time period (Molloy, Moore, Sohoglu, & Amitay, 2012). The shorter time period also ensured that participants complied with training. In this research, compliance with training was excellent.

Although the training was focussed on processing speed and control of inhibition, due to the design of the exercises intended to keep them interesting, cognitive functions such as working memory and response latency may also have been given some training. Additionally, incidental training of participants to employ ‘top-down’ support may have also been used, strengthening their supporting role for an individual. This may particularly have been the case for older individuals who appear to rely on extra support as they age. Training tasks that tap only the cognitive functions that are to be investigated would allow for better isolation of the involvement of that particular function; however, it was felt that this may reduce training compliance and ‘active’ participation in the training.

In summary, therefore, although the training was aimed at improving inhibition control and processing speed, there may have been some incidental training of other functions, which may have lent some support to task improvements. On balance, though, keeping tasks interesting to ensure training compliance, even though doing so may tap other functions, is preferable to presenting tasks that focus purely on one function but may result in boredom, potentially reducing training compliance. An alternative would be to put some support mechanisms in place (such as group training sessions, which have been found to be successful already in many cases when dealing with hearing loss (Hawkins, 2005)) or to break the training up with other activities that do not directly train any cognitive function.

3.6. Conclusion

Computer training aimed specifically at certain cognitive skills in order to improve speech-in-noise perception did not demonstrate any advantages in terms of improving discrimination. There was also no clear evidence of improvement in cognitive functions after training. There was a significant interaction between the auditory training pathway and QuickSIN performance, but as this was the only significant result, its reliability is somewhat questionable. There was also no clear pattern of significant improvement in 'on-task' performance for the cognitively trained pathways when compared to an active control pathway. As such, it appears that cognitive training did not seem to offer any particular advantage.

There was some very weak evidence that cognitive training is more appropriate when improvements are sought in a particular modality, so for auditory tasks, auditory training potentially gives a greater advantage. The same would appear to hold true for the visual training of cognitive tasks, as 'on-task' improvements are better achievable if the training is in the same modality. The overall lack of success of cognitive training in delivering 'on-task' improvements or in delivering far transfer improvements suggests that global cognitive function training does not offer the desired gains, certainly when it comes to improvement of speech-in-noise discrimination.

This chapter did, however, give further evidence that cognitive functions partly explain some of the variation in speech-in-noise performance. The total explanation of QuickSIN performance in the hierarchical regression models was 44%, which is considerable but does still leave some unexplained variation. There was a range of variation in performance of the different cognitive tasks within groups, so it would appear that different individuals use different coping/support strategies to improve their performance when listening to speech-in-noise. Although this chapter did not give evidence of working memory providing key support for speech-in-noise performance, there is a good body of evidence that suggests it plays a key role (Rudner et al., 2011), alongside other evidence presented in this thesis. As such, it also needs to be considered as a cognitive function that may offer some advantage if it is trained. A better way of identifying where a particular individual's strengths or weaknesses lie with regards to their ability to understand speech-in-noise needs to be established. Exploration of whether a more tailored cognitive training program, after an assessment of an individual's needs, provides more substantial benefit would seem worthy of further research. This may deliver better outcomes, as the training in this chapter was not effective. In the overall conclusions of this thesis, a suggested assessment/training package will be discussed.

CHAPTER FOUR

4. Does auditory training lead to on-task improvements with resulting far transfer gains and a release of cognitive resources?

4.1. Introduction

As discussed in Chapter 1, there are many factors that may play a role in the understanding of speech-in-noise. In the previous chapter, cognitive training was not able to demonstrate any significant improvement in speech-in-noise performance for any of the trained groups over simple practice effects. As audition has been reported as the main predictor of speech-in-noise performance (Akeroyd, 2008; Humes & Dunbo, 2010), then as has been investigated by previous authors, auditory training may offer improvements (Burk & Humes, 2008; Stacey & Summerfield, 2008). Additionally, if auditory training truly leads to enhanced discrimination, then this improvement should extend beyond passive learning from simple exposure to the training material or basic practice effects. It is this deeper level of any advantages of training and how this may be transferred to performance of other, non-trained tasks that forms the focus of this chapter.

As stated in the Introduction, any gains on performance of the trained task itself are considered as being ‘on-task’. Transfer of improved performance to other, similar (or related) tasks is considered ‘near transfer’. Improvement on unrelated or novel tasks is known as ‘far transfer’. In evaluating the potential for near transfer, this chapter sought evidence of improved performance in the discrimination of untrained speech. Far transfer was explored by assessing whether there was any release of cognitive resources following speech discrimination improvement and if this also allowed for information to be better encoded into secondary memory.

In an example of both on-task and near transfer gains following auditory training, Humes et al. (2009) trained a group of 16 older adults with hearing loss, using words that frequently occur in American English. After the training period, 75 to 80% of the older adults were better able to discriminate speech in a background of fluctuating speech noise; this improvement applied both to the speech presenters used during the training and to novel speakers. It was also the case that these improvements were noted on new sentence material. There was, however, no control group of similar older adults with hearing loss recruited for this research, so the effect of simple learning cannot be totally discounted as contributing to any improved performance. The improved performance of novel and different tasks, however, does lend some weight to the improvements being more than just practice effects.

There are still areas of debate surrounding any auditory training regimen. Evidence on the duration of training sessions in general (Humes et al., 2014; Wright & Sabin, 2007) would appear to indicate that doing too much may be counterproductive. The Humes et al. research indicates that training on the order of two to three times a week for 5–15 weeks is sufficient to yield benefits. Further, most gains are achieved early in the training program, with later training sessions adding little advantage (Molloy et al., 2012). Other research involving training a visuomotor skill has demonstrated that allowing time for some consolidation of training rather than training all at once was also advantageous (Goedert & Miller, 2008). In view of the evidence presented above, the training planned for this research involved short daily sessions to allow for training consolidation and to maximise improvements. It was hypothesised that delivering shorter-intensity training sessions and allowing time between sessions would allow for the building of new phonological representations of the degraded speech. If these new representations are well established, they should enable a listener to match words quickly and accurately when exposed to new speech material degraded in the same manner, thus using less cognitive processing resources. This is explained by the rapid automatic multimodal binding of phonology (RAMBPHO) as presented in the ease of language understanding model (Rönnberg et al., 2013).

It is somewhat unclear what percentage of training gains were retained and whether any top-up training was necessary to maintain these gains. Two reports that used LACE training, by Sabes and Sweetow (2007), working with hearing-aided patients, and Song et al. (2012) with non-hearing aid users both reported that participant performance was maintained at post-training sessions. This evaluation was 2 months post-training for Sabes and Sweetow and 6 months for Song et al. Sabes and Sweetow also included a rather anecdotal report that at 6 months there was some decrease in scores but some of the improved performance was maintained. Unfortunately, there was no evidence presented to support this. The positive reporting of retained speech-in-noise performance by Song et al. and the rather ambiguous report from the Sabes and Sweetow research may be due to the age of the participants. Song et al.'s study had recruited participants aged 19–35 years; Sabes and Sweetow, aged 28–85 years. It is very possible to hypothesise that the retention of improved performance was superior for the Song et al. research due to the younger participants in that study having greater plasticity in their neural systems. This plasticity allows for the newly established neural encoding evoked by the training to become more hard-wired, and it could be this that accounts for the difference in the reporting of these two studies on long-term training benefits.

Burk and Humes (2008) carried out a study with a small sample size of 8 participants and demonstrated that even with older adults, training gains could be maintained. They hypothesised that the reasons behind declining performance would be either simply due to time or interference coming from new materials. The research participants undertook an initial 12 weeks of training, then were monitored for a further 14 weeks to determine whether any trained improvements were maintained. During the initial 12-week training on 75 hard words, performance on those words demonstrated an average improvement of 40%. This improvement was maintained when trained on a new set of 75 easy words. It was also reported that when tested in an open set condition, the improved performance for all of the words was maintained for upwards of 3 and a half months.

A potential reason for the difference in the performances reported for this group of older participants and for those recruited for the Sabes and Sweetow research is the training involved. The LACE training employed by Sabes and Sweetow (explained below in Section 4.1.2 and more fully in Chapter 5) is not solely auditory training, as there are some cognitive exercises involved. The Burk and Humes training, however, was based on their well-established auditory training method (Humes et al., 2009), which is purely focussed on auditory training. It also uses hard-to-discriminate words in the training, and there is the perception that in order to demonstrate the greatest gains, training must be a challenge to the individual being trained.

Other reports on training retention that employed a different training regime, have also demonstrated that training gains can be maintained. A study by Burk et al. (2006) reported that at 6 months after training, performance was still significantly improved (62.9% compared to 37.6% at baseline) but had significantly declined compared to immediately post-training (83.5%). Oba et al. (2011) reported that for both 'on-task' and 'generalisable' improvements, performance was retained at 1 month with no decrease compared to immediate post-training scores. The amount of training needed to allow for retention of any training advantages needs quantifying, as does the need for any top-up training to maintain performance. A longitudinal study exploring these factors would help greatly with planning and developing new training packages.

The transferability of any training gains is also intuitively appealing, as it would seem feasible that skills acquired after training for one task may lead to cross-over benefits on other, more loosely associated tasks. It is also attractive from the standpoint that if training gains extend to other functions, the training is a better investment. Transferability of gains is an area of some contention in terms of both broader cognitive training (Redick et al., 2013)[33] and auditory training, with some research reporting benefits (Sweetow & Sabes, 2006a) and others not (Saunders et al., 2016a). The possibility of transfer of gains is explored in the research for this chapter, specifically exploring the likelihood that enhanced speech representations offer the opportunity for cognitive resource release.

It is not clear whether analytic or synthetic training offers the greater advantage for improving speech-in-noise discrimination performance. As described in Section 1.2.6, analytic auditory training, which was utilised by Humes et al. (2009), requires a listener to make judgements

---

[33] This has been explored in further detail in Chapter 3.

and attempts to improve their recognition of speech sounds (syllables/phonemes). Listeners may also have to discriminate between words. Synthetic training is based on completion of sentence-type tasks, asking the individual to listen to a complete sentence and make a judgement about what they heard. Woods et al. (2015) and Davis et al. (2005) suggest that the dominant role of top-down processing in synthetic training allows for the greatest transfer of benefit to novel speech material.

This chapter set out to determine if a short course of auditory training could offer ‘on-task’, near transfer and far transfer advantages for those who undertook the training. The evidence surrounding the advantages of different types of training, the potential for transfer of any gains and the mechanism for any cognitive resource release are considered in the sub-sections below, following an overview of language processing and perception.

4.1.1. Understanding language

Language can be seen as having three constituent parts that are interlinked. For a successful exchange of information, the language must initially be perceived by the receiver, be understood, and a response must be produced, if needed (Dell, 1986). Speech can be partitioned into four main areas: the semantic level of the meaning of what is being said; the syntactic level of grammatical structure; the morphological level of the units used in making the sentence (morphemes); and the basic units at the phonological level (phonemes), which are the sounds within the speech. By understanding the building blocks of language, it is possible to envisage the areas that can potentially be targeted by auditory training.

If useful, meaningful communication, as opposed to surface perception, is to occur, the initial step of perception of incoming speech needs to be followed by comprehension of what is being said. The process for extracting the elements of speech is well summarised by Eysenck (2005), who outlines five stages of the comprehension of speech. The first of these stages is the detection of the phonemes in the speech, which then allows for syllables to be identified, leading to identification of words. The last two stages operate at the syntactic and grammatical levels to make sense of the sentence.

A further consideration is that for language to be easily understood, the information extracted from speech signals should already be well-represented in long-term memory (Rönnberg et al., 2013; Rönnberg et al., 2008). The better the match between the incoming speech and its representation in long-term memory, the less cognitive resources are necessary to recognise it (that is, processing can be done relatively automatically). Processing resources are stretched when there is a mismatch, as in suboptimum listening conditions (Rönnberg et al., 2013), which can be caused by hearing loss or a difficult listening environment. This can also be the case for hearing-impaired listeners, for whom the amplification and compression strategies employed in hearing aids adjust speech such that understanding language is not as easy as it might otherwise be (Foo et al., 2007). By targeting any deficit with a known training strategy, it may be possible to reduce the burden on the cognitive processing resources through the development of better phonological representations of speech.

4.1.2. Auditory training

Auditory training, which is defined in Section 1.2.6, can be summarised as the way in which the contribution of an individual's audition to their perception can be increased. The largest group that stands to benefit from auditory training is individuals with hearing loss, as difficulty discriminating speech from background noise is their most common complaint (Ferguson & Henshaw, 2015b). There is, however, a need for commitment from individuals to any training regime, as the time required to achieve any auditory/communication improvement can be substantial. Therefore, for participants to feel they are benefitting from their commitment, there needs to be an advantage for them in terms of improved performance of speech discrimination in 'real world' situations.

In this chapter, it is important to distinguish between the more broadly defined communication training and auditory training itself. Some of the commercially available communication training packages, and the research surrounding them, do give evidence of improved speech-in-noise performance. However, these trainings often include cognitive exercises along with the auditory ones. For example, Listening and Communication Enhancement (LACE; (Sweetow & Sabes, 2006a) uses speech-in-noise exercises in conjunction with processing speed, working memory and inhibition training. The package is composed of 20 sessions, with listeners working at the limit of their performance level. An alternative training package, eARena (Siemens), uses much the same methodological approach as LACE in that it challenges its users to work at the limits of their ability and, like LACE, uses both auditory and memory span exercises. Research involving this more multi-faceted communication-based training (using LACE) is described in Chapter 5 of this thesis. This current study, although not conducted with participants with hearing loss, set out to explore transfer and improvements by utilising degraded speech in an auditory training paradigm. The training did not include a cognitive component, thus distinguishing it from broader communication training.

Auditory training is not new (for a review see Kricos and McCarthy (2007)). However, the advent of home-based, computerised auditory training has transformed the field, moving auditory training from a professional, staff-intensive activity to one that can be made available to many more individuals, including those with hearing loss—both with and without hearing aids or cochlear implants—as well as adults and children with speech and/or auditory processing disorders.

Henshaw and Ferguson (2013) conducted a review of the efficacy of computerised auditory training for adults with hearing loss. An overview of their comments on the quality of the research reviewed was given in Section 1.2.6. In brief, they found problems associated with the lack of a control group (or employing an active control group) and with the proper calculation of necessary sample sizes. They did, however, report that on review of the evidence, compliance with computer-based auditory training was high, and retention of learning was demonstrated at post-training follow-ups. The retention of any significant improvement was not reported for all studies; however, when it was, it was divided into 'on-task' and 'generalisable' (near transfer) improvements.

This is encouraging from the point of view of this thesis, as training needs to be generalisable to be considered truly beneficial, but any gains achieved also need to be maintained. On a note of caution, Henshaw and Ferguson, the authors of the review, considered the published evidence to be of very low to moderate quality, concluding that the published evidence around computerised auditory training for adults with hearing loss is not robust.

In the research for this chapter, the main criticisms reported in the review by Henshaw and Ferguson are addressed. This necessitated calculating adequate sample sizes and recruiting both an active and an additional control group. Further, strong attempts were made to document compliance with any training that was undertaken.

4.1.3. Evidence from previous auditory training research.

This sub-section further summarises some of the research that has employed auditory training. It reports on the success or failure of auditory training and gives evidence of any associated cognitive resource release. Much of the research reported in this section utilised degraded speech (vocoded and speech-in-noise). It reports on studies that employed both analytic and synthetic training and provides evidence of their relative success in improving listeners' discrimination performance. A key hypothesis for this chapter is that auditory training can lead to improvements in on-task performance that can also be transferred to improved performance for other types of speech. There may also be some degree of release of cognitive resources following the training. In order to evaluate this hypothesis, the most appropriate training was necessary, which was selected after a review of the evidence presented here.

Stacey and Summerfield (2008) examined the effectiveness of three different types of training. The research involved three groups of normally hearing adults undertaking either sentence-based (analytic), word-based (synthetic) or phoneme-based (synthetic) training. The speech material used was noise-vocoded. The sample size was small (n = 18 (6 per group), age 18–

28 years), and the training involved nine 20-minute sessions over a 2–3 week period. At the end of the training period, the effectiveness of the training program was assessed by the ability of the participants to distinguish between noise-vocoded consonants and vowels and the ability to discern words in noise-vocoded sentences.

For the ‘noise-vocoded vowel and consonant discrimination’ task, the phoneme and sentence training groups demonstrated no particular improvement, whereas the word-based training group did. The groups that undertook word and sentence training performed significantly better than the phoneme-trained group on the ‘target word in vocoded sentence’ task. For the ‘consonant and vowel’ task, the phoneme-trained group had been predicted to exhibit superior performance, as targeted training generally seems to offer the greatest success but this was not the case. However, as expected, the synthetic-trained sentence group achieved the best performance on the ‘words in sentences’ task. Since the research for this present chapter explores whether training on one type of degraded speech could transfer to improved performance on other types of degraded speech sentences, synthetic-based auditory training was employed.

Two additional pieces of research also explored analytic auditory training, this time with speech-in-noise. The first involved auditory training on different syllables or phonemes (Ferguson et al., 2014); the second distinguished between consonants (Woods et al., 2015). Ferguson et al.’s research was reported in detail in Section 1.2.6. and thus is not reviewed again here. In summary, their findings were of on-task improvements with no significant clear evidence of any near transfer to other speech tasks. There was, however, some evidence of transfer to cognitive tasks.

The Woods et al. (2015) research reported on training involving consonant identification, in which consonant-vowel-consonant sets were presented at different signal-to-noise ratios and the hearing-impaired listeners[34] attempted to identify consonant-vowel trigrams. After training

---

[34] Participants (N = 16) aged 61–81 years with mild to moderate hearing loss.

for a total of 40 hours (5 days a week for 1 hour per day), consonant threshold identification improved by 9.1 dB, and 94% of participants exhibited a statistically significant benefit from taking the training. Performance evaluations involved both familiar speakers (the speakers from the training) and unfamiliar speakers. Post-training, participants performed significantly better on the trigram task regardless of the familiarity of the speaker. There was, however, less recorded improvement in the separate task of sentence reception thresholds (SeRTs).

Woods et al.'s findings are similar to previous work by Davis et al. (2005), which used 6-channel noise-vocoded speech instead of consonants in noise. That work included five different experiments with varying group sizes, and training was performed within the sessions. Experiment 1 indicated improved discrimination of noise-vocoded sentences as listeners became accustomed to the degraded speech material. Experiment 2 is of particular relevance for this thesis (and the planned research in this chapter), as listeners were given feedback after sentences during training to see if this improved subsequent sentences. The feedback was given in the form of either clear-distorted-clear or distorted-distorted-clear sentences. The distorted-clear-distorted condition showed significantly better performance, producing a 'pop out' effect where, after a clear presentation on second listening of the degraded sentence, the sentence became clearer and learning could occur. Importantly, other experiments using non-words did not show the same improvements, leading the authors to conclude (similarly to Woods et al. (2015) that top-down, lexically driven learning is key in the perception of degraded speech.

The studies by Woods at al. and Davis et al. both indicate that training gains are best achieved by targeting the desired outcome. In the former study, that outcome is consonant identification, and in the latter, understanding degraded speech sentences.

In considering the points raised above, this present auditory training research was planned without a cognitive training element to ensure that it was investigating the building of new speech representations. It was designed based on synthetic sentence-based training principles, meaning that the training sentence material was presented to the different groups in a distorted/clear/distorted fashion after Davis et al. (2005). Based on the existing evidence, synthetic training appears to offer more consistent performance gains and thus was chosen as the preferred training method. As the end goal of training should be to transfer performance to general communication, being able to perceive and process sentence-based information would appear to offer the most benefit for individuals. To determine if training on one type of speech offers advantages in understanding other types of degraded speech, this chapter involved groups training on two types of vocoded speech, sine and noise,[35] and a third group trained on speech-in-noise. By training three groups on different types of degraded speech material (and having two controls), the transfer of any benefit after training could be determined.

Further, if it can be determined that sentence-based training does offer significant improvements then this may be considered as the preferred training method for future auditory training packages.

It was hypothesised that after a short training period, participants would be able to build new phonological representations of a particular type of vocoded speech (or speech-in-noise). This would then allow for more rapid identification of incoming speech when presented in the degraded speech format. If the model of Rönnberg et al. (2013) holds true, then this would potentially free some cognitive resources, allowing for better performance on auditory tasks that require some degree of cognitive processing. It may further allow for incoming information to be better processed into secondary memory. The assessment tasks for the research were chosen to help differentiate whether training gains were purely perceptual or allowed for the deeper processing of auditorily delivered information.

4.1.4. Aims and hypotheses

If auditory training is to be truly successful, evidence is required of transfer to different degraded speech material as well as targeted improvements in the trained speech material. This would indicate that transfer to real-world performance is possible. Furthermore, working

---

[35] The process of vocoding speech is described in Section 1.7

memory appears to offer support in understanding speech-in-noise, at least for some individuals; therefore, if listening becomes easier due to training, some cognitive resources previously used to assist with understanding of speech-in-noise will be released. The freed resources may allow for better encoding of received information into the secondary memory for later use.

The primary aim, therefore, is to investigate whether training in one type of degraded speech leads to discrimination improvements in that particular type of degraded speech (‘on-task’) as well as in other types of degraded speech (near transfer). A further aim is to investigate whether there is a release of cognitive resources when an individual’s discrimination performance improves. There is also the theoretical possibility that the simple act of undertaking any training will be cognitively stimulating and that this will allow for improved performance on unrelated tasks.

The hypotheses to be tested are as follows:

H4.1. If working memory plays a positive role in discriminating speech-in-noise, then there will be a significant relationship between participants’ performance in the speech-in-noise task and working memory at baseline, and the relationship will be stronger for a listening span (LSPAN) than for a reading span (RSPAN) task.

H4.2. If training in discriminating degraded speech leads to improved phonological representations, then speech discrimination performance for that type of degraded speech will significantly improve, above the improvement of groups offered different training.

H4.3. If training allows for improved discrimination performance in one type of degraded speech, then the process of auditory training may also allow for significantly improved performance in discriminating other types of degraded speech, thus demonstrating a degree of near transfer.

H4.4. If discriminating degraded speech becomes less effortful and more automatic, then any released cognitive resources will allow for significantly improved performance of speech-based working memory tasks in the same type of degraded speech.

H4.5. If the act of undertaking a training regimen itself is cognitively stimulating, then there will be a significant improvement in performance of non-related cognitive tasks following a training period.

H4.6. If discriminating degraded speech becomes less effortful and more automatic, then processing of heard information will become simpler, allowing for better memory of the delivered information.

4.2. Methods

4.2.1. Ethics

Aston University Research Ethics Committee (UREC) approved the proposed research (Ethics number 552).

4.2.2. Participants

There have been no previous studies using these methods as employed in this research. The use of two different types of vocoded training alongside a speech-in-noise trained group is unique to this study. As such, a power analysis using G*Power software (Faul et al., 2007) was used to calculate sample size based on previous work by Loebach, Pisoni, and Svirsky (2010), which used a similar, though not identical design. Loebach et al.'s research measured the discrimination performance of noise-vocoded speech between a trained and a control group of participants who all had normal hearing.

Loebach et al. reported that the trained group performed significantly better (M = 0.66, SD = 0.10) than the untrained control group (M = 0.47, SD = 0.14). Using the figures above, Cohen's d (Cohen, 1992) was calculated for groups with independent means using G*Power (Faul et al., 2007). This resulted in a participant number of 11 per group at a power of 80% and alpha at .05 (for two-tailed significance). As there were five groups[36], a total of 55 participants would be necessary to determine whether any statistically significant differences existed between trained groups and control groups. In order to allow for an attrition rate of 20%, 14 participants per group were recruited.

The participants all had normal hearing as defined by the British Society of Audiology (BSA, 2011) and were aged 18-50 years. Their hearing thresholds were averaged from 500-4000Hz in octave bands and no participant had a threshold below 20dBHL at any frequency in either ear. They were either recruited from the research participation system (SONA)[37], or were colleagues or associates of the researcher, and no payment was given for taking part, although the SONA participants were awarded course credits. This age range was selected to ensure that there was no expected deterioration in the peripheral hearing structures as there is evidence that even though hearing thresholds may not alter, this may still have an effect on speech discrimination (Dubno et al., 1984; Pichora-Fuller et al., 1995). Participants were assigned to either one of three auditory training groups, one control group that did visual training exercises and one that did no training. A one-way ANOVA indicated no significant difference in age between the groups. There were no drop-outs between the training and follow up. Participants were randomly assigned to groups (14 per group) using a computer program website that enables randomisation (GraphPad). Table 1 below provides the group details.

Table 1
Means and standard deviations and gender balance of the training groups

Training Group	Mean Age	Standard Deviation	Gender	Total
Sine Vocoded	25.8	10.8	6 Men 8 Women	14
Noise Vocoded	28.5	10.9	4 Men 10 Women	14
Speech-in-Noise	27.5	9.5	7 Men 7 Women	14
Active control	28.7	12.4	4 Men 10 Women	14
Control	28.0	10.7	11 Men 3 Women	14

---

[36] 1. Speech-in-noise trained. 2. Sine vocoded. 3. Noise vocoded. 4. Active control. 5. Control.

[37] Students are granted course credits for participating in psychology research

4.2.3. General design

Correlational analysis was used initially to investigate Hypothesis H4.1, which states that working memory plays a role in speech understanding in noise, so the two working memory tasks at baseline were compared with speech-in-noise performance and also with each other.

A within- and between-training group design was used for the main part of the research with all participants performing all the speech tests (sine vocoded, noise vocoded and speech-in-noise) before and after a training period. This design was used to explore Hypotheses H4.2 and H4.3, where it was hypothesised that the establishing of new phonological patterns may allow for improved ‘on-task’ speech discrimination. Further, the training would also allow for a transfer of improved performance between speech types. This may be due in part to the training act itself and/or the establishment of the new phonological patterns. The dependent variables were the percentages of correctly identified words scored for the sine-vocoded speech, noise-vocoded speech and speech-in-noise speech tasks. Data analysis was by MANOVA, as there were three dependent variables. In order to investigate H4.4, H4.5 and H4.6 regarding whether working memory and recognition would improve after auditory training, mixed-methods ANOVAs were performed.

4.2.4. Test materials

4.2.4.1. Audiometry

A calibrated pure tone audiometer as part of the aurical+ testing and assessment system (Otometrics) was used to perform the hearing assessment.

4.2.4.2. Discrimination tasks

Speech can be artificially made more challenging to comprehend in a variety of ways. The speech material chosen for the research for this chapter and the way that the speech was degraded is described below.

A novel set of speech test material (Patel, 2011) was used for the research for this chapter, based on the original Bamford-Kowal-Bench (BKB; (Bench et al., 1979) sentences. These sentences were developed primarily to be used with children with hearing loss (although they have become widespread in their use) and they, therefore, have a simple structure. In their original form, there were 21 lists of 16 sentences. The novel set of sentences was based on the original BKB sentences and uses simple language and construction. Keeping the speech material to basic construction has the effect of limiting any advantage of individuals who possess superior language lexicons.

For the sine- and noise-vocoded tasks, two-channel vocoding was used. This made the speech very challenging to understand as information from just two channels gives very limited phonological information. Two channels were chosen as it was one of the overall hypotheses of this thesis that making training challenging allows for the possibility of better transfer of performance. Hypothetically, this may be due to the fact that making training challenging allows for better development of new neural representations (or the strengthening of existing networks) than when training is easier.

Vocoding of the bank of new BKB sentences for all assessments was performed by the present researcher using AngelSim (2012). The frequencies that made up the speech sentences were divided into two bands, 200-1425.8Hz and 1425.8- 7000Hz. The filter bands had a roll-off of 24dB per octave. The amplitude modulation in the 2 filtered bands of the speech was extracted, then either sine waves at the centre of the two filter bands were amplitude modulated for sine-wave vocoded speech, or a white noise carrier was amplitude modulated for the two filter bands for noise- vocoded speech.

Speech-shaped noise was used as the masking noise for the speech-in-noise task affording limited glimpses of speech as there are no dips or troughs in the masking noise. The entire speech-in-noise test was delivered at a demanding non-adaptive SNR of -8dB SNR, meaning the speech was 8dB quieter than the masking noise and being non-adaptive maintained a consistent level of difficulty. This was felt to be more consistent with the vocoding tasks (as outlined below), which did not change their amount of vocoding so that all speech tasks remained consistently difficult.

The order of presentation of the three lists (speech-in-noise, sine vocoded and noise vocoded) for the discrimination tasks was manipulated to prevent order effects. The sentences used were randomly presented from the new bank of 1,628 BKB sentences and were spoken by a male speaker. The sentences were normalised for loudness (The Mathworks, 2006b).

48 sentences (3 lists of 16 sentences) were presented in each of the degraded manners as indicated above. The participants repeated the words that they heard immediately after each presented sentence and were encouraged to guess if they could not discriminate the words clearly. The tester recorded the number of key words correctly repeated by the participant. As an example, ‘The **clown** has a **funny face**’ is a standard BKB sentence, with the letters in bold being the scoring words. The individual performing the task would listen to the sentence presented (in whichever type of degraded speech was currently being tested) and then immediately repeat aloud what they heard to the researcher. If they did not hear the entire sentence, they were encouraged to say any single words they heard. A loose scoring method was employed where a correct mark is given even when only the route of the word is correct (i.e., ‘run’ would score a mark even if the target word was actually ‘running’), the same as in Section 2.2.4.2. Participants’ scores for the three types of degraded speech tasks were recorded as a percentage of the number of key words they were able to repeat.

4.2.4.3. Working memory tasks

Working memory span is a measure of the ability of an individual to hold and manipulate information. Individuals with better working memory spans can hold and manipulate more pieces of information than individuals with poorer ones. The task can be made more complex by introducing a secondary task or a variable that has a distracting influence on what is trying to be kept ‘live’ in working memory. The two working memory tasks employed are described below.

4.2.4.3.1. Listening span task (LSPAN)

The listening span (LSPAN) task for the speech-in-noise group was set at an SNR of -4dB and for the vocoded-trained groups 8-channel vocoded sentences were used. Since the goal of the listening span task was to assess sentence span performance rather than discrimination, the adjustment of the signal-to-noise ratio from the speech assessment task for the speech-in-noise LSPAN was made to allow for the speech discrimination to be somewhat challenging but that the words were clearly audible to the listener with some effort. This adjustment was confirmed by giving listeners 5 sample sentences.

In similar fashion, the sine- and noise-vocoded speech material when presented to the respective training groups, both had extra filter channels added (8 filter bands for each) to make discrimination challenging while audibility remained reasonably straightforward with effort. Participants for the vocoded LSPAN were also presented 5 sample sentences.

The LSPAN task involved participants listening to a sentence in full and then recalling the last word which they said aloud. If the word was correctly identified, two sentences were then presented and, after these had played in full, participants were asked to recall both of the last words in the sentences in the order that they had heard them. For example if Sentence 1 was:

- She cut with her knife.

The participant would say **knife.** They would then hear two more sentences, for example:

- Children like strawberries.
- The house had nine rooms.

The participant would then say **strawberries** and **rooms**. If their response was correct, they would receive 3 sentences, e.g.

- They’re buying some bread.
- The green tomatoes are small.
- He played with his train.

The participant would then say **bread, small** and **train.** This would carry on until the participant made an error. Once the participant failed on one or more of the words, they were played one sentence less than the number of sentences they had failed. If they then got the next word list correct, the number of sentences was increased again until the next failure. This stair-casing procedure could develop as in this example:

1 correct, 2 correct, 3 correct, 4 correct, 5 (one wrong), 4 (one wrong), 3 correct, 4 correct, 5 correct, 6 (one wrong), 5 correct, 6 (one wrong) - this would then be recorded as 5.

If a participant had two correct answers at one span level but only one at another during the stair-casing, the level at which they were able to achieve a repeat performance was recorded. This method was employed to prevent any lack of concentration at the start of the test affecting the results and to ensure a consistent level of response was recorded. Sentences 2, 3 and 4 of the original BKB speech material were used.

4.2.4.3.2. Reading sentence span (RSPAN)

A complex reading span task (RSPAN) was also performed with all participants (based on the work of Daneman and Carpenter (1980)). This was the same as originally described in Section 2.2.4.3. but as in Chapter 3 was performed as a reading span (RSPAN) task rather than an LSPAN.

4.2.4.4. Auditory recognition task

The auditory recognition task was conducted at the end of the testing session. Participants were informed at the start of the listening span task that there would be a recognition test at the end of the session. The recognition task involved identifying 10 of the sentences that were heard during the listening span task. The target sentences were presented in written format to the participant alongside three distractor sentences in a multiple choice format. The participant selected the one they thought they had heard by ringing it on the presented task sentence sheet.

4.2.5. Procedure

An initial ear examination and otoscopy were performed (BSA, 2016). If any adverse findings were discovered, participants were advised where to seek further medical assistance. Since the researcher was an audiologist, identification of ear health was straightforward, and no participants were found to have a condition that would disqualify them from the study. A hearing test was performed to BSA protocol (BSA, 2011) to confirm the participants' hearing fell within the normal threshold range (levels of <25dBHL at all frequencies from 500-4000Hz in octave bands). This test was only performed at the initial visit, and all participants were found to have hearing within the normal range. No participant had any threshold at any frequency over 20dBHL. The above hearing test and all the cognitive tests were performed in a well-lit sound proof room at Aston University with a lux reading between 320 and 360. Detailed instructions were given before the start of any of the tasks and checks that these had been understood were performed.

As an overview of the assessment procedure, Table 2 illustrates when the tests were performed and the order they were presented in for the three training groups.

Table 2
Assessment and training sessions for participants

Assessment 1	**Training**	**Assessment 2**
All groups Otoscopy Pure Tone Audiometry 3 Degraded BKB sentence discrimination tasks Auditory sentence span task Reading span task Recognition task from auditory sentence span task	**Training groups** 5 sessions a week of listening to 64 sentences for two weeks for the auditory training groups. Training was delivered in the degraded speech of the participants' training group. **Active control** 5 sessions a week of performing 20 minutes of cognitive exercises from a workbook. **Passive control** No activity	**All groups** 3 Degraded BKB sentence discrimination task Auditory sentence span task Reading span task Recognition task from auditory sentence span task

During the baseline assessments all participants performed the three types of degraded speech tests (noise-, sine-vocoded and speech-in-noise). The presentation for the different tests was counterbalanced to prevent order effects. The sentences were played to the participants through a Samsung personal computer and an Edirol UA-25 audio capture box with Sennheiser HD250 headphones. The Edirol soundbox output was calibrated using a Brüel and Kjær 2250 sound level meter using a speech calibration signal. The output through the Sennheiser headphones was calibrated to deliver the speech at 65dBSPL.

Both the working memory span tasks and the recognition task were also performed at baseline. The LSPAN was performed first and presented through the same set-up as the speech discrimination tasks with the speech output set at 65dBSPL. The LSPAN task was administered only in the groups own type of degraded speech (both the control groups performed this as a speech-in-noise task at both visits as they would have had little exposure to vocoded speech).

For the RSPAN, the test was administered using a program written by Millisecond software (Inquisit) and presented on a Toshiba Satellite Pro laptop computer. As was outlined in the materials section, participants were informed that if they did not make sufficiently accurate judgments of the sense of the sentences (thereby avoiding the complexity of the task) and their correct identification of sentences fell below 85%, their results would be excluded. None of the participant fell below this mark so all results were included. The final test was the auditory recognition test which was purposefully delivered at the end of the assessment session in order to ascertain how much of the auditory presented information used in the LSPAN had been encoded into, and could be recognised from, its store in secondary memory.

4.2.5.1. Training

Following the testing in Session 1, if participants were allocated to a training pathway, they completed exercises provided on a compact disc with the training material in the form of degraded BKB sentences. The type of degraded speech depended on the assigned group.

This either involved training on two-channel sine-vocoded speech, two-channel noise-vocoded speech or speech-in-noise at an SNR of -8dB. The active control group completed a cognitive exercise book for the same amount of time as the auditory training (5 days x 20 minutes a day for two weeks) or they did nothing (passive control).

The training material consisted of 128 BKB sentences split into 64 sentences for Week 1 training and 64 different sentences for Week 2 training. These sentences were drawn from lists 5-12 of the original BKB material (Bench et al., 1979).

Training was delivered in an unclear/clear/unclear format, as distorted sentences seem to become clearer after they have been listened to in this sequence, known as the 'pop out' effect (Davis et al., 2005). The participants were asked to listen to the sentences 5 times a week for 2 weeks and encouraged to attempt to determine for themselves what the unclear speech sentence was before checking with the clear sentence. Each session took approximately 20 minutes, and sessions were tracked on the training sheet provided. The active control group performed the same number and length of training sessions working through their cognitive training workbook. The training period for all the trained groups (and the active control) was kept at two weeks, as there is good evidence that overtraining does not necessarily prove beneficial (Molloy et al., 2012).

4.2.5.2. Post-training assessment

The procedure order for performing the tests at the second visit was the same as the baseline visit. The degraded speech discrimination tasks again used the new set of BKB sentences (Patel, 2011) and the presentations were counterbalanced as at the pre-training session to prevent order effects. The LSPAN task was repeated using the same procedure as for Session 1 but employing different BKB sentence lists. The computerised reading span task on the computer was repeated to keep the testing procedure the same as the first visit. As the lists were the same for each of the groups, task difficulty was maintained.

The final test was the recognition task involving sentences used on the second LSPAN task.

4.3. Results

The mean scores for the discrimination performance for each of the training groups on each of the degraded speech tasks at baseline and after training are given below in Table 3.

The data below show the means and standard deviations for the pre- and post-training speech discrimination scores. In general, all participants found discriminating the noise-vocoded sentences to be the most challenging, as evidenced by the low scores for this test.

Table 3
Means and standard deviations for the pre- and post-training scores on perception of BKB sentences for each of the training groups.

Speech test	Training Group	Baseline Mean % words correct	Standard Deviation	Post-training Mean % words correct	Standard Deviation
Sine-Vocoded	Sine-Vocoded	32.0	19.7	65.0	20.8
	Noise-Vocoded	30.8	21.3	56.0	16.4
	Speech-in-Noise	38.8	17.6	55.6	16.0
	Active Control	39.7	14.6	46.7	17.5
	Control	41.1	14.8	46.6	12.8
Noise-Vocoded	Sine-Vocoded	15.5	16.2	31.2	19.6
	Noise-Vocoded	16.5	18.0	46.4	21.8
	Speech-in-Noise	13.9	13.0	30.9	19.1
	Active Control	17.2	17.6	23.1	16.9
	Control	16.2	16.2	20.4	16.8
Speech-in-Noise	Sine-Vocoded	34.1	19.7	49.1	14.0
	Noise-Vocoded	41.5	21.3	57.5	11.0
	Speech-in-Noise	43.9	17.6	61.3	8.8
	Active Control	39.7	17.8	45.4	14.9
	Control	41.1	16.5	47.3	14.4

4.3.1. Analysis plan

When the performances of the five groups on the different types of speech were explored for normality, the recorded performances indicated a normal distribution on the basis of both a visual examination of the normality plots and Shapiro-Wilk testing (all *p* >.05).

As the recorded data for the speech tests had a normal distribution both before and after training, and as there was more than one dependent variable, MANOVA testing was used as an omnibus test to explore for differences in performance on the speech tests among the five groups of participants. If significant differences were found, post-hoc analysis would be performed to explore where any differences lay.

4.3.2. Correlation at baseline between working memory and speech-in-noise performance

Initial data analysis was performed prior to training to examine the correlation between performance on the two sentence span tasks (LSPAN and RSPAN) and the speech-in-noise discrimination scores. Only data from the participants who undertook the LSPAN presented as speech-in-noise (n = 42; speech-in-noise group and both controls) were used to perform the correlational analysis with the working memory tasks, as other groups completed the LSPAN in their own particular types of degraded speech. Table 4 below shows the performance data for this sub-group of 42 participants.

Table 4
Means and standard deviations for the pre-training scores of the sub-group of 42 participants that performed the LSPAN with sentences presented as speech-in-noise (n = 42).

Test	Mean	Standard deviation
BKB speech-in-noise score at baseline	41.54	13.82
Auditory span (LSPAN) score at baseline	3.52	1.04
Absolute reading span (RSPAN) score at baseline	29.92	12.88

Analysis of the data demonstrated normal distributions for the RSPAN and the speech-in-noise task but not for the LSPAN for the 42 participants. In view of this, the non-parametric Spearman’s rho was employed to explore any relationships. The relationships are given below in Table 5.

Table 5
Spearman's rho correlations of working memory performance with BKB speech-in-noise performance; N = 42.

	BKB speech-in-noise score at baseline	**Listening span (LSPAN) score at baseline**
BKB speech-in-noise score at baseline	-	
Auditory span (LSPAN) score at baseline	0.433*	-
Absolute reading span (RSPAN) score at baseline	0.020	0.193

Correlation of BKB Speech-in-Noise Score with Pre-Training LSPAN and RSPAN Scores
Key: * significant at $p < 0.05$.

The visual working memory task (RSPAN) was not significantly correlated with the speech-in-noise task, but the auditory working memory task (LSPAN) was (Spearman's rho = 0.433, $p < .01$). Hypothesis 2.1 was, therefore, only partially supported, as both the working memory tasks did not correlate significantly with speech-in-noise performance. Hypothesis 2.1 had also hypothesised that the LSPAN relationship would be stronger, so this part of the hypothesis was supported.

Since all pathway groups received different training, the correlations of the LSPAN and RSPAN scores with speech-in-noise performance for the other two groups could only have been measured on a group-by-group basis and would, therefore, have resulted in too few participants for meaningful interpretation.

4.3.3. Differences between groups in degraded speech perception before and after training

The individual speech test result data had demonstrated a normal distribution. When exploring the different groups' performances on the different speech tests, equality of covariance matrices (Box's test, $p = .758$) and homogeneity of variance (Levene's test, $p > 0.05$) were also demonstrated.

The performances on the three different speech tests by the different training groups were explored for any significant differences at baseline. A MANOVA with 14 participants per group and 3 dependent variables (df = 12) demonstrated that although the speech tests themselves were significantly different,[38] there was no overall significant difference between the five groups in performance on each of the speech tasks ($F(12,166) = 0.738$, $p = 0.713$; Wilks Lambda = 0.872; partial $n^2 = 0.045$).

For the post-training results, assumptions of equality of covariance (Box's test, $p = 0.369$) and homogeneity of variance (Levene's test, $p > 0.05$) were also met. As there was no significant difference among groups in performance on the speech tests at baseline, no covariate was necessary, so a MANOVA test was used to test the hypothesis that following the training pathway, there would be a significant difference in performance on the different speech tasks.

The MANOVA indicated a significant difference in performance on the speech tasks after training among the five different groups($F(12,166) = 4.383$, $p < 0.001$; Wilks Lambda = 0.485; partial $n^2 = 0.215$). To ascertain where the differences were more specifically, follow-up univariate ANOVA tests were performed. This produced the following results;

- There was a significant difference among groups for the post-training speech-in-noise test ($F(4, 62) = 10.1$, $p < 0.001$).
- There was a significant difference among groups for the post-training sine-vocoded test ($F(4, 62) = 10.7$, $p < 0.001$).
- There was a significant difference among groups for the post-training noise-vocoded test ($F(4, 62) = 10.9$, $p < 0.001$).

These results indicated that significant differences did exist between at least some of the groups for all of the speech tests after training. On examination of Figures 1, 2 and 3, the greatest differences between groups after training appeared in the type of speech that the group was trained on, e.g., the group trained on noise-vocoded speech performed better on the noise-vocoded speech test than the other groups. There was also an indication that in addition to improvements in their own type of trained speech, there was transfer to improved

---

[38] The noise-vocoded speech was significantly more difficult to discriminate.

performance on other types of degraded speech performance. This was greatest for the speech-trained groups and not the controls.

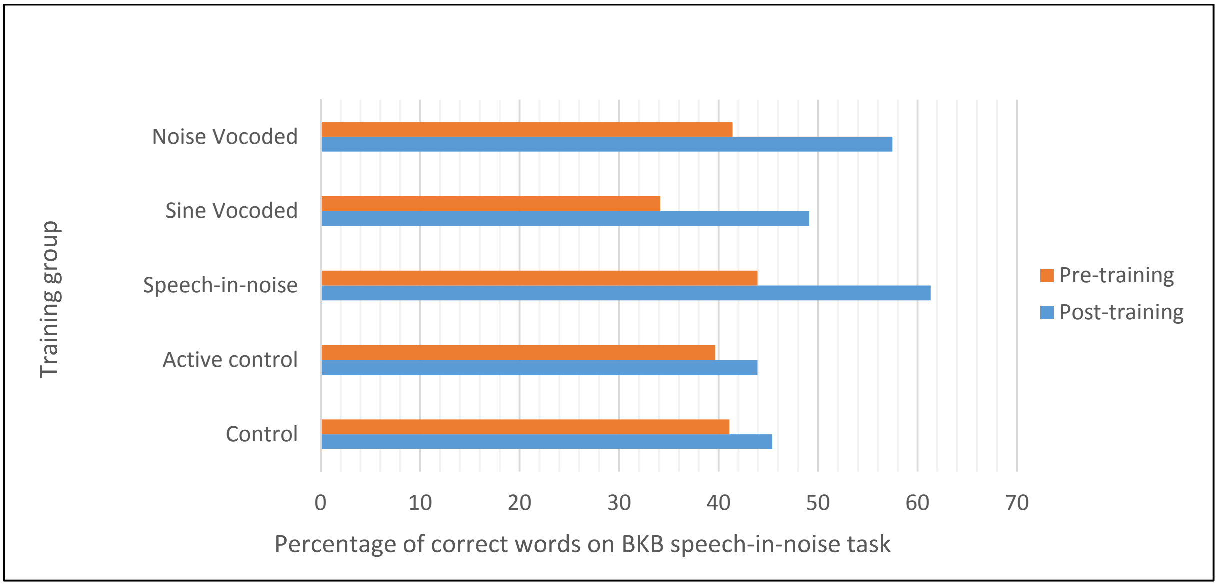


Figure 1. Pre- and post-training speech discrimination scores for speech-in-noise for the five training groups.

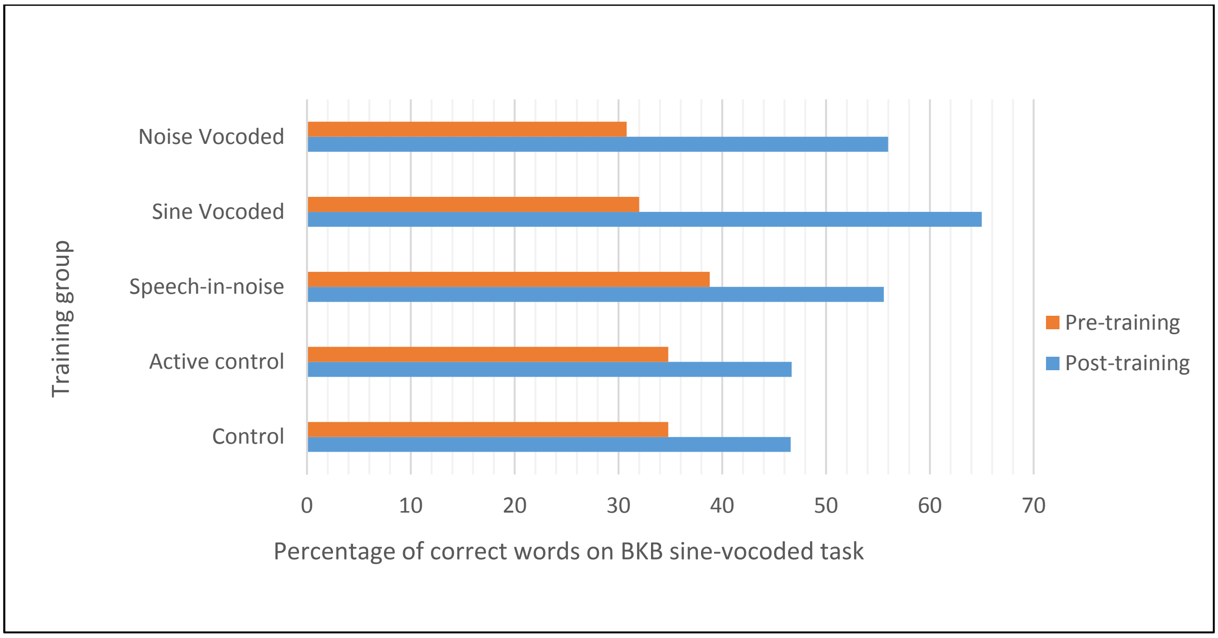


Figure 2. Pre- and post-training speech discrimination scores for sine-vocoded speech for the five training groups.

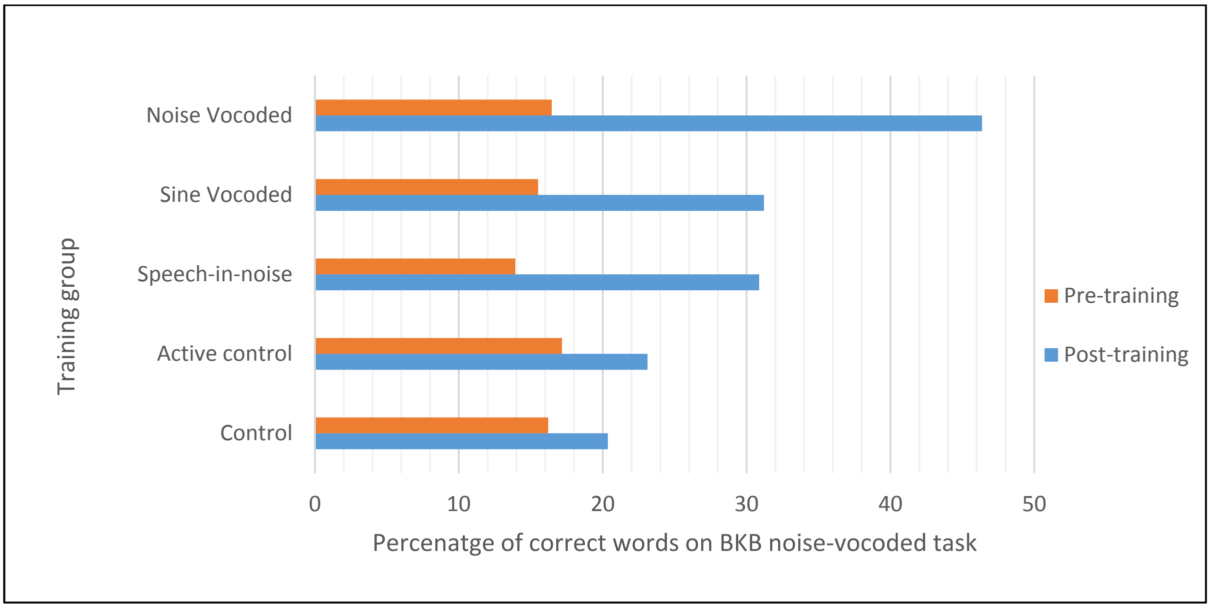


Figure 3. Pre- and post-training speech discrimination scores for noise-vocoded speech for the five training groups.

To explore for significant differences in post-training task performance for the different types of speech, multiple comparisons were made. These are discussed below.

Speech-in-noise discrimination performance for the different groups after training.

Multiple comparisons following the univariate ANOVA demonstrated that for post-training speech-in-noise test performance, significant differences were found between the speech-in-noise trained group and the active control and control groups, both at $p < 0.05$. There were no other significant differences.

Sine-vocoded discrimination performance for the different groups after training.

Multiple comparisons following the univariate ANOVA demonstrated that the sine-vocoded speech trained group performed significantly better than the control, active control and speech-in-noise trained groups, all at $p < 0.05$. There were no other significant differences.

Noise-vocoded speech discrimination performance for the different groups after training.

Multiple comparisons following the univariate ANOVA demonstrated on the noise vocoded speech tasks after the training period indicated that the noise-vocoded speech trained group performed significantly better than the passive control group ($p < 0.01$ level) and the active control group ($p < 0.05$). There were no other significant differences.

For within group performance changes on the speech tasks over time, a series of paired sample t-tests were performed and these indicated that all the groups performed all of the speech tests significantly better at the post training session compared to baseline (all p < .05).

The above summary of the results for between-group performance on the different speech tests indicates that a short course of training in a particular type of degraded speech can produce significant improvement for the trained group over other groups. This supports Hypothesis 4.2, in that training in a particular type of degraded speech offers a significant improvement for listeners after the training period. Hypothesis 4.3, however, is not supported, as the speech-trained groups did not demonstrate any near-transfer, i.e., they did not perform any degraded speech tests (other than the one that they were trained in) significantly better than other groups.

4.3.4. Working memory span

As speech discrimination scores improved for each of the groups (particularly in their respective trained speech material), it was hypothesised that there would be an associated freeing up of cognitive resources (H4.4). The mean LSPAN data at baseline and after training, recorded below in Table 6, were used to evaluate this hypothesis.

Table 6
Mean scores for the LSPAN and RSPAN at baseline and after training.

Memory span task	**Training Group**	**Mean span pre-training (baseline)**	**Standard deviation**	**Mean span post-training**	**Standard deviation**
Auditory (LSPAN)	Sine-Vocoded	4.0	0.475	4.86	0.66
	Noise-Vocoded	3.93	0.475	4.71	0.61
	Speech-in-Noise	3.36	1.15	3.93	0.10
	Active Control	3.93	1.07	4.00	1.21
	Control	3.29	0.825	3.64	0.745
Absolute Reading Span (RSPAN)	Sine0Vocoded	29.71	18.41	30.14	12.56
	Noise-Vocoded	24.57	15.40	31.21	15.38
	Speech-in-Noise	30.07	15.95	34.14	15.39
	Active Control	29.36	13.42	35.21	14.05
	Control	32.71	10.94	34.71	17.64

Shapiro-Wilk testing demonstrated a normal distribution for the reading span tasks both pre- and post-training but not for the LSPAN tasks. However, a visual inspection of the Q-Q plots for the LSPAN tasks suggested a reasonably normal distribution. To determine whether there had been any significant increase in LSPAN or RSPAN performance after the training period, mixed-methods ANOVA testing was employed, as this has generally proven to be reasonably robust when dealing with deviations from normality (Blanca et al., 2017).

Listening span (LSPAN) results

The mixed methods ANOVA gave the following results for the LSPAN.

There was a significant training group*time interaction on performance of the LSPAN task ($F(4, 65) = 3.331$, $p = 0.015$, partial $\eta^2 = 0.170$).

- There was a statistically significant difference between the training groups for LSPAN performance ($F(4, 65) = 4.756$, $p = 0.002$).
- The within-participant effect demonstrated a statistically significant effect of time on LSPAN performance ($F(1, 65) = 16.52$, $p < 0.001$, partial $\eta^2 = 0.203$).

The effect of training on LSPAN performance is shown below in Figure 4, demonstrating an increased LSPAN for all training groups at their second visit. Post-hoc analysis was performed to assess where any significant differences in LSPAN performance lay. At baseline, there was no statistically significant difference in LSPAN performance between the groups, as they all had similar working memory spans, all at $p > 0.05$. After training, post-hoc analysis demonstrated that working memory span as evidenced by LSPAN performance was significantly greater for the sine-vocoded speech trained group than for the control group or the active control group, both at $p < 0.01$, and marginally better than the speech-in-noise trained group ($p = 0.05$). Additionally, the noise-vocoded speech trained group also performed significantly better on the LSPAN after training than the control group and the active control group (both at $p < 0.05$). There were no other statistically significant results between training groups.

In terms of within-group improvements, post-hoc analysis revealed significant improvements in LSPAN performance for the sine-vocoded speech trained group between baseline and post training ($p < 0.001$) as well as for the noise-vocoded group ($p < 0.01$). The other training groups did not exhibit a statistically significant change in LSPAN performance over time. Hypothesis 4.4 was, therefore, supported, at least for the noise-vocoded and sine-vocoded speech trained groups.

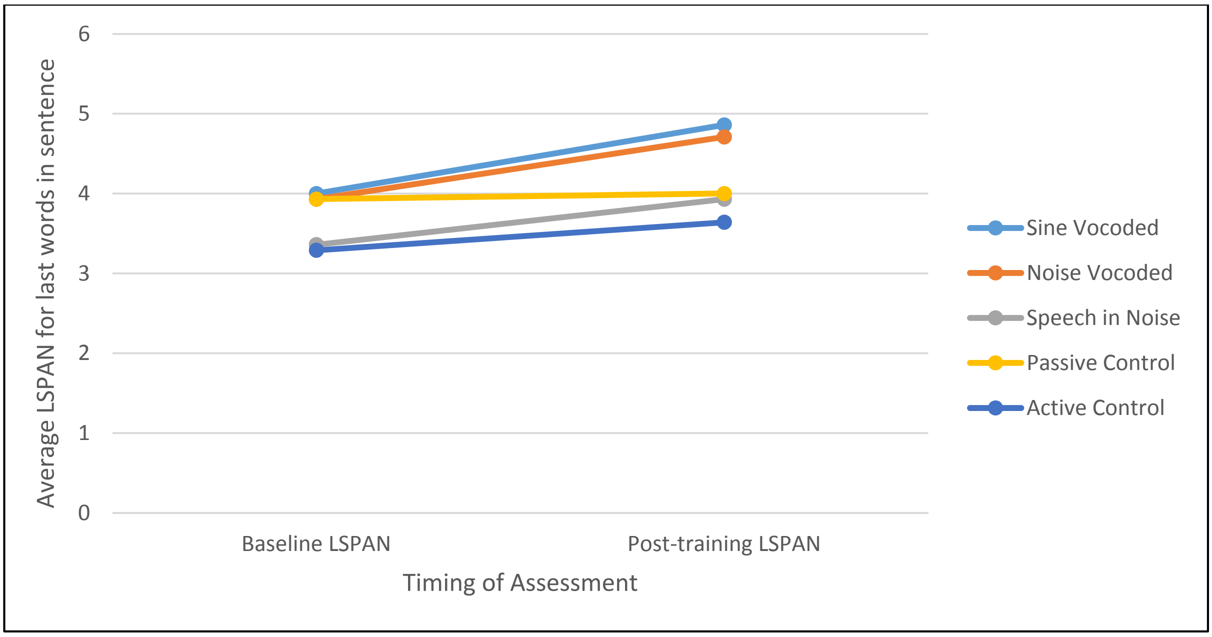


Figure 4. Mean LSPAN scores for the different training groups before and after training.

Reading span (RSPAN) results

All of the groups in this study improved their RSPAN scores by the end of the training period. As with the LSPAN, any differences in RSPAN performance between or within groups were explored by a mixed-methods ANOVA, for the same reasons mentioned in the LSPAN section above. Analysis of the RSPAN results was done to determine whether Hypothesis 4.5 could be supported. The ANOVA results indicated the following:

- There was no significant training group*time interaction on performance of the RSPAN task after training ($F(4, 65) = 0.596$, $p = 0.667$).

- There was no statistically significant effect of training group (*F*(4, 65) = 0.390, *p* = 0.815).
- There was, however, a significant effect of time (*F*(4, 65) = 6.372, *p* = 0.014).

As there was a significant effect of time on performance, a post-hoc analysis using a series of dependent sample *t*-tests was performed. For the sine-vocoded, speech-in-noise, active, and passive control groups, there was no significant difference. However, the difference was significant for the noise-vocoded speech trained group (*p* = 0.038). These results are shown below in Figure 5.

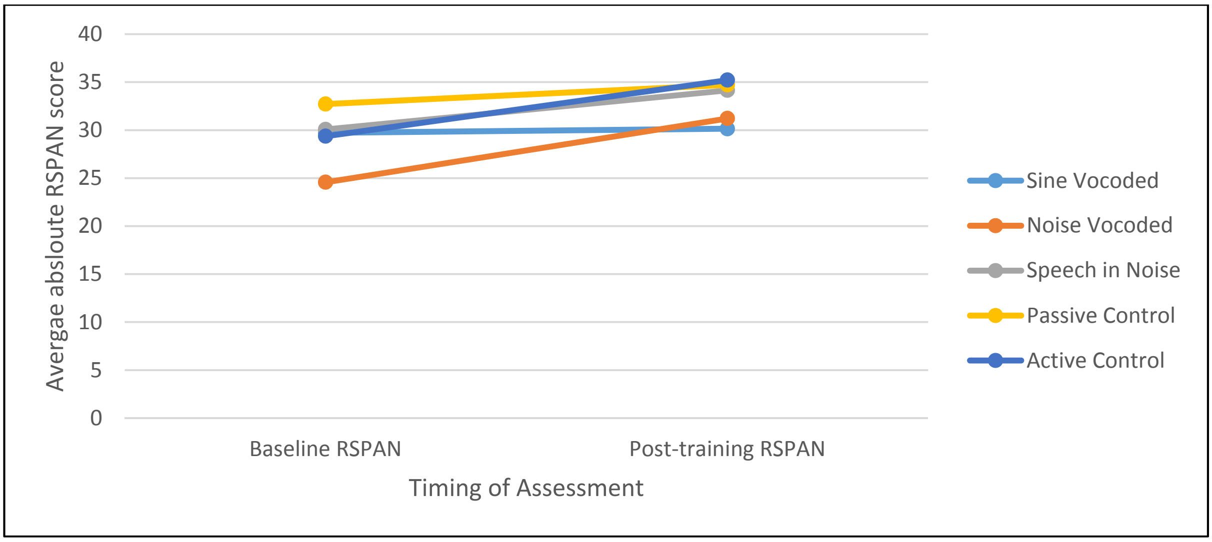


Figure 5. Mean RSPAN scores by training group before and after training.

The results of the RSPAN offer partial support to Hypothesis 4.5 that the act of undertaking training itself, although not directly aimed at improving a cognitive task, may still generate significant improvement. This, however, only holds true for the noise-vocoded speech trained group (Figure 5).

4.3.5. Recognition tests

The ability of participants to recognise sentences used for the auditory LSPAN task was evaluated both before and after training, as displayed in Table 7 below by training group.

Table 7

Means and standard deviations for pre- and post-training sentence recognition.

Time of Assessment	Training Group	Mean sentences correctly identified	Standard Deviation
Pre-training test	Sine-Vocoded	6.79	1.81
	Noise-Vocoded	6.29	1.90
	Speech-in-Noise	5.93	1.5
	Active Control	6.57	1.6
	Control	6.07	1.82
Post-training test	Sine-Vocoded	5.64	1.87
	Noise-Vocoded	6.07	2.13
	Speech-in-Noise	6.57	1.65
	Active Control	5.79	1.81
	Control	6.86	2.14

On general observation, as seen in Table 7, although some recognition scores (for the control, active control and speech-in-noise groups) improved, the noise- and sine-vocoded groups' recognition scores fell. The scores for each of the groups was tested for normality both pre- and post-training, and visual examination and statistical analysis showed normal distributions.

A two-way mixed ANOVA was performed in a similar fashion to the working memory analysis in order to assess for significant differences between and within groups after training that would support H4.6. That hypothesis stated that auditory training may allow for freeing up of cognitive resources, which would be demonstrated by differences between the trained and non-trained groups and/or within-group improvements. The ANOVA demonstrated the following:

- There was a statistically significant time*pathway interaction for sentence recognition and time ($F(4, 65) = 3.139$, $p = 0.020$, partial $\eta^2 = 0.162$).
- There was no statistically significant effect of training group ($F(4, 65) = 0.079$, $p = 0.989$, partial $\eta^2 = 0.005$).
- There was no statistically significant effect of time ($F(4, 65) = 0.442$, $p = 0.508$, partial $\eta^2 = 0.007$).

Post-hoc analysis found no significant differences in recognition of the sentences that were used for the LSPAN task, as the performance differences between training groups were all $p > 0.05$. Recognition performance varied for the five training groups, with some showing improved performance after training and others performing more poorly (see Figure 6). Although there was noted a significant interaction indicating that the trends were significantly different between the groups, this did not equate to any difference between the means on post hoc analysis. As there were no statistically significant differences between or within the groups, Hypothesis 4.6 could not be supported.

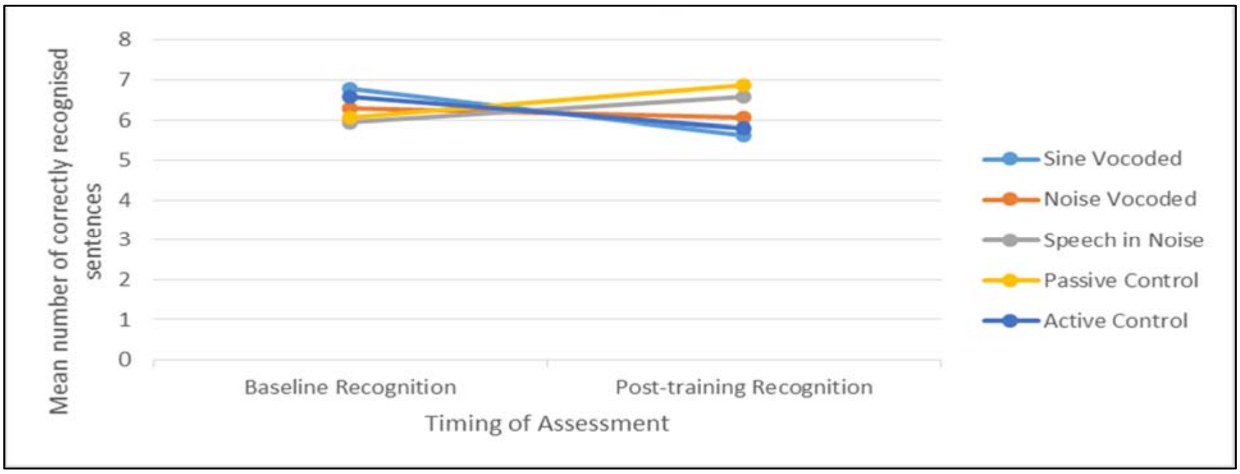


*Figure 6***.** Mean number of recognised sentences before and after training.

4.4. Discussion

As had been hypothesised, a significant relationship was demonstrated between working memory (as measured by the LSPAN task) and speech-in-noise recognition performance when undertaken at baseline (H4.1). This is a finding that has not been consistently demonstrated for younger listeners with normal hearing (Füllgrabe & Rosen, 2016b) and may be due either to the speech-in-noise test used in this chapter being consistently difficult or to the use of the LSPAN rather than the RSPAN for the working memory task. The LSPAN was made somewhat more complex by being performed with degraded speech (albeit at an easily understood SNR).

One potential reason for a relationship between the speech-in-noise task and the LSPAN, rather than the RSPAN, is that listeners still need to discriminate words for the LSPAN, which may necessitate some cognitive support. Individuals, therefore, who can better discriminate speech will need to use less cognitive support to do so, leaving more of their pool of cognitive resources available for other functions such as working memory. If the participants' discrimination ability was poorer, they would need to use cognitive ‘back-up’ to assist with discrimination, hence performing less well on the LSPAN as their cognitive resources are unavailable for working memory (in this case because they are being employed to assist with discrimination).

There is the possibility that the converse is true, in that individuals with better working memory may be better equipped for speech-in-noise or vocoded speech discrimination performance. However, if this were true, then it could have been anticipated that the relationship of speech-in-noise/vocoded speech discrimination performance should also correlate significantly with working memory when presented in a visual task, in this case the RSPAN. As there was no relationship with the RSPAN, it would appear that when using a speech-in-noise task with younger adults at a consistently difficult SNR, it is discrimination ability that predicts working memory performance, not the reverse. Therefore, although H4.1 is supported, it may be the case that it is discrimination that, to some extent, predicts working memory performance.

Hypothesis 4.2 stated that auditory training would lead to significantly improved performance in the trained material itself. This was indeed the case, as significant differences were demonstrated between all of the trained groups in their own types of degraded speech when compared to the two control groups. This replicated the findings of two previous pieces of research, Hervais-Adelman et al. (2011) and Loebach, Pisoni, and Svirsky (2009), in which training on vocoded speech led to improved ‘on-task’ performance with an indication that transfer to other tasks may be possible.

The results obtained from the present study, also along similar lines to those of Hervais-Adelman et al. and Loebach et al., reported that training did offer an indication that transfer may be possible. Figures 1, 2 and 3 in Section 4.3.3 all demonstrate that the speech-trained groups’ final performances were better for all of the speech tasks compared to the control groups. The differences between final performances, however, were not significant, all $p > 0.05$. Therefore, Hypothesis 4.3, which theorised there would be significant evidence of near transfer performance gains, could not be supported. It had been anticipated that transfer of improved performance between the different types of vocoded speech may occur, as new speech representations may be similar enough to still allow better recognition, but this was not evidenced.

In contrast, researchers using analytic training (whilst not employing vocoded speech) have reported that training using lexically difficult words can lead to near transfer improvements (Burk & Humes, 2007). Indeed, Burk and Humes' later work (Burk & Humes, 2008; Humes et al., 2009) reported, along similar lines to their previous work, that training could demonstrate 'on-task' improvements but also transfer to novel words and speakers. Oba et al. (2011) also demonstrated the possibility of significant near transfer improvements using analytic training, albeit with cochlear implant patients.

It may be that in addition to the type of training offered (synthetic vs. analytic) and the different speech test materials, the training regime may account for the reports of varying success of near transfer. For this present research, the training was over a short period and involved 3 hours and 20 minutes of training time. This may not have been sufficient to establish easily accessible strong phonological representations of the new speech that would allow for near transfer to be exhibited. The research by Oba et al. (2011) and Burk and Humes (2007) included total training times of 10 and 15 hours, respectively.

As the type of training in this chapter employed a synthetic approach to training and this was unable to demonstrate any significant near transfer improvements, consideration needs to be given to word-based (analytic) training, as this may offer some advantages for an individual that synthetic training may not. This raises the question of whether both types of training could be employed when designing any auditory training program. Potentially, this should be based around an individual's speech assessment performance, where any deficits in listening skills could be identified.

Although there were no significant near transfer gains in speech discrimination performance evidenced, auditory training did appear to provide an improved working memory span, as evidenced by performance on the LSPAN. Thus, Hypothesis 4.4 was supported. Although this was not the case for all training groups, both the sine-vocoded and noise-vocoded speech trained groups did demonstrate this improvement. This was seen in their significantly better performances on the LSPAN compared with the control groups. Further, within-group analysis also demonstrated statistically improved performance on the LSPAN for the sine- and noise-vocoded groups after training. The theory is that as ease of access and understanding of the vocoded speech becomes more automatic, this leaves more cognitive resources available to hold information in working memory.

It could be argued that as the changes were small (although significant) that instead of real changes in working memory span, that improvements were due to participants finding different strategies to improve their performance rather than any actual memory span improvement. However, as test/re-test reliability of the LSPAN is good, the final span measure was achieved by 'staircasing' and all groups were balanced with no significant difference in performance at baseline this would give added credence to the results. All participants only did the LSPAN test twice (once at baseline and once after training) and, therefore, they all had the same opportunity to improve performance by adjusting strategy. In view of this, it does lend support to the fact that the groups that did demonstrate significant improvements did so as a result of the auditory training that they had undertaken.

The ease of language model (Rönnberg et al., 2013) offers an explanation for this improvement. When less explicit processing is necessary (as the degraded speech is matched more automatically with the newly established representations), there are extra cognitive resources available for other tasks, such as working memory. This relationship has not been commonly reported in other research and is perhaps exhibited in this chapter due to the test conditions for assessment and training being particularly challenging.

Henshaw and Ferguson (2014) suggest that benefits from auditory training are more evident when listening is challenging, which would fit the pattern of being trained on and listening to noise- and sine-vocoded speech material. Indeed, it was the participants trained on these vocoded pathways who demonstrated improved working memory performance on the LSPAN.

Further, it was also found that RSPAN performance improved significantly, but only for the noise-vocoded speech trained group. This offers some support to Hypothesis 4.5, that undertaking any type of training may be cognitively stimulating and allow for the potential improvement of cognitive functions (Henshaw & Ferguson, 2014). However, the fact that the other trained groups in the present research, particularly the sine-vocoded speech trained group, did not show such improvements makes this argument less convincing. It does, however, raise the potential that when reporting on the success or otherwise of rehabilitation programs aimed at improving communication, the focus should not be solely on improved audition.

It had also been hypothesised that as discrimination improves, there would be additional cognitive resources available, such that when performing the LSPAN, information could be more easily processed into secondary memory for later recall (Hypothesis 4.6). There was, however, no evidence of this on the recognition task. The mixed pattern of results did not lend any evidence to an improvement in information processing. As the training focused purely on developing new auditory representations, although primary encoding in working memory could be improved, the new representations did not allow for encoding into secondary memory for later recognition. Potentially, to demonstrate such improvements, training may have to be more intense or targeted more specifically at secondary memory encoding.

Compliance with training for the research for this chapter was high. Ferguson and Henshaw (2015a) reported that compliance for computerised auditory training is also generally high. It is unfortunately the case that some trainees find that training at their limit of performance is disheartening, as they make frequent mistakes and have to exert great effort. However, if the findings from this chapter are replicated, it does appear that making training challenging does elicit the greatest advantages. If this is the case, rather than simplifying the training, better methods to encourage compliance, such as utilising training as an add-on to group counselling support sessions (Hawkins, 2005), would appear to be a more advantageous route to follow. Further, as hearing aid users report that hearing aids often fall into disuse because the benefit from them does not outweigh the perceived trouble in using them (McCormack & Fortnum, 2013; Mizutari et al., 2013), then some form of structured activity with proven benefits may improve hearing aid uptake and use. The use of communication training for hearing aid users is explored further in Chapter 5.

There were some limitations to this study; these are explained in the following section.

4.5. Limitations

The first limitation of this study was that the auditory training undertaken, although logged by participants, could not be readily verified. It would have been advantageous to have a computer-based 'on-line' system that could log training for participants or to have participants undertake their training at the university so that adherence could have been more closely monitored.

The length and amount of training could have been increased, but based on evidence from previous and present research, longer training periods do not always seem warranted (Wright & Sabin, 2007). Retention of any improved performance could also have been evaluated (e.g., follow-up visit at 6 months); however, as speech-in-noise was used as one of the training regimes and participants would be exposed to that but not to vocoded speech during their daily lives, it is unclear what conclusions, if any, could have been drawn from an additional visit.

4.6. Conclusions

After a short training course, on-task improvements for the discrimination of novel (or degraded) speech were readily achievable, as were far transfer improvements in working memory span (LSPAN). There was, however, no clear evidence of near transfer to discrimination of different types of degraded speech.

For the younger individuals without hearing loss in this research, working memory does appear to offer some support when listening to degraded speech. Conversely, it may be that it is listening ability that accounts for performance variations and that discrimination is easier for some individuals, necessitating less cognitive support. In that case, those individuals could use the extra resources to hold more information in their working memory when tested on their LSPAN performance using degraded speech material. After auditory training, as discrimination improves, the ability to hold even more information in memory also shows evidence of improving. There does not, however, appear to be any subsequent improvement in encoding information from working memory into secondary memory.

The relationship between the complex working memory task and speech-in-noise discrimination performance would further suggest that when testing for memory span, using the sensory modality under investigation (so in auditory mode when exploring any association with speech-in-noise discrimination performance) may provide additional insight into relationships that non-sensory-specific assessment does not.

This thesis, to this point, has focussed on participants with measured normal hearing or who have presented with no hearing problems. The next chapter focusses on individuals with hearing loss and explores the potential benefits offered by a communication training package.

CHAPTER FIVE

5. The role of a commercial communication training package (LACE) in the rehabilitation of newly fitted hearing-aided patients.

5.1. Introduction

Previous research in this thesis was conducted with individuals self-presenting with either normal hearing (although they had audiograms performed, some of which demonstrated mild hearing losses) or measured normal hearing thresholds. It has been demonstrated with these individuals that cognitive training (Chapter 3) involving processing speed and control of inhibition failed to demonstrate improved performance of the cognitive functions themselves or any transfer to significant improvement in speech perception. Auditory training (Chapter 4), conversely, did offer some advantages, as following synthetic auditory training, performance improvements were observed in speech discrimination, both in noise and when speech was degraded by vocoding. Additionally, as a result of these improvements, there was a degree of cognitive resource release.

This chapter focuses on understanding the role of a commercial training package and how it may benefit newly fitted hearing aid patients. Specifically, the question addressed was whether a communication training package offers advantages over the standard practice of normal acclimatisation. To determine whether training using the Listening and Communication Enhancement (LACE) package really delivered any advantage, three pathways were used for comparison purposes: the LACE pathway, a standard acclimatisation pathway and an active control pathway that involved performing listening exercises.

Previous research from the training package developers (Sweetow & Sabes, 2006a) has reported advantages for listeners after undertaking the training. Other authors have been more cautious in reporting improved performance (Olson, Preminger, & Shinn, 2013) or have not found any evidence that improvements occur following LACE training (Saunders et al., 2016a). A thorough critique of these studies follows in Section 5.1.3. It is, however, important to note at this point that an aim of the research for this chapter was to standardise the research methods, in particular the participants' degree of hearing loss, type of hearing aids fitted, fitting formula and most importantly that all patients were fitted and tested at the same time before any acclimatisation to the hearing aids took place. The large-scale randomised trial by Saunders et al. (2016b), whilst of high quality, did define ‘new’ hearing aid users as those having hearing aids for more than 4 weeks and less than 6 months. This wide range means that some listeners who enrolled in the research as new users had already undergone a period of acclimatisation, whereas others had not. The period necessary for hearing aid acclimatisation, or whether indeed such a period is necessary, is a matter of some contention. Some authors have reported that speech perception improves over time after hearing aid fitting (Munro & Lutman, 2003; Yund et al., 2006) and some that it does not (Humes et al., 2002). In either case, to control for any potential acclimatisation effects, all patients in the research for this chapter were baseline-tested at the time of fitting, so all had the same exposure to non-linear amplified sound.

As discussed in the introduction to this thesis, the concept of auditory training is by no means new. The methods of inducing communication enhancement follow one of two broadly defined pathways: analytic (phonemic/word based) and synthetic (sentence based). Many of the communication enhancement training programs available commercially also involve some form of cognitive training alongside any auditory training exercises.

Previous reports suggest that combined communication training can lead to improved listening skills and communication for patients newly fitted with hearing aids (Miller, Watson, Kistler, Wightman, & Preminger, 2008; Sweetow & Sabes, 2006a). Much of the research in this area, however, has lacked rigor (for a review, see Henshaw and Ferguson (2013)). Further, some of that research has been sponsored or conducted by the proprietary companies that manufacture and market the training software (Sweetow & Sabes, 2006a), bringing its objectivity into question. For the research in this chapter, to ensure that training was advantageous for the participant beyond the normal exposure to sounds following routine hearing aid fitting, a randomised controlled trial was conducted. It was also ensured that the participants in the different training pathways were balanced in terms of ages and years of education (as an indicator of potential intelligence), as there has been some weak evidence that these factors may influence speech-in-noise performance (van Rooij & Plomp, 1990).

It has been noted that the rate of successful effective uptake of hearing aids is low (Smeeth et al., 2002), with many recipients reporting problems and being inconsistent with their use (McCormack & Fortnum, 2013). Smeeth et al.'s study reported that 40% of patients fitted with NHS hearing aids were not using them regularly. McCormack and Fortnum identified five key areas in which hearing aids were not meeting expectations:

- Aids not providing enough benefit
- Poor fit and comfort of the hearing aid
- Problems with care and maintenance of the aid
- Patients could hear well enough without a hearing aid
- The aid possessed certain technical shortcomings.

If a communication training package is able to improve some of these perceived failings, for example, by increasing the benefits that the aid provides, then increased aid use and retention should follow. It could also be argued that by improving uptake and use of hearing aids, this training would have a positive effect on seeking out social interactions as communication becomes easier. This is turn would have other positive health benefits.

There are proposed consequences that may arise from hearing loss. Work by Lin et al. (2013) and others (Lin, Ferrucci, et al., 2011; Lin, Metter, et al., 2011) arising from the Baltimore longitudinal study of ageing (previously discussed in Section 1.1.4.) has explored the link between hearing loss and decline in cognitive ability. Wayne and Johnsrude (2015) have suggested four causal mechanisms that may account for this relationship (also introduced in the same 1.1.4 section). Further, by developing an ever improving understanding of this relationship, enhanced evaluations of rehabilitative interventions (including hearing aid use) could be better explored.

There is some early evidence that hearing aids may offer some protection against these cognitive declines (Amieva et al., 2015; Deal et al., 2015; Qian et al., 2016). This is particularly important if the cascade theory (Sekuler & Blake, 1987), that sensory underload drives cognitive decline, holds true. If this decline is reversible or possible to mediate against, then successful aiding would allow for not only better access to audition but also improved cognitive well-being. If there is anything, therefore, that can be done to improve communication ability both with and without hearing aids and consequently improve patient outcomes, it would be both cost-effective in terms of health care resources as well as clinically and socially beneficial.

Previous studies exploring the potential efficacy of communication training have generally only examined immediate perception of information, as opposed to the processing and use of that information for communication, recall and reasoning. The research in this chapter was designed to examine whether any potential gains in listening skills may also lead to deeper communication and processing benefits beyond simple surface-level improvements in speech perception.

A report of the Committee on Hearing and Bioacoustics and Biomechanics in America (CHABA, 1988) on speech understanding and ageing has suggested three hypotheses regarding the difficulties faced by older listeners:

1) The peripheral hypothesis: The impact that declining hearing thresholds have on speech perception is related to structural changes that occur in the cochlea with age.

2) The central auditory hypothesis: Declining performance in speech perception stems from changes in neural connectivity in the brainstem and the auditory cortical areas of the brain and in the way that speech and sounds are processed.

3) The cognitive hypothesis: Speech-in-noise perception becomes more difficult with age due to the role that cognitive functions play in the understanding of speech, particularly that of speech-in-noise.

There would appear to be an interplay between these three ideas. Chapter 2 of this thesis demonstrated that cognitive functions do add to the explanation of speech-in-noise performance variation; however, in Chapter 3, training of these same cognitive functions was not successful in improving speech perception performance. Chapter 4 demonstrated that discrimination improvements were possible following auditory training and that release of cognitive resources could be evidenced as speech discrimination improved.

Hearing aids are able, to some extent, to ameliorate some of the problems with accessing auditory signals that are caused by peripheral hearing loss. There is also evidence that hearing aids can induce plastic neural changes and improve neural connectivity (Giroud, Lemke, Reich, Matthes, & Meyer, 2017). It would, therefore, seem likely that any communication training packages that target cognitive functions and auditory training could possibly offer communication advantages. It may be, however, that this scatter-gun approach, in which individuals perform all potentially helpful exercises instead of employing targeted training, is not necessarily the most effective method of improving communication. This would appear to be borne out by the fact that throughout this thesis, auditory training, where it is targeted has produced the desired outcomes. Therefore, identifying where problems exist for an individual in terms of communication/cognitive ability would allow for the construction of tailored training that is better suited to deliver the support that they require.

5.1.1. Auditory/communication training

Auditory training can be undertaken on an individual basis in private sessions (Rubinstein & Boothroyd, 1987) or with a group (Hawkins, 2005). One of the difficulties with both these types of support is that they are very time-consuming for staff. The advent of affordable personal computing and reasonably priced access to the internet has opened up the possibility of auditory training becoming less staff-intensive and more accessible to patients with and without hearing loss. As such, a number of internet/computer-based training programs have become available on the market. The success of these in assisting patients fitted with hearing aids has been reviewed by Henshaw and Ferguson (2013), amongst others.

Auditory training may not necessarily be for the sole use of the hearing-impaired, as communication training programs have also been demonstrated to offer benefits to children with delayed language (Tallal et al., 1996) and to individuals without hearing loss but who carry a diagnosis of auditory processing disorder (APD;(Bamiou, Musiek, & Luxon, 2001; Cameron et al., 2012; Chermak & Musiek, 2002; Musiek, 2004; Weihing, Chermak, & Musiek, 2015). The evolving opportunities offered by these on-line communication training programs has the potential to offer readily available support to a much greater number of individuals with a wide range of communication problems.

In designing the on-line LACE training program, (Sweetow & Sabes, 2006a) acknowledged that older adults have fewer cognitive resources available (Craik & Byrd, 1982) and that there is, therefore, a resultant decline in cognitive skills with age (Cepeda, Kramer, & de Sather, 2001; Wingfield, Tun, & McCoy, 2005). Amongst the key cognitive skills that show evidence of decline are those of processing speed and working memory (Cepeda et al., 2001). Sweetow and Sabes (2006a) argued that listening to speech in noise is multi-faceted and that listeners have to integrate several contributing skills. Specifically, they report that these skills may include "cognition, auditory memory, auditory closure, auditory learning, alongside the use of pragmatics, semantics, grammatical shape, metalinguistic, localisation, visual cues, repair tactics and effective communication strategies" (Sweetow & Sabes, 2006a, p. 543). In view of this, the LACE training was designed to include cognitive training alongside listening exercises. The concept of cognitive training support has also been addressed by other authors, such as Pichora-Fuller (2003a) and Kiessling et al. (2003a).

In two reviews of the impact of cognitive skills on speech-in-noise perception (Akeroyd, 2008; Besser et al., 2013a), the strongest predictor of performance variations in speech perception in noise, apart from hearing thresholds, has been working memory. Evidence of associations of speech-in-noise perception with both intelligence and processing speed (Desjardins & Doherty, 2013; Vaughan, Storzbach, & Furukawa, 2006) has also been demonstrated. The roles of control of inhibition and information processing speed in hearing speech in challenging situations (Sommers & Danielson, 1999; Taler et al., 2010) were explored in Chapter 2, and both skills showed some evidence of a relationship with speech-in-noise performance. Training of these functions gave no advantage in discrimination performance, however, as Chapter 3 had recruited normal or near-normal hearing individuals, the same lack of any significant improvement cannot be inferred for older patients/individuals with hearing aids.

Overall, it would seem that in addition to auditory training exercises that set out to create new or adjusted representations of speech (Kraus, 2012), there is a strong argument that there may need to be a complementary focus on cognitive training. Linked with this is the fact that improved recognition of speech sounds allows for easier access to what is being said (as reported in Chapter 4), such that language understanding may become a simpler task (Rönnberg et al., 2013). Given that when language becomes easier to understand there should be a subsequent freeing up of cognitive resources that are no longer needed to assist with discriminating speech, these additional available resources could help with secondary encoding of the spoken information.

In addition to the lack of research quality identified in the review by Henshaw and Ferguson (2013), they also present the criticism that there has been little exploration of any improvement in secondary encoding (the deeper processing of received information). This was investigated in Chapters 3 and 4 of this thesis, where no improvement was noted in secondary encoding after either cognitive or auditory training. This will also be explored in this chapter by determining whether information delivered during listening assessments has been encoded into new memories. The ability to process newly delivered information more efficiently after a training period would give evidence of training's ability to improve deeper communication for individuals.

The following literature review focuses on LACE training, followed by a critique that discusses the quality of the research. Although one of the primary aims of LACE is to achieve improved speech perception in noise, the current research in this chapter also aimed to assess any release of cognitive resources associated with improved listening. The existing literature on

LACE was, therefore, examined for evidence of speech perception improvement but also any cognitive performance changes.

5.1.2. LACE literature review.

A search of online literature resources was performed for studies published between 2005 and the present, as LACE was not developed until 2005. The Preferred Reporting Items for Systematic Reviews (PRISMA) checklist was used, and the search engines were SCOPUS and Web of Science. Only two search terms were used: ‘hearing’ and ‘LACE’. The initial search returned 22 papers, and the process for accepting the papers is detailed below in Figure 1.

To be included in the literature review, a study’s participants had to have hearing loss and be new or existing hearing aid users. Research papers were excluded if the participants had normal hearing or did not wear hearing aids.

The one paper in the PRISMA review from an additional source (Martin, 2007) was a paper that had been suggested to the author of this thesis. None of the six records excluded at the screening stage involved hearing research. Of the three full-text articles that were excluded, one was a reprint of research reported the year before (Sweetow & Sabes, 2007), one was an editorial review (Jerger, 2006) and one did not involve hearing aid users (Song et al., 2012).

The findings of the remaining six papers are summarised in Table 1, which follows the PRISMA review.

PRISMA Flow Diagram

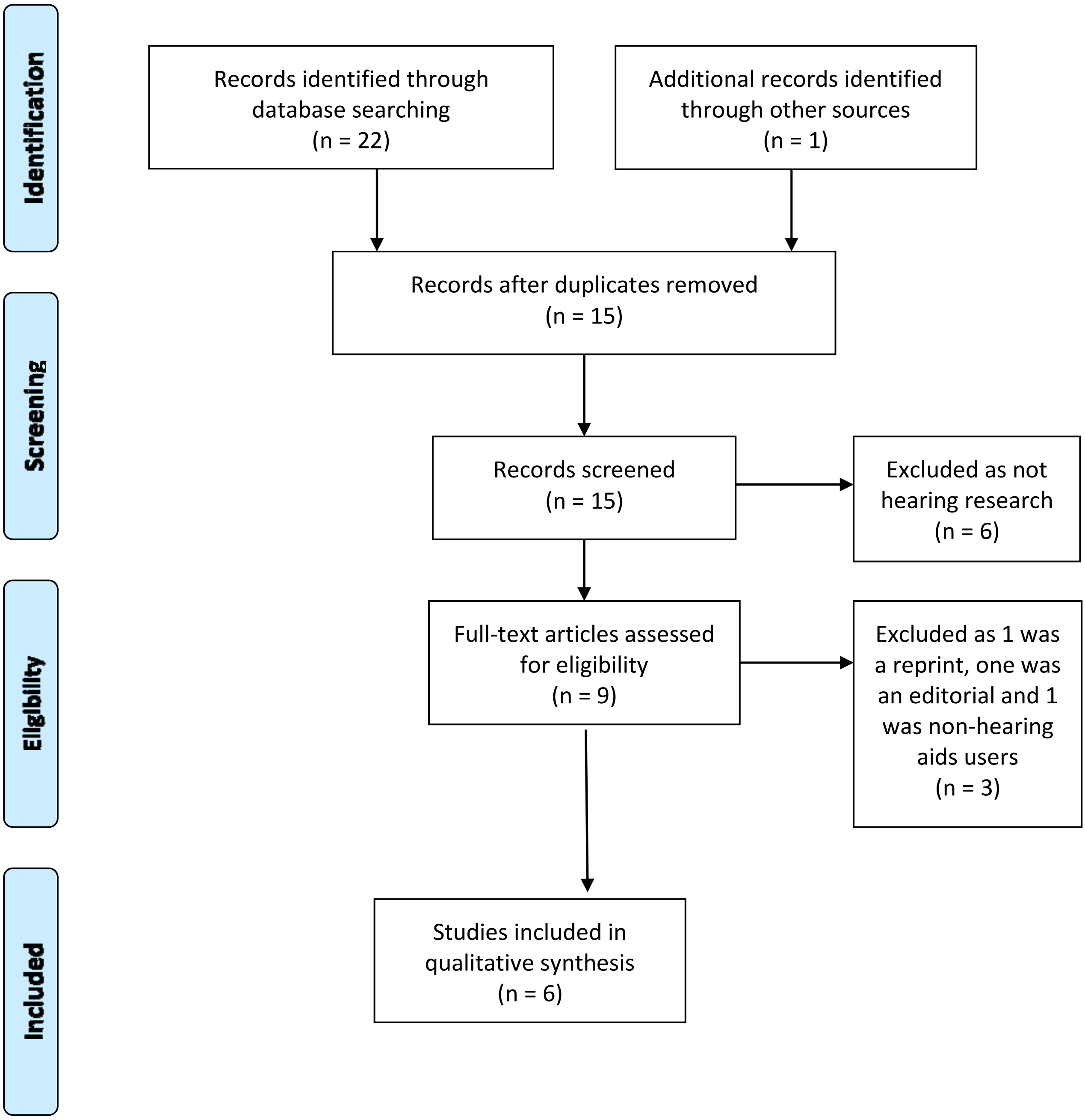


Figure 1. PRISMA Flowchart to identify research papers in which LACE has been used with individuals with hearing aids

Table 1
Review of LACE in published literature

Reference	Design	Training/Control	Outcome Measures	Key Results	Discussion
Sweetow and Sabes (2006a)	Random Crossover design	65 participants (9 without hearing aids) in a crossover design (28-85 years) with Group 1 completing 4 weeks training and Group 2 as a control. Measures were administered at 4 weeks for all participants. Group 2 then also undertook LACE training and completed the assessments at the end of the training.	On-task assessment was QuickSIN, time-compressed speech, competing talker, auditory working memory and missing word tasks (a test of semantic knowledge to fill in blanks in a sentence). Off-task assessment was Hearing in Noise Test (HINT), Listening span and a Stroop task. Two questionnaires were used: Hearing Handicap Inventory for the Elderly (HHIE) and the Communication Scale for Older Adults (CSOA).	LACE-trained participants showed significant improvements for all on-task improvements between the start of training and the end of training. For off-task assessments, on the QuickSIN the LACE-trained group showed an improvement of 1.5dBSNR ($p < .001$ within-group) and the control was 0.2dBSNR) at 70dB Speech. The HINT showed no significant differences between the improvement in the LACE-trained and the control group (0.9dB ($p = 0.43$ within group) compared to 0.4dB SNR change respectively). For the Listening span there was a change in the trained group of .5 sentences ($p < .05$ within-group change) compared to 0 in the control. The Stroop task demonstrated an improvement in the Stroop interference timing performance of 3.1 seconds ($p < .05$ within-group change) compared to the control at 0.7. Both of the questionnaires demonstrated significant improvements for the LACE group (a change of 7.5 points in the HHIE ($p < .005$ change within-group) and of .14 points in the CSOA-s ($p < .005$ change within-group), a subtest of the CSOA, and also an improvement in the CSOA-a with a .06 ($p < .005$ within-group) improvement. The control group showed no significant change between test sessions. There was no information published on the standard deviations of these tests	Results were reported for significant differences between the start of training and after training at 4 and 8 weeks. The differences were reported as group's scores pre and post training (LACE and Control) rather than between the groups. The overall conclusion was that LACE training provided significant benefit on a number of the tasks when compared to the control group.

Sabes and Sweetow (2007)	Analysis of trends. There was no real design just a modelling of who did and who didn't benefit most from LACE training.	The same 65 participants as used in the above study.	The dependent variable was the recorded improvement in LACE performance. The multiple regression correlative analyses were performed on all baseline measures, as listed in the research above. The outcome measures were the strength of relationships of the variables at baseline and their association with the predictive value of undertaking the training.	The following results all indicated a significant relationship at p < .05. Baseline test performance and training compliance indicated that younger adults took longer to complete the training (r = -.036) and that greater loss was related to increased probability of completing the training early or on time (r = -0.48). Participants with lower baseline scores were also more likely to perform their training earlier (r = -.038). Speech-in-babble improvements were greatest for those with the largest degree of hearing loss. Poor QuickSIN performance at baseline was associated with improvements after training (r = .44), and the HINT baseline score was associated with improved QuickSIN score (r = .42). The poorer the performance at baseline testing, the greater the improvement. In attempting to develop a model to predict which patients would benefit the most from LACE, severity of hearing loss, CSOA-s (r = 0.31), Speech-in-babble performance (r = 0.31), competing speech (r = 0.43) time compression (r = 0.28) target word (r = -0.26) were examined. The HHIE questionnaire on hearing handicap also predicted overall improvement (r = .48)	This research was designed to explore which patients may be suitable for LACE training. Those with the poorest scores on baseline tests (speech-in-noise and competing speech), those with greater degree of hearing loss and those ranking their performance as poor were the most likely to demonstrate the greatest improvement overall. However, there is much variability, and some patients who demonstrated small gains gave positive subjective reports.
Martin (2007)	Observational study	Patients self-selected whether to perform LACE training. Outcomes from 625 patients were analysed (173 undertook LACE training and 452 did not). There was no information on patient demographics.	Hearing aid return rate was the only predictor of success. Logistic regression was done predicting hearing aid returns.	3.5% of patients that undertook LACE training returned their hearing aids compared to 13.1% who did not perform LACE. On logistic regression, only hearing threshold and LACE were significant predictors of hearing aid returns.	Whilst not a paper measuring hearing or cognitive performance, LACE training did have a significant effect on hearing aid users persisting with aid use.

Chisolm et al. (2013)	Data from LACE training database.	50 LACE-trained hearing aid users (50-85 years old, mean 66.4, SD 7.6) Tests were administered pre and post training.	Measures were the 5 LACE training tasks tested for on-task outcomes, (speech-in-babble, time compression, competing speaker, auditory memory and missing word). Off-task outcomes were measured using the WIN test, Rapid Speech (processing speed), (inhibition control) (NU-6 45% and 65% compression), Competing speaker (NU-20 test; processing speed), Digit Span (modified version of WAIS-III; working memory), Linguistic context (using revised speech perception in noise test with 25 low predictability words and 25 high predictability words. The other variable was compliance with training and the effect this had.	84% (42 out of 50) users completed all the training sessions. The results for the 5 LACE tasks all indicated a significant improvement over the training period for all the on-tasks. The training period was split into 4 quarters. Quarters 2, 3 and 4 were all significantly better than the 1st quarter measure. One-way repeated measures ANOVA indicated that the main effect of the quarter was statistically significant ($p < .000$). Maximum improvement was between 10 and 20 sessions. For off-task, significant improvement between the compliant training group and non-compliant was seen for 4 out of 7 tasks. For compliant LACE trainees (completed 20 sessions), these were words in noise, (LACE group (L) 8.9, non-compliant (NC) 9.5, p < .05), both compressed speech measures (L 75.4, NC 65.3 p < .05) and L 55.1, NC 46.6, p < .01) competing sentences (L 4.6, NC 6.2, p <.001). There was no significant improvement for non-compliant trainees (<20 sessions)	There was no control group for comparison as results were measured before and after training. On-task results were compared to the original Sweetow and Sabes (2006) research and were similar. Non-compliance with training resulted in less improvement on the off-task measures.

Olson, Preminger, and Shinn (2013)	Randomised controlled trial with repeated measures.	29 participants between 50 and 81 years were recruited: 8 new hearing aid users undertaking LACE training, 7 new hearing aid users with no training and 14 established hearing aid users undertaking LACE training. LACE training was provided by DVD over 10 sessions instead of the traditional on-line/CD based training.	QuickSIN test, compressed speech at 45% compression rate, Synthetic Sentence Identification (competing speakers; SSI) and 3 questionnaires (IOI-HA, the IOI-AI and the SSQ).	After controlling for hearing loss, repeated-measures ANCOVA results were not significant across time (*p* = .056) or between groups for the QuickSIN. Compressed speech also did not demonstrate a significant interaction between group/time, although there was a significant interaction for the SSI group*time interaction, *p* < .05, with post hoc analysis indicating the difference lay between the control group and the trained groups. The SSQ indicated a significant change over time (p < .05), but the interaction of group*time was not significant.	Although there was no significant interaction of training group and time on QuickSIN performance, there was a large effect size for LACE training for this test (Cohens d = 1.5) after two weeks of training for new hearing aid users compared to 0.8 for the control group. The same Cohens d of 0.8 was also reported for existing hearing aid users at 2 weeks. The effect size at 4 weeks for new hearing aid users was 1.9.
Saunders et al. (2016)	Randomized controlled trial.	279 hearing aid users (both new and experienced), split into four training groups: LACE DVD, LACE computer, placebo auditory training, and educational counselling. There was no age range given but the mean age of the participants was approximately 70 years.	Words-in-noise (WIN test), Rapid Speech at 45% compression rate, Competing speaker (NU-20 test using NU-6 words), Digit Span (modified version of WAIS-III), Linguistic context (using revised speech perception in noise test with 25 low predictability words and 25 high predictability words) and the APHAP questionnaire.	No statistically significant interactions or main effects were found for use of LACE.	The outcome measures were all selected to demonstrate possible performance transfer to similar tasks (near transfer) for both hearing and the cognitive measures. No transfer gains were seen.

5.1.3. Critique of LACE studies.

There are a variety of weaknesses in the majority of the studies in this review, which are explored below. Further, a general weakness of all of the research in this area is their duration, as short-term studies such as these may not necessarily demonstrate changes over a short period of time. As some potential changes, resulting from LACE or other training programs may be subtle these may only detectable in longer-term studies. The principle being that any new phonological representations may require time to become fully developed.

A specific weakness of many of the reviewed LACE studies is that training compliance has not always been closely monitored for the intervention pathways. This was the case for LACE itself in some cases, as LACE DVD training cannot be monitored (Olson et al., 2013; Saunders et al., 2016a), although the on-line version of LACE was able to be recorded via the training portal. Further, for some of the research, the control group was simply passive (Olson et al., 2013; Sweetow & Sabes, 2006b). This lack of an active control potentially biases any end results. The control group needs to have the same expectation of improvement as the experimental group; though this is achieved better with an active compared to a passive control, even an active control arrangement may not completely achieve parity in expectation (Boot et al., 2013). The current research in this chapter addressed this by having online training monitoring for the LACE group and requiring the active control group to complete exercises while listening to CDs. Their logbook completion was evidence of their training. It involved the same time investment and auditory exposure as undertaking the LACE training.

Further, the analysis in the original Sweetow and Sabes (2006a) research had reported differences in performance on a series of outcome measures before and after training alongside reporting on a control group’s performance. This was instead of employing a mixed methods 2 x 2 ANOVA design. Such an approach fails to properly explore whether there were actual significant differences between the groups. This same research also used a wide age range, from 28 to 85, and all participants either were established hearing aid users[39] or did not use an aid at all but expressed a problem hearing in background noise. This mismatch of patients between groups makes between-groups analysis significantly less robust. There was also no information on types of hearing aids worn or hearing aid fitting protocol, again compromising interpretation of the results. To address these issues, in the present research, all patients had bilateral fittings and were fitted to the same algorithm. Also, mixed-methods ANOVAs were used to examine the interaction of time and pathway within and between groups to explore if particular training pathways offered different (potentially more successful) outcomes. Patient recruitment was also much more stringently controlled to ensure less variation between participants.

Research by Olson et al. (2013) and Saunders et al. (2016a) recruited only hearing-aided participants, who in the former study were aged 50 to 81 years and in the latter had an average age of 68 years, although the authors failed to report the age range. There was no control over the type of hearing aid used by patients in the Olson et al. study, as some had in-the-canal aids, some in-the-ear and some behind-the-ear open fittings. The majority of aids used in the Saunders et al. study were from one manufacturer, but this was not exclusively the case. Hearing aid fittings for both studies were all verified against the NAL NL1 fitting algorithm (Byrne, Dillon, Ching, Katsch, & Keidser, 2001), and this was used to quantify speech audibility. Both studies used both new[40] and experienced hearing aid users.[41]

Neither the Olson et al. nor the Saunders et al. study demonstrated a significant improvement in speech-in-noise performance following LACE training, although the Olson et al. research did identify a trend towards significant improvement for the LACE-trained group on the QuickSIN compared to a passive control. In the Saunders et al. (2016a) research, there was a much larger number of participants recruited, 263 in total, and they used four groups, including two control groups (one active placebo trained and one counselled on hearing aid

---

[39] 85% were bilateral hearing aid users, and the remainder were unilateral.

[40] Had worn aids for >4 weeks but < 6 months.

[41] In the Olson study they had used aids for > 2 years and in the Saunders study for > 6 months

use). It could be argued that unlike the Olson study, the better-designed Saunders et al. research did not report any trends towards significant differences because the activities of the control groups offered more stimulation.

A further complication of the Olson et al. and Saunders et al. studies was in their definition of a new hearing aid user. They both considered individuals who had been using hearing aids for greater than 4 weeks but less than 6 months as new users. This wide range would therefore include some users who had acclimatised to the sound of their hearing aid (Dawes & Munro, 2017; Munro & Lutman, 2003; Yund et al., 2006) and so potentially would be less likely to derive maximum benefit from training, as in some cases, new sound representations may already have been established.

As summarised in Table 1, one study demonstrated significant improvements between the start point and various assessment time points after undertaking LACE training (Sweetow & Sabes, 2006a). They report that the LACE-trained group performed significantly better on both speech based and cognitive tasks than a control group that undertook no training. The reported improvements in assessment performance, however, were measured within groups and not between groups. The two other pieces of research that involved assessment of outcomes (Olson et al., 2013; Saunders et al., 2016a) did not show significant improvements in the LACE-trained group over alternative interventions for the majority of the cognitive based tasks. Indeed, the Saunders et al. study reported no statistically significant main effects for any LACE--trained outcome measure compared to a placebo-trained or a control group.

Summarising the above, there is no general agreement that LACE-trained participants gain an advantage over non-trained or placebo-trained groups. Many of the methodological weaknesses of previous studies described above were addressed in the research for this chapter. In addition, there was a particular focus on any differences between ‘on-task’ performance and performance on tasks involving near and far transfer. The previous studies reviewed here focused more on general performance outcomes and less on the opportunity for the release of cognitive resources as speech discrimination improved. Again, this opportunity is a key hypothesis for this chapter in that as patients get used to their hearing aids and discriminate speech better there is a resultant decrease in the reliance on cognitive support, allowing for the reallocation of these freed resources to other cognitive functions such as working memory and encoding of information into secondary memory.

5.1.4. Aims and hypotheses

As previous chapters have explored whether cognitive and auditory training offer an individual advantages in improving their speech-in-noise perception, this chapter tested the theory that offering a combination approach of the two types of training to newly fitted hearing aid users may offer success. The computer-based Listening and Communication Enhancement (LACE) training package aims to improve an individual's hearing ability using such a combination approach. The primary aim of this chapter is to determine not just whether LACE training results in 'on-task' improvements for newly fitted hearing aid patients, but also whether any improvements lead to improvements in deeper processing (the encoding of auditory information into secondary memory) and the release of cognitive resources. This will be tested by assessing whether recognition of auditory material and working memory performance improve after LACE training. If improvements in speech perception do release cognitive resources previously used for discrimination purposes, this should allow for improved working memory span and secondary encoding.

An additional aim of this section is to assess which (if any) of three care pathways (LACE training, placebo training and no training), after hearing aid fitting and training, is associated with an increase in participants' self-ratings on a 12-item questionnaire that assesses different listening skills,. Furthermore, the self-rating judgements can be compared with the individuals' performance on the different speech-in-noise tasks. It is anticipated that participants on all pathways may self-rate as improving, as undertaking any training or receiving a new intervention (such as hearing aids) may allow for the anticipation of success (Crow et al., 1999).

If the gains are real, however, then any change in self-rating performance will correlate strongly with improved task performance post-training.

The final aim is to establish which cognitive functions explain the variation in speech-in-noise performance for hearing aid users, after controlling for age and hearing loss. This is also expected to reveal whether the relationship of cognitive functions with speech-in-noise perception differs from those noted in the portions of this thesis that recruited unaided participants with normal (or near normal) hearing. It is anticipated that older patients with hearing loss will utilise more of their remaining cognitive resources to support their listening, as although hearing aids give patients better levels of audition, the aids are still amplifying these speech sounds to an inner ear that has undergone physiological changes due to age (presbycusis). As a result, although speech may be heard, the broader tuning peaks on the basilar membrane may provide reduced clarity, such that phonological mismatches occur with previously stored information, necessitating explicit cognitive processing to support discrimination. This would be exhibited by a greater range of cognitive functions demonstrating a relationship with speech-in-noise performance compared with the ranges from other chapters where participants' hearing levels were normal (or close to it).

Two types of speech-in-noise tests were employed: adaptive speech in babble and constantly challenging speech in speech-shaped noise. The different cognitive support mechanisms for listening under these different conditions can thus also be explored. The cognitive functions examined are working memory, information processing speed (a response latency task), and control of inhibition. To control as much as possible for individual variation, variables such as type and degree of hearing loss, type of hearing aid, length of time of hearing aid usage, and fitting strategy have all been standardised.

The hypotheses being tested to achieve the above aims were as follows:

H5.1. If auditory and cognitive training, as delivered by LACE, correctly target the areas needed to improve speech-in-noise performance, then the LACE-trained pathway group will perform the speech-in-noise tasks significantly better at the second assessment compared to the other training pathways.

H5.2. If speech discrimination improves following training, then there will be a subsequent release of cognitive resources that allows for significantly improved performance on auditory-based memory span tasks (both simple and complex) and a speech recognition task of previously delivered auditory information.

H5.3. If LACE training offers significant advantages over placebo and no communication training, then there will be a significant difference in self-rating improvements between the pathways, both in general and for specific listening situations such as competing noise, sound quality, spatial awareness, and effort.

H5.4. If improved performance affects perceived hearing efficacy following hearing aid fitting, then the self-rating improvements noted by the participants in each of the pathways will correlate with the participants' speech-in-noise task performance.

H5.5. If cognitive functions assist with discriminating speech-in-noise for newly fitted hearing aid users, then the tasks assessing the different cognitive functions will significantly correlate with speech-in-noise performance.

H5.6. When testing H5.5, different sets of significant relationships will be identified for the two types of speech-in-noise tests, as steady-state speech noise and fluctuating noise may necessitate the support of different cognitive functions.

H5.7. If the speech performance variation is due to cognitive functions themselves and not raised hearing thresholds or the effects of age, then once age and hearing level have been accounted for, the tasks assessing the cognitive functions will significantly add to the explanation of speech performance variation.

5.2. Methods

5.2.1. Ethics

The East Midlands - Leicester Central Research Ethics Committee approved the research (Reference Number 15/EM/0496) after Aston governance had been granted.

5.2.1.1. Grants

An initial pump priming grant to explore the methods used from Aston Research Centre for Healthy Ageing (ARCHA) and a BSA grant (Project Number 60292) supported the research. The grants were used to purchase LACE training CDs, workbook training materials and pay for participant travel expenses.

5.2.2. Participants

Since there have been no studies with this exact design (as no previous studies have used a three group design of newly fitted hearing aid participants), a sample size calculation using G power (Faul et al., 2007) was performed based on research by Olson et al. (2013). As the primary aim of this chapter was to explore for significant differences within and between different intervention pathways after hearing aid fitting, the power calculation was based on the sample size necessary to demonstrate such a difference. The LACE research by Olson used three groups and reported a large effect size for speech-in-noise performance after a hearing aid acclimitisation period (a control group) for newly fitted hearing aid users of 0.8. For the LACE trained group the effect size was nearly double at 1.5. Therefore, when considering the difference in effect size between these two groups (ignoring the established hearing aid user group recruited by Olson et al. as no established user group was recruited for this present study) the difference is quite marked. As this present study was not a replication of Olson et al. and to be conservative, a medium effect size (0.5) was used for the calculation to approximate the difference in performance between a trained and an untrained group. The sample size necessary for a within-between mixed-methods ANOVA, with a power of .80 and an alpha level set at 0.05, was 30. With three groups, this meant 10 participants were needed per group. Participants were recruited from the Audiology department at the Great Western Hospital, Swindon. Four participants consented to join the research at their initial audiological assessment appointment, and the rest of the participants were recruited whilst awaiting a hearing aid fitting appointment. In order to identify these participants, the records of 1,850 patients were reviewed. The necessary inclusion criteria were an established mild/moderate sensorineural hearing loss[42] and having opted for bilateral slim-fit hearing aids. In total, 236 invitation letters were sent, and 37 patients met all the necessary criteria and consented to participate. The goal was to recruit a total of 60 participants, as the effect size had been estimated from research involving two groups and without a calculated *f* value. Having extra participants would also allow for dropouts, which were anticipated to be at 20% of total recruits. Stratified randomisation was used (Suresh, 2011) to allocate participants to groups whilst keeping an age balance between the groups, the same as in Section 3.2.2. This was performed using the computer software program (GraphPad). The training groups are summarised below and the detail of what their training involved is given in Section 5.2.6.

- Group 1 pathway. LACE training (LACE)
- Group 2 pathway. Structured Listening Activity (SLA) (Active Control)
- Group 3 pathway. Control (Standard Acclimitisation) (Control)

Unfortunately, recruitment of new participants had to cease after 37 had been recruited as no more LACE registrations were possible. This led to an imbalance in the groups at the point where the research ended. One participant dropped out of the research and one found the tasks too challenging to perform. This left 35 participants who had been randomly allocated to the pathways up to the point where LACE would not accept new registrants. The change in LACE registration and training is explained more fully in Section 5.5. The participants were divided as follows: 11 LACE, 13 Structured Listening Activities (SLA) and 11 Control.

---

[42] As defined by the British Society of Audiology, averaged across the pure tone frequencies of 500, 1000, 2000 and 4000Hz (BSA, 2011). Mild loss is classified as between 20-40dBHL and moderate as between 41 -70dBHL.

The average time between participant appointments is recorded in the Table 2 along with the bilateral PTA hearing thresholds averaged across 500-4000Hz. The final education level of the participants is recorded and calculated as 1= GCSE, 2 = A level, 3 = BSc Degree, 4 = Postgraduate.

The ability to record compliance with LACE training was unfortunately affected due to the servers on the Neurotone website ceasing to log training sessions after January 31, 2017, and no new participants being able to register after this date. This also had an effect on the ability to recruit any new participants into the LACE pathway.

The patient demographics are given below in Table 2.

Table 2
Patient demographics

Training Pathway	**Mean Age**	**Age Standard Deviation**	**PTA average 500-4000Hz**	**Final education level**	**Gender**	**Average Time between appointments (days)**	**Total**
LACE	69.18	6.63	32.14	2.18	7 Men 4 Women	44	11
SLA	66.92	7.33	36.00	2.38	8 Men 5 Women	46	13
Control	68.73	5.64	32.27	2.09	7 Men 4 Women	45	11

Key: LACE = Listening and Communication Enhancement, SLA = Structured Listening Activity.
Note: There was homogeneity of variances between the groups as assessed by Levene's test (all at *p* >.05). One-way ANOVA testing between the groups of age, PTA, and time between appointments were all at *p* >.05, indicating no significant differences between the groups at baseline.

Below in Figure 1 are the average thresholds for each frequency (with the standard deviations) for the participants that took part in the research for this chapter. No participant had a threshold at any frequency greater than 65dBHL.

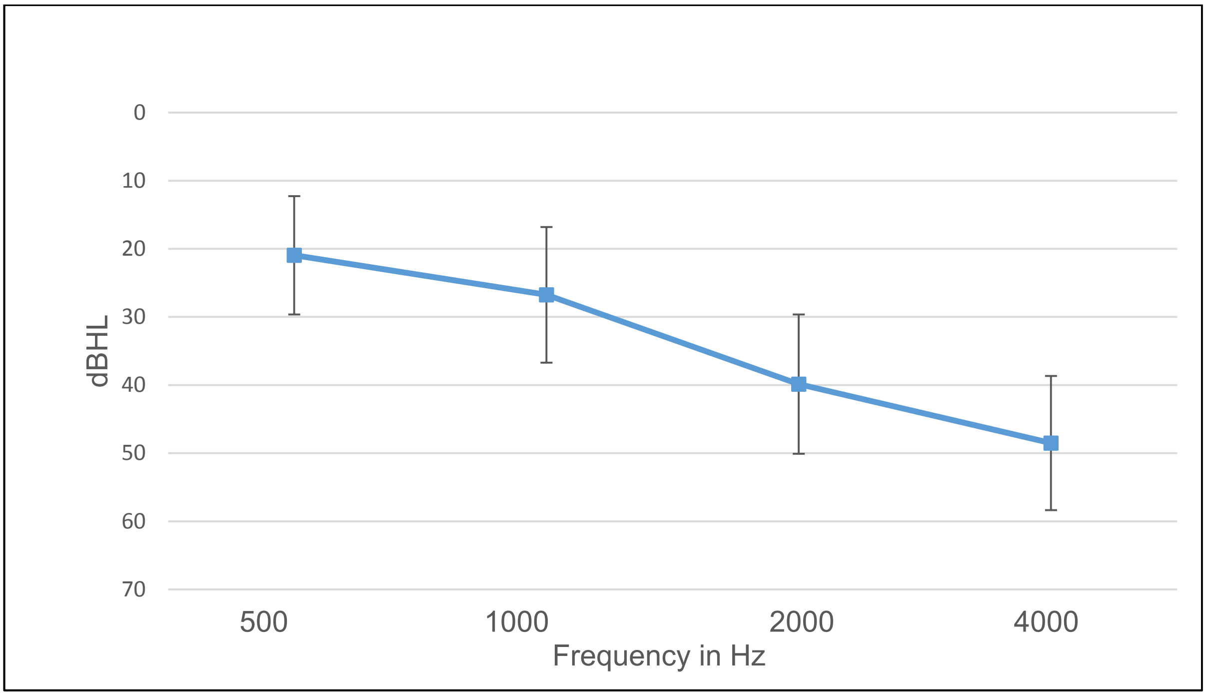


Figure 2. Average pure tone thresholds at each frequency with standard deviations

Further inclusion criteria for the research were that the participants needed to have a computer and access to the internet. Although randomised allocation meant that not all participants would be in the LACE training group, all who enrolled needed to have access to the same facilities. They also needed to be free of colour blindness, to have English as their first language and to have no reported memory issues.

5.2.3. General design

The study was designed to analyse participants' performance differences within and between three care pathways. It was designed to ascertain if one of the care pathways offered a significant advantage over the others. This was, therefore, analysed using a two-way mixed-method ANOVA design (3 x 2 design). The between-participants' factor was the pathway followed: LACE training, structured listening activity (SLA) and normal acclimitisation (control). Participant performance of the two speech-in-noise tests, a BKB in speech-shaped-noise, and a QuickSIN test in babble served as the dependent variables. The cognitive functions measured were assessments to explore working memory, speed of information processing, recognition and the control of prepotent inhibition[43]. These were all analysed before and after the training period.

There were two factors, care pathway (three levels) and time (two levels, baseline and follow-up). All of the variables tested were explored for general difference, the effect of time and for a group*time interaction. If any significant differences were indicated by the ANOVA, post hoc analysis would be performed to analyse where these differences lay.

The second aim of the research was correlational research to explore the demographic and cognitive factors that explain the variation in speech-in-noise performance for individuals with hearing loss. Hierarchical regression modelling was subsequently used to determine if any cognitive factors additionally explained speech-in-noise performance differences, after partialling out age and hearing level.

The study was further designed to evaluate the methods used in the research, including ease of participant recruitment and pathway intervention compliance.

5.2.4. Test materials and hearing aid fitting

Much of the testing material described below was delivered through a loudspeaker set-up for the research in this chapter as the patients were wearing hearing aids. The configuration of this delivery system is given in procedures in Section 5.2.5.

5.2.4.1. Audiometry

Participants had already undergone pure tone audiometry performed by qualified audiologists at the Great Western Hospital. A diagnosis had been given, and the patients had decided that they wished to have bilateral hearing aids fitted.

---

[43] Inhibition control in 'real-time' was not assessed in the research for this chapter as the QuickSIN used speech in meaningful babble which was already assessing to some extent this 'real-time' inhibition control

Otoscopy was performed to identify if real ear measurements could be performed on the participants (BSA, 2016).

5.2.4.2. Hearing aid fitting

All participants had bilateral Oticon Synergy aids. This type of slim fit behind-the-ear hearing aid has been reported by Kochkin (2011) to demonstrate an increasing trend in patient satisfaction. All the aids were fitted to the NAL-NL 2 fitting algorithm and all patients had real ear measurements performed by the researcher to verify the amount of amplification being delivered per frequency. All fittings were within the tolerances of the calculated NAL-NL2 targets of the participants' hearing loss, as delineated by MHAS[44]. The hearing loss had to be suitable for a slim tube fitting.

Once the fitting had been verified and the participant was instructed on hearing aid handling and acclimatisation, the baseline measures outlined below were performed with the hearing aids in place.

5.2.4.3. Discrimination tasks

QuickSIN test

The material for the QuickSIN (Killion et al., 2004) test was the same as in Chapter 3 (Section 3.2.4.2), utilising the same sentence lists and scoring method.

BKB test

The BKB (Bench et al., 1979) sentences and scoring as explained in Section 2.2.4.2 were used. The background speech-shaped noise was kept at a consistently difficult SNR of -4dB. The easing of the SNR compared to other chapters in this thesis (where it was set at -8dBSNR for normally hearing participants) acknowledged that with a hearing loss, and hearing aids, the -8dBSNR was too challenging.

---

[44] 5 to -5dB from 250Hz to 2000Hz and within 8 to -8dB at 3000 and 4000Hz.

5.2.4.4. Auditory sentence span (LSPAN) and recognition task

The testing material and procedure was the same for the LSPAN as in Section 4.2.4.3.1. The background noise was speech-shaped, and the SNR was set at +10dB. The reason for the high SNR was to endeavour to make the speech reasonably discernible but to ensure there was some interference to make the task more complex than a simple memory listening task. The participant was asked to recall the last word of a sentence, and the number of sentences and hence words to be remembered was increased and then staircased (see 4.2.4.3.1.)

5.2.4.5. Auditory digital span

A forward and reverse digital span task presented auditorily was employed. The patient heard digits spoken by a female speaker at the calibrated speech level of 65dBSPL. At the end of a sequence of numbers the patient entered the numbers they heard into a computer in the order they heard them. If they got them all correct the list length was increased, if they made an error the list was shortened. Using an adaptive scoring algorithm the length of numbers they were able to get correct 50% of the time was calculated.

After performing this in a forward paradigm, participants also performed the task in reverse order such that 2, 4, 7 would be repeated as 7, 4, 2. The same ‘staircasing’ method was used to compute the reverse digit span as for the forward digit span described above.

5.2.4.6. Stroop Test

The same Stroop test was employed as used in Section 2.2.4.4.

5.2.4.7. Auditory recognition task

This test was the same as employed in Section 4.2.4.4.

5.2.4.8. SSQ12 questionnaire

The original Speech, Spatial and Qualities of Hearing scale (SSQ; (Gatehouse & Noble, 2004) involves 49 questions aimed at investigating an individual’s self-rating of hearing performance in a variety of listening situations. The SSQ endeavours to reflect listening performance in real-world situations. As the title implies, there are three main components:

1. Listening to speech in different situations

2. Awareness of where sounds are coming from (spatial)

3. Sound quality.

There is a fourth, smaller component, which concerns listening effort. The original 49-item SSQ is lengthy so a briefer version, the SSQ12 was developed (Noble, Jensen, Naylor, Bhullar, & Akeroyd, 2013). The SQ12 covered the same areas as the full SSQ and on evaluation, Noble reported that the SSQ12 provides results similar to the SSQ49.

Participants were asked to rate their performance in a range of listening conditions on the SSQ12 on a scale from 1-10. For example, the first question reads: ‘You are talking with one other person and there is a TV on in the same room. Without turning the TV down, can you follow what the person you're talking to says?’

The rating scale ranges from ‘not at all’ to ‘perfectly’. When participants returned for subsequent appointments, they completed the SSQ12B, which asks the same question above, having the participant rate the following statement: ‘Comparing your ability now with your ability before getting your hearing aid/s’ using the options of -5 to +5 through, with 0 in the middle. A rating of -5 means ‘much worse’, 0 is ‘unchanged’ and +5 is ‘much better’.

5.2.5. Procedure

All the testing was carried out in a well-lit (lux 340-380) soundproof room at the Great Western Hospital, Swindon. Patients were given detailed instructions about what was expected of them and were given opportunities to ask questions and have breaks as the test session proceeded.

All the testing and the SSQ12 questionnaire (apart from the real ear measurements used to verify the hearing aid fitting) were performed at baseline and at follow-up after the training pathway. Pure tone audiometry was not performed as part of this research as it had been previously performed before the hearing aid fitting by audiologists at the Great Western Hospital.

The order of tasks after the initial ear examination and filling in of the SSQ 12 questionnaire was: 1. Hearing aid fitting, 2. Speech discrimination performance (BKB and QuickSIN), 3. Working memory tasks (LSPAN and forward and backward digit span), 4. Control of inhibition (Stroop interference measure) 5. Processing speed (Stroop congruent speed measure) and 6. Response latency (visual reaction time task). The same order of testing was used before and after training but with no fitting of hearing aids at the patient’s second visit although their hearing aids were checked for their performance before the second set of assessment tasks was undertaken.

Where the material was delivered auditorily for the discrimination tasks, the LSPAN and the digit spans, then the testing apparatus was a Samsung Personal Computer R530 and an Edirol UA-25 audio capture box. The sound output was played through an Edifier loudspeaker. The Edirol soundbox output was calibrated using a Brüel and Kjær 2250 sound level meter to deliver the speech at 65dBSPL. The patient sat 1 metre in front of the loudspeaker at a set calibrated distance. For all the tasks, the patients wore both their hearing aids at their prescription matched target settings.

For the Stroop task where the material was not auditorily presented the test was administered using a program written by Millisecond software (Inquisit) and presented on a Toshiba Satellite Pro laptop computer. The auditory recognition task procedure involved the patient being given two sheets of paper with the correct sentences on that they had heard during the LSPAN in amongst distracting sentences and they selected (ringed) what they remembered as being the presented sentences.

5.2.6. Training pathways and training regime

As discussed in the Section 5.2.2, participants were randomly assigned to one of the three care pathways.

In addition to the training to be undertaken, participants were also asked to acclimatise to their hearing aids as directed by the audiologist and wear their hearing aids as they went about their normal lives. The overview of the testing and training undertaken is given in Table 3.

Table 3
Overview of assessments and training undertaken by patients on the three training pathways

Group	Hearing aid fitting	Testing with hearing aids (Baseline)	Training: 4 - 8 weeks	Testing with hearing aids (1 month)
1. LACE	Bilateral hearing aids fitted using real ear measurements to NAL NL2 fitting targets	Listening performance, inhibition, memory span and recognition tasks. SSQ 12 Questionnaire	LACE: 20 sessions over 4 weeks	Listening performance, inhibition, memory span and recognition tasks. SSQ 12 Questionnaire
2. Structured Listening Activities: Active Control (SLA)	Bilateral hearing aids fitted using real ear measurements to NAL NL2 fitting targets	Listening performance, inhibition, memory span and recognition tasks. SSQ 12 Questionnaire	Structured listening activities: 20 sessions over 4 weeks	Listening performance, inhibition, memory span and recognition tasks. SSQ 12 Questionnaire
3. Control Pathway	Bilateral hearing aids fitted using real ear measurements to NAL NL2 fitting targets	Listening performance, inhibition, memory span and recognition tasks. SSQ 12 Questionnaire	No structured activity	Listening performance, inhibition, memory span and recognition tasks. SSQ 12 Questionnaire

5.2.6.1. Introduction to LACE training

LACE was originally designed as a 20-session training program to be completed by a listener (normally a hearing aid user, but not exclusively) over a 4-week period. A researcher/clinician or the participant visits Neurotone's (makers and distributors of LACE) website and uses a security code to register the LACE training software. This allows both the participant and the researcher/clinician to log in to view training progress. The LACE program is on a compact disc (CD), so participants needed a computer with a CD player and access to the internet to be able to log and upload their training. LACE training is adaptive in that it determines at what point an individual is finding the task they are undertaking challenging and focuses around this point to encourage improved performance. For each of the training areas given below, LACE will adapt to make the training challenging.

LACE training is split into different constituent parts. The largest focus area for training is detecting speech in babble (30%). This is similar in nature to the QuickSIN test. The trainee starts at a signal-to-noise ratio of 10dBSNR and then, following an adaptive paradigm depending on whether or not they can understand what was said, the task becomes more or less difficult. The step size changes shift from large to smaller ones as the test proceeds.

Time-compressed speech makes up 20% of the training. This is an attempt to train listeners' processing speed ability. The initial speed compression is 85%, meaning a sentence is time-compressed to 85% of its original time. The training adapts depending on performance, so speech will either speed up or slow-down in set increments.

Competing speaker is another task undertaken by LACE trainees (20%). This task trains real-time inhibition and task switching to a lesser extent, thus exercising executive control. Three voices are used: a man, a woman and a child. Listeners are told to target one of the two competing speakers and ignore the other. The procedure is adaptive and is similar to the speech-in-babble task where initially one speaker is louder than the other by 10dB at the start of the training, changing depending on performance in the same way the speech in babble task adjusts. For the above three tasks, when starting a new training session, LACE would start the task two levels above where the last session was finished so working at challenging levels would occur more quickly for subsequent training after the initial session.

The last two tasks are designed to train working memory through span tasks and processing speed (alongside the use of linguistic and contextual clues) using missing word exercises. These each make up 15% of the overall training. For the former, the listener is given a target word in a sentence and then needs to respond with the word that came directly before the target. Adapting the task to make it more challenging is done by using longer sentences and by using two target words. For the latter task, listeners have to make judgements on what would have been an appropriate word to fill in in a spoken sentence. They are given options to choose from and asked to do the task as quickly as possible. There is no adaptive element to this training.

Whilst undertaking the LACE training for many of the tasks, individuals will be utilising and improving their top-down processing as well as the skills themselves which could allow for additional benefit from the training.

5.3. Results

5.3.1. Intervention pathway exercise compliance

Before the results were analysed, intervention compliance for the three intervention pathways was examined.

Of the 11 LACE participants that were registered, eight completed all 20 sessions,[45] and the remaining three completed between 10 and 19 sessions. As the greatest performance improvements delivered by LACE are achieved after completion of the first 10 sessions (Sweetow & Sabes, 2006a), the results from all 11 participants were included in the data analysis.[46] For the SLA pathway, the 13 participants all completed at least 19 of the 20 sessions. One participant could not complete one training session due to a faulty track on their CD. Compliance with training was recorded in the specially created exercise-training book, and all training books were returned to the researcher with the required information filled in, indicating that the listening activities had been completed.

The Control, LACE, and SLA groups all attended a hearing aid follow-up appointment, which took place immediately prior to their second assessment session. All were checked for hearing aid handling and their use of the aids.

---

[45] Three participants had to demonstrate this by bringing in their personal computers to show that the training had been completed.

[46] The replacement training offered by LACE (LACE Online) now suggests an intervention of 11 sessions.

5.3.2. Descriptive statistics

The general descriptive statistics for participant demographics are given in Table 2 in Section 5.2.2. above. There were no significant differences found between the groups in terms of any of the variables assessed at baseline when tested by a series of one-way ANOVAs.

The mean scores for Assessment Sessions 1 (baseline) and 2 (post-training) are given below in Table 4. For some measures, a decrease in value indicates improved performance, such as on the QuickSIN where a smaller number indicates that a listener can still discriminate speech even when it is only just above the noise, meaning they can hear at a reduced signal-to-noise ratio. The measures for reaction time and processing speed are similar in as much that the smaller the value, the quicker the reaction time, hence the better the performance.

The clearest difference in performance was the change observed in participants' QuickSIN scores. The LACE-trained group exhibited a comparatively large change, indicating that they improved on this task more than the other pathway groups did.

Table 4
Mean scores for the assessment tasks at baseline and at the second, post-training visit.

Pathway group	Assessment	Pre-pathway score	Standard deviation	Post-pathway score	Standard deviation
LACE	QuickSIN	5.54	2.72	4.14	2.12
	BKB	40.30	16.95	49.33	12.07
	LSPAN	3.73	.65	4.09	.83
	Digits Forward	5.91	1.22	5.64	1.29
	Digits Backward	4.55	1.37	5.00	1.00
	Stroop IM	160.00	84.10	175.87	135.30
	Stroop PS	818.78	212.55	752.64	142.43
	SSQ12 overall	5.65	.95	2.61 (Change)	-
	Reaction time	301.85	36.93	284.15	33.13
	Recognition task	5.55	1.44	4.00	1.41
SLA	QuickSIN	5.82	2.94	6.15	2.49
	BKB	35.31	20.58	38.87	14.24
	LSPAN	4.15	.80	4.00	1.16
	Digits Forward	6.23	1.36	6.54	.967
	Digits Backward	4.62	.96	4.92	1.12
	Stroop IM	60.64	55.74	98.96	62.57
	Stroop PS	877.03	164.06	773.31	124.27
	SSQ12 overall	5.60	1.61	2.47 (Change)	-
	Reaction time	312.25	51.78	292.11	43.38
	Recognition task	5.92	1.61	4.38	1.50
Control	QuickSIN	6.55	2.24	6.39	3.31
	BKB	25.70	11.38	33.51	11.33
	LSPAN	3.91	.831	4.09	.701
	Digits Forward	6.00	1.48	6.00	1.18
	Digits Backward	4.55	1.86	4.73	1.20
	Stroop IM	134.85	69.57	146.53	90.27
	Stroop PS	835.34	112.07	784.43	76.08
	SSQ12 overall	6.27	1.45	2.39 (Change)	-
	Reaction time	344.41	82.08	299.01	48.83
	Recognition task	5.36	1.63	4.55	1.86

Key: Stroop IM = Stroop interference measure (Congruent condition - Incongruent condition). Stroop PS = Processing speed, time in msec taken to identify colour in congruent condition.

Analysis plan

To explore Hypothesis 5.1, that there would be significant changes following different care pathways, a mixed-methods analysis of variance was performed. The findings are reported in Section 5.3.4. This allowed for analysis of the effect of any care pathway*time interaction on performance over time in order to examine significant differences between the groups after the care pathway. The exploration of whether improved speech discrimination performance allowed for the freeing up of cognitive resources (H5.2) was explored using the same mixed methods ANOVA method but using LSPAN performance as the dependent variable.

To explore whether LACE training offered a significant advantage in improving self-rated performance (compared to the other pathways), one-way ANOVA testing of the improvement in self-rated scores was employed. This was performed for the overall SSQ12 score and its subsections to investigate Hypothesis 5.3. Hypothesis 5.4 stated that if patient judgements were accurate, then any change in their own self-ratings would be correlated with a change in their speech-in-noise performance. Correlational analysis was used to explore whether there were any significant relationships between the change in speech-in-noise performance and overall and sub-section changes in the SSQ12 scores.

To investigate Hypothesis 5.5, which stated that a range of tested cognitive functions would demonstrate a significant relationship with the speech tests, correlational analysis was performed, the results of which are displayed in Table 7 and discussed in Section 5.3.4. To investigate whether the relationships varied for different types of speech-in-noise (Hypothesis 6.6), the correlations were explored for each type of speech test. Hypothesis 6.7 was investigated using hierarchical regression to determine whether the addition of the significantly correlated cognitive functions contributed significantly to explaining the performance differences on the speech tasks. This was done after partialling out age and hearing loss.

5.3.3. Two-way mixed-methods ANOVAs

Previous methods investigating differences between groups trained using LACE and control groups that have undertaken no or placebo training have used different statistical methods to assess for differences. These have included mixed-methods ANCOVAs with a covariate of hearing level and ANCOVAs with a co-variate of pre-training scores. These are two of the main variables that have been identified as having a potential effect on training outcomes (Henderson Sabes & Sweetow, 2007).

For this chapter, as reported in Section 5.3.2 (Table 3), ANOVA testing demonstrated no significant differences in age or hearing level among the three care pathway groups. Furthermore, baseline performances on the QuickSIN, BKB and tests of other cognitive variables investigated also indicated no significant differences among the groups, all at $p > .05$.[47]

On exploration for data normality, the two digit span tasks (forward and backwards) and the LSPAN demonstrated a non-normal distribution on Shapiro-Wilk testing, all at $p < .05$. However, as ANOVAs are reasonably robust to more minor deviations from normality (Blanca et al., 2017), mixed-methods ANOVAs were used to explore for performance differences. In addition to the tests for normality, Levene's test for homogeneity of variances and Box's test for equality of covariance matrices were computed before analysis. For all of the following ANOVAs in this section, Levene's and Box's tests were all at $p > .05$, indicating that there was no violation of these assumptions.

Hypothesis 5.1 suggested that the participants on the LACE trained pathway would perform speech-in-noise perception tasks better following their training. A mixed methods ANOVA gave the following results:

- There was no significant difference overall among the groups after training ($F(2, 32) = 1.194$, $p = .316$).
- There was no significant effect of time ($F(1, 32) = 2.239$, $p = .144$).
- There was, however, a significant interaction of training pathway*time on QuickSIN performance ($F(2, 32) = 3.511$, $p < .05$).

To explore where any difference may lay, post hoc analysis was performed within the individual groups. The within-group changes in QuickSIN performance indicated a significant improvement for the LACE pathway ($t(10) = 3.244$, $p < .01$) but not for the SLA ($t(12) = -.731$,

---

[47] Apart from the Stroop interference measure, so when assessing potential change in performance for the Stroop, an ANCOVA was used with starting level as the co-variate.

p = .479) or control pathways (*t*(10) = .290, *p* = .778). The within-group changes demonstrated that the QuickSIN performance significantly improved for the LACE-trained group, whereas it did not for the other pathways. Thus, for the QuickSIN, Hypothesis 5.1 was supported.

The significant interaction of training pathway*time is shown below in Figure 3. There is a decline after training in the SNR needed to discriminate speech for the LACE trained group, evidenced by the negative slope of the blue line. A lower SNR score indicates superior listening performance. The performances for the SLA and control groups were largely unaltered, with the SLA group showing slightly decreased performance and the control group slightly improved performance. Post-hoc analysis demonstrated no significant differences between the care pathways after training.

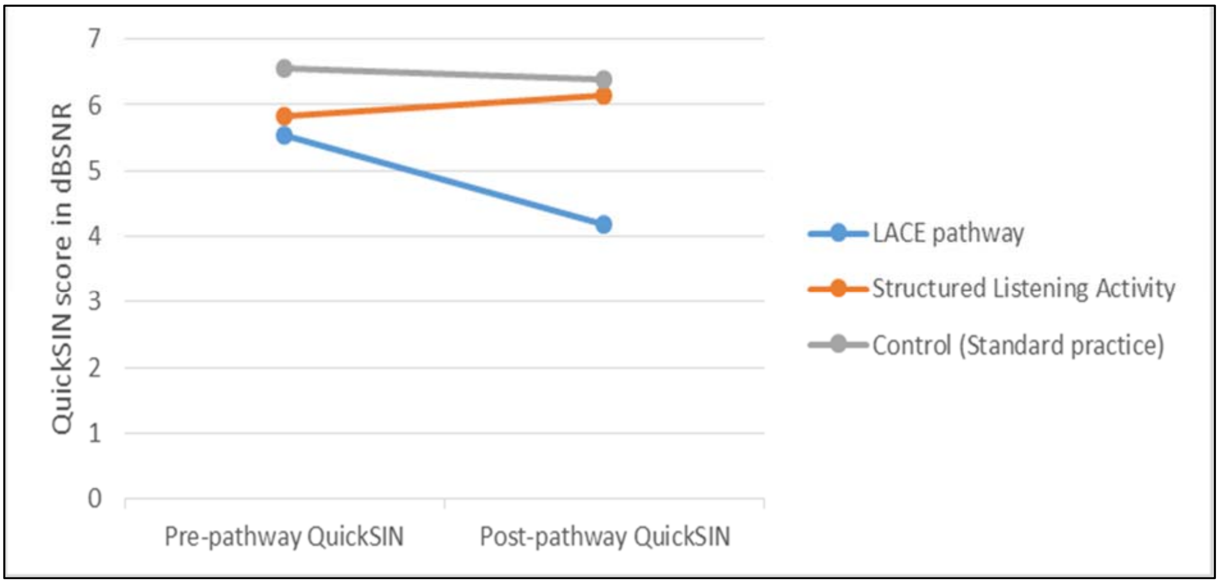


Figure 3. Relationship between assessment time and dB SNR threshold for the QuickSIN task for the three pathways of LACE, SLA and Control.

For the BKB speech test, the relationship between time and the three pathway groups is shown below in Figure 4. There was no significant interaction of time and pathway (*p* = .532). However, there was a main effect of time, as scores significantly improved across the two assessment sessions (*F*(1, 32) = 11.428, *p* < .005). A higher score on the BKB indicates better performance.

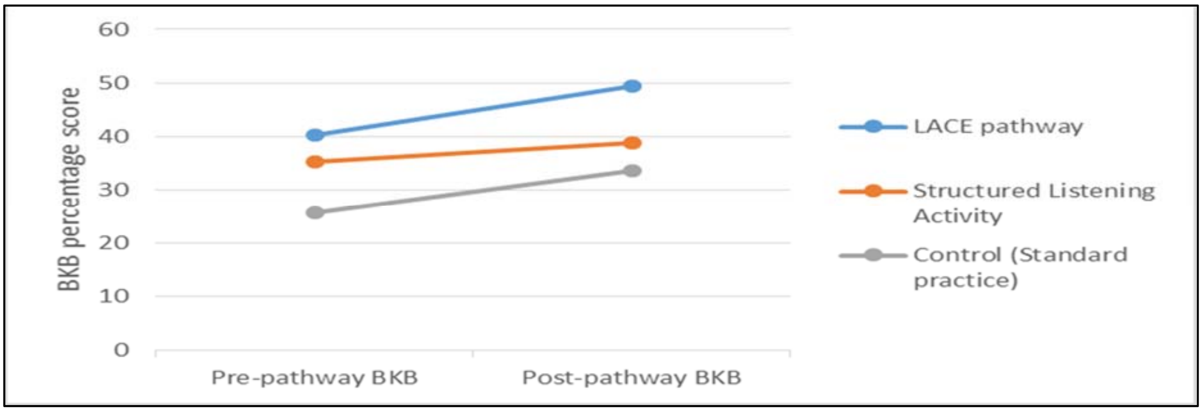


Figure 4. Change in BKB scores for the three care pathways across the 2 assessment sessions.

As there was a change over time for BKB performance within-group post-hoc analysis was performed and this indicated no significant improvement for the LACE pathway ($t(10) = -2.05$, $p = .067$) or for the SLA pathway ($t(12) = -1.29$, $p = .221$) but for the control pathway the improvement was significant. ($t(10) = -2.79$, $p = <.05$).

Overall, Hypothesis 5.1 could only be partially supported in that there was a pathway*time interaction for QuickSIN performance but not for BKB performance.

Mixed ANOVAs were also performed for all of the other variables in Table 4 for which there were pre- and post-training assessments. There were, however, no other significant interactions or main effects apart from the auditory based recognition task which although not giving an overall significant difference between groups ($F(2, 32) = .272$, $p = .763$), did demonstrate a main effect of a decrease in performance over time ($F(1,32) = 16.291$, $p < .001$, Figure 5 below).

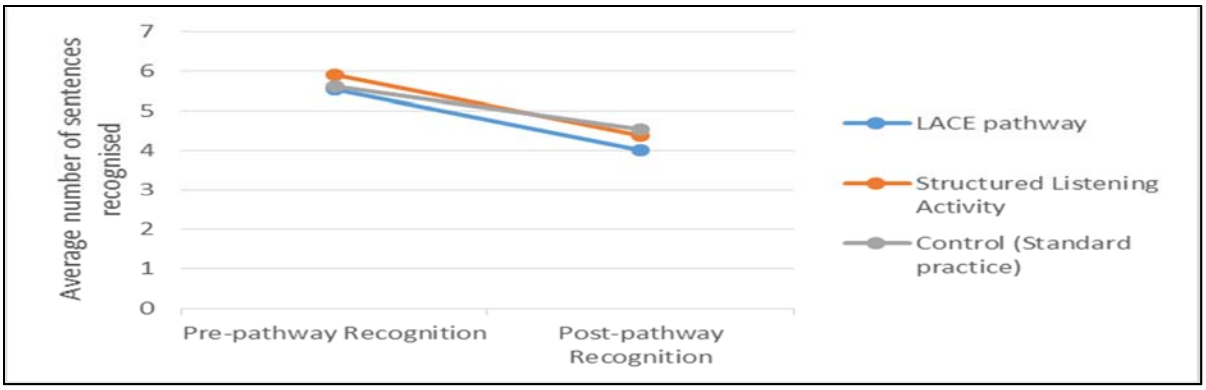


Figure 5. Performance on the recognition of sentences task before and after the training pathways.

Post-hoc analysis of the difference in performance of the recognition task over time were computed and indicated that there was a significant decreases in performance for the LACE pathway (*t*(10) = 3.02, *p* < .05) and for the SLA pathway (*t*(12) = 3.06, *p* < .05) but not for the control pathway (*t*(10) = 1.24, *p* = = .24).

There was, therefore, no clear supporting evidence to support Hypothesis 5.2, as none of the auditory cognitive tasks involving listening skills (auditory sentence span (LSPAN), forward and backward digits) demonstrated improved performance, as there were no significant effects of time or interactions of time*pathway. Indeed, of the cognitive functions assessed the only significant change was in the auditory recognition task and for the pathways that undertook structured training this demonstrated a fall in performance.

The only variable that was significantly different between groups when assessed at baseline was the Stroop interference measure. In view of this, when assessing for any changes in Stroop performance, the starting Stroop level was used as a covariate to control for this difference. Therefore, to investigate potential differences between the care pathways, an ANCOVA was used with the baseline Stroop measure as a covariate. This test did not demonstrate any significant overall difference in Stroop performance between the care pathways after training (*F*(2,31) = .057, *p* = .945).

5.3.4. SSQ12 results for the different care pathways and correlation with post training speech perception performance.

The SSQ12 was divided into its constituent parts of hearing with competing speech (questions 1–5), quality of speech (questions 6–8), spatial awareness of sounds (questions 9–11) and listening effort (question 12). The self-rated participant scores at baseline and the amount of self-rated improvements after training are below in Table 5.

Table 5

Overall SSQ12 ratings and ratings by section of SSQ measures.

Pathway group	SSQ measure	Pre-pathway rating	Standard deviation	Post-pathway improvement rating
LACE	Overall	5.65	.95	2.61
	Competing speech	4.58	1.38	2.48
	Sound quality	5.93	1.05	2.82
	Spatial awareness	6.89	1.43	2.66
	Effort	4.45	1.85	2.32
SLA	Overall	5.60	1.61	2.46
	Competing speech	4.73	1.88	2.61
	Sound quality	6.79	2.03	2.45
	Spatial awareness	6.53	1.61	2.11
	Effort	4.30	2.41	2.58
Control	Overall	6.26	1.45	2.60
	Competing speech	4.85	1.71	2.60
	Sound quality	7.40	2.03	2.50
	Spatial awareness	7.26	2.28	2.11
	Effort	5.36	2.28	2.18

As seen in Table 5, after using their hearing aids for the 4–8 week period of the research, all participants on average gave improved self-rated scores irrespective of which pathway they had followed. The rating scale was such that 0 would indicate no improvement, +1 to +5 measures of improvement and conversely -1 to -5 measures of decreased performance. As a two-way ANOVA could not be used because the SSQ12 gives a change in score rating rather than a second score, a one-way ANOVA was used to explore differences between groups after the pathways had been followed. There were, however, no significant differences in SSQ12 ratings between the training pathways for either the self-rated overall change or any of the sub-sections. Thus, there were no significant differences between the pathways for the SSQ12, meaning that Hypothesis 5.3 was not supported.

Correlational analysis between the post speech-in-noise performance for QuickSIN and BKB and the change in rating level on the SSQ12 for each pathway (both overall rating and subsection ratings as above in Table 5) indicated no significant correlations (after the Holm-Bonferroni correction had been applied) apart from performance of the two speech tests with each other. As there were no significant correlations the table is not included here. Accordingly, Hypothesis H5.4 could not be supported.

5.3.5. Full correlation analysis

Correlational analysis was performed on the pre-training pathway scores to explore relationships between the measured variables. As some of the assessment tasks demonstrated a non-normal distribution at baseline as reported in Section 5.3.3., non-parametric analysis was performed using Spearman's rank correlation coefficient to explore for any significant relationships. A Holm-Bonferroni (Holm, 1979) correction was also used as in Chapters 2 and 3. This utilised a changing critical *p* value, and since there were 10 variables explored for correlation, the correction applied changed the alpha value of .05 from alpha/10 to alpha/1. The changing alpha values are applied to the strongest relationship first, then to the second and so on. The analysis aimed to determine the strength and direction of any relationships; the results are given below in Table 6.

Following the correlation table, there are 5 scatterplots demonstrating the relationships of the QuickSIN test with other speech tests and with other demographic factors or cognitive functions.

Table 6

Correlation matrix of performance on tasks with Spearman's rank correlation coefficients (N = 36).

	Age	**PTA Ave**	**QuickSIN**	**BKB**	**LSPAN**	**Stroop IM**	**Stroop PS**	**Digit span forward**	**Digit span back**	**RT**
Age	—									
PTA Ave	.243	—								
QSIN	.455*	.666*	—							
BKB	-.243	-.366	-.724*	---						
LSPAN	-.046	-.066	-.400*	.179	—					
Stroop IM	.117	-.010	-.068	.264	-.280	—				
Stroop PS	.234	-.038	.001	.013	.069	-.194	—			
Digit span forward	-.119	-.357	-.467*	.157	.533*	-.412*	.160	---		
Digit span backward	.181	-.063	-.042	-.034	.356	-.204	.125	.501*	---	
RT	.345	-.231	.166	-.216	-.047	.065	.461*	-.081	.111	---

Key: PTA = Pure tone average threshold across 500-4000Hz, QSIN = QuickSIN speech in variable babble, BKB = Speech in constant speech-shaped noise at -4 dB SNR, LSPAN = Listening span (last word in sentence), Stroop IM = Interference Measure (measure of inhibition control), Stroop PS = Processing speed of word/colour in congruent condition, RT = Reaction time.
Correlations: *Correlation is significant at the 0.05 level (2-tailed), Holm-Bonferroni correction applied.

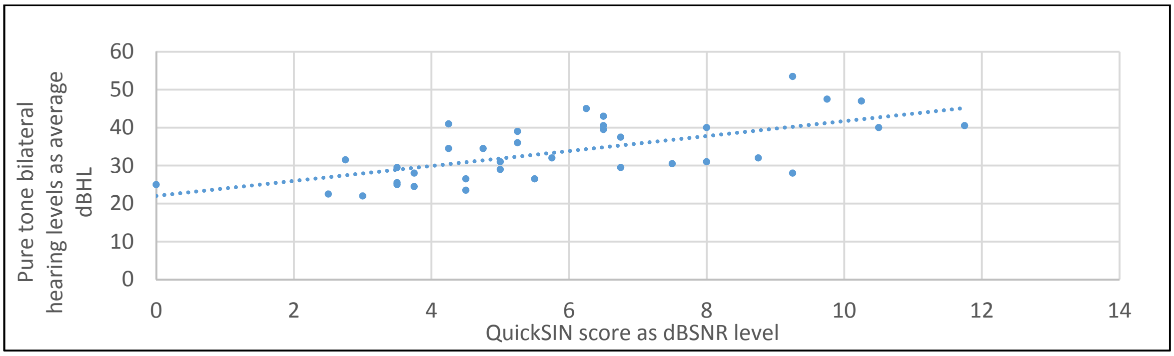


Figure 6. The correlation of QuickSIN performance and hearing level (PTA average). Spearman's rho = 0.666.

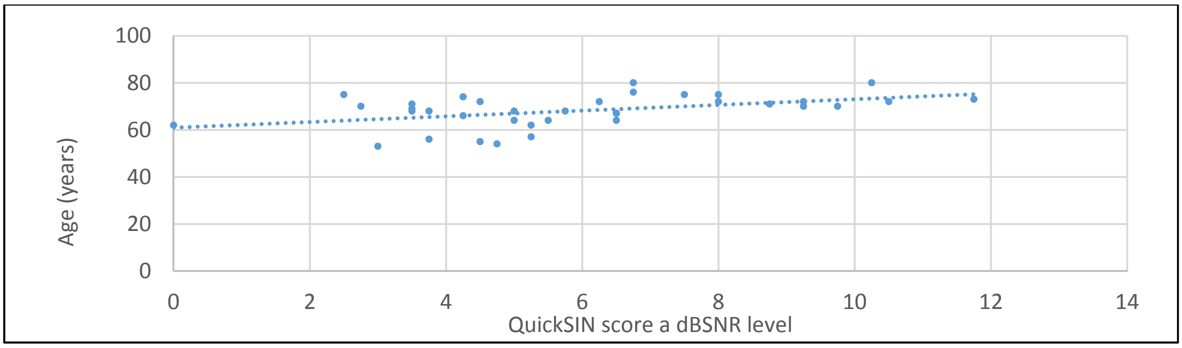


Figure 7. The correlation of QuickSIN performance and age in years. Spearman's rho = 0.455.

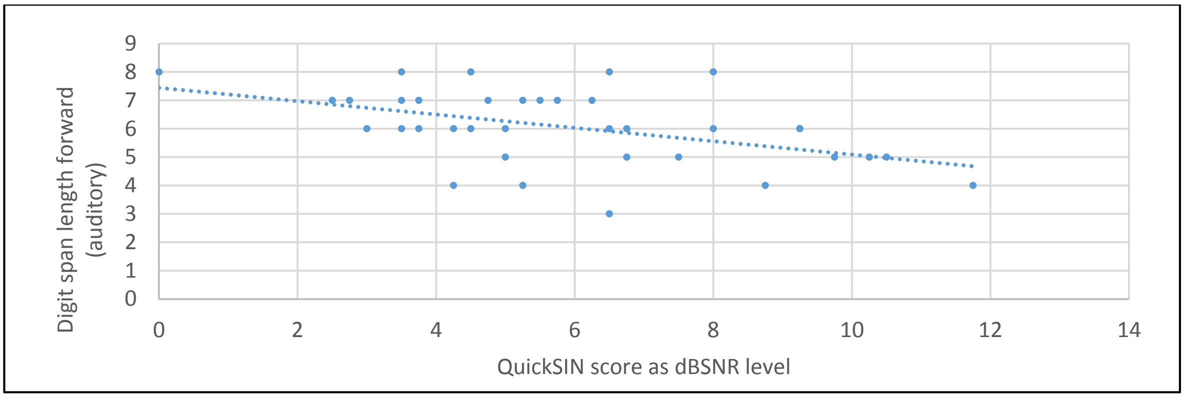


Figure 8. The correlation of QuickSIN performance with digit span forward performance. Spearman's rho = - .467.

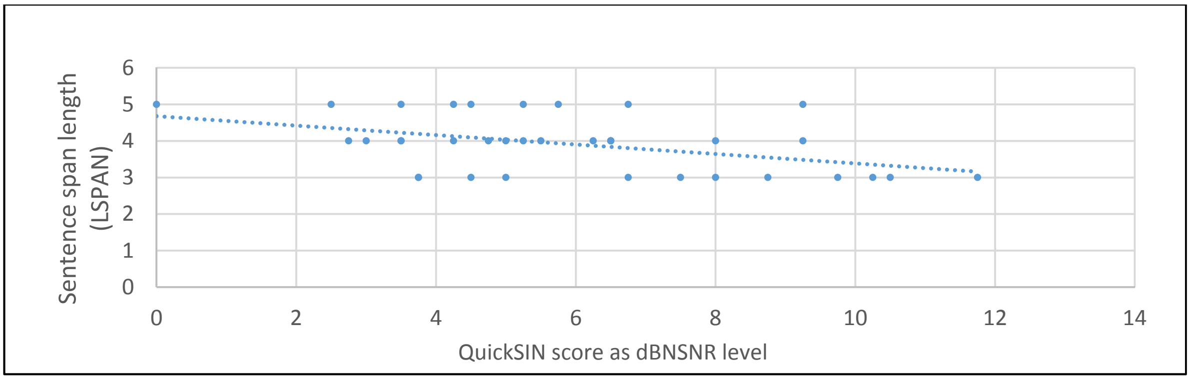


Figure 9. The correlation of QuickSIN performance and working memory performance (LSPAN). Spearman's rho = -.400.

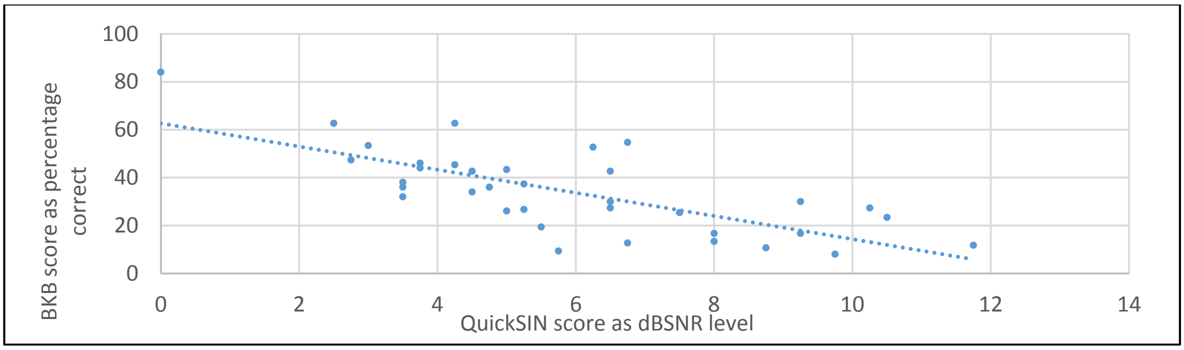


Figure 10. The correlation of QuickSIN performance and BKB performance. Spearman’s rho = -.724.

Hypothesis 5.5 was supported for the QuickSIN, in that there were two cognitive functions that significantly correlated with QuickSIN speech-in-noise performance, but was not supported for the BKB test. The relationships between the two speech tests and cognitive functions did, however, support Hypothesis 5.6, which stated that the two speech tests would have different relationships with cognitive functions. The relationships are explained below.

The QuickSIN and BKB speech-in-noise tests were strongly correlated (*ρ* = -.724). Therefore, participants who performed well on one speech test also performed well on the other, even though one involved listening in variable speech babble and one in speech-shaped noise at a fixed, challenging SNR.

The relationships of the two speech tests with the other variables did not show similarly strong relationships. For the QuickSIN, the correlational analysis demonstrated that the PTA average, calculated as a participant’s average hearing threshold across 500–4000 Hz, showed the strongest significant relationship. As a lower dB SNR value on the QuickSIN corresponds to better performance, the correlations of QuickSIN with age and PTA average (for which a lower value also indicates better hearing) are positive. Other significant relationships with the QuickSIN were the two listening span memory tasks, last word in sentence and auditory digits (forward). These had correlations of *ρ* = -.400 and *ρ* = -.467, respectively.

The BKB, by contrast, was only correlated with the QuickSIN task, thus supporting H6.6 in that there are different relationships between the different speech tests and performance on other tasks.

5.3.6. Hierarchical regression

The hierarchical regression analysis was developed with the same rationale as in Chapters 2 and 3. The hypothesis to be tested was, do cognitive functions add significantly to performance variations in QuickSIN tests after age and hearing levels have been factored out. In view of the previous research age and PTA made up the first two models with working memory being added third. The functions of speed of information processing and control of inhibition were added together in model 4 as neither had demonstrated a significant relationship with the QuickSIN test but previous chapters and other researchers had indicated that they may add to the explanation of the variance. A model was not created for the BKB speech test as there were no significant correlations between cognitive functions and BKB performance, the only variable that correlated was the performance of the QuickSIN.

The hierarchical regression model below (Table 8) was developed for the QuickSIN test.

Table 7
Hierarchical regression model to predict QuickSIN performance.

Model	**Predictors**	***R^2***	***R^2* change**	**Significance *F* change (*p* value)**
Model 1	Variation due to age	.189	.189	.005
Model 2	Variation due to PTA	.476	.294	.000
Model 3	Variation due to memory span (last word in sentence and forward auditory digit span)	.633	.169	.002
Model 4	Variation due to speed of information processing (reaction time) and inhibition control	.662	.045	.113

The hierarchical model was computed to determine whether the addition of the working memory tasks improved the prediction of QuickSIN task performance over and above prediction based only on age and PTA average. Table 8 contains the full prediction of the QuickSIN task scores. The full model including age, PTA and working memory tasks (last word sentence performance and forward digit task) was statistically significant ($R^2 = .633$, $F(4,31) = 16.071$, $p < .005$). The addition of the two working memory measures (Model 3) led to a statistically significant increase in R² of .169 ($F(2, 31) = 8.047$, $p < .005$). The further addition of model 4 adding information processing speed and control of inhibition did not add significantly to the final model. Hypothesis 5.7 was, therefore, partially supported as some of the cognitive functions did contribute significantly to perceiving speech-in-noise after age and hearing level had been partialled out.

Table 8
Hierarchical regression predicting QuickSIN test scores from age (model 1), from age and PTA (model 2), and from age, PTA, working memory (LSPAN and digital span forward) (model 3) and information processing speed and control of inhibition.

	b	***SE b***	***ß***	***R²***	***R² change***	**Significance of *F* change**
Model 1				.189	.189	.005
Age	0.175	0.058	**0.460***			
Model 2				.476	.294	.000
Age	0.117	0.048	**0.308***			
Bilateral PTA average	0.185	0.042	**0.563****			
Model 3				.633	.169	.002
Age	0.118	0.040	**0.312***			
Bilateral PTA average	0.168	0.038	**0.513****			
LSPAN	-1.244	0.402	**-0.367****			
Digital span (forward)	-0.165	0.252	-0.083			
Model 4				.662	.045	.113
Age	0.130	0.040	**0.343****			
Bilateral PTA average	0.157	0.037	**0.477****			
LSPAN	-1.274	0.386	**-0.376****			
Digital span (forward)	-0.394	0.264	-0.197			
Reaction time	0.000	0.005	-0.11			
Strop interference	-0.008	0.004	**-0.243***			

Note: * $p < .05$, ** $p < .005$

The final model accounted for 66% of speech-in-noise performance variance. Individually, age, hearing level (as measured by the PTA average), memory span

(LSPAN) and inhibition control (Stroop interference measure) all contributed significantly to the final model.

5.4. Discussion

Previous research has demonstrated transfer gains that resulted in speech-in-noise improvements on two different speech tasks following communication training using LACE (Song et al., 2012). In addition, evidence of cognitive release after auditory training has also been reported in Chapter 4 of this thesis, where working memory significantly improved after a training period. For both of those pieces of research, the participants were normally hearing young adults (for Song et al., ages were 19–35 years, and for Chapter 4, 18–50 years). This differs from the participant demographics of previous research that had found no successful transfer (Olson et al., 2013; Saunders et al., 2016a)[48] as well as the research in this current chapter[49]. Additionally, where these older participants/patients were recruited, they all had hearing loss and hearing aids.

The difference in outcomes reported by the researchers who explored the potential for transfer to other speech tasks and/or cognitive release could, therefore, be due to the age and hearing status of the participants/patients recruited. Hypothetically, this may be due to two different but potentially linked processes. Firstly, where auditory training was used with young adults (Song et al., 2012) and Chapter 4), the greater plasticity of these young adults' neurological systems allowed for clearer phonological representations of speech sounds to be more readily developed. This in turn allowed for the development of sufficiently strong phonological representations to support better speech performance, even on novel speech tasks. For the older groups, as the older patients/participants had some degree of hearing decline and a poorer-functioning neural system, this may limit how efficiently new phonological representations are made, even with better audition.

---

[48] Olson et al. 50-81 years. Saunders. et al. average 68 years.
[49] Chapter 5. 50-80 years.

Secondly, this more malleable neural system of younger, unaided, normally hearing participants allowed for the release of cognitive resources for other tasks as audition became easier. As LACE also employs general cognitive exercises, and simply undertaking the training itself assists in neural activation and stimulation, there is an additional strengthening of the neural networks associated with the cognitive tasks that LACE employs. It is potentially the case, therefore, that older adults with hearing loss and less plastic neural systems may need more intense or longer-duration training before any advantages in terms of transfer gains are exhibited, especially if there is any longstanding sensory underload (Sekuler & Blake, 1987). This second point is not to say that transfer is not possible for older hearing aided patients, but it may take longer.

Overall, it can be seen from the present study and past research that on-task improvement in the QuickSIN speech task may be achievable. However, as it was only the QuickSIN results in the current research that showed a significant interaction with LACE, and LACE specifically targets speech in babble, this does illustrate that it is targeted training that offers the greatest advantage. This reinforces that fact that demonstrating improvements in the performance of even closely associated tasks (speech in a different type of noise) is much more difficult to confirm. The implication is that if training is to be successful, training that more closely mimics the desired improvement will yield better results. For example, if an individual struggles to hear in a noisy office, basing their training on listening to speech in background noise spectrally shaped to represent office noise would seem to offer the greatest advantage on an individual basis.

Great efforts were made for the research in this chapter to ensure that all participants had the same hearing aids and hearing aid fitting set up and that real ear measurement verification was performed to ensure good access to speech sounds. Further, as baseline measures were all performed at Day 1, they had all been exposed to hearing aid use for the same length of time. It could be argued that there should have been some acclimatisation time for the patients before the baseline assessments; however, it is often the early exposure to novel/different representations of sounds that induce change. If hearing users are classified as ‘new users’ but have had hearing aids for between 4 weeks and 6 months (Olson et al., 2013; Saunders et al., 2016a), any training after this initial period may not demonstrate the same effects. This may then have an undue influence on any reported results, particularly speech tasks. The research for this chapter did indeed demonstrate that LACE may offer a significant advantage for the targeted speech task, whilst the Olson et al. and Saunders et al. research did not. This may be due to the training being introduced on Day 1 and to ensuring better comparability during patient selection in this research.

The overall SSQ12 score demonstrated that participants felt they had improved on a range of listening scenarios after their training pathway, as evidenced by all groups giving positive average improvement scores. The LACE group did record the largest improvement score (2.61) compared to SLA and Control (2.32 and 2.22, respectively), but this difference was not statistically significant. These improvements were not evidenced by any measured task improvements for the SLA or control groups.

It had been expected that as the LACE-trained group demonstrated a significant improvement in the QuickSIN task, there would also be a significant self-rated improvement in the ‘listening in noise’ subsection of the SSQ12 among that group compared to the other training groups. However, this was not the case. Although overall, the correlation values for the LACE care pathway group did indicate that there was a greater tendency towards a significant relationship with the global SSQ12 score and the SSQ12 subsections, none of these correlations was at a statistically significant level. Unfortunately, due to training pathway sizes (only 11-13 participants per pathway) the correlations reported may not fully represent the possibility of any potential relationships.

For the other training groups, although the QuickSIN score did not improve significantly following those training pathways, participants still reported that they felt they were hearing better in competing noise situations. There were, however, no significant correlations of speech performance with self-rated improvements for either the BKB or the QuickSIN. This suggests either that the tasks themselves are not sufficiently representative of how an individual makes judgements about listening in real life situations or that individuals on self-rated questionnaires do not make the best judges of changes in performance.

There was a decline in the recognition task for all groups, which was found to be a significant effect over the two assessment sessions. A possible interpretation of this is that with newly fitted hearing aids, participants focus on the listening aspect more than the storing and processing of information. This is an area that is in need of further investigation. Neither of the active training pathways (LACE and SLA) in this present research involved recall/recognition exercises. If the results observed here are replicated in further studies, this decline is an area that would need addressing, as ‘attended to’ information that is being heard by an individual needs to be successfully processed into longer-term memory for future use (Baddeley, 1986). It could be argued that the use of the working memory LSPAN task sentences for the recognition task meant that listeners may have focussed on one task at the detriment of the other, but the task was the same at the two assessment visits, which would not account for the performance decrement. Without proper encoding of information into long term memory, real communication is not taking place; therefore, any communication training package may need to consider incorporating some exercises involving encoding information into secondary memory.

In the final hierarchical regression model for this chapter, age, hearing level (measured by PTA), working memory (LSPAN), and control of inhibition all contributed significantly to the variation in speech-in-noise performance (66%). Speed of information processing did not offer a significant contribution. Hearing level was the main contributor in the final model which is in agreement with previous research as being the major predictor for speech-in-noise performance (van Rooij & Plomp, 1990) The findings from this chapter would also agree with previous research on working memory (Rudner et al., 2011) where it has also been reported that working memory plays a significant role for older individuals with hearing loss in the support of discriminating speech-in-noise. It was also the case that as in Chapter 2 that control of inhibition does offer some support after other factors have been considered.

The hierarchical regression model in this chapter offered the greatest explanation of speech-in-noise performance variation compared to other chapters (where hearing levels were better and participants had a younger average age). It appears, therefore, that when hearing thresholds are worse (in this case associated with a hearing decline due to age) that hearing itself accounts for a greater amount of variation but additionally these same older individuals are also more likely to rely on cognitive support to assist with their hearing in background noise.

5.5. Limitations

Obtaining participants to perform the research was a particular challenge, as evidenced by the large number of records reviewed for suitability and the limited number of participants consenting after contact. The selection criteria were rigid in order to minimise bias and variability, as recruiting participants who did not have a computer or internet access into groups other than LACE would have been far simpler but would have had an effect on research quality. A multi-centre site approach would have offered faster recruitment and should be considered for future research. The lack of participants had the further effect of limiting the statistical power of the research. Both the ANOVAs and the correlations among care pathways would have benefitted from greater participant numbers.

The largest challenge encountered was that of Neurotone, the developers of LACE, changing its operational system to a purely online registration and training. As this happened well into the course of the research, it meant that no further participants could be registered and any extra training sessions undertaken by existing participants were no longer recorded on the training website. Thus, research recruitment had to end when the random allocation reached the next LACE participant, although the robust effect size showing a significant interaction for the QuickSIN test indicates that a sufficient number of participants had been recruited. As the purely online training was a different program than the CD-based training, there was no possibility of continuing the research as it was designed following the change.

A further factor was that of LACE training compliance, as 2 participants only completed 10 LACE sessions. However, as argued previously, the advantage of undertaking the training appears to take place mainly in the first 10 sessions. Thus, the results of all participants were included in the data analysis.

5.6. Conclusions

For the group of older participants with hearing loss who were recruited for the research in this chapter, the correlations demonstrated that performance on working memory tasks contributed to explaining the differences in speech-in-noise perception performance. This was after partialling out age and hearing loss. Working memory does seem to play a role for older individuals with hearing loss (Füllgrabe & Rosen, 2016b), but only when listening to speech in certain types of noise and/or certain sentence lengths (in this case, QuickSIN in speech babble and not the shorter BKB in a constant speech-shaped noise). The relationship was readily demonstrable when using both a simple, auditorily presented memory task and a more complex LSPAN in which the sentences, including the target words, were presented in noise. Therefore, the choice of speech test does influence whether significant relationships are demonstrated. Older listeners with hearing loss may already be applying cognitive resources to assist with speech-in-noise perception. Therefore, when the task is challenging with longer sentences and when parts of words can be more readily heard (as in speech glimpses within babble), it is individuals who have higher cognitive functioning and who efficiently reallocate their remaining resources who will perceive speech-in-noise better.

The results demonstrated significant improvements on the QuickSIN speech-in-noise task for participants who followed the LACE pathway. Still, there was no demonstration of any gains being transferred to a different speech test or to improvements on any cognitive tasks. It had been anticipated that there may be evidence of release of cognitive resources, since less processing (and thus less cognitive resources) would be needed as discrimination became more automatic. This was not observed in the current study, as no significant improvement was demonstrated, particularly for the auditory-based working memory tasks. It appears, therefore, that targeted training works well, but release of cognitive resources and far transfer improvements are not indicated.

The evidence surrounding any advantage of undertaking LACE, therefore, remains somewhat equivocal, as significant advantages over the other pathways were limited to just one speech test. Overall, therefore, there is no great indication that employing LACE training at the start of acclimitisation to hearing aids (so immediately post-fitting) offers any significant advantages. However, following hearing aid fitting, all patients did self-rate themselves as improving in listening in noise and in other measures, as indicated by the increased SSQ12 scores, with the LACE group rating their self-improvement highest. It may be, therefore, that some of the tasks used to judge the success of hearing aid fitting and subsequent intervention pathways were not sensitive enough to specifically measure any real gains. This also offers the suggestion that a ‘one-size-fits-all’ training package may not be the best way to improve an individual’s communication.

CHAPTER SIX

6. General discussion

This final chapter begins with a summary of the research undertaken. It describes the general findings and makes suggestions for the potential direction of any future research. The specific findings and limitations of the research reported in previous chapters are not reiterated here. Instead, following the research summary, it goes on to describe the shifting relationship between cognitive functions and speech discrimination with changing hearing level, how targeted training delivers near and far transfer improvements, and the potential for cognitive release. It also makes suggestions for the most effective way to investigate cognitive factors that play a role in speech-in-noise perception performance. The chapter concludes with suggestions for future research and the potential development of an assessment package that could be utilised by listeners who report struggling when listening in challenging situations. Based on the outcome of the assessment package, tailored training could be offered on an individualised basis.

This thesis had at its core the importance of the ability of individuals to communicate and its key role in social interaction. For many individuals, this will entail being able to listen to and process spoken information, although this may not be the case for individuals with a visual language. In many situations, the environment in which communication takes place will not be ideal and listening is effortful. For some listeners in these environments, particularly older ones, there may be a reliance on cognitive support mechanisms in order to efficiently process imparted information. There appears to be a wide range in the efficiency with which individuals can listen in these degraded listening environments as well as in the support mechanisms that they use. Each individual will potentially vary in their reliance on the different support mechanisms.

For older listeners, the problem may be exacerbated due to presbycusis, for which they may or may not have sought help. Listening in noisy situations is more challenging for all listeners and particularly for individuals with hearing loss. As listening becomes more challenging, there may be a tendency for older individuals to avoid these noisier situations, but unfortunately, these are often the situations where social engagements take place. There is a growing bank of evidence surrounding the negative health effects of social withdrawal, with an increasing link being demonstrated between declining hearing, mental health wellbeing and the incidence of dementia (Lin, Metter, et al., 2011).

This thesis examined the problems of speech perception in background noise and explored underlying reasons for the variations noted in performance of speech-in-noise tasks. This research was split into attempting to identify two core phenomena: the cognitive processes that underlie or help to explain these variations, and the effects of training (speech or cognitive) on performance.

A summary of the main findings follows in Table 1. The three key areas that have arisen from the research in this thesis are then discussed in terms of the general overall findings.

Table 1.
Summary of studies

Study	Sample	Aim(s)	Measures	Results	Conclusion summary
Chapter 2: Speech discrimination in noise and correlation with control of inhibition, speed of information processing and working memory	70 participants with no self-reported hearing loss. Aged 18-76 years. 22 men and 48 women	1. To examine which of the cognitive functions of control of inhibition, speed of processing, response latency and working memory significantly correlated with a speech-in-noise task set at a constantly challenging signal-to-noise ratio. 2. To examine which of the correlating cognitive factors best explained the differences in speech-in-noise performance. 3. To explore the changing relationship of cognitive skills in relation to age in explaining the differences in performance of speech- in-noise tasks	Pure tone audiogram Speech-in-noise test at -8dB. Cognitive tests: An LSPAN, working memory test, Stroop, auditory selective attention task, random number generation task, go/no-go task.	A significant correlation of cognitive functions with speech-in-noise performance (BKB) was only seen for speed of information processing as measured by the go/no-go choice reaction task (CRT). Hierarchical regression modelling (after hearing level and age had been partialled out, Model 1) and working memory, measured by an LSPAN (model 2), demonstrated a significant extra explanation in performance variance when the speed of information processing task (CRT) was added in Model 3. Model 4 involved the addition of control of inhibition (as measured by a Stroop and a selective auditory attention task). This also gave a significant extra explanation of the speech-in-noise performance. The total variance explained was 29%. Individually, the only factors that predicted performance variance after the others had been accounted for were speed of information processing (the go/no-go task) and control of inhibition (the Stroop interference measure).	Two cognitive functions did assist in explaining the variation in speech-in-noise performance at a particularly challenging SNR. These were speed of information processing and control of inhibition. Participants in this study had normal (or only a mild hearing loss) and hearing thresholds and age did not correlate strongly with speech-in-noise performance. It appears that when the speech task is highly challenging that the ability to process any glimpsed information quickly or to be able to quickly inhibit unwanted noise offers performance advantages. As the total variance of the dependent speech-in-noise scores explained by the measures used was low, the conclusion was that there were other variables not used in this study that may further explain performance variations.
Chapter 3: Does inhibition and processing speed training improve	48 participants with no self-reported hearing loss.	1. To explore relationships between cognitive functions and speech-in-noise performance.	Two speech tests, the QuickSIN in adaptive background babble and a BKB in speech-shaped noise	Significant correlations were seen for the QuickSIN performance with the UfoV selective attention task, the Stroop interference measure, auditory compressed speech and auditory	On the correlational analysis before training, hearing loss and age, correlated with QuickSIN speech-in-noise performance. The cognitive functions also

performance in speech-in-noise and other associated cognitive tasks?	Aged 18-80 years. 27 women and 21 men Randomly assigned to one of three training pathways: visual or auditory processing speed and inhibition training or a general cognitive training pathway.	2. To investigate if a short training course aimed at improving either processing speed or control of inhibition would improve a listener's performance of two different speech tasks. 3. Would any potential gains be sensory modality-specific for the speech tasks and for a range of cognitive tasks? 4. To explore the changing relationship of cognitive skills in relation to age in explaining the variations in performance of speech-in-noise tasks.	at a constant -8dB SNR were performed. The cognitive functions assessed were control of inhibition (measured by random number generation, Stroop interference, selective auditory selective attention and a useful field of view (UfoV), working memory (measured by a complex reading span task) and processing speed (measured by UfoV and auditory compressed sentences task). The participants were randomly assigned to one of three training pathway groups. These were a workbook trained group involving general cognitive activities, or 2 groups that performed on-line either auditory or vision training tapping the cognitive skills of processing speed and inhibition.	selective attention task. There was no significant relationship with performance of the BKB. However, the auditory compressed speech task did tend towards significance. Hierarchical regression modelling demonstrated that age and hearing level significantly explained QuickSIN performance variance. Once this had been accounted for (Model 1), working memory was added in Model 2 bud did not add significantly to the explanation. Speed of information processing added significantly to the variance (Model 3). The addition of control of inhibition measures, although they added to the explanation of performance variance in model 4 this was not significant. Overall, the variance accounted for by all these measures was 44%. There was only one significant interaction of time and training group and this was for the QuickSIN with the auditory trained inhibition and processing speed group showing the largest improvement on the QuickSIN. Post hoc analysis though showed no significant difference between the groups after training. The differences in performance of the BKB test and all the other cognitive tasks generally demonstrated no significant advantage of undertaking one type of training compared to another.	helped to explain performance variations in the QuickSIN. The performance of different speech tasks with different types of background noise demonstrated that the cognitive skills that help explain performance variations for one task may not do so for another, implying that different skills are needed depending on the speech material used and the type of background noise. Processing auditory information quickly appeared as a key skill in QuickSIN performance and was the only factor that suggested any relationship close to significance correlation with the BKB task, reinforcing the important role this cognitive skill plays. It appears that training focussing on speed of processing and inhibition control do not give evidence that training leads to improved speech-in-noise performance.

Chapter 4: Does auditory training lead to on-task improvements with resulting far transfer gains and a release of cognitive resources?	70 participants with no self-reported hearing loss. Aged 18-50 years. 24 men and 46 women. Randomly allocated to one of five groups.	1. To explore if there was an improvement in speech perception of degraded speech after a short training course in the degraded speech. 2. Did any improvement demonstrate near transfer to a different type of degraded speech? 3. Was there any far transfer of any improvement in performance, indicating a release of cognitive resources in an associated working memory span task?	Pure tone audiogram Degraded speech perception scores for speech-in-noise, sine vocoded and noise vocoded speech. Working memory span scores measured by a listening span (LSPAN) in the training group's own type of degraded speech. Reading span task (RSPAN).	Groups trained in a particular type of degraded speech performed significantly better after training in that type of speech than a control and an active control group. The noise-vocoded trained group also performed better than the control and active control for all the different types of degraded speech after training. The performance of the LSPAN task for the noise and sine vocoded groups demonstrated a significant improvement post training. The RSPAN task did not correlate significantly with speech-in-noise performance or show any signs of improving after training.	Training can lead to significant improvements in speech performance. The improvements were most evident for the type of speech participants were trained in. There was no significant evidence of near transfer improvement to performance of other speech tasks. Where the speech was novel (noise and sine vocoded) there was a release of cognitive resources as the speech discrimination improved, as evidenced by the significantly improved LSPAN scores after training, thus demonstrating a degree of far transfer.
Chapter 5: The role of a commercial communication training package (LACE) in the rehabilitation of newly fitted hearing-aided patients.	37 hearing impaired participants who had opted for binaural hearing aid fitting. Aged 53-80 years. 16 men and 19 women. Randomly assigned to one of three care pathways: a LACE care pathway, a	1. To investigate relationships between cognitive functions and speech-in-noise performance for a group of older adults with hearing loss. 2. To investigate if one of the care pathways led to improvement in the two speech-in-noise tasks undertaken along with any changes in	Two speech tests, the QuickSIN and a BKB in speech-shaped noise at -4dB SNR were performed. The cognitive test battery included a listening span (LSPAN) task (last word of a sentence presented in noise at +10dB SNR), a Stroop response latency and	Age and pure tone average hearing threshold both correlated significantly with QuickSIN performance. Two of the memory span tasks, the last word listening span task and the forward digit task (presented auditorily) also correlated significantly. Although the two speech tests correlated significantly with each other the only variable that correlated with the BKB test was hearing thresholds (the PTA average). Hierarchical regression modelling for QuickSIN performance used age as Model 1 and average hearing threshold for Model 2.	The explanation of performance variation due to working memory measures (the LSPAN) and the forward digit span) indicates that for older listeners with hearing loss, working memory is a key skill for the QuickSIN. This was not the case for BKB performance. It would seem this is due in part to QuickSIN sentences being longer than BKB sentences and the BKB task being considerably more difficult.. The significant interaction time*care pathway demonstrated

	cognitive workbook care pathway and a control.	associated cognitive tasks. 3. To explore near and far transfer of any potential improvements offered by undertaking a care pathway.	interference measure and forward and backward digit span presented auditorily in quiet. A recognition task of sentences used in the LSPAN was also undertaken. An SSQ12 questionnaire that measures how well listeners self-rate their performance in different situations. This was completed before and after the care pathway.	Model 3 consisted of the memory span tasks of last word in sentence and auditory digit span. Model 4 explored if control of inhibition and speed of information processing added further to the final explanation of speech-in-noise performance variance but their addition was not significant. The total hierarchical model accounted for 66% of performance variance. There was a significant interaction of care pathway and time (pre and post-pathway), with the LACE pathway being the only pathway to demonstrate a significant improvement. This was not replicated for the BKB test, although the LACE pathway did display the largest improvement. There was a main effect of time for the BKB test, however, with all the pathway groups performing the task better at their 2nd visit. There were no other significant interactions or main effects for any of the other cognitive variables. The recognition task displayed a main effect of time, demonstrating a fall in recognition over time. The SSQ12 self-rating questionnaire did not show any significant difference between the pathways, although overall participant ratings were higher for the LACE training pathway.	that LACE training can improve QuickSIN performance, although it should be noted that QuickSIN tasks form part of the LACE training package. Therefore, it perhaps can be assumed that training on the test material does lead to improvements in performance. It remains unclear whether it is practice of the QuickSIN that leads to improvements or contributions of any of the cognitive functions that are also trained. There was no demonstration of far transfer to BKB task performance, which does not form part of LACE training, so it appears training does have to be targeted for significant improvements to be evidenced. The lack of transfer to questionnaire improvement in real-world listening situations also demonstrates the difficulty in achieving far transfer gains.

6.1. The changing influences of cognitive input and hearing level when listening to speech in background noise.

The research undertaken for this thesis demonstrated that cognitive functions often do correlate significantly with speech-in-noise perception performance. The general trend was for any relationships to be weak, though significant. This was further demonstrated by the hierarchical regression models that were developed for the research chapters. For the participants with no self-reported hearing loss in Chapters 2 (ages 18–76 years) and 3 (18–80 years), the hierarchical model accounted for 29% and 44% of performance variance, respectively. The speech tests for these chapters, however, differed, with BKB in speech-shaped noise utilised in Chapter 2 and QuickSIN in four-person babble and BKB in speech-shaped noise in Chapter 3. The BKB test was significantly more difficult, and hearing threshold (PTA) and age did not offer a significant explanation of any variation in performance. Working memory offered a limited, non-significant explanation, but when age, PTA and working memory had been partialled out, both speed of information processing (11%) and control of inhibition (9%) significantly added to the explanation. For this challenging speech-in-noise task, it could be interpreted that as long as hearing thresholds are not drastically reduced, it is individuals with faster processing speed and better inhibition control who perform best.

When exploring performance of the QuickSIN in Chapter 3, as more speech could potentially be ‘glimpsed’ amongst the noise for this test than for the BKB, the hearing level (PTA) became more important (combined with age, this explained 28% of the variance). This is in accordance with previous research where hearing levels have been found to be the main predictors of speech perception (Humes & Dunbo, 2010; Van Rooij & Plomp, 1992b). The cognitive functions did contribute as well (speed of information processing and control of inhibition added 13% to the explanation of variance), but these were secondary to hearing level.

In Chapter 5, when all of the participants were older first-time hearing aid users, a regression model was constructed based on significant correlations with the same QuickSIN performed as in Chapter 3 (age, hearing threshold and working memory; auditory last word in sentence in background noise and forward digit span), and the final model accounted for 66% of performance variance. In this case, age and hearing level alone accounting for 47% of variance. Working memory did also add a significant extra explanation (16%) even after age and hearing level had been partialled out. This is in agreement with previous observations that working memory appears to support older individuals with hearing loss much more than younger individuals (Füllgrabe & Rosen, 2016b). Additionally, when all of the other factors had been considered, control of inhibition also contributed significantly to the final explanation model. The pattern, therefore, is of a potential changing reliance on hearing level and different cognitive skills when discriminating speech-in-noise. It is also the case that the type of speech-in-noise being attended to will also have an influence.

If declining sensory facility is an indicator of a general decline of brain architecture and cognitive functioning (Humes et al., 2013; Lindenberger & Baltes, 1994), it would imply that older adults with hearing loss, who rely more on cognitive support to discriminate speech-in-noise, will be the group of individuals who have the least cognitive resources available. Further, if this group have to reallocate more of the cognitive resources they have left to support speech discrimination, this will leave less for other important cognitive processes.

The implication of this is that with a smaller pool of cognitive resources, older individuals may only process incoming information at a ‘surface level’. The information being imparted to them may, therefore, be lost before being attended to and/or encoded into secondary memory. If this is the case, then true communication has not really taken place. At a basic level, in terms of the interaction of older, hearing-impaired individuals visiting an audiology department, if they are given lots of information verbally, they may not process it well. The same could hold true for any medical or non-medical interaction where full understanding of the implications of the imparted information needs to be properly processed. Linked with the general mental well-being that social interaction brings, it would seem appropriate that any strategies that could improve this communication should be further explored.

Working memory was a key cognitive skill for understanding speech-in-noise for the older listeners with hearing loss in Chapter 5. This result is similar to the findings of Füllgrabe and Rosen (2016b), whose meta-analysis concluded that for younger listeners with normal hearing, working memory was found to be relatively unimportant. Conversely, they did report a more significant role of working memory for older adults with hearing loss, along similar lines as Rudner et al. (2011). It may have been anticipated that for the thesis chapters involving participants across a wide age range with no self-reported hearing loss (although some minor losses were present on threshold testing), if they had all been older, relationships between cognitive functions and speech-in-noise task performance may have been more significant.

For the participants presenting with normal hearing or mild hearing loss in Chapter 2, relationships were demonstrated between speech-in-noise performance under very challenging conditions (BKB in speech shaped noise at -8dBSNR) and two cognitive functions. It was found that on average, individuals with faster information processing speed and better inhibition control performed more highly on the speech-in-noise task. It should also be noted that the inhibition task (a Stroop) did involve a degree of task switching alongside the inhibition, thus involving more executive control. Additionally, it is suggested that the Stroop task involves a strong element of speed of information processing (Uttl & Graf, 1997). Although control of inhibition and processing speed may tap similar processes, as both functions added to the explanation of speech-in-noise performance variance individually, they appear to offer different aspects of support.

It may also be the case that there may be physiological changes occurring in older adults' inner ear systems, that while not giving evidence of any significant change in hearing thresholds may still make discrimination of speech under challenging conditions more problematic. It has been suggested that this may be due to cochlear synaptopathy, although the evidence is not convincing (Guest, Munro, Prendergast, Millman, & Plack, 2018).

6.2. Directed training offers the greatest improvements in discriminating degraded speech signals, with limited evidence of near and far transfer and release of cognitive resources.

In Chapter 4 of this thesis, there was evidence that targeted specific auditory training, even after a relatively short period of exposure, was able to generate significant improvements in performance on the discrimination of degraded speech. This was illustrated in Chapter 4, as the type of degraded speech in which training was undertaken consistently demonstrated the greatest improvement in the same type of speech when compared to other speech training. There was, however, no clear evidence of near transfer to other speech tasks, although there was evidence of far transfer, as individuals' auditory working memory span (an LSPAN) increased when the working memory task was performed in the same particular type of trained degraded speech.

This result links to the wider discussion presented later in this chapter in Section 6.4, as working memory changes are only readily demonstrated when the working memory task is performed in the assessment material of the trained speech. This would suggest that following a period of auditory training, new speech memory patterns are developed, making the matching of perceived speech in the auditory working memory task less cognitively demanding. This is congruent with the Ease of Language Understanding (ELU) model (Rönnberg et al., 2013), but it was only evident in this thesis when auditory working memory tasks were performed under challenging conditions, such as degraded speech. There was no significant improvement noted for other cognitive tasks following auditory training despite audition improving. Indeed, there was no clear evidence that cognitive training improved cognitive performance, as no significant differences were consistently found between the cognitively trained groups and a control (Chapter 3). It could be that both the auditory and cognitive training were not intense or long enough to promote the desired changes. It does, however, appear that specific targeted auditory training offers improvements in the task trained and potentially some cognitive release. If an individual struggles to hear in background noise, then training to understand speech in that particular type of background noise should offer both improved audition and the potential to better process the incoming information, as the cognitive resources needed to understand the speech are reduced. However, the lack of any evidence that encoding into secondary memory is improved following training is a further concern (Chapters 4 and 5). This thesis did not demonstrate that this could occur with targeted auditory or cognitive function training. This would suggest that any communication training packages should consider offering some form of training to enhance memory encoding. This would be especially the case if, on assessment, an individual is found to be having problems with transfer of information from working to longer-term memory.

It may further be the case that any training to achieve the best outcomes should be performed in the sensory faculty that the desired outcome is seeking to improve, e.g., auditory training when seeking auditory improvements. The same may apply for any cognitive training. Although there was no clear evidence in Chapter 3 that training cognitive functions demonstrated either on-task improvements or far transfer improvements to a speech-in-noise task, the rationale for training cognitive functions seems reasonable. Cognitive functions have been demonstrated to be either correlated with or are able to explain some speech-in-noise performance variance. Training to improve these functions should, therefore, theoretically offer an individual more support. Other researchers reporting on cognitive training have found that improvement in the function being trained is very possible (Ball et al., 2002; Kueider, Parisi, Gross, & Rebok, 2012). What is not clear is whether any training can result in any far transfer of improvements in more loosely associated processes/tasks (Redick et al., 2013).

As cognitive functions play a more minor role in assisting with speech-in-noise task performance, far transfer was potentially beyond the scope of the research in Chapter 3. However, as they do offer significant support in the perception of degraded speech (Chapters 2, 3, 4 and 5), training them needs further exploration. Further, it makes sense that given that auditory training demonstrates improvements (Chapter 4), training cognitive functions auditorily would also seem to be the best course to pursue.

Sensory-specific training is easier to deliver for some cognitive functions than others. Working memory training fits easily into this type of design as, when exploring the memory model of Baddeley (2000), there are clear pathways for vision training (the visuospatial sketchpad) or auditory training (via the phonological store). Although there are clear links and crossovers between these two slave systems, they can be envisaged as having different pathways. It is also the case that information processing speed and the individual components of executive functioning, such as control of inhibition, may also allow for sensory-specific training fairly readily. Training executive functions in a sensory-specific manner is more problematic, as some of the functions potentially rely on different individual inputs and training strategies. In order to potentially achieve this type of training, there may be a need to devise sensory-specific tasks/training that tap these executive functions.

6.3. Different cognitive functions offer support for speech-in-noise discrimination in different types of background noise.

This topic has already been introduced as an area for discussion in Section 6.1 but is expanded on here. As in Heinrich, Henshaw, and Ferguson (2015), this thesis has demonstrated that performance of different speech assessments demonstrate different relationships with cognitive functions depending on the type of background noise. In the research for this thesis, two types of background noise were used: speech-shaped noise (for BKB assessments) and background four-talker babble (for QuickSIN assessments). The speech material was also slightly different for the two types of noise, with the QuickSIN sentences being slightly longer. BKB sentences are 3–6 words in length, whereas QuickSIN are 6–9.

Attempting to discriminate speech in the two types of background noise brings different challenges. The background speech babble used in the QuickSIN is informational and as such is harder to inhibit (Koelewijn et al., 2012), as there is a greater tendency to listen to what the babble is saying. It does, however, have gaps between the babble that allows for clearer glimpses of speech, which the listener can use to discriminate (or use support mechanisms to interpret) what was being said. The speech-shaped noise used to mask BKB sentences in this thesis has no gaps, but as the noise does not carry informational speech, it may be easier to inhibit.

In the chapters where both types of speech assessments were used, although the tests correlated significantly with each other in terms of participant performance, the cognitive functions evaluated to explain any performance differences were not consistent. Working memory, unsurprisingly, demonstrated a more consistent relationship when the longer QuickSIN sentences were used.

In Chapter 2, where significant relationships were demonstrated between speech-in-noise discrimination performance and the cognitive functions of information processing speed and prepotent inhibition control, no significant correlation was found for working memory. This may have been because only the shorter BKB sentences were used and not the QuickSIN. This was by design, in that the chapter specifically set out to investigate what functions, other than working memory may assist in highly challenging speech-in-noise performance. There were few auditory cues and the sentences were short. It was anticipated that the functions of processing speed and the ability of an individual to use their top-down processing would offer more support than working memory, and this was indeed the case. In the real world, different environments with diverse background noises and various speakers will pose different challenges to listeners. This will mean that as an individual moves from one environment to another, they themselves will have to potentially employ different cognitive functions (or a different mix of them) and be able to switch between them (Miyake et al., 2000). If tailored training is to work assessing performance with a range of speech material in different background noises may be necessary.

Continuing this theme, the type of working memory assessment would also appear to have a strong bearing on whether a relationship is demonstrated between working memory and speech-in-noise perception performance. Working memory demonstrated the most consistent relationship with speech-in-noise performance in this thesis when the working memory task involved listening to sentences with at least some degree of signal degradation. To quantify how much influence working memory has for any individual, it is proposed that working memory tasks should be presented auditorily in a marginally degraded fashion, such as a high SNR, as this appears to better demonstrate the type of cognitive support that the individual uses when listening to sentences in noise.

Further, it would seem appropriate that the other cognitive functions that have been identified in this thesis as demonstrating a significant relationship in at least one type of listening condition also need to be assessed if a bespoke training package is to be developed for an individual. This would, therefore, include assessments for information processing speed, both prepotent and on-task inhibition control, and an assessment exploring executive function. Executive functioning has not been directly explored in this thesis, but it is proposed that it should be considered as an essential extra to be included in any future test battery.

6.4. Proposed further research.

From the research in Chapters 4 and 5, there was evidence of improved on-task performance of discrimination of degraded speech following auditory training. There was also some partial evidence of far transfer after the training period, although this was limited. This could have been due to insufficient or inadequate training. However, given that improvements are possible (albeit small), even these small changes are worth pursuing if they enhance an individual's communication.

A hidden advantage of integrating an additional training program with any planned rehabilitation program involving hearing aids is that usage appears to increase (Martin, 2007). Although this may address the problem for older individuals with declining hearing due to presbycusis (and for some older and younger individuals with hearing loss acquired due to other pathologies) who have sought out hearing aids, it still leaves a large group of individuals of all ages with hearing difficulties who have eschewed any help. Alongside these will be further individuals with normal hearing but who have problems hearing in noise. For all of these groups, there is the potential that their poor communication may lead them to withdraw from social interaction, with the long-term social and mental health consequences that this may bring.

There are many reports appearing from longitudinal studies highlighting the links between hearing loss, cognition, hearing aids and social well-being (Amieva et al., 2015; Dawes et al., 2015; Deal et al., 2015; Lin, Ferrucci, et al., 2011). It can be hoped that these and other studies will continue to report with further information on how individuals' audition may change over a period of years, whilst also including reports on cognitive and emotional well-being and the protection offered against declines by aiding hearing and/or other intervention strategies.

It is proposed that alongside these large, long-term studies, in order to seek more detailed information, a smaller longitudinal study obtaining both quantitative and qualitative data on intervention strategies—including auditory/cognitive training—would provide valuable extra information. Ideally, this would be run as a randomised control trial (RCT), although ethical difficulties might arise if it appears that treatment is being withheld (e.g., not issuing hearing aids or restricting access to auditory training). These smaller studies could employ a mixed methods design to further explore the reasons for hearing aid non-use and the link with social isolation/depression. As a long-term study, this may also help to demonstrate which of the four hypotheses offered by Wayne and Johnsrude (2015) on the link between hearing loss and cognitive decline may hold the most validity; it would also explore whether any sensory underload can be reversed if the cascade hypothesis holds true (Sekuler & Blake, 1987). Indeed, even if the common cause hypothesis (Lindenberger & Baltes, 1994) were true, there would still be a case to be made for enriching communication either with or without amplification and/or training, as improving audition or cognitive functioning would still lead to improved outcomes.

Having proposed the above, it is the further assertion of this author that the studies should be based on the principle that communication training works best when it is targeted. Before using a ''scatter-gun'' or trial-and-error approach to training, areas of strong or weak listening and cognitive performance need to be identified by an assessment battery. Indeed, it is argued by

Cameron and Dillon (2011), albeit for assessing and training children with spatial APD problems, that unless the correct deficit is identified, training other areas may well not be of any great assistance. It could also potentially be seen as a dis-service, as the patient feels that the problem they have presented with is being addressed when indeed it may not be. Assessment using a variety of speech tests with different types of background noise and sentence lengths, tests of spatial hearing and auditory closure, alongside an assessment of executive control and the various cognitive skills that have been identified as potentially able to enhance or support a listener's ability to hear in noise, is proposed. This will entail assessment of the following cognitive functions:

- Processing speed
- Working memory
- Real time and prepotent control of inhibition
- Simple response latency
- Recall/recognition
- Executive control (to explore updating and task-switching ability)

These will form the basis of any assessment battery. It is suggested that the majority of assessments and all designed training packages will, wherever possible, be delivered/presented auditorily.

The possibility of audiologists administering cognitive screening has been introduced in a paper by Shen, Anderson, Arehart, and Souza (2016). Shen at al.'s study reported there were many potential benefits that could be anticipated for audiologists administering one of the cognitive assessment test batteries such as MMSE or Min-cog. As audiologists receive counselling training, they could employ these skills to debrief and appropriately manage patients needing onward referral for further assessment. Indeed, it may be that older adults who present at audiology clinics may have a higher incidence of mild cognitive impairment (MCI), which could be anticipated with the proven link between hearing loss and cognitive decline/dementia (Lin et al., 2013). The authors conclude by proffering the idea that cognitive interventions could offer the opportunity to optimise the outcomes of audiological interventions.

That is not to say that this type of screening test for MCI plays a role in assessing communication and training needs, as the testing is different from such an assessment, as is the end goal. One is intended to design training to improve communication and the other to spot early signs of dementia.

The training model for communication suggested here in Figure 1 is not aimed at identifying MCI, but there will need to be safeguards in place for onward referral of patients who give evidence of poor cognitive functioning. This training model would first need a feasibility study and then a pilot to test the methods. As the older population becomes more adept with information technology, ready access to individualised training packages may offer an opportunity for improved communication and the accompanying audition and social interaction improvements that this brings.

It remains unclear whether training should be targeted at weak areas of an individual’s performance or if it is preferable to work on enhancing listening/cognitive skills that they already perform well. Not training the skills that a listener relies upon may have the effect of reducing their efficiency as a support mechanism. Alternatively, training up weaker skills may be beneficial, but if these skills are resistant to improvement, then the listener will not derive any benefit from the training. Work surrounding these areas need to be included in the pilot study.

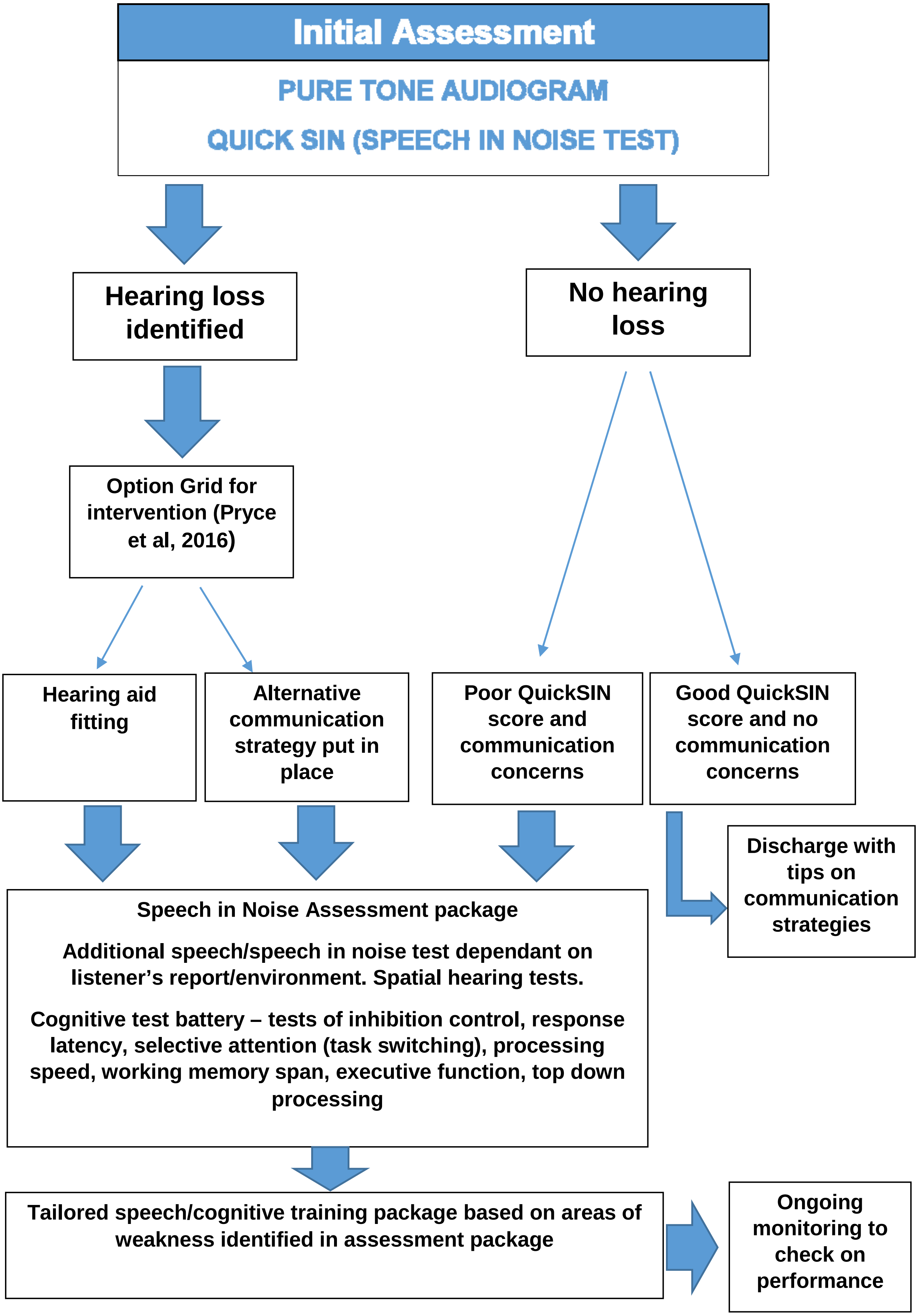


Figure 1. Proposed assessment pathway to establish need for cognitive training.

REFERENCES

Acar, B., Yurekli, M. F., Babademez, M. A., Karabulut, H., & Karasen, R. M. (2011). Effects of hearing aids on cognitive functions and depressive signs in elderly people. *Archives of Gerontology and Geriatrics, 52*(3), 250-252. doi: 10.1016/j.archger.2010.04.013

Action on Hearing Loss. (2015). Facts and figures on hearing loss, deafness and tinnitus. Retrieved 25th July, 2016, from http://www.actiononhearingloss.org.uk

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *Int J Audiol, 47 Suppl 2*, S53-71. doi: 10.1080/14992020802301142

Amieva, H., Ouvrard, C., Giulioli, C., Meillon, C., Rullier, L., & Dartigues, J. F. (2015). Self-reported hearing loss, hearing AIDS, and cognitive decline in elderly adults: A 25-year study. *Journal of the American Geriatrics Society, 63*(10), 2099-2104. doi: 10.1111/jgs.13649

Anderson, S., Parbery-Clark, A., Han-Gyol, Y., & Kraus, N. (2011). A Neural Basis of Speech-in-Noise Perception in Older Adults. *Ear and Hearing, 32*(6), 750-757. doi: 10.1097/AUD.0b013e31822229d3

Anderson, S., Parbery-Clark, A., White-Schwoch, T., & Kraus, N. (2012). Aging Affects Neural Precision of Speech Encoding. *Journal of Neuroscience, 32*(41), 14156-14164. doi: 10.1523/jneurosci.2176-12.2012

Anderson, S., Parbery-Clark, A., White-Schwoch, T., & Kraus, N. (2013a). Auditory Brainstem Response to Complex Sounds Predicts Self-Reported Speech-in-Noise Performance. *Journal of Speech Language and Hearing Research, 56*(1), 31-43. doi: 10.1044/1092-4388(2012/12-0043)

Anderson, S., Parbery-Clark, A., White-Schwoch, T., & Kraus, N. (2013b). A dynamic auditory-cognitive system supports speech-in-noise perception in older adults. *Hearing Research, 300*, 18-32. doi: 10.1016/j.heares.2013.03.006

AngelSim, T. (2012). Retrieved from http://www.tigerspeech.com/angelsim/angelsim

Arlinger, S., Gatehouse, S., Bentler, R. A., Byrne, D., Cox, R. M., Dirks, D. D., . . . Willott, J. F. (1996). Report of the Eriksholm Workshop on auditory deprivation and acclimatization. *Ear Hear, 17*(3 Suppl), 87s-98s.

Au, J., Buschkuehl, M., Duncan, G. J., & Jaeggi, S. M. (2016). There is no convincing evidence that working memory training is NOT effective: A reply to Melby-Lervåg and Hulme (2015). *Psychonomic Bulletin and Review, 23*(1), 331-337. doi: 10.3758/s13423-015-0967-4

Au, J., Sheehan, E., Tsai, N., Duncan, G. J., Buschkuehl, M., & Jaeggi, S. M. (2015). Improving fluid intelligence with training on working memory: a meta-analysis. *Psychonomic Bulletin & Review, 22*(2), 366-377. doi: 10.3758/s13423-014-0699-x

Audacity, T. (2017). Free Audio Editor and Recorder (Version 2.1.3) [Computer Program]. Retrieved from http://audacity.sourceforge.net/

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417-423. doi: 10.1016/S1364-6613(00)01538-2

Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.

Baddeley, A. D., & Hitch , G. (1974). Working memory. *In G.H. Bower (Ed.), The pyschology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47-89).

Bajo, V. M., & King, A. J. (2012). Cortical modulation of auditory processing in the midbrain. *Frontiers in Neural Circuits, 6*, 114. doi: 10.3389/fncir.2012.00114

Baldwin, C. L., & Ash, I. K. (2011). Impact of Sensory Acuity on Auditory Working Memory Span in Young and Older Adults. *Psychology and Aging, 26*(1), 85-91. doi: 10.1037/a0020360

Ball, K., Berch, D. B., Helmers, K. F., Jobe, J. B., Leveck, M. D., Marsiske, M., . . . Willis, S. L. (2002). Effects of cognitive training interventions with older adults: A randomized controlled trial. *Journal of the American Medical Association, 288*(18), 2271-2281. doi: 10.1001/jama.288.18.2271

Bamiou, D. E., Musiek, F. E., & Luxon, L. M. (2001). Aetiology and clinical presentations of auditory processing disorders - a review. *Archives of Disease in Childhood, 85*(5), 361-365. doi: 10.1136/adc.85.5.361

Barcroft, J., Sommers, M. S., Tye-Murray, N., Mauze, E., Schroy, C., & Spehar, B. (2011). Tailoring auditory training to patient needs with single and multiple talkers: Transfer-appropriate gains on a four-choice discrimination test. *International journal of audiology, 50*(11). doi: 10.3109/14992027.2011.599868

Bench, J., Kowal, A., & Bamford, J. (1979). The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *British journal of audiology, 13*(3), 108-112. doi: 10.3109/03005367909078884

Besser, J., Koelewijn, T., Zekveld, A. A., Kramer, S. E., & Festen, J. M. (2013a). How linguistic closure and verbal working memory relate to speech recognition in noise--a review. *Trends Amplif, 17*(2), 75-93. doi: 10.1177/1084713813495459

Besser, J., Koelewijn, T., Zekveld, A. A., Kramer, S. E., & Festen, J. M. (2013b). How Linguistic Closure and Verbal Working Memory Relate to Speech Recognition in Noise-A Review. *Trends in amplification, 17*(2), 75-93. doi: 10.1177/1084713813495459

Besser, J., Zekveld, A. A., Kramer, S. E., Ronnberg, J., & Festen, J. M. (2012). New Measures of Masked Text Recognition in Relation to Speech-in-Noise Perception and Their Associations With Age and Cognitive Abilities. *Journal of Speech Language and Hearing Research, 55*(1), 194-209. doi: 10.1044/1092-4388(2011/11-0008)

Bialystok, E., Craik, F., & Luk, G. (2008). Cognitive Control and Lexical Access in Younger and Older Bilinguals. *Journal of Experimental Psychology: Learning Memory and Cognition, 34*(4), 859-873. doi: 10.1037/0278-7393.34.4.859

Blanca, M., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017). *Non-normal data: Is ANOVA still a valid option?* (Vol. 29).

Boot, W. R., Simons, D. J., Stothart, C., & Stutts, C. (2013). The Pervasive Problem With Placebos in Psychology: Why Active Control Groups Are Not Sufficient to Rule Out Placebo Effects. *Perspectives on Psychological Science, 8*(4), 445-454. doi: 10.1177/1745691613491271

. BrainHQ. (2012): Posit Science. Retrieved from BrainHQ.com

Brehmer, Y., Westerberg, H., & Bäckman, L. (2012). Working-memory training in younger and older adults: training gains, transfer, and maintenance. *Frontiers in Human Neuroscience, 6*, 63. doi: 10.3389/fnhum.2012.00063

Brouns, K., El Refaie, A and Pryce, H. (2011). Auditory Training and Adult Rehabilitation:
A Critical Review of the Evidence.

BSA. (2011, 6th February 2012). Pure-tone air-conduction and bone-conduction threshold audiometry with and without masking. Retrieved 15th January, 2017, from http://www.thebsa.org.uk/resources/pure-tone-air-bone-conduction-threshold-audiometry-without-masking/

BSA. (2016). Ear examination. Retrieved 15th January, 2017, from http://www.thebsa.org.uk/resources/recommended-procedure-ear-examination/

Burk, M. H., & Humes, L. E. (2007). Effects of training on speech recognition performance in noise using lexically hard words. *Journal of Speech Language and Hearing Research, 50*(1), 25-40. doi: 10.1044/1092-4388(2007/003)

Burk, M. H., & Humes, L. E. (2008). Effects of long-term training on aided speech-recognition performance in noise in older adults. *Journal of Speech Language and Hearing Research, 51*(3), 759-771. doi: 10.1044/1092-4388(2008/054)

Burk, M. H., Humes, L. E., Amos, N. E., & Strauser, L. E. (2006). Effect of training on word-recognition performance in noise for young normal-hearing and older hearing-impaired listeners. *Ear and Hearing, 27*(3), 263-278. doi: 10.1097/01.aud.0000215980.21158.a2

Byrne, D., Dillon, H., Ching, T., Katsch, R., & Keidser, G. (2001). NAL-NL1 procedure for fitting nonlinear hearing aids: Characteristics and comparisons with other procedures. *Journal of the American Academy of Audiology, 12*(1), 37-51.

Cabeza, R. (2002). Hemispheric asymmetry reduction in older adults: The HAROLD model. *Psychology and Aging, 17*(1), 85-100. doi: 10.1037//0882-7974.17.1.85

Cameron, S., & Dillon, H. (2011). Development and Evaluation of the LiSN & Learn Auditory Training Software for Deficit-Specific Remediation of Binaural Processing Deficits in Children: Preliminary Findings. *Journal of the American Academy of Audiology, 22*(10). doi: 10.3766/jaaa.22.10.6

Cameron, S., Glyde, H., & Dillon, H. (2012). Efficacy of the LiSN & Learn auditory training software: randomized blinded controlled study. *Audiology Research, 2*(1), e15. doi: 10.4081/audiores.2012.e15

Carabellese, C., Appollonio, I., Rozzini, R., Bianchetti, A., Frisoni, G. B., Frattola, L., & Trabucchi, M. (1993). SENSORY IMPAIRMENT AND QUALITY-OF-LIFE IN A COMMUNITY ELDERLY POPULATION. *Journal of the American Geriatrics Society, 41*(4), 401-407.

Carlson, R., Fant, G., & Granstroem, B. (1974). 2-FORMANT MODELS, PITCH, AND VOWEL PERCEPTION. *Acustica, 31*(6), 360-362.

Cepeda, N. J., Kramer, A. F., & de Sather, J. (2001). Changes in executive control across the life span: Examination of task-switching performance. *Developmental Psychology, 37*(5), 715-730. doi: 10.1037//0012-1649.37.5.715

CHABA. (1988). SPEECH UNDERSTANDING AND AGING .99. *Journal of the Acoustical Society of America, 83*(3), 859-895.

Chermak, G. D., & Musiek, F. E. (2002). Auditory training: Principles and approaches for remediating and managing auditory processing disorders. *Seminars in Hearing, 23*(4), 297-308. doi: 10.1055/s-2002-35878

Choi, A. Y., Shim, H. J., Lee, S. H., Yoon, S. W., & Joo, E. J. (2011). Is Cognitive Function in Adults with Hearing Impairment Improved by the Use of Hearing Aids? *Clinical and Experimental Otorhinolaryngology, 4*(2), 72-76. doi: 10.3342/ceo.2011.4.2.72

COGMED, R. Working memory training. Retrieved from http://www.cogmed.com/program

Cohen, J. (1992). A POWER PRIMER. *Psychological Bulletin, 112*(1), 155-159. doi: 10.1037/0033-2909.112.1.155

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). ON THE CONTROL OF AUTOMATIC PROCESSES - A PARALLEL DISTRIBUTED-PROCESSING ACCOUNT OF THE STROOP EFFECT. *Psychological Review, 97*(3), 332-361. doi: 10.1037//0033-295x.97.3.332

Conway, A. R. A., & Getz, Sarah J. (2010). Dispatch: Cognitive Ability: Does Working Memory Training Enhance Intelligence? *Current Biology, 20*, R362-R364. doi: 10.1016/j.cub.2010.03.001

Craik, F. I. M. (1983). ON THE TRANSFER OF INFORMATION FROM TEMPORARY TO PERMANENT MEMORY. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 302*(1110), 341-359. doi: 10.1098/rstb.1983.0059

Craik, F. I. M. (1986). A functional account of age differences in memory. In F. Klix & H. Hagendorf (Eds.), *Human Memory and Cognitive Capabilities, Mechanism and Performance*. Amsterdam: Elsevier.

Craik, F. I. M., & Byrd, M. (1982). Aging and cognitve deficits: the role of attentional resources. In F. I. M. Craik & S. E. Trehub (Eds.), *Aging and Cognitive Processes*. New York: NY: Plenum.

Crow, R., Gage, H., Hampson, S., Hart, J., Kimber, A., & Thomas, H. (1999). The role of expectancies in the placebo effect and their use in the delivery of health care: A systematic review. *Health Technology Assessment, 3*(3), iii-48.

Cutler, A., & Clifton, C. (1999). The neurocogntion of language. In C. M. Brown & P. Hagoort (Eds.), (pp. 123-166). Oxford: Oxford University Press.

Daneman, M., & Carpenter, P. A. (1980). INDIVIDUAL-DIFFERENCES IN WORKING MEMORY AND READING. *Journal of Verbal Learning and Verbal Behavior, 19*(4), 450-466. doi: 10.1016/s0022-5371(80)90312-6

Darowski, E. S., Helder, E., Zacks, R. T., Hasher, L., & Hambrick, D. Z. (2008). Age-related differences in cognition: The role of distraction control. *Neuropsychology, 22*(5), 638-644. doi: 10.1037/0894-4105.22.5.638

Davis, M. H., Johnsrude, I. S., Hervais-Adelman, A., Taylor, K., & McGettigan, C. (2005). Lexical information drives; Perceptual learning of distorted speech: Evidence from the comprehension of noise-vocoded sentences. *Journal of Experimental Psychology-General, 134*(2). doi: 10.1037/0096-3445.134.2.222

Dawes, P., Emsley, R., Cruickshanks, K. J., Moore, D. R., Fortnum, H., Edmondson-Jones, M., . . . Munro, K. J. (2015). Hearing Loss and Cognition: The Role of Hearing Aids, Social Isolation and Depression. *Plos One, 10*(3). doi: 10.1371/journal.pone.0119616

Dawes, P., & Munro, K. J. (2017). Auditory Distraction and Acclimatization to Hearing Aids. *Ear & Hearing (01960202), 38*(2), 174.

Dawes, P., Munro, K. J., Kalluri, S., & Edwards, B. (2014). Acclimatization to hearing aids. *Ear Hear, 35*(2), 203-212. doi: 10.1097/AUD.0b013e3182a8eda4

Deal, J. A., Sharrett, A. R., Albert, M. S., Coresh, J., Mosley, T. H., Knopman, D., . . . Lin, F. R. (2015). Hearing impairment and cognitive decline: A pilot study conducted within the atherosclerosis risk in communities Neurocognitive Study. *American Journal of Epidemiology, 181*(9), 680-690. doi: 10.1093/aje/kwu333

Deary, I. J., Der, G., & Ford, G. (2001). Reaction times and intelligence differences - A population-based cohort study. *Intelligence, 29*(5), 389-399. doi: 10.1016/s0160-2896(01)00062-9

Dell, G. S. (1986). A SPREADING-ACTIVATION THEORY OF RETRIEVAL IN SENTENCE PRODUCTION. *Psychological Review, 93*(3), 283-321. doi: 10.1037//0033-295x.93.3.283

Desjardins, J. L., & Doherty, K. A. (2013). Age-Related Changes in Listening Effort for Various Types of Masker Noises. *Ear and Hearing, 34*(3), 261-272. doi: 10.1097/AUD.0b013e31826d0ba4

Doherty, K. A., & Desjardins, J. L. (2015). The benefit of amplification on auditory working memory function in middle-aged and young-older hearing impaired adults. *Frontiers in psychology, 6*(JUN). doi: 10.3389/fpsyg.2015.00721

Dubno, J. R., Dirks, D. D., & Morgan, D. E. (1984). EFFECTS OF AGE AND MILD HEARING-LOSS ON SPEECH RECOGNITION IN NOISE. *Journal of the Acoustical Society of America, 76*(1), 87-96. doi: 10.1121/1.391011

Duquesnoy, A. J. (1983). THE INTELLIGIBILITY OF SENTENCES IN QUIET AND IN NOISE IN AGED LISTENERS. *Journal of the Acoustical Society of America, 74*(4), 1136-1144. doi: 10.1121/1.390037

Era, P., Jokela, J., Ovarnberg, Y., & Heikkinen, E. (1986). PURE-TONE THRESHOLDS, SPEECH UNDERSTANDING, AND THEIR CORRELATES IN SAMPLES OF MEN OF DIFFERENT AGES. *Audiology, 25*(6), 338-352.

Eysenck, M. W. a. K., M.T. (2005). *Cognitive Psychology: A Students handbook* (5th ed.). Hove: Psycholgy Press Limited.

Fant, G. (1960). *Acoustic Theroy of Speech Production*. The Hague, Netherlands: Mouton & Co.

Faul, F., Erdfelder, E., Lang, A., -G. , & Buchner, A. (2007). G*Power 3: A flexible statisic power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Fellowes, J. M., Remez, R. E., & Rubin, P. E. (1997). Perceiving the sex and identity of a talker without natural vocal timbre. *Perception & Psychophysics, 59*(6), 839-849. doi: 10.3758/bf03205502

Ferguson, M., & Henshaw, H. (2015a). Computer and Internet Interventions to Optimize Listening and Learning for People With Hearing Loss: Accessibility, Use, and Adherence. *American journal of audiology, 24*(3), 338-343. doi: 10.1044/2015_aja-14-0090

Ferguson, M., & Henshaw, H. (2015b). How Does Auditory Training Work? Joined-Up Thinking and Listening. *Seminars in Hearing, 36*(4), 237-249. doi: 10.1055/s-0035-1564456

Ferguson, M., Henshaw, H., Clark, D., & Moore, D. (2014). Benefits of Phoneme Discrimination Training in a Randomized Controlled Trial of 50- to 74-Year-Olds With Mild Hearing Loss. *Ear and Hearing, 35*(4), E110-E121.

Ferguson, M. A., & Henshaw, H. (2015c). Auditory training can improve working memory, attention, and communication in adverse conditions for adults with hearing loss. *Frontiers in psychology, 6*. doi: 10.3389/fpg.2015.00556

Field, A. (2016). *An Adventure in Statistics: The Reality Enigma*. London: Sage.

Foo, C., Rudner, M., Rönnberg, J., & Lunner, T. (2007). Recognition of speech in noise with new hearing instrument compression release settings requires explicit cognitive storage and processing capacity. *Journal of the American Academy of Audiology, 18*(7), 618-631. doi: 10.3766/jaaa.18.7.8

Fu, Q.-J., & Galvin, J. J., 3rd. (2007a). Perceptual learning and auditory training in cochlear implant recipients. *Trends in amplification, 11*(3), 193-205. doi: 10.1177/1084713807301379

Fu, Q.-J., & Galvin, J. J., III. (2007b). Computer-assisted speech training for cochlear implant patients: Feasibility, outcomes, and future directions. *Seminars in Hearing, 28*(2), 142-150. doi: 10.1055/s-2007-973440

Fuellgrabe, C., Moore, B. C. J., & Stone, M. A. (2015). Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition. *Frontiers in Aging Neuroscience, 6*. doi: 10.3389/fnagi.2014.00347

Füllgrabe, C., & Rosen, S. (2016a) Investigating the role of working memory in speech-in-noise identification for listeners with normal hearing. *Vol. 894. Advances in Experimental Medicine and Biology* (pp. 29-36).

Füllgrabe, C., & Rosen, S. (2016b). On the (un)importance of working memory in speech-in-noise processing for listeners with normal hearing thresholds. *Frontiers in psychology, 7*(AUG). doi: 10.3389/fpsyg.2016.01268

Galbraith, G. C., Arbagey, P. W., Branski, R., Comerci, N., & Rector, P. M. (1995). Intelligible speech encoded in the human brain stem frequency following response. *Neuroreport, 6*(17), 2363-2367. doi: 10.1097/00001756-199511270-00021

Galbraith, G. C., Gutterson, R. P., Levy, D. S., Mussey, J. L., Sabatasso, F. A., & Wasserman, R. I. (2004). Correlated brain stem and cortical evoked responses to auditory tone change. *Neuroreport, 15*(17), 2613-2616. doi: 10.1097/00001756-200412030-00010

Gatehouse, S., & Noble, I. (2004). The Speech, Spatial and Qualities of Hearing Scale (SSQ). *International journal of audiology, 43*(2), 85-99. doi: 10.1080/14992020400050014

Gates, N. J., & Sachdev, P. (2014). Is Cognitive Training an Effective Treatment for Preclinical and Early Alzheimer's Disease? *Journal of Alzheimers Disease, 42*, S551-S559. doi: 10.3233/jad-141302

Gianopoulos, I., Stephens, D., & Davis, A. (2002). Follow up of people fitted with hearing aids after adult hearing screening: the need for support after fitting. *British Medical Journal, 325*(7362), 471-471. doi: 10.1136/bmj.325.7362.471

Ginsburg, N., & Karpiuk, P. (1994). RANDOM GENERATION - ANALYSIS OF THE RESPONSES. *Perceptual and Motor Skills, 79*(3), 1059-1067.

Giroud, N., Lemke, U., Reich, P., Matthes, K. L., & Meyer, M. (2017). Research Paper: The impact of hearing aids and age-related hearing loss on auditory plasticity across three months – An electrical neuroimaging study. *Hearing Research, 353*, 162-175. doi: 10.1016/j.heares.2017.06.012

Goedert, K. M., & Miller, J. (2008). Spacing practice sessions across days earlier rather than later in training improves performance of a visuomotor skill. *Experimental Brain Research, 189*(2), 189-197. doi: 10.1007/s00221-008-1414-9

Goodin, D. S., Squires, K. C., Henderson, B. H., & Starr, A. (1978). AGE-RELATED VARIATIONS IN EVOKED-POTENTIALS TO AUDITORY-STIMULI IN NORMAL HUMAN SUBJECTS. *Electroencephalography and Clinical Neurophysiology, 44*(4), 447-458. doi: 10.1016/0013-4694(78)90029-9

Gopinath, B., Hickson, L., Schneider, J., McMahon, C. M., Burlutsky, G., Leeder, S. R., & Mitchell, P. (2012). Hearing-impaired adults are at increased risk of experiencing emotional distress and social engagement restrictions five years later. *Age and Ageing, 41*(5), 618-623. doi: 10.1093/ageing/afs058

GraphPad. GraphPad software. Retrieved from https://www.graphpad.com/

Greenberg, S. (1980). WPP, no 52: Temporal neural coding of pitch and vowel quality. *Working Papers in Phonetics*.

Greenwood, P. M. (2007). Functional Plasticity in Cognitive Aging: Review and Hypothesis. *Neuropsychology, 21*(6), 657-673. doi: 10.1037/0894-4105.21.6.657

Guest, H., Munro, K. J., Prendergast, G., Millman, R. E., & Plack, C. J. (2018). Impaired speech perception in noise with a normal audiogram: No evidence for cochlear synaptopathy and no relation to lifetime noise exposure. *Hearing Research, 364*, 142-151. doi: 10.1016/j.heares.2018.03.008

Hallberg, L. R. M., Hallberg, U., & Kramer, S. E. (2008). Self-reported hearing difficulties, communication strategies and psychological general well-being (quality of life) in patients with acquired hearing impairment. *Disability and Rehabilitation, 30*(3), 203-212. doi: 10.1080/09638280701228073

Harrison Bush, A. L., Lister, J. J., Lin, F. R., Betz, J., & Edwards, J. D. (2015). Peripheral Hearing and Cognition: Evidence From the Staying Keen in Later Life (SKILL) Study. *Ear Hear, 36*(4), 395-407. doi: 10.1097/aud.0000000000000142

Hasher, L., & Zacks, R. T. (1979). AUTOMATIC AND EFFORTFUL PROCESSES IN MEMORY. *Journal of Experimental Psychology-General, 108*(3), 356-388. doi: 10.1037//0096-3445.108.3.356

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation* (pp. 193-225). New York: Academic Press.

Hawkins, D. B. (2005). Effectiveness of counseling-based adult group aural rehabilitation programs: a systematic review of the evidence. *Journal of the American Academy of Audiology, 16*(7), 485-493. doi: 10.3766/jaaa.16.7.8

Hegarty, M., Shah, P., & Miyake, A. (2000). Constraints on using the dual-task methodology to specify the degree of central executive involvement in cognitive tasks. *Memory & Cognition, 28*(3), 376-385. doi: 10.3758/bf03198553

Heinrich, A., Henshaw, H., & Ferguson, M. A. (2015). The relationship of speech intelligibility with hearing sensitivity, cognition, and perceived hearing difficulties varies for different speech perception tests. *Frontiers in psychology, 6*. doi: 10.3389/fpsyg.2015.00782

Helfer, K. S., & Freyman, R. L. (2014). Stimulus and listener factors affecting age-related changes in competing speech perception. *Journal of the Acoustical Society of America, 136*(2), 748-759. doi: 10.1121/1.4887463

Henderson Sabes, J., & Sweetow, R. W. (2007). Variables predicting outcomes on listening and communication enhancement (LACE) training. *International journal of audiology, 46*(7), 374-383.

Henshaw, H., & Ferguson, M. A. (2013). Efficacy of Individual Computer-Based Auditory Training for People with Hearing Loss: A Systematic Review of the Evidence. *Plos One, 8*(5). doi: 10.1371/journal.pone.0062836

Henshaw, H., & Ferguson, M. A. (2014). *Assessing the benefits of auditory training to real-world listening: identifying appropriate and sensitive outcomes.* Paper presented at the ISAAR 2013: Auditory Plasticity - Listening with the Brain. 4th Symposium on Auditory and Audiological Research Nyborg.

Hervais-Adelman, A., Davis, M. H., Johnsrude, I. S., & Carlyon, R. P. (2008). Perceptual learning of noise vocoded words: Effects of feedback and lexicality. *Journal of Experimental Psychology-Human Perception and Performance, 34*(2). doi: 10.1037/0096-1523.34.2.460

Hervais-Adelman, A. G., Davis, M. H., Johnsrude, I. S., Taylor, K. J., & Carlyon, R. P. (2011). Generalization of Perceptual Learning of Vocoded Speech. *Journal of Experimental Psychology-Human Perception and Performance, 37*(1). doi: 10.1037/a0020772

Holland, C. A., & Rabbitt, P. M. A. (1991). AGING MEMORY - USE VERSUS IMPAIRMENT. *British Journal of Psychology, 82*, 29-38.

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics, 6*, 65-70.

Humes, L. E. (2002). Factors underlying the speech-recognition performance of elderly hearing-aid wearers. *Journal of the Acoustical Society of America, 112*(3), 1112-1132. doi: 10.1121/1.1499132

Humes, L. E., Burk, M. H., Strauser, L. E., & Kinney, D. L. (2009). Development and efficacy of a frequent-word auditory training protocol for older adults with impaired hearing. *Ear and Hearing, 30*(5), 613-627. doi: 10.1097/AUD.0b013e3181b00d90

Humes, L. E., Busey, T. A., Craig, J., & Kewley-Port, D. (2013). Are age-related changes in cognitive function driven by age-related changes in sensory processing? *Atten Percept Psychophys, 75*(3), 508-524. doi: 10.3758/s13414-012-0406-9

Humes, L. E., & Dunbo, J. R. (2010). *Factors Affecting Speech Understanding in Older Adults.* (Vol. 34). New York: Springer.

Humes, L. E., Kinney, D. L., Brown, S. E., Kiener, A. L., & Quigley, T. M. (2014). The effects of dosage and duration of auditory training for older adults with hearing impairment. *Journal of the Acoustical Society of America, 136*(3), EL224-EL230. doi: 10.1121/1.4890663

Humes, L. E., Lee, J. H., & Coughlin, M. P. (2006). Auditory measures of selective and divided attention in young and older adults using single-talker competition. *Journal of the Acoustical Society of America, 120*(5), 2926-2937. doi: 10.1121/1.2354070

Humes, L. E., Wilson, D. L., Barlow, N. N., Garner, C. B., & Amos, N. (2002). Longitudinal Changes in Hearing Aid Satisfaction and Usage in the Elderly Over a Period of One or Two Years After Hearing Aid Delivery*,* 428.

IEEE. (1969). IEEE Recommended Pratice for Speech Quality Measurements. *IEEE Transactions on Audio and Electroacoustics, 17*(3), 225-246. doi: 10.1109/TAU.1969.1162058

Inquisit. Millisecond. Retrieved from http://www.millisecond.com/

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proc Natl Acad Sci U S A, 105*(19), 6829-6833. doi: 10.1073/pnas.0801268105

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Shah, P. (2011). Short- and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences, 108*(25), 10081-10086. doi: 10.1073/pnas.1103228108

Janse, E., & Jesse, A. (2014). Working memory affects older adults' use of context in spoken-word recognition. *Quarterly Journal of Experimental Psychology, 67*(9), 1842-1862. doi: 10.1080/17470218.2013.879391

Jerger, J. (2006). Auditory training. *Journal of the American Academy of Audiology, 17*(8).

Jerger, J., Jerger, S., & Pirozzolo, F. (1991). CORRELATIONAL ANALYSIS OF SPEECH AUDIOMETRIC SCORES, HEARING-LOSS, AGE, AND COGNITIVE-ABILITIES IN THE ELDERLY. *Ear and Hearing, 12*(2), 103-109. doi: 10.1097/00003446-199104000-00004

Kail, R., & Salthouse, T. A. (1994). PROCESSING SPEED AS A MENTAL-CAPACITY. *Acta Psychologica, 86*(2-3), 199-225. doi: 10.1016/0001-6918(94)90003-5

Kalluri, S., & Humes, L. E. (2012). Hearing Technology and Cognition. *American journal of audiology, 21*(2), 338-343. doi: 10.1044/1059-0889(2012/12-0026)

Keselman, J. C., Lix, L. M., & Keselman, H. J. (1996). The analysis of repeated measurements: A quantitative research synthesis. *British Journal of Mathematical and Statistical Psychology, 49*(2), 275-298. doi: doi:10.1111/j.2044-8317.1996.tb01089.x

Kidd, G. R., Watson, C. S., & Gygi, B. (2007). Individual differences in auditory abilities. *Journal of the Acoustical Society of America, 122*(1), 418-435. doi: 10.1121/1.2743154

Kiessling, J., Pichora-Fuller, M. K., Gatehouse, S., Stephens, D., Arlinger, S., Chisolm, T., . . . von Wedel, H. (2003a). Candidature for and delivery of audiological services: Special needs of older people. *International journal of audiology, 42*(Supplement 2), 2S92-92S101.

Kiessling, J., Pichora-Fuller, M. K., Gatehouse, S., Stephens, D., Arlinger, S., Chisolm, T., . . . von Wedel, H. (2003b). Candidature for and delivery of audiological services: special needs of older people. *International journal of audiology, 42*, S92-S101.

Killion, M. C., Niquette, P. A., Gudmundsen, G. I., Revit, L. J., & Banerjee, S. (2004). Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America, 116*(4), 2395-2405. doi: 10.1121/1.1784440

Klatt, D. H. (1980). SOFTWARE FOR A CASCADE-PARALLEL FORMANT SYNTHESIZER. *Journal of the Acoustical Society of America, 67*(3), 971-995. doi: 10.1121/1.383940

Knight, S., & Heinrich, A. (2017). Different Measures of Auditory and Visual Stroop Interference and Their Relationship to Speech Intelligibility in Noise. *Frontiers in psychology, 8*. doi: 10.3389/fpsyg.2017.00230

Kochkin, S. (2011). MarkeTrak VIII Mini-BTEs tap new market, users more satisfied. *The Hearing Journal, 64*(3), 17-18,20,22,24. doi: 10.1097/01.HJ.0000395478.70959.b1

Koelewijn, T., Zekveld, A. A., Festen, J. M., Ronnberg, J., & Kramer, S. E. (2012). Processing load induced by informational masking is related to linguistic abilities. *International journal of otolaryngology, 2012*, 865731-865731. doi: 10.1155/2012/865731

Kraus, N. (2012). Biological impact of music and software-based auditory training. *Journal of Communication Disorders, 45*(6). doi: 10.1016/j.jcomdis.2012.06.005

Kricos, P. B., & McCarthy, P. (2007). From ear to there: A historical perspective on auditory training. *Seminars in Hearing, 28*(2), 89-98. doi: 10.1055/s-2007-973435

Kueider, A. M., Parisi, J. M., Gross, A. L., & Rebok, G. W. (2012). Computerized Cognitive Training with Older Adults: A Systematic Review. *Plos One, 7*(7), e40588. doi: 10.1371/journal.pone.0040588

Lehrl, S., Funk, R., & Seifert, K. (2005). The first hearing aid increases mental efficiency. *Hno, 53*(10), 852-862. doi: 10.1007/s00106-005-1253-y

Lesperance, M. M., & Burmeister, M. (2008). Commentary on "Occupational noise, smoking and a high body mass index are risk factors for age-related hearing impairment and moderate alcohol consumption is protective: A european population-based multicentre study" by Fransen et al., J. Assoc. Res. Otolaryngol. DOI 10.1007/s10162-008-0123-1. *Jaro-Journal of the Association for Research in Otolaryngology, 9*(3), 261-263. doi: 10.1007/s10162-008-0128-9

Liberman, A. M. (1995). *Speech: A special code*. Cambridge, Massachusetts: MIT.

Lin, F. R., Ferrucci, L., Metter, E. J., An, Y., Zonderman, A. B., & Resnick, S. M. (2011). Hearing Loss and Cognition in the Baltimore Longitudinal Study of Aging. *Neuropsychology, 25*(6), 763-770. doi: 10.1037/a0024238

Lin, F. R., Metter, E. J., O'Brien, R. J., Resnick, S. M., Zonderman, A. B., & Ferrucci, L. (2011). Hearing Loss and Incident Dementia. *Archives of Neurology, 68*(2), 214-220. doi: 10.1001/archneurol.2010.362

Lin, F. R., Yaffe, K., Xia, J., Xue, Q.-L., Harris, T. B., Purchase-Helzner, E., . . . Hlth, A. B. C. S. G. (2013). Hearing Loss and Cognitive Decline in Older Adults. *Jama Internal Medicine, 173*(4), 293-299. doi: 10.1001/jamainternmed.2013.1868

Lindenberger, U., & Baltes, P. B. (1994). SENSORY FUNCTIONING AND INTELLIGENCE IN OLD-AGE - A STRONG CONNECTION. *Psychology and Aging, 9*(3), 339-355. doi: 10.1037/0882-7974.9.3.339

Lindenberger, U., Scherer, H., & Baltes, P. B. (2001). The strong connection between sensory and cognitive performance in old age: Not due to sensory acuity reductions operating during cognitive assessment. *Psychology and Aging, 16*(2), 196-205. doi: 10.1037//0882-7974.16.2.196

Liston, R., Solomon, S., & Banerjee, A. K. (1995). PREVALENCE OF HEARING PROBLEMS, AND USE OF HEARING-AIDS AMONG A SAMPLE OF ELDERLY PATIENTS. *British Journal of General Practice, 45*(396), 369-370.

Livingston, G., Sommerlad, A., Orgeta, V., Costafreda, S. G., Huntley, J., Ames, D., . . . Mukadam, N. (2017). Dementia prevention, intervention, and care. *The Lancet*. doi: 10.1016/S0140-6736(17)31363-6

Loebach, J. L., Pisoni, D. B., & Svirsky, M. A. (2009). Transfer of auditory perceptual learning with spectrally reduced speech to speech and nonspeech tasks: Implications for cochlear implants. *Ear and Hearing, 30*(6), 662-674. doi: 10.1097/AUD.0b013e3181b9c92d

Loebach, J. L., Pisoni, D. B., & Svirsky, M. A. (2010). Effects of Semantic Context and Feedback on Perceptual Learning of Speech Processed Through an Acoustic Simulation of a Cochlear Implant. *Journal of Experimental Psychology-Human Perception and Performance, 36*(1), 224-234. doi: 10.1037/a0017609

Logan, G. D. (2004) Cumulative progress in formal theories of attention. *Vol. 55. Annual Review of Psychology* (pp. 207-234).

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing, 19*(1), 1-36. doi: 10.1097/00003446-199802000-00001

Lunner, T. (2003). Cognitive function in relation to hearing aid use. *International journal of audiology, 42*, S49-S58.

Macleod, C. M. (1991). HALF A CENTURY OF RESEARCH ON THE STROOP EFFECT - AN INTEGRATIVE REVIEW. *Psychological Bulletin, 109*(2), 163-203. doi: 10.1037//0033-2909.109.2.163

Macleod, C. M. (2005). The Stroop Task in Cognitive Research. In A. Wenzel & R. D.E (Eds.), *Cognitive methods and their application to clinical research*. Washington, DC: APA books.

Martin, M. (2007). Software-based auditory training program found to reduce hearing aid return rate. *Hearing Journal, 60*(8), 32-35.

Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes, 27*(7-8), 953-978. doi: 10.1080/01690965.2012.705006

McCabe, D., Langer, K. G., Borod, J. C., & Bender, H. A. (2011). Practice Effects. In J. S. Kreutzer, J. DeLuca & B. Caplan (Eds.), *Encyclopedia of Clinical Neuropsychology* (pp. 1988-1989). New York, NY: Springer New York.

McCormack, A., & Fortnum, H. (2013). Why do people fitted with hearing aids not wear them? *International journal of audiology, 52*(5), 360-368. doi: 10.3109/14992027.2013.769066

McCoy, S. L., Tun, P. A., Cox, L. C., Colangelo, M. L., Stewart, R. A., & Wingfield, A. (2005). Hearing loss and perceptual effort: Downstream effects on older adults' memory for speech. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 58*(1), 22-33. doi: 10.1080/02724980443000151

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology, 49*(2), 270-291. doi: 10.1037/a0028228

Melby-Lervåg, M., & Hulme, C. (2016). There is no convincing evidence that working memory training is effective: A reply to Au et al. (2014) and Karbach and Verhaeghen (2014). *Psychonomic Bulletin and Review, 23*(1), 324-330. doi: 10.3758/s13423-015-0862-z

Miller, J. D., Watson, C. S., Kistler, D. J., Wightman, F. L., & Preminger, J. E. (2008). Preliminary evaluation of the speech perception assessment and training system (SPATS) with hearing-aid and cochlear-implant users. *Proc Meet Acoust, 2*(1), 1-9. doi: 10.1121/1.2988004

Millward, K. E., Hall, R. L., Ferguson, M. A., & Moore, D. R. (2011). Training speech-in-noise perception in mainstream school children. *International Journal of Pediatric Otorhinolaryngology, 75*(11). doi: 10.1016/j.ijporl.2011.08.003

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49-100. doi: 10.1006/cogp.1999.0734

Mizutari, K., Michikawa, T., Saito, H., Okamoto, Y., Enomoto, C., Takebayashi, T., . . . Nishiwaki, Y. (2013). Age-Related Hearing Loss and the Factors Determining Continued Usage of Hearing Aids among Elderly Community-Dwelling Residents. *Plos One, 8*(9). doi: 10.1371/journal.pone.0073622

Molloy, K., Moore, D. R., Sohoglu, E., & Amitay, S. (2012). Less Is More: Latent Learning Is Maximized by Shorter Training Sessions in Auditory Perceptual Learning. *Plos One, 7*(5). doi: 10.1371/journal.pone.0036929

Moore, D. R., & Amitay, S. (2007). Auditory training: Rules and applications. *Seminars in Hearing, 28*(2), 99-109. doi: 10.1055/s-2007-973436

Moore, D. R., Edmondson-Jones, M., Dawes, P., Fortnum, H., McCormack, A., Pierzycki, R. H., & Munro, K. J. (2014). Relation between Speech-in-Noise Threshold, Hearing Loss and Cognition from 40-69 Years of Age. *Plos One, 9*(9). doi: 10.1371/journal.pone.0107720

Morrison, A., & Chein, J. (2011). Does working memory training work? The promise and challenges of enhancing cognition by training working memory. *Psychonomic Bulletin & Review, 18*(1), 46.

Mulrow, C. D., Aguilar, C., Endicott, J. E., Tuley, M. R., Velez, R., Charlip, W. S., . . . Denino, L. A. (1990). QUALITY-OF-LIFE CHANGES AND HEARING IMPAIRMENT - A RANDOMIZED TRIAL. *Annals of Internal Medicine, 113*(3), 188-194.

Munro, K. J., & Lutman, M. E. (2003). The effect of speech presentation level on measurement of auditory acclimatization to amplified speech. *Journal of the Acoustical Society of America, 114*(1), 484-495. doi: 10.1121/1.1577556

Murphy, D. R., Craik, F. I. M., Li, K. Z. H., & Schneider, B. A. (2000). Comparing the effects of aging and background noise on short-term memory performance. *Psychology and Aging, 15*(2), 323-334. doi: 10.1037/0882-7974.15.2.323

Musiek, F. E. (2004). Auditory processing disorders. *Journal of the American Academy of Audiology, 15*(6), 462-463. doi: 10.3766/jaaa.15.6.7

Ng, E. H. N., Classon, E., Larsby, B., Arlinger, S., Lunner, T., Rudner, M., & Rönnberg, J. (2014). Dynamic Relation Between Working Memory Capacity and Speech Recognition in Noise During the First 6 Months of Hearing Aid Use. *Trends in Hearing, 18*, 2331216514558688. doi: 10.1177/2331216514558688

Noble, W., Jensen, N. S., Naylor, G., Bhullar, N., & Akeroyd, M. A. (2013). A short form of the Speech, Spatial and Qualities of Hearing scale suitable for clinical use: The SSQ12. *International journal of audiology, 52*(6), 409-412. doi: 10.3109/14992027.2013.781278

Oba, S. I., Fu, Q.-J., & Galvin, J. J., III. (2011). Digit Training in Noise Can Improve Cochlear Implant Users' Speech Understanding in Noise. *Ear and Hearing, 32*(5). doi: 10.1097/AUD.0b013e31820fc821

Office of National Statisitics. (2015). National Population Projections: 2014-based

Statistical Bulletin. *Website.* Retrieved 26th July, 2016, from https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationprojections/bulletins/nationalpopulationprojections/2015-10-29

Olson, A. D., Preminger, J. E., & Shinn, J. B. (2013). The Effect of LACE DVD Training in New and Experienced Hearing Aid Users. *Journal of the American Academy of Audiology, 24*(3), 214-230. doi: 10.3766/jaaa.24.3.7

Otometrics. Retrieved 08.04.2018

Parbery-Clark, A., Anderson, S., Hittner, E., & Kraus, N. (2012). Musical experience offsets age-related delays in neural timing. *Neurobiology of Aging, 33*(7), 1483.e1481-1483.e1484. doi: 10.1016/j.neurobiolaging.2011.12.015

Parbery-Clark, A., Skoe, E., & Kraus, N. (2009). Musical Experience Limits the Degradative Effects of Background Noise on the Neural Processing of Sound. *Journal of Neuroscience, 29*(45), 14100-14107. doi: 10.1523/jneurosci.3256-09.2009

Patel, M. (2011). *An extended BKB sentence test appropriate for a wide range of speech perceptin abilities.*

Pichora-Fuller, M. K. (2003a). Cognitive aging and auditory information processing. *International journal of audiology, 42*(Supplement 2), 2S26-22S32.

Pichora-Fuller, M. K. (2003b). Cognitive aging and auditory information processing. *International journal of audiology, 42*, S26-S32.

Pichora-Fuller, M. K. (2008). Use of supportive context by younger and older adult listeners: Balancing bottom-up and top-down information processing. *International journal of audiology, 47*, S72-S82. doi: 10.1080/14992020802307404

Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M. (1995). HOW YOUNG AND OLD ADULTS LISTEN TO AND REMEMBER SPEECH IN NOISE. *Journal of the Acoustical Society of America, 97*(1), 593-608. doi: 10.1121/1.412282

Pisoni, D. B., Goldinger, S. D., & Luce, P. A. (1988). ACOUSTIC-PHONETIC PRIMING IN SPOKEN WORD RECOGNITION - A TEST OF THE NEIGHBORHOOD ACTIVATION MODEL. *Bulletin of the Psychonomic Society, 26*(6), 505-506.

Pronk, M., Deeg, D. J. H., Festen, J. M., Twisk, J. W., Smits, C., Comijs, H. C., & Kramer, S. E. (2013). Decline in Older Persons' Ability to Recognize Speech in Noise: The Influence of Demographic, Health-Related, Environmental, and Cognitive Factors. *Ear and Hearing, 34*(6), 722-732.

Pronk, M., Deeg, D. J. H., & Kramer, S. E. (2013). Hearing Status in Older Persons: A Significant Determinant of Depression and Loneliness? Results From the Longitudinal Aging Study Amsterdam. *American journal of audiology, 22*(2), 316-320. doi: 10.1044/1059-0889(2013/12-0069)

Pryce, H., & Gooberman-Hill, R. (2012). 'There's a hell of a noise': living with a hearing loss in residential care. *Age and Ageing, 41*(1), 40-46. doi: 10.1093/ageing/afr112

Pryce, H., Hall, A., Laplante-Levesque, A., & Clark, E. (2016). A qualitative investigation of decision making during help-seeking for adult hearing loss. *International journal of audiology, 55*(11), 658-665. doi: 10.1080/14992027.2016.1202455

Qian, Z. J., Wattamwar, K., Caruana, F. F., Otter, J., Leskowitz, M. J., Siedlecki, B., . . . Lalwani, A. K. (2016). Hearing Aid Use is Associated with Better Mini-Mental State Exam Performance. *Am J Geriatr Psychiatry, 24*(9), 694-702. doi: 10.1016/j.jagp.2016.03.005

Rabbitt, P. (1991). MILD HEARING-LOSS CAN CAUSE APPARENT MEMORY FAILURES WHICH INCREASE WITH AGE AND REDUCE WITH IQ. *Acta Oto-Laryngologica*, 167-176.

Ralph, P., Kopp, G., & Green, H. (1947). *Visible speech*: D. Van Norstrand.

Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., . . . Engle, R. W. (2013). No Evidence of Intelligence Improvement After Working Memory Training: A Randomized, Placebo-Controlled Study. *Journal of Experimental Psychology-General, 142*(2), 359-379. doi: 10.1037/a0029082

Remez, R. E., Fellowes, J. M., & Rubin, P. E. (1997). Talker identification based on phonetic information. *Journal of Experimental Psychology-Human Perception and Performance, 23*(3), 651-666. doi: 10.1037/0096-1523.23.3.651

Reuter-Lorenz, P. A., & Cappell, K. A. (2008). Neurocognitive Aging and the Compensation Hypothesis. *Current Directions in Psychological Science, 17*(3), 177-182. doi: 10.1111/j.1467-8721.2008.00570.x

Ridderinkhof, K. R., van den Wildenberg, W. P. M., Wijnen, J., & Burle, B. (2004). *Cognitive Neuroscience of Attention* (M. Posner Ed.). New York: Guildford Press.

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., . . . Rudner, M. (2013). The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Frontiers in Systems Neuroscience, 7*. doi: 10.3389/fnsys.2013.00031

Rönnberg, J., Rudner, M., Foo, C., & Lunner, T. (2008). Cognition counts: A working memory system for ease of language understanding (ELU). *International journal of audiology, 47*, S99-S105. doi: 10.1080/14992020802301167

Rubinstein, A., & Boothroyd, A. (1987). EFFECT OF 2 APPROACHES TO AUDITORY TRAINING ON SPEECH RECOGNITION BY HEARING-IMPAIRED ADULTS. *Journal of Speech and Hearing Research, 30*(2), 153-160.

Rudner, M., & Rönnberg, J. (2008). The role of the episodic buffer in working memory for language processing. *Cognitive Processing, 9*(1), 19-28. doi: 10.1007/s10339-007-0183-x

Rudner, M., Ronnberg, J., & Lunner, T. (2011). Working Memory Supports Listening in Noise for Persons with Hearing Impairment. *Journal of the American Academy of Audiology, 22*(3), 156-167. doi: 10.3766/jaaa.22.3.4

Russo, N. M., Nicol, T. G., Zecker, S. G., Hayes, E. A., & Kraus, N. (2005). Auditory training improves neural timing in the human brainstem. *Behav Brain Res, 156*(1), 95-103. doi: 10.1016/j.bbr.2004.05.012

Sabes, J. H., & Sweetow, R. W. (2007). Variables predicting outcomes on listening and communication enhancement (LACE (TM)) training. *International journal of audiology, 46*(7), 374-383. doi: 10.1080/14992020701297565

Sabri, M., Melara, R. D., & Algom, D. (2001). A confluence of contexts: Asymmetric versus global failures of selective attention to stroop dimensions. *Journal of Experimental Psychology-Human Perception and Performance, 27*(3), 515-537. doi: 10.1037//0096-1523.27.3.515

Saito, H., Nishiwaki, Y., Michikawa, T., Kikuchi, Y., Mizutari, K., Takebayashi, T., & Ogawa, K. (2010). Hearing Handicap Predicts the Development of Depressive Symptoms After 3 Years in Older Community-Dwelling Japanese. *Journal of the American Geriatrics Society, 58*(1), 93-97. doi: 10.1111/j.1532-5415.2009.02615.x

Salthouse, T. A. (1994). AGING ASSOCIATIONS - INFLUENCE OF SPEED ON ADULT AGE-DIFFERENCES IN ASSOCIATIVE LEARNING. *Journal of Experimental Psychology-Learning Memory and Cognition, 20*(6), 1486-1503. doi: 10.1037//0278-7393.20.6.1486

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103*(3), 403-428. doi: 10.1037/0033-295x.103.3.403

Salthouse, T. A. (Ed.). (1985). *A Theory of Cognitive Aging* (Vol. 28): Elsevier.

Sanders, D. A. (1971). *Auditory Rehabilitation*. Englewood Cliffs. NJ: Prentice-Hall.

Saunders, G. H., Smith, S. L., Chisolm, T. H., Frederick, M. T., McArdle, R. A., & Wilson, R. H. (2016a). A Randomized Control Trial: Supplementing Hearing Aid Use with Listening and Communication Enhancement (LACE) Auditory Training. *Ear Hear*. doi: 10.1097/AUD.0000000000000283

Saunders, G. H., Smith, S. L., Chisolm, T. H., Frederick, M. T., McArdle, R. A., & Wilson, R. H. (2016b). A Randomized Control Trial: Supplementing Hearing Aid Use with Listening and Communication Enhancement (LACE) Auditory Training. *Ear and Hearing, 37*(4), 381-396. doi: 10.1097/aud.0000000000000283

Schmeidt, R. A. (2010). The ageing auditory system: Perceptual characterization and neural basis of presbycusis,. In S. Gordon-Salant, Frisina, R.D., Popper, A. and Fay, D. (Ed.), *Handbook of Auditory Research.* (pp. 9-38). Berlin: Springer.

Schneiders, J. A., Opitz, B., Tang, H., Deng, Y., Xie, C., Li, H., & Mecklinger, A. (2012). The impact of auditory working memory training on the fronto-parietal working memory network. *Frontiers in Human Neuroscience, 6*. doi: 10.3389/fnhum.2012.00173

Schoof, T., & Rosen, S. (2014). The role of auditory and cognitive factors in understanding speech in noise by normal-hearing older listeners. *Frontiers in Aging Neuroscience, 6*. doi: 10.3389/fnagi.2014.00307

Schuknecht, H. F., & Gacek, M. R. (1993). COCHLEAR PATHOLOGY IN PRESBYCUSIS. *Annals of Otology Rhinology and Laryngology, 102*(1), 1-16.

Sekuler, R., & Blake, R. (1987). SENSORY UNDERLOAD. *Psychology Today, 21*(12), 48-51.

Sheldon, S., Pichora-Fuller, M. K., & Schneider, B. A. (2008). Priming and sentence context support listening to noise-vocoded speech by younger and older adults. *Journal of the Acoustical Society of America, 123*(1), 489-499. doi: 10.1121/1.2783762

Shen, J., Anderson, M. C., Arehart, K. H., & Souza, P. E. (2016). Using Cognitive Screening Tests in Audiology. *American journal of audiology, 25*(4), 319-331. doi: 10.1044/2016_aja-16-0032

Sheppard, L. D., & Vernon, P. A. (2008). Intelligence and speed of information-processing: A review of 50 years of research. *Personality and Individual Differences, 44*(3), 535-551. doi: https://doi.org/10.1016/j.paid.2007.09.015

Siemens. eARena. Retrieved 31st October, 2016, from https://www.bestsoundtechnology.com/pro/knowledge-base/earena/

Skoe, E., & Kraus, N. (2010). Auditory brainstem response to complex sounds: a tutorial. *Ear and Hearing, 31*(3), 302-324. doi: 10.1097/AUD.0b013e3181cdb272

Skoe, E., & Kraus, N. (2013). The Brainstem Toolbox. Version 2013. from www.brainvolts.northwestern.edu

Smeeth, L., Fletcher, A. E., Ng, E. S. W., Stirling, S., Nunes, M., Breeze, E., . . . Tulloch, A. (2002). Reduced hearing, ownership, and use of hearing aids in elderly people in the UK - the MRC Trial of the Assessment and Management of Older People in the Community: a cross-sectional survey. *Lancet, 359*(9316), 1466-1470. doi: 10.1016/s0140-6736(02)08433-7

Smith, E. E., & Jonides, J. (1999). Storage and Executive Processes in the Frontal Lobes. *Science, 283*(5408), 1657-1661. doi: 10.1126/science.283.5408.1657

Smith, G. E. (2016). Healthy Cognitive Aging and Dementia Prevention. *American Psychologist, 71*(4), 268-275. doi: 10.1037/a0040250

Smith, G. E., Housen, P., Yaffe, K., Ruff, R., Kennison, R. F., Mahncke, H. W., & Zelinski, E. M. (2009). A Cognitive Training Program Based on Principles of Brain Plasticity: Results from the Improvement in Memory with Plasticity-based Adaptive Cognitive Training (IMPACT) Study. *Journal of the American Geriatrics Society, 57*(4), 594-603. doi: 10.1111/j.1532-5415.2008.02167.x

Smith, S. L., & Pichora-Fuller, M. K. (2015). Associations between speech understanding and auditory and visual tests of verbal working memory: effects of linguistic complexity, task, age, and hearing loss. *Frontiers in psychology, 6*. doi: 10.3389/fpsyg.2015.01394

Sohmer, H., & Feinmesser, M. (1967). Cochlear action potentials recorded from the external ear in man. *The Annals of otology, rhinology, and laryngology, 76*(2), 427-435.

Sohoglu, E., Peelle, J. E., Carlyon, R. P., & Davis, M. H. (2012). Predictive Top-Down Integration of Prior Knowledge during Speech Perception. *Journal of Neuroscience, 32*(25). doi: 10.1523/jneurosci.5069-11.2012

Sommers, M. S. (1996). The structural organization of the mental lexicon and its contribution to age-related declines in spoken-word recognition. *Psychology and Aging, 11*(2), 333-341. doi: 10.1037/0882-7974.11.2.333

Sommers, M. S., & Danielson, S. M. (1999). Inhibitory processes and spoken word recognition in young and older adults: The interaction of lexical competition and semantic context. *Psychology and Aging, 14*(3), 458-472. doi: 10.1037//0882-7974.14.3.458

Song, J. H., Skoe, E., Banai, K., & Kraus, N. (2012). Training to Improve Hearing Speech in Noise: Biological Mechanisms. *Cerebral Cortex, 22*(5). doi: 10.1093/cercor/bhr196

Souza, P., & Arehart, K. (2015). Robust relationship between reading span and speech recognition in noise. *International journal of audiology, 54*(10), 705-713. doi: 10.3109/14992027.2015.1043062

Stacey, P. C., Raine, C. H., O'Donoghue, G. M., Tapper, L., Twomey, T., & Summerfield, A. Q. (2010). Effectiveness of computer-based auditory training for adult users of cochlear implants. *International journal of audiology, 49*(5), 347-356. doi: 10.3109/14992020903397838

Stacey, P. C., & Summerfield, A. Q. (2008). Comparison of word-, sentence-, and phoneme-based training strategies in improving the perception of spectrally distorted speech. *Journal of Speech Language and Hearing Research, 51*(2), 526-538. doi: 10.1044/1092-4388(2008/038)

Stough, C., Nettelbeck, T., & Cooper, C. (1993). THE RELATIONSHIP BETWEEN SPEED OF INFORMATION-PROCESSING MEASURES AND PSYCHOMETRIC INTELLIGENCE. *Behavior Genetics, 23*(6), 567-567.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643-662.

Stuss, D. T., & Knight, R. T. (2002). *Pricniples of Frontal Lobe Function* (1 ed.). Oxford, New York: Oxford University Press.

Sullivan, J. R., Thibodeau, L. M., & Assmann, P. F. (2013). Auditory training of speech recognition with interrupted and continuous noise maskers by children with hearing impairment. *The Journal of the Acoustical Society of America, 133*(1), 495-501. doi: 10.1121/1.4770247

Suresh, K. P. (2011). An overview of randomization techniques: An unbiased assessment of outcome in clinical research. *Journal of Human Reproductive Sciences, 4*(1), 8-11. doi: 10.4103/0974-1208.82352

Surprenant, A. M., & Watson, C. S. (2001). Individual differences in the processing of speech and nonspeech sounds by normal-hearing listeners. *Journal of the Acoustical Society of America, 110*(4), 2085-2095. doi: 10.1121/1.1404973

Sweetow, R. W., & Sabes, J. (2006a). The need for and development of an adaptive Listening and Communication Enhancement (LACE (TM)) program. *Journal of the American Academy of Audiology, 17*(8), 538-558. doi: 10.3766/jaaa.17.8.2

Sweetow, R. W., & Sabes, J. H. (2006b). The need for and development of an adaptive Listening and Communication Enhancement (LACE (TM)) program. *Journal of the American Academy of Audiology, 17*(8), 538-558. doi: 10.3766/jaaa.17.8.2

Sweetow, R. W., & Sabes, J. H. (2007). Listening and communication enhancement (LACE). *Seminars in Hearing, 28*(2), 133-141. doi: 10.1055/s-2007-973439

Taler, V., Aaron, G. P., Steinmetz, L. G., & Pisoni, D. B. (2010). Lexical Neighborhood Density Effects on Spoken Word Recognition and Production in Healthy Aging. *Journals of Gerontology Series B-Psychological Sciences and Social Sciences, 65*(5), 551-560. doi: 10.1093/geronb/gbq039

Tallal, P., Miller, S. L., Bedi, G., Byma, G., Wang, X., Nagarajan, S. S., . . . Merzenich, M. M. (1996). Language Comprehension in Language-Learning Impaired Children Improved with Acoustically Modified Speech. *Science, 271*(5245), 81-84. doi: 10.1126/science.271.5245.81

Tesch-Römer, C. (1997). Psychological Effects of Hearing Aid use in Older Adults. *The Journals of Gerontology: Series B, 52B*(3), P127-P138. doi: 10.1093/geronb/52B.3.P127

The Mathworks, I. (2006b). MATLAB. Natick, MA: The MathWorks, Inc.

Towse, J. N., & Neil, D. (1998). Analyzing human random generation behavior: A review of methods used and a computer program for describing performance. *Behavior Research Methods Instruments & Computers, 30*(4), 583-591. doi: 10.3758/bf03209475

Troyer, A. K., Graves, R. E., & Cullum, C. M. (1994). Executive functioning as a mediator of the relationship between age and episodic memory in healthy aging. *Aging, Neuropsychology, and Cognition, 1*(1), 45-53. doi: 10.1080/09289919408251449

Turner, C. W., Fabry, D. A., Barrett, S., & Horwitz, A. R. (1992). DETECTION AND RECOGNITION OF STOP CONSONANTS BY NORMAL-HEARING AND HEARING-IMPAIRED LISTENERS. *Journal of Speech and Hearing Research, 35*(4), 942-949.

Uttl, B., & Graf, P. (1997). Color-word stroop test performance across the adult life span. *Journal of Clinical and Experimental Neuropsychology, 19*(3), 405-420. doi: 10.1080/01688639708403869

Valderrama, J. T., de la Torre, A., Alvarez, I., Segura, J. C., Thornton, A. R. D., Sainz, M., & Vargas, J. L. (2014). A study of adaptation mechanisms based on ABR recorded at high stimulation rate. *Clinical Neurophysiology, 125*(4), 805-813. doi: 10.1016/j.clinph.2013.06.190

van Hooren, S. A. H., Anteunis, L. J. C., Valentijn, S. A. M., Bosma, H., Ponds, R., Jolles, J., & van Boxtel, M. P. J. (2005). Does cognitive function in older adults with hearing impairment improve by hearing aid use? *International journal of audiology, 44*(5), 265-271. doi: 10.1080/14992020500060370

van Rooij, J. C., & Plomp, R. (1990). Auditive and cognitive factors in speech perception by elderly listeners. *Acta oto-laryngologica. Supplementum, 476*, 177-181.

van Rooij, J. C., & Plomp, R. (1992a). Auditive and cognitive factors in speech perception by elderly listeners. III. Additional data and final discussion. *J Acoust Soc Am, 91*(2), 1028-1033.

van Rooij, J. C. G. M., & Plomp, R. (1989). Auditive and cognitive factors in speech perception by elderly listeners. I: Development of test battery. *Journal of the Acoustical Society of America, 86*(4), 1294-1309. doi: 10.1121/1.398744

Van Rooij, J. C. G. M., & Plomp, R. (1992b). AUDITIVE AND COGNITIVE FACTORS IN SPEECH PERCEPTION BY ELDERLY LISTENERS III. ADDITIONAL DATA AND FINAL DISCUSSION. *Journal of the Acoustical Society of America, 91*(2), 1028-1033.

Vaughan, N., Storzbach, D., & Furukawa, I. (2006). Sequencing versus nonsequencing working memory in understanding of rapid speech by older listeners. *Journal of the American Academy of Audiology, 17*(7), 506-518. doi: 10.3766/jaaa.17.7.6

Veneman, C. E., Gordon-Salant, S., Matthews, L. J., & Dubno, J. R. (2013). Age and Measurement Time-of-Day Effects on Speech Recognition in Noise. *Ear and Hearing, 34*(3), 288-299. doi: 10.1097/AUD.0b013e31826d0b81

Verhaeghen, P., & De Meersman, L. (1998). Aging and the stroop effect: A meta-analysis. *Psychology and Aging, 13*(1), 120-126. doi: 10.1037/0882-7974.13.1.120

Vermiglio, A. J., Soli, S. D., Freed, D. J., & Fisher, L. M. (2012). The relationship between high-frequency pure-tone hearing loss, hearing in noise test (HINT) thresholds, and the articulation index. *J Am Acad Audiol, 23*(10), 779-788. doi: 10.3766/jaaa.23.10.4

Versfeld, N. J., & Dreschler, W. A. (2002). The relationship between the intelligibility of time-compressed speech and speech in noise in young and elderly listeners. *The Journal of the Acoustical Society of America, 111*(1), 401-408. doi: 10.1121/1.1426376

Wayne, R. V., Hamilton, C., Huyck, J. J., & Johnsrude, I. S. (2016). Working Memory Training and Speech in Noise Comprehension in Older Adults. *Frontiers in Aging Neuroscience, 8*. doi: 10.3389/fnagi.2016.00049

Wayne, R. V., & Johnsrude, I. S. (2015). A review of causal mechanisms underlying the link between age-related hearing loss and cognitive decline. *Ageing Research Reviews, 23*, 154-166. doi: 10.1016/j.arr.2015.06.002

Weihing, J., Chermak, G. D., & Musiek, F. E. (2015). Auditory Training for Central Auditory Processing Disorder. *Seminars in Hearing, 36*(4), 199-215. doi: 10.1055/s-0035-1564458

Werker, J. F., & Tees, R. C. (1992). THE ORGANIZATION AND REORGANIZATION OF HUMAN SPEECH-PERCEPTION. *Annual Review of Neuroscience, 15*, 377-402. doi: 10.1146/annurev.neuro.15.1.377

Wingfield, A., & Tun, P. A. (2007). Cognitive supports and cognitive constraints on comprehension of spoken language. *J Am Acad Audiol, 18*(7), 548-558.

Wingfield, A., Tun, P. A., & McCoy, S. L. (2005). Hearing loss in older adulthood - What it is and how it interacts with cognitive performance. *Current Directions in Psychological Science, 14*(3), 144-148. doi: 10.1111/j.0963-7214.2005.00356.x

Woods, D. L., Doss, Z., Herron, T. J., Arbogast, T., Younus, M., Ettlinger, M., & Yund, E. W. (2015). Speech Perception in Older Hearing Impaired Listeners: Benefits of Perceptual Training. *Plos One, 10*(3). doi: 10.1371/journal.pone.0113965

Wright, B. A., & Sabin, A. T. (2007). Perceptual learning: How much daily training is enough? *Experimental Brain Research, 180*(4), 727-736. doi: 10.1007/s00221-007-0898-z

Yang, C. H., Schrepfer, T., & Schacht, J. (2015). Age-related hearing impairment and the triad of acquired hearing loss. *Frontiers in Cellular Neuroscience, 9*. doi: 10.3389/fncel.2015.0026

Young, E. D., & Sachs, M. B. (1979). REPRESENTATION OF STEADY-STATE VOWELS IN THE TEMPORAL ASPECTS OF THE DISCHARGE PATTERNS OF POPULATIONS OF AUDITORY-NERVE FIBERS. *Journal of the Acoustical Society of America, 66*(5), 1381-1403. doi: 10.1121/1.383532

Yund, E. W., Roup, C. M., Simon, H. J., & Bowman, G. A. (2006). Acclimatization in wide dynamic range multichannel compression and linear amplification hearing aids*,* 517.

Zekveld, A. A., Festen, J. M., & Kramer, S. E. (2013). Task Difficulty Differentially Affects Two Measures of Processing Load: The Pupil Response During Sentence Processing and Delayed

Cued Recall of the Sentences. *Journal of Speech Language and Hearing Research, 56*(4), 1156-1165. doi: 10.1044/1092-4388(2012/12-0058)

Zekveld, A. A., George, E. L. J., Kramer, S. E., Goverts, S. T., & Houtgast, T. (2007). The development of the text reception threshold test: A visual analogue of the speech reception threshold test. *Journal of Speech Language and Hearing Research, 50*(3), 576-584. doi: 10.1044/1092-4388(2007/040)

Zekveld, A. A., Rudner, M., Johnsrude, I. S., Festen, J. M., van Beek, J. H. M., & Roennberg, J. (2011). The Influence of Semantically Related and Unrelated Text Cues on the Intelligibility of Sentences in Noise. *Ear and Hearing, 32*(6), E16-E25. doi: 10.1097/AUD.0b013e318228036a

APPENDIX 1

Appendix 1

The following sections in Appendix 1 report on the potential role the accurate encoding of speech sounds into neural signals may play in predicting speech-in-noise performance. The general hypothesis being that the larger and more faithful the representations of speech sounds are, the better the speech-in-noise performance will be. Although not forming a chapter within the main thesis, the study is still worth reporting here as it gives evidence that for the age range of the participants recruited for this current study, larger amplitude neural responses (evoked by a speech sound) are not related to better speech-in-noise performance.

Introduction to neural representations of speech sounds and the potential link to hearing speech-in-noise

In Chapter 2 of this thesis, there was a focus on the roles of different cognitive factors in the discrimination of speech-in-noise. On hierarchical regression, response latency, working memory and control of inhibition all demonstrated some degree of significance when predicting speech-in-noise performance. Overall, however, variance explanation, remained low at 22%. The representation of speech sounds in the auditory neural pathway was introduced earlier in this thesis in Section 1.4.1. What follows in this Appendix 1 is background information on some of the literature surrounding neural representations of speech sounds and their potential relationship with speech in noise discrimination ability.

Previous researchers (Anderson, Parbery-Clark, White-Schwoch, & Kraus, 2013a, 2013b) have hypothesised that the better the representation of a speech sound in the auditory neural pathway, the better speech perception in noise will be, as richer neural encoding may result in superior discrimination.

Frequencies found in speech sounds

Within speech sounds there are certain frequencies that have larger amplitudes of activity relative to other frequencies. The different frequencies present in an utterance can be visually presented on a speech spectrograph, as originally described by Ralph, Kopp, and Green (1947). The large amplitude frequencies present in a speech sound were described as ‘Formants’ by Fant (1960). The amount of energy in each of the formants will rise and fall in a cyclic pattern as the vocal chords open and close. Formants from monophthong vowels (where there is only one sound to the vowel e.g. ‘i’ and nasals (a sound released through the nose as in n and m) do not generally change and have spectral continuity. For this reason they can also be more readily recorded in the human brainstem as they generate following responses (sustained neural activity).

It tends to be the frequencies of the first two or three formants (F1, F2 and F3) in a vowel that distinguish one vowel from another vowel (Carlson et al., 1974). All voiced phonemes will have formants but monophthong vowels are the easiest to identify on spectrograms. It should also be remembered that there are a number of other speech sounds, such as the plosives (p as in pit) and fricatives (f as in fan), that do not have formants.

When considering the formants in vowels, it is possible to display how different frequencies make up the different vowels on the speech spectrogram. For example, when comparing vowels spoken by the same speaker, the /æ/ vowel, as in the word ‘bat’, has frequencies in the region of F1 at 690Hz and F2 at 1,660Hz. The vowel /ɒ/, as in ‘box’, has an F1 of 710Hz and F2 of 1100Hz. Depending on the speaker, these formant frequencies have a range of potential different values. This is because the formants arise from resonances in the vocal tract and, therefore, the size of the tract impacts variations. Talkers are able to change the resonant frequencies for different sounds by moving the ‘articulators’ in the speech process. The articulators, such as the lips, jaws, tongue, and soft palate, can change the dimensions of the resonance cavities in the vocal tract.

The neural activity generated by speech sounds can be recorded in the auditory brainstem. This principle is based on the early seminal work of Sohmer and Feinmesser (1967) who described the recording of auditory evoked potentials using scalp electrodes. The knowledge that it is the first few formants of a speech sound that are key to understanding, allows a researcher to examine their representation in brainstem recordings and can provide a useful insight into how well a stimulus is being transferred to a listener. One hypothesis is that the better the formants are represented, the better the listener will be at performing listening tasks.

Therefore, there has been growing interest in the ways that the brainstem can encode complex speech stimuli. The recordings may help to explain participants' differences in performance of listening tasks and how representation may change through ageing (Anderson, Parbery-Clark, Han-Gyol, & Kraus, 2011; Anderson, Parbery-Clark, White-Schwoch, & Kraus, 2012; Anderson et al., 2013b). Having knowledge, therefore, of the frequencies present in the speech stimulus (used to evoke the response) enables a researcher to compare the recorded responses in the brainstem (Skoe & Kraus, 2010). This allows for exploration of how faithfully a stimulus is being represented by neural firing.

Speech sounds represented in the auditory neural pathway

Previous experimental work on cats has shown that formants within speech can be recorded in the neural discharge patterns of the auditory nerve (Young & Sachs, 1979). This research, when examined alongside work by Greenberg (1980), carrying out neural recordings with humans, demonstrates the relationship between temporal encoding of complex auditory stimuli (speech sounds) and their representation in the auditory brainstem. This early work demonstrated the faithful representation of vowel formants within speech in the neural discharge patterns in the auditory brainstem.

Expanding further on the research described above, Galbraith, Arbagey, Branski, Comerci, and Rector (1995), in research that was introduced in Section 1.4.1, demonstrated the hi-fidelity representation that can be achieved by neural firings in the human auditory brainstem. Galbraith et al. utilised complete words (ten high and ten low probability) and recorded the evoked responses of these 20 words on two human participants. To ensure good quality recordings, the responses were evoked by 1000 repetitions of each word. The frequencies of the generated electrical activity encoded in the two participants' neural responses of the 20 words, in their respective brainstems were then reconstituted as an acoustic sound. The newly reconstituted sound representations (of the original 20 words) were then played to 80 normally hearing listeners who were asked to attempt to articulate what, if any, words they heard. This was done both with and without contextual clues. The result for the most recognisable word was that it could be identified by 92% of the listeners demonstrating the hi-fidelity preservation of the original speech in the auditory nerve. The least recognisable word could only be identified 5% of the time which could be down to simple chance.

There is also evidence that neural encoding in the peripheral auditory system demonstrates a relationship with later cortical representation (Galbraith et al., 2004). Although not using a speech stimulus, the Galbraith et al. study used a pitch change between two tones. The pitch change of the two tones (as is often found in speech) evoked a cortical response and the two tones evoked a frequency following response (FFR) recorded in the brain stem. The analysis confirmed that the spectral intensity differences of the FFR of the two tones was significantly correlated with the recorded amplitude recorded in the auditory cortex. Whilst not the focus of the research in this chapter, the central auditory cortex that deciphers incoming speech sounds is the final destination for the peripherally generated neural signals. The correlation discussed above with the neural coding in the periphery, indicates that if speech signals are clearly represented initially, this will lead to a better final cortical representations. This should offer a listener the prospect of a richer representation of any incoming sounds, which can then be better matched to previously stored information.

Speech and other stimuli used in complex ABR’s

A published tutorial covering different stimuli, recording parameters and different methods of analysis (Skoe & Kraus, 2010) has led to growing interest in the exploration of differences in individuals’ neurally encoded representations of complex sounds. The potential of having an objective measure to explain performance variations in listening to and understanding speech is an exciting development for researchers.

Despite a range of different stimuli being available, it is generally the /da/ syllable that has proven to be the stimulus of choice for many researchers. There is a variety of reasons for choosing /da/ as a stimulus with the most basic reason being that it does tend to evoke reliable, replicable and clear complex auditory brainstem responses. These responses are reliable between and within participants (Russo, Nicol, Zecker, Hayes, & Kraus, 2005). The /da/

stimulus is also found in most European languages allowing research to be carried out (as a non-novel speech sound) in many countries. This, therefore, increases the knowledge database surrounding the implications of these complex recordings. It also has acoustic properties that make it a good choice for the investigation of neural encoding. These consist of the fact that /da/ evokes a transient response at its onset, elicited by the stop burst portion of /da/ and there is a following sustained evoked response generated by the formant transition. This later rich, harmonic periodic segment allows for analysis of the phase-locking abilities of the neural system to the stimulus. For example, it examines how accurately the timing of the speech stimulus that is presented to the auditory system can be represented by the neural firing rate to faithfully represent the original speech signal. It is therefore possible to measure the neural phase-locking to the fundamental frequency F0 and to examine encoding of the formants within the stimulus. For the above reasons, it was the choice of stimulus for the research in this chapter.

A further reason for choosing /da/ is that stop consonants are more difficult for hearing impaired listeners to perceive (Turner, Fabry, Barrett, & Horwitz, 1992). A stop consonant is characterised by having a low amplitude and fast aspect, and is low in amplitude compared to the following vowel as in the case of /da/. When being listened to in the presence of background noise, this makes stop consonants such as d and b more difficult to discriminate for a listener with full hearing ability and even more so for hearing impaired individuals. It was for this reason that /da/ was utilised for this study in an attempt to explain differences in performance, where the amplitude and accuracy of the neural encodings could be compared to the speech perception in noise test results.

For the research in this chapter, it was hypothesised that individuals with stronger neural representations would achieve better speech-in-noise scores as these stronger representations would allow for easier identification of speech sounds and words. In addition to superior speech-in-noise performance, it was further hypothesised that there may also be a relationship of neural representations with complex working memory (as given by an RSPAN task) which is known to play a potential role in hearing speech-in-noise (Schoof & Rosen, 2014).

The direction of any such relationship between working memory and neural representation however could not be predicted. For example, strong neural representation may necessitate less working memory capacity for individuals to understand speech so freeing cognitive resource to perform working memory tasks. Conversely, weak neural encoding of a speech sound could be evidence that the individual does not rely on the neural encoding to perceive speech-in-noise but instead relies on a generally superior working memory performance, which is a known cognitive function contributor to perceiving speech-in-noise. The implication being that any relationship between amplitude of neural encoding and working memory for these individuals would be in a negative direction.

Refractoriness and corticofugal influences on evoked recordings

Responses in the auditory brainstem are relatively unaffected by refractoriness. In that on repeated presentations, the evoked ABR is not overly susceptible to attenuation. Where there is refractoriness, this is more commonly seen when inter-stimulus timing is very short (Valderrama et al., 2014). For the short duration (40msec) /da/ stimulus delivered at 7.1 per second for the research in this chapter, this allowed for a reasonably long inter-stimulus gap such that refractoriness was not a particular issue.

As was introduced in Section 1.4.1 the ascending route up which the neural activity travels has many different nuclei and the pathway is far from unidirectional as there is a profuse returning efferent pathway. It has been established that the auditory neural pathway is in fact made up of a series of dynamic loops such that changes in higher activity in the auditory cortex can, therefore, influence activity in subcortical nuclei (Bajo & King, 2012). Indeed, it could be hypothesised that the modulating influence has an effect on the final recorded ABR. In which case the accuracy of encoding of the speech stimulus into the neural code may have as much to do with refined corticofugal control as initial encoding.

Complex ABR's (using /da/) and hearing speech-in-noise

Work by various authors has attempted to demonstrate a relationship between the encoding of the speech stimuli /da/ and speech-in-noise performance. Anderson et al. (2013a), described the relationship of the encoding of the 40msec /da/ and the self-reporting of speech difficulties using the hearing in noise questions of the Speech, Spatial and Qualities of Hearing Scale (SSQ) (Noble et al., 2013). In Anderson et al.'s study it was identified that for the 111 middle to older age participants (104 being accepted for analysis), the variability of the brainstem recordings contributed to the explanation of the variation of the self-reported speech-in-noise subtest of the SSQ when considered alongside an actual Quick Speech-in-noise (QuickSIN), age, and hearing level. The model of the four variables accounted for 30% of the variance of the speech-in-noise subtest of the SSQ12. When the cABR recordings were removed from the model, this fell to 19%. The breakdown of the contribution of the different components of the cABR are discussed below, alongside a more detailed analysis of the regression model.

The ABR measures used are based on several components of the morphology of an expected recorded complex ABR waveform. Figure 1 below[50], demonstrates the measurements of the different variables that were recorded in the Anderson et al. (2013a) study. The brainstem variables used to attempt to explain the variations in self-reported difficulty in hearing speech-in-noise were the latency of the onset response (Vlat), the onset of peak O (Olat), the slope of the VA complex (VAslope) and the overall response morphology (STRr) (in comparison to the original speech stimulus to the recorded cABR morphology). For any of the older adults who had some degree of hearing loss, a compensatory amplification of the response was utilised.

---

[50] Taken from the study reported in this chapter and is displayed for demonstration purposes

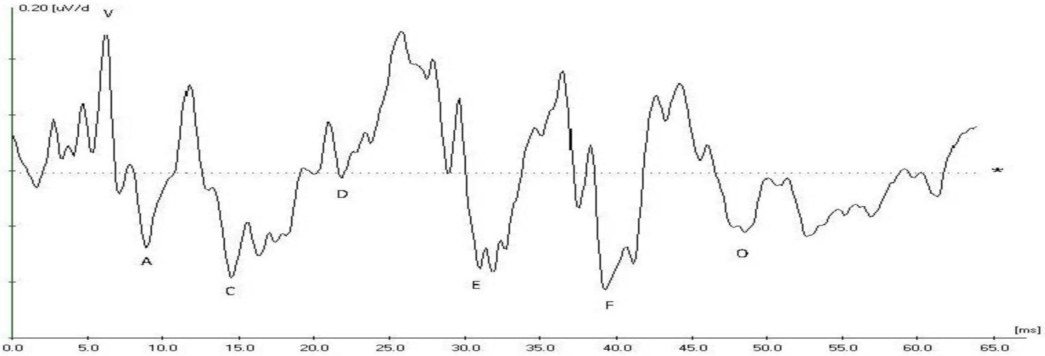


Figure 1. *ABR recorded to /da/ stimulus indicating the waveform labels used for results analysis in Anderson et al research 2013 (this figure was generated from a participant in the current research for this chapter*

In explaining the variation in self-reported difficulty, as evidenced by the subset of the SSQ dealing with speech-in-noise (as the dependent variable), a linear regression model was developed by Anderson et al. in the following order: the measures of speech-in-noise (SPIN) scores; PTA average, Age, and the complex ABR variables. It was found that the brainstem variables were the best predictor of the self-reported perception of speech-in-noise. Overall, the model was a reasonable fit for the data, R^2.306 when including all variables. However, on an individual basis it was only the Olat and the STRr that were significant predictors of self-reported difficulty.

On exploring the inter-correlations in this same study, it was reported that the QuickSIN test itself demonstrated a significant correlation with PTA, age, the self-reported SSQ and the Olat. The authors also discuss the fact that the only ABR variable giving a significant relationship with age was the Olat, which could have been anticipated due to neural encoding speeds demonstrating a deterioration with age (Parbery-Clark, Anderson, Hittner, & Kraus, 2012). Indeed, Parberry-Clark et al. reported that neural encoding delays can be demonstrated in people as young as 47 years old and, as the youngest participant in the Anderson study was 45 years old[51] they may well all have been giving evidence of some delays in latency timing. As the research in this present chapter used younger participants (<50 years), it was

---

[51] The age range of the study was 45-78 years.

anticipated that latency differences may have been minimal and, therefore, using the amplitude of the frequency following responses generated by the speech stimulus was anticipated to give a better variability between participants in recordings.

When considering complex ABR recordings it is possible to compare the amplitude of recorded responses at certain frequencies as defined by the speech stimulus. This is in contrast to exploring timing latency differences as explored in the Anderson et al. (2013a) study discussed above. These amplitude measures are the recorded neural activity in the brainstem at the fundamental (F0), and other formants (F1) and (F2) frequencies found in the evoking speech stimulus. By carefully controlling the stimulus, it is possible to analyse the representation of the frequencies recorded in the ABR compared with the stimulus and a correlation of the amplitude of the frequency responses with speech-in-noise performance (rather than peak waveform latencies). Anderson et al. (2011), utilised a synthesised /da/ of 170msecs and recorded complex ABR's on a group of normally hearing older adults (28 adults in total ranging from 60-73 years of age). The participants also performed an adaptive Hearing in Noise test (HINT). Anderson et al. reported a significant Pearson correlation of the amplitude of the recorded F0 and the HINT scores whether the ABR was recorded in quiet or in noise (both less then p

$<.05$). The result demonstrated that the larger the amplitude of the neural representation of the F0, the better the hearing in noise score.

In addition, Anderson et al. split the group along their HINT scores such that there were a group of participants defined as 'poor speech-in-noise discriminators' compared to 'better speech-in-noise discriminators'. On examination of the ABR traces, there was a recordable difference in the amplitude of the F0 of the brainstem between the two respective groups. The group that had the better speech discrimination had significantly larger F0 responses when the ABR's were recorded in quiet and had a tendency towards a significant difference when the ABR's were recorded in noise, although not quite reaching a significant level. It was further reported that the top HINT group had a greater root mean square (RMS), the overall amplitude of ABR response, which was significantly different from the low scoring HINT group, when the ABR response was recorded in both quiet and noise.

Parbery-Clark, Skoe, and Kraus (2009), reported on research with younger adults age 19-30 who were musicians and non-musicians. Using the longer 170msec /da/, they reported a significant relationship between the speech reception threshold, the level at which 50% of words can be correctly identified, of participants in quiet and the summed harmonics of an ABR to the /da/ stimulus also presented in quiet. The same was found to be the case when the speech reception was tested in noise and the ABR was recorded in noise. The conclusion was that the better the harmonic representation in the ABR, the better the speech perception. When reporting on the brainstem measures in quiet, however, the onset peak of the ABR, the transition peak, and the stimulus to response correlation did not evidence any relationship with either a Quick Speech-in-noise test (QuickSIN) or a hearing in noise test (HINT).

For this chapter, the recorded neural encodings of the evoked cABR's were analysed for any relationships with speech-in-noise performance and working memory span. Since the participants recruited were all younger, between the ages of 18-50, using latency measures was not anticipated to produce significant differences. Therefore, Fourier's analysis of the recorded complex ABR's and comparing the magnitude of the F0, the F1 and the RMS with speech-in-noise performance was used instead. Instead of using the longer /da/ stimulus, the shorter 40msec /da/ was used.

As all participants also performed sine and noise vocoded BKB speech discrimination tasks the research in this chapter was also able to explore if better neural encoding also allowed listeners to better discriminate novel types of degraded speech. The hypothesis being that even though the degraded speech was new to the listener, neural systems with a richer neural hi-fidelity representation of speech sounds would still allow for better discrimination. Theoretically, although there are not stored phonological representations (of the novel degraded speech) to match against as described in the ELU model of Rönnberg et al. (2013)

individuals better neural encoding may still have better access to make partial matches and therefore be able to discriminate better.

This current research additionally split participants on the grounds of their high and low speech discrimination performance in similar fashion to Anderson et al. 2013. This allowed for any significant differences between the groups cABR’s to be examined. It was anticipated that this may then allow for a further exploration of the different strategies that participants may use which could account for their speech perception performance.

Aims and hypotheses

Previous research has demonstrated that both amplitude and latency of neural representations of speech sounds may assist in explaining some of the differences noted in understanding speech-in-noise. This chapter sought to specifically explore the amplitude of neural responses generated by the fundamental and first formant and overall amplitude of response of the speech sound /da/ and any relationship with speech perception in noise.

A further aim was to investigate if better auditory neural representations equate to a generally well-functioning neural system. Since cognitive functions rely on neural transmission, better auditory neural representation may be associated with longer working memory span as for both the neural system may allow for superior performance. There exists the possibility of an alternative relationship in that if better neural representations allow for better speech-in-noise performance, then those individuals may rely less on other cognitive functions, such as working memory. These participants may therefore have lower working memory capacity but still have good speech perception in noise.

The hypotheses to be tested are labelled here as A1, A2 and so on so as to distinguish these hypotheses from the other hypotheses in the main body of the thesis. The hypothesis were:

A1. If better neural representations in the brainstem of the fundamental and first formant of the sound /da/ allow for better access to speech sounds when speech is challenging to listen to, then participants' with the largest neural representations will demonstrate the best speech-in-noise discrimination scores.

A2. If larger amplitude evoked neural responses are an indicator of strong cognitive functioning then the amplitude of the neural responses will correlate significantly with working memory performance.

A3. The alternative hypothesis to the above will also be tested to the above, which is if the amplitude of the neural responses is small, this may indicate that listeners' with these smaller responses do not rely on the neural encoding as they use their working memory as their support. In this case an inverse relationship will be exhibited.

Methods used for study

Ethics

The Aston University Research Ethics Committee (UREC) approved the proposed research as part of research reference number 701.

Participants

Participants were recruited by opportune sampling and through the Psychology Research Participation System at Aston (SONA). All the participants in the research were a subset of the participants recruited for the research in Chapter 2, so some of the following sub sections are a summary of what they had already undertaken. The section below in methods outlines the procedures used to record and interpret the complex auditory brainstem responses (cABR).

Research by Anderson et al. (2011) on a group of older individuals aged 60-73, reported on a relationship of the magnitude of response in microvolts of the fundamental frequency in the brainstem in quiet (evoked by a /da/ syllable of 170msec) with speech-in-noise performance (using the HINT). The relationship reported by Anderson et al. was $r = -0.553$, $p = 0.0002$. Although the participants in this present research were younger than in Anderson et al.'s research, the methods being employed were very similar and, as such, the sample size calculation was based on the relationship established in the Anderson work. The *r* value of

.553 and an alpha value of *p* = .05 with a probability of 80% was used with G*Power sample size calculator (Faul et al., 2007). This produced a figure of 23 participants.

The participants recruited were from the 70 participants that participated in the research conducted in Chapter 3 and described in Section 3.2.2. These participants were those who consented to have a complex speech ABR performed alongside the other testing. Additionally, the participants needed to sit reasonably calmly for the test duration in order to take good quality ABR recordings. This was measured by monitoring an individual's background electroencephalography (EEG) activity during each individual recording epoch. If the activity exceeded +/- 35 microvolts, the reject level suggested by Skoe and Kraus (2010), that response would be discarded. If too many were being discarded (>10%) of the total collections, the participant was presumed to be unable to settle, so they were excluded from the data analysis. Of the 52 participants that consented to have the extra test, four were not able to settle for long enough (as described above). The exclusions left 48 participants in the study. There were 14 men and 34 women, the mean age of the participants was 27.38 years with a standard deviation of 10.01. None of the participants reported any language disorders and none were trained musicians.

General design

A correlational study was used and none of the variables were manipulated. The amplitude of the evoked neural responses of the fundamental, first formant and overall amplitude of response of the /da/ speech stimulus was correlated with the three speech tests (speech-in-noise, sine and noise vocoded).

Correlational analysis was also performed between both the BKB speech-in-noise test and the evoked neural responses and a working memory task (an RSPAN).

Test material

This chapter involved participants that were a sub-group of those that had performed the test battery for the research in Chapter 2. The initial examination, non-adaptive speech-in-noise (speech-shaped) at -8dB SNR and cognitive tests are not re-iterated here as the test material and procedures are all explained in that chapter.

The materials and procedure used for recording the additional complex auditory brainstem response (cABR) recordings on these participants is given below.

Complex ABR recordings

The recordings of the complex Auditory Brainstem Response (cABR) were performed at the end of the testing session to allow participants to relax and not to have to worry about performing further assessments. As discussed above, it is possible for the ABR to represent both the transient features of speech sounds and also the frequency following response. The /da/ syllable was chosen for this research as it had such properties.

The stimulus was a synthesised 40msec duration, five formant version of the /da/ speech stimulus (Klatt, 1980). The stimulus, therefore, had five harmonics of the original fundamental frequency. In general, the more formants, the better one vowel can be distinguished from other vowels sounds. However, in reality the higher formants do not add greatly to this discrimination. The first two are sufficient, as described in Section 4.1.1. The anticipated recorded waveform generates typical identifiable peaks (Figure 2 below: This is the same as Figure 1 but is illustrated again here to explain how the recorded activity relates to the evoking speech stimulus). The early peak, labelled V, is analogous to the response evoked by a click ABR stimulus. However, the C dip indicates where the point of transition occurs from the stop burst of the consonant, which is aperiodic, to the periodic formant transition, which is voiced. On closer examination of the waveform, the F0 (fundamental frequency), of the speech sound is represented by D, E and F. The O at the far right is the offset of the stimulus. Knowing the aspect of the stimulus that generate the response, it is then possible to judge the activity in the brainstem by analysing in the frequency domain or by looking at inter peak intervals of D, E and F for the F0. The activity that encodes the first formant (F1) is a result of smaller changes or voltage fluctuations that occur between the larger peaks of D, E and F as shown in Figure 2.

The /da/ stimulus has a F0 in the range of 103-121Hz and the F1 of the stimulus was in the frequency range 220-720Hz. The stimulus was presented monaurally at a level of 70dBSPL and at a rate of 7.1 /da/s per second.

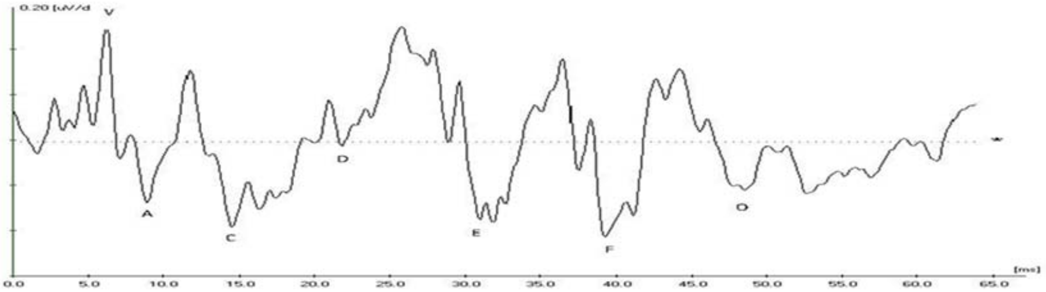


Figure 2: ABR recorded to /da/ stimulus

The participant had three electrode sites prepared, the high forehead and the reverse side of both ear lobes (Skoe & Kraus, 2010). Site preparation was performed using Nu-prep gel and the electrodes used were silver silver/chloride. The impedances of the electrodes were all <6kΩ and within 2kΩ of each other. Recordings were made from the active electrode to the reference electrode on the right reverse earlobe with the left earlobe used as the ground. The Skoe and Kraus tutorial reported that using earlobe electrode sites assisted in reducing recorded myogenic artefacts.

The stimulus was delivered to the right ear through an inserted EAR25 earphone. The recording parameters used were similar to those suggested by Skoe and Kraus (2010). The participants were seated in a comfortable chair and asked to relax as much as possible. There were two collection runs of 2000 presentations of the /da/ in an alternating polarity in an attempt to minimise stimulus artefact. The two recorded traces were compared and an average waveform for the two runs was derived. If a patient was too unsettled during the ABR collection (artefact rejection at >10%) they were excluded from the study. Recording times took on average 12-15 minutes, which allowed for a period for the participants to settle. A selection of the traces evoked by the /da/ stimulus for 6 of the participants are displayed in Figure 3. As was described above, there are two recordings for each of the participants and these were combined to give an average waveform.

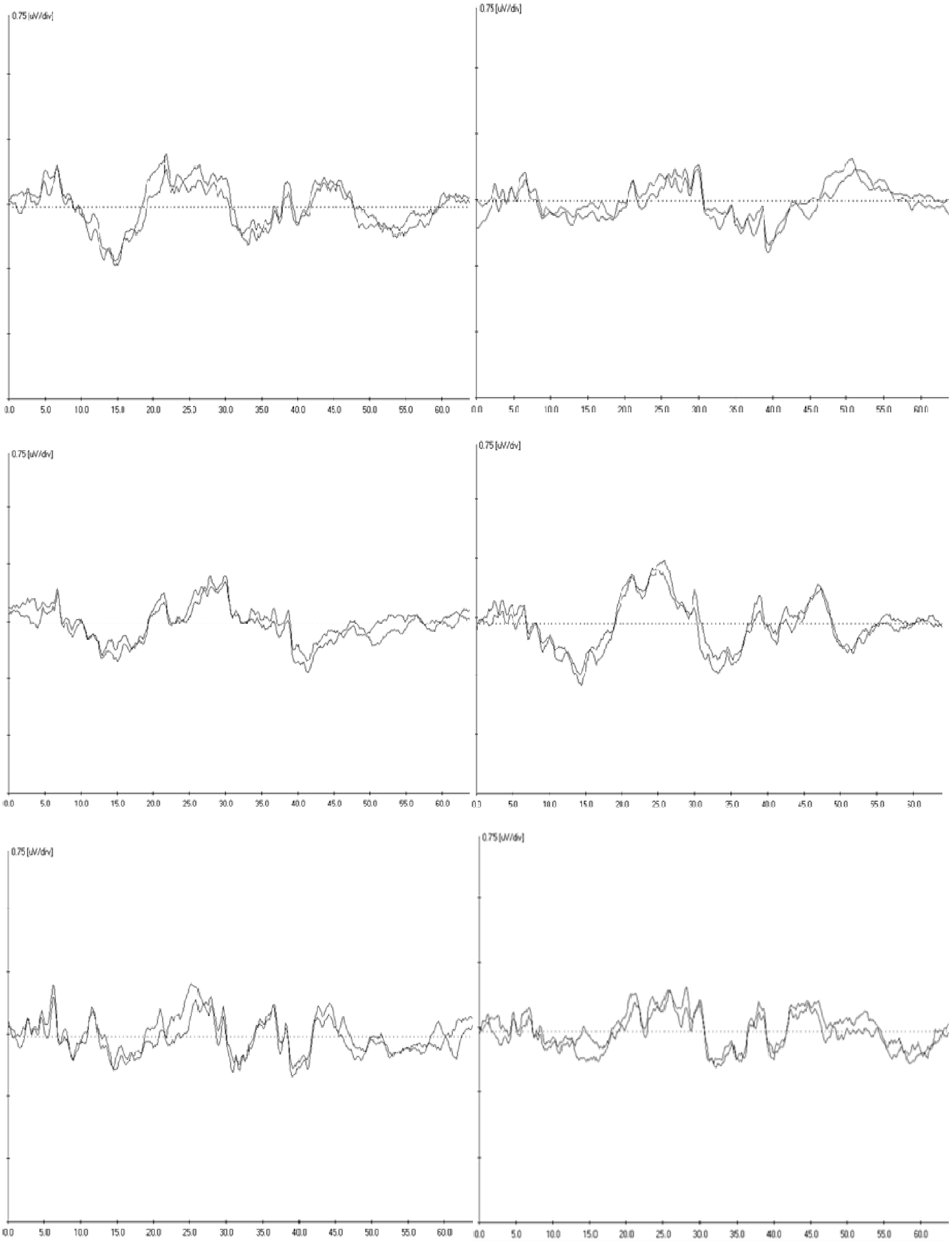
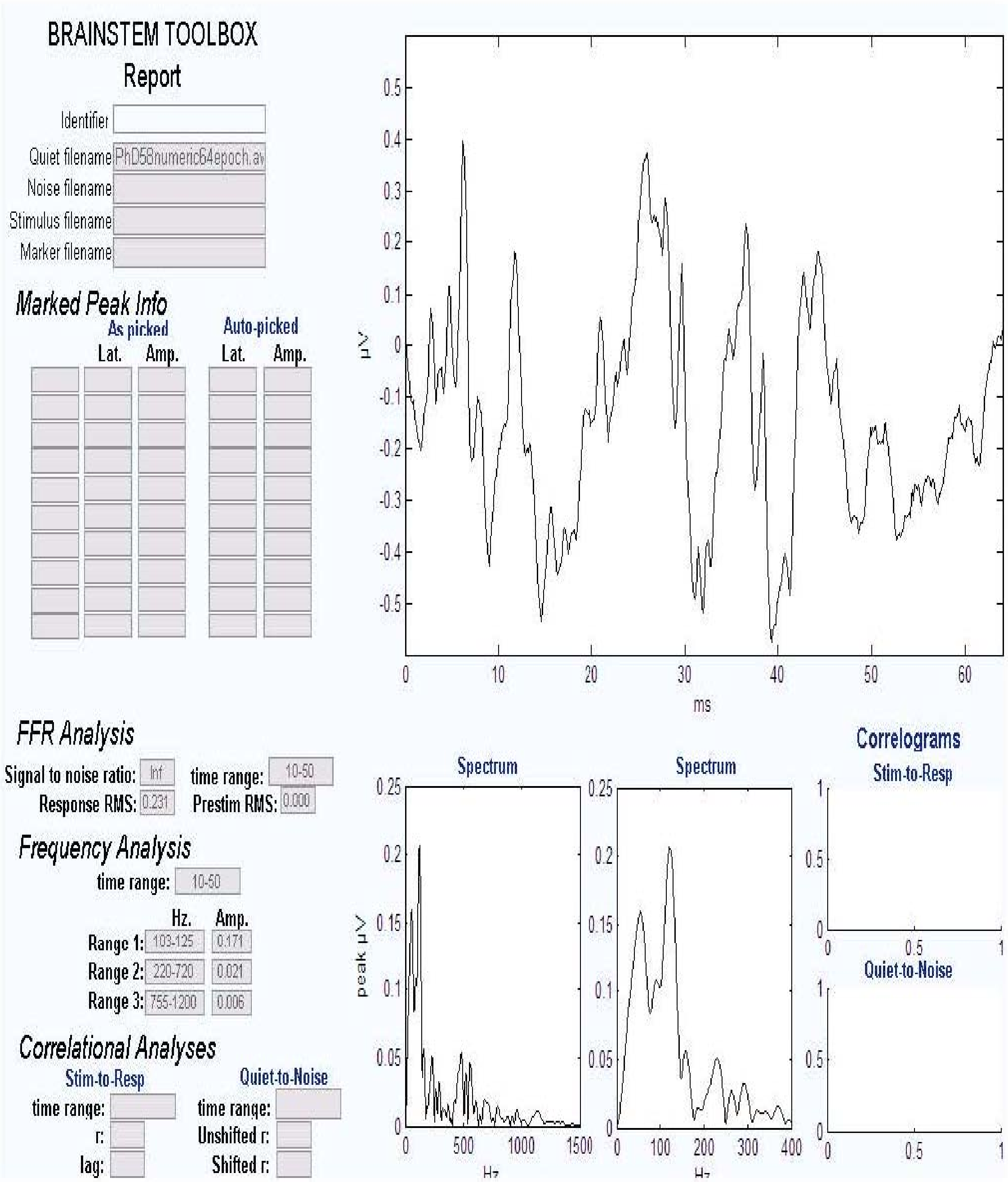


Figure 3. Six evoked ABR traces to the /da/ stimulus

The derived average waveform was exported in a .txt file using AEP2ASCII and the text files were then converted to .avg files in MATLAB (The Mathworks, 2006b). The conversion to .avg and the description of the analysis that follows were all performed using The Brainstem Toolbox (Skoe & Kraus, 2013) in the same version of MATLAB as above. The .avg files were analysed by the bt_gui module of the Brainstem Toolbox responsible for performing frequency and time domain analysis. The output from the toolbox for the same participant used as an example in Figure 2, is shown below in Figure 4.

Following the successful recording of the cABR traces and their conversion and analysis in the Brainstem Toolbox, correlational analysis between the representation of the /da/ stimulus in the ABR recording (in terms of the F0, the F1 and the RMS amplitude, all in μV) and discrimination of the three types of degraded speech was performed. Alongside these, the participant's performance of a working memory task (an RSPAN) was also analysed. Further statistical testing involved participants being split into high performers and low performers for the different degraded speech scores.



BRAINSTEM TOOLBOX
Report

Identifier:
Quiet filename: PhD58numeric64epoch.av
Noise filename:
Stimulus filename:
Marker filename:

Marked Peak Info

As picked Lat.	As picked Amp.	Auto-picked Lat.	Auto-picked Amp.

FFR Analysis

Signal to noise ratio: Inf
time range: 10-50
Response RMS: 0.231
Prestim RMS: 0.000

Frequency Analysis

time range: 10-50

	Hz.	Amp.
Range 1:	103-125	0.171
Range 2:	220-720	0.021
Range 3:	755-1200	0.006

Correlational Analyses

Stim-to-Resp
time range:
r:
lag:

Quiet-to-Noise
time range:
Unshifted r:
Shifted r:

Figure 4. The Brainstem Toolbox output for a participant in the study showing the results of the frequency analysis over the requested time range of 10-50msecs. The values used in this research are the overall Root Mean Square (RMS) response (0.231 in this example), the F_O found in range 1 (0.171) and the F_1 found in range 2 (.021) all given in µV.

Results

SPSS was used to explore the data. The general descriptive statistics are given below in Table 1. The data was assessed for normality for the performance of the different speech tests and for the fundamental, first formant frequency and the overall (RMS) amplitude. On statistical testing and analysis, the majority of the data was found to be abnormally distributed. Noise vocoded sentences and the fundamental frequency recorded in the cABR to the /da/ speech stimulus demonstrated non-normal distributions.

The non-parametric Spearman's rho correlational analysis was, therefore, performed in order to investigate hypothesis A1, that the representation of any of the evoked neural responses (F0, F1 and RMS) would be significantly correlated with the performance of any of the three type of degraded speech tasks. Spearman's rho correlational analysis was also employed to investigate hypothesis A2 and A3, to investigate if significant relationships existed between working memory (RSPAN) and neural representations. These results are given below in Table 2.

Table 1
General descriptive statistics for the three speech tests and the fundamental frequency and first formant for the sub-group of 48 participants that performed the tasks in Chapter 2

Measure	**Mean**	**Standard deviation**	**SE of the mean**
BKB speech-in-noise %Score	40.73	14.23	2.05
BKB sine vocoded %Score	32.43	18.75	2.71
BKB noise vocoded %Score	17.36	16.84	2.43
Fundamental frequency F_0	.096	.042	.006
First formant frequency F_1	.018	.005	.0007
RMS overall amplitude	.344	.185	.027
RSPAN – absolute value	29.19	14.41	2.08

Key: *BKB scores are percentages of correct words repeated out of 150 presentations in 48 sentences. The Fundamental (F_0) and First format (F_1) are measures of the frequency activity measured in µV in the 10-50 msecs after stimulus onset. The ranges analysed were F_0 = (103-125Hz) and F_1 = (220-720Hz).*

As is demonstrated in Table 1 above, the most challenging discrimination task undertaken by the participants was the noise vocoded BKB sentences where the average percentage of correctly identified words was only 17.36%. Although some participants performed this test poorly this did not give evidence of being a floor effect.

Table 2

Correlation matrix with Spearman's rho values for the degraded speech tests and the fundamental frequency and first formant for the participants (N=48)

	BKB Speech-in-noise	BKB Sine vocoded	BKB Noise vocoded	F_0	First F_1 formant	RMS value	RSPAN
BKB Speech-in-noise	—						
BKB Sine vocoded	.453**	—					
BKB Noise vocoded	.442**	.692**	—				
Fundamental frequency F_0	-.126	.027	-.037	—			
First F_1 formant	-.050	-.114	-.167	.165	—		
RMS value	.112	.029	.086	.092	.471**	---	
Reading span (RSPAN)	-.208	.134	.152	.015	.012	-.117	---

Note: *** Correlation is significant at the 0.01 level (2 tailed) * Correlation is significant at the 0.05 level (2 tailed) Key: The BKB scores used for correlation were the recorded percentages and the formants were the activity in μV in the frequency ranges indicated above in Table 1 for F_0 and F_1 and overall for the RMS value. The Reading Span Score (RSPAN) was the absolute value of numbers of letters*

Table 2 indicates that the different speech tasks correlated well with each other. Therefore, if a participant was able to discriminate one type of degraded speech well they also performed well on other degraded speech tasks. These were all found to be at the *p* <.001 level. The only other significant correlation found was between the first formant activity recorded in μV in the brainstem and the overall RMS activity in μV. The experimental hypothesis A1 was not supported as there was no significant relationship between the performance of any of the speech discrimination tasks and any of the measures of brainstem activity evoked by the /da/ stimulus. Furthermore, Hypotheses A2 and A3 were not supported as there was no significant relationship between working memory performance and any of the measures of brainstem activity evoked by the /da/ stimulus. It was in view of this lack of any significant relationship that this study was moved to the appendix.

Discussion

There was no significant correlation observed for any of the brainstem measures with any of the speech discrimination scores used in this research. Therefore, all the experimental hypotheses based on finding significant correlations were rejected (A1, A2 and A3). This may not be an unexpected result concerning the two types of vocoded speech tests as they are using novel speech for the listener to interpret. As such, there would not be any established neural representations to give any advantage in discrimination performance. It could however be argued that even when the type of speech sentences (vocoded) under test were novel to the listener, a stronger neural representation may still have allowed for quicker or better access to speech sounds. This may be due to phonologic matches being found even to degraded signals or to necessitate less explicit processing, leaving more cognitive resource to enable sentence discrimination (Rönnberg et al., 2013). This use of cognitive repair strategies and a listener's semantic knowledge may allow for better discrimination. This, however, was not supported by the correlation analysis.

There is a stronger argument that there should be a significant relationship between the speech-in-noise perception results and neural encoding. The Anderson et al. (2011) research demonstrated that high performing listeners have a richer hi-fidelity neural representation in relation to the stimulus and are better able to use the cues they get from this speech encoding in the neural system. This, in turn, allows for better speech perception in noise. There was, however, no significant relationship established for the speech-in-noise task in this present research.

Previous research has indicated a relationship with the frequency representation of /da/ in the brainstem. Some of the key differences between previous research and this current research may account for the difference in outcome. The Anderson et al. (2011) research involved older adults and older adults are starting to show evidence of neural slowing from the age of 47 (Parbery-Clark et al., 2012). It may be therefore that it is not until individuals are above this age that differences in neural representations become significant. With older adults, this in effect may actually be a measure of neural changes due to age, even if not exhibiting any hearing loss, rather than a measure of the fine temporal analysis or tuning of individuals different neural systems.

This current research used a different stimulus and neural representation analysis to previous research with young adults (Parbery-Clark et al., 2009) which has not been researched before, the findings, however, were similar. When recording the ABR in quiet there was no significant relationship between the fine encoding of the speech stimulus as measured by the amplitude of the F0, F1 or the RMS, with speech perception in noise performance (or for the two vocoded speech tests in this present research). It could be hypothesised that in order to demonstrate a significant relationship, the cABR needs to be performed in the conditions in which the potential relationship is being explored. For example, if exploring speech-in-noise performance and a relationship with neural responses evoked in the cABR, then the cABR also needs to be performed in noise. A further hypothesis could be put forward that it is the modulation of the evoked response from the higher cortical areas that is more important rather than the initial encoding. This may explain the different relationships demonstrated when the cABR recordings were made in quiet compared to noise (Parbery-Clark et al., 2009). The reason being that in challenging listening situations there will be a greater demand on the higher cortical areas to modulate the incoming signal to make it more salient, so focussing on encoding the signal and not the noise. This will in turn be demonstrated by a better neural encoding but only after the ‘top-down’ modulation has taken effect (Bajo & King, 2012). Designing research to explore which may have the greater influence would give a useful insight into the complex nature of initial encoding and any modulation effects.

Previous research has demonstrated the importance of the F0 pitch cue on hearing in background noise, as it allows a listener to ‘tag’ a target voice (Anderson et al., 2011). Particularly in situations where there are gender differences between talkers, fine detail phonetic information is highly important in judging the gender of a talker and in searching for phonetic resemblances in longer term memory (Fellowes, Remez, & Rubin, 1997; Remez, Fellowes, & Rubin, 1997). It would seem possible to imagine the use of this in real world situations where trying to follow one talker amongst others. This spatial awareness ability of locating the talker is a key skill when listening in noisy group situations. This in turn would allow for better streaming of the different auditory inputs so the target talker can be attended to and the other streams inhibited. If the F0 pitch cue plays a role in this, then on review of the research undertaken for this chapter, it may well be that a stronger relationship would have been exhibited between a large amplitude F0 neural encoding and speech perception in noise score when the speech was presented in babble. This would be due to the ability to lock into the target single voice whilst inhibiting the other informational background speech babble.

It is also possible that the shorter /da/ of 40msecs, chosen as it was the amplitude of the response that was being investigated, does not explore the hi-fidelity processing of the neural system in as much detail as the longer 170msec /da/ and that some greater standardisation of the stimuli of choice may be necessary for different groups of listeners for further research. For listeners presenting with similar hearing levels and younger age grouping, it may be necessary to extract more detailed information to display any significant relationships or differences. The longer stimulus may allow for this.

For younger listeners (18-50 years), where hearing levels are all within normal limits discriminating degraded speech demonstrates a reasonable wide variation of performance. The brainstem amplitude activity evoked by a short duration speech phoneme /da/ does not significantly explain performance variations (this chapter) and working memory has also been demonstrated as not having a significant role (Füllgrabe & Rosen, 2016b). If neural encoding does have a role to play in explaining any performance variation, the speech stimuli used to evoke responses and recording conditions (such as recording speech ABR's in noise) need further exploration as potentially stressing the neural system may give a greater insight into the mechanisms involved.

Limitations

The research in this chapter was designed specifically to explore the relationship of neural representation of a speech stimulus in younger individuals when speech signals were degraded. Having a wider age range and a greater number of participants would have allowed for groups to be split along age lines and allowed for analysis of changing relationships of neural encoding and degraded speech.

The lack of any significant relationships in this chapter's research compared to previous research puts forward an argument that potentially performing the cABR in noise to generate more potential predictors and to stress the neural system may have added value. Performing cABR's in both quiet and noise to explore relationships with degraded speech may therefore have added benefit.

Conclusions

There was no clear evidence that for younger individuals, the amplitude of responses in the brainstem evoked by the speech sound /da/ predicted performance of understanding speech in degraded listening conditions. This was the case for all the different constituent parts of /da/ that were investigated, the fundamental frequency, first formant and overall RMS amplitude. It does not appear, therefore, that neural encoding of a short duration speech sound in quiet is a useful predictor of degraded speech performance in this younger cohort.

As an overall significant relationship had not been demonstrated between neural representations and speech-in-noise performance, the participants had been split into high and low performers for each of the speech tests. This was to investigate if there were subtle differences in neural encoding that accounted for the better listening performance of the high performing group compared to the lower performing group. On the data analysis there were no significant differences observed.

It would appear that in order to utilise neural encodings as predictors of degraded speech performance that either;

1. The neural recordings need to be performed when the neural system is under stress (recording the cABR in background noise).

2. The participants under investigation need to have a potential wider range of neurological differences. For example to include those with declining neurological speeds, such as older participants, or those with suspected neurological deficit, such as participants with auditory processing disorders, or those with enhanced neural representations such as musicians.