UNIVERSITY OF SOUTHAMPTON

Auditory acclimatisation to amplified speech in adults

Kevin James Munro

A thesis submitted for the degree of

Doctor of Philosophy

Institute of Sound and Vibration Research

Faculty of Engineering and Applied Science

March 2002

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING AND APPLIED SCIENCE
INSTITUTE OF SOUND AND VIBRATION RESEARCH

Doctor of Philosophy

AUDITORY ACCLIMATISATION TO AMPLIFIED SPEECH IN ADULTS

By Kevin James Munro

With the provision of amplification, the listener receives newly available speech cues that were previously inaudible. In addition, cues that were previously audible are shifted towards the higher intensity end of the neural representation. These changes may immediately confer greater intelligibility. However, it is possible that the provision of amplification may induce reorganisation within the auditory system resulting in perceptual learning and an improvement in performance over time. Studies measuring speech recognition, intensity discrimination and loudness perception have confirmed the existence of auditory acclimatisation. However, some recent studies have failed to demonstrate an acclimatisation effect. This has led some researchers to state that the acclimatisation effect is small or non-existent. These conflicting findings suggest that evidence of acclimatisation may only be apparent under certain test conditions. The aim of this thesis was to improve understanding of the conditions required to measure acclimatisation. This should result in a more robust methodological framework for measuring acclimatisation in future studies.

Three experiments were undertaken on separate groups of 16 subjects with moderate bilateral sensorineural hearing impairments. Subjects were fitted monaurally with a linear hearing instrument that provided approximately 20 dB real ear insertion gain at 2-4 kHz. The not-fitted ear was used as the control. The self-reported use of the hearing instrument was typically 8-12 hours per day. The main outcome measure was the Four Alternative Auditory Feature speech recognition test. Acclimatisation was defined as an improvement over time in recognition score in the aided ear relative to the unaided ear.

In the first experiment, subjects were tested with speech presented at an overall level of 60 dB SPL at three-week intervals over a post-fitting period of 24 weeks. The results revealed an improvement in performance in both ears over time but there was no consistent evidence of acclimatisation. In the second experiment, new subjects were tested with speech presented at an overall level of 60 dB SPL at four-week intervals over a post-fitting period of 24 weeks. In an attempt to reduce the practice effect observed in the first experiment, subjects were not given trial-by-trial feedback about their performance. On this occasion, there was a non-significant trend of improving performance in the fitted ear with no improvement in the control ear. It was hypothesised that the amplified level of speech in these experiments was not sufficiently different from that experienced in everyday life, prior to aiding, to show perceptual learning.

In the final experiment speech was presented at overall levels of 55, 62 and 69 dB SPL and new subjects were tested at six-week intervals over a post-fitting period of 12 weeks. There was a significant improvement in performance over time with a trend towards greater acclimatisation at the highest presentation level. The findings are consistent with a process of reorganisation of the auditory cortex in the intensity domain resulting in an increased representation to behaviourally important speech sounds that are perceived at a higher intensity level than experienced before aiding. In order to detect acclimatisation, it will be necessary for future studies to use an amplified speech signal that is more intense than commonly experienced in everyday life, prior to aiding.

CONTENTS

FIGURES

Chapter One

Chapter Two

Chapter Three

Chapter Four

Chapter Five

Chapter Six

TABLES

Chapter One

Chapter Two

Chapter Three

Chapter Four

Chapter Five

Chapter Six

ACKNOWLEDGEMENTS

I am grateful to Professor ME Lutman, Institute of Sound and Vibration Research, for suggesting auditory acclimatisation as a suitable area of study; in addition, I am indebted to him for his guidance and support as project supervisor.

I would like to thank Professor ARD Thornton, MRC Institute of Hearing Research and Dr IH Flindell, Institute of Sound and Vibration Research, for their comments and advice during the review board process.

During the course of the thesis I received advice from a number of national and international colleagues: many authors of previous acclimatisation studies provided clarification about their own work and offered helpful suggestions about the present study. In particular, I wish to express my appreciation to Professor S Gatehouse, MRC Institute of Hearing Research, who was always helpful and courteous, despite my frequent questions. The following individuals provided comments on an earlier draft of the thesis: Dr LE Humes, Indiana University, Dr JE Marriage, University of Cambridge, Professor DR Moore, University of Oxford and Dr EW Yund, Veterans Affairs Medical Centre, Martinez, CA.

I am grateful to the authors and publishers who gave me permission to reproduce figures from their work. I am also grateful to the MRC Institute of Hearing Research for providing a recording of the speech materials used in the experiments.

I would like to thank Nicola Hatton and Debra Graumann for their involvement in data collection during the early stages of the project. I would also like to acknowledge the generous financial support provided by the NHS Research and Development Directorate.

I would like to use this opportunity to record my appreciation of a number of individuals who have been instrumental in my career. They include my parents, Kenneth and Anne Munro, Mrs IJ West, Royal Infirmary Edinburgh, and Professor RC Seewald, University of Western Ontario, Canada.

This thesis is dedicated to the memory of Simon Stanley [1962-2000], a devoted family man and schoolteacher whose friendship and good company is greatly missed.

ABBREVIATIONS

ABR	auditory brainstem response
AGC	automatic gain control
ANOVA	analysis of variance
BKB	Bamford-Kowal-Bench
BTE	behind-the-ear
CANS	central auditory nervous system
CID	Central Institute for the Deaf
CNS	central nervous system
CST	connected speech test
CUNY	Central University of New York
DLI	difference limen for intensity
DSL	Desired Sensation Level
FAAF	Four Alternative Auditory Feature
FDL	frequency difference limen
fMRI	functional magnetic resonance imaging
GHADP	Glasgow Hearing Aid Difference Profile
HAPI	hearing aid performance inventory
HINT	Hearing In Noise Test
HHIE	Hearing Handicap Inventory in the Elderly
HPI	Hearing Performance Inventory
ILD	interaural level difference
ITD	interaural time difference
MCL	most comfortable level
MEG	magnetoencephalography
MMN	miss-match negativity
MRT	Modified Rhyme Test
NAL-R	National Acoustics Laboratory-Revised

NST	Nonsense Syllable Test
PB	phonetically balanced
PET	positron emission tomography
PHAB	Profile of Hearing Aid Benefit
RAU	randomised arcsine unit
REAR	real-ear aided response
RECD	real-ear-to-coupler difference
REIG	real-ear insertion gain
REUR	real-ear unaided response
SII	Speech Intelligibility Index
SPIN	Speech in Noise test
SNR	signal-to-noise ratio
SPAC	Speech Pattern Auditory Contrast
SRS	speech reception score
SRT	speech reception threshold
SVT	Sentence Verification Test

CHAPTER ONE
INTRODUCTION

Many studies have confirmed the existence of neural plasticity in the central auditory nervous system [CANS] of adult mammals [Palmer *et al.*, 1998]. The consequences of this plasticity are unclear but these may have important functional, behavioural, and perceptual effects that are relevant to hearing impairment and subsequently the use of a hearing instrument.

A model that illustrates how these changes might occur is shown in Figure 1.1. This model represents a conceptual framework that may explain the mechanism by which changes in performance can occur following regular use of a hearing instrument. The normal relationship is shown in Fig 1.1a with conversational speech [S; left column] being mapped to the middle of the neural representation [S; right column]. This mapping is represented by the horizontal line connecting the two columns. The purpose of the shaded area is to demonstrate that neural resources are concentrated around the position of conversational speech within the CANS. In a hearing-impaired subject, the level of conversational speech is unchanged although the speech signal is now attenuated by the peripheral auditory system. As a result, speech is located in a different position within the CANS [as shown by the downward facing arrow in Fig 1.1b]. However, reorganisation has occurred within the CANS so that resources have become concentrated at the lower end of the neural representation [to reflect the position of conversational speech within the subject's auditory range]. Immediately after provision of appropriate amplification, the input level of conversational speech increases and this alters the location of speech within the CANS [as shown by the upward facing arrow in Fig 1.1c]. However, there is a mismatch between the concentration of neural resources located at the lower end of the neural representation and the location of conversational speech within the CANS. After longstanding use of amplification there is reorganisation of the neural map to accommodate the shift in the intensity domain; neural resources become concentrated at a higher position in the neural representation where speech is located [as shown in Fig 1.1d]. It is possible that an improvement in performance over time is due to this reorganisation. [The effect of this reorganisation on quiet and raised speech is discussed below]. The shift in neural

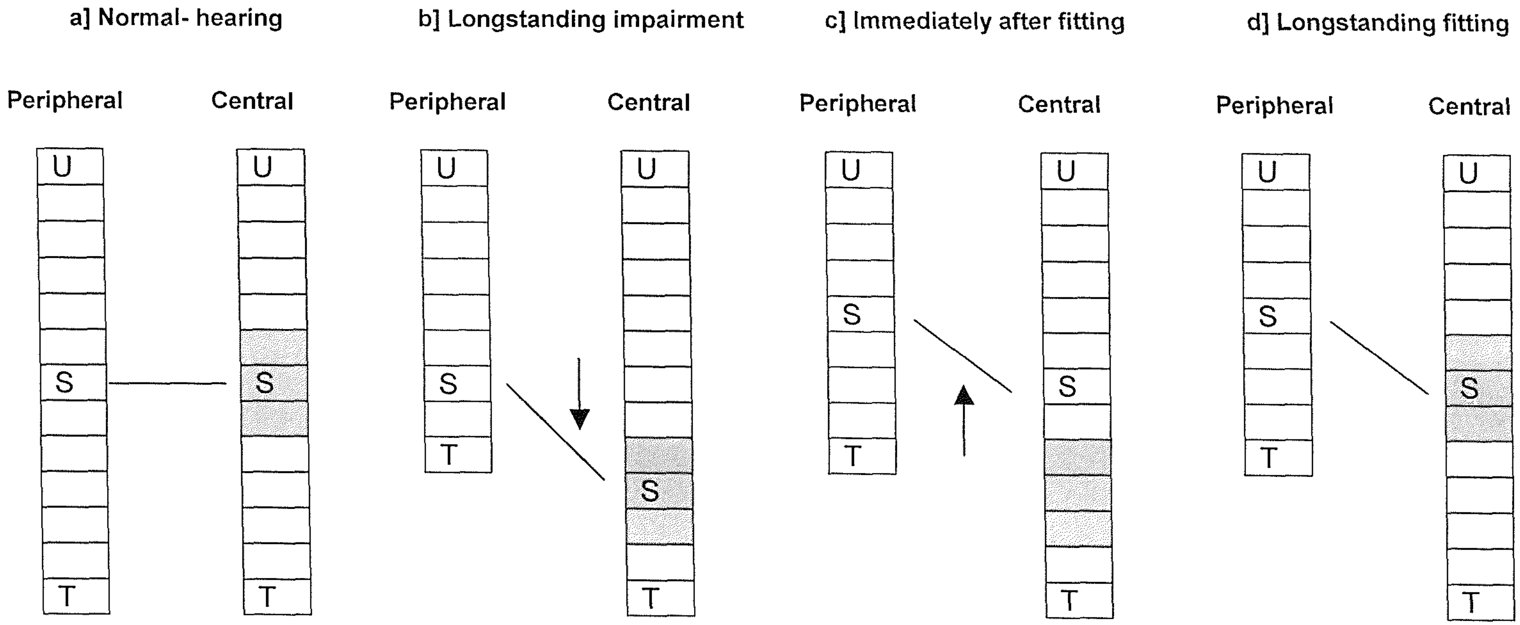


Figure 1.1. A model showing the mapping of intensity between the peripheral auditory system and the neural representation within the central nervous system [CNS]. The symbols T, U and S represent the sound pressure level at threshold, uncomfortable loudness level and conversational speech respectively. The normal relationship is shown in [a]. After a longstanding hearing impairment, the neural representation acclimatises as shown in [b]. Immediately after fitting, there is a miss-match between the concentration of neural resources and the location of speech within the CNS as shown in [c]. After regular use of a hearing instrument the neural representation re-acclimatises so that the concentration of neural resources within the CNS occurs for speech as shown in [d].

resources is consistent with the findings of the intensity discrimination study by Robinson and Gatehouse [1996]; in subjects with a bilateral hearing impairment who were fitted monaurally, the fitted ear was superior at higher presentation levels but inferior at lower presentation levels [see section 2.5.3]. The change after aiding may immediately confer greater intelligibility through the provision of extra information. However, it is possible that listeners will also improve in their ability to use these speech cues over time. Neural plasticity may be the underlying physiological mechanism responsible for this process of perceptual learning.

The terminology that has been used to document changes in auditory performance following fitting of a hearing instrument is diverse. The early studies reporting changes over time were concerned with the apparent decrement in speech recognition performance in the not-fitted ear of subjects who had been fitted monaurally. This has become known as the **auditory deprivation effect** and has been defined as '*a systematic decrease over time in auditory performance associated with the reduced availability of acoustic information*' [Arlinger *et al.,* 1996]. This is a general definition that makes no assumption about the aspect of auditory performance that may change and is independent of the severity of the hearing impairment and the symmetry of the acoustic information available to the listener. The implication is that deprivation is a consequence of providing amplification to one ear with the loss of relative importance of the not-fitted ear. In the context of this thesis, the auditory deprivation effect refers to the reduction in speech recognition performance that occurs in adults who have a bilateral sensorineural hearing impairment and have long-term use of a single hearing instrument. In these subjects, there is an increase in acoustic information in the fitted ear relative to the not-fitted ear. This may lead to greater allocation of neural resources to areas representing the fitted ear, and consequently a paucity of resources available to the not-fitted ear. Auditory deprivation has been shown to occur, over a period of years, in adults and children with sensorineural and conductive hearing impairment and, in a number of case studies, recovery has been demonstrated when subjects were fitted with binaural hearing instruments [Neuman, 1996]. However, it appears that a large-scale prospective study reporting incidence, magnitude and time course of the deprivation effect has yet to be conducted.

More recent studies have investigated improvements in performance over time in the ear that has been fitted with the hearing instrument [Turner *et al.*, 1996]. This change in performance has been associated with a variety of terminological labels including maturation of hearing aid benefit, learning effects, training, adaptation, habituation and acclimatisation. The term that has become largely accepted within the field is **auditory acclimatisation** and has been defined as '*a systematic improvement in auditory performance with time, linked to a change in the acoustic information available to the listener. It involves an improvement in performance that cannot be attributed purely to task, procedural or training effects*' [Arlinger *et al.*, 1996]. This is a general definition that makes no assumptions concerning the aspects of auditory performance that may change over time nor does it make any assumption about how the auditory information has been changed; however, it does make it clear that the improvement in performance is not simply due to procedural aspects of the experiment unrelated to changes in other auditory experience. This point is illustrated conceptually in Figure 1.2. After fitting of a hearing instrument, the subject returns for four test sessions [designated periods T1-4 on the top line]. Changes in performance may occur due to the experience of listening with the hearing instrument during the testing sessions [designated as periods Ia-d on the middle line] and in everyday life [designated as periods IIa-d on the bottom line]. Although the change in performance in everyday life will be restricted to the ear that is normally fitted with the hearing instrument, the experience gained during the testing period may be sufficient to change performance, even in an ear that is only fitted with the hearing instrument for the testing period. In the context of this thesis, which involves the monaural fitting of subjects with a bilateral sensorineural hearing impairment, auditory acclimatisation is defined operationally as the difference in speech recognition performance over time between the fitted ear and the not fitted control ear [i.e., the difference in performance that occurs over time only as a result of wearing the hearing instrument in everyday life]. In previous studies that show an acclimatisation effect, this becomes evident from around 4-6 weeks post-fitting.

The conceptual model that was used to illustrate the possible mechanism for acclimatisation suggests that this effect would be greatest for conversational [or raised] speech. This is in contrast to the commonly held view that acclimatisation is due to the subject learning to make use of newly audible speech sounds since this suggests that the effect will be greatest for quiet speech. This can be explained by considering the location of quiet and raised speech within the CANS immediately after aiding [Fig

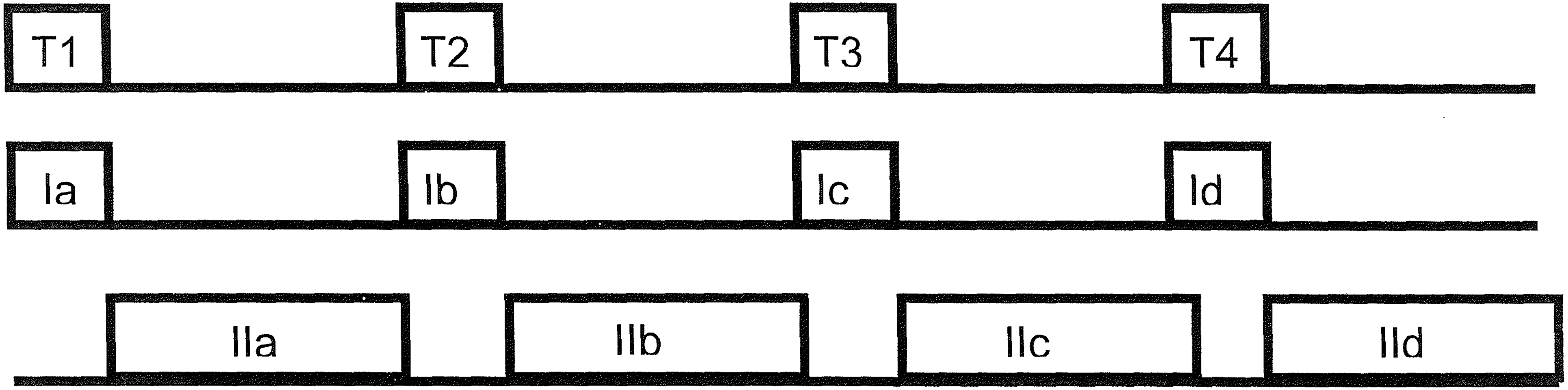


Figure 1.2. Conceptual illustration of auditory acclimatisation. In this example, the subject returns for four test sessions as shown on the top line [designated periods T1-4]. Changes in performance due to listening experience with the hearing instrument during the test session are shown on the middle line [designated periods Ia-d]. Changes in performance due to listening experience with the hearing instrument in everyday life is shown on the bottom line [designated periods IIa-d].

1.1c]. Quiet speech may fall within the upper boundary of concentrated neural resources but raised speech will be well above the upper boundary. This means that immediately after aiding, the concentration of neural resources will favour quiet speech. However, over time, the concentration of neural resources will shift as shown in Figure 1.1d. This means that performance may decrease over time for quiet speech but increase over time for average and raised speech [as the concentration of neural resources shift within the CANS].

Auditory acclimatisation is of considerable relevance to the researcher and the clinician; in addition to extending our knowledge of the auditory system, the effectiveness of a hearing instrument prescription, or comparison between different signal processing strategies should take place after any changes in perceptual learning have occurred. A number of studies have been published that have specifically investigated the phenomenon of auditory acclimatisation in adults fitted with an acoustic hearing instrument. The main motivation for most of these studies has been the potential relevance to clinical practice. Since there is no consensus as to what outcome procedure provides a definitive measure of hearing instrument benefit, the studies have used a variety of performance and self-report procedures.

The review in the next chapter has been divided into a number of sections commencing with a discussion of the evidence for plasticity in the central nervous system [CNS] in adult mammals. This is followed by a section on perceptual learning. Perceptual learning is the ability to improve performance over time by learning to extract information from the stimulus that was previously unused; the premise of this definition is that perceptual learning benefits an individual by tailoring the information gathering process to the use of the information. The characteristics of auditory acclimatisation suggest that this is perceptual learning occurring within the auditory system. The perceptual learning of interest in this thesis is that occurring during everyday hearing instrument use and hence for sounds heard via the fitted ear. The focus is not on perceptual learning taking place during the test sessions, as described above.

The review then concentrates on studies investigating changes in hearing instrument benefit over time. It begins with a review of late-onset auditory deprivation and is followed by auditory acclimatisation. For the purposes of the review, the studies on auditory acclimatisation have been split into those that appear to demonstrate an improvement over time from those that do not show an improvement. The penultimate section reviews studies that provide supporting evidence for the existence of auditory acclimatisation. This includes studies that have shown a change in intensity discrimination as well as studies showing a change in perception of loudness after hearing instrument fitting.

The main conclusion of the review is that there is irrefutable evidence for the existence of auditory acclimatisation. Despite this conclusion, many recent studies have failed to measure an acclimatisation effect. This has led some researchers to state that the acclimatisation effect is small or non-existent. These conflicting findings suggest that evidence of acclimatisation may only be apparent under certain test conditions. A review of the literature suggests that the following five conditions required for the measurement of acclimatisation:

1. Subjects should have a sufficient degree of hearing impairment so that the audibility of important signals such as speech is improved with aiding. This may relate to re-mapping, as illustrated in Figure 1.1. As yet, influences of the age of the subject and the degree/configuration of the hearing impairment are unknown.

2. The hearing instrument should make a change to important information such as speech. It is not clear if this means that the hearing instrument should improve the audibility [i.e., change from inaudible to audible] of conversational speech or if it results in re-mapping of the relationship between intensity and loudness as described in Figure 1.1 [in which case, it could conceivably occur in normal hearing individuals by increasing the overall sound level of speech that is already audible]. The implication of points one and two is that the greater the change in information, the greater will be the improvement in performance over time.

3. The outcome measure should be sensitive to the changes that occur as a result of acclimatisation. For example, in subjects with an age related hearing

impairment that have been provided with high frequency amplification, the speech recognition material must be sensitive to changes in the audibility or re-mapping of this high frequency information [see Figure 1.1c].

4. The speech material must be reliable enough to show the effect size and to be free from ceiling/floor effects. The variability in performance associated with repeated speech testing is well known and can be overcome by, for example, increasing the number of test items. In addition, the test should be configured so that initial performance is such that changes over time [increases or decreases] can be measured. For example, noise can be used to limit maximum performance.

5. There should be a control condition that will allow improvements over time due to use of the hearing instrument to be differentiated from any other improvements. For example, improvements due to practice with the test material can be identified in control subjects who have not been fitted with a hearing aid or, alternatively, in individuals who are already experienced hearing instrument users.

These five conditions are summarised in Table 1.1.

Table 1.1 Conditions required for the measurement of auditory acclimatisation

1. hearing-impaired subjects
2. amplification which makes a significant change to important signal such as speech
3. outcome measure should be sensitive to the changes that occur as a result of acclimatisation
4. outcome measure should be reliable enough to show the effect size and be free from floor/ceiling effects
5. there should be a control condition so that improvements due to wearing the hearing instrument can be differentiated from other improvements

It is not possible to report the prevalence and effect size of acclimatisation, its full time course, or the effect of binaural versus monaural amplification until the conflict over the conditions required for acclimatisation to be measured is resolved. The aim of this thesis was to improve understanding of the conditions required to measure acclimatisation and hence, explain the conflict concerning the existence, or otherwise, of acclimatisation. This should result in a more robust methodological framework for future studies on auditory acclimatisation.

Auditory acclimatisation was investigated in three experiments using separate groups of new hearing instrument users. The methodology that was common to these experiments is reported in chapter three. The first two experiments are reported in chapter four and the final experiment, which builds on the earlier findings, is reported in chapter five.

Throughout the thesis, the ear fitted with the hearing instrument is referred to as the 'fitted' ear while the ear not normally fitted with the hearing instrument is referred to as the 'not-fitted' ear, the latter being considered as the control ear. These terms were used to avoid confusion between the aided and unaided test condition, which could apply to each ear. In addition, hearing instrument benefit refers to the difference between the ear's performance with and without the hearing instrument.

CHAPTER TWO
BACKGROUND

2.1 Plasticity in the adult central nervous system

The central nervous system [CNS] in mammals is characterised by an orderly representation of neurones that exist between the peripheral receptors and the higher levels within the brain. This ordered relationship permits the plotting of 'neural maps' within the CNS. These maps have long been known to change and reorganise in developing animals. This plasticity is responsible for the alteration of structure and function of the brain over time, during development and in response to environmental change [Kass, 1995]. Until recently, there was widespread belief that sensory systems were highly plastic only over a short developmental period termed a 'critical period'. Experiments using microelectrode-mapping techniques in mature animals have now provided evidence that receptive areas in the cortex of adult mammals also change with learning and experience.

Many of the early experiments on neural plasticity in adults were concerned with reorganisation within the somatosensory system. These initial experiments used major sensory deprivations [such as nerve cuts or crushes] to produce map changes that could be measured effectively. An example of this type of experiment is the 1984 study by Merzenich *et al.* who demonstrated reorganisation of the sensory map of the hand in monkeys after surgical removal of a digit. Several months later, the adjacent digit activated the cortical area that had previously responded to the amputated digit. A number of subsequent studies in monkeys and other mammals have demonstrated plasticity within the somatosensory map in adult mammals. More recently, studies have extended this finding to include humans. Mogilner *et al.* [1993] used a technique known as magnetoencephalography [MEG] to demonstrate changes in the area of cortical activity before and after surgical separation of webbed fingers in two adult subjects. The pre-surgical maps displayed somatosensory activity over a reduced cortical area; however, within weeks of surgery, cortical reorganisation (i.e., activity now occurred over a larger area) correlating with the new functional status of the separated fingers was evident. These experiments have demonstrated that plasticity occurs in the somatosensory cortex of adult mammals including humans. Subsequent experiments have been carried out on other sensory systems including the auditory system. A variety of techniques have been used to induce a sensory hearing impairment when investigating plasticity in the primary auditory cortex and a selection of these is summarised in Table 2.1.

Table 2.1. Examples of the techniques used to investigate central auditory plasticity. Adapted from Palmer *et al.* [1998].

Technique	Author	Species	Findings
Mechanical	Robertson & Irvine [1989]	guinea pig	Auditory cortex responds to adjacent frequencies
Ototoxic	Schwaber *et al.* [1993]	macaque monkey	Auditory cortex responds to adjacent frequencies
Noise	Kaltenbach *et al.* [1992]	hamster	Refinement of frequency representation within DCN that normally correspond to the damaged area of cochlea
Age-related	Willot [1984, 86, 96a] Willot *et al.* [1988, 1993]	C57 mice	Neurones in high frequency region become more responsive to mid and low frequencies
Training	Recanzone *et al.* [1993]	owl monkey	Cortical representation greater at trained frequency

Robertson and Irvine [1989] produced a mechanically induced lesion in a restricted region of the cochlea. Some 1-3 months later, the deprived area of the primary auditory cortex became responsive to frequencies adjacent to the frequency region damaged in the cochlea. Schwaber *et al.* [1993] used ototoxic drugs to induce a high frequency hearing impairment in monkeys. Two to three months later, recordings in the primary auditory cortex showed that regions formerly responsive to high frequency tones were now responsive at normal hearing thresholds to lower frequencies. Kaltenbach *et al.* [1992] caused cochlear damage in hamsters by exposure to intense noise. Some 1-3 months later, there was a shift in the frequency representations within the dorsal cochlear nucleus. The studies described above have induced a hearing impairment mechanically, ototoxically or via excessive noise exposure. Willott has used a mutant strain of mouse with accelerated ageing [Willot 1984, 1986, 1996a; Willot *et al.*, 1988, 1993]. This makes these mice ideal for the study of plasticity associated with age-related hearing impairment. The mice develop a high-frequency sensorineural hearing impairment commencing from 2-3 months of age. The outcome is that neurones in the high-frequency tonotopic region of the inferior colliculus and auditory cortex become more responsive to the middle and lower frequencies. The tonotopic relationship reported in all of these studies is summarised schematically in Figure 2.1. The normal tonotopic map is shown in illustration (a). After damage to the high frequency region of the cochlea, the direct link to that area of the cortex is damaged [illustration (b)]. After a period of reorganisation, the neurones that previously responded to the high frequencies now become activated by the adjacent frequencies [illustration (c)].

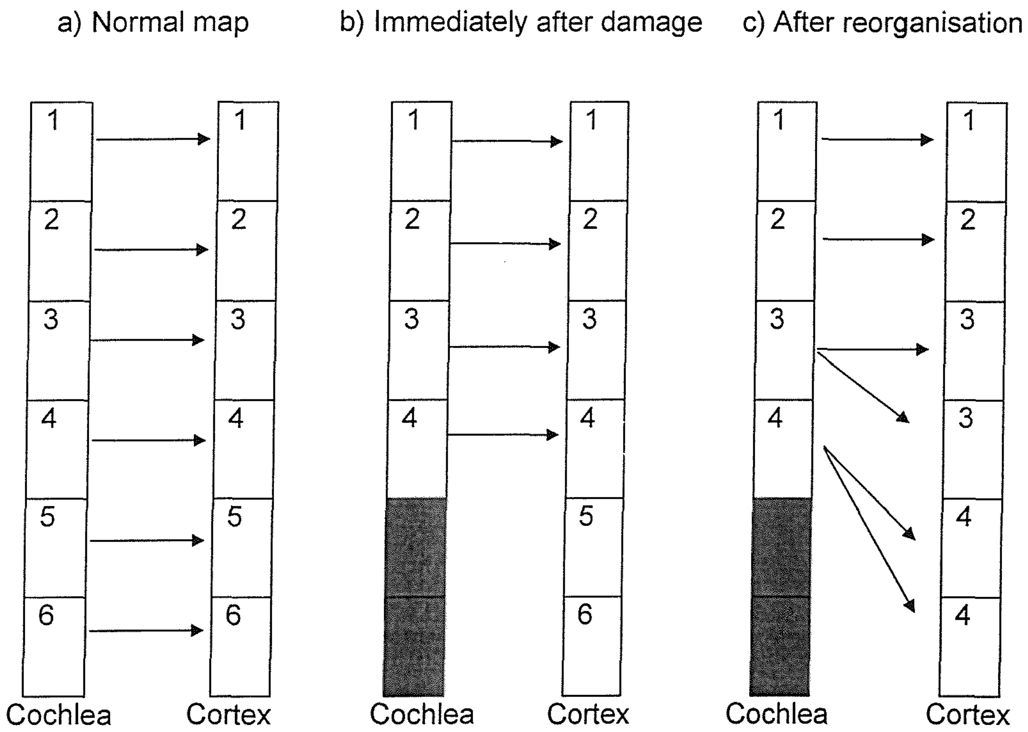


Figure 2.1. A schematic representation of the tonotopic map before and after reorganisation. The normal ordered relationship between the cochlea and the auditory cortex is shown in (a). After damage, the normal relationship is disrupted (b). After reorganisation (c) the cortex responds to adjacent frequencies resulting in an enlarged representation for these frequencies.

Vasama and Makela [1995] used MEG to record auditory evoked magnetic fields in response to acoustic stimuli of rapid onset in human subjects. They compared the response of eight adult human subjects 2-5 years after the onset of a sudden, unilateral, sensorineural hearing impairment with a control group of eight normally hearing adult subjects. The normal response has a latency of around 100 ms. The authors noted a variety of differences between the groups [for example, the latency was shorter in four of the hearing impaired subjects]. This suggests that hearing impairment may be responsible for modifying the CNS of adult subjects. A prospective study using subjects who are recently hearing-impaired has yet to be published.

A phenomenon that is due, at least in part, to plasticity is tinnitus evoked by eye movements and/or sustained lateral gaze after surgical removal of a cerebellopontine angle tumour [Lockwood *et al.*, 1999]. Cacace *et al.* [1994] have speculated that this may be due to cross-modality plasticity; axons from the oculomotor nerve nucleus invade the adjacent auditory nerve nucleus so that every time a signal is sent from the brain to move the eyes, the command is inadvertently sent to the auditory nerve nucleus and is perceived as tinnitus. Recently, Lockwood *et al.* [2001] used positron emission tomography [PET] to confirm the presence of this cross-modality plasticity. A related case study of a hearing impaired man who experienced tinnitus evoked by movement of one finger has been reported by Cullington [2001].

There is a substantial body of literature documenting cross-modal plasticity. Finney *et al.* [2001a] cite many articles that suggest that the visual cortex may come to serve other sensory modalities [i.e., tactile or auditory] when deprived of its normal [i.e., visual] input in blind humans. Parallel to this, Finney *et al.* [2001b] used functional magnetic resonance imaging [fMRI] to show that the auditory cortex of congenitally hearing impaired humans can be activated by visual stimuli: cortical activity caused by moving dot patterns produced significantly greater activity in the auditory cortex of six severely hearing impaired subjects than in six healthy controls.

In addition to creating peripheral lesions, the effect of training has also been used to study plasticity in the CNS. Recanzone *et al.* [1993] trained monkeys in an auditory discrimination task for several weeks and then mapped the tonotopic representation of the primary auditory cortex after improvements in discrimination performance were measured. The response was compared to two control groups of monkeys: those receiving no training and those that received auditory stimulation while concurrently engaged in an unrelated tactile discrimination task. The trained monkeys showed a statistically significant improvement in perceptual discrimination that paralleled an enlargement of cortical representations for the trained frequencies. There were no such changes in the control groups. The specific nature of the training stimulus was assessed in two monkeys. The improvement with training did not extend to other stimulus frequencies and, in one monkey, there was a decrement in performance at a previously trained frequency. The specificity of the training stimulus is discussed in more detail in the next section on perceptual learning.

In addition to demonstrating the effect of training, the results are also consistent with the notion that reorganisation requires an element of salience or importance: of the two food-deprived groups who received auditory stimulation, cortical reorganisation and improved behavioural performance occurred only in the group where the task was relevant i.e., they were motivated to complete the task because they were given a banana-flavoured food pellet reward. Little is known about the mechanism that allows the cortex to reorganise to behaviourally important stimuli while ignoring irrelevant stimuli.

In a more recent study, Kraus *et al.* [1995] have measured plasticity associated with speech discrimination training in adult humans using the mismatch negativity [MMN] response. The MMN is a cortical response to a change in an acoustic stimulus occurring in a repetitive sequence. It is present even when the change is barely perceived behaviourally. Thirteen healthy normal-hearing adult subjects were trained to discriminate between two similar sounding synthetic speech stimuli. Behavioural results indicated a significant improvement in pre- versus post-training discrimination of the two speech stimuli. This was stable one month after the final training session. The behavioural results are shown in Figure 2.2. MMN responses were recorded immediately before the first training session and immediately after the final session and these are summarised in Figure 2.3. The box under the difference wave indicates a statistically significant difference between the responses to standard and deviant stimuli. The region of statistical significance was larger after training. This study demonstrates that experience results in a neurophysiological change in the MMN event-related potential. This suggests that intense training [at least in normal adult subjects] can alter central auditory neurophysiological responses.

The physiological mechanisms responsible for the changes in sensory reorganisation are not fully understood. It is possible that small or rapid changes in sensory reorganisation could be the result of existing [but previously ineffective] connections substituting for deactivated, removed or unused connections. Rapid changes may also result from the release of neuromodulators that alter the effectiveness of pre-existing synapses. On the other hand, many types of reorganisation take time to fully emerge and are long lasting so it is likely that structural changes have taken place. These may involve, for example, synapses being modified in size or there may be local sprouting of axons and dendrites. Despite these changes, there is reason to believe that the basic structural framework is maintained since reorganisation can be reversed if normal activation returns. For example, Merzenich *et al.* [1983] demonstrated that the somatosensory map reverts to the normal tonotopic map after recovery from a nerve crush. This reversibility is consistent with behavioural evidence that normal tactile abilities return after regeneration of crushed nerves in humans.

The functional consequences of sensory reorganisation are not always clear [Willot, 1996b]. There is some evidence that adding more neurones may enhance existing capacity. For example, Recanzone *et al.* [1992] reported that monkeys trained in making a tactile discrimination of the frequency of a vibrating probe on a specific digit improved over time, and this was accompanied by an increase in the size of the cortical representation of that digit. More recently, McDermott *et al.* [1998] have measured frequency discrimination in five adult human subjects having a steeply sloping high-frequency sensorineural hearing impairment. As expected, frequency difference limens [FDL] were elevated compared to normal hearing subjects but, interestingly, four subjects showed a local reduction in FDL near the edge of the cut-off frequency[1]. The authors suggest that these subjects may have undergone cortical reorganisation resulting in an increase in the spatial representation of the cut-off frequencies and resulting in improved frequency discrimination for these frequencies.

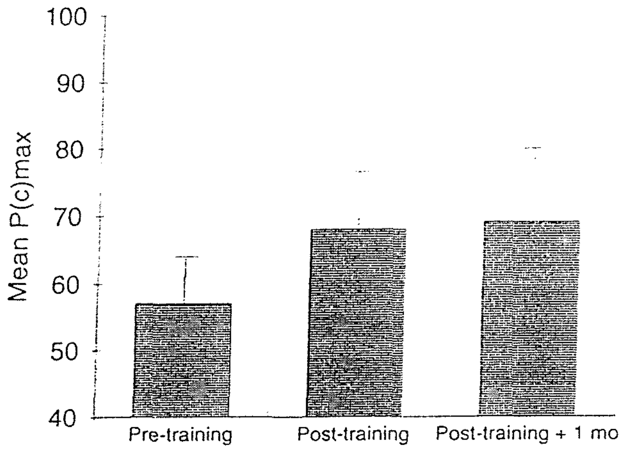


Figure 2.2. Behavioural discrimination scores pre-training, immediately after training, and one month post-training. There was a significant improvement in pre vs. post-training but no significant difference between post-training scores. Reprinted from Kraus *et al.* [1995], with permission.

---

[1] The cut-off frequency was defined as the lowest frequency for which the threshold was greater [worse] than 15 dB HL, and the slope of the hearing impairment [in free field dB SPL] at the high frequencies was 50 dB/oct.

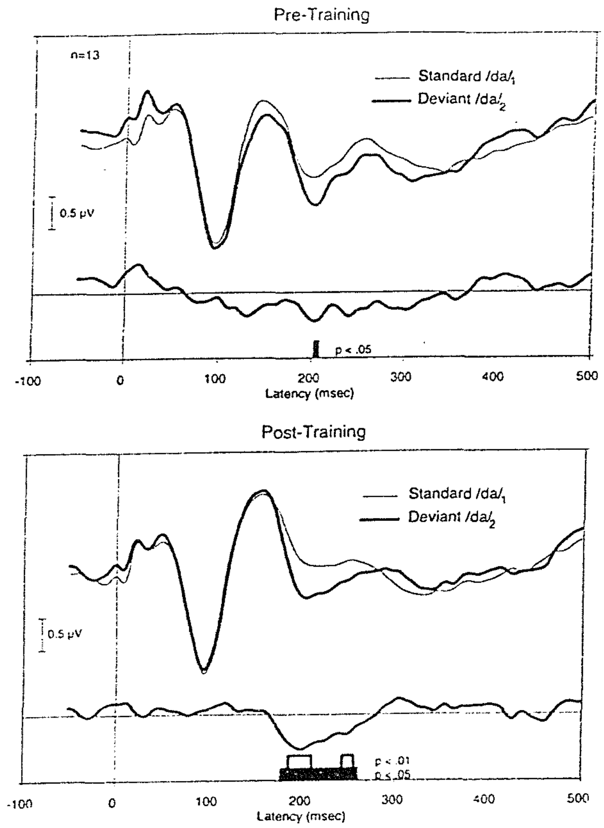


Figure 2.3. Mean mismatch negativity response to standard and deviant stimuli. Pre-training, top graph; Post-training, bottom graph. There is a greater difference between the post-training curves. Reprinted from Kraus *et al.* [1995], with permission.

On the other hand, it is possible that adding new neurones may degrade performance. For example, most amputees have phantom sensations of missing limbs [Melzack, 1990] suggesting that central representations are being activated and continue to cause the perception of the missing limb. Likewise, Jastreboff [1990] has postulated that the phenomenon of tinnitus may be the result of plastic changes to synapses within the central nervous system due to reduced peripheral input.

The auditory system could benefit from increased numbers of neurones responding to sounds that continue to be audible to the hearing impaired subject. If secondary plasticity readily occurs, perhaps as a result of amplification, auditory acclimatisation might occur with the neural map returning to nearer its pre-impaired state. Alternatively, if the reorganisation of the tonotopic map has been beneficial, the provision of a hearing instrument could be contra-indicated. The reorganisation that accompanies a hearing impairment could also cause central problems if high-frequency neurones start to respond to lower frequencies. In addition to any changes that may occur in the cortical representations of the ear that has been fitted with a hearing instrument, changes may occur in the cortical representations of the not-fitted ear since this will effectively receive less stimulation.

In summary, recent experiments have provided evidence that behaviourally important stimuli can cause the CANS of adult humans to reorganise. This may result in improvements in performance if the reorganisation enables the subject to extract previous unused information from the signal. Thus, cortical reorganisation may be directly related to the changes in performance that has been reported to occur after provision of a hearing instrument.

2.2 Perceptual learning

Learning has been defined as the process of acquiring information [Goldstone, 1998]. Robinson and Summerfield [1996] discussed adult auditory learning and training and how these might relate to late-onset auditory deprivation and acclimatisation. They refer to perceptual learning as 'stimulus' learning since the process involves learning specific features of the stimulus, such as newly audible high frequency acoustic cues. They argue that both late-onset auditory deprivation and auditory acclimatisation can be viewed as long-term stimulus learning.

Learning can generally be classified as either procedural, perceptual or task learning. Procedural learning refers to the learning associated with the response demands of the task. For example, Theodoridis and Schoeny [1990] demonstrated the effect of procedural learning on a speech recognition task. The subjects performance improved over time as they gained experience and adapted to the requirements of the task, the experimental setting and other procedural factors. The extent and rate of procedural learning is influenced by 'task learning'. For example, the task of recognising a word is different from a discrimination task that requires the subject to simply know if two words sound the same or different.

Perceptual learning has been defined by Goldstone [1998] as '*a relatively long-lasting change to an organism's perceptual system that improves its ability to respond to its environment'*. This is achieved by extracting information from the stimulus that was previously unused. The premise of this definition is that perceptual learning benefits an organism by tailoring the information gathering process to the organism's use of the information. Goldstone suggests that there are four mechanisms that explain how organisms learn. These mechanisms are known as attention weighting, stimulus imprinting, differentiation and unitisation.

Attention weighting involves the organism increasing the level of attention paid to features that are important and/or decreasing the level of attention to irrelevant features. Categorical perception is one example of attention weighting. While it is recognised that some categorical perception effects are innate [for example, Eimas *et al.* (1971) showed that discrimination of the voiced and voiceless stop consonants /b/ and /p/ by infants was better across the adult phonemic boundary than within the adult phonemic category], a number of studies clearly show that sound categories can be learned; for example, Lively *et al.* [1993] report training procedures that allow Japanese speakers to discriminate between the phonemes /r/ and /l/ that are not present in their native language. In addition, experienced musicians can show a pronounced categorical effect for relative pitch differences that is not found in inexperienced musicians.

The second model to explain perceptual learning is 'imprinting'. The stimulus detecting mechanism is shaped by repeated stimulation and this results in increased speed and accuracy in processing the stimuli. For example, Palmeri *et al.* [1993] have shown that individuals can identify words more accurately when they are spoken by familiar voices. This is consistent with reports that listeners familiar with the abnormal speech of deaf children can interpret this more accurately than unfamiliar listeners [Most *et al.,* 1997; McGarr, 1983]. In a review of the literature, Weinberger [1993] identified evidence that cells within the auditory cortex become 'tuned' to the frequency of often-repeated tones. This is consistent with the perceptual learning of specific features that occur within a stimulus. In addition to imprinting particular stimuli, there is some evidence that dimensions of a stimulus have to be learned. Goldstone [1998] reports that loudness perception is disorganised in babies and this dimension has to be learnt and organised over time.

The third model to explain perceptual learning is for stimuli to become increasingly differentiated from each other. Once separated, discriminations can be made between stimuli that were originally indistinguishable. In many cases, exposure to the stimuli may be all that is necessary to promote differentiation. An improvement in performance over time, after provision of a hearing instrument, could be interpreted as perceptual learning that has taken place without providing formal training. However, learning to differentiate between stimuli is typically accelerated by training. Discrimination training is often highly specific to the task. In recent years there has been considerable interest in the specificity of some forms of visual learning. For example, Karni and Sagi [1991] demonstrated improvements in the discrimination of visual texture that was specific to a particular eye, to a particular location in the visual field, and to a particular orientation of the stimulus. Learning did not transfer to other target orientations, to other locations in the visual field, or to the other eye.

A limited number of studies have been published that investigate the specific nature of auditory training tasks. The study by Racanzone *et al.* [1993], discussed in the previous section, demonstrated the specific nature of the training stimulus used in a frequency discrimination task on owl monkeys. Two studies have investigated the specificity of frequency discrimination training in humans. The first of these was Demany [1985] who showed that improvements in training to a 0.2 kHz stimulus extended to higher frequencies although training at 6 kHz did not extend to lower frequencies. In the second study, Irvine *et al.* [2000] showed that the improvement at the trained frequency was greater than that observed at other frequencies. In 1997, Wright *et al.* studied the specificity of temporal discrimination training. The task was to discriminate between a standard gap of 100 ms and a longer interval. They observed an improvement at the standard interval but this did not extend to gaps of 50, 200 or 500 ms [although the improvement extended to other frequencies]. In a more recent study, Wright and Fitzgerald [2001] investigated changes in interaural level differences [ILDs] and interaural time differences [ITDs]. ILD cues are analysed in the lateral superior olive and ITD cues are analysed in the medial superior olive. Integration of these localisation cues start at the level of the inferior colliculus. The stimuli, presented over headphones, for the ILD and ITD were tones at 4 kHz and 0.5 kHz respectively. An improvement in performance with training only occurred for the ILD cues suggesting that the training had its effect at a relatively low neural level before integration. The improvement also occurred if the ILD was altered from 0 to 6 dB but not if the frequency was changed to 0.5 or 6 kHz. This implies that training occurred at a level where frequency is analysed separately. These studies on auditory training tasks show that there is some degree of specificity to the training stimulus. This suggests that any learning associated with the provision of a monaural hearing instrument may be specific to that ear and may not transfer to the opposite ear. It is possible that, when hearing instruments are fitted binaurally, effects will be observed under binaural testing but not present on monaural testing. Contradictory findings come from studies that show that perceptual training can transfer across sensory modalities. Children with auditory attention deficits tend to also show deficits in visual selective attention tasks [Quittner *et al.*, 1994]. A possible explanation for this apparent contradiction is given by Sagi and Tanne [1994] and by Sireteanu and Rettenbach [1995] who suggest that learning involving a change to early perceptual processes [i.e., at level of brainstem] will be less generalisable than learning that takes place at a higher level within the CNS.

The final mechanism that may be responsible for perceptual learning is 'unitisation'. This involves the construction of a single functional unit that can be triggered when a complex configuration arises. As a result, a task that originally required detection of several parts can be accomplished by detecting a single unit. Whereas differentiation divides whole structures into clearly separated component parts, unitisation integrates these component parts into a single whole structure. Many researchers have argued that whole words are perceived as single units because of the individual's life long experience with them. They argue that recognition occurs at levels higher in the CNS than the individual phonemes. Although unitisation appears to be at odds with differentiation, Goldstone argues that perceptual learning is a result of a combination of mechanisms that collectively serve to differentiate stimuli.

In summary, perceptual learning improves the organism's ability to respond to its environment. A number of mechanisms have been proposed to explain this learning process. Robinson and Summerfield [1996] refer to perceptual learning as stimulus learning and argue that both late-onset auditory deprivation and acclimatisation can be viewed as long-term stimulus learning.

2.3 Late-onset auditory deprivation

Silman *et al.* [1984] were the first to report the phenomenon of late-onset auditory deprivation. They undertook a retrospective study on two groups of subjects who had a mild-to-moderate bilateral sensorineural hearing impairment and had been fitted with a hearing instrument. One group of subjects consisted of 44 male adults [mean age 59 years] who had been fitted monaurally. The second group consisted of 23 male adults [mean age 57.9 years] who had been fitted binaurally. The subjects performed a speech recognition test before hearing instrument fitting and again 4-5 years after fitting. The test material consisted of Central Institute for the Deaf [CID] W-22 phonetically balanced [PB] words presented over headphones at 40 dB *re*: speech reception threshold [SRT]. SRT is defined as the lowest level that the subject is able to score 50% correct. Since the SRT was around 40-50 dB HL, the presentation level of speech was approximately 80-90 dB HL. The results are summarised in Figure 2.4. The filled columns correspond to the mean pre-fitting performance score and the open columns correspond to the mean post-fitting score. Only the results from the right ear of the binaurally aided subjects are shown but similar results were obtained for the left ear. There was a mean decrease of 18.2% in the speech recognition score [SRS] for the not-fitted ear of the monaurally fitted subjects [$p<0.01$] but no significant difference [$p>0.05$] in the fitted ear, or in either ear of the binaural users.

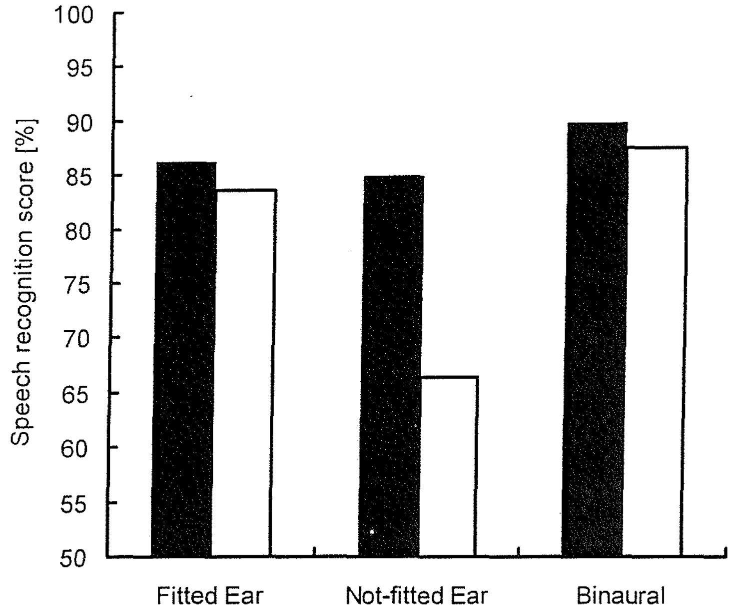


Figure 2.4. Mean speech recognition scores in percent correct for CID W-22 word lists in individuals fitted monaurally and binaurally. The black columns correspond to the pre-fitting score and the open columns correspond to the post-fitting score. Only the mean scores from the right ear of the binaurally aided subjects are shown but similar results were obtained for the left ear. Adapted from Silman *et al.* [1984].

Using a similar methodology, Gelfand *et al.* [1987] repeated the above study but also included a control group of hearing-impaired subjects who had not been fitted with a hearing instrument. The monaural group consisted of 48 subjects, the binaural group 19 subjects and the not-fitted control group 19 subjects. The mean age of the subjects at the time of fitting was 52 years, and the duration between the pre and post-fitting assessment was typically 7-8 years. The findings are summarised in Figure 2.5. Once again, the filled columns correspond to the initial speech performance score and the open columns correspond to the post-fitting score. The 7.2% reduction in the SRS of the not-fitted ear of the monaural group was statistically significant [p=0.01]. The scores for the other subjects decreased by around 3-4% but this difference was not statistically significant [p>0.05]. Despite similar methodologies, the magnitude of the deprivation effect in the not-fitted ear of the monaural subjects differs markedly between the two studies [18.2% and 7.2% respectively]. Gelfand *et al.* suggested that this might reflect sampling differences between

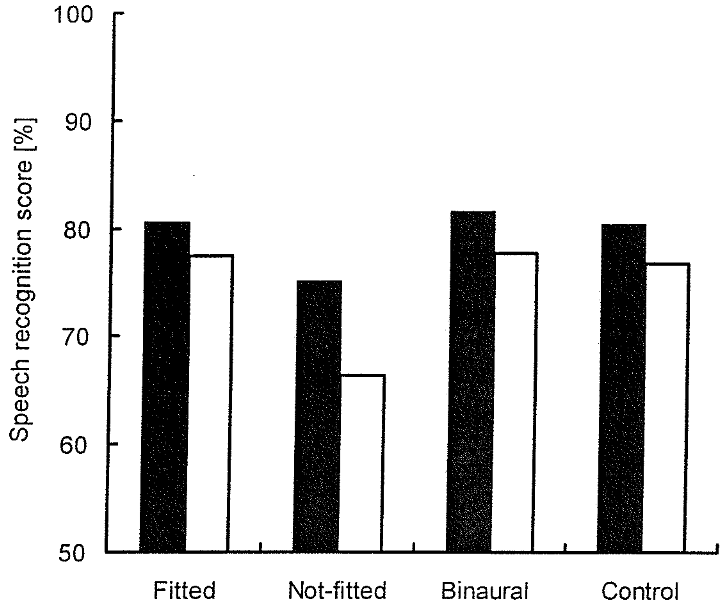


Figure 2.5. Mean speech recognition scores in percent correct on CiD W-22 word lists in individuals fitted monaurally and binaurally as well as control subjects. The filled columns correspond to the pre-fitting score and the open columns correspond to the post-fitting score. Adapted from Gelfand *et al.* [1987].

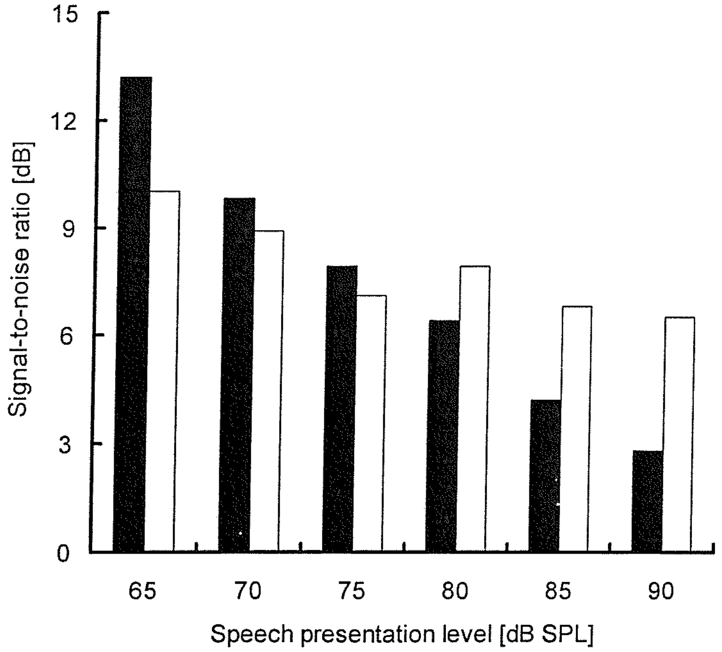


Figure 2.6. Mean signal-to-noise ratio (dB) required to achieve 50% performance on the Four Alternative Auditory Feature test as a function of presentation level in individuals fitted monaurally. All testing was administered monaurally over headphones in noise using an adaptive strategy. The filled columns correspond to the normally aided ear while the open columns correspond to the normally unaided ear. Adapted from Gatehouse [1989].

the groups; in particular, the not-fitted ear in the monaural subjects in their study had a wider range of pre-fitting speech recognition scores than Silman *et al.* [mean score (± 1 SD) was 75.7% (±16.1) and 84.9% (±11.8) respectively]. An alternative explanation is that there may have been differences in gain provided by the hearing instrument in each study; the asymmetry between ears will be larger with higher levels of gain and this may influence the magnitude of the deprivation effect.

Gatehouse [1989] suggested that the previous findings might be related to the presentation level of the test material; an ear that is use to receiving a high level of stimulation will acclimatise to the pattern of cues presented and be most effective at analysing at high presentation levels. This hypothesis was tested in a group of hearing-impaired individuals by measuring the signal-to-noise ratio [SNR] required to achieve 50% performance on the Four Alternative Auditory Feature [FAAF] test as a function of speech presentation level. Subjects comprised 24 monaurally fitted adults [mean age 59.3 years] with a mild-to-moderate bilateral sensorineural hearing impairment. All testing was administered monaurally over headphones. Speech recognition scores were not available for the subjects at the time of fitting. However, the subjects displayed symmetrical pure tone thresholds so there was no reason to suspect any systematic asymmetry between the ears. The subjects were tested at a mean duration of 4.8 years post-fitting and reported wearing their hearing instruments, on average, for 8.6 hours per day. A summary of the results is given in Figure 2.6. The filled columns correspond to the fitted ear and the open columns correspond to the not-fitted ear. The results show the expected improvement in performance for both ears as the presentation level of speech increases [a less favourable SNR was required at 90 dB SPL than at 60 dB SPL]. More importantly, the results reveal a significantly higher SNR in the fitted ear at low presentation levels [3.2 dB] and in the not-fitted ear at high presentation levels [3.0 dB]. These differences are substantial and represent an equivalent of around 20% difference in score since mean scores change at a rate of 6% per decibel change in SNR over the range 40-80% [Foster and Haggard, 1987]. The results at the higher presentation levels replicate the findings of Silman *et al.* and Gelfand *et al.*; the fitted ear performs more efficiently than the not-fitted ear at high presentation levels, while the opposite is true at lower presentation levels. Gatehouse interpreted this finding as evidence that the ear performs most efficiently at a presentation level that is assumed to represent its normal listening level for speech. This may partly explain the decreased performance in the not-fitted ear reported in the studies discussed earlier; the not-fitted ear will be exposed to lower speech levels after monaural aiding because speakers [including television and radio] will not have to raise their voice to pre-aiding levels in order to become audible [in the fitted ear]. This study was carried out using headphones and not with the subjects own hearing instrument [or with headphones using a frequency response shaped to match the hearing instrument response]. In addition, it would be helpful to undertake a prospective study to confirm that these differences occur over time as a result of aiding.

Silman *et al.* [1993] undertook a prospective study investigating changes in performance over time in adults with a mild-to-moderate symmetrical bilateral sensorineural hearing impairment. The study involved subjects fitted binaurally [n=28] and monaurally [n=19]. There was also a control group of hearing-impaired subjects [n=19] who were not hearing instrument users. Speech testing using CID PB wordlists, the CUNY Nonsense Syllable Test [NST] and the Speech in Noise [SPIN] test, was undertaken at 6-12 weeks post-fitting and again at one-year post-fitting. The CUNY NST test consists of seven subgroups, each of which contains seven to nine nonsense syllables either consonant-vowel or vowel-consonant format. The SPIN test words are 25 monosyllables that are presented in 50 sentences with half having high contextual information [Kalikow *et al.*, 1977; Bilger *et al.*, 1984]. Once again, testing was carried out at 40 dB referenced to the SRT. The mean SRTs in the fitted and not-fitted ear of the monaural group were 33.4 and 28.4 dB HL respectively [hence speech was typically presented at 73.4 and 68.4 dB HL]. Figure 2.7 shows the outcome of the three speech tests in the monaurally aided group. The filled columns correspond to the mean score obtained at 6-12 weeks post-fitting and the open columns correspond to one-year post-fitting. For all tests, the initial performance scores are slightly poorer in the fitted ear but this is probably a reflection of the slightly poorer air-conduction pure tone thresholds and SRTs [around 5 dB at the start of the study]. The results for CID words are shown in Figure 2.7a. At the start of the study, the mean performance in the fitted ear was 4.63% less than the not-fitted ear. However, by the end of the study the mean performance was 4.53% higher in the fitted ear. This difference was statistically significant on a paired *t*-test [$p<0.05$]. A similar finding was observed with the nonsense syllable test [Figure 2.7b]. At the beginning of the study, mean performance was 9.84% lower in the fitted ear but this was 6.3% higher at the end of the study. This difference was also statistically significant on a paired *t*-test [$p<0.05$]. Figure 2.7c shows the mean difference in SNR for the SPIN test. The fitted ear required a mean SNR that was more favourable than the not-fitted ear by 5.94 dB at the start of the study. By the end of the study the difference was only 2.82 dB although the change over time was not statistically significant [$p>0.05$]. There was no statistically significant difference over time for the binaurally aided group or the control group [$p>0.05$].

Observation of Figure 2.7 reveals that the changes over time are more apparent on the NST, which has less linguistic redundancy compared to the CID word lists. In addition, the mean decrement in the not-fitted ear appears to be greater than the mean improvement in the fitted ear [approximately 10% and 6% respectively for NST]. The authors suggest that it is likely that more time is required for a significant acclimatisation effect to emerge in the fitted ear[s] of both the monaurally and binaurally fitted subjects. However, there are at least two alternative explanations. Firstly, the presentation level of speech to the fitted ear at the initial and retest session was around 73.4 and 75.9 dB HL respectively. The corresponding levels in the not-fitted ear were 68.4 and 68.9 dB HL. During routine hearing instrument use, it is likely that the sound pressure level [SPL] of amplified speech would be somewhat higher than these presentation levels. Therefore, the fitted ear may not have been tested at its familiar level where, presumably, it performs at maximum efficiency. Secondly, it is possible that acclimatisation may have commenced before data collection at the initial 6-12 week test session. However, the study did show a trend towards an improvement in the fitted ear that was not apparent in either of the two previous studies where the post-fitting duration was substantially longer. The decrease in performance in the not-fitted ear is consistent with the findings in the earlier studies;

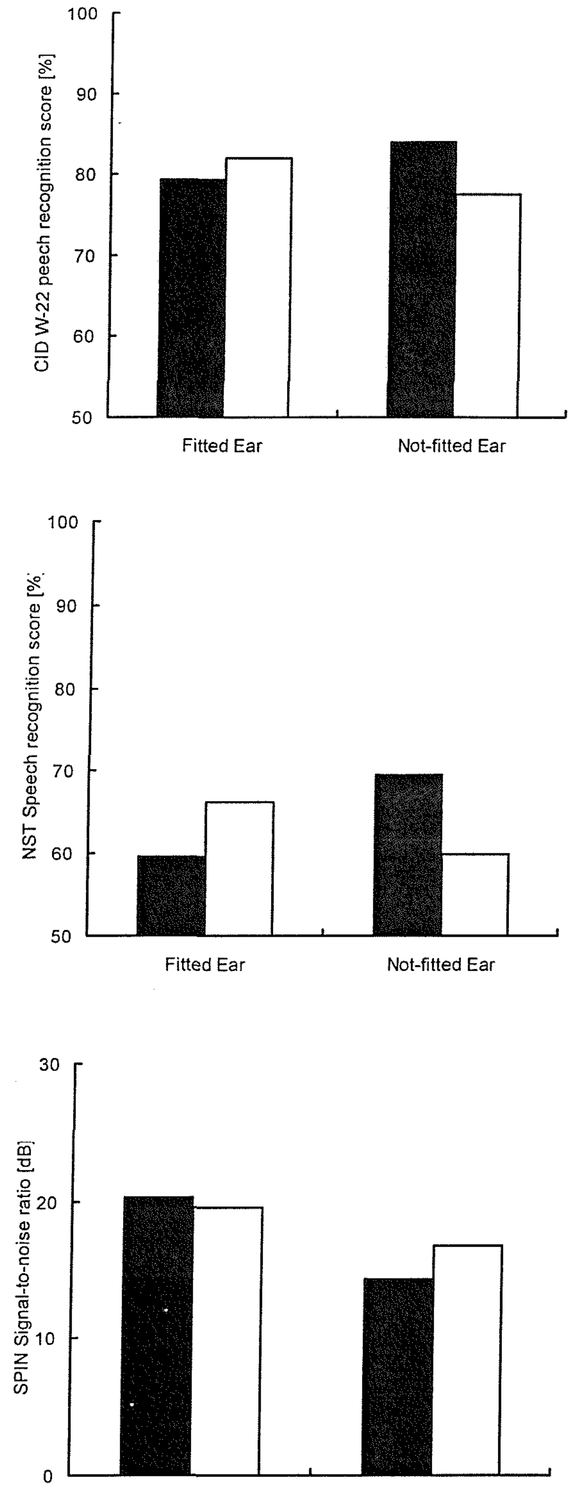


Figure 2.7. Mean performance on speech tests in individuals fitted monaurally. Black columns correspond to initial performance and open columns to retest performance. CID W-22 speech recognition score, top graph; NST speech recognition score, middle graph; Signal-to-noise ratio for 50% performance on SPIN, bottom graph. Adapted from Silman *et al.* [1993].

performance decreases in the not-fitted ear as a result of receiving relatively less useful stimulation in everyday life compared to the fitted ear.

A limitation of the initial studies by Silman *et al.* and Gelfand *et al.* was that subjects were all male with a presumed noise-induced hearing impairment. In numerous follow-up studies by other investigators, the reported effects have been replicated in populations consisting of men and women, adults and children, and sensorineural and conductive hearing impairment. Many of these studies were published in a special issue of the Journal of the American Academy of Audiology in 1993 [Volume 4, number 5] and also reviewed by Neuman [1996]. These studies provide convincing evidence of the late-onset auditory deprivation effect for both mean data and in a substantial number of individual subjects. In addition, a number of case studies demonstrate recovery from deprivation in some individuals with use of amplification in the previously not-fitted ear. However, a large-scale prospective study reporting incidence, magnitude and time course of the deprivation effect has yet to be published.

2.4 Auditory acclimatisation

2.4.1 Studies showing an improvement over time

As early as 1940, it was noted that the aided speech recognition scores of some individuals improved over time [Watson and Knudson, 1940]. During the past decade there has been considerable interest in improvements in hearing instrument benefit over time. A review article by Turner *et al.* [1996] summarises articles that were available in the mid-1990s for acoustical hearing instruments. These articles were also discussed in an article by Palmer *et al.* [1998] on functional and physiological changes in the adult auditory system.

The studies investigating changes over time are summarised in Table 2.2. Of the 16 studies listed in this table, eight report a change over time. The first recent study that reported improvements in hearing instrument benefit over time was by Cox and Alexander [1992]. This study used both performance and self-report outcome measures to assess

Table 2.2. Studies on auditory acclimatisation in adults [arranged chronologically] using acoustical hearing instruments. Evidence of acclimatisation indicated as 'Yes' or 'No' in performance and self-report outcome column.

Study	Performance	Self-report	Comments
Cox & Alexander [1992]	Yes [CST]	Yes [PHAB]	n=10,no controls, duration 10 weeks, benefit scores but no aided or unaided scores, subjects adjusted gain.
Gatehouse [1992a]	Yes [FAAF]		n=4, control [non-test ear], duration 12 weeks, benefit scores and aided and unaided scores.
Bentler *et al.* [1993a,b]	No [SPIN, NST]	No [HPI]	n=39 [new & experienced users], control [n=26 experienced users], duration 12 months, aided scores only, subjects altered gain, wide range of hearing level, hearing aids not always worn, poor match with target.
Gatehouse [1993]	Yes [FAAF, SVT]		n=36, duration 16 weeks, subjects adjusted gain, NHS compared with new prescription, Aided scores only.
Taylor [1993]	No [NU-6]	No [HHIE]	n=58, no controls, duration 12 months, aided scores only, mild impairment, low gain setting, gain control not fixed.
Humes *et al.* [1996]	No [NST, HINT]	No [HAPI, HHIE]	n=10, Controls [n=10 experienced users], Duration 12 months, Binaural fittings but tested monaurally.
Ovegard *et al.* [1997]		Yes [quality judgements]	n=10, Improvements in some sound qualities.

[CONT'D OVERLEAF]

Cox *et al.* [1996]	Yes [CST, SPAC]		n=22, Control [n=5 experienced users], Duration 12 weeks, Gain held constant Benefit increased in subgroup but this was due to poorer performance unaided.
Horwitz and Turner [1997]	Yes [NST]	Yes [PHAB]	n=13 Control [n=13 experienced users], Fixed and user-adjusted gain settings, Duration 18 weeks.
Neuman *et al.* [1997]	No [NST, HINT]		n=7, Limited information available from poster presentation.
Saunders & Cienkowski [1997]	No [spondee, HINT]		n=48 [new & experienced users], Duration 3 months but only scores from 1 month included in analysis, 6 different hearing aid configurations.
Surr *et al.* [1998]	No [CST]	No [PHAB]	n=15 Experienced users fitted with WDRC, Compared 6 weeks with 2 years.
Arlinger *et al.* [1999]	Yes [sentence material]	No [PHAB, quality judgements]	n=23, Experienced users fitted with digital aid, Compared 1 month with 1 year,
Bentler *et al.* [1999]	No [NST]		n=26, Controls included a group of new subjects and a group of experienced hearing aid users, Limited information available concerning methodology.
Kuk [2000]	Yes [SPIN]		n=20, No controls Experienced users refitted with WDRC
Humes *et al.* [In press]	Yes [CST], No [NST]	No HAPI, HHIE	n= 88 new users fitted binaurally & tested binaurally

benefit. The performance test material was the 150-item [six passages each with 25 key words] Connected Sentence Test [CST]. The CST provides an estimate of speech communication ability in daily life situations [Cox *et al.*, 1988]. This was carried out for four simulated listening conditions that were designed to mimic real-life environments such as a living room or cocktail party [Cox *et al.*, 1991a]. Benefit was defined as the difference between the aided and unaided scores. The percent correct scores were transformed into rationalised arcsine units [RAU; Studebaker, 1985] in order to minimise the relationship between performance and variability that is seen with percentage scoring. [Within the range 12 to 88, raus correspond closely to percentages.] The initial benefit score was measured immediately after fitting and before the hearing instrument was used routinely by the subject. The final benefit score was the difference in the aided score at ten weeks and the unaided score at nine weeks post-fitting. Twelve subjects [mean age 67 years] completed the performance tests; eight were new users and four were experienced users. A graph of the mean audiogram shows hearing thresholds sloping from around 30 dB HL at 0.5 kHz to around 75 dB HL at 4 kHz. All subjects were fitted with new hearing instruments; nine [75%] were monaural and three [25%] were binaural. The hearing instrument prescriptions used were the Memphis State University [MSU] Hearing Instruments Prescription Procedure [Cox, 1988] and the Revised National Acoustics Laboratory [NAL-R] prescription approach [Byrne and Dillon, 1986]. The MSU 2-cc coupler gain target approximately matches the NAL-R target at 4 kHz but provides less gain at lower frequencies. The frequency response of the hearing instrument was reported to provide a 'reasonable' match to the prescribed values although no data were provided. It was not reported if the four existing users were already experienced and familiar with the newly prescribed frequency response. During the course of the study, the experienced and new users reported around eight and four hours of daily use respectively.

The mean benefit on the performance test at the start of the study was significantly greater [$p<0.05$] for the experienced users [8.8 rau] compared with the new users [1.9 rau]. These results were pooled because there was no significant interaction with other variables. The results [see Figure 2.8] show that benefit increased during the first ten weeks of use for some environments. The clearest improvement was seen for environment A. This environment is typical of the living-room type environment in which the subjects would have most listening experience. This was the only environment where the change in benefit [ca 5.5%] was statistically significant [$p<0.05$]. It was not reported if this increase in benefit was due to a decrease in the unaided score or an increase in the aided score. No control group was included as a check for practice effects although this would seem unlikely given that the increase in benefit did not occur in all conditions.

The self-report outcome material was the 48-item Profile of Hearing Aid Benefit [PHAB] questionnaire, which assesses performance in a variety of everyday situations [Cox *et al.*, 1991b]. Ten of the original 12 subjects [7 old and 3 new] completed the PHAB at both two and ten weeks post-fitting. There was a significant difference [$p<0.05$] in the overall mean benefit score at two and ten weeks post-fitting [22.9 and 28.9% respectively]. The initial and long-term benefit scores for the five sub-scales of the PHAB are shown in Figure 2.9. The largest increase in mean benefit occurred on the familiar talkers and the ease of communication sub-scales [ca 8 and 12% respectively]. These correspond most closely with the living-room type environment. However, the interaction between subscales and measurement time was not statistically significant, indicating that there was no differential change between subscales over time.

The trend of increasing self-reported benefit over time may be due to acclimatisation. However, there are two more prosaic explanations. Firstly, subjects may initially underestimate the problems they experience unaided and this may result in a relatively small initial benefit score. After a period of regular hearing instrument use the subject's perception of unaided difficulties may increase and this will result in a larger benefit score. It is not clear if subjects were allowed to reassess their unaided responses after a period of hearing instrument use [Cox, personal communication]. Secondly, it is possible that the self-reported disability may reduce over time when aided because subjects become more adept at adjusting the gain control to their preferred setting in different listening environments. The combination of new and experienced users, monaural and binaural fittings, and two different prescriptions may be a weakness because it is not known if auditory acclimatisation will generalise across these conditions. However, this study was the first in recent times to suggest the existence of auditory acclimatisation.

An elaborate experiment that overcame some of the limitations of the Cox and Alexander study [for example, lack of control condition] was published by Gatehouse [1992a]. Changes in aided and unaided data were obtained on ten occasions over a period of 12 weeks commencing from the time of the hearing instrument fitting. The measurements were limited to a single 80-item run of the Four Alternative Auditory Feature [FAAF] test described by Foster and Haggard [1979, 1987]. This is a closed set word discrimination in noise task with no visual cues. Discrimination of the FAAF words is strongly dependent on the audibility of high frequency speech sounds.

Four new hearing instrument users [age range 55-70 years] took part in the study. They were typical of first-time hearing instrument users in the UK with pure tone thresholds around 30 dB HL at 0.5 kHz and 65 dB HL at 4 kHz. Each subject was fitted monaurally with a linear hearing instrument. The prescription method was not reported but examination of the raw data shows that the hearing instruments were very effective at providing high frequency amplification: compared to the NAL-R prescription formula, the real ear insertion gain [REIG] was approximately 10 and 5 dB greater than the target at 2 and 4 kHz respectively. There was no systematic change in the user-gain setting after five weeks post-fitting. The amount of daily use made of the hearing instrument was not reported so it is difficult to know just how familiar the subjects were with amplified speech. The subjects were tested in the sound field and also under a variety of headphone conditions, one of which simulated the electroacoustic characteristics of the hearing instrument response. Using headphones and a variety of frequency response conditions reduced the possibility of the subject being motivated to provide a better score with the condition that simulated the hearing instrument. The not-fitted ear was used to control for improvements due to practice effects.

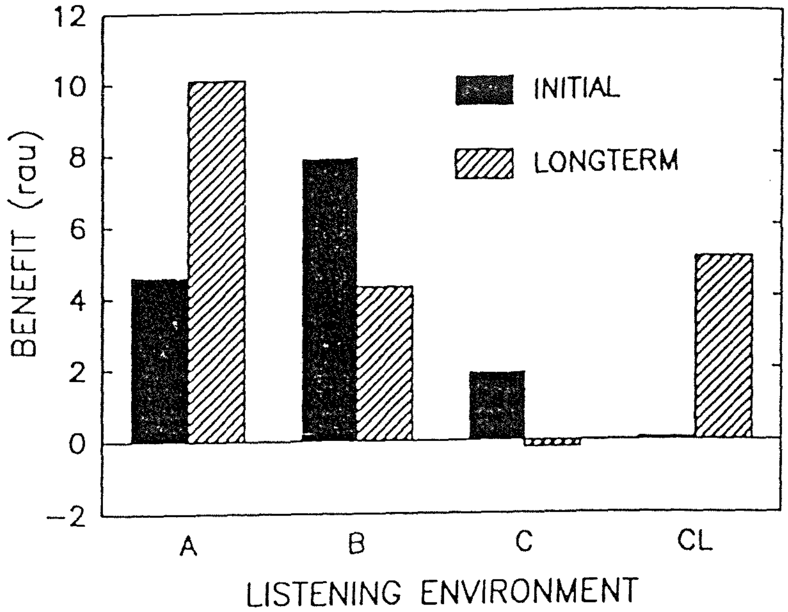


Figure 2.8. Mean initial and long-term benefit measured on the Continuous Sentence Test for four simulated listening environments. Listening environment A) 55 dBA, SNR +5 dB; B) 63 dBA, SNR + 8 dB, overall reverberation time 1s; C) 64 dBA, SNR+2 dB; CL) as for C. All test conditions included visual cues except CL. Reprinted from Cox and Alexander [1992], with permission.

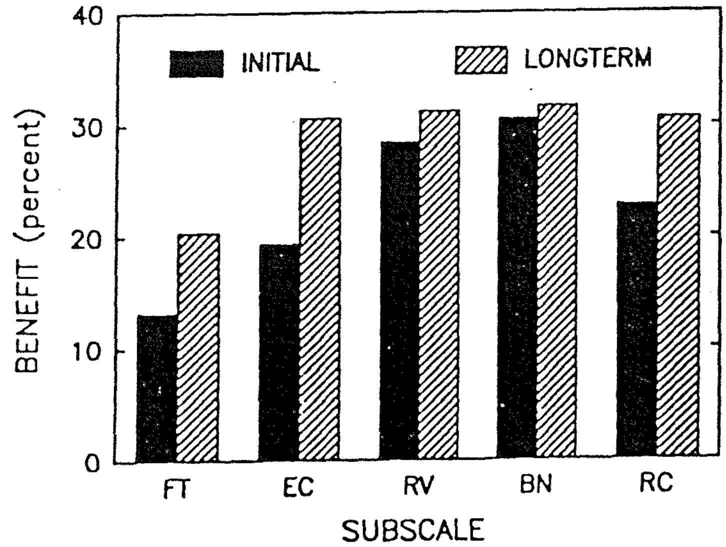


Figure 2.9. Mean initial and long-term benefit for each of the Profile of Hearing Aid Benefit speech communication subscales. Listening to familiar talkers in quiet environment, FT; easy listening EC; reverberation, RV; background noise, BN; reduced visual cues, RC. Reprinted from Cox and Alexander [1992], with permission.

The change in benefit as a function of post-fitting time for both the sound field and the headphone condition that attempted to simulate the normal sound field condition are shown in Figure 2.10. These two graphs show essentially similar findings. The benefit score for the control ear remains relatively stable at around 5% while the fitted ear increases from 5% to greater than 15% when the study terminated at 12 weeks post-fitting. This increase commenced from around 4-6 weeks post-fitting. The increase over time was more marked for the headphone condition that showed the lower initial benefit; this may be related to the fact that headphone performance was monaural while the sound field presentation was binaural [the not-fitted ear contributing to the unaided performance]. A repeated-measures ANOVA was performed on the data from each subject separately. While there was no significant change over time in the not-fitted control ear of any of the subjects, three [75%] subjects showed a statistically significant increase over time in the fitted ear. The lack of increase in the control ear means that this finding is not due to a practice effect with repeated exposure to the test material.

Figure 2.11 shows the aided and unaided data for the fitted ear with headphone presentation. The improvement in benefit is due to a combination of an increase in the aided score and a corresponding decrease in the unaided score. The aided and unaided scores were analysed separately in a one-factor repeated-measures analysis of variance [ANOVA] to determine if performance changed significantly over time. The difference over time was significant [$p<0.01$] for both the aided and unaided scores. While the change over time was discussed in the context of the contribution to the overall change in benefit, the implications for dependency on the hearing instrument were not discussed.

Specifically, it would appear that the subject would be in a worse position than before hearing instrument fitting when not wearing the hearing instrument. This decrement in the fitted ear is in addition to any late-onset auditory deprivation that may occur in the not-fitted control ear. This finding has been observed from the raw data of other studies [e.g., Cox *et al.*, 1996] and is discussed later in this section.

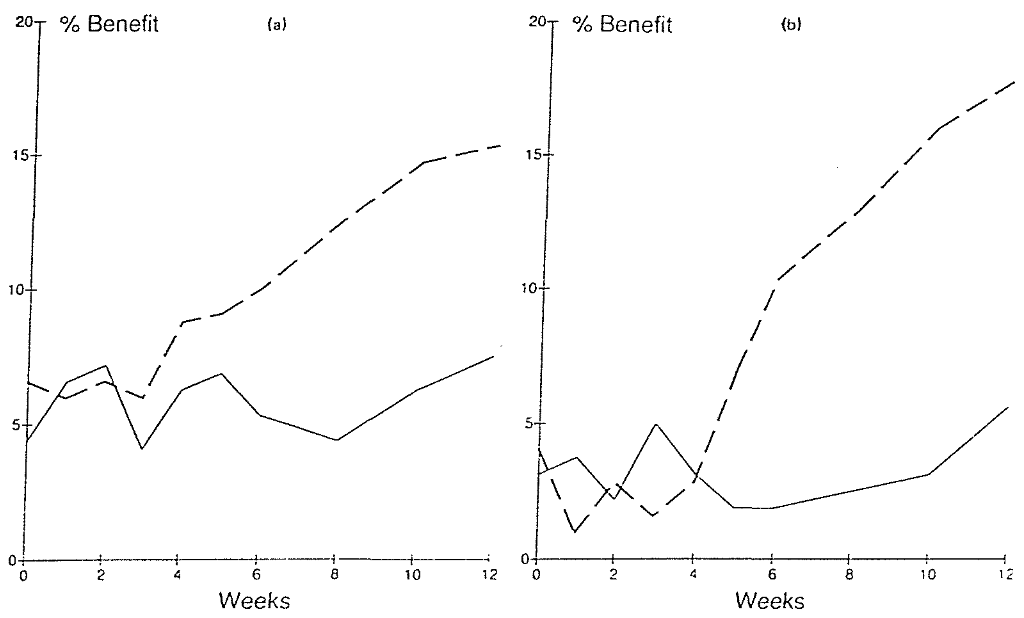


Figure 2.10. Mean benefit for Four Alternative Auditory Feature test [aided minus unaided score] as a function of post-fitting time. Sound field presentation is shown in panel [a] and headphone presentation is shown in panel [b]. The broken line is the fitted ear and the solid line is the not-fitted ear. Reproduced from Gatehouse [1992], with permission.

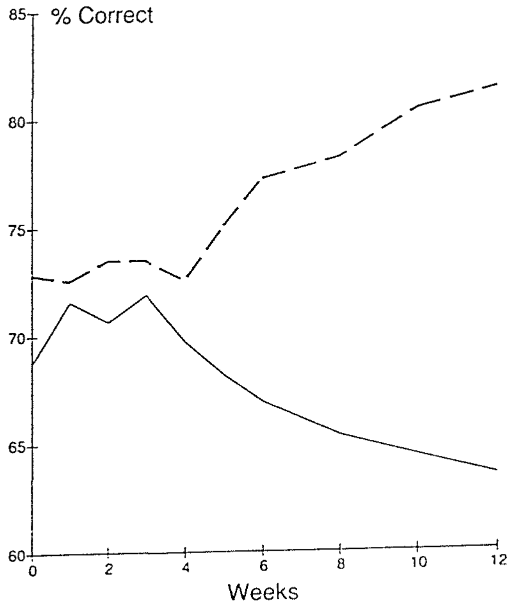


Figure 2.11. Mean recognition score in percent correct for Four Alternative Auditory Feature test with headphone presentation as a function of post-fitting time. The broken line corresponds to the aided score and the solid line corresponds to the unaided score. From Gatehouse [1992], with permission.

A further finding of the Gatehouse study was the difference in score obtained with flat and tailored [i.e., additional high frequency] amplification with headphone presentation. This is shown in Figure 2.12 and reveals that the advantages of high frequency amplification were not apparent until at least six weeks post-fitting. The aided scores for both the flat and the tailored amplification were subjected to a separate one-factor repeated-measures ANOVA . The difference over time was significant [$p<0.01$] for the tailored but not the flat amplification.

Gatehouse published another article in 1993 that compared two hearing instrument frequency responses over a period of 16 weeks. Technically speaking, this later study did not measure benefit because only aided scores were reported. Measurements were carried out at week 0, 8 and 16 using the FAAF test as well as the more recently developed Sentence Verification Test [SVT]. The SVT is a speech in noise test but, in addition to scoring word identification, the response time for the subject to verify that the sentence is 'silly' or 'sensible' is also recorded [Gatehouse, 1992b]. The study involved 36 monaurally fitted subjects with mean age of 64 years and mean pure tone thresholds sloping from 31 dB HL at 0.5 kHz to 57 dB HL at 4 kHz. The subjects had been previously fitted with a standard UK NHS linear hearing instrument for 12-15 months and it was assumed that auditory acclimatisation would be essentially complete by this time. This fitting failed to match the NAL-R targets for REIG by 11, 17 and 20 dB at 2, 3 and 4 kHz respectively. The subjects were re-fitted with a new hearing instrument that was within 3 dB of the NAL-R target at 0.25-2 kHz and within 5 dB at 3 and 4 kHz. In effect, the study was comparing differences in high frequency amplification. It is not clear how the gain of the hearing instruments were set on each test session although it is likely that the UK NHS fitting was tested at the user-gain recorded on week 0 and the NAL-R was tested at the fixed target-gain setting. There was no difference in the patterns of daily use between the different hearing instruments as assessed by simple self-report although the actual duration of use was not reported.

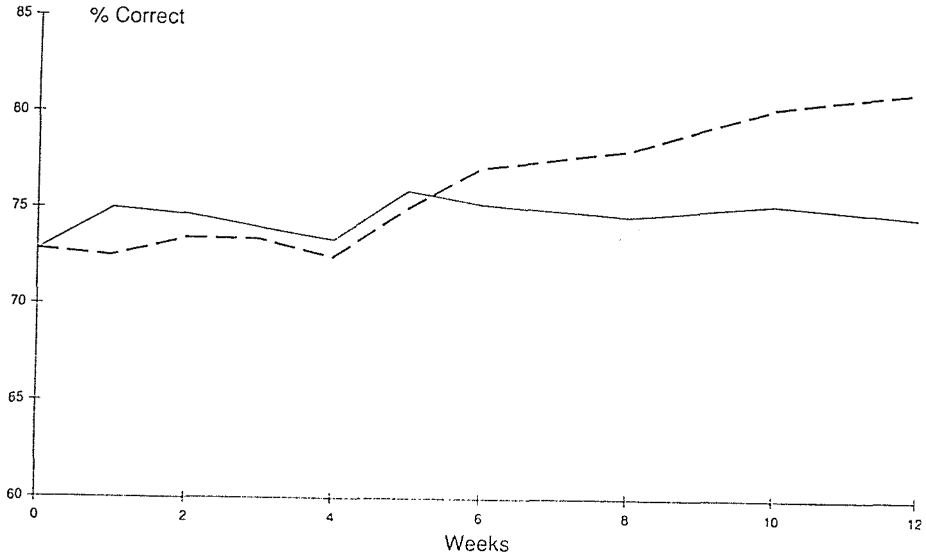


Figure 2.12. Mean recognition score in percent correct for Four Alternative Auditory Feature test as a function of post-fitting time for two simulated frequency responses. The broken line corresponds to tailored high frequency amplification and the solid line corresponds to flat amplification. Reproduced from Gatehouse [1992], with permission.

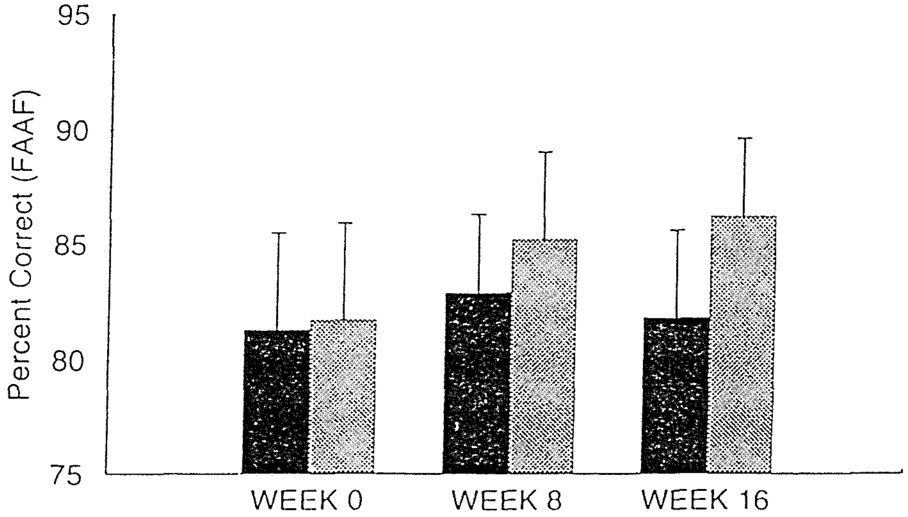


Figure 2.13. Mean aided recognition score in percent correct for Four Alternative Auditory Feature test as a function of NAL-R post-fitting time. The black columns correspond to the previous NHS hearing aid and the grey columns correspond to the new NAL-R fitting. The bars show 2 SEM. Reproduced from Gatehouse [1993], with permission.

The mean aided FAAF scores are shown in Figure 2.13. There was no difference in the aided scores at week zero. However, by week 8 the difference was 2.3% and by week 16 it was 4.4%. A similar finding occurred on the verification component of the SVT. All three of the performance indices showed statistically significant advantages for the new hearing instrument. This finding is consistent with the previous study by Gatehouse in demonstrating that differences in frequency response are measurable in the laboratory, but only after time has been allowed for acclimatisation to occur. While the most likely explanation for the improved performance is the additional amplification in the 2-4 kHz region, there may have been other differences between the fittings, such as sound quality or differences in subject motivation to perform better with the newer hearing instrument.

Cox *et al.* [1996] investigated acclimatisation using two performance measures of benefit. The first was the CST that was used in the earlier study of auditory acclimatisation by Cox and Alexander [1992]. The second measure was the 48-item Speech Pattern Contrast [SPAC] test reported by Boothroyd in 1985. This is designed to score performance in terms of eight different speech features [e.g., vowel height and place]. Both of these tests were performed as audio-visual tasks under conditions that mimic everyday social events. This environment was used because the main complaint of hearing instrument wearers is usually related to understanding speech in noisy environments. This was also one of the environments used by Cox and Alexander [1992] when they obtained a borderline significant improvement [p=0.07] in benefit over a 10 week post-fitting period.

The experimental group comprised 22 new hearing instrument users with a mean age of 72 years. Although six individuals had tried a hearing instrument in the past, this had resulted in limited success and none were current users. The hearing thresholds of each individual were plotted on the same pure tone audiogram and show a typical hearing threshold level around 35 dB at 0.5 kHz sloping down to 60 dB at 4 kHz. All subjects were fitted monaurally with one of three behind-the-ear programmable hearing instruments [3M, n=8; Widex Quattro Q8, n=6 and Phonak Phox, n=8]. The NAL-R prescription method was used to generate frequency response targets. The achieved frequency responses were very close to the real ear target for all of the models except at 4 kHz with the Phox where REIG was some 13 dB less than the target value. The individuals were diligent hearing instrument users with an average daily use of eight hours from week three onwards. Five experienced subjects were used to control for changes in performance that may have resulted from practice effects with the test material. The gain was set at the users preferred listening level on the initial test session and held constant for all subsequent test sessions.

The aided and unaided scores for the CST were obtained at 0 and 12 weeks post-fitting. The mean aided and unaided CST scores are shown in Figure 2.14. There was a statistically significant increase in mean benefit [$p<0.05$] from an initial level of around 4% to an eventual level of 8% at 12 weeks post-fitting in the experimental group. There were no statistically significant changes over time in the control group. This study demonstrates that there are increases in hearing instrument benefit over time that are not simply due to practice effects with the test material.

The earlier study by Gatehouse [1993] suggests that the increase in benefit is due to increased audibility of high frequency speech cues since this was the main difference between the two frequency responses. The 3M and Widex hearing instrument users should therefore obtain greater changes over time than the Phox hearing instrument users. The results revealed that 3M users showed little subsequent benefit while the Phox and Widex showed an increase of around 6%. Unfortunately, the statistical power was low and the differences across devices were not found to be significant. In addition, it is possible that high frequency amplification is important but this may only need to extend to 2-3 kHz.

The SPAC measurements were carried out at 3, 6, 9 and 12 weeks post-fitting. The mean aided composite SPAC score showed a statistically significant increase at nine weeks post-fitting compared to all earlier test sessions for the fitted ear [$p<0.05$]. This time span is similar to that reported by Gatehouse [1993]. No other test condition differed significantly over time. However, the SPAC subscores failed to show that the change over time was due to increased use of high frequency information.

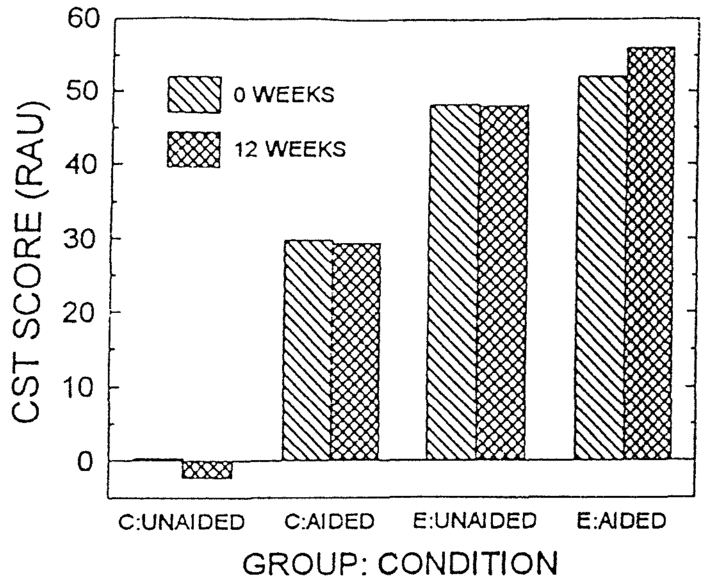


Figure 2.14. Mean recognition score in rau for Continuous Sentence Test for the experimental and control subjects. Aided and unaided scores are given at the time of fitting and twelve weeks post-fitting. Reproduced from Cox *et al.* [1996], with permission.

There was one other interesting finding in the Cox *et al.* article. A subgroup of subjects [n=7] was reviewed six months after the end of the main study. The mean aided and unaided scores for this subgroup are shown in Figure 2.15. The results show an increase in benefit over the extended time period; however, the increase in benefit is due to a rather dramatic decrease in the unaided scores and not an increase in the aided score. This differs from the original group who showed an increase in the mean aided score but no change in the mean unaided score at 12 weeks post-fitting. Gatehouse [1992a] showed an increase in the aided score and a corresponding decrease in the unaided score. The explanation for these differences is unclear but may be related to the amount of amplification: this is discussed in chapter five. It is possible that a combination of poorer unaided performance in the fitted ear and late-onset auditory deprivation in the not-fitted ear may account for the anecdotal reports that individuals become dependent on their hearing instrument.

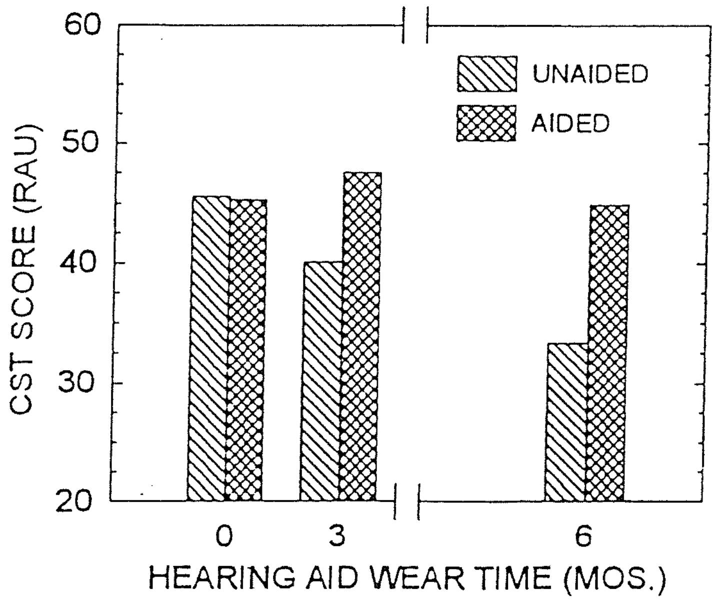


Figure 2.15. Mean recognition score in rau for Continuous Sentence Test for a subgroup of experimental subjects. Aided and unaided scores are given at the time of fitting and three and six months post-fitting. Reproduced from Cox *et al.* [1996], with permission.

Horwitz and Turner [1997] measured performance and self-report benefit over an 18-week post-fitting period. The performance test material was the UCLA recording of the 192-item NST presented at a level of 70 dB SPL. This consists of 16 consonants paired with the vowel 'ee' in three lists [two lists are consonant-vowel-consonant (CVC) and one list is vowel-consonant-vowel (VCV)]. Each list was repeated four times to give 192 items. This test is particularly sensitive to the listener's ability to detect and interpret high-frequency information. The self-report measure was the 48-item PHAB questionnaire that was used in the earlier acclimatisation study by Cox and Alexander [1992]. Measurements were carried out at three-week intervals from the time of fitting, except for the PHAB questionnaire, which did not commence until three weeks post-fitting. Since changes in aided performance over time could be confounded by changes in gain setting, each aided measurement was performed twice; once with the gain setting selected at the time of fitting and also with the current user-gain setting.

Subjects comprised 13 new monaural users with a mean age of 68 years. The individual pure tone audiograms show a sloping high frequency hearing impairment from around

25 dB HL at 0.5 kHz to around 65 dB HL at 4 kHz. Eight subjects received compression hearing instruments and five received linear hearing instruments. Specific details of the fitting prescription were not reported but it was noted that the mean 4 kHz REIG was 11 dB less than the NAL-R prescription target. The hearing instruments were worn for more than eight hours per day during the course of the study [Horwitz, personal communication]. The control group comprised 13 matched individuals who had been wearing a hearing instrument for a mean of four years.

Figure 2.16 shows the mean NST benefit as a function of time for the fixed-gain setting. There was a statistically significant increase in benefit over time of 7 rau in the new users. This represents a doubling of the benefit score measured at the time of fitting. There was no change in the control group over time. Thus, the study confirms that benefit increases with time for new hearing instrument users. This cannot be explained in terms of a practice effect with the test material since no change occurred in the control group. A change in audibility can be ruled out as an explanation for the improvement over time since the mean user-gain setting decreased by two decibels over the course of the study. In addition, the increase in benefit at the user-gain setting was similar to that observed at the fixed-gain setting. The authors did not find a relationship between high frequency amplification and acclimatisation although this may be because the mean gain at 4 kHz was well below the prescription target value.

The clinician is interested in predicting significant differences in benefit over time for individual hearing instrument users. While eleven [85%] individuals showed a tendency for benefit to increase over the course of the study, it would appear from the data that only three [23%] of these were greater than the 95% critical difference for benefit scores. Identifying the sources of variation among individuals is an important area of research that has still to be investigated.

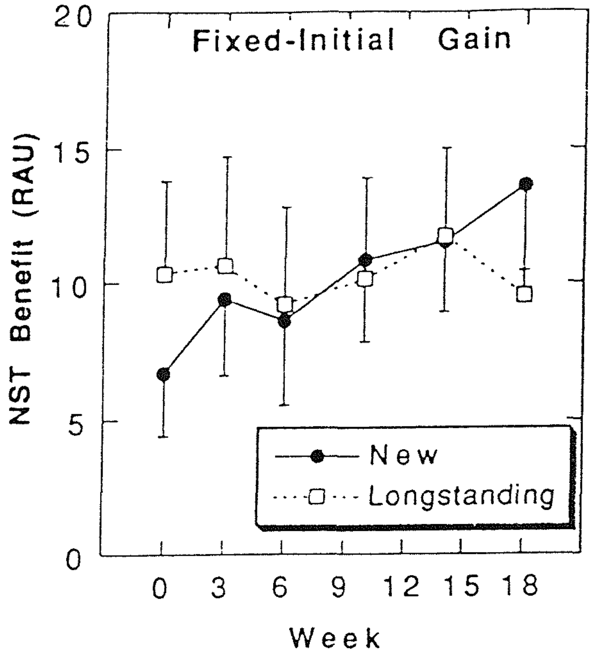


Figure 2.16. Mean benefit for Nonsense Syllable Test (aided minus unaided score) as a function of post-fitting time. The solid line corresponds to the new subjects and the broken line corresponds the long-standing subjects. Reproduced from Horwitz and Turner [1997], with permission.

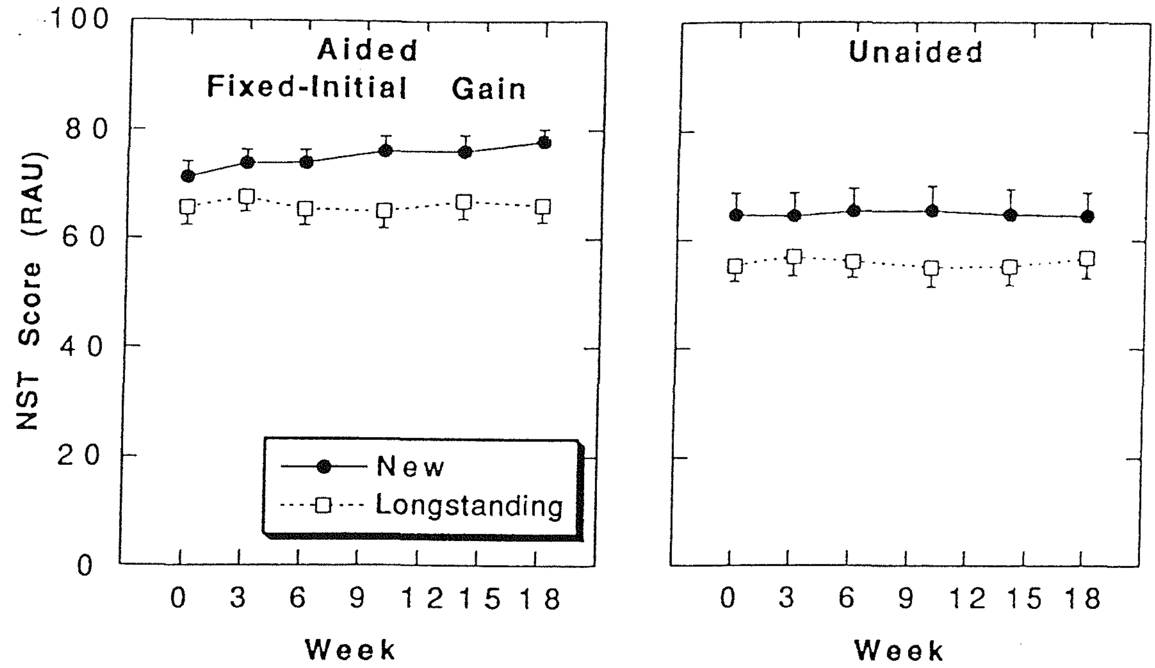


Figure 2.17. Mean recognition score in percent correct for Nonsense Syllable Test as a function of post-fitting time. The aided scores are shown on the left and the unaided scores are shown on the right. The solid line corresponds to the new subjects and the broken line corresponds to the experienced subjects. Reproduced from Horwitz and Turner [1997], with permission.

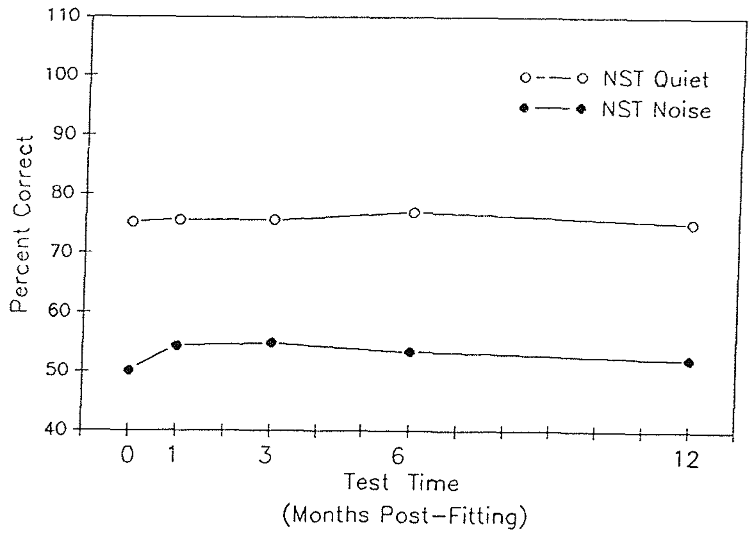


Figure 2.19. Mean aided recognition score in percent correct for Speech Perception in Noise as a function of time. The black symbols correspond to the low predictability items and the open symbols correspond to the high predictability items. Reproduced from Bentler *et al.* [1993a], with permission.

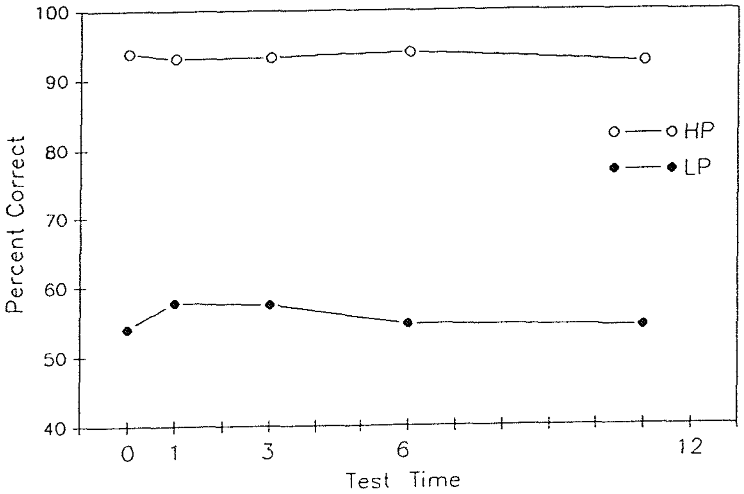


Figure 2.20. Mean aided recognition score in percent correct for Nonsense Syllable Test [NST] as a function of time. The black symbols correspond to presentation in noise and the open symbols correspond to presentation in quiet. Reproduced from Bentler *et al.* [1993a], with permission.

over time for the new hearing instrument users. There was no change in the mean unaided scores. Thus, the change over time was due to an increase in benefit for the aided condition. While this may be interpreted as supporting evidence for acclimatisation, it is possible that this is a reflection of the subject becoming more familiar with the hearing instrument [for example, become more proficient with adjustments of the gain control in different listening situations] and this causes the self-reported benefit to increase.

Kuk [2000] presented the results of a study involving 20 severe-to-profoundly hearing-impaired subjects who were refitted with a digital hearing instrument. The new instrument had a low compression threshold so substantially more gain was provided for quiet speech than with the previous instrument. The hypothesis was that this new audibility would result in an acclimatisation effect. The subjects were assessed using the SPIN test. The mean performance in quiet at 50 dB SPL showed a statistically significant increase from 23% at the time of fitting to 31% some three months later. It is difficult to understand why there should be an acclimatisation effect since the higher gain for quiet speech means that the overall SPL reaching the ear would not be too dissimilar from amplified speech at average listening levels. Since there was no control condition it is possible that the improvement over time was simply a practice effect although Kuk [personal communication] suggests that this is unlikely since the increase in performance was greater for the low predictability sentences than for the high predictability sentences. On the other hand, it is possible that changes other than audibility occurred with the new hearing instrument. For example, there may have been changes in temporal cues due to the compression circuitry and this may have required a period of acclimatisation.

In 1999, Arlinger *et al.* published the results of a one-year follow up of 29 subjects who had been fitted with a digital hearing instrument. The findings at one year were compared with the one-month post-fitting data. All of the subjects had been using a conventional analogue hearing instrument for several years prior to commencing the study. The mean pure tone hearing thresholds of the original group of 33 subjects [incorporating the 29 who were assessed over one year] showed a gently sloping high frequency hearing impairment

[Arlinger *et al.,* 1998]. Fifteen subjects had binaural fittings and 14 had monaural fittings. The mean age at one-year follow-up was 65 years. The study used a variety of outcome measures including the abbreviated version of the PHAB [APHAB], the Gothenberg Profile, a subjective quality judgement task and a speech recognition task. The speech recognition task involved measuring the SNR required to achieve 40% performance on low-redundancy sentences presented at a level of 60 and 75 dB (C).

After one year, the performance criterion on the speech recognition test was achieved with a less favourable SNR: a mean change of 1.41 and 0.65 dB at 60 and 75 dB (C) respectively. The authors report that the change in SNR at 60 dB (C) is equivalent to a 20-25% increase in performance. This improvement is larger than that reported in the earlier acclimatisation studies using analogue hearing instruments. A possible explanation for this apparent improvement is a calibration error. This seems unlikely since the mean change in SNR observed with a subgroup of subjects wearing their convention analogue hearing instrument was 0.1 dB [+0.63 and −0.78 at 60 and 75 dB [C] respectively]. It is unclear if the acclimatisation effect was due to a change in the frequency response with the new hearing instrument or if it relates to some other aspect of the digital processing. It is interesting to speculate why the difference in SNR over time should be more marked at a speech presentation level of 60 dB [C] than at 75 dB [C]. It is possible that this is simply random variation in the sample; however, it may have occurred because the 60 dB [C] presentation level corresponds more closely with the subject's familiar listening situation. In addition, this finding is not consistent with the results of the final experiment in the present thesis that suggests that the acclimatisation effect should be greatest for the higher presentation level. The reported change in SNR is so large that it does place some doubt over the reliability of the finding. There was little change on most of the self-report measures including the sub-scales of the APHAB. Cox and Alexander [1992] reported a significant increase on the complete PHAB but not for individual sub-groups.

In 1997, Ovegard *et al.* published the results of a study that investigated changes in perceived sound quality over time. Subjects comprised ten adults [mean age, 76.8 years] with a sloping mild-to-moderate high frequency sensorineural hearing impairment [30 dB HL at 0.5 kHz and 55 dB HL at 4 kHz]. All subjects were first time hearing instrument users who were fitted with behind-the-ear hearing instruments using the NAL-R prescription target. The subjects were asked to make repeated judgements in three situations within their own home environment. The three listening situations were speech in quiet, speech in noise and music. Examples of these three environments include listening to speech on television, speech in the presence of a noisy washing machine and music on the radio respectively. Each listening condition was judged on six quality scales: softness, clarity, brightness, fullness, loudness and total impression. The time period between the first and last judgement varied across subjects but was generally around 100-200 days. The subjects also varied in the number of occasions they made the judgement ratings. The subjects were responsible for setting the gain of the hearing instrument on each occasion that the quality judgements were carried out. The authors report that there was no evidence to suggest that there was a systematic change in gain during the period of study. Although there was considerable variation within subjects across time, analysis of the group data confirmed several trends. Firstly, there was a positive trend in perceived sound quality on the brightness scale when all three listening conditions were combined [$p<0.05$]. When the background noise condition was removed from the analysis [i.e., combining speech in quiet with the music environment] a positive trend over time was present on the clarity and total impression quality scales. The improvement in clarity over time is consistent with the increase in speech recognition scores reported in the studies discussed earlier. The relationship between quality judgements and time was analysed for each individual subject using regression analysis. Six of the ten subjects showed a significant trend over time for at least one quality judgement [$p<0.05$] although there was no clear pattern. In most cases, the changes over time were larger for the aided than the unaided condition.

2.4.2 Studies that do not show an improvement over time

The studies by Cox and Alexander [1992] and Gatehouse [1992] prompted renewed interest in improvements in hearing instrument benefit over time. However, other studies have cast doubt on the auditory acclimatisation phenomenon. The first of these studies was published by Bentler *et al.* [1993a,b]. Their study used both performance [1993a] and self-report [1993b] outcome measures with the results reported in separate articles. The performance measures included the SPIN test and the 62-item NST, which has since been used in subsequent acclimatisation studies [Humes *et al.*, 1996; Horwitz and Turner, 1997]. Speech was presented at a level between 50-60 dB HL. SNRs of +5 and +8 dB were used for SPIN and NST respectively. Measurements were carried out at the time of fitting and 1, 3, 6 and 12 months later. Only aided measurements were reported with no unaided control for practice effects.

Table 2.3. Profile of experimental subjects included by Bentler *et al.* [1993a,b]

Factor	Characteristic
Gender	43 male, 22 female
Age	63.8 years [range 21-84]
Style of hearing aid	46 ITE, 19 BTE
Number of hearing aids	55 monaural, 10 binaural
Hearing aid experience	39 new, 26 experienced
Hearing aid use	13 part-time, 52 full-time
Degree of impairment	37 mild, 28 moderate
Audiometric configuration	19 flat, 33 gently sloping, 13 steeply sloping
Circuit type	9 adaptive filter, 11 frequency-dependent input compression, 10 adaptive compression, 20 Zeta noise reduction, 15 no noise-reduction

The subjects used in this study were probably the most heterogeneous group used in any auditory acclimatisation study [see Table 2.3]. There were 65 subjects ranging in age from 21-84 years. Pure tone audiometry revealed 37 [57%] subjects with a mild hearing impairment and 28 [43%] with a moderate hearing impairment. Audiometric configurations ranged from flat to steeply sloping. There were 26 [40%] subjects who were experienced users of amplification and some of the remaining 39 [60%] 'new' subjects had been using a hearing instrument for up to 12 months prior to commencing the study. A wide range of hearing instruments was used ranging from linear with peak clipping to input compression and some with noise reduction circuits. The mean frequency response showed a good match to the NAL-R target up to 2 kHz but it was 6.2 and 7.7 dB below the 3 and 4 kHz targets. With the exception of the initial measurements when the gain control was fixed, subjects adjusted the gain to their preferred setting on all subsequent visits. The biggest discrepancy occurred at six months post-fitting when the 3 and 4 kHz gains were 10 dB and 13 dB below the NAL-R target. Although most subjects reported an average daily use of greater than nine hours, 13 [20%] subjects used the instrument for less than four hours per day. A total of 55 [85%] subjects received monaural hearing instruments and ten [15%] received binaural hearing instruments.

The results of the SPIN test and NST are shown in Figures 2.19 and 2.20 respectively. There was no statistically significant change over time although the low predictability SPIN sentences and the NST in noise showed an increase of around 4% at 1-3 moths post-fitting. It is possible that this would have been statistically significant if a repeated-measures ANOVA had been restricted to the data up to 12 weeks post-fitting. It would have been difficult to measure any increase in the mean score of the high predictability SPIN items since the initial score was in excess of 90%. There was no significant interaction with factors such as experience, degree and configuration of hearing impairment, reported daily use or circuit type.

It could be argued that auditory acclimatisation is not likely to be present in individuals with a mild hearing impairment, who may not be good hearing instrument users, and when the frequency response of the hearing instrument may be providing less than optimum amplification. Recognition of NST material is dependent on high frequency cues but the hearing instruments failed to meet the high frequency prescription target by up to 10-13 dB. It is also possible that not all of the experienced users were provided with new acoustic information after re-fitting so acclimatisation would not be expected to occur.

The same group of subjects also completed a variety of self-report outcome measures including the 38-item 'Understanding Speech' subsection of the Hearing Performance Inventory [HPI], a qualitative judgement task and a satisfaction questionnaire. The HPI was developed by Giolas *et al.* [1979] as a self-report inventory of daily listening problem areas. The 'Understanding Speech' subsection can be divided into different listening environments: a] fairly quiet, b] background music, and c] background speech. The measurements were carried out at 0, 6 and 12 months post-fitting. The results are shown in Figure 2.21. The only score to change significantly over time [$p<0.05$] was the 'fairly quiet' environment where there was a perceived improvement. This agrees with Cox and Alexander [1992] who showed most improvement occurred in the living-room environment. There was no change in the other self-report measures over time. Taken as a whole, the results of this study do not show evidence of an increase in benefit over time.

In 1993, Taylor reported a study that assessed changes in performance and self-report outcome in 58 subjects after hearing instrument fitting. The performance test material was the 50-item Northwestern University Auditory Test 6 [NU-6] presented at 50 dB HL in quiet and in the presence of noise [+10 dB SNR]. The self-report measurement was the 25-item Hearing Handicap Inventory for the Elderly [HHIE] developed by Ventry and Weinstein [1982]. The measurements were carried out at three weeks post-fitting and repeated at 3, 6 and 12 months post-fitting. Subjects had a mean age of 72 years and it is reported that they had a high-frequency hearing impairment with an average high frequency threshold of 34.5 dB HL. Thirty-seven [64%] subjects were fitted monaurally and 21 [36%] binaurally. The specific prescription target was not reported but examination of the mean functional gain data reveals that hearing thresholds improved by 15 dB at 3 kHz. The functional gain measurements were very stable over the year suggesting that the mean user-gain setting did not change. The amount of daily use made of the hearing instrument was not reported.

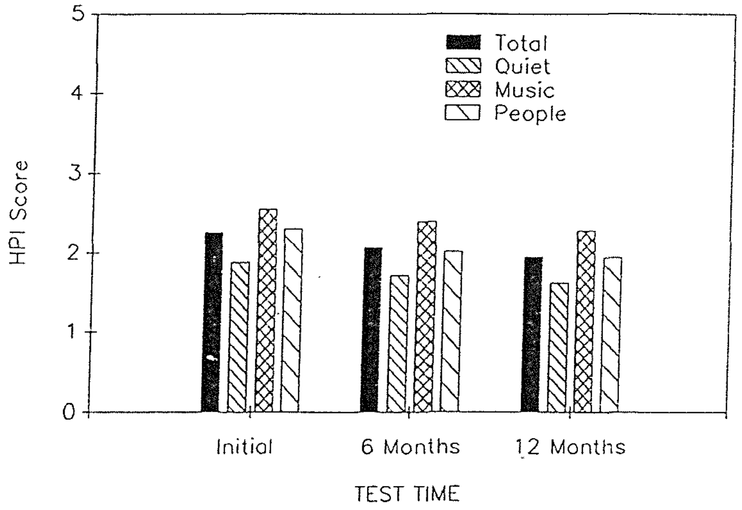


Figure 2.21. Mean scores for each of the Hearing Profile Inventory speech understanding categories as a function of time. Reproduced from Bentler *et al.* [1993b], with permission.

The mean aided scores for speech in quiet and speech in noise at each test session were 84-85% and 58% respectively with no change over time. The high scores obtained for speech in quiet may have prevented detection of further increases over time. In addition, the speech in quiet condition is not particularly challenging and there may be little need for perceptual learning to take place in this condition. The recognition of monosyllabic words relies on frequency information above 1 kHz but they are not particularly sensitive to very high frequencies where the hearing impairment was greatest. It is not known if the presentation level of speech or the hearing instrument gain setting was representative of the subject's normal listening experience.

The self-report measure showed a significant decline in handicap at three weeks post-fitting. However, the handicap score at three months was significantly higher that at any other time period. The return to higher handicap may have occurred for several reasons. For example, the increase in audibility at the time of fitting may result in a more favourable perception of residual difficulty and it may take the individual some time to appreciate the ongoing difficulty in background noise. This change in perception could result in an apparent increase in handicap over time. In summary, this study failed to demonstrate any improvement in performance and self-report measurements over time. However, compared to other studies, individuals had a very mild hearing impairment and received relatively little gain from the hearing instruments. This means there may have been no consistent improvement in audibility and hence no acclimatisation. Also, the results were obtained from a combination of monaural and binaural fittings and there have been no studies demonstrating acclimatisation in subjects with binaural fittings.

Humes *et al.* [1996] investigated the reliability and stability of several hearing instrument outcome measures over a six-month post-fitting period and discussed the relevance of the findings with reference to auditory acclimatisation. Performance measures included the NST [used previously by Horwitz and Turner, 1997] presented in quiet at 70 dB SPL and the 60-item Hearing In Noise Test [HINT; Nilsson *et al.*, 1994] presented at a signal-to-babble ratio of +8 dB. This consists of 250 short and meaningful sentences and is sensitive to audibility of frequencies below 3 kHz. The latter was administered as a sentence-based measure of speech recognition and is more sensitive to low and mid frequency speech information. Therefore, it more closely represents real-world communication abilities than the NST. Self-report measures included the HHIE, previously used in the Taylor [1993] study, as well as an abbreviated version [37-item] of the Hearing Aid Performance Inventory [HAPI] designed for use in the elderly population [Schum, 1993]. The measurements were carried out at 0, 7, 15, 30, 60, 90 and 180 days post-fitting. Due to technical problems, scores from the HINT were not reported after 60 days post-fitting.

Twenty subjects [mean age 71.5 years] completed the study. The median hearing threshold was around 30 dB HL at 0.5 kHz and 65 dB HL at 4 kHz. Prior to commencement of the study, ten [50%] subjects were already hearing instrument users. Four [40%] of the remaining subjects had previously worn hearing instruments intermittently but not at all in the past two years. All subjects were fitted with binaural programmable hearing instruments that matched the NAL-R targets except at 4 kHz where the mean gain was 5-10 dB below the target; this was done deliberately to reduce acoustic feedback. The NAL-R response programme [i.e., gain fixed at target] was used most of the time but the exact details of programme use was not analysed [Humes, Personal Communication]. The data logging facility of the hearing instrument revealed that they were worn for around 6-7 hours per day during the course of the study.

The NST data are shown in Figure 2.22. The hearing instrument improved the score by around 15%. The initial unaided score was used to calculate benefit. Any change in benefit over time was small and not statistically significant on a repeated-measures ANOVA [$p>0.05$]. No information was provided concerning the statistical power of the analysis and no numerical data [effect size and standard deviation] were given by the authors. The mean unaided score for the HINT was in excess of 80% at the start of the study. The initial aided score was in excess of 90% so there was little opportunity to detect further improvements over time. There was no change in the HHIE or the HAPI over time.

An improvement in NST over time might have been anticipated since the subjects were making good use of their hearing instruments and they were providing significant high frequency amplification. One explanation is that 50% of subjects were experienced hearing instrument users who may have conceivably acclimatised to the NAL-R frequency response before the study commenced. Thus, only half of the subjects may have been in a position to show changes over time. In addition, if acclimatisation to binaural aiding is to be investigated, it may have been inappropriate to test the subjects monaurally [the specificity of perceptual learning was discussed in section 2.2]. The experimental design was configured to study hearing instrument performance rather than acclimatisation and this probably explains why the performance tests were carried out monaurally.

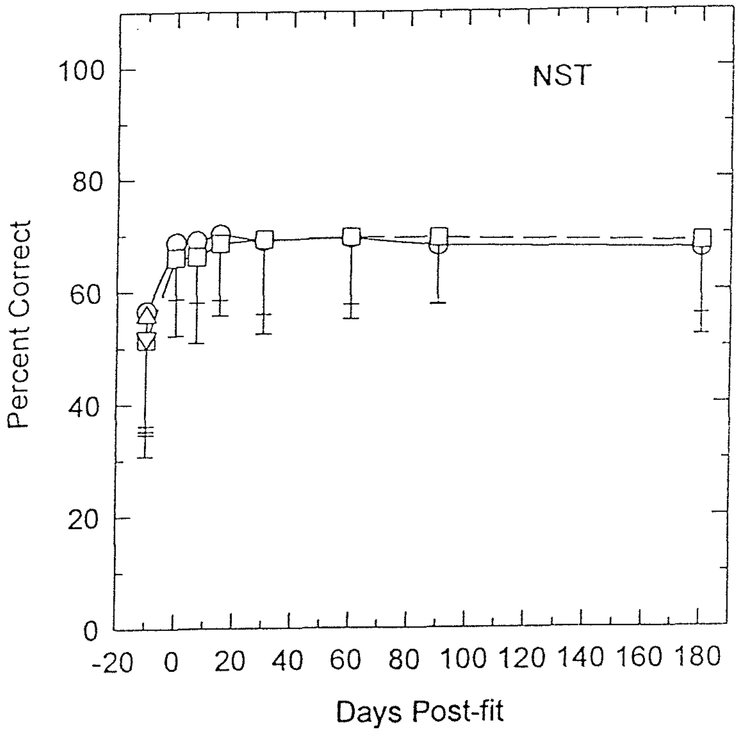


Figure 2.22. Mean recognition score in percent correct for Nonsense Syllable Test as a function of post-fitting time. The circles and squares correspond to the right and left ear respectively. The triangles correspond to an unaided retest carried out at the end of the study. The bars correspond to one standard deviation. Reproduced from Humes *et al.* [1996], with permission.

Saunders and Cienkowski [1997] published a study that reported hearing instrument benefit over a post-fitting period of three months. The study was not designed specifically to investigate acclimatisation but rather to investigate the effectiveness of different programmable hearing instruments. The design included 48 subjects, 24 were new users and 24 experienced users. Subjects were reported as having a mild-to-moderate, symmetrical, sensorineural hearing impairment with a presumed aetiology of presbyacusis and noise damage. The subjects were split into six groups. The individuals in each group were fitted binaurally with a different hearing instrument configuration [some groups used the same hearing instrument but different electroacoustic characteristics]. The SRT in quiet was measured using CID W-1 spondees. This test material is sensitive to the audibility of frequencies below 1 kHz. The SRT in noise was measured using the full 250 sentence HINT. The SRT in noise was measured adaptively using two scoring methods. In one method the noise was adjusted to obtain 50% correct recognition. In the second method, the noise was increased or decreased by the subject according to whether they believed they could understand the test material [irrespective of their true performance]. The presentation level for speech on the HINT was set at most comfortable listening level [MCL] for the unaided and the aided listening condition on the day of fitting. This resulted in mean presentation levels of 74 and 62 dB SPL respectively. Testing was carried out at 0, 30, 60 and 90 days post-fitting. A summary of the data is given in Figure 2.23. All results were expressed as the mean difference between aided and unaided test scores.

The mean aided SRT in quiet was some 8 dB lower than the mean unaided presentation level. Subjects required a mean unaided SNR approximately 0.5-1.5 dB more favourable than when aided. The authors noted that mean performance at the time of fitting was high compared to subsequent visits. They suggested that the initial scores were artificially high because of a combination of better-aided performance [perhaps because of increased motivation] and worse unaided performance [perhaps because of procedural learning on subsequent test sessions]. For this reason, statistical analysis did not include the data collected at the time of fitting. There was a trend for performance to improve over time [i.e., criterion performance was obtained with a less favourable HINT SNR but this was not statistically significant ($p>0.05$)]. The authors did not give the statistical power of the analysis: the number of subjects is known and the effect size can be estimated from the graph but there is no indication of the within-subject standard deviation required to calculate the statistical power. There was a significant difference over time for the speech in quiet data but this was due to reduced performance at 60 days post-fitting. The data were separated for new and experienced users and re-analysed. The trend of increasing performance over time in noise was present for both new and experienced users. There was no significant difference between new and experienced users or any significant interaction over time.

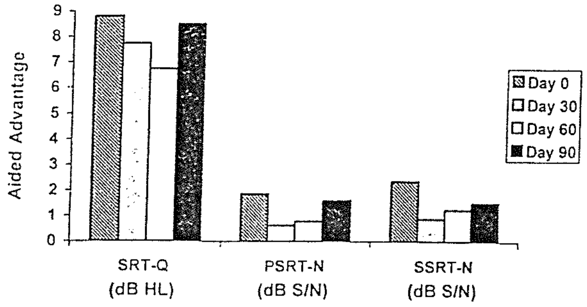


Figure 2.23. Mean aided improvement as a function of time. The improvement in speech reception threshold in quiet is given in dB HL. The improvement in self-report and performance in noise is given in dB signal-to-noise ratio. Reproduced from Saunders and Cienkowski [1997], with permission.

There are several potential weaknesses in this study. The initial 30 days were excluded from the analysis so any acclimatisation effect occurring before this time was missing from the analysis. Six different hearing instrument configurations were used in the study and it is not known if different signal processing strategies result in different rates of acclimatisation. Finally, although the authors demonstrated significant improvements in audibility when aided, it is not known if the hearing instruments were worn on this setting or how much use was made of the hearing instrument in the subject's own environment.

The authors suggest that previous studies demonstrating acclimatisation have used subjects fitted monaurally and tested them monaurally with the unaided ear occluded. [This is not strictly correct since Cox and Alexander (1992) used a combination of monaural and binaurally aided subjects.] They suggest that the asymmetry in amplification may enhance the acclimatisation effect perhaps by reducing the initial benefit score. Consistent with this point is the study by Gatehouse [1992a] that demonstrated a larger acclimatisation effect with monaural headphone presentation than in the sound field [the non-test ear was not occluded]. However, acclimatisation occurred with both monaural headphone presentation and bilateral sound field stimulation. The hypothesis requires testing in a single experiment in which the protocol allows these parameters to be contrasted: monaurally fitted subjects would need to be tested with the control ear occluded as well as unoccluded: binaurally fitted subjects would need to be tested monaurally as well as binaurally.

In 1997, Neuman *et al.* presented the results of an acclimatisation study on seven subjects with a mild to moderate hearing impairment. The subjects were fitted with monaural hearing instruments using the NAL-R target values. Speech testing included spondees and monosyllables [in quiet and in noise], HINT and NST. Aided and unaided speech recognition testing was carried out at monthly intervals over a five month post-fitting period. It is assumed [although not stated] that the non-test ear was plugged and muffed. There was no control condition to check for a practice effect. The benefit scores were expressed in terms of improvement in SNR [except for the NST, which was given as a percentage improvement in score]. Performance improved when aided on most outcome measures but there was no statistically significant difference in benefit over time. Given the inherent variability of speech tests, the statistical power of the analysis [which was based on a total of seven subjects] must have been low although observation of the benefit scores does not show any trend over time. The NST was the only test material that relied on high frequency acoustic cues. Numerical data concerning the hearing threshold levels, the closeness of fit to the NAL-R targets and the daily use of the hearing instrument were not given in the presentation so it is possible that conditions were not optimal for improvements over time.

Surr *et al.* [1998] compared short- and long-term changes on the PHAB and the CST. Both of these measures had previously been used to show changes in benefit over time by Cox and Alexander [1992]. The benefit measures were carried out at six weeks post-fitting and again some 16-22 months later. The 15 subjects had a mean age of 67.3 years [range, 55-

75] and had mean hearing thresholds of 30 dB HL at 0.5 kHz and 60 dB HL at 4 kHz. The subjects had between 3-12 years experience with linear hearing instruments before commencing the study. At the start of the study, they were fitted with binaural wide dynamic range compression [WDRC] hearing instruments and were reported to be 'full-time' hearing instrument users. The gain/frequency response characteristics were not reported.

The mean results for the PHAB are shown in Figure 2.24. The initial benefit scores on PHAB were in excess of 40% and there was no statistically significant increase over time. The authors noted that the large initial improvement in performance might have resulted in ceiling effects. Cox and Alexander [1992] obtained a significant increase in the overall PHAB score that was initially around 20%. The results for the four CST conditions are shown in Figure 2.25. The mean difference between the mean short- and long-term score was less than 2 rau. This difference was not statistically significant. The authors conclude that a six-week period is sufficient to allow for acclimatisation with WDRC hearing instrument fittings. Like the Kuk [2000] study reported earlier, it is not clear why there should be an acclimatisation effect unless the level of amplified speech experienced by the subjects differs from the range previously experienced with the linear hearing instrument. In theory the amplification of average conversational speech should be the same and the range of SPL covering quiet and loud speech should be narrower with the WDRC hearing instrument so there may be no new information available. It is not known if WDRC processing is sufficiently different from linear processing to have a marked effect on acclimatisation. A study has yet to be published which specifically investigates the effect of speech presentation level and SNR in new hearing instrument users. In addition, acclimatisation may have occurred before the initial measurement at six weeks post-fitting [or transient improvements may have occurred during the two-year interval between the two test sessions].

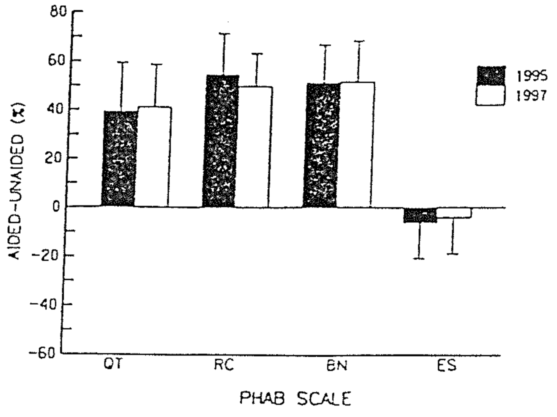


Figure 2.24. Mean initial and long-term benefit for each Profile of Hearing Aid Benefit sub-scale (pre-fitting minus post-fitting score). Listening in quiet environment, QT; reduced visual cues, RC; background noise, BN; environmental sounds, ES. Reproduced from Surr *et al.* [1998], with permission.

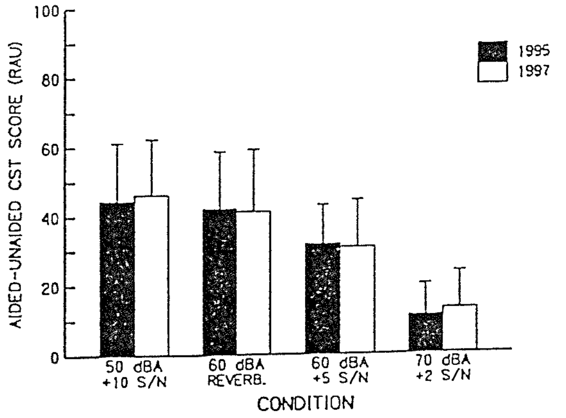


Figure 2.25. Mean benefit in rau for Continuous Sentence Test (aided minus unaided score) as a function of post-fitting time. The black columns correspond to initial benefit at six weeks post-fitting. The open columns correspond to benefit two years post-fitting. Reproduced from Surr *et al.* [1998], with permission.

In a recent article, Bentler *et al.* [1999] summarised an unpublished study by Holte [1997]. This study was undertaken to measure changes in benefit in a group of 26 new binaural hearing instrument users over a six-month post-fitting period. Subjects were in the age range 46-76 years and had bilateral hearing impairment ranging from mild to severe. The hearing instruments were set according to NAL-R targets although the closeness of the match was not reported. Benefit was assessed at monthly intervals using the NST. The presentation level and SNR were not reported. Subjects were also required to indicate their preferred frequency response [NAL-R versus flat amplification] at monthly intervals when listening to running speech and music. Details about the user-gain setting as well as the gain setting used at each test session were not provided. Also, the daily use of the hearing instrument was not reported. It is assumed that all testing was binaural although this was not specified. There were two control groups: one group of experienced users and one group of subjects yet to be fitted with hearing instruments.

Although the mean benefit scores were different for each group, there is no obvious increase over time. It would have been helpful to know the mean aided scores to ensure that there was not a ceiling effect. The results of the subjective preference using running speech showed that the experienced users preferred the response to which they had been fitted [i.e., NAL-R]. However, the preference for NAL-R increased over time for the new users [and for the individuals yet to be fitted who only wore a hearing instrument for the test sessions]. In summary, there was no evidence of acclimatisation on speech tests but there were changes in listener's preference over time. Further details of the experimental design are required before the study can be critically reviewed.

In a study that has yet to be published, Humes *et al.* [In press] investigated performance at intervals up to 3 years post-fitting. This study used a larger sample size, included more outcome measures and used a longer post-fitting time than most previous acclimatisation studies. In addition, this is one of the few acclimatisation studies that have tested subjects fitted binaurally. Performance was measured on CST and NST at 65 dB SPL in the presence of cafeteria noise [SNR +8 dB]. The CST was also presented at 50 dB SPL in the quiet. Self-report was measured using the HAPI and HHIE. The number of new hearing instrument users who completed the assessment at 1, 2 and 3 years was 88, 31 and 10 respectively. Subjects were typically 70 years of age with a gently sloping sensorineural hearing impairment [1, 2 and 4 kHz average of around 50 dB HL]. Each subject was fitted with binaural linear ITE hearing instruments fitted to the NAL-R targets. Although aided performance testing was carried out at the target values, the mean user gain was some 6-9 dB below the target. For example, the match to target at 2 kHz in the 2-cc coupler was 25 dB but at the mean user setting was nearer 16 dB. The initial unaided and aided testing was carried out at 2 and 4 weeks post-fitting respectively. The CST in noise showed an increase in mean benefit, when tested binaurally, from around 5% at the time of fitting to nearer 10% at one year post-fitting. This was due to an increase in aided performance and a slight decrease in unaided performance. A similar pattern of results was obtained for the smaller groups of subjects who completed testing at 2 and 3 years post-fitting. Benefit on the other performance test conditions was relatively stable with no clear evidence of acclimatisation. In addition, the self-report measures showed, if anything, a decrease in benefit over time. Thus, there was little evidence of acclimatisation.

In a review article on acclimatisation, Turner *et al.* [1996] commented that most studies on late onset auditory deprivation had also failed to show any evidence of acclimatisation in the fitted ear. However, testing [primarily CID at 40 dB re: SRT] was typically presented via headphones and did not specifically test the situation of acclimatisation under more realistic conditions of listening to newly amplified speech via the subjects' hearing instrument. Gatehouse [1992a] has demonstrated that appropriate frequency shaping of the test material is required to mimic the subjects' normal listening situation [i.e., the hearing instrument gain and frequency response characteristics]. In addition, speech should be presented at a level that is representative of the subjects' normal listening level.
Turner and Bentler [1998] have been critical of a review article by Palmer *et al.* [1998], which paints a favourable picture of acclimatisation. In a letter to the editor they state that '*the acclimatisation effect is small or non-existent, is obscured by large variability in the measurement tools and has not been reliably demonstrated to be important*'.

2.5 Supporting evidence of auditory acclimatisation

There are a number of studies that have measured hearing instrument benefit over an extended period of time. Although these studies have not been designed specifically to investigate auditory acclimatisation, they do provide supporting evidence for it. For convenience, the studies that have used speech recognition tests have been separated into studies using conventional acoustical hearing instruments and studies using cochlear implants. With the exception of speech recognition performance, the only psychoacoustic abilities that have received specific attention are intensity discrimination and perception of loudness. These topics are reviewed in the penultimate subsection. The final subsection briefly reviews studies on auditory localisation and lateralisation.

2.5.1 Acoustic hearing instruments

Foust and Gengel [1973] compared conventional and frequency-transposition hearing instruments. The subjects were nine moderate-to-profoundly hearing-impaired students, eight of whom were linear hearing instrument users. Monosyllabic word recognition scores were obtained for the Modified Rhyme Test [MRT] and the CID phonetically balanced word lists. The MRT is somewhat similar to the FAAF test but there are six alternatives in a closed-set format. When words were wrongly identified, they were repeated, together with the word given in response by the listener. This procedure enabled listeners to concentrate on the auditory features that could be used to identify the test items. Testing was undertaken at weekly intervals over a period of six weeks. The frequency transposition instrument was not worn outside the laboratory. The median MRT score as a function of test session is shown in Figure 2.26. This shows that there was an initial decrement in performance with the transposition instrument but, after six weeks of testing, the aided score increased from around 46% to nearer 70%. The corresponding change over time with the conventional hearing instrument [effectively the control condition for practice effects] was around 6%.

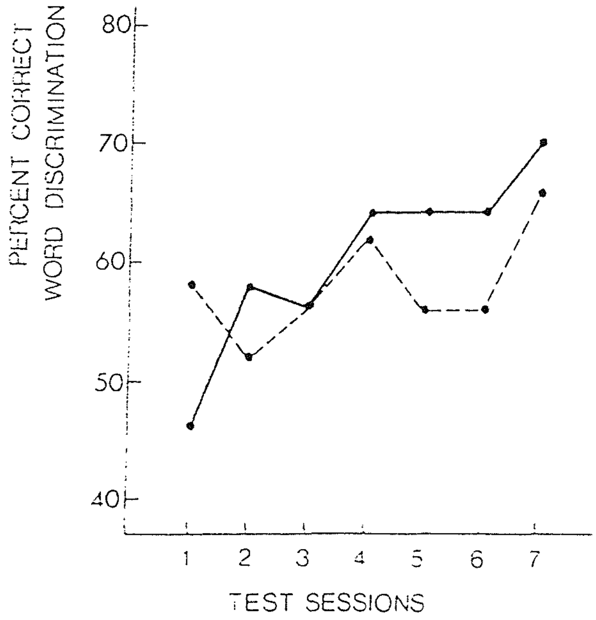


Figure 2.26. Median aided recognition score in percent correct for Modified Rhyme Test as a function of transposed post-fitting time. The solid line corresponds to the transposed hearing aid and the broken line corresponds to the conventional hearing aid. Reproduced from Foust and Gengel [1973], with permission.

Yund and Buckles [1995] carried out a series of laboratory experiments on 15 hearing-impaired subjects over a period of one year. The experiments were designed to study different parameters of multichannel compression signal processing. The subjects had a mean age of 65 years [range, 48-79 years] and ten were previous hearing instrument users. The pure tone thresholds were typically 30 dB HL at 0.5 kHz and 70 dB HL at 4 kHz. The same stimulus [NST] was used across three experiments resulting in a total of 150 hours listening experience. There was no control condition [for example, testing with the subjects own linear hearing instrument] to check for practice effects but most subjects had already taken part in a previous experiment using the NST and the authors report that improvements due to practice were complete before the experiments commenced. In addition, the subjects also practised the NST before commencing these new experiments.

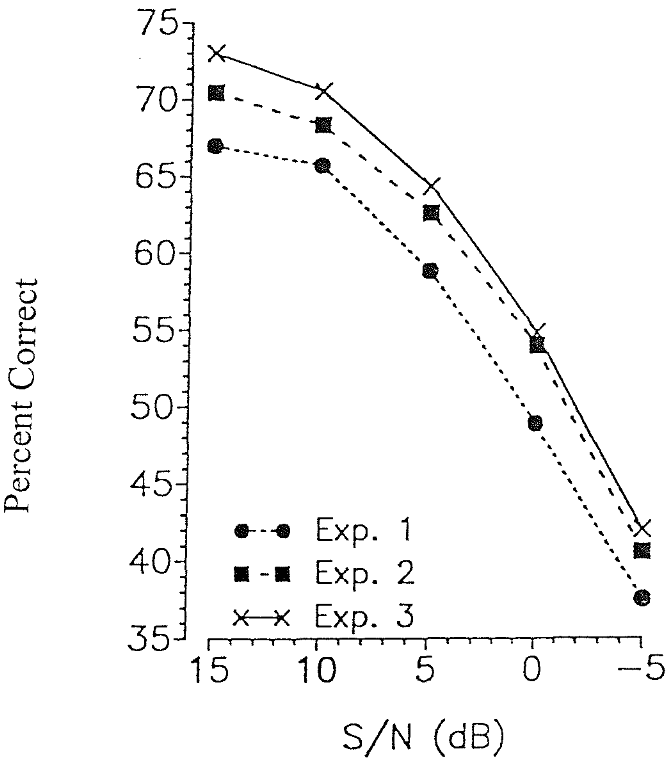


Figure 2.27. Mean recognition score in percent correct as a function of signal to noise ratio for three experimental conditions. The circles, squares and crosses correspond to experiment one, two and three respectively. Reproduced from Yund and Buckles [1995], with permission.

Figure 2.27 shows the mean results for the three successive experiments. Irrespective of SNR, it can be seen that the scores increase on successive experiments. The horizontal distance between experiment one and three [i.e., change in SNR] at the mean recognition score of 58.5% was +3 dB. The mean increase in recognition score between the first and third experiment was 4.8%. Fourteen of 15 individual subjects showed significant improvement across experiments [significant positive slope of regression line across experiments]. These results indicate that the subjects learned to make better use of the information provided by the multichannel compression hearing instrument as the experiments proceeded. The authors commented that the increase was specific to multichannel compression processing because it did not generalise to frequency-shaped linear amplification used in later experiments. This is consistent with the very specific nature of stimulus learning discussed in section 2.2.

Kiessling and Steffens [1993] compared a three-channel AGC hearing instrument in 26 experienced users with their own single-channel AGC hearing instrument. Rhyme test material was used to measure performance over a range of SNR from +15 to −5 dB. In four subjects, the SRS was measured at the time of fitting and after a trial of 20 days. The pure tone thresholds for this subgroup were not reported but they came from a larger pool of subjects who had a mean hearing threshold level around 45 dB at 0.5 kHz and 60 dB at 4 kHz. The new hearing instrument provided similar gain to the existing instrument. The mean score for the four subjects is shown in Figure 2.28. The authors did not plot standard deviations because of the small number of subjects and the data were not subjected to statistical analysis. The scores were higher after the trial period than at the time of fitting. The improvement was typically around 10%. The difference was less marked at the more favourable SNR suggesting that perceptual learning may be greatest in a more adverse listening environment. However, there was no control condition so it is not possible to rule out practice effects.

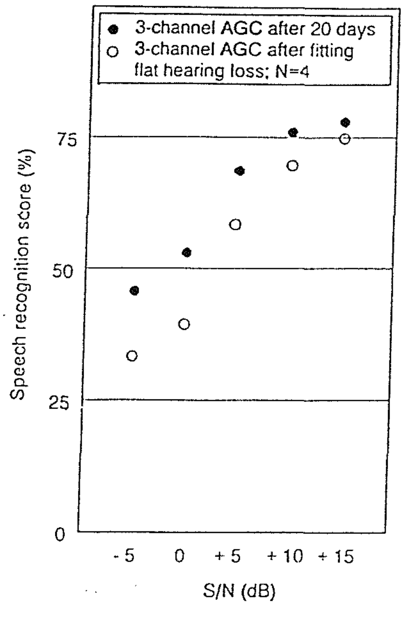


Figure 2.28. Mean recognition score in percent correct on the German Rhyme Test as a function of signal to noise ratio. The open circles correspond to the score with the three-channel AGC instrument at the time of fitting and the closed circles correspond to the score at 20 days post-fitting. Reproduced from Keissling and Steffens [1993], with permission.

2.5.2 Cochlear implants

Cochlear implants differ from conventional hearing instruments because they bypass the cochlear mechanics and stimulate the auditory nerve electrically. An electrode array is placed in the basal 1½ turns of the cochlea. When stimulated, the auditory nerve fibres send neural impulses to the brain and these are interpreted as sound. An individual with a very severe hearing impairment who receives limited benefit from a conventional acoustic hearing instrument may be suitable for cochlear implantation. Cochlear implants represent an extreme form of treatment for individuals with a very severe hearing impairment; these individuals experience an extreme change from hearing nothing to hearing something. This makes it more likely that re-learning will be required before maximum benefit is achieved.

There are a number of studies that have tracked the performance of subjects after cochlear implantation. Dorman and Loizou [1997] presented speech recognition data as a function of time in a single cochlear implant subject. There were significant increases in consonant and vowel recognition over time. There were also significant changes in sentence material but not single monosyllabic words. This is consistent with the improvements, over time, in consonant and vowel recognition. The small change in word recognition suggests that the changes in sentence recognition arose out of an interaction between consonant and vowel information and sentence context. That is, small changes in consonant and vowel recognition can lead to large changes in sentence context. Dorman *et al.* [1990] demonstrated improvements over a period of two years in 27 adult-implanted subjects. The median recognition score for spondees increased from 10% at one month post-fitting to 56% by two years. Spivak and Waltzmann [1990] compared speech recognition performance on 15 subjects over a period of three years. Subjects exhibited greatest improvement in the first three months of implant use but improvements continued for three years. Fryauf-Bertschy *et al.* [1992] demonstrated improvements over time in children who received a cochlear implant. The post-lingually impaired children exhibited significant improvements after six months whereas the rate of improvement in the congenitally impaired children was slower. Tyler and Summerfield [1996] compared the mean speech recognition score for three different devices. There was a considerable increase in benefit over time [see Figure 2.29]. Similar to the Dorman and Loizou [1997] study, there were greater changes in sentence material than on a consonant test. One device showed rapid and greatest benefit but this was not apparent initially. This finding agrees with the study by Gatehouse [1993] who demonstrated that benefit increases over time and that differences in hearing instrument response only became apparent after a period of acclimatisation.

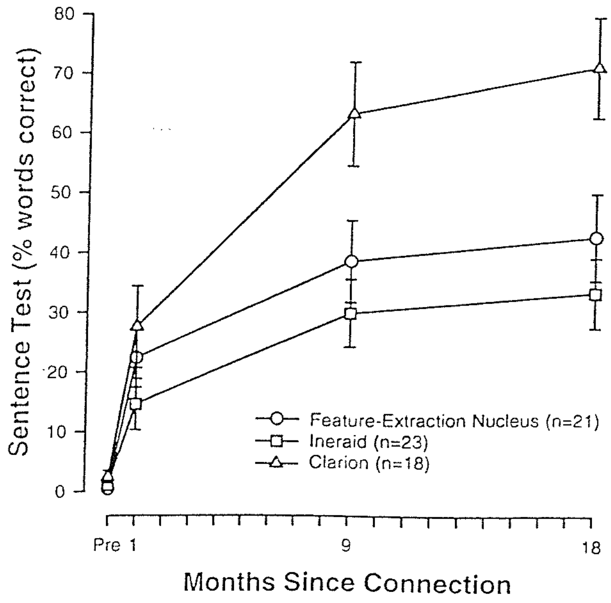


Figure 2.29. Mean recognition score in percent correct on Iowa Sentence Test as a function of post-connection time in cochlear implant subjects. The squares, circles and triangles correspond to the Ineraid, Nucleus and Clarion device. Reproduced from Tyler and Summerfield [1996], with permission.

Tyler *et al.* [1997] expanded on this earlier study by providing data on a larger number of subjects [49 adults] and over a longer time period [see Figure 2.30]. Post-implantation performance was superior to pre-implantation performance at 9 months but continued to improve up to 18-30 months post-fitting.

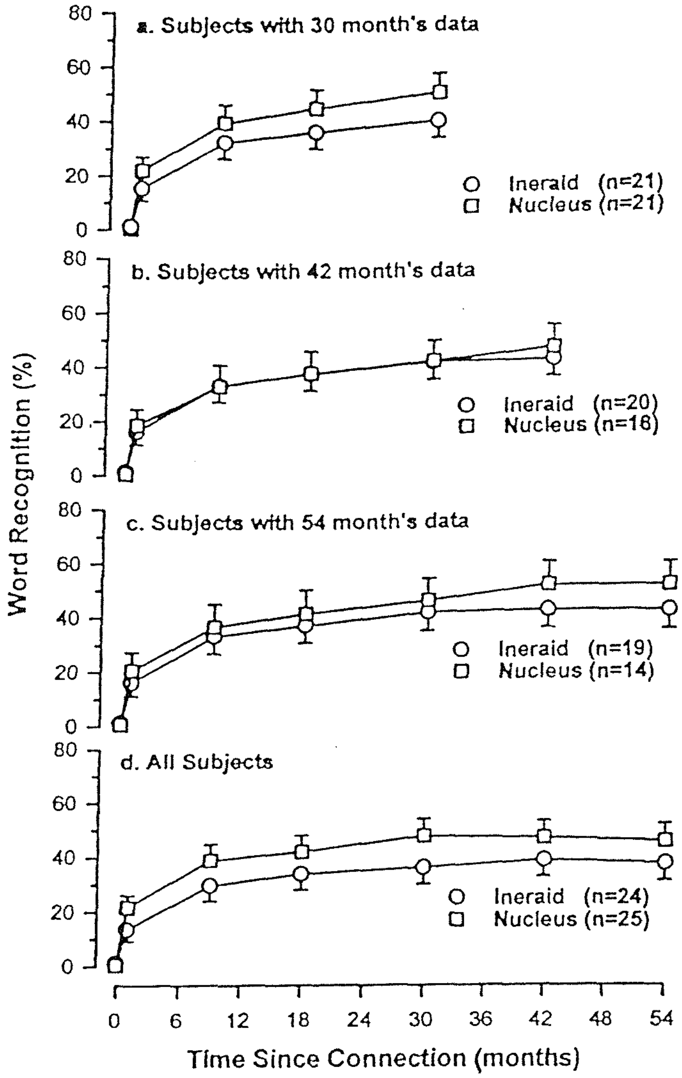


Figure 2.30. Mean recognition score in percent correct on Iowa Sentence Test as a function of post-connection time in cochlear implant subjects. The top three panels show group data for the same subjects. Vertical bars correspond to one SEM. The bottom panel shows estimated mean scores over time for all available data. Vertical bars correspond to 1 SEM based on a large sample approximation. The squares and circles correspond to the Ineraid and Nucleus. Reproduced from Tyler *et al.* [1997], with permission.

Summerfield and Marshall [1996] presented data of environmental sound recognition in cochlear implant subjects. The stimulus set was 20 sounds and these were presented at 1, 9 and 18 months post-operation. The initial score of 39% at 1 month increased by 12% at 9 months and then a further 6% at 18 months. It is possible, however, that the scores increased because of practice with the same limited stimulus set.

The definition of acclimatisation provided by Arlinger *et al.* [see chapter one] requires the occurrence of an improvement that cannot be attributed purely to task, procedural or training effects. In these cochlear implant studies, it is not possible to completely rule out these factors since the non-implanted ear cannot be used as the control condition. However, Tyler and Summerfield reported that lip-reading scores did not improve significantly over time and this tends to rule out changes due to a practice effect with the test material. It therefore appears that the major improvement in performance with a cochlear implant over time does represent acclimatisation. The magnitude of improvement is considerably greater than in the studies using conventional acoustical hearing instruments. This is consistent with the need for greater learning because of the unfamiliarity arising from the incomplete representation of the speech signal and its transposition to the basal turns of the cochlea.

2.5.3 Intensity discrimination and perception of loudness

Several studies have specifically investigated changes in intensity discrimination following hearing instrument fitting [Robinson and Gatehouse, 1995, 1996]. The impetus for this work was the 1989 study by Gatehouse that showed that the fitted ear [tested unaided] performed better than the not-fitted control ear at high presentation levels, while at lower levels the converse was true. Robinson and Gatehouse [1995] investigated the difference limen for intensity [DLI] at different frequencies in four hearing-impaired individuals. The subjects [age range 54-82 years] had a mean pure tone threshold of 24 dB HL at 0.25 kHz and 58 dB HL at 3 kHz. Each subject had been fitted with a linear monaural hearing instrument for a mean of 2 years and 7 months. The REIG was 0 dB at 0.25 kHz and 19 dB at 3 kHz. The control group consisted of five normal-hearing individuals in the 18-35 year age range.

The stimuli were tone complexes centred at 0.25 and 3 kHz. DLI were measured using the gated pedestal method with an adaptive, three-interval, forced-choice procedure [Levitt, 1971]. This was performed for stimuli at 65, 80 and 95 dB SPL. The DLI was calculated as the Weber fraction10 log [$\Delta I/I$].

The results are summarised in Figure 2.31. The only statistically significant difference in gradient was at 3 kHz, where the fitted ear gradient was steeper than the not-fitted ear and the normal group [$p<0.05$]. This means that DLI in the fitted ear was poorer at low presentation levels but better at high presentation levels compared to the not-fitted ear. The simplest explanation for this finding is that intensity discrimination changed as a result of exposure to amplified sound at 3 kHz since performance was best at the frequencies where amplification was greatest. This parallels the findings of Gatehouse [1989] for speech identification in noise.

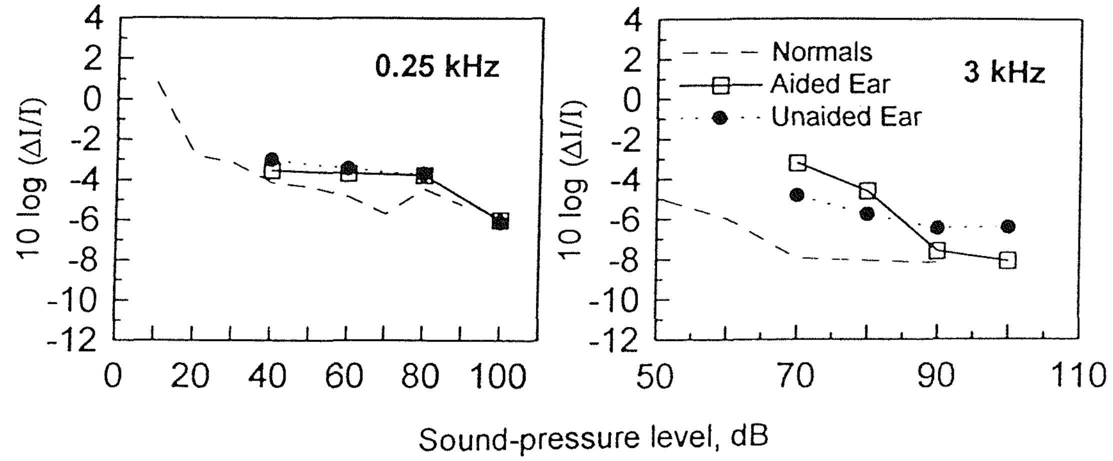


Figure 2.31. Mean intensity discrimination threshold as a function of absolute level for hearing-impaired subjects at 0.25 and 3 kHz. The open squares correspond to the fitted ear and the filled circles correspond to the not-fitted ear. For comparison, mean results for normal-hearing subjects are shown as a dashed line. Reproduced from Robinson and Gatehouse [1995], with permission.

Since it is possible that the change at 3 kHz was present before hearing instrument fitting, Robinson and Gatehouse [1996] carried out a prospective study of intensity discrimination in five individuals [age range 38-83 years] who were fitted with a monaural, linear, peak-clipping hearing instrument. The subjects had a bilateral sensorineural hearing impairment with a mean pure tone threshold [in the fitted ear] of 23 dB HL at 0.25 kHz and 69 dB HL at 3 kHz. The not-fitted ear of one subject was considerably better than the fitted ear and this resulted in a mean 3 kHz threshold of 59 dB HL in the not-fitted ear. The hearing instruments were fitted according to the NAL-R target for REIG. There was 1 dB gain provided at 0.25 kHz [target 0 dB] and 19 dB at 3 kHz [target 19 dB]. The stimuli and equipment were the same as those used in the previous study. Measurements were carried out at 0-4, 6-12 and 15-18 weeks post-fitting. The results are shown in Figure 2.32.

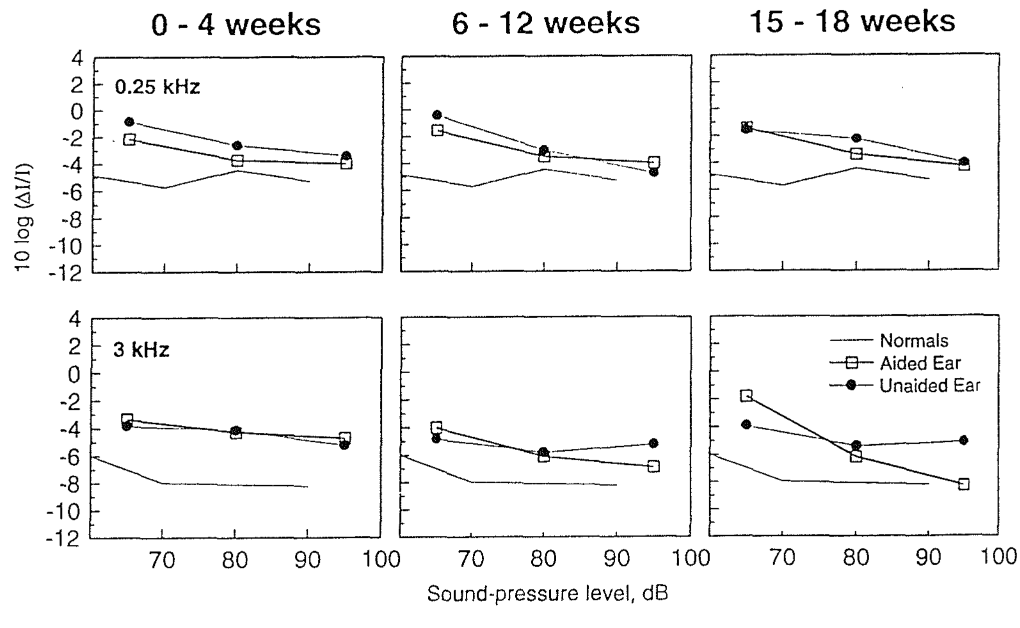


Figure 2.32. *Mean intensity discrimination threshold as a function of absolute level showing the effect of post-fitting time. The upper graphs show the 0.25 kHz stimulus and the lower graphs show the 3 kHz stimulus. The open squares correspond to the fitted ear and the filled circles correspond to the not-fitted ear. For comparison, mean results for normal-hearing subjects are shown as a thin line. Reproduced from Robinson and Gatehouse [1996], with permission.*

There is a progressive influence of hearing instrument experience on the 3 kHz DLI with a statistically significant interaction between duration and stimulus level [$p<0.05$]. Discrimination was significantly better in the fitted ear at high presentation levels, 15-18 weeks post-fitting [$p<0.05$]. This study shows that the fitted ear becomes progressively better able to discriminate intensity at the highest sound pressure level for frequencies that are normally amplified by the hearing instrument. Some 15-18 weeks of experience with the hearing instrument was required before this was observed. There was little or no change over time at lower sound pressure levels, at frequencies not amplified by the hearing instrument, or in the not-fitted control ear. This finding is similar to the speech identification in noise results reported by Gatehouse in 1989. The one subject with an asymmetric hearing impairment [30 dB difference at 3 kHz] showed the largest change in DLI although the significance of this finding is unclear.

In the studies that have demonstrated increased benefit on speech recognition tests, this has been measured from around six weeks post-fitting as opposed to the15-18 weeks for the DLI study. It may well be that there are slower changes occurring in a range of auditory abilities that collectively can be revealed earlier on speech recognition tests. Improvements in speech performance may continue to occur until all of these individual abilities reach an asymptote. It is not known if changes in intensity discrimination form the basis of the auditory acclimatisation phenomenon reported on speech recognition tasks or if there are also changes in other auditory capabilities. Experiments have still to be performed that investigate changes in frequency resolution and temporal resolution over time for different intensities.

Gatehouse and Robinson [1996] studied the perception of loudness in four listeners with symmetrical sensorineural hearing impairment who were established monaural hearing instrument users. The subjects were between 67 and 83 years of age and pure tone thresholds were around 30 dB HL at 0.5 kHz and 65 dB HL at 2 kHz. The REIG was around 3 dB at 0.5 kHz and 25 dB at 2 kHz. The loudness scaling technique used a restricted response scale with anchor points that aim to produce results with good stability [for further details, see Gatehouse and Robinson, 1996]. The data for each subject were analysed separately and are shown in Figure 2.33. None of the subjects showed any systematic difference in loudness rating between ears at 0.5 kHz where there was little amplification. However, there was a divergence in the loudness ratings between the ears of all subjects at 2 kHz. The individual subject's results were subjected to an analysis of variance with the object of detecting differences between the ears at different frequencies and presentation levels. There were no significant differences between the ears at 0.5 kHz for any of the subjects. However, there were significant differences at 2 kHz for every subject. This finding demonstrates that the loudness function differs between ears at the frequency for which the hearing instrument provided material gain.

In the same article, Gatehouse and Robinson reported a single case study showing changes in DLI as well as changes in the amplitude of the slow vertex auditory evoked potential over time. Data were collected at stimulus levels of 65, 80 and 95 dB SPL. The results are presented in Figure 2.34. The intensity discrimination results are very similar to those reported earlier by Robinson and Gatehouse. There was no apparent difference at the lower frequency. At the higher frequency, the fitted ear exhibits better intensity discrimination abilities at high presentation levels. The electrophysiological data also showed no difference at the lower frequency. However, at the higher frequency, the fitted ear shows the highest amplitude at the highest presentation level. Thus, for this single subject, the electrophysiological data exhibited similar behaviour to the psychophysical data. This study provides further evidence that systematic changes occur in the auditory system following the fitting of a single hearing instrument.

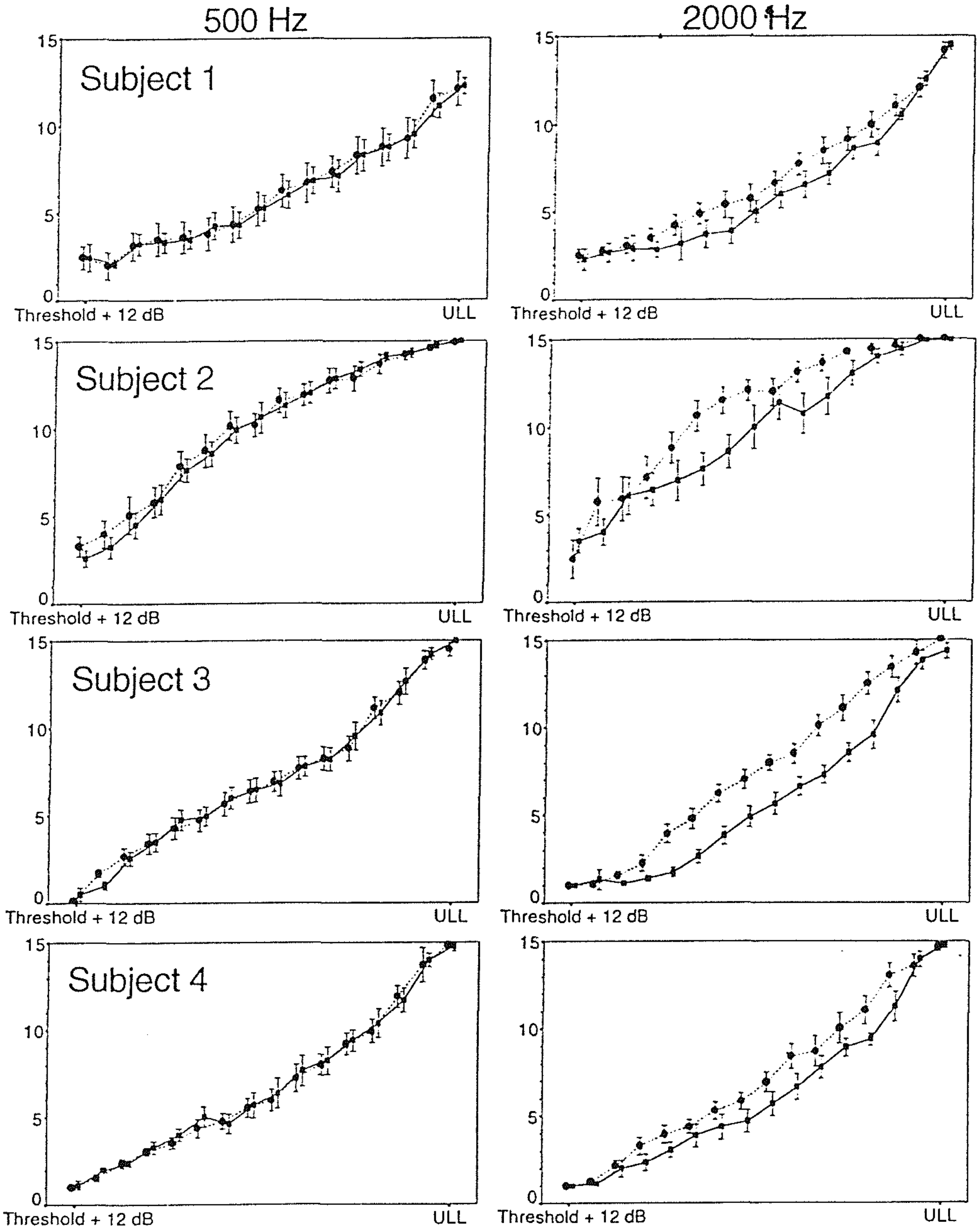


Figure 2.33. Mean loudness rating at 0.5 and 2 kHz in four hearing impaired subjects. The filled squares correspond to the normally aided ear and the filled circles correspond to the normally unaided ear. Vertical bars show 2 SEM. Reproduced from Gatehouse and Robinson [1996], with permission.

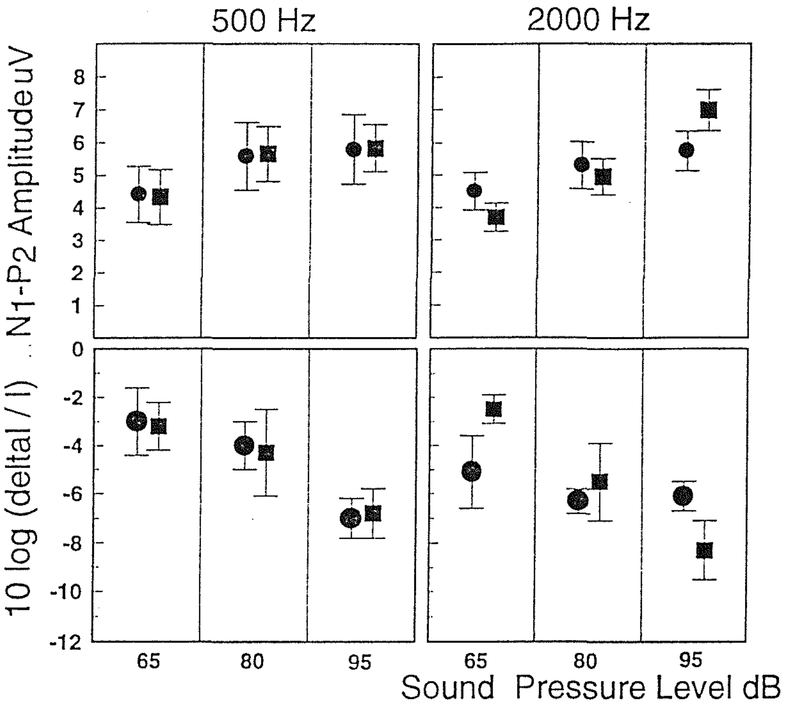


Figure 2.34. Amplitude of the SVR and the difference limen for 0.5 and 2 kHz at 65, 80 and 95 dB SPL in a single subject (subject 4 in previous figure). The filled squares correspond to the normally aided ear and the filled circles correspond to the normally unaided ear. Vertical bars show 2 SEM. Reproduced from Gatehouse and Robinson [1996], with permission.

In a conference presentation by Philibert *et al.* [2001a], they reported changes in intensity discrimination and perception of loudness in eight subjects after hearing instrument fitting; intensity discrimination improved at 2 kHz while sound was categorised as 'less loud'. In a further presentation, Philibert *et al.* [2001b] reported that the latency of the auditory brainstem response [ABR] evoked by a 90 dB nHL click decreased in a group of five subjects after hearing instrument fitting. While full details of both of these studies have yet to be published, the results are consistent with the studies by Robinson and Gatehouse.

Olsen *et al.* [1999] also compared loudness perception in long-term hearing instrument users with a group of control subjects who had not worn a hearing instrument. There were 18 subjects [age range, 24-65 years] in each group. Hearing thresholds were around 40 dB HL at 0.5 kHz and 65 dB HL at 4 kHz. The median duration of hearing instrument use was nine years [range, 0.5-27 years]. Ten subjects were fitted binaurally and eight were fitted monaurally. No information was provided about REIG. Categorical loudness scaling was performed unaided. The intensity levels were selected at random and the subject had to allocate them to one of six categories ranging from 'very soft', to 'too loud'. There was a statistically significant difference in the mean value allocated to the 'loud' category [$p < 0.05$]. In the non-users the value was 102.9 dB HL but this was 4.5 dB higher in the hearing instrument users. This change reflects a higher tolerance to loud sounds in ears that are experienced at listening to high presentation levels. The result suggests that there may be an acclimatisation effect leading to greater tolerance of loud sounds. However, the groups were not matched for hearing level, age and test frequency: a carefully controlled prospective study of the effect in new hearing instrument users has yet to be reported.

2.5.4 Localisation and lateralisation

Byrne and Dirks [1996] reviewed studies on auditory localisation and lateralisation that provide evidence that experience with a particular type of hearing instrument fitting may result in improved performance. For example, Markides [1977] compared the aided localisation performance of experienced individuals fitted with monaural and binaural hearing instruments. Under bilateral conditions, individuals who were normally aided binaurally performed best. Markides also showed that monaurally aided subjects did better than normal subjects under monaural conditions. Noble and Byrne [1990, 1991] demonstrated that aided subjects localise best with their normal hearing instrument.

Bauer *et al.* [1966] studied changes in localisation for normal subjects who wore an earplug in one ear. They reported that subjects were able to re-orient themselves to localisation cues and that this could be speeded up with training. Butler [1987] also demonstrated that monaural localisation could be improved with training but this was specific to the signal used for training. This finding agrees with the perceptual learning framework whereby acclimatisation occurs for the stimulus with which the subject is familiar. These studies provide evidence that localisation abilities can be influenced by acclimatisation.

Florentine [1976] carried out a study on lateralisation. She showed that individuals with an asymmetric hearing impairment are able to lateralise in the middle of their head. This result suggests that these individuals had learned to lateralise with sound levels less loud in the poorer ear than in the better ear. When fitted with an earplug, normal subjects required several days of use before they were able to lateralise in the same manner as the hearing-impaired subjects. After removing the earplug, they continued to lateralise with a lower intensity for seven days. This study is evidence that lateralisation abilities can change over time in normal hearing subjects.

Wilmington *et al.* [1994] examined changes in various auditory abilities of subjects after surgery for correction of a conductive hearing impairment. At four weeks post-surgery, subjects performed abnormally on localisation and masking level difference tests. However, when the tests were repeated at 24 weeks post-surgery, 5 [71%] subjects had improved localisation and 2 [50%] subjects had improved masking level difference. This finding is consistent with acclimatisation occurring during the interval between the tests.

In summary, the studies reviewed in this section suggest that acclimatisation effects are not restricted to speech recognition tests alone but may affect a wide range of auditory abilities.

2.6 Summary and Aims

In summary, recent experiments have provided evidence that behaviourally important stimuli can cause the CANS of adult humans to reorganise. This may result in improvements in performance if the reorganisation enables the subject to extract previous unused information from the signal. Speech is a behaviourally important signal and this can be changed by provision of amplification [for example, changes in frequency response and overall level]. Thus, provision of amplification may induce reorganisation within the auditory system resulting in perceptual learning and a change in performance. This may explain late-onset auditory deprivation effects and auditory acclimatisation.

Many retrospective studies have demonstrated that performance on a speech recognition task decreases in the unaided ear of subjects with a bilateral hearing impairment who have been fitted monaurally; however, a large-scale prospective study reporting incidence, magnitude and time course of this late-onset deprivation effect has yet to be published.

The situation regarding improvements, over time, in the newly fitted ear is less clear. Improvements in performance have been demonstrated in a number of studies using a variety of performance and self-report outcome measures. Changes over time have also been shown to occur in psychoacoustic abilities such as intensity discrimination and loudness perception. These studies show irrefutable evidence for the existence of auditory acclimatisation. Despite this conclusion, some studies have failed to measure an acclimatisation effect. It is not always clear that these studies have provided a change to the speech signal. For example, there may have been infrequent use of a low gain hearing instrument. Alternatively, the test conditions may not be appropriate for the, sometimes, very specific nature of perceptual learning. It is likely that the rate of learning is affected by factors such as degree of hearing impairment and the amount of change brought about by amplification. The conflicting findings have led some researchers to state that the acclimatisation effect is small or non-existent. Thus, there are substantial uncertainties concerning the test conditions required to measure auditory acclimatisation and there are many outstanding research questions. It is not known under what conditions acclimatisation exists, how prevalent it is, the exact effect size, its full time course, or the effect of binaural versus monaural amplification. It is difficult to address the outstanding research questions until the uncertainties surrounding the measurement of acclimatisation are resolved.

The aim of the thesis is to improve understanding of the conditions required to measure auditory acclimatisation. This should result in a more robust methodological framework for future studies on auditory acclimatisation. The first consideration is to ensure that the hearing instrument provides new information that could benefit the subject since this is a prerequisite of perceptual learning.

CHAPTER THREE
GENERAL METHODOLOGY

3.0 Introduction

Details of the methodology that is common to all three experiments in the thesis are provided in this chapter. This includes information on subject selection, outcome measures and test procedures. Methodological issues that are specific to each individual experiment are reported in subsequent chapters.

3.1 Subjects

Subjects were all first time hearing instrument users recruited from the local hospital audiology service. They presented complaining of hearing disability and pure tone audiometry revealed a symmetrical, mild-to-moderate, sloping, high frequency sensorineural hearing impairment. Exclusion criteria included an asymmetry in air conduction thresholds of greater than 15 dB at two or more frequencies, an air-bone gap on either ear of greater than 15 dB at any test frequency and abnormal middle ear function assessed using oto-admittance audiometry. Prior to hearing instrument management, subjects were interviewed and informed that the aim of the experiment was to investigate the benefit provided by a single hearing instrument on different settings and on different tests in order to determine the 'best' settings. However, they were naive to the changes expected over time. A new group of subjects was recruited for each experiment. The sample size was calculated for paired data with a mean difference of 4% on the FAAF test and a standard deviation of the difference of 5.0%. Fifteen subjects were required for a statistical power of 80% at a two-tailed significance level of 5% on Student's *t*-test but 16 were recruited to allow for attrition.

The mean hearing threshold and uncomfortable loudness levels as a function of audiometric frequency for the subjects recruited to experiment one, two and three are given in Table 3.1-3.3 respectively [see also Fig 3.1]. The results show a symmetrical, high frequency sensorineural hearing impairment with a reduced dynamic range. The standard deviations are relatively narrow and reflect the homogeneity of the group.

Table 3.1. Summary of audiometric data for the subjects recruited for experiment one. The table includes the mean air conduction, not-masked bone conduction hearing threshold levels and uncomfortable loudness levels, in decibels hearing level. One standard deviation is given in brackets [n=16].

Frequency [Hz]	250	500	750	1000	1500	2000	3000	4000	6000
Air conduction [dB HL]									
Fitted ear	25 [11]	33 [8]	38 [8]	42 [8]	46 [8]	53 [8]	58 [7]	61 [7]	64 [9]
Control ear	24 [10]	30 [10]	35 [11]	41 [12]	45 [13]	53 [8]	58 [10]	62 [8]	66 [10]
Bone conduction [dB HL]									
Not-masked		32 [11]	27 [9]	33 [13]	40 [12]	56 [8]	53 [11]	60 [8]	
Uncomfortable loudness level [dB HL]									
Fitted ear		93 [16]		94 [16]		97 [16]		98 [14]	
Control ear		90 [19]		94 [17]		95 [15]		101 [15]	

Table 3.2 Summary of audiometric data for the subjects recruited for experiment two. The table includes the mean air conduction, not-masked bone conduction hearing threshold levels and uncomfortable loudness levels, in decibels hearing level. One standard deviation is given in brackets [n=16].

Frequency [Hz]	250	500	750	1000	1500	2000	3000	4000	6000
Air conduction [dB HL]									
Fitted ear	36[9]	33 [7]	39 [5]	42 [7]	47 [8]	51 [9]	56 [87]	60 [7]	66 10]
Control ear	33[8]	36 [8]	39 [7]	41 [7]	44 [7]	49 [9]	57 [9]	63 [9]	68 [7]
Bone conduction [dB HL]									
Not-masked		32 [9]	33 [8]	32 [11]	45 [6]	48 [9]	49 [9]	52 [10]	
Uncomfortable loudness level [dB HL]									
Fitted ear		98 [13]		97 [15]		101 [15]		106 [13]	
Control ear		99 [11]		97 [11]		102 [13]		106 [12]	

Table 3.3. Summary of audiometric data for the subjects recruited for experiment three. The table includes the mean air conduction, not-masked bone conduction hearing threshold levels and mean uncomfortable loudness levels, in decibels hearing level. One standard deviation is given in brackets [n=16].

Frequency [Hz]	250	500	750	1000	1500	2000	3000	4000	6000
Air conduction [dB HL]									
Fitted ear	33 [12]	34 [16]	38 [14]	44 [12]	50 [10]	53 [8]	60 [9]	64 [6]	71 [9]
Control ear	34 [11]	36 [13]	38 [13]	44 [11]	49 [9]	52 [11]	58 [10]	62 [8]	67 [12]
Bone conduction [dB HL]									
Not-masked		30 [13]		35 [8]		50 [11]		55 [7]	
Uncomfortable loudness level [dB HL]									
Fitted ear		102 [11]		101 [10]		104[12]		108 [14]	
Control ear		101 [10]		100 [11]		105[11]		111 [13]	

Table 3.4. Mean change in air conduction hearing threshold level [initial minus final measurement in decibels] in the test ear at the beginning and the end of the study. One standard deviation is given in brackets [n=16].

Frequency [Hz]	250	500	1000	2000	4000
Experiment One	–2 [6]	0 [6]	–1 [4]	–1 [3]	–1 [5]
Experiment Two	4 [8]	-1 [5]	-2[7]	-1[7]	-1[8]
Experiment Three	0[6]	0[6]	1[6]	0[8]	2[8]

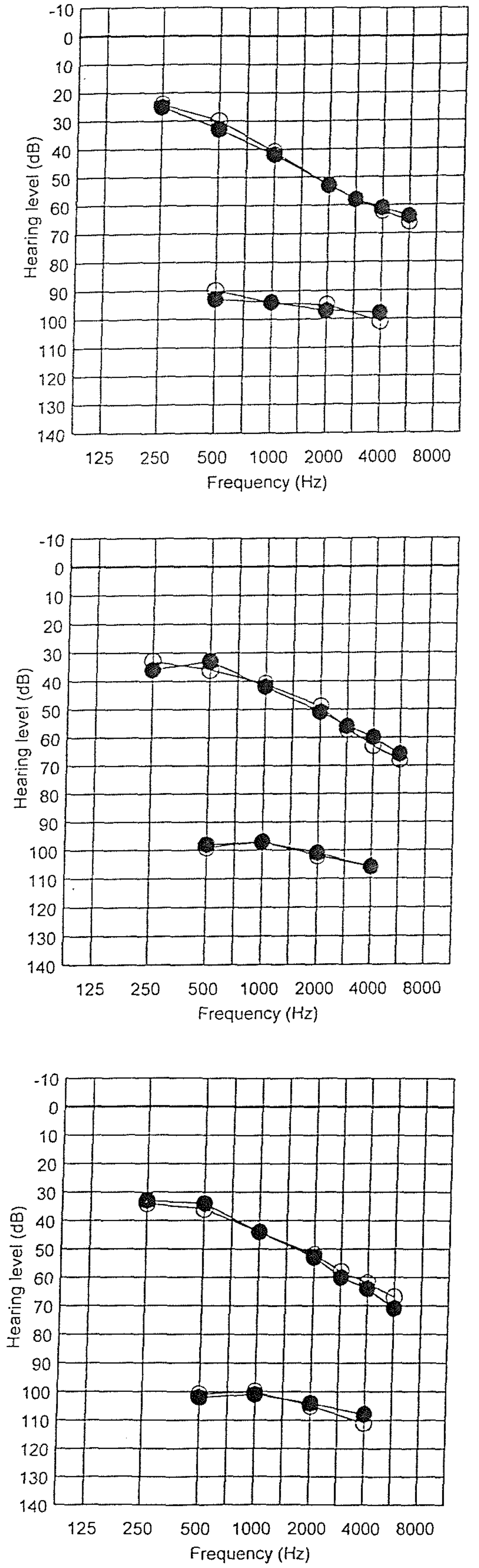


Figure 3.1. Pure tone hearing threshold levels [upper symbols] and uncomfortable loudness levels [lower symbols] for the subjects used in experiment 1 [top], two [middle] and three [bottom]. The filled symbols show the fitted ear and the open symbols show the control ear.

The mean change in hearing threshold level between the beginning and end of each experiment [initial minus final measurement] for the fitted ear is shown in Table 3.4. Inspection of the raw data confirmed that over 80% of hearing thresholds changed by less than 5 dB at each audiometric frequency and 90% changed by less than 10 dB. This degree of repeatability is consistent with the test-retest differences reported in the literature [for example, Robinson, 1991].

3.2 Hearing Instrument Fitting

Following recruitment into the study, each subject had bilateral impressions taken for a standard shell earmould, to which a parallel 0.8 mm vent was added. The same model of hearing instrument was used for all subjects: Phonak Sono-Forte 331X-L PiCS. This is a miniature, high gain, digitally programmable BTE hearing instrument. This model has a three-way filter that allows considerable flexibility when tailoring the frequency response. The hearing instrument can also store three programmes that can only be accessed using a hand control. The subjects were not issued with the hand control and were therefore not able to change programme during the study; however, they were able to control volume (except in experiment three where this was disabled). The fittings were all monaural as this allowed the not-fitted ear to be used as the control condition. The not-fitted control ear underwent the same test protocol as the fitted ear. It was important to have a control condition because any increase in performance in the control ear could be explained by a practice effect due to repeated exposure to the test material. Only the difference between the increase in performance in the fitted ear and the increase in performance in the not-fitted control ear was assumed to be a measure of acclimatisation. Some subjects expressed a clear preference for ear to be fitted: the fitted ear in the remaining subjects was selected at random. The earmould for the not-fitted control ear was retained in the laboratory and not given to the subject.

Subjects were informed that they would be allowed to keep the hearing instrument if they completed the study although they were given the choice of changing to a standard National Health Service hearing instrument when the study terminated, or sooner if they decided to withdraw from the study. No subject opted to change to a standard NHS hearing instrument. Travelling expenses were paid to each subject. Approval was obtained from the local research ethics committee of each hospital where subjects were recruited. Approval was also obtained from the Human Experimentation Safety and Ethics Committee at the Institute of Sound and Vibration Research where the study was undertaken.

3.3 Outcome measures

The main speech recognition test used in the study was a digitised version of the Four Alternative Auditory Feature [FAAF] test. This was selected because the subjects were provided with high frequency amplification and this test is strongly dependent on the audibility of high frequency speech sounds. This was the reason the FAAF test was also used in the acclimatisation studies reported by Gatehouse [1992a, 1993]. A copy of the FAAF test was obtained on compact disc from the MRC Institute of Hearing Research. The FAAF test is a forced-choice word recognition task based on the rhyme test principle, described by Foster and Haggard [1979, 1987]. The material consists of 20 sets of four minimally paired words; each based on two binary auditory/phonetic distinctions, giving an 80-item vocabulary. The items in a set differ on either initial or final consonant, and hence the test is particularly sensitive to high frequency auditory capabilities. The key words occur in the context of the carrier phrase, 'Can you hear [*keyword*] clearly?' One item was presented acoustically in the carrier phrase [spoken by a male speaker] and the subject's task is to select the appropriate word from the choice of four presented on a computer monitor touch screen. A list of the key words along with their frequency spectrum is shown in Appendix A. The frequency spectrum of conversational speech [ANSI, R1997] has been included for comparison. The FAAF key words show slightly more emphasis at the high frequencies compared to conversational speech.

All tests were carried out in a quasi-free sound field [ISO 8253-2] with reverberation times < 0.3 s, measured in octave intervals from 0.125 to 4 kHz. The ambient noise levels measured at the reference point[1], with the test equipment switched on but subject absent, was sufficiently low to allow measurement of binaural hearing thresholds in the sound field of 0 dB [maximum uncertainty +5 dB] from 0.25 to 4 kHz [ISO 8253-2][2].

---

[1] Defined in ISO 8253-2 as the midpoint of a straight line connecting the listener's ear canal openings when positioned in the listening point in the sound field.

[2] Occasional extraneous noise exceeded the permissible ambient levels quoted in ISO 8253-2.

The FAAF test was played from a Sound Blaster 16-bit sound card, at a sample rate of 20 kHz, routed via a Grayson-Stadler GSI 61 clinical audiometer to a Fostex 6301B loudspeaker. It was presented with steady noise that was filtered to give a long-term spectrum similar to the key words. The noise spectrum is also shown in Appendix A. The speech and noise were presented from the same loudspeaker located at an azimuth of 0^0 and a distance of 1.5 m from the reference point.

The speech and the noise were calibrated at the reference point by measuring the overall SPL of two corresponding wave files: the speech file contained the concatenated key words[3] and the other contained the filtered steady noise. A correction factor was then applied to the audiometer dial reading to compensate for the difference between the dial reading and the measured SPL. Subjects were instructed to sit comfortably facing the loudspeaker with their head positioned directly under the calibration marker. Head position was monitored visually throughout each test session. The subjects responded by selecting a key word on a touch-screen monitor. In the first experiment, the correct response flashed on the monitor giving feedback concerning the correct response. This was considered helpful since it might speed up any practice effect and reduce the number of test sessions required before subjects were fitted with the hearing instrument. The feedback was also considered useful because it may maintain subject motivation. This visual feedback was removed for experiments two and three; the explanation for this decision is given in section 4.5.

The overall level of the key words used in the first two experiments was set to 60 dB SPL[4]. The level was changed in the third experiment for the reasons given in section 5.1.3. A fixed SNR was used throughout testing but, first, this had to be determined for each subject. The FAAF test was presented monaurally using the adaptive strategy described by Lutman and Clark [1986] to estimate the signal-to-noise ratio [SNR] required for each subject to obtain a score of 71% when aided. [In some subjects it was found to be just as efficient to present FAAF words at a few predetermined SNRs in order to meet the criterion performance level]. The noise commenced 300 ms before the start of the carrier phrase and continued until 300 ms

---

[3] The key words had previously been adjusted in level to equalise intelligibility.

[4] When the level of the FAAF test was checked before commencing the third experiment, the level of speech was found to be 1-2 dB less than expected. This means that the speech level used in the earlier experiments may have been nearer 58-59 dB SPL.

after the end of the carrier phrase. The noise onsets and offsets were controlled by linear ramps with a duration of 50 ms. The SNR was defined as the difference between the overall SPL of the key words and the overall SPL of the filtered steady noise. The subject then practised the test, with the investigator adjusting the SNR to maintain the score around 71%. This typically involved increasing the FAAF-shaped noise by 1-3 dB over the course of several practice sessions. Performance was measured during the study at the fixed SNR. Subject performance was assessed as the number of words correctly identified and was expressed as percent correct. Since the study involved multiple presentations of the speech material, a random numbers table was employed to generate sets of unique ordering of the 80 items.

3.4 Procedures

Each subject attended for several sessions [usually between 4-6 with each lasting approximately two hours] before obtaining his or her hearing instrument. This time was spent tailoring the frequency response of the hearing instrument and determining the SNR required for presentation of the speech material. It also enabled the subject to become familiar with the testing procedure and materials. Once the subjects had completed these practice sessions, they were fitted with the hearing instrument for everyday use.

The number of test sessions, and the interval between sessions, varied across experiments; details specific to each experiment are provided in later sections. Each test session commenced with a general discussion about use and progress with the hearing instrument since the previous visit. The subjects returned a diary that included a section for self-reported hours of daily use of the hearing instrument. The categories available to the subject were less than 4 hours, 4-8 hours, 8-12 hours and greater than 12 hours. Subjects were instructed at the beginning of the study that they would be required to wear the hearing instrument for at least 6-8 hours per day. The daily use of the hearing instrument was summarised by noting the most frequently selected category between each test session. Although subjects tend to over-estimate their daily use of the hearing instrument [Brooks [1981] this is marginal if they receive extensive support.

Next, a listening check was made on the instrument and then electroacoustic tests were carried out to ensure that hearing instrument performance did not change. The tolerance used was within ±4 dB of the full-on gain [measured with an input level of 50 dB SPL] and OSPL90 measured at the start of the study. In addition, inter-modulation distortion was assessed using a composite speech-weighted input signal. With the gain control at the fixed target gain setting, the input level was increased until the response curve began to break up, reflecting inter-modulation distortion. None of the instruments tested failed the electroacoustic tests and none were operating with observable inter-modulation distortion for the settings and input levels used in the present study. Coherence measurements were made with a broadband speech-shaped signal in order to obtain an overall measure of distortion. There was no change in coherence over the range 60-90 dB SPL. One subject in the final experiment reported a non-functioning device between test sessions and a replacement model was fitted the following day.

Subjects performed the FAAF test on each test session. The first FAAF list of 80 words was used as a practice run on each test session but was not scored. Two FAAF lists [i.e., 160 words] were used for every test condition. All testing was performed monaurally with the non-test ear plugged and muffed. Benefit scores were obtained by subtracting the unaided score from the aided score. The complete set of FAAF tests took 90-120 minutes to complete. A 20-minute comfort break was provided during the test session.

After the last test session was completed, and usually on a different day, pure tone audiometry was repeated. This was to check for any change in hearing sensitivity that could confound the interpretation of the test results. For example, aided performance could decrease if the hearing impairment increased during the course of the study.
It took approximately 12 months to recruit subjects and complete the data collection for each of the first two experiments and 6 months for the third experiment.

3.5 Statistical analysis

The data were inspected before analysis to confirm it was appropriate to use parametric tests. Statistical analysis consisted primarily of a repeated-measures analysis of variance [ANOVA] using SPSS version 10.0, to determine if performance changed significantly over the course of the study. Factors are transformed into contrasts for analysis with the default contrast of polynomial used for within-subject factors. The factors that were treated as repeated measures were: ear, time and gain setting [in experiment three the gain setting was fixed but a third repeated measure was presentation level]. An acclimatisation effect would result in a differential change over time between the fitted ear and the not-fitted control ear. This should result in a statistically significant interaction between ear and time. The degrees of freedom were modified using the Greenhouse-Geisser correction when there was a statistically significant deviation from sphericity on Mauchly's test. The raw data from all experiments are listed in Appendix D and a summary of the statistical analysis is given in Appendix E.

CHAPTER FOUR
PERFORMANCE OVER A SIX-MONTH POST-FITTING PERIOD

4.0 Experiment One

Of all the studies that have demonstrated an auditory acclimatisation effect, the most dramatic of these was the study reported by Gatehouse [1992]. His study fitted new subjects with hearing instruments that provided more gain than any other study. This finding is consistent with the hypothesis that the subjects learnt to extract additional information from the newly audible, behaviourally important, speech signal resulting in an improvement in performance over time. Thus, the Gatehouse study was used as a starting point for establishing a methodological framework for demonstrating auditory acclimatisation.

Gatehouse tested a small number of subjects [n=4] extensively on one outcome measure [FAAF] over a post-fitting period of 12 weeks. The objective of the present study was to use a larger number of subjects, with a similar degree and configuration of hearing impairment, to extend the range of outcome measures to include sentence material and self-report measures and to extend the post-fitting time period to 24 weeks. The latter point was included because the duration required for acclimatisation to reach an asymptotic level has yet to be established: most previous studies have terminated at around 12 weeks post-fitting with none extending beyond 18 weeks. With the exception of Horwitz and Turner [1997], few studies have strictly controlled for the effect of changes in gain on the outcome data. This may have obscured the extent to which improvements over time were due to acclimatisation effects or to increases in speech audibility. Thus, in the present study, the subjects were tested at two gain conditions: fixed target gain [for a recognised prescription method] and user-adjusted gain: the former was held constant while the latter was adjusted to the users preferred gain setting on each session. Since learning is likely to be related to length of experience, the subjects were asked to record the hours of daily use that was made of the hearing instrument. Finally, the intention was to gain experience with a crossover design since this may have been used in future experiments when comparing hearing instruments or signal processing strategies; however, for technical reasons [see later], the difference between the two selected gain/frequency response settings was relatively small. Thus, although a randomised, double blind[1], crossover design was used, the subjects experienced only a small change in frequency response over the course of the study. The relevant features of the design were:

1. a homogeneous group of subjects with a hearing impairment consistent with those used in successful acclimatisation studies was used since the generality of acclimatisation for different degrees and configurations of hearing impairment is not known
2. a larger group of subjects was used than most previous studies
3. subjects were fitted with the same model of linear hearing instrument since it is not known if acclimatisation effects are the same for different amplification and processing strategies
4. a recognised prescription target was used to ensure that there was a change in audibility
5. testing was undertaken at both a fixed target gain and a user gain setting
6. the hours of use with the hearing instrument were recorded since acclimatisation is likely to be related to experience [i.e., more use gives more opportunity to learn and improve]
7. speech recognition tests were configured so that performance would not be obscured by a ceiling effect
8. a speech recognition test [FAAF] that is sensitive to changes in high frequency audibility was used
9. additional outcome measures that have a high face validity were included
10. measurements were made over a longer post-fitting time period in order to determine the full time course of acclimatisation

The aim of the study therefore, was to measure benefit in new hearing instrument users, at different gain settings using three outcome measures, over a post-fitting period of six months. The null hypothesis was that there would be no change in benefit over time.

---

[1] The experiment was free from bias by the subject and the investigator because the same programmable hearing instrument was used throughout the study and an independent tester was responsible for the setting and monitoring of the electroacoustic performance of the hearing instrument.

The alternative hypothesis was that there would be an increase in benefit over time. Methodological issues that were not covered in the general methodology chapter are provided in the next section.

4.1 Additional methodology for experiment one

Sixteen subjects were recruited [9 male, 7 female] with a mean age of 74 years [± 5.4]. Six subjects were fitted in the right ear and ten subjects were fitted in the left ear. All subjects completed the study.

The Desired Sensation Level fitting method [DSL; Seewald *et al.*, 1993] was used for target gain/frequency response values. This is one of the most popular prescription methods [Martin *et al.*, 1998]. It was selected because it prescribes more high frequency gain than most other prescription methods and this should emphasise the difference between the aided and unaided condition.

The hearing instrument was matched to the DSL real-ear aided response [REAR] target using the DSL protocol on the Audioscan RM500 probe-tube microphone system. This involves deriving the REAR by measuring the subject's real-ear to coupler difference [RECD] acoustic transform and adding it to the 2-cc coupler value. The hearing instrument response was tailored using the three-way audio filter and a combination of venting, damping and horn tubing. The greatest difficulties involved matching the 4 kHz target as well as removing the resonance peaks in the mid-frequencies, as is commonly found in clinical practice. The maximum output of the hearing instrument was set close to the subject's uncomfortable loudness level and modified if the subject reported loudness discomfort. The fitting was entered into memory one of the hearing instruments since this was the default memory used by the subjects in their home environment.

The target REAR and the actual REAR for each subject were converted to REIG in order to calculate the Speech Intelligibility Index [SII; ANSI R1997]. The SII was used to provide a guide to the change in audibility before and after amplification. The SII is described in Appendix F along with a worked example of how to calculate the SII. The REIG was obtained by subtracting the REUR reported by Shaw and Vaillancourt [1985] from each subject's REAR. The mean REIG values for the fitted ear are shown in Figure 4.1 and summarised numerically in Table 4.1. There is good agreement between the target and actual fitted response. Also shown is the alternative frequency response that was used in the crossover design. A technical fault meant that the difference between the two frequency responses at the high frequencies was not as large as originally intended. This was because of a fault with the Rastronics RM2000 probe-tube microphone system that was used for measuring this second frequency response. It was not until all subjects had commenced the study that this error was detected. Eight subjects were fitted with the DSL prescription and eight subjects were fitted with the alternative prescription. This was reversed after 12 weeks. The main difference at crossover was that there would be a small change in overall gain although this was minimal at the high frequencies where amplification was greatest.

Table 4.1 Mean real ear insertion gain [in decibels] for the fitted ear. The target and the fitted values are given for both prescriptions.

Frequency [Hz]	500	750	1000	1500	2000	3000	4000
			DSL				
Target	15	18	22	24	29	32	34
Fitted	12	18	23	26	29	33	33
			Alternative response				
Target	9	15	20	22	22	23	23
Fitted	5	13	19	25	28	31	27

In addition to the FAAF test, the opportunity was taken to include the Bamford-Kowal-Bench [BKB] sentences since this test material gives a more valid indication of communication in natural listening conditions. A copy was obtained on compact disc from the MRC Institute of Hearing Research. The BKB sentences measure the accuracy with which key words can be identified in fluently spoken 'everyday-type' sentences [Bench *et al.*, 1979]. The standard BKB sentences consist of 21 lists, each having 50 key words [three or four per sentence] such as He's holding his nose. The BKB sentence lists are reproduced in Appendix B. Subjects were instructed to repeat the sentence immediately it was presented to them, even if they only heard part of the sentence. The key words were scored as percent correct. The BKB sentences were presented at 60 dB SPL [with corrections to each sentence for equal intelligibility] in the presence of filtered noise with the same long-term spectrum as the BKB sentences.

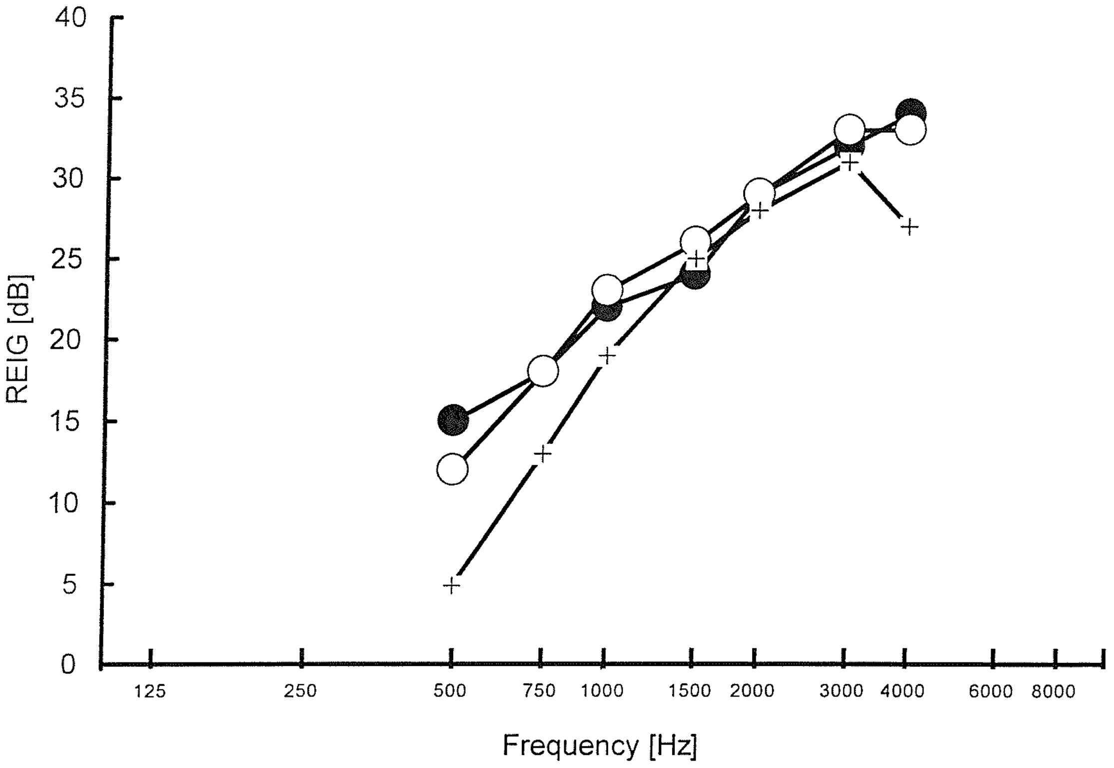


Figure 4.1. Mean real-ear insertion gain for the fitted ear. The target values for DSL are shown as solid circles and the match is shown as open circles. The alternative fitted response is shown as crosses.

The overall level of the noise was set to be the same as the BKB sentences; this meant that that the noise could be used to calibrate the BKB sentences. An adaptive strategy was used to determine the SNR for criterion performance of around 71%. A random numbers table was used to determine the order of presentation of the sentence lists.

Self-report was measured using a modified version of the Glasgow Hearing Aid Difference Profile [GHADP]. This is similar to the Glasgow Profile of Hearing Aid Benefit that was developed by Gatehouse [1999] except questions on benefit and satisfaction were rephrased to make a comparison with an alternative period of time instead of with the unaided condition. It was used to assess changes in benefit and satisfaction over time. The profile has scores for pre-intervention disability and handicap, and post-fitting hearing instrument use, residual disability, and changes in benefit and satisfaction [see appendix C]. The scores are obtained over a combination of pre-specified listening circumstances as well as additional situations nominated by the subject. The additional benefit and satisfaction scales were modified to compare current benefit and satisfaction with those reported at the previous test session [i.e., any changes that had occurred since the previous visit]. The GHADP was completed on the day of [or just before] fitting in order to obtain the pre-fitting disability and handicap scores. The post-fitting scores were obtained at each post-fitting test session.

Subjects performed the FAAF test at each test session. Both ears were tested individually in the aided [fixed target gain setting] and unaided condition. The fixed target gain setting means that audibility with amplification was held constant across test sessions. Since this may not be representative of the subject's normal [and hence familiar] gain setting, the fitted ear was also tested at a second gain position that was designed to be more representative of the subject's normal gain setting. However, any change in performance may be confounded by a change in audibility; performance may increase if more gain is selected over time. The user-gain setting was determined for the fitted ear of each subject at each test session. The FAAF test was presented at the subject's fixed SNR and they were instructed to adjust the hearing instrument gain control, using a bracketing procedure, in order to find the preferred user-gain position. The investigator then measured the gain in an HA2 coupler at the preferred gain position. The subject performed the adjustment three times and the average of the three settings was used for the user-gain condition on that test session. The mean difference between the fixed-gain and the user-gain setting at 2 kHz in a 2-cc coupler is shown in Table 4.2. The mean user-gain was typically 7 dB below the fixed-gain setting and this was relatively stable from week three onwards.

Table 4.2. Mean difference in hearing instrument gain [in decibels] between the fixed-gain setting [i.e., best match to target] and the user-gain setting at 2 kHz in the 2-cc coupler as a function of post-fitting time. A positive value indicates the extent to which the fixed-gain exceeds the user-adjusted gain. One standard deviation is given in brackets.

Post-fitting time [weeks]	0	3	6	9	12	15	18	21	24
Fixed-gain – user gain [dB]	10 [6]	8 [7]	7 [5]	7 [6]	5 [6]	7 [5]	6 [7]	6 [6]	6 [6]

In order to reduce the number of repetitions of the BKB lists, sentence testing was performed on alternate visits [i.e., six-week intervals] and aided testing was performed at the fixed gain setting only. To reduce the variability associated with a single sentence list, four lists were used for each test condition. A random numbers table was used to determine the presentation order of the test conditions. No sentence list was repeated within the same test session.

The modified GHADP was completed using a paper and pencil response format on each test session. The investigator checked and clarified the subject's responses. This took around 15 minutes to complete and was carried out midway through each test session in an attempt to reduce subject fatigue during the speech recognition tests.

4.2 Results

Subjects were issued with a diary to record the hours of daily use of the hearing instrument. The categories available to the subject were: less than 4 hours, 4-8 hours, 8-12 hours or greater than 12 hours. The mean self-reported daily use of the hearing instrument as a function of post-fitting time is shown in Table 4.3. More than 90% of subjects reported using the hearing instrument for more than 4 hours per day; for virtually all intervals, at least half used their instrument for at least 8 hours per day.

The mean SII for the aided and unaided conditions are shown in Table 4.4. The SII was calculated using the octave band procedure. The calculations were based on the mean hearing threshold level, in decibels, in the fitted ear. The equivalent speech spectrum level of the concatenated FAAF words was based on an L_{eq} of 60 dB SPL. The equivalent noise spectrum of the FAAF-shaped noise spectrum was based on an L_{eq} of 58 dB SPL [i.e., a signal to noise ratio of +2 dB]. The band importance function for nonsense syllables was used in the calculations. The SII values show that the hearing instrument makes a substantial improvement to the audibility of the FAAF test when presented at 60 dB SPL.

Table 4.3. Mean self-reported daily use of the hearing instrument, in hours, as a function of post-fitting time. Subjects were issued with a diary that required them to indicate the hours of daily use of the hearing instrument. The categories available to the subject were: less than 4 hours, 4-8 hours, 8-12 hours or greater than 12 hours. The daily hours of use for each subject was taken as the most frequently ticked category over each of the three-week periods. The percentage of responses in each category is given in brackets [n=16].

Post-fit [weeks]	0-3	4-6	7-9	10-12	13-15	16-18	19-21	22-24	Mean
< 4 hours	1 [6]	1 [6]	1 [6]	2 [12]	0 [0]	2 [12]	1 [6]	1 [6]	1 [7]
4-8 hours	5 [31]	4 [25]	4 [25]	4 [25]	9 [56]	5 [31]	1 [6]	3 [19]	4 [27]
8-12 hours	8 [50]	9 [56]	8 [50]	7 [44]	7 [44]	8 [50]	9 [56]	7 [44]	8 [49]
> 12 hours	0 [0]	0 [0]	0 [0]	1 [6]	0 [0]	0 [0]	2 [13]	3 [19]	1 [5]
Missing data	2 [13]	2 [13]	3 [19]	2 [13]	0 [0]	1 [6]	3 [19]	2 [13]	2 [12]

Table 4.4. Mean Speech Intelligibility Index for the aided and unaided conditions in the test ear. The Speech Intelligibility Index was calculated using the octave band procedure. The calculations were based on the mean hearing threshold level, in decibels, in the fitted ear. The equivalent speech spectrum level of the FAAF test was based on an Leq of 60 dB SPL. The equivalent noise spectrum of the FAAF shaped noise spectrum was based on an Leq of 58 dB SPL [i.e., a signal to noise ratio of +2 dB]. The band importance function for nonsense syllables was used in the calculation.

	UNAIDED	TARGET	FIXED	USER GAIN (at each post-fitting test session)								
				0	3	6	9	12	15	18	21	24
DSL	0.126	0.519	0.519	0.419	0.454	0.471	0.471	0.503	0.471	0.487	0.487	0.487
Alt	0.126	0.456	0.519	0.343	0.399	0.418	0.418	0.453	0.418	0.436	0.436	0.436

4.2.1 Outcome measured with the FAAF test

The mean FAAF benefit, as a function of post-fitting time, derived using the fixed-gain aided scores is shown in Figure 4.2. The benefit score was derived from the difference between the aided [fixed-gain] and unaided speech recognition score. The filled circles correspond to the fitted ear and the open circles to the not-fitted control ear. The mean benefit score is around 14% and 12% in the fitted and not-fitted control ear respectively. The numerical data are given in Table 4.5. The mean change in benefit,

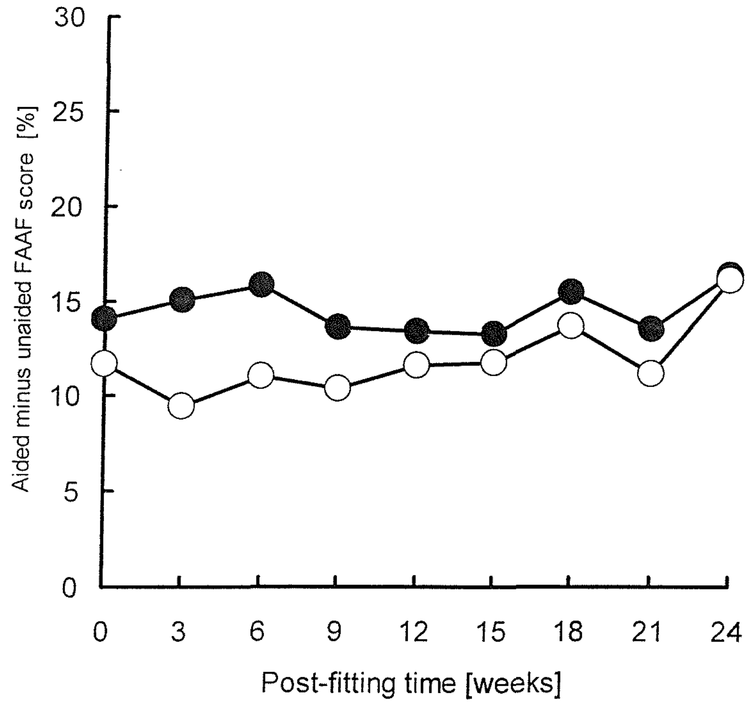


Figure 4.2. Mean FAAF benefit [aided minus unaided score] as a function of post-fitting time at the fixed-gain setting. The filled circles correspond to the fitted ear and the open circles to the not-fitted control ear.

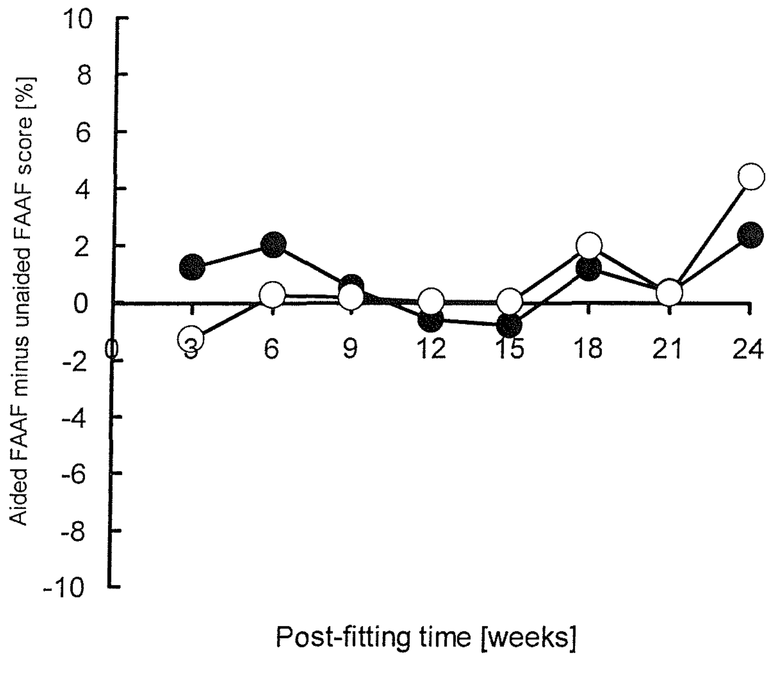


Figure 4.3. Mean change in FAAF benefit [aided minus unaided score] relative to the time of fitting at the fixed-gain setting. The filled circles correspond to the fitted ear and the open circles to the not-fitted control ear.

relative to the time of fitting, is shown in Figure 4.3 [and Table 4.6]. There is no apparent increase in benefit over time in the fitted ear.

Table 4.5. Mean FAAF benefit score [aided minus unaided score] as a function of post-fitting time at the fixed-gain setting. The number of subjects used to calculate each mean score varied from 13 to 16. One standard deviation is shown in brackets.

Time [wks]	0	3	6	9	12	15	18	21	24
Fitted ear	14.0	15.0	15.8	13.6	13.4	13.2	15.4	13.5	16.4
	[10.8]	[9.5]	[11.4]	[12.0]	[12.2]	[9.3]	[11.3]	[10.5]	[11.1]
Control ear	11.7	9.5	11.0	10.4	11.6	11.7	13.7	11.2	16.1
	[9.6]	[8.0]	[10.0]	[8.0]	[8.4]	[9.4]	[8.7]	[9.6]	[11.0]

Table 4.6. Mean change in benefit score relative to the initial benefit score for the fixed-gain setting. The number of subjects used to calculate each mean score varied from 13 to 16. One standard deviation is shown in brackets.

Post-fitting time [weeks]	3	6	9	12	15	18	21	24
Fitted ear	1.2	2.0	0.5	−0.6	−0.8	1.2	0.4	2.3
	[9.0]	[9.1]	[11.6]	[10.3]	[9.5]	[11.6]	[7.8]	[10.2]
Control ear	−1.3	0.3	0.2	−0.0	0.0	2.0	0.3	4.4
	[3.8]	[9.6]	[4.1]	[8.2]	[10.8]	[8.4]	[7.5]	[5.9]

One of the 16 subjects was excluded from of all the FAAF analysis because he missed three consecutive test sessions. The remaining data were smoothed [using linear interpolation] for four subjects who missed a single test session. This enabled them to be included in the repeated-measures analysis. A summary of the repeated-measures analysis of variance [ANOVA] is shown in Table I, Appendix E. Two factors were treated as repeated-measures: ear [fitted and not-fitted control] and post-fitting time [0, 3, 6, 9, 12, 15, 18, 21, 24 weeks]. The effects of ear, post-fitting time and the interaction between ear and post-fitting time were not statistically significant. The lack of interaction indicates that the ears were not differentially affected by post-fitting time. Thus, the null hypothesis was not rejected.

Results for the benefit scores obtained using the user-gain condition are shown in Figure 4.4 with the relative difference shown in Figure 4.5 [see also Table 4.7 and 4.8]. The filled circles correspond to the fitted ear and the open circles to the not-fitted control ear. There is a transient increase of 5% at six-weeks post-fitting in the fitted ear. The repeated-measures ANOVA revealed a statistically significant mean difference between ears, but not post-fitting time. There was no statistically significant interaction [see Table III, Appendix E] indicating that the ears were not differentially

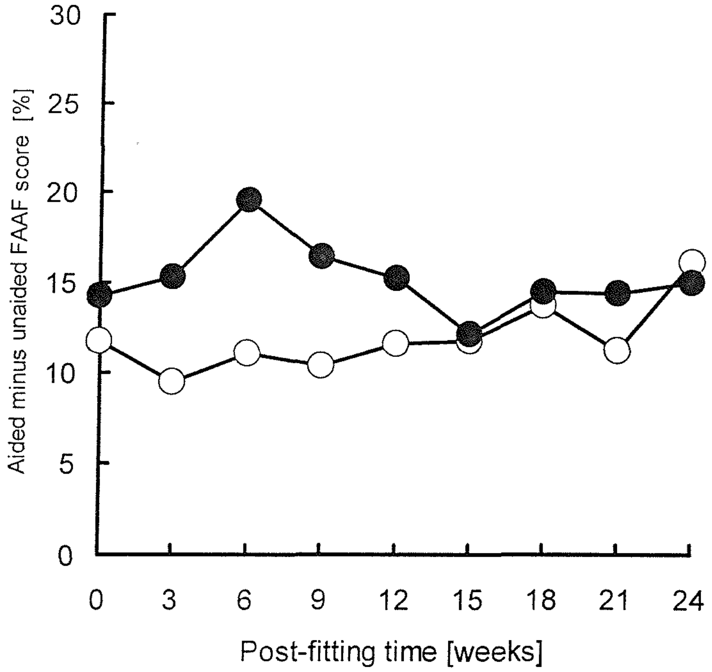


Figure 4.4. Mean FAAF benefit [aided minus unaided score] as a function of post-fitting time at the user-gain setting in the fitted ear and the fixed gain setting in the control ear. The filled circles correspond to the fitted ear and the open symbols to the not-fitted control ear.

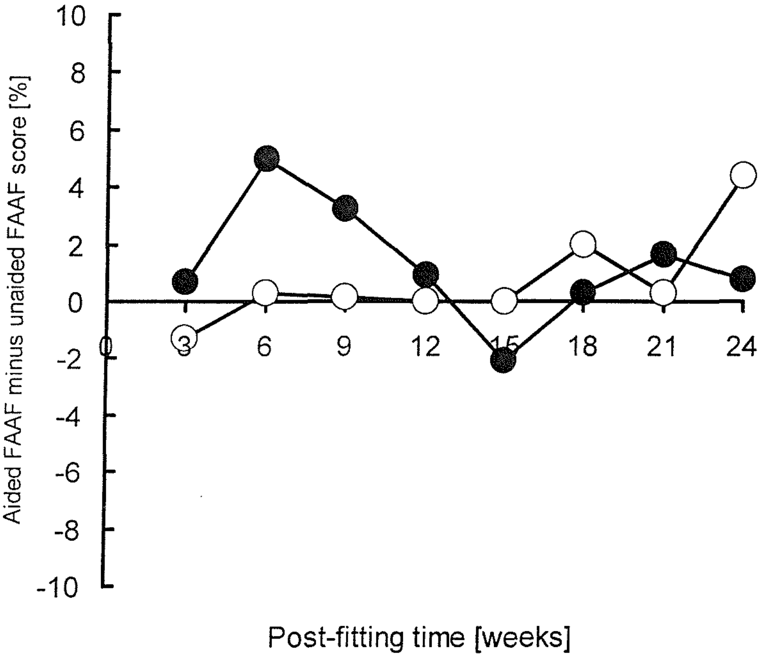


Figure 4.5. Mean change in FAAF benefit [aided minus unaided score] relative to the time of fitting at the user-gain setting in the fitted ear and the fixed gain setting in the control ear. The filled circles correspond to the fitted ear and the open symbols to the not-fitted control ear.

affected by post-fitting time. The mean difference between ears was examined for each test session using paired Student *t*-tests. The 0.05 significance level was adjusted to 0.006 using the Bonferonni correction to account for the nine comparisons. The results of the analysis are shown in Table III, Appendix E. The difference in benefit between ears at six weeks post-fitting was statistically significant.

Table 4.7. Mean FAAF benefit score [aided minus unaided score] as a function of post-fitting time in the fitted ear at the user-gain setting. The number of subjects used to calculate each mean score varied from 13 to 16. One standard deviation is shown in brackets.

Time [weeks]	0	3	6	9	12	15	18	21	24
	14.2	15.3	19.6	16.4	15.2	12.2	14.4	14.4	15.0
	[11.4]	[10.3]	[11.0]	[9.9]	[10.5]	[10.0]	[10.0]	[10.5]	[12.4]

Table 4.8. Mean change in benefit score relative to the initial benefit score in the fitted ear for the user-gain setting. The number of subjects used to calculate each mean score varied from 13 to 16. One standard deviation is shown in brackets.

Post-fitting time [weeks]	3	6	9	12	15	18	21	24
	0.7	5.0	3.2	0.9	-2.1	0.3	1.6	0.8
	[10.3]	[9.5]	[11.0]	[9.4]	[10.9]	[10.3]	[10.4]	[11.8]

Table 4.9. Mean FAAF speech recognition score, in percent correct, as a function of post-fitting time in the unaided condition. The number of subjects used to calculate each mean score varied from 13 to 16. One standard deviation is given in brackets.

Time [weeks]	0	3	6	9	12	15	18	21	24
Fitted ear	56.3	56.3	55.6	58.3	59.2	60.9	59.6	59.7	58.8
	[10.2]	[10.0]	[9.9]	[10.1]	[10.0]	[10.0]	[8.8]	[10.1]	[9.4]
Control ear	56.1	58.3	60.0	58.8	58.3	61.6	59.8	62.1	57.1
	[9.3]	[8.6]	[9.0]	[7.8]	[8.7]	[9.7]	[9.2]	[10.4]	[11.0]

The primary focus of the study was on hearing instrument benefit, derived by subtracting the unaided speech recognition score from the aided score. It was of interest to examine benefit in terms of the aided and unaided scores. The mean unaided scores are shown in Figure 4.6. The filled circles correspond to the fitted ear and the open circles to the not-fitted control ear. The numerical data are presented in Table 4.9. The mean unaided scores show a general trend of increasing performance over time. The results of the repeated-measures ANOVA are summarised in Table IV, Appendix E. The two factors that were treated as repeated-measures were ear [test and control] and post-fitting time [0, 3, 6, 9, 12, 15, 18,21 and 24 weeks]. The data were smoothed using linear interpolation to allow subjects with occasional missing data to be included in the analysis. The observed increase in scores over time was statistically significant

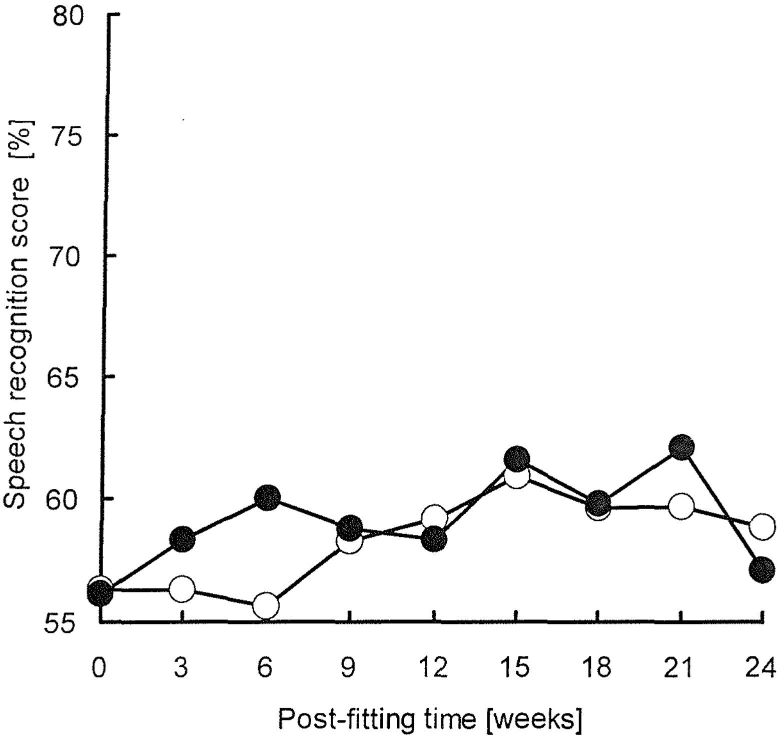


Figure 4.6. Mean unaided BKB speech recognition scores in percent correct as a function of post-fitting time. *The filled circles correspond to the fitted ear and the open circles to the not-fitted control ear.*

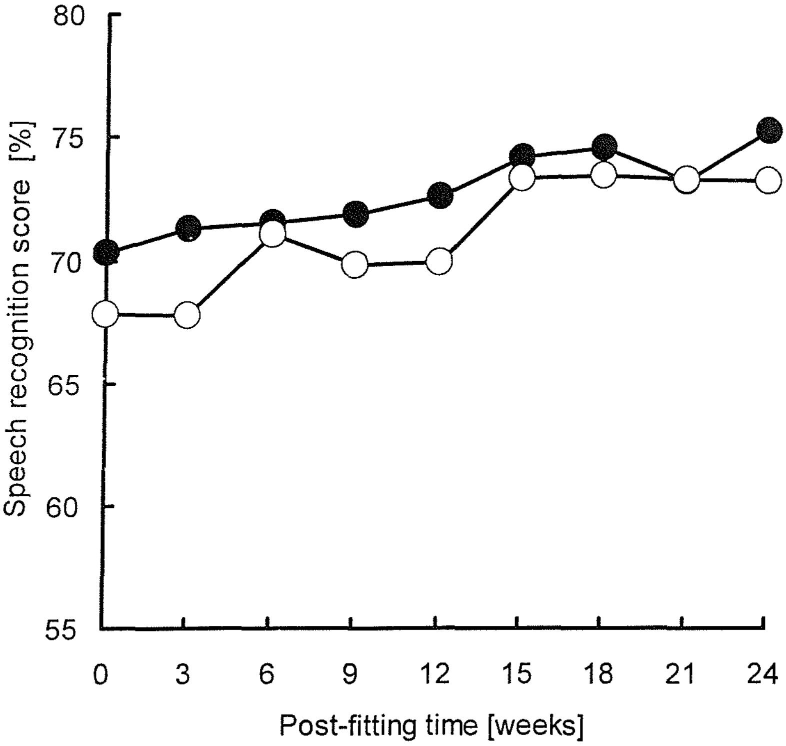


Figure 4.7. Mean aided speech recognition scores in percent correct as a function of post-fitting time for the fixed-gain setting. The filled circles correspond to the fitted ear and the open circles to the not-fitted control ear.

for ear. It was the linear component of the orthogonal polynomial breakdown that was significant [F(1,14)=11.2; p<0.01]. Since this occurs for both ears and without amplification, the most likely explanation is a learning effect due to repeated exposure to the test material.

The mean aided scores at the fixed-gain setting are shown in Figure 4.7 and tabulated in Table 4.10. The filled circles correspond to the fitted ear and the open circles to the not-fitted control ear. There is a general trend of increasing performance over time and this was statistically significant on a two-factor repeated-measures ANOVA [Table V, Appendix E]. It was only the linear component that was significant [F(1,14)=18.4; p<0.01]. This finding is consistent with the practice effect observed on the unaided condition. The mean aided scores at the user-gain setting are shown in Figure 4.8 and tabulated in Table 4.10. The filled circles correspond to the fitted ear and the open circles to the not-fitted control ear. These results were used to derive the user-gain benefit data for the test ear as shown in Figure 4.4. There is a jump in performance in the fitted ear from six-weeks onwards whereas the increase in the control ear is more gradual. The repeated-measures ANOVA confirmed that the changes observed over time were statistically significant [Table V, Appendix E]. Only the linear component was significant [F(1,14)=15.8; p<0.01]. In addition, there was a significant difference between ears as well as a significant interaction between ear and post-fitting time. This time the orthogonal polynomial breakdown revealed that it was only the fifth order contrast that was significant [F(1,14)=8.0; p=0.01].

A one-factor repeated-measures ANOVA was performed on the data from each ear separately to look at the simple main effect of post-fitting time. The 0.05 significance level was adjusted to 0.025 using the Bonferonni correction to account for the two comparisons. The results are summarised in Table VI, Appendix E. The change in score over time was statistically significant for each ear and for both the aided and unaided condition. It was only the linear component of the polynomial breakdown that was significant for each of the simple main effects. The mean difference between ears on each test session was examined using multiple paired Student *t*-tests. The 0.05 significance level was adjusted to 0.006 using the Bonferonni correction to take account of the nine comparison tests. The results are shown in Table VIII, Appendix E. There

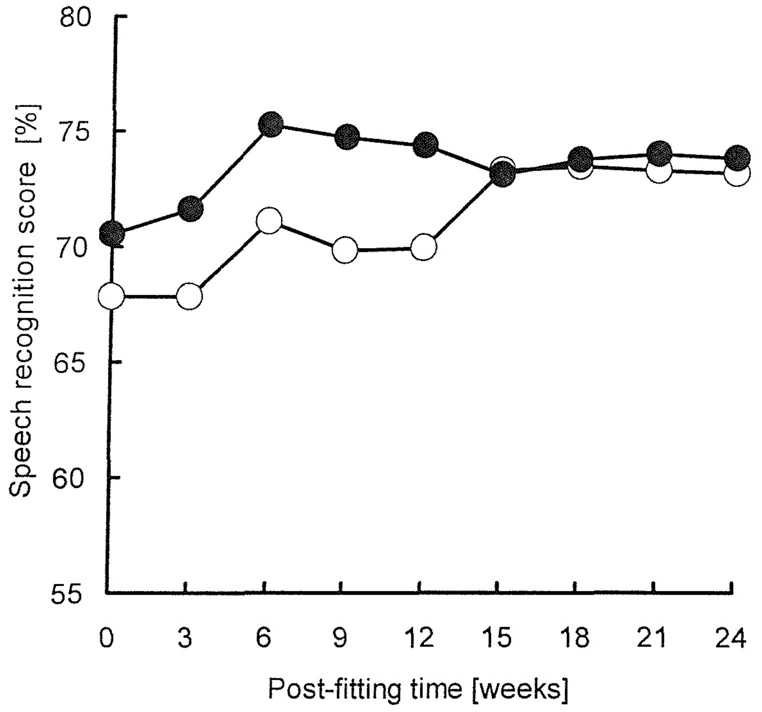


Figure 4.8. Mean aided speech recognition scores in percent correct as a function of post-fitting time for the user-gain setting. The filled circles correspond to the fitted ear and the open symbols to the not-fitted control ear.

is a 4-5% difference in mean scores between ears from 6-12 weeks but this difference was not statistically significant when the significance level was corrected using the Bonferonni correction for multiple paired comparisons.

On every test session, the FAAF test was also performed with the subjects using the second frequency response [at the fixed-gain setting]. The results were very similar to those obtained with the subjects own frequency response at the fixed-gain setting. This is not surprising given the similarity between the two frequency responses.

Table 4.10. Mean FAAF speech recognition score in percent correct as a function of post-fitting time in the aided condition. The number of subjects used to calculate each mean score varied from 13 to 16. One standard deviation is given in brackets.

Time [weeks]	0	3	6	9	12	15	18	21	24
Fitted ear, fixed-gain	70.3	71.3	71.5	71.9	72.6	74.1	74.6	73.2	75.2
	[2.5]	[3.4]	[3.8]	[5.8]	[5.8]	[5.0]	[6.7]	[6.0]	[5.0]
Fitted ear, user-gain	70.5	71.6	75.2	74.7	74.3	73.1	73.8	74.0	73.8
	[4.7]	[5.4]	[4.8]	[5.9]	[5.3]	[6.4]	[5.5]	[6.8]	[7.4]
Control ear, fixed-gain	67.8	67.8	71.0	69.8	69.9	73.3	73.4	73.3	73.2
	[4.0]	[4.6]	[4.8]	[2.8]	[4.9]	[5.5]	[4.6]	[5.7]	[4.5]

Since two FAAF wordlists were used for each test condition, it was possible to investigate the test-retest difference. Table 4.11 shows the mean test minus retest difference as a function of post-fitting time. The mean difference was generally very close to zero with a standard deviation of around 5%.

Table 4.11. Reliability of FAAF recognition score [test minus retest] as a function of post-fitting time. The table shows the mean difference with one standard deviation given in brackets. The speech recognition scores for the unaided and aided conditions varied with ear and post-fitting time but were typically 59% and 71% respectively.

Post-fitting time [weeks]	0	3	6	9	12	15	18	21	24
			Control ear						
Unaided	−1.6	−0.7	0.5	−0.5	1.6	0.6	0.1	0	4.3
	[5.1]	[5.5]	[6.0]	[4.5]	[5.6]	[4.6]	[4.7]	[6.1]	[4.7]
Aided	2.5	0.6	0.5	0.9	−1.8	−0.4	−1.8	−1.0	1.6
	[5.2]	[6.0]	[7.6]	[5.0]	[5.0]	[6.0]	[7.7]	[5.4]	[5.6]
			Test ear						
Unaided	−2.6	0.9	−0.8	1.9	2.9	−1.9	−2.0	0	−1.6
	[6.3]	[5.8]	[6.3]	[9.4]	[7.3]	[8.0]	[6.9]	[6.6]	[7.4]
Aided	−0.3	−0.3	−2.3	−1.4	−1.6	0.4	−0.3	0.4	0.2
	[6.3]	[6.0]	[6.3]	[4.7]	[4.6]	[5.6]	[5.4]	[5.8]	[8.1]

In summary, the lack of statistically significant interaction shows that the ears do not change differentially over time. Thus, the null hypothesis cannot be rejected. The only statistically significant difference in mean benefit occurred at six weeks post-fitting but only when the benefit score was derived using the user-gain condition. There was a statistically significant increase over time in the mean aided and unaided scores for both ears.

4.2.2 Outcome measured with the BKB sentences

Figure 4.9 shows mean BKB benefit as a function of post-fitting time. Unlike the FAAF test, the BKB test was carried out at six-week intervals and only with the hearing instrument at fixed-gain. The filled circles correspond to the fitted ear and the open circles to the not-fitted control ear. The benefit score was derived from the difference between the aided and unaided speech recognition scores. The numerical data are presented in Table 4.12. The mean benefit score is around 25%. There is no increase in benefit in the fitted ear over time.

Table 4.12. Mean BKB benefit score [aided minus unaided] in percent correct as a function of post-fitting time. Results are given for the fitted and control ear. The number of subjects used to calculate each mean score was 13 for week 0, 14 for weeks 12 and 15 for weeks 18 and 24. One standard deviation is given in brackets.

Post-fitting time [weeks]	0	6	12	18	24
Test ear	26.9 [15.6]	27.3 [17.5]	23.0 [16.3]	26.0 [14.9]	21.8 [17.4]
Control ear	20.9 [17.1]	18.9 [16.3]	26.3 [12.0]	25.3 [17.0]	26.4 [17.8]

A two-factor repeated-measures ANOVA was performed on the data. The two factors that were treated as repeated-measures were ear and post-fitting time. Two of the 16 subjects were excluded from the analysis because they missed the first or last test session. The remaining data were smoothed using linear interpolation on four subjects who missed a single test session as this allowed them to be included in the repeated measure analysis. A summary of the repeated-measures ANOVA is shown in Table VIII, Appendix E. The effects of ear, post-fitting time and the interaction between ear and post-fitting time were not statistically significant. The lack of interaction indicates that the change in performance of the fitted ear, over time, was similar to the performance in the not-fitted control ear.

The aided and unaided scores that were used to derive the benefit scores were examined. The mean unaided scores are shown in Figure 4.10 and the mean aided

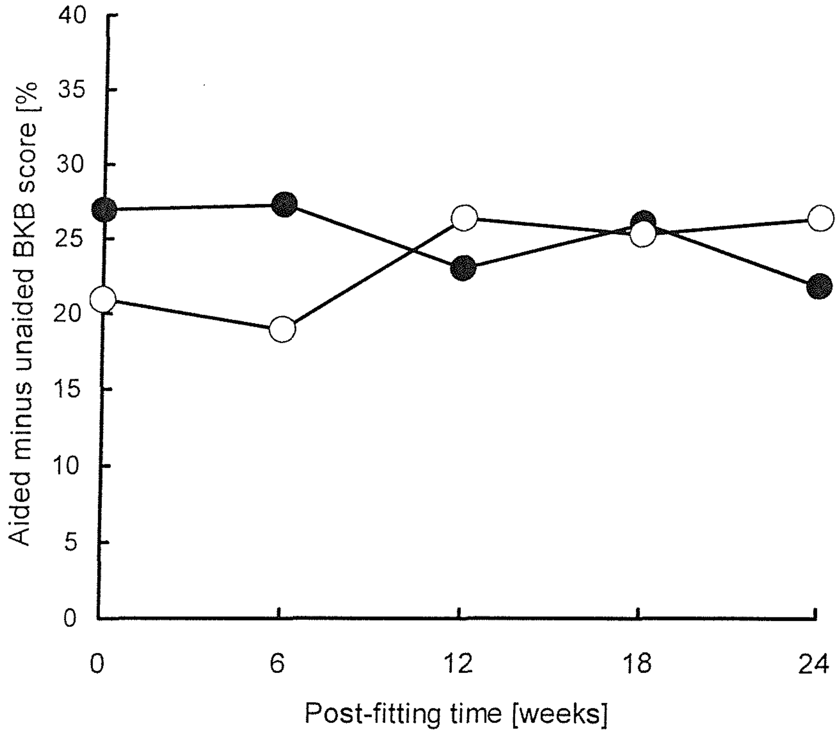


Figure 4.9. Mean BKB benefit [aided minus unaided score] as a function of post-fitting time. The filled circles correspond to the fitted ear and the open symbols to the not-fitted control ear.

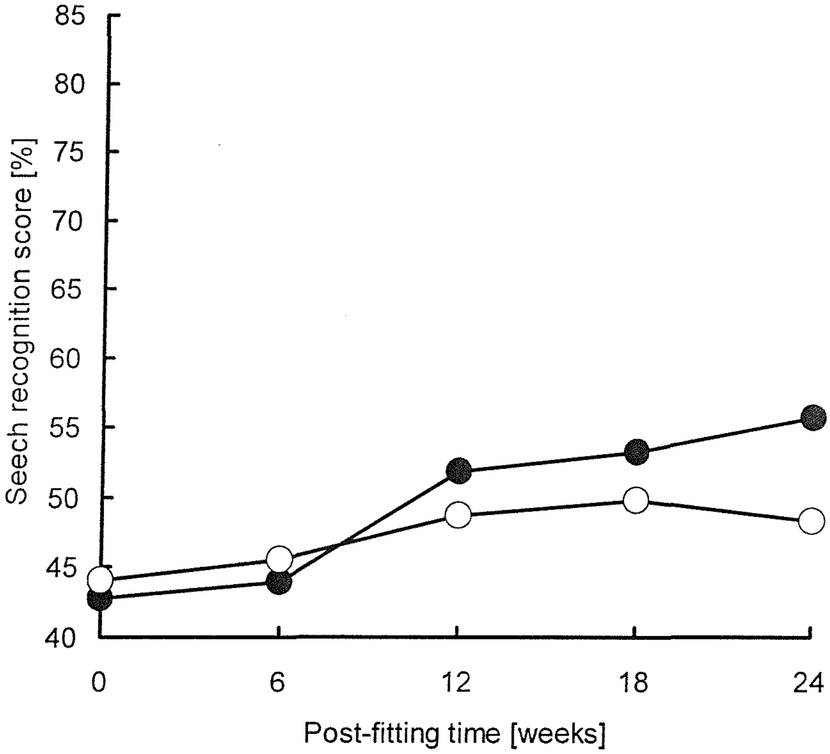


Figure 4.10. Mean unaided speech recognition scores in percent correct as a function of post-fitting time. The filled circles correspond to the fitted ear and the open circles to the not-fitted control ear.

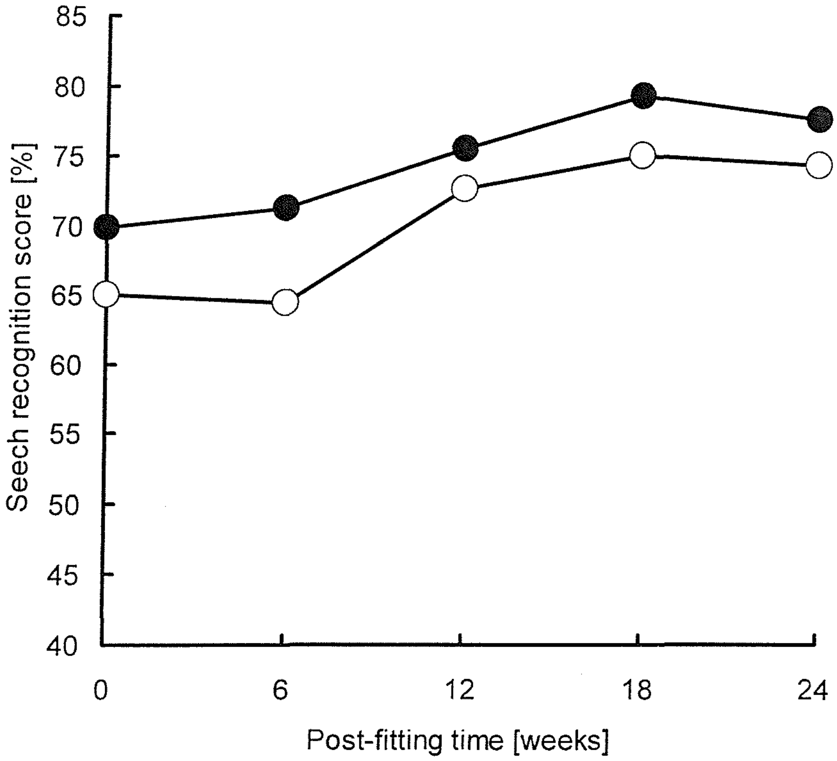


Figure 4.11. Mean aided BKB speech recognition scores in percent correct as a function of post-fitting time. The filled circles correspond to the fitted ear and the open circles to the not-fitted control ear.

scores are shown in Figure 4.11. The filled circles correspond to the fitted ear and the open circles to the not-fitted control ear. The numerical data are given in Table 4.13. The component scores show a general trend of increasing performance over time. Since this occurred for both ears and for the aided and unaided conditions, the most likely explanation is a practice effect. The data were smoothed, as before, and analysed using a two-factor repeated-measures ANOVA. The two factors that were treated as repeated-measures were ear and post-fitting time. The results of the analysis are summarised in Table IX and X, Appendix E, for unaided and aided performance respectively. The observed increase in scores over time was statistically significant for the aided condition [linear component, F(1,12)=6.8; p=0.02] but not for the unaided condition when the ears were grouped together. A one-factor repeated-measures ANOVA was performed on the aided scores of each ear separately to look at the simple main effect of post-fitting time. The 0.05 significance level was adjusted to 0.025 using the Bonferonni correction to account for the two comparison tests. The results of the analysis are summarised in Table XI, Appendix E. The change in score over time was statistically significant for the not-fitted control ear [linear component, F(1,12)=8.0; p=0.02] but not for the fitted ear. In summary, there is evidence of improving performance in the aided and unaided test condition. There was no differential increase in the fitted ear relative to the not-fitted control ear.

Table 4.13. Mean BKB score in percent correct as a function of post-fitting time. Data are presented for the aided and unaided conditions. The number of subjects used to calculate each mean score was 13 for week 0, 14 for weeks 4 to 12 and 15 for weeks 18 and 24. One standard deviation is given in brackets.

Post-fitting time [weeks]	0	6	12	18	24
		Test ear			
Unaided	42.8 [13.0]	44.0 [17.6]	51.9 [20.2]	53.2 [18.9]	55.7 [22.2]
Aided	69.9 [8.7]	71.3 [11.7]	80.2 [12.3]	79.3 [13.9]	77.6 [15.1]
		Control ear			
Unaided	44.1 [13.1]	45.5 [17.7]	48.7 [18.9]	49.8 [17.7]	48.3 [18.2]
Aided	65.0 [11.0]	64.4 [9.6]	72.6 [11.1]	75.0 [11.6]	74.3 [13.6]

4.2.3 Self-report using the modified GHADP

Table 4.14 shows the distribution of disability and handicap scores on the modified GHADP prior to hearing instrument fitting. A high score represents greater disability and handicap. The median values for disability and handicap are 50 and 42.7% respectively.

Table 4.14. Distribution of pre-fitting disability and handicap scores on the Glasgow Hearing Aid Difference Profile [n=16].

Percentile	Initial disability	Initial handicap
25th	40.4	35.6
50th	50.0	42.7
75th	56.9	59.3

The four post-fitting sub-scales are *Residual Disability, Use of Hearing Instrument, Additional Benefit* and *Additional Satisfaction.* The distribution of the sub-scale data show marked skewing. An example of this is illustrated in Figure 4.12 for the data collected at 12 weeks post-fitting. The negative tail on the *Use of Hearing Instrument* sub-scale means that a few subjects were using the hearing instrument less frequently than most subjects. The other sub-scales show a positive tail. This means that compared to most subjects, a small number of subjects had a higher *Residual Disability* but greater *Additional Benefit* and *Additional Satisfaction.*

The distribution of the scores on these sub-scales is given in Table 4.15 and the median values are plotted in Figure 4.13. The *Use of Hearing Instrument* sub-scale is typically around 90%. This means that the subjects report wearing the hearing instrument for 90% of the time they are in a listening situation that, prior to fitting, lead to hearing difficulty. The *Residual Disability* sub-scale is a measure of the extent to which things are better post-fitting compared to pre-fitting. This is typically 10-15% and does not show any particular pattern of change over time. The *Additional Benefit* sub-scale refers to the change in benefit that has occurred since the previous visit. A score of greater than 50% indicates that the subjects report more benefit whereas a score of less than 50% indicates less benefit. The median score falls from 67% at three weeks post-fitting to 50% from week fifteen onwards. This suggests that there was additional benefit up to three months post-fitting and this level of benefit was sustained until the end of the study. A similar interpretation applies to the *Additional Satisfaction* sub-scale. The scores fall from 70% at three-weeks post-fitting to 50% from week nine onwards. This suggests that there was additional satisfaction up to six weeks post-fitting and this level of satisfaction was maintained until the end of the study.

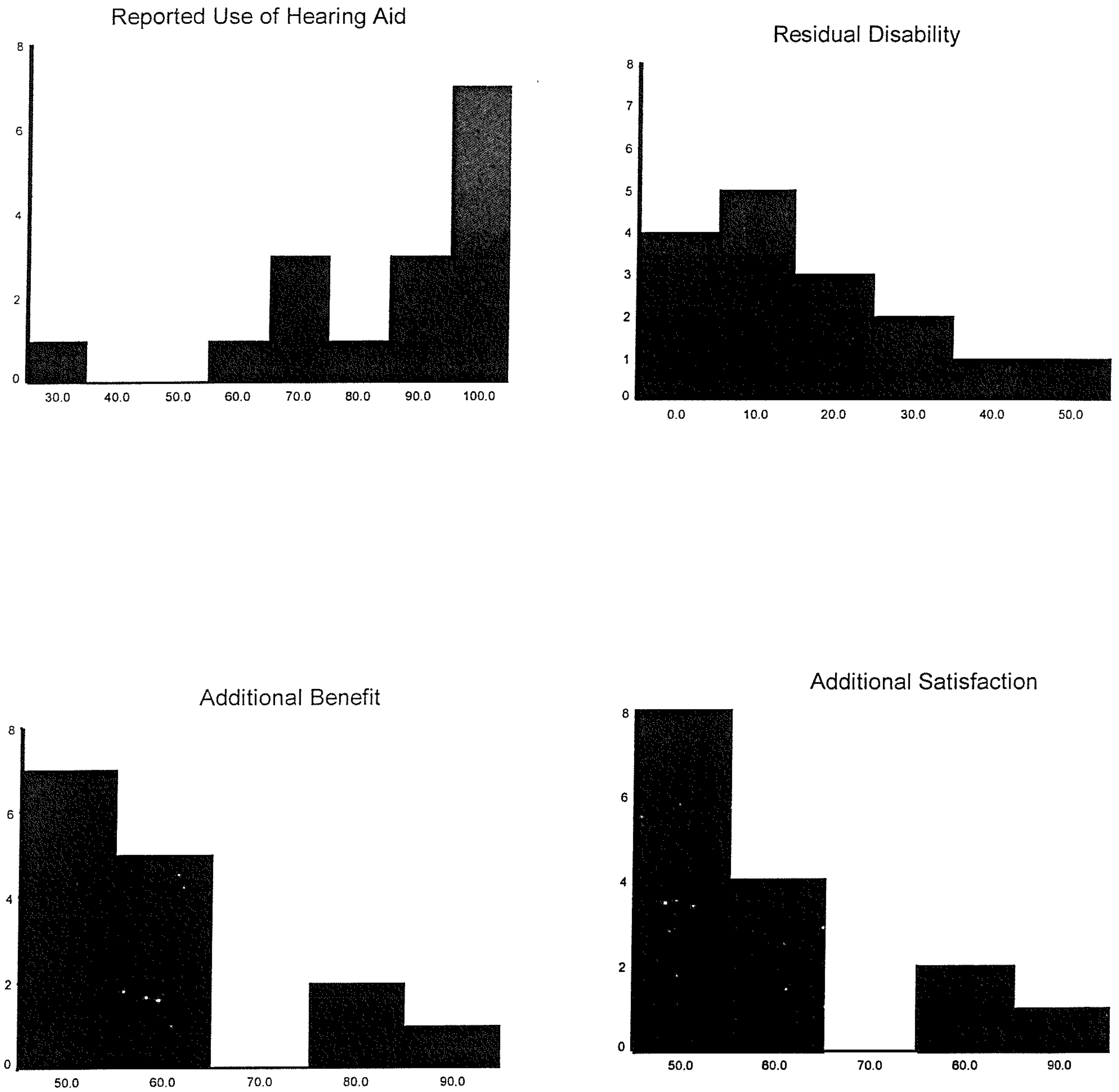


Figure 4.12. Distribution of Glasgow Hearing Aid Difference Profile sub-scales at 12-weeks post-fitting.

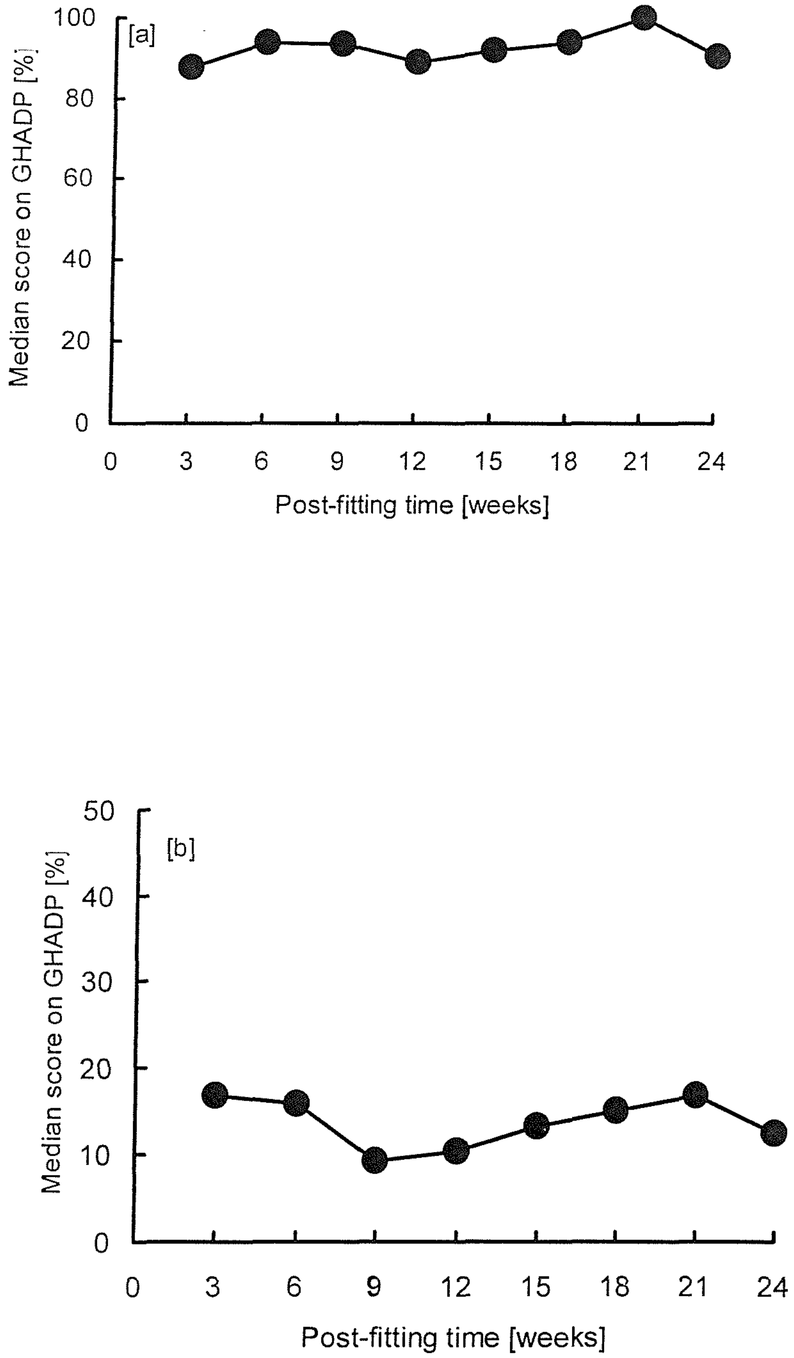
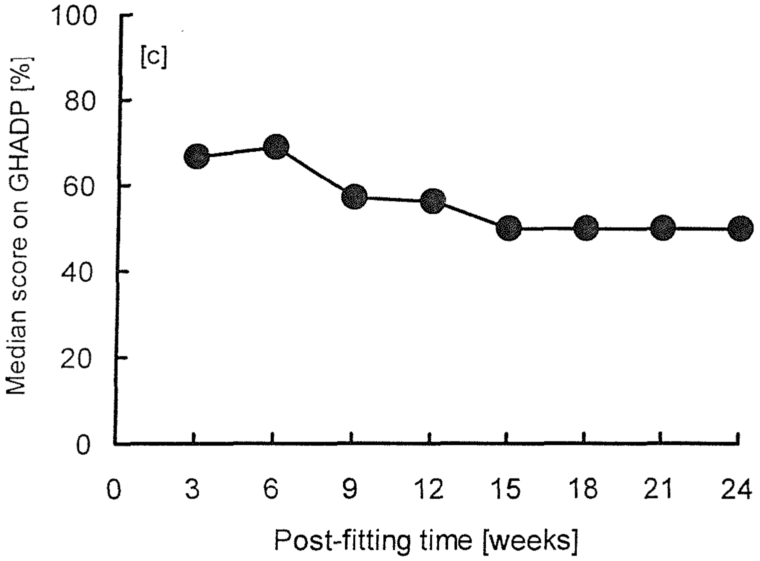


Figure 4.13. Median scores on the post-fitting sub-scales of the modified Glasgow Hearing Aid Difference Profile [n=13]. Use of Hearing Aid and Residual Disability are shown in [a] and [b] respectively. The scores for Additional Disability and Additional Satisfaction are shown in [c] and [d] respectively [overleaf].



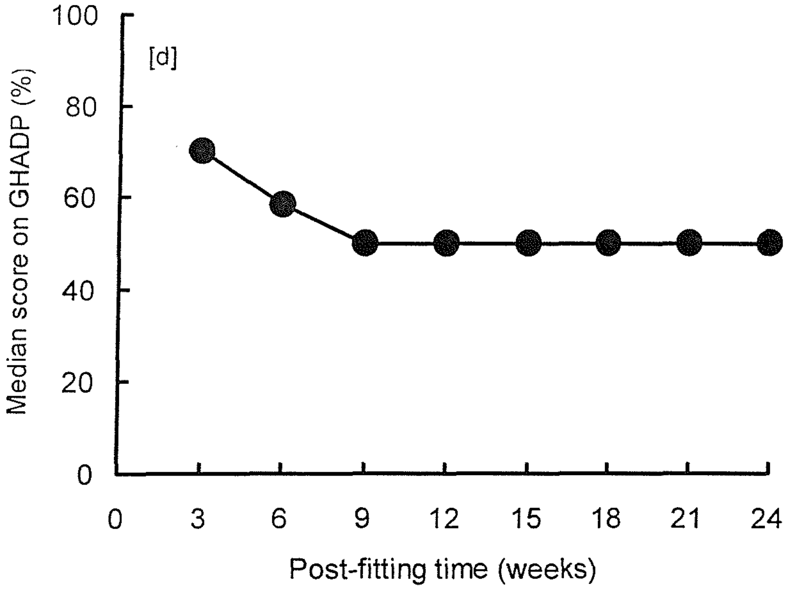


Figure 4.13 cont'd

The focus of the analysis was concerned with the difference in sub-scales over time. The difference data conformed to a normal distribution and were analysed using parametric statistics. One of the 16 subjects was excluded from the analysis because he did not complete the questionnaire on three successive occasions. Linear interpolation was used to smooth the remaining data to account for any missed test sessions. A one-factor repeated-measures ANOVA was performed on each sub-scale. The degrees of freedom were adjusted for univariate tests using the Greenhouse-Geisser correction for the factors that were statistically significant on Mauchly's test of sphericity. A summary of the ANOVA table is given in Table XII, Appendix E. The *Use of Hearing Instrument* and *Residual Disability* sub-scales did not show any statistically significant change over time. However, the increase in benefit and satisfaction were statistically significant. For both of these, it was only the linear component that was statistically significant [additional benefit, F(1,14)=15.6; p<0.01, additional satisfaction, F(1,14)=5.4; p<0.05].

Table 4.15. Distribution of scores on the Glasgow Hearing Aid Difference Profile post-fitting sub-scales. Scores are given for Residual Disability, Use of Hearing Aid, Additional Benefit and Additional Satisfaction since previous visit. A low Residual Disability is indicated by a low score. A high score on Use of Hearing Aid indicates that the hearing instrument is being used in a listening situation that, prior to fitting, lead to hearing difficulty. Additional Benefit refers to the change in benefit that has occurred since the previous visit 3-weeks earlier. A score greater than 50% means that additional benefit has been reported whereas a score of less than 50% means that benefit has decreased since the previous visit. The Additional Satisfaction category is interpreted in the same manner as the addition benefit category. The number of subjects used to obtain each mean value ranged from 13 to 16. Incomplete data were due to subjects unable to attend or lack of time at the end of the test session to complete the questionnaire.

Time [weeks]	3	6	9	12	15	18	21	24
Residual Disability								
25th	0.0	0.0	0.0	1.6	2.1	0.0	2.5	5.3
50th	16.7	15.8	9.2	10.4	13.3	15.0	16.7	12.5
75th	25.0	26.6	17.2	23.4	31.3	31.3	25.0	23.8
Use of Hearing Aid								
25th	66.7	87.5	81.2	67.1	81.8	90.0	92.7	75.0
50th	87.5	93.7	93.3	88.8	91.9	93.8	100.0	90.4
75th	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Additional Benefit								
25th	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
50th	66.7	68.8	57.3	56.3	50.0	50.0	50.0	50.0
75th	95.3	80.3	67.2	62.5	56.2	75.0	70.0	50.0
Additional Satisfaction								
25th	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
50th	70.0	58.3	50.0	50.0	50.0	50.0	50.0	50.0
75th	77.1	77.1	69.1	56.3	55.9	58.3	68.8	50.0

4.2.4 Summary of statistical analysis

The results of the statistical analysis on the mean data can be summarised as follows:

FAAF test

- The benefit scores in the fitted ear and the not-fitted control ear were not differentially affected over time.
- There was a statistically significant difference in benefit score between ears at six weeks post-fitting but only when benefit was calculated for the user-gain setting in the fitted ear [the fitted ear increased by 5% whereas the not-fitted control ear increased by 0.3%].
- The aided and unaided scores increased over time for both ears by around 5%.

BKB sentences

- The benefit scores in the fitted ear and the not-fitted control ear were not differentially affected over time.
- There was a trend of increasing score in the aided and unaided conditions over time but only the aided scores in the not-fitted control ear were statistically significant.

Modified GHADP

- The use of the hearing instrument and the residual disability after hearing instrument fitting did not change over time.
- There was a statistically significant increase in benefit and satisfaction over time.

4.3 Discussion

The experiment assessed hearing instrument benefit, at repeated intervals, over a post-fitting time period of 24 weeks. The aim was to determine if benefit increased over time. Subjects were fitted monaurally as this allowed the not-fitted ear to be used as the control condition to check for practice effects. Benefit was defined as the difference between the aided and the unaided speech score.

4.3.1 Outcome measured with the FAAF test

Benefit was assessed for two conditions of hearing instrument gain setting. In the first condition, the gain setting was fixed to provide a constant level of audibility across test sessions. In the second condition, the subject adjusted the gain setting at each visit prior to data collection in order to reflect any changes in preferred gain that may have occurred with experience. There was no differential increase in the fitted ear relative to the not-fitted control ear for either of the gain settings. This indicates that the benefit scores in each ear were not differentially affected by time.

The lack of evidence to support an increase in benefit at the fixed-gain setting and the lack of sustained increase in benefit over time for the user-gain setting are surprising. The present study has many similarities to previous studies that have demonstrated changes over time. For example, the FAAF test was used as the outcome measure by Gatehouse [1992a, 1993]. The most likely explanation for the conflicting findings is methodological differences between the studies.

The hearing threshold levels of the subjects used in the present study showed a moderate, gently sloping hearing impairment. This degree and configuration of hearing impairment is similar to that of subjects used in previous studies that have demonstrated acclimatisation [for example, Gatehouse, 1992a and Horwitz and Turner, 1997]. The similarity in hearing threshold level across individuals in the study [as shown by the relatively narrow standard deviations in Table 3.1] demonstrates that the subjects form a relatively homogeneous audiometric group. This is advantageous because some previous studies that have failed to demonstrate acclimatisation [for example, Bentler, 1993a, 1993b] have used a relatively heterogeneous group of subjects. The magnitude and prevalence of acclimatisation for different degrees and patterns of hearing impairment has not yet been established. In terms of audiometric data, therefore, the subjects in the present study appear to be appropriate for a laboratory investigation of auditory acclimatisation.

All subjects were fitted with the same programmable hearing instrument [providing linear amplification] thereby reducing potential differences due to different signal processing strategies. All of the subjects in the Gatehouse [1992a] study and a considerable number of subjects in the Horwitz and Turner [1997] study were also provided with linear amplification. The hearing instrument provided considerable amplification, especially at the higher frequencies. The reported daily use of the hearing instrument was typically 8-12 hours. Not all of the previous studies have specified the hours of daily use of the hearing instrument but it is unlikely that they exceed the levels reported in the present study. In terms of providing and using amplification, therefore, the present study appears to be ideally suited for the investigation of auditory acclimatisation.

Ceiling effects in the performance on the speech tests can be ruled out since the mean aided score at the time of fitting was around 70%. It is possible that acclimatisation did not occur at the fixed-gain setting because this was not representative of the subject's familiar setting. Robinson and Summerfield [1996] have argued that stronger effects will be observed when the test condition parallels the normal experience of the listeners. The mean user-gain was relatively stable from three-weeks post-fitting onwards and was typically some 5-8 dB less than the fixed-gain setting. The difference in SII between the fixed-gain and user-gain was of the order of 0.05. It is not known if this relatively small difference in audibility is sufficient to explain the lack of acclimatisation at the fixed-gain setting although it would seem unlikely. Gatehouse [1998] has provided evidence to show that a rich sampling of the range of auditory environments experienced by the listener is required to adequately reflect the diverse range of listening environments that are experienced in everyday situations. This would tend to suggest that acclimatisation should have been measurable at the fixed-gain setting since this does not differ markedly from the user-gain setting.

The user-gain setting is expected to be more characteristic of the normal listening condition. The relatively stable setting from week three onwards has helped to reduce any confounding changes in benefit over time due to differences in gain [and hence audibility]. Once again, there was no systematic increase in benefit in the fitted ear. However, there was a statistically significant increase in benefit at six weeks post-fitting. The increase in benefit score at the user-gain setting at six weeks post-fitting is primarily due to an increase in the aided score although there is also a slight reduction in the unaided score. This is consistent with the findings of Gatehouse [1992a] who showed that performance increased for the aided listening condition to which the ear was accustomed and decreased in the unaided condition. The studies by Cox *et al.* [1996] and Horwitz and Turner [1987] showed an increase in aided score with little change in the unaided score.

Figure 4.14 shows a scatter plot of benefit scores for the fitted and not-fitted ears at six weeks post-fitting. If there were no systematic differences between ears, the results would lie close to the diagonal [lowered by 2-3% to account for the asymmetry in benefit scores between ears at the start of the study]. Most subjects show an increase in benefit in the fitted ear relative to the not-fitted control ear. The mean difference is 8.6% [SD, 10.3]. Only one subject showed a difference greater than two standard deviations [this subject showed an unusually high benefit score in the fitted ear of 35%]. Even when this subject was removed from the analysis, the mean difference of 6.4% [SD, 6] was statistically significant on a paired *t*-test [t=3.9, p<0.05].

The mean aided score at six weeks post-fitting was 75.2% at the user-gain setting and 71.5% at the fixed-gain setting. Given the well-documented relationship between audibility and speech performance, performance would have been expected to be higher on the fixed-gain setting. Despite the reduced audibility, subjects as a group performed better at the user-gain setting. It is possible that there was less upward spread of masking at the lower gain setting and this resulted in the better performance despite the lower audibility. However, if this were the case, then presumably the difference in performance would have been greatest at the time of fitting when the difference between the user and fixed-gain was 10 dB.

It is also puzzling why the difference between ears at six weeks post-fitting did not persist given that the mean user-gain setting was relatively stable; the mean gain setting was similar at week 6, 9 and 15 and the remaining weeks were within 1-2 dB of this setting. The studies reported by Gatehouse [1992a, 1993] and Horwitz and Turner [1997] showed an increase in performance over the duration of their respective studies and there was no evidence of an asymptote when the studies terminated. There is no way of checking that the user-gain selected by the subject on subsequent visits was representative of the gain used in their own home environment. However, it is difficult

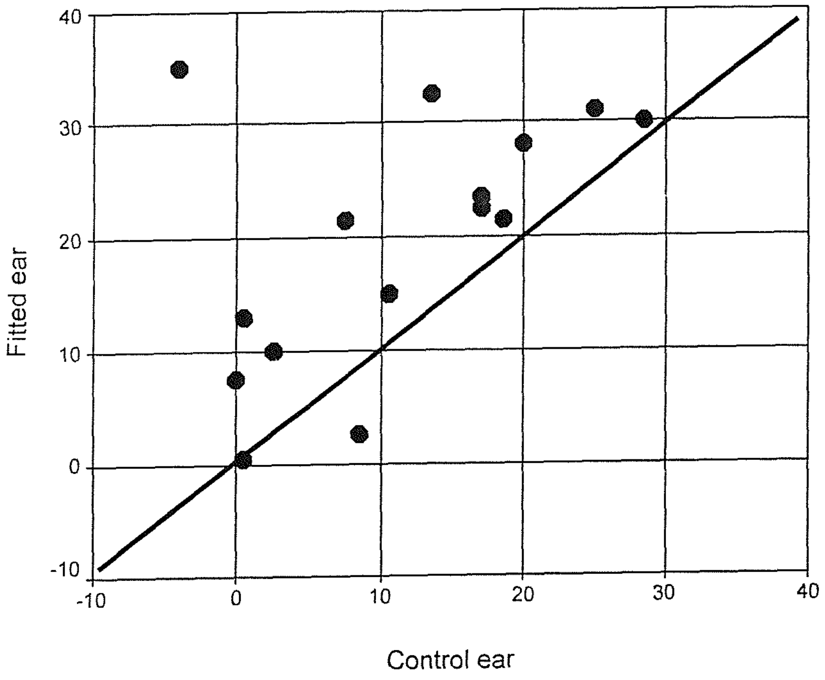


Figure 4.14. Scatter plot of benefit scores for FAAF at six- weeks post-fitting for the fitted ear [at user-gain] and the control ear.

to believe that the subjects correctly selected their preferred user-gain at six weeks post-fitting but were unable to repeat this on subsequent visits when their experience with manipulation of the gain control had presumably increased. The increase in benefit at six weeks post-fitting is smaller than that reported by Gatehouse [1992a] and Horwitz and Turner [1997]. This finding, along with the lack of sustained increase in initial benefit, means that, despite providing new and useful information to a homogeneous group of subjects, the experience of amplification provided by the linear hearing instrument [fitted to a recognised prescription target] was not sufficient to detect acclimatisation.

The aided and unaided scores for both ears show a general increase over time. An increase of 5% on an initial aided score of 71% represents an increase in the number of words scored correctly from 56 to 60. The most likely explanation for this improvement over time is that there was a practice effect with the test material. A practice effect did occur when determining the SNR before hearing instrument fitting and several sessions were required before this appeared to stabilise. However, the FAAF test is reported to be robust to pure response learning [Foster and Haggard, 1984]. Gatehouse [1992a] did not show any systematic increase in the aided or unaided scores in the control ear despite intensive testing on each subject. A possible explanation for the difference between studies is that the software used in the present study incorporated visual feedback to the subject. The correct word flashed on the touch screen monitor after the subject had made their response. This feedback was not present in the Gatehouse study [Gatehouse, personal communication]. It is possible that the feedback trained the subject to concentrate on subtle acoustic features that were present, but not used, at the start of the study. In addition, subjects could use the feedback to indicate that a particular item out of the 80 in a FAAF list had been used and this would restrict the range of alternatives that were available on future presentations. If this strategy were to increase over time then this could potentially contribute towards a systematic learning effect. Whatever the explanation, it is important to identify and remove the source of this practice effect before further studies are carried out since this may hamper the interpretation of the aided and unaided scores that are used to derive the benefit scores. The interpretation of acclimatisation is less open to criticism if no improvement in performance occurs in the not-fitted control ear.

4.3.2 Outcome measured with BKB sentences

The BKB sentence test was used to give an indication of communication in a more natural listening condition than the single words used in the FAAF test. There was no statistically significant interaction between ears and post-fitting time. This indicates that the benefit scores in each ear were not differentially affected by time.

The BKB material consists of 21 sentence lists. This limited test set restricted testing to two conditions: unaided and aided at fixed-gain. It could be argued that acclimatisation did not occur because the aided condition was not representative of the normal user-gain setting. In addition, band importance functions for sentence material are less loaded at the high frequencies where amplification [and hence the need for learning] was greatest. A more challenging listening condition may be required to demonstrate that the subject is required to learn to extract subtle cues from the speech signal.

The variability of score in a BKB sentence list is intrinsically higher than the FAAF test because there are fewer scoring words [50 instead of 80]. In addition, contextual cues are available in the sentences that mean the true number of independent scored items in a list is less than 50. In order to reduce test variability [and hence observe small increases in benefit over time], four sentence lists were presented for each of the four test conditions [aided and unaided in each ear]. Of the 21 sentence lists, 16 were used on each of the five test sessions. On average, each sentence list was used on four occasions. It is likely that the repetition of the sentence material was responsible for the increase in the aided and unaided scores that occurred over the course of the study. These limitations suggest that the BKB sentences may not be an appropriate outcome measure when designing a study that requires repeated measurements over time.

4.3.3 Self-report using the modified GHADP

The modified GHADP produces both pre- and post-intervention scores on a range from 0 to 100. The median scores on the pre-intervention scales of disability and handicap were 50 and 42.7 respectively. This corresponds to 50% of the subjects reporting 'moderate difficulty' and 50% reporting a handicap just short of moderate in severity. The corresponding scores reported by Gatehouse [1999] during development of the profile were 44 and 25.8 respectively. There is broadly similar agreement in the disability scores but greater handicap [essentially a categorical difference on the profile from *'a little'* to *'a moderate'* handicap score] in the present study. The subjects in both studies were of a similar age and had a similar degree and configuration of hearing impairment. However, the subjects used in the present study were a rarefied subgroup of the clinical population. They were recruited because they were highly motivated individuals who were experiencing significant difficulty and were keen to wear a hearing instrument on a regular basis. The median handicap score is consistent with those reported in clinical trials conducted by Lutman and Payne [1999] and Wood and Lutman [2001].

The median score for self-reported use of the hearing instrument showed small variations across time but was typically around 90. This corresponds to 50% of subjects using their hearing instruments for greater than 90% of the time when they find themselves in listening situations that prior to receiving the hearing instrument resulted in hearing difficulty. This supports the finding that subjects reported 8-12 hours daily use of the hearing instrument. Given the high motivation of the subjects, it is not surprising that the use score is greater than the median score of 70.2 reported by Gatehouse [1999].

The median scores for residual disability showed small variations across time but were around 10-15. This corresponds to a self-report of '*none*' or '*slight*' residual disability. These subjects were fitted with programmable hearing instruments and received substantial pre- and post-intervention support. This probably explains the more favourable outcome compared to the 20.9 reported by Gatehouse.

There is a potential danger when using self-assessment to investigate relatively short-term changes over time. For example, the initial increase in audibility at the time of fitting may result in a more favourable response than when the individual has some experience of ongoing difficulty in background noise. This may result in a lower score [disability and handicap] soon after fitting but increases over time as the subject becomes aware of the limited benefit in difficult listening situations. The present study did not show any rebound in residual disability when measured at three-week intervals over a post-fitting period of six months. Taylor [1993] demonstrated a reduction in handicap on the HHIE at three weeks which was followed by a significant increase in handicap at six months before stabilising at a level somewhere between the two scores. A similar finding was reported by Malinoff and Weinstein [1989] who also compared the HHIE at the same time periods. Other authors have not reported this 'rebound' effect. For example, Mulrow *et al.* [1992] reported a significant decline in handicap at four months and this was still present at 8 and 12 months.

The final two sub-scales were concerned with differences in benefit and satisfaction over time. The subjects were asked to compare their present level of benefit and satisfaction with the levels experienced on the previous visit. There was a statistically significant increase in benefit over the first 12 weeks from 66.7 to 56.3 before stabilising at 50. This corresponds to the subjects reporting a marginal increase in benefit, week on week, before settling at 12 weeks post-fitting. The results for the additional satisfaction sub-scale follow the same pattern with a statistically significant increase from 70 at three weeks post-fitting to 80 from nine weeks onwards.

Although there was a perceived improvement in benefit and in satisfaction, this was not mirrored by a reduction in residual disability. This may be because of the relative insensitivity of the residual disability score due to floor effects; essentially, there was not enough room to reveal a further reduction in disability. Alternatively, it may be that the relative change in benefit and satisfaction is a more sensitive indicator of the subtle changes that have occurred over time. The change in benefit and satisfaction was made relative to the perceived state on the previous visit three weeks earlier. An alternative approach would have been to compare change relative to those experienced around the time of fitting. This would have meant that any change would be referenced to a single anchor in time. It is not known if these alternative approaches would result in the same outcome.

A number of previous studies have used self-report to measure changes over time. Cox and Alexander [1992] reported a significant increase in benefit over a 10-week period using the PHAB. Although the group consisted of new and experienced users, they were all fitted with new hearing instruments so changes due to acclimatisation may have been possible. Similarly, Horwitz and Turner [1997] reported a significant reduction in the frequency of difficulties experienced for the aided condition in new users over an 18-week period using the PHAB. They also showed a consistent trend towards an increase in aided benefit that was not statistically significant on a two-factor repeated-measures ANOVA [but may have been so if the new and experienced users had been analysed separately using a one-factor repeated-measures ANOVA]. Bentler *et al.* [1993] used the understanding speech subsection of the HPI and reported a significant improvement for conversation in quiet [with a trend in noise] over a 12-month period in a group of 65 new and experienced users. Using the HHIE, Taylor [1993] reported no significant reduction in handicap between three weeks and one year post-fitting in a group of 58 new users.

An alternative explanation for any changes noted over time is that the improvement may simply be a reflection that the subject has become more familiar with the hearing instrument [for example, has become more proficient with adjustment of the gain control in different listening situations]. There is no easy and reliable method to untangle these two possibilities. For this reason, the use of self-report is likely to be of limited use in studies investigating changes in benefit over time that are specific to acclimatisation. As with the residual disability score, there was no evidence of a rebound effect due to subjects being overly optimistic at the time of fitting.

In the present study, there were a few occasions when the subjects did not find themselves in a particular listening environment between each test session. For example, the subject may have been unable to attend church due to ill health so this listening environment would have contributed intermittently to the self-report scores across test sessions. It is also possible that providing a hearing instrument may enable subjects to enter environments that they otherwise would avoid. This highlights fundamental problems for any longitudinal study where lifestyle changes across time. The results reported in this study included all the listening environments when they occurred. However, the data were analysed after excluding the listening environments that occurred intermittently [or ceased to be important]. Fortunately, this did not lead to any difference in the outcome for the group data.

4.4 Conclusions

The mean hearing instrument benefit, calculated using the fixed-gain aided condition, was stable over the 0-24 week post-fitting period when measured with the FAAF test and BKB sentences. However, there was an increase in mean performance in the aided and unaided condition for both ears. The most likely explanation for this improvement was a practice effect with the test material, either because the feedback allowed the subject to concentrate on subtle acoustic features or because it reduced the number of key words yet to appear. Therefore, it is important to identify and remove this practice effect before undertaking further studies on auditory acclimatisation.

There was a difference between the fitted and not-fitted control ear when performance was measured on the FAAF test for the user-gain condition but this was only statistically significant at six weeks post-fitting. Thus, the results do not show evidence of sustained increases in improvement in performance specifically in the fitted ear over time. Despite providing new and useful information to a homogeneous group of subjects, the experience of amplification provided by the linear hearing instrument fitted to a recognised prescription target was not sufficient to detect acclimatisation.

There was a self-reported improvement in both benefit and satisfaction over the initial 9-12 week post-fitting period. This is consistent with auditory acclimatisation although it could be due to factors such as increased familiarity with the hearing instrument controls.

Given the difficulty measuring acclimatisation with both performance tests, it is probably appropriate to restrict further experiments to measuring performance with only the FAAF test since this is more strongly dependent on high frequency speech sounds than the BKB sentence material.

4.5 Experiment Two

The results from the previous experiment revealed that performance on the FAAF test increased, over time, in both ears and in the aided and the unaided test conditions. The most likely explanation for this increase across all test conditions is a practice effect due to repeated exposure to the test material. One approach would be to dissociate the acclimatisation effect from the practice effect by subtracting the improvement in the test ear [practice and acclimatisation] from the improvement in the not-fitted control ear [practice]. However, this assumes that there is no transfer of acclimatisation to the control ear and no interaction between learning the test procedure and perceptual learning. The interpretation of acclimatisation is less open to criticism if no improvement in performance occurred in the control condition. In addition, since the magnitude of this practice effect [approximately 5%] is similar to the magnitude of the acclimatisation effect reported in some of the previous studies [for example, Cox *et al.*, 1997] this could also hamper the detection of acclimatisation, particularly since the variability of speech testing is known to be relatively high. Therefore, it would be helpful to identify and remove the source of the practice effect. Foster and Haggard [1984] have reported that the FAAF test is robust to pure response learning and the study by Gatehouse [1992a] did not show any improvement in the control ear after extensive testing. However, an increase in performance was observed on the practice sessions preceding hearing instrument fitting. This meant that in order to maintain aided performance at around 71%, the noise had to be increased by 1-3 dB. The 5% improvement that occurred over the 24 week post-fitting period is equivalent to a further change in SNR of around 1 dB [Lutman and Clark, 1986]. In the previous study, the subjects selected their response from the multiple choice on the touch screen monitor. They were then alerted to the correct response that flashed on the monitor. The studies by Foster and Haggard [1984] and Gatehouse [1992a] did not provide this feedback. While the use of feedback may have helped to maintain subject motivation, it may also have enabled subjects to improve their performance for the reasons discussed in section 4.3.1. Thus, the practice effect may be related to the visual feedback given to the subject. One aim of Experiment Two was to determine if disabling the feedback would reduce the practice effect. The null hypothesis was that there would be no change in performance in the not-fitted control ear over time. The alternative hypothesis was that there would be an improvement in performance in the control ear over time.

Experiment One also revealed an increase in mean aided performance at six weeks post-fitting for the user-gain setting but not the fixed-gain setting. The user-gain setting was more representative of the subject's normal listening level but the difference between the two gain settings was small: the user-gain was, on average, 7 dB less than the fixed-gain setting. It is not known if this relatively small difference in audibility is sufficient to explain the lack of acclimatisation at the fixed-gain setting. A further aim of the present study was to test the subjects at fixed-gain and user-gain to determine if this finding was repeatable. The null hypothesis was that there would be no improvement over time in the fitted ear at the user-gain setting: the alternative hypothesis was that there would be an improvement in performance over time.

If the improvement in performance at the user-gain setting is repeatable, then it should also be present at the higher fixed-gain setting if the speech input level is reduced to keep the amplified level constant. For example, if the user-gain is 7 dB below the fixed-gain setting, the same amplified level of speech can be achieved at the fixed-gain setting with the speech level reduced by 7 dB. Since the increase in gain is offset by an equivalent decrease in input level, this should not cause any non-linearity in the hearing instrument. This was included as a further test condition. The null hypothesis was that there would be no improvement in performance over time in the fitted ear at the simulated fixed-gain setting (i.e., user gain and higher speech level). The alternative hypothesis was that there would be an improvement in performance over time. Testing was also undertaken at the user-gain setting with the speech input level increased to simulate the fixed-gain condition. The null hypothesis was that there would be no improvement in performance over time at the simulated user-gain setting: the alternative hypothesis was that there would be an improvement over time.

A further finding from the previous experiment was the improvement in self-reported benefit and satisfaction that occurred over the first 9-12 weeks after hearing instrument fitting. This information was obtained by asking the subjects to compare their present level of benefit and satisfaction with the level they experienced on the previous test session three weeks earlier. An alternative approach would have been to make the comparison relative to the benefit and satisfaction at the time of fitting. This approach would have meant that any change that occurred would be referenced to a single anchor point in time. Since the GHADP only takes approximately 15 minutes to complete, it was decided to include it in the present study with this alternative wording. It is important for methodology to know if the two approaches result in the same outcome. In addition, if improvements in benefit and satisfaction were also reported in the present study then this would add confidence to the use of self-reported improvements over time. The null hypothesis was that there would be no improvements over time: the alternative hypothesis was that there would be an improvement over time.

Little use has been made of sound quality judgements to investigate changes in preference over time. It is possible that a preference for the fitted response may increase over time. Ovegard *et al.* [1997] reported a significant difference in mean data over time when comparing 'clarity' and 'total impression'. Ching *et al.* [1997, 1999] showed a preference for the fitted response but it is not known if this preference was present at the time of fitting. Because of the ease involved in switching between hearing aid memories, a paired comparison of sound quality judgements for two frequency responses was also included. The null hypothesis was that there would be no increase in preference for the fitted response over time: the alternative hypothesis was that the preference for the fitted response would increase over time.

4.6 Additional methodology for Experiment Two

Sixteen subjects were recruited [8 male, 8 female] with a mean age of 70.8 years [±4.9]. Twelve subjects were fitted in the right ear and four subjects were fitted in the left ear. All 16 subjects completed the study.

The same fitting targets were used as in experiment one. Again, the greatest difficulty was achieving the 4 kHz target and removing the resonance peaks in the mid-frequencies. The REAR from each subject was then converted to REIG in order to calculate the SII values. The mean REIG values for the fitted ear are shown in Figure 4.15 and summarised numerically in Table 4.16. Also shown is the alternative frequency response that was used for the paired comparison test. This frequency response has less high frequency emphasis. It was selected because the DSL fitting method recommends greater amplification than most other prescription methods. The mean high frequency gain reductions were 8, 13 and 16 dB at 3, 4 and 6 kHz respectively. This is generally similar to the difference between DSL and NAL-RM

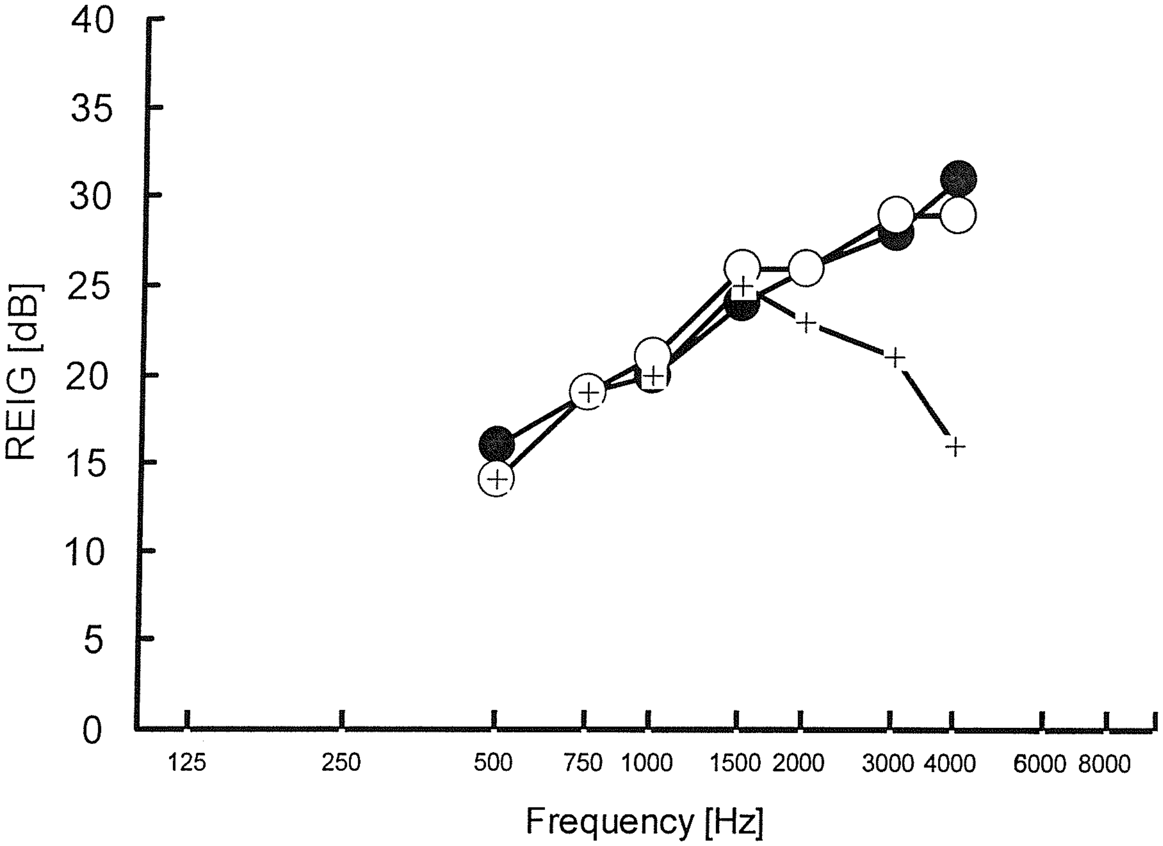


Figure 4.15. Mean real-ear insertion gain for the fitted ear. The target values for DSL are shown as filled circles and the match is shown as open circles. The alternative response, used in the paired comparison, is shown as plus signs.

targets for a typical age-related hearing impairment [although DSL also recommends greater gain at the low frequencies].

Table 4.16. Mean real-ear insertion gain [dB] for the fitted ear. This was obtained by subtracting the real-ear un-amplified speech signal from the data in the previous table. One standard deviation is given in brackets.

Frequency [Hz]	500	750	1000	1500	2000	3000	4000	6000
Target	16 [4]	19 [3]	20 [4]	24 [5]	26 [6]	28 [6]	31 [5]	44 [8]
Match	14 [4]	19 [3]	21 [5]	26 [5]	26 [7]	29 [6]	29 [5]	29 [9]
Alternative	14 [4]	19 [3]	20 [5]	25 [5]	23 [7]	21 [6]	16 [6]	13 [11]

The same FAAF test material was used as in the earlier experiment. The only difference was that the subject was not given visual feedback after making their response. The modified GHADP was also used but with different wording from experiment one: the additional benefit and satisfaction scales were obtained by comparing current benefit and satisfaction with what the subjects could recollect at the time of fitting.

Table 4.17 Magnitude preference scale

Memory 1 MUCH better than Memory 2	[6]
Memory 1 MODERATELY better than Memory 2	[5]
Memory 1 SLIGHTLY better than Memory 2	[4]
Memory 2 SLIGHTLY better than Memory 1	[3]
Memory 2 MODERATELY better than Memory 1	[2]
Memory 2 MUCH better than Memory 1	[1]

Subjects performed a paired comparison for sound quality dimensions of comfort and clarity as well as overall preference. In order to provide additional information regarding the strength of the preference, the paired comparison technique incorporated a magnitude preference scale along the lines described by Dillon [1984] and previously used by Balfour and Hawkins [1992]. The response options are shown in Table 4.17. The values assigned to each response option are shown in brackets. Using this approach, the point where the two frequency responses would be equivalent would be 3.5. In order to test the null hypothesis that the mean values do not differ significantly from zero, a constant value of 3.5 was subtracted from each data entry. The test material used for the paired comparison consisted of BKB sentences presented in three backgrounds: quiet, filtered steady noise and babble. The filtered noise was used in the previous experiment and had the same long-term spectrum as the BKB sentences. The babble was supplied with the BKB sentences and consisted of CUNY sentences from 20 different talkers [male and female] mixed at approximately equal levels [Foster, personal communication]. The BKB sentences were presented at a SNR that had previously been determined for each subject in the practice sessions to give an aided score in the region of 71%. Both frequency responses were presented at the gain setting that was selected by the subject at the start of each test session while listening to the BKB material. The investigator alternated the frequency responses until the subject made a judgement. The quality dimensions were rated three times on each visit and the responses were then averaged. All subjects practised using the quality dimension scale before data collection commenced.

Outcome measures were taken on the day of fitting and again at 4, 8, 12, 16, 20 and 24 weeks post-fitting. The previous study used intervals of three weeks but this was extended to four weeks in the present study; although the number of test sessions decreased, the overall number of FAAF lists used was the same in both studies. This was considered appropriate when comparing differences across studies due to practice effects.

Table 4.18. Mean difference [dB] between the fixed gain and the user-gain setting as a function of post-fitting time at 2 kHz in 2-cc coupler. A positive value indicates the extent to which fixed-gain exceeds the user-gain. One standard deviation is given in brackets.

Post-fitting time [weeks]	0	4	8	12	16	20	24
Fixed minus user gain [dB]	7 [6]	6 [6]	4 [6]	3 [7]	3 [6]	4 [6]	5 [6]

The FAAF test was carried out on each session at the fixed-gain and the user-gain setting in the fitted ear. Only the fixed-gain setting was used for the not-fitted control ear. This was the same as the previous study. It was not necessary to carry out unaided testing since the practice effect could be assessed using the aided scores alone. Any change in the control ear could be attributed to a practice effects and not auditory acclimatisation. The FAAF test was presented at the subject's fixed SNR and they were instructed to adjust the hearing instrument gain control, using a bracketing procedure, in order to find their preferred user-gain position. The investigator then measured the gain in an HA2 2-cc coupler at the user-gain setting. The subject performed the adjustment three times and the average of the three settings was used for the user-gain condition on that test session. The mean difference between the fixed gain and the user gain as a function of post-fitting time is shown in Table 4.18. The mean user-gain was 3 to 7 dB less than the fixed-gain.

The design included two additional test conditions where the speech level was adjusted to compensate for the difference between gain settings. These were known as the simulated user-gain and the simulated fixed-gain conditions. The speech input level was adjusted so that the simulated user-gain was equivalent to the fixed-gain condition and vice versa. This resulted in a total of five test conditions and these are summarised in Table 4.19.

Table 4.19. Test conditions and overall presentation level [dB SPL] for FAAF

	Test ear	Control ear
Aided at fixed-gain	60	60
Aided at user-gain	60	
Aided at simulated fixed-gain	60 + correction	
Aided at simulated user- gain	60 - correction	

Correction = Fixed gain – User gain

The modified GHADP was completed using a paper and pencil response format. The investigator checked and clarified the responses. This took approximately 15 minutes to complete and was carried out midway through each test session. The subjective judgement task was performed at the end of each visit using the paired comparison and magnitude estimation scale. The order of quality dimensions was randomised across session.

4.7 Results

The mean self-reported daily use of the hearing instrument as a function of post-fitting time is shown in Table 4.20. No subjects reported using the hearing instrument for less than 4-8 hours per day; more than half reported wearing the hearing instrument for >8 hours per day.

The difference in audibility between the fixed and the user gain was predicted using the SII procedure outlined in Appendix F. The mean SII values are given in Table 4.21. The mean unaided SII was 0.110. The mean aided SII at fixed-gain was 0.499. The mean SII at user-gain varied between 0.434 and 0.501 depending on the test session. The hearing instrument therefore made a substantial improvement to the audibility of the FAAF test when presented at 60 dB SPL.

Table 4.20. Mean self-reported daily use of the hearing instrument [hours] as a function of post-fitting time. Each subject was issued with a diary that required them to indicate the hours of daily use of the hearing aid. The categories available to the subject were: less than 4 hours, 4-8 hours, 8-12 hours or greater than 12 hours. The daily hours of use for each subject was taken as the most frequently ticked category over each of the three-week periods. The percentage of responses in each category is given in brackets [n=16].

Time [weeks]	0-4	4-8	8-12	12-16	16-20	20-24	Mean
< 4 hours	0 [0]	0 [0]	0 [0]	0 [0]	0 [0]	0 [0]	0 [0]
4-8 hours	7 [44]	2 [12.5]	3 [19]	5 [31]	4 [25]	3 [19]	4 [25]
8-12 hours	5 [31]	9 [56]	5 [31]	7 [44]	6 [37.5]	5 [31]	6 [37.5]
> 12 hours	4 [25]	4 [25]	4 [25]	4 [25]	5 [31]	6 [37.5]	5 [31]
Missing data	0 [0]	1 [6]	4 [25]	0 [0]	1 [6]	2 [12.5]	1 [6]

Table 4.21. Mean Speech Intelligibility Index for the aided and unaided conditions in the fitted ear. The Speech Intelligibility Index was calculated using the octave band procedure. The calculations were based on the mean hearing threshold level, in decibels, in the fitted ear. The equivalent speech spectrum level of the FAAF test was based on an Leq of 60 dB SPL. The equivalent noise spectrum of the FAAF shaped noise spectrum was based on a SPL of 58 dB [i.e., a signal to noise ratio of +2 dB]. The band importance function for nonsense syllables was used in the calculation.

UNAIDED	AIDED TARGET	AIDED MATCH	USER ADJUSTED						
			0	4	8	12	16	20	24
0.110	0.498	0.499	0.434	0.451	0.485	0.501	0.501	0.485	0.465

4.7.1 Changes in performance over time in the control ear

The change in mean FAAF score in the control ear, relative to the time of fitting, is shown in Figure 4.16 with the numerical data in Table 4.22. The difference between the mean performance at week 0 and week 24 is +0.6%. This means that performance improved by less than 1%. Figure 4.17 shows the results of the control ear from experiment one. The difference between week 0 and week 24 was +5.4%. Since the first experiment involved measurements at 3-week intervals, an alternative comparison

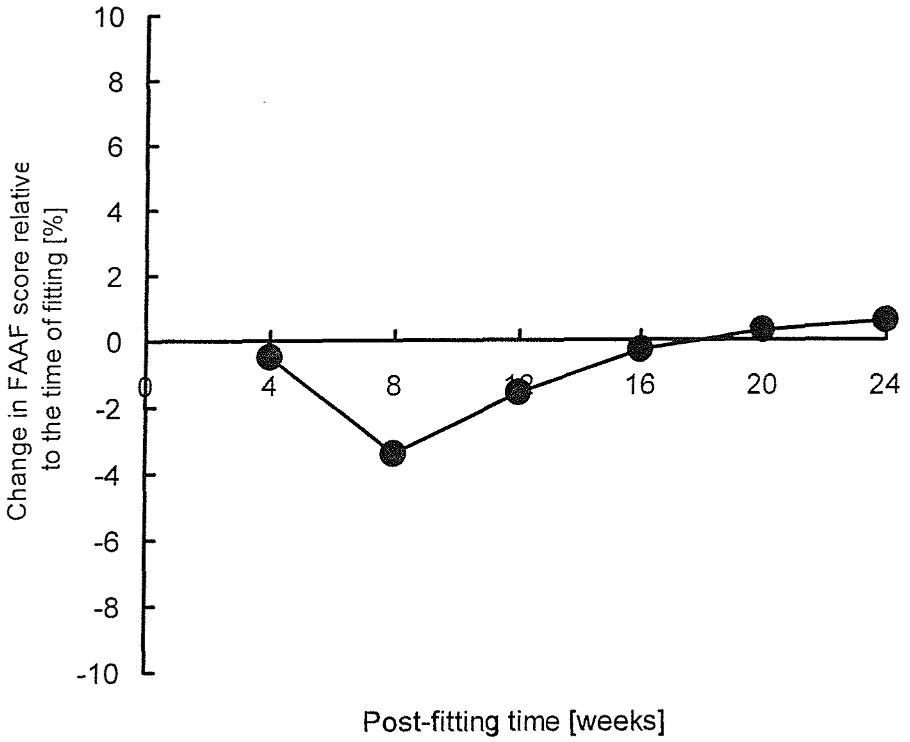


Figure 4.16. Mean change in aided FAAF score in the not-fitted control ear relative to the initial score at the time of fitting.

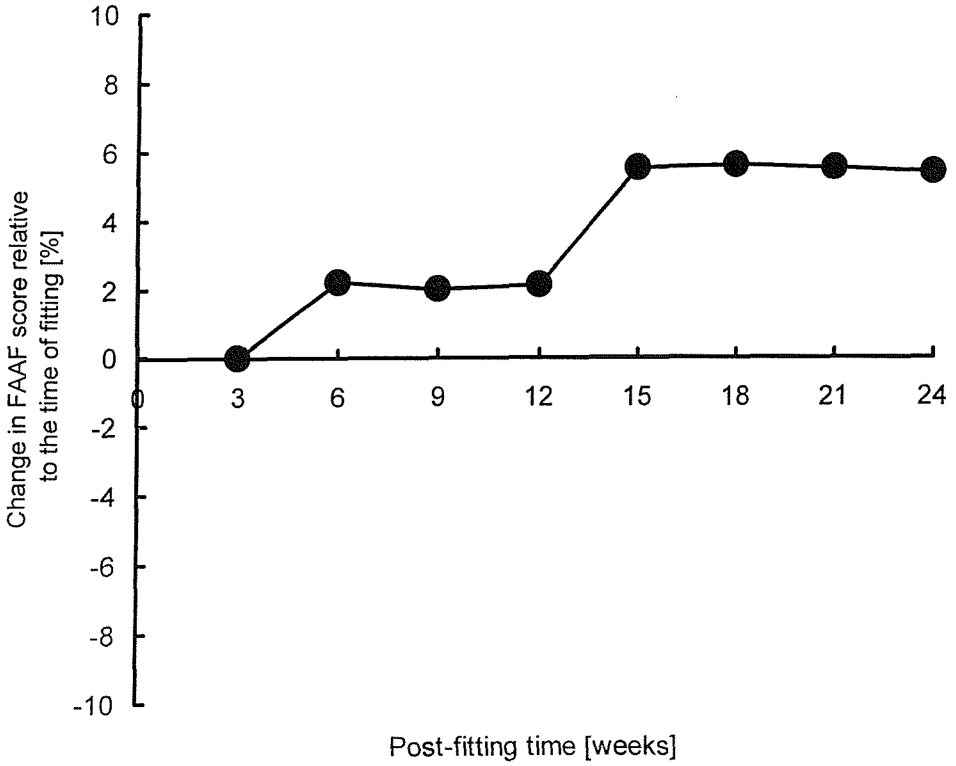


Figure 4.17. Mean change in aided FAAF score in the not-fitted control ear relative to the initial score at the time of fitting from experiment one which included visual feedback.

is the difference at week 0 and week 18 [i.e., after the same number of test sessions]. The difference this time was +5.6%. The present study does not show the increase over time that was observed in the first experiment. In fact, there was a reduction in performance at eight weeks post-fitting; this is discussed later.

Table 4.22. Mean change in aided FAAF score, in percent, relative to the time of fitting [weeks] in the not-fitted control ear. The scores from one subject who missed several consecutive test sessions in each of the experiments have not been included. One standard deviation is given in brackets [n=16].

Time [weeks]	4	8	12	16	20	24	Mean
	−0.5 [3.8]	−3.4 [5.6]	−1.6 [3.0]	−0.3 [3.7]	0.3 [4.0]	0.6 [3.1]	−0.7 [2.6]

A one-factor repeated-measures ANOVA was performed on the data from the present study. The data were smoothed, using linear interpolation, for five subjects who missed a single test session. This enabled all of the subjects to be included in the repeated-measures analysis. A summary of the analysis is given in Table XIII, Appendix E. The table includes details of the orthogonal polynomial breakdown. There was a statistically significant difference over time with a significant cubic relationship. When week eight was removed from the analysis, the relationship was no longer statistically significant [F(4,60) =1.3; p=0.27]. Thus, the null hypothesis can be accepted because there was no systematic improvement in performance in the control ear over time.

4.7.2 Performance in the fitted ear at the fixed-gain and the user-gain settings

The mean FAAF score at the fixed-gain and the user-gain in the fitted ear, as a function of post-fitting time, is shown in Figure 4.18. The filled circles correspond to the fixed-gain condition and the open circles correspond to the user-gain condition. The triangles correspond to the control ear and are included for comparison. The numerical data are given in Table 4.23. The mean scores for the two gain conditions are similar with an increase over time of around 2%.

The data were analysed using a repeated-measures ANOVA. The two factors that were treated as repeated-measures were post-fitting time and aided condition [fixed-gain in each ear and user-gain in the fitted ear]. The data were smoothed [using linear

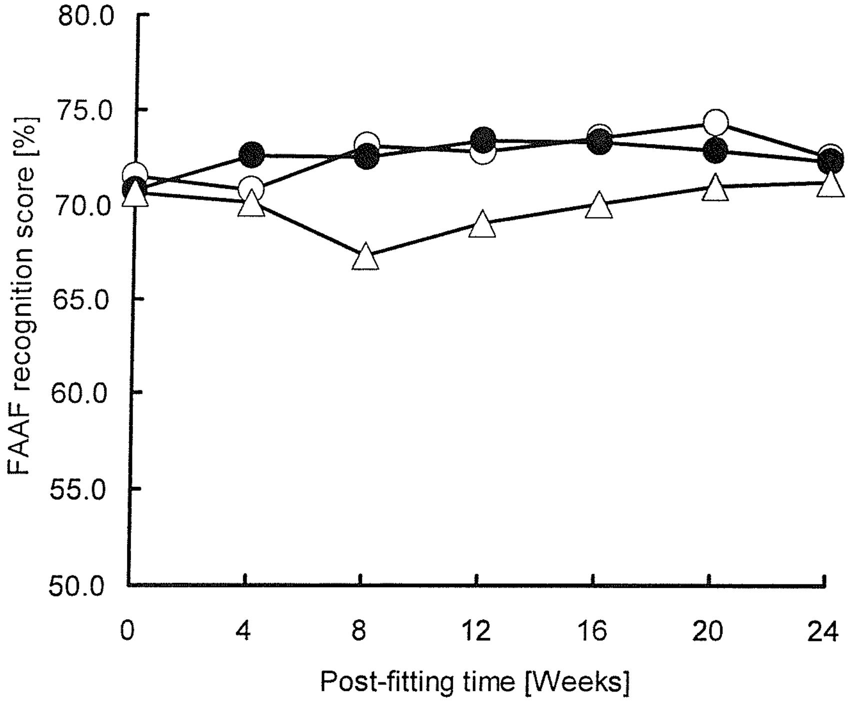


Figure 4.18. Mean aided FAAF recognition score, in percent, as a function of post-fitting time. The circles correspond to the fitted ear [filled symbols at fixed-gain and open symbols at user-gain] and the triangles correspond to the not-fitted control ear.

interpolation] for five subjects who missed a single test session. This allowed all subjects to be included in the repeated-measures analysis. A summary of the analysis is shown in Table XIV, Appendix E. The mean difference over time was not statistically significant. There was a statistically significant difference across gain condition and inspection of the orthogonal polynomial breakdown revealed a significant linear relationship [F(1,15)=9.7; p< 0.01]. This was not significant when the control ear was removed from the analysis [see Table XV, Appendix E]. This means that there was a statistically significant difference between ears but not between the two gain conditions in the fitted ear. Thus, the null hypothesis can be accepted. The lack of interaction between ear and post-fitting time indicates that the ears were not differentially affected by post-fitting time.

Table 4.23. Mean FAAF score, in percent, as a function of post-fitting time. One standard deviation is given in brackets [n=16].

	0	4	8	12	16	20	24	Mean
Fixed gain	70.8 [2.2]	72.6 [3.7]	72.5 [5.1]	73.4 [4.8]	73.3 [4.1]	72.9 [4.1]	72.3 [5.7]	72.5 [4.3]
User gain	71.5 [2.7]	70.8 [4.0]	73.1 [6.0]	72.8 [4.8]	73.5 [5.3]	74.3 [4.4]	72.5 [6.1]	72.6 [4.8]
Simulated fixed	71.7 [4.0]	73.5 [5.7]	71.9 [4.5]	73.4 [5.9]	73.0 [5.3]	73.7 [4.7]	72.1 [6.2]	72.8 [5.3]
Simulated user	71.2 [3.9]	70.8 [4.2]	71.4 [6.8]	71.0 [5.4]	71.0 [6.9]	71.9 [5.6]	72.9 [4.8]	71.5 [5.3]
Control Ear	70.6 [2.2]	70.1 [2.4]	67.3 [4.1]	69.0 [3.3]	70.0 [3.7]	71.0 [3.7]	71.2 [3.1]	70.0 [3.4]
Mean	71.2 [3.0]	71.6 [4.2]	71.2 [5.6]	71.9 [5.1]	72.2 [5.2]	72.8 [4.6]	72.2 [5.2]	71.9 [4.8]

4.7.3 Performance in the fitted ear at the simulated gain settings

The mean FAAF scores for the fixed-gain and the simulated fixed-gain setting [i.e., user-gain with increased speech level] as a function of post-fitting time, are shown in Figure 4.19. The filled circles correspond to the fixed-gain condition and the open circles correspond to the simulated gain condition. The numerical data are shown in Table 4.23. The mean difference in FAAF score between the two conditions is less than 1%. The data were analysed using a repeated-measures ANOVA. The two factors that were treated as repeated-measures were post-fitting time and gain condition [fixed-gain and simulated fixed-gain]. The data were smoothed [using linear interpolation] for five subjects who missed a single test session. This enabled them to be included in the repeated-measures analysis. The degrees of freedom were adjusted for univariate tests

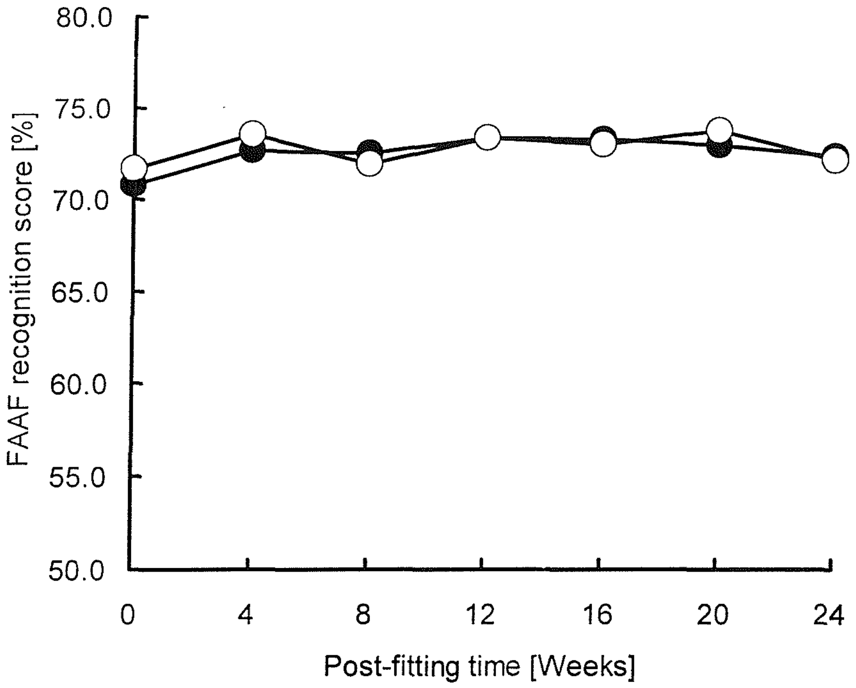


Figure 4.19. Mean aided FAAF recognition score, in percent, as a function of post-fitting time. The filled circles correspond to the fixed-gain condition and the open circles correspond to the simulated fixed-gain condition.

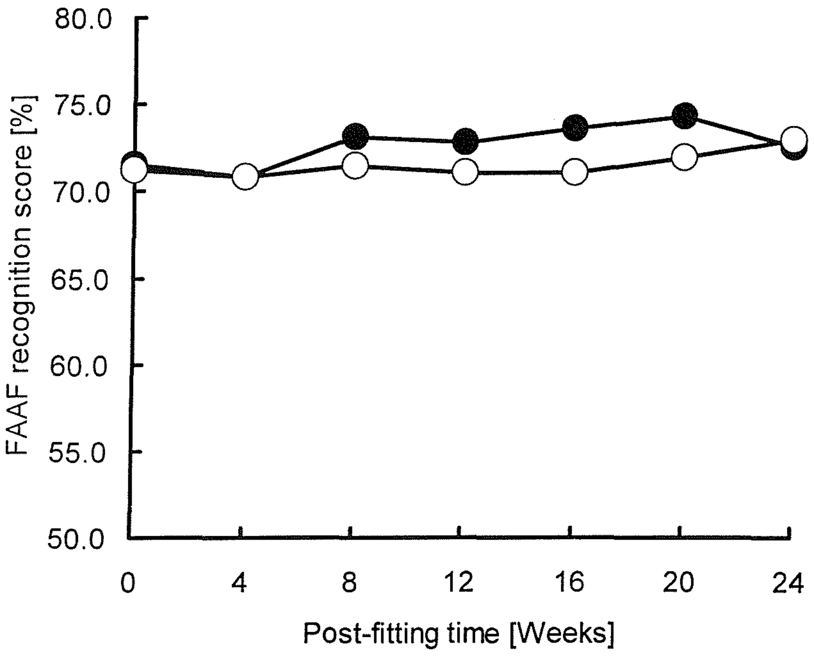


Figure 4.20. Mean aided FAAF recognition score, in percent, as a function of post-fitting time. The filled circles correspond to the user-gain condition and the open circles correspond to the simulated user-gain condition.

using the Greenhouse-Geisser correction for the factors that were statistically significant on Mauchly's test of sphericity. A summary of the analysis is shown in Table XVI, Appendix E. There was no statistically significant difference for time or gain condition, nor was there a significant interaction. The lack of interaction indicates that the gain conditions are not differentially affected by post-fitting time.

The data for the user-gain and the simulated user-gain setting are shown in Figure 4.20 [with the numerical data in Table 4.23]. The mean difference in FAAF score between these two conditions ranges from 0% at week four to 2.6% at week 16. The data were analysed using a repeated-measures ANOVA and a summary of the analysis is shown in Table XVII, Appendix E. There was no statistically significant difference in performance across post-fitting time but there was a significant difference for gain condition; the mean performance at the user-gain setting was around 1% higher than with the simulated user setting. Inspection of the orthogonal polynomial breakdown revealed a significant linear relationship [$F(1,15)=5.8$; $p<0.05$]. This finding is discussed in section 4.8.3. Despite the difference between gain settings, there was no significant interaction between post-fitting time and gain condition. The lack of interaction indicates that, while the gain settings are different from each other, they are not differentially affected by post-fitting time. Thus, the null hypothesis can be accepted.

4.7.4 Self-reported outcome using the modified GHADP

Table 4 .24 shows the distribution of disability and handicap scores on the modified GHADP prior to hearing instrument fitting. A high score represents a greater disability and handicap. The median values for pre-intervention disability and handicap are 60% and 58% respectively.

The distributions of the subscale data show the same marked skewing as in the previous experiment. The distribution of the scores on these subscales is given in Table 4.25 and the median values are plotted in Figure 4.21 The *Use of Hearing Aid* subscale is in excess of 90%. This means that the subject reports wearing the hearing instrument for at least 90% of the time they are in a listening situation that, prior to fitting, led to hearing difficulty. The *Residual Disability* subscale is a measure of the extent to which

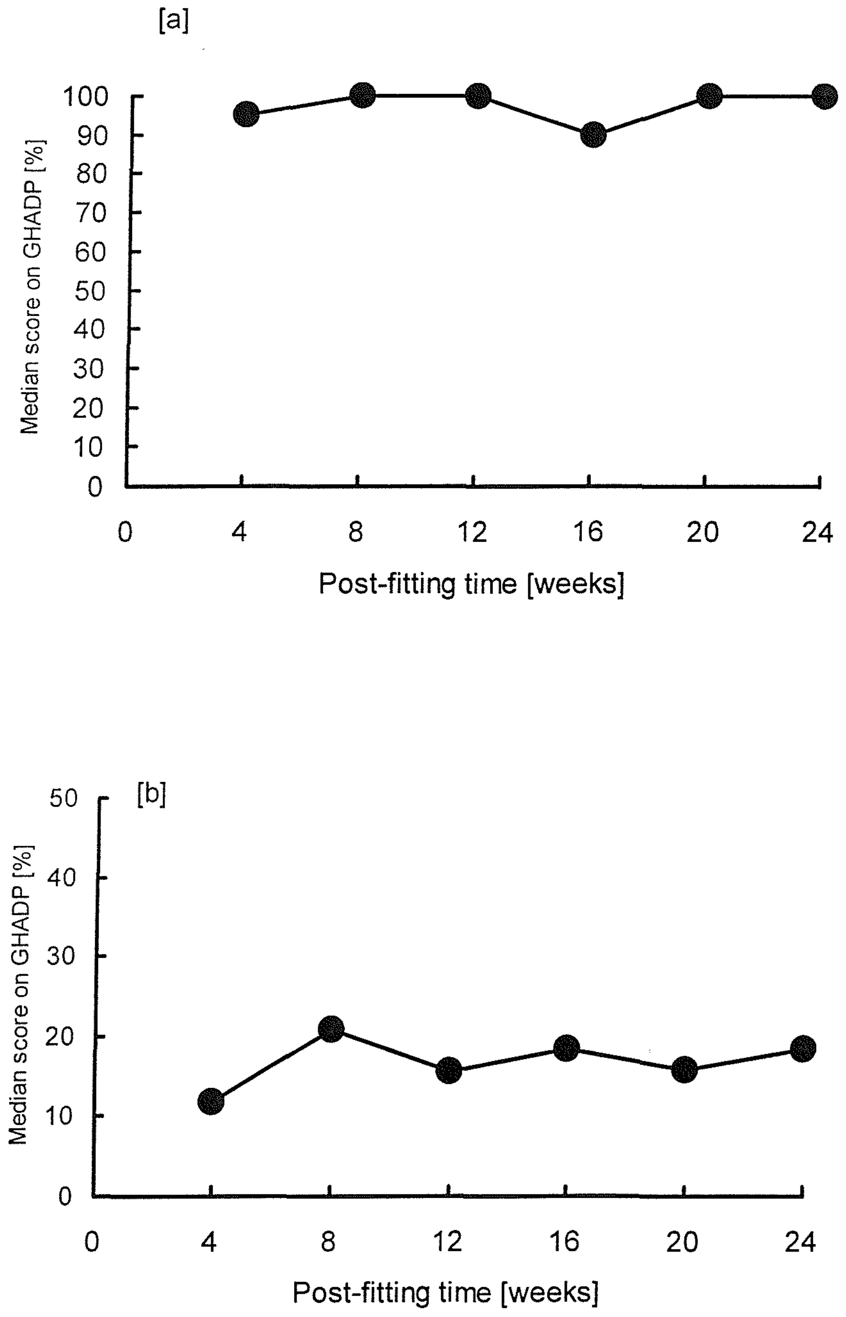
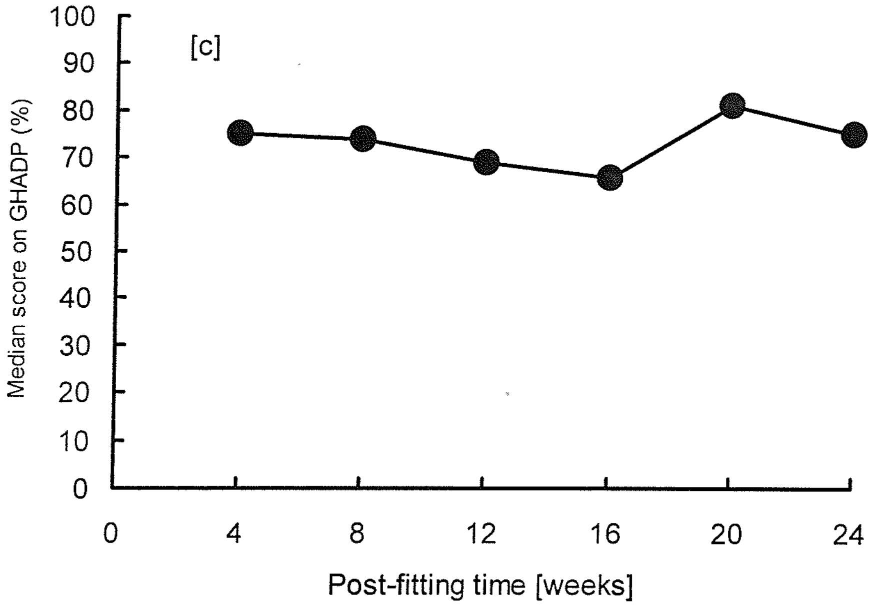


Figure 4.21. Median scores on the post-fitting scales of the modified Glasgow Hearing Aid Difference Profile. Use of Hearing Aid sub-scale is shown in panel [a]. The Residual Disability sub-scale is shown in panel [b]. The Additional Benefit (panel [c]) and Additional Satisfaction sub-scales (panel [d]) are shown overleaf.



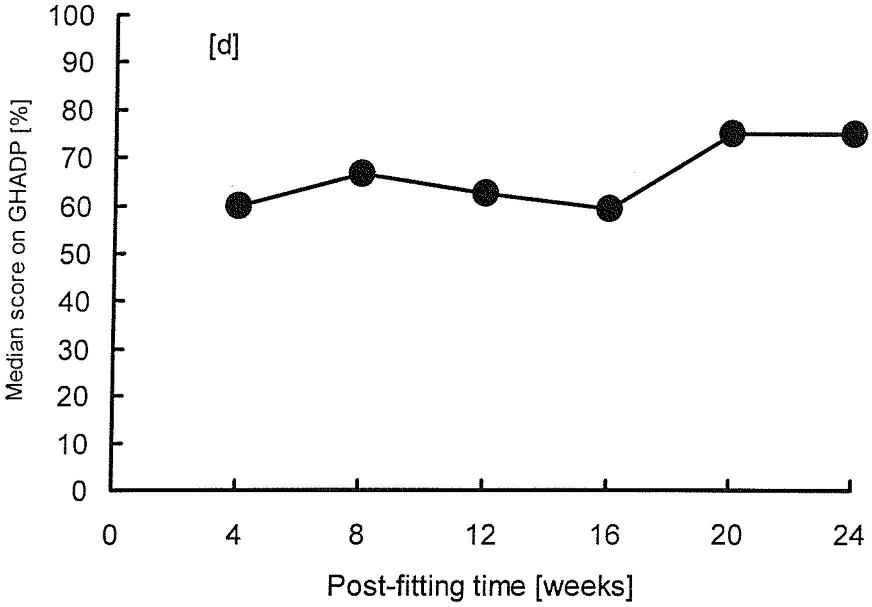


Figure 4.21. Cont'd

things are better post-fitting compared to pre-fitting. This is typically 10-20% and does not show any particular pattern of change over time. The *Additional Benefit* subscale refers to the change in benefit that occurred since the time of fitting. A score greater than 50% indicates that the subject reported more benefit whereas a score of less than 50% indicates less benefit. The median score is always greater than 50% suggesting that there was additional benefit present after four weeks of hearing instrument use but there were no further increases over time. A similar interpretation applies to the *Additional Satisfaction* subscale although the additional satisfaction is stable from week 4 to 16 but then increases for the remainder of the study.

Table 4.24. Distribution of pre-fitting disability and handicap scores on the GHADP [n=16].

Percentile	Initial disability	Initial handicap
25th	50	42
50th	60	58
75th	66	70

Table 4.25. Distribution of scores on the GHADP post-fitting subscales. Scores are given for Residual Disability, Use of Hearing Aid, Additional Benefit and Additional Satisfaction [n=16].

Post-fitting time [weeks]	4	8	12	16	20	24
Residual Disability						
25	3.8	6.7	4.7	1.6	10.0	13.5
50	11.8	20.8	15.6	18.3	15.8	16.7
75	20.0	34.2	26.1	28.0	25.0	28.6
Use of Hearing Aid						
25	70.0	92.5	96.9	64.6	78.8	95.0
50	95.0	100.0	100.0	90.0	100.0	100.0
75	100.0	100.0	100.0	100.0	100.0	100.0
Additional Benefit						
25	69.2	57.2	54.7	50.0	60.9	62.5
50	75.0	75.0	68.8	65.6	81.3	75.0
75	85.6	100.0	90.6	87.5	100.0	95.0
Additional Satisfaction						
25	56.0	52.5	55.9	50.0	50.0	50.0
50	60.0	66.7	62.5	59.4	75.0	75.0
75	75.0	97.5	81.9	93.8	90.6	100.0

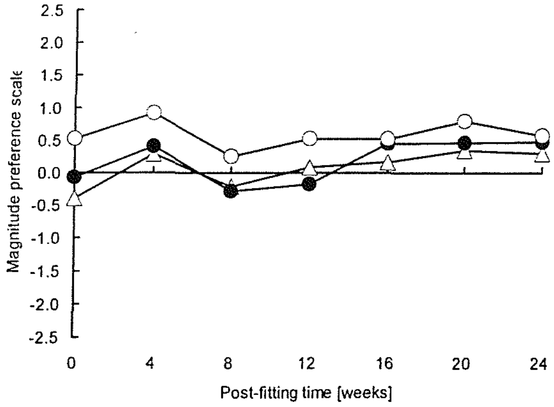
The focus of the analysis was concerned with the differences in subscales over time. These data conformed to a normal distribution and were analysed using parametric statistics. Linear interpolation was used to smooth the data to account for eight subjects who did not complete the questionnaire on a single test session. A one-factor repeated-measures ANOVA was performed on each subscale. The degrees of freedom were adjusted for univariate tests using the Greenhouse-Geisser correction for the factors that were statistically significant on Mauchly's test of sphericity. A summary of the analysis is given in Table XVIII, Appendix E. There was no statistically significant change over time for *Residual Disability, Additional Benefit* or *Additional Satisfaction*. Thus, the null hypothesis can be accepted. However, there was a statistically significant difference in the smoothed mean data for self-reported use of the hearing aid; inspection of the orthogonal polynomial breakdown revealed a significant cubic relationship [$F(1,15)=10.4$; $p<0.01$]. Figure 4.21 suggests that this is due to a slight reduction in use at 16 weeks post-fitting.

4.7.5 Paired comparison of sound quality judgements

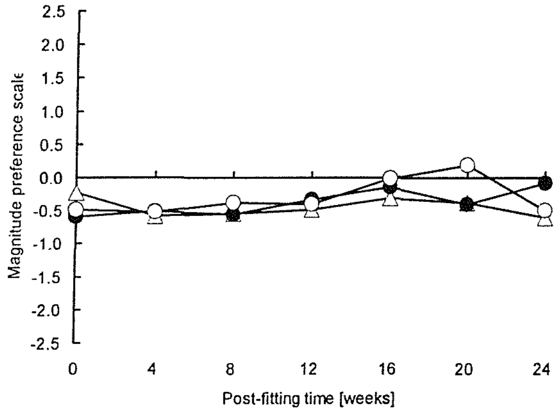
The paired comparison test involved the familiar frequency response versus a frequency response with less high frequency emphasis. Subjects were allowed to adjust the gain of each frequency response. Although the mean difference between the 4 kHz target values in a 2-cc coupler was 13 dB, the actual difference [after adjusting the gain] between the two frequency responses was 10-12 dB [see Table 4.26]. Table 4.27 gives the mean preference judgements for clarity, comfort and overall preference as a function of post-fitting time. The results are illustrated in Figure 4.22. The judgements were made in quiet, as well as in a background of steady speech-shaped noise and speech babble. The rating can range from +2.5 to –2.5. A positive value indicates the extent to which the subject prefers the fitted frequency response. All mean preference ratings were small. The subjects preferred the fitted frequency response when rating clarity but not when rating comfort or overall preference.

Table 4.26. Mean difference in 2-cc coupler gain at 4 kHz [in decibels] between the familiar frequency response and the alternative frequency response as a function of post-fitting time. A positive value indicates the extent to which the gain of the fitted response exceeds the alternative response. One standard deviation is given in parenthesis.

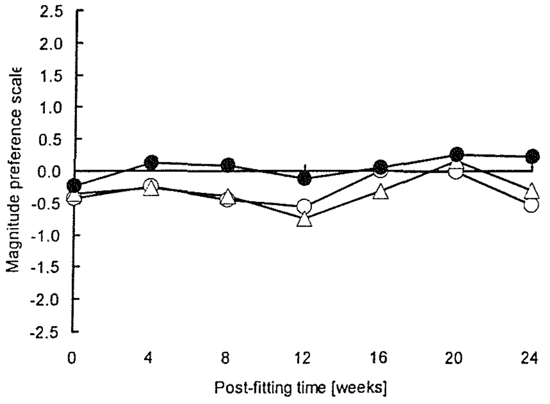
Post-fitting time [weeks]	0	4	8	12	16	20	24
Difference	11.9 [5.0]	12.2 [4.6]	12.1[4.2]	12.0 [3.3]	10.1 [3.4]	10.2 [4.0]	10.6 [3.9]



Clarity preference in quiet [triangles], steady noise [filled circles] and babble [open circles]



Comfort preference in quiet [triangles], steady noise [filled circles] and babble [open circles]



Overall preference in quiet [triangles], steady noise [filled circles] and babble [open circles]

Figure 4.22. Mean preference judgement for the fitted frequency response versus the alternative frequency response as a function of post-fitting time. Each subjective judgement was measured with speech in quiet [triangles], steady noise [filled circles] and babble [open circles]. A positive value indicates the extent to which the subject preferred the familiar frequency response. The top graph shows the preference for clarity, the middle graph shows preference for comfort and the bottom graph shows overall preference.

After confirming that the data conformed to a normal distribution, the results were analysed using a two-factor repeated-measures ANOVA for each of the three quality dimensions. The repeated-measures were background [quiet, speech-shaped noise and babble] and post-fitting time. The data from five subjects were smoothed using linear interpolation to account for a missing test session. The degrees of freedom were adjusted using the Greenhouse-Geisser correction for the interactions between background and post-fitting time since these were statistically significant on Mauchly's test of sphericity. The results of the analysis are summarised in Table XIX, Appendix E. There was no statistically significant difference for post-fitting time nor was there a significant interaction between background and post-fitting time. Thus, the null hypothesis can be accepted. However, there was a significant difference between backgrounds for clarity and overall preference. Observation of the mean data in Table 4.27 suggests that the preference in noise was different from the other background conditions. The subjects preferred the fitted frequency response when rating clarity in babble and when rating overall preference in steady noise.

Table 4.27. Mean paired comparison data with magnitude estimation for clarity, comfort and *overall preference for the familiar versus the alternative frequency response condition.* Each subjective judgement was measured with speech in quiet, speech in noise and speech in babble. A positive value indicates the extent to which the subject preferred the familiar frequency response condition. One standard deviation is given in parenthesis, n=16.

Post-fit [weeks]	0	4	8	12	16	20	24	Mean
			Clarity					
Quiet	–0.4 [1.0]	0.3 [1.2]	–0.2 [1.1]	0.1 [0.8]	0.2 [0.8]	0.3 [0.7]	0.3 [1.0]	0.1 [1.0]
Noise	–0.1 [0.9]	0.4 [0.8]	–0.3 [1.3]	–0.2 [1.0]	0.5 [0.7]	0.5 [0.8]	0.5 [1.0]	0.2 [1.0]
Babble	0.5 [1.2]	0.9 [1.0]	0.3 [1.2]	0.5 [0.8]	0.5 [0.8]	0.8 [0.6]	0.6 [0.9]	0.6 [0.9]
Mean	0.0 [1.1]	0.5 [1.1]	–0.1 [1.2]	0.1 [0.9]	0.4 [0.8]	0.5 [0.7]	0.4 [1.0]	0.3 [1.0]
			Comfort					
Quiet	–0.2[1.3]	–0.6 [1.1]	–0.5 [0.9]	–0.5 [0.9]	–0.3 [0.9]	–0.4 [0.8]	–0.6 [1.2]	–0.4 [1.0]
Noise	–0.6 [1.0]	–0.5 [1.2]	–0.6 [0.9]	–0.3 [1.1]	–0.1 [0.9]	–0.4 [1.0]	–0.1 [1.3]	–0.4 [1.0]
Babble	–0.5 [1.2]	–0.5 [0.9]	–0.4 [0.8]	–0.4 [1.1]	0.0 [0.6]	0.2 [1.0]	–0.6 [1.3]	–0.3 [1.00
Mean	–0.4 [1.1]	–0.5 [1.1]	–0.5 [0.9]	–0.4 [1.0]	–0.2 [0.8]	–0.2 [1.0]	–0.4 [1.1]	–0.4 [1.0]
			Overall preference					
Quiet	–0.4 [1.0]	–0.3 [1.1]	–0.5 [0.8]	–0.6 [0.9]	0.0 [0.6]	0.0 [1.0]	–0.5 [1.1]	–0.3 [0.9]
Noise	–0.4 [1.1]	–0.3 [1.4]	–0.4 [0.5]	–0.7 [0.9]	–0.3 [1.0]	0.2 [1.2]	–0.3 [1.0]	–0.3 [1.1]
Babble	–0.2 [1.1]	0.1 [1.2]	0.1 [0.6]	–0.1 [0.9]	0.0 [1.0]	0.3 [1.1]	0.2 [1.1]	0.1 [1.0]
Mean	–0.3 [1.0]	–0.1 [1.2]	–0.3 [0.7]	–0.5 [1.0]	–0.1 [0.9]	0.1 [1.1]	–0.2 [1.0]	–0.2 [1.0]

4.7.6 Summary of the statistical analysis

The results of the statistical analysis can be summarised as follows:

FAAF test

- There was no statistically significant increase in the not-fitted control ear over time: the difference between week 0 and 24 was +0.6%.
- The mean scores in the fitted ear at the fixed-gain and user-gain setting increased over time by around +2% but this was not statistically significant.
- There was no statistically significant difference between performance with the fixed-gain and the simulated fixed-gain setting.
- There was a statistically significant difference in performance between the user-gain and the simulated user-gain setting.

Modified GHADP

- There was no statistically significant difference over time on disability, benefit or satisfaction subscales.

Sound Quality Judgements

- There was no statistically significant difference over time for any of the quality judgements.
- The subjects preferred the fitted response when rating clarity in babble and overall preference in steady noise.

4.8 Discussion

This study examined the mean aided FAAF score over a post-fitting time course of 24 weeks. Subjects were fitted monaurally and the not-fitted ear was used as the control condition. Unlike the previous experiment, subjects were not given any visual feedback about their performance. One main aim of the study was to determine if disabling the feedback would reduce the practice effect. A lack of systematic increase in the mean aided score in the control ear would rule out a practice effect. A second aim was to establish if the difference in performance between the gain settings that was observed in the previous experiment was repeatable. Related to this last point, the subjects were also tested at both gain settings but with adjustments made to the speech input level in order to mimic the amplified speech level obtained with the other gain setting. Thirdly, the subjects completed the modified GHADP with changes over time reported relative to the time of fitting. In the previous study, there were self-reported improvements in benefit and in satisfaction when the comparison was made relative to the previous visit. If the same improvements could be obtained with both approaches then this would add confidence in the interpretation of self-reported changes over time. Finally, the opportunity was also taken for subjects to perform a paired comparison of the fitted response with an alternative response with less high frequency emphasis.

4.8.1 Changes in performance in the control ear

The not-fitted ear was used as the control condition to check for a practice effect. A practice effect may be the result of familiarisation with the task and/or materials. For example, subjects may learn to sit more quietly [reducing interfering noise and distractions] or anticipate the timing of the test items [so that they know when to settle and focus attention on the test's auditory cues]. When assessing true perceptual learning it is important to control for potentially confounding test-learning issues.

Testing in the control ear was performed with the hearing instrument at the fixed-gain setting. The mean increase in the FAAF test score over time, relative to the time of fitting, was less than 1%. Thus, the practice effect observed in the first experiment was greatly diminished. The changes that occurred in the fitted ear [discussed in the next section] were also less than those observed in the previous study. This is consistent with the conclusion that removing the visual feedback has reduced the practice effect. Previously, the correct word flashed on the subject's monitor after they made their response. It is likely that the use of feedback trained the subject for the reasons discussed in section 4.3.1. This suggests that visual feedback should be used with caution in studies that require multiple presentations of the test material over time. An alternative would have been to include additional pre-fitting sessions to allow this practice effect to reach an asymptote before fitting. A simpler solution would be to ensure that feedback is not used in future studies. The results of the present study are consistent with the earlier reports that the FAAF test is robust to pure response learning provided visual feedback is not used. An unexpected finding was the decrease in performance at week eight relative to the time of fitting: this is discussed in the next section.

4.8.2 Performance in the fitted ear at the fixed-gain and user-gain setting

Aided FAAF scores were obtained in the fitted ear for two gain settings. For one setting, the gain control was fixed to provide a constant level of speech audibility across all test sessions. For the second setting, the subjects were allowed to adjust the gain control on each visit prior to data collection. This was done to reflect any changes in preferred gain that may have occurred with experience. There was no statistically significant interaction between ears and post-fitting time for either of the gain settings. This indicates that the aided scores in each ear were not differentially affected by time.

Despite the lack of statistical significance, there was a trend for the scores in the fitted ear to increase over time. For example, the mean score at the time of fitting for the fixed-gain condition was 70.8% but the overall mean score [averaged over all test sessions] was 72.5%. There was also a difference between ears. For example, at the fixed-gain setting, the initial difference in mean score between ears was 0% but by week 8 it was 5.4% and by week 12 it was 4.4%. If the post-fitting time period had been restricted to 12 weeks [as in the Gatehouse study] then there would have been a statistically significant difference in performance between the two ears and there would have also been a significant interaction between ear and post-fitting time. The difference over time in the fitted ear as well as the difference between ears, at least over the initial 12-week period, is consistent with auditory acclimatisation although the effect size is smaller than that reported in previous studies. There is also an indication that a similar finding was present on the first experiment: the difference between the mean aided score for the fitted ear [at user-gain] and the control ear was 2.7% at the time of fitting but this had increased by a further 2% some 9-12 weeks later. The findings discussed above are consistent with a process of auditory acclimatisation but are smaller than those reported by others in the literature.

The biggest difference between ears occurred at eight weeks post-fitting. Figure 4.23 shows a scatter plot of the aided scores at the time of fitting and at eight weeks post-fitting. Two subjects did not complete the week eight test session so only 14 data points are present. The open symbols correspond to the not-fitted control ear and the filled symbols correspond to the fitted ear. If there was no change in mean FAAF test score between the two test sessions then all of the points would have fallen on the diagonal line. There is a tendency for the individual scores in the control ear to decrease and for the scores in the fitted ear to increase. The mean changes in score in each ear were –3.6% [±5.4] and 1.8% [±4.6] respectively. Seven (50%) subjects showed a difference of greater than two standard deviations from the week zero values. There were no obvious differences between the seven subjects who showed the largest change and the remaining subjects. For example, both groups of subjects had similar hearing threshold levels, insertion gain and self-reported daily use of the hearing instrument. Bentler *et al.* [1993a] did not show any statistically significant increase in performance on speech recognition tests over a post-fitting time period of 12 months. However, a close examination of their data [see their Table 5] shows a trend of increasing performance on both the low predictability SPIN sentences as well as the NST in noise. There was a mean increase in the initial aided score of around 4-4.5% at 1 and 3 months post-fitting before returning to within 2.5% of the initial score from 6 months onwards. Thus, this pattern mirrors that obtained in the fitted ear in the present study. Figure 4.24 shows a scatter plot of the difference in score between ears (test minus control) at week zero and week eight. The mean differences between week 0 and 8 were 0% [±2.6] and 5.4% [±6.5] respectively. These scatter plots also confirm that the decrement in score in the control ear is not simply due to a few outliers.

The only published study that has used the not-fitted ear for the control condition was Gatehouse [1992a]. That study showed a statistically significant decrement in aided performance in the control ear of around 3% using a headphone simulation of the hearing aid response[2]. The decrement occurred within the initial 6-8 week post-fitting period but there was no sign of recovery in performance in the control ear when the

---

[2] Performance in the sound field using the hearing aid did not reveal a statistically significant change over time. However, the sound field presentation involved stimulation to both ears, and, therefore, did not attempt to study changes over time between the ears.

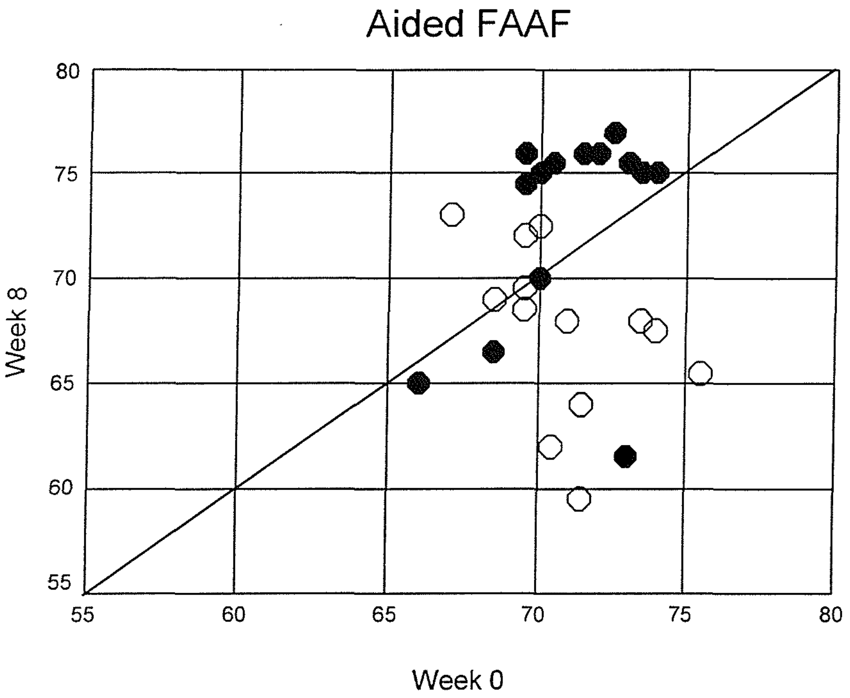


Figure 4.23. Scatter plot of aided FAAF score at week zero and week eight. The open circles correspond to the not-fitted control ear and the filled circles correspond to the fitted ear.

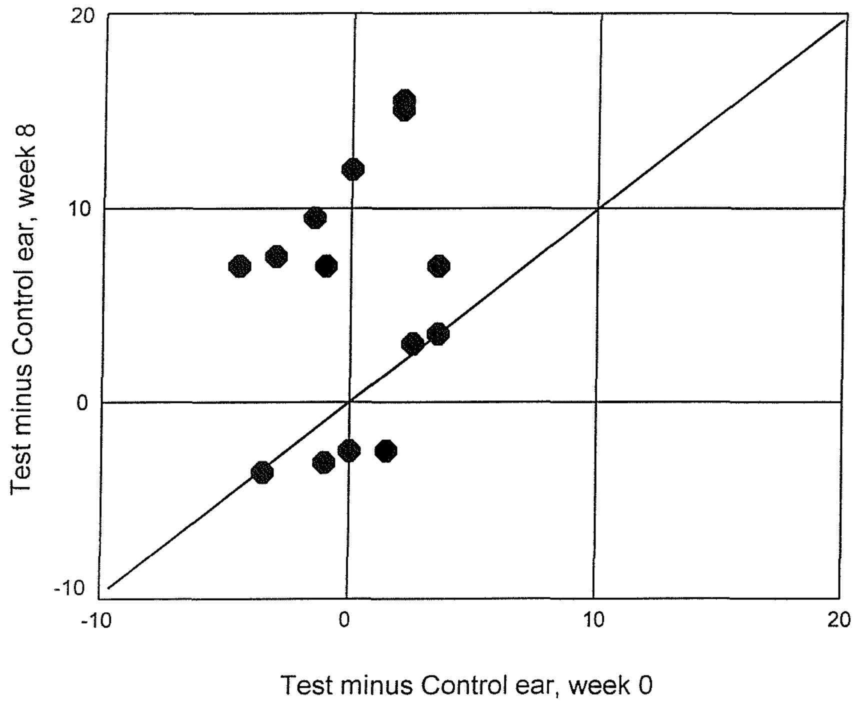


Figure 4.24. Scatter plot of the inter-aural difference score [Fitted ear minus the not-fitted control ear] at week zero and week eight.

study terminated at 12 weeks post-fitting. Studies on auditory deprivation also show a decrement in performance in the not-fitted ear but the time period is usually considerably longer, typically years, and there is no spontaneous recovery.

One explanation for the decrease in performance in the control ear is that subject motivation was poor at eight weeks post-fitting. However, if this was the case, then performance may have been expected to decrease in both ears. Perhaps the transient reduction in the control ear is the result of the auditory system 'over reacting' to the sudden inter-aural asymmetry in auditory stimulation and this results in an initial decrease in performance in the not-fitted ear.

Table 4.28. Summary of acclimatisation studies showing the initial, subsequent and final benefit scores. Some of the scores were estimated from graphs in the relevant publications. Studies that did not give benefit scores have been excluded.

Study	Initial	Subsequent	Overall
Cox *et al.* [1996]	5	5	10
Gatehouse [1992]			
Sound field	6	9	15
Simulated under phones	4	13	17
Taylor [1993]	14	0	14
Humes *et al.* [1995]	15	0	15
Cox and Alexander [1992]			
subjects tested over 12 weeks	4	6	10
sub group tested over 6 months	0	8	8
Horwitz and Turner [1997]			
fixed gain	7	7	14
daily adjusted gain	6	7	13
Neuman *et al.* [1997]	16	0	16
Surr *et al.* [1997]			
50, +10 SNR	40	0	40
60, reverberation	40	0	40
60, +5 SNR	30	0	30
70, +2 SNR	10	0	10
Bentler *et al.* [1999]	10	0	10
Munro [first experiment]	14	0	14

A possible explanation for the smaller than expected effect size could be related to the speech presentation level. Table 4.28 summarises the benefit scores from previous studies [with some data extracted from graphs]. There are other studies on auditory acclimatisation but these do not provide benefit scores. Gatehouse [1992] configured the FAAF test by setting the SNR so that the initial benefit score was of the order of 5% for sound field testing and around 2% for headphone testing. Horwitz and Turner [1997] showed a mean initial benefit score of around 6-7 rau that increased to around 14 rau at the end of the study. Cox *et al.* [1996] showed an initial benefit score of around 4-5 rau that doubled after 12 weeks. On the other hand, Humes *et al.* [1996] reported a mean initial benefit score in excess of 10-15% and this did not show any subsequent increase over time. Similarly, Neuman *et al.* [1997] showed an initial benefit score of 16% and the scores of Surr *et al.* [1998] were even greater. The data in Table 4.28 show that it is only studies that report a small initial benefit score that subsequently show an increase in benefit over time. The one exception to this trend was the negative finding of Holte *et al.* [cited by Bentler *et al.*, 1999] who showed an initial benefit score around 10%. However, the potential weaknesses of this study were discussed in section 2.4.2. Although the unaided score was not measured in experiment two, the methodology was similar to experiment one where the initial benefit score was 14% with no further increase over time. This is consistent with the conclusions drawn from Table 4.23; that is, changes over time may only be measurable when the difference between the initial aided and unaided scores are small.

A model that explains how the difference in benefit scores are related to presentation level is illustrated in Figure 4.25. This is based on hypothetical data but illustrates the influence of presentation level on performance. The filled circles show the aided scores and the open circles show the unaided scores. Performance increases in the unaided condition as the presentation level increases. This is because a greater proportion of the speech signal is audible at these relatively high presentation levels. Performance in the aided condition does not change with increasing presentation level because the fixed SNR means that the same proportion of the speech signal is audible across all levels. In fact, some studies have shown a rollover in aided speech performance at high presentation levels and the SII model supports this finding. The studies that have shown an improvement in performance over time are located on the right side of the graph. This suggests that acclimatisation is related to presentation level: it can be measured at relatively high presentation levels but not for relatively low presentation levels. If this is correct, then auditory acclimatisation cannot be related to changes in audibility *per se*, since the smallest change in audibility occurs with high presentation

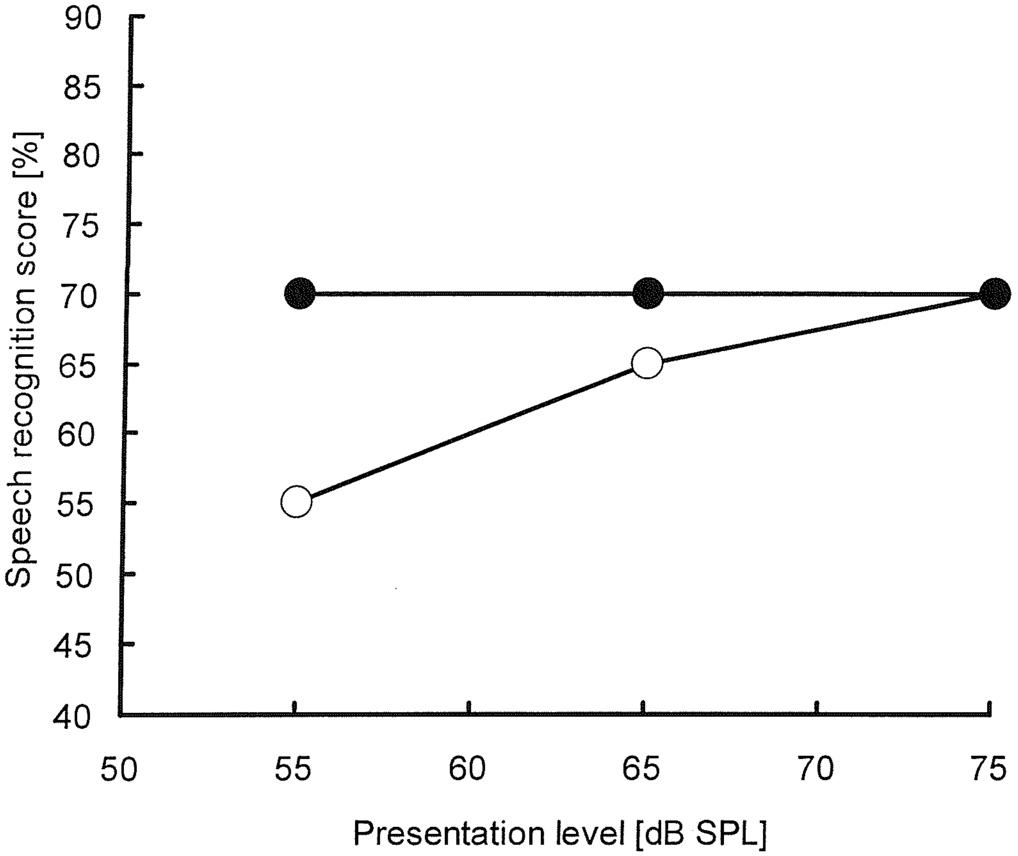


Figure 4.25. Model relating speech recognition scores to presentation level. The open circles correspond to the unaided condition and the filled circles correspond to the aided condition.

levels. The important factor may be the new experience offered by the hearing instrument. For example, consider a subject who experiences speech in everyday life over the range 55-75 dB SPL before fitting and 75-95 dB SPL after fitting i.e., the hearing instrument provides 20 dB gain. [For simplicity, the shape of the gain/frequency response curve is not considered here.] If the subject is then tested aided with speech at 60 dB then there is very little difference in the overall level of speech that was experienced before providing amplification [since amplified speech will be around 80 dB SPL]. However, if testing is performed with speech at 70 dB SPL then this will be amplified to a level not previously experienced by the subject. The important point is not simply that the hearing instrument improves audibility but that performance is measured with speech amplified to a level that was not experienced, in everyday life, prior to aiding. This may explain why improvements in performance, over time, are more likely to be detected at higher rather than lower presentation levels.

In experiment one there was a statistically significant difference between the scores obtained in the fitted ear at the two gain settings six weeks post-fitting. The fixed-gain setting provided, on average, 7 dB more gain but the mean aided score was 3.7% lower than at the user-gain setting. The reason for this difference was unclear. In experiment two the greatest difference in aided scores between the two gain conditions occurred at week four. The fixed-gain setting provided, on average, 6 dB greater gain than the user-gain setting but, this time, the mean score was 1.8% higher. It is possible that acclimatisation was present in the previous two experiments but the presentation level was not ideal and hence, only small and transient changes were measured. Thus, it was not possible to replicate the change over time between gain settings that occurred in the previous experiment.

4.8.3 Performance in the fitted ear at the simulated gain settings

In experiment one, a higher aided FAAF score was obtained at the lower user-gain setting. Although the difference in gain was small, it was unclear if the higher score was a reflection of acclimatisation to the normal listening condition. If this was correct, the score obtained at the fixed-gain setting should be similar to that obtained at the user-gain setting with the speech input level adjusted to maintain the same amplified speech level. The mean aided FAAF score at the fixed-gain setting was compared with the score obtained at the lower user-gain setting but with the speech input level increased to maintain the same amplified speech level. Essentially, the speech input level was increased by around 5 dB to account for the reduction in gain. The scores for both conditions were very similar and there was no statistically significant difference between gain condition, post-fitting time and no interaction.

A similar comparison was made with the scores obtained at the user-gain setting and at the simulated user-gain setting [fixed-gain setting with the speech level reduced by around 5 dB]. Although the scores were similar across most test sessions, there were differences that were statistically significant. The largest difference was 2.6% and this occurred at week 16. On close inspection of the raw data, it was noted that one subject [subject number 32] preferred a user gain some 16-22 dB below the fixed-gain setting and this was most marked at week 16. The score for this subject at week 16 was 72% at the user-gain setting but only 52.5% at the simulated user-gain setting. The poorer score for the simulated condition may have been the result of masking by instrument noise since the presentation level of speech was reduced from 57 to 35 dB SPL. The speech signal may have been masked by environmental sounds such as electrical hum from the amplifier. When the statistical analysis was repeated with this subject excluded, the difference in scores at week 16 reduced from 2.6% to 1.5% and the ANOVA was no longer statistically significant [$F(1,14)=4.21$, $p>0.05$].

4.8.4 Self-reported outcome using the modified GHADP

The GHADP produces both pre- and post-intervention scales on a range from 0 to 100. The median scores on the pre-intervention scales of disability and handicap were 60 and 58 respectively. Report of ‘moderate’ and ‘great’ difficulty would result in scores of 50% and 75% respectively. These median scores are at the upper end of the inter-quartile range in experiment one but this is probably due to the small sample size, and hence wide confidence interval, in each experiment. Once again, the median scores are higher than those reported by Gatehouse [1999].

The median score for self-reported use of the hearing aid was is in excess of 90%. This corresponds to 50% of subjects using their hearing aids for almost all of the time they find themselves in listening situations that prior to receiving the hearing instrument was resulting in hearing difficulty. This agrees with the subjects reporting 8-12 hours daily use of the hearing instrument. Given the high motivation of the subjects and their participation in this study, it is not surprising that the hearing aid use score is greater than the median score of 70.2 reported by Gatehouse [1999].

The median score for residual disability after intervention showed small variations across time but was between 12 and 21. This corresponds to a self-report of 'none' or only 'slight' residual disability. The subjects were fitted with programmable hearing aids and received substantial pre- and post-intervention support. This probably explains why the outcome is a little more favourable that that reported by Gatehouse but consistent with the clinical trials reported by Lutman and Payne [1999] and Wood and Lutman [2001].

The final two subscales were concerned with changes in benefit and satisfaction over time. The subjects were asked to compare present level of benefit and satisfaction with the levels experienced immediately after fitting. The additional benefit score varies between 66 and 75%. This corresponds to 'a little more' benefit than immediately after fitting [no change would have resulted in a score of 50%]. This additional benefit is present from four weeks post-fitting and shows no statistically significant increase over time. The results for additional satisfaction vary between 60 and 75%. Although there is a trend of increasing satisfaction over time, it was not statistically significant. A difference of 25% is required for a categorical change [for example, 50% means 'no change' whereas 75% means 'a little better']. Once again, the small additional benefit is present from four weeks post-fitting and shows no statistically significant increase over time.

The perceived improvement in benefit and satisfaction was not mirrored by a reduction in residual disability over time. This may be because of the relative insensitivity of the residual disability score due to a floor effect; essentially, there was not enough room to reveal a further reduction in disability. Alternatively, it may be that the relative change in benefit and satisfaction is a more sensitive indicator of the subtle changes that occurred during the initial four weeks post-fitting. However, there is concern about the reliability and validity of gathering retrospective information.

Paykel [1983] published a review article that questioned the reliability and validity of information that was gathered retrospectively. The reliability of reporting major events in life such as illness was surprisingly poor even when the interval between self-report was as short as two weeks. For example, only 50% of events reported by subjects at the start of the study were reported when the exercise was repeated 10 months later. In addition, when the number of self-reported events was compared with those reported by another individual [usually a family member], the agreement was only 35%. The reliability and validity were improved by using interview techniques instead of using self-report questionnaires. However, this required detailed probing with recorded interviews [that lasted as much as half a day] with judgements on the content made later. This suggests that a comparison of present situation with a previous occasion may be unreliable.

The additional benefit and satisfaction scores in experiment two are different from experiment one where there was evidence of continuing improvement over a 12-week post-fitting period. In experiment one, the subjects were asked to make a comparison with the previous visit three weeks earlier. It is not known which of these approaches result in the correct response although they may both be equally unreliable. A number of previous studies have used self-report to measure changes over time. While these studies have suggested that additional improvements have occurred, they are subject to the same concerns over reliability and validity. Additionally, a change over time may simply be due to the subject becoming more familiar with the hearing aid [for example, the subject has become more proficient with adjustment of the volume control in different listening situations]. There is no easy and reliable method to untangle these two possibilities. For this reason, the use of self-report is likely to be of limited benefit in studies investigating changes in benefit that are specific to auditory acclimatisation.

4.8.5 Paired comparison of sound quality judgements

Judgements for clarity, comfort and overall preference were performed with the fitted frequency response and an alternative response with less high frequency emphasis. In general terms, the subjects preferred the familiar high frequency response when rating clarity but not for comfort or overall preference. There was no difference in preference rating over time for any of the quality dimensions or background noise conditions. Therefore, there is no evidence of acclimatisation to different frequency responses. This may be due to the small difference between the two frequency responses.

4.9 Conclusions

The FAAF test was presented to the subject without visual feedback. There was no evidence of a practice effect in the control ear. Visual feedback should not be used in studies that require multiple presentation of the test material over time. Alternatively, greater practice could be given with the test material before hearing instrument fitting. The FAAF test can be used in further experiments without interpretation of the results being hampered by a practice effect.

There was no statistically significant interaction between ear and post-fitting time for the fixed-gain or the user-gain condition. However, there was a trend of increasing performance in the fitted ear but decreasing in the control ear. If the statistical analysis had been restricted to 12 weeks instead of 24 weeks then there would have been a significant interaction. Only aided performance was measured in this experiment although the test conditions were similar to the previous study where mean benefit was around 14%. Only studies that report a small initial benefit score have subsequently demonstrated improvements over time. This suggests that measurement of auditory acclimatisation might be related to presentation level; the important point is not simply that the hearing instrument improves audibility but that performance is measured with speech amplified to a level that was not experienced, in everyday life, prior to aiding. This possibility is investigated in Experiment Three.

There was no change in outcome on the modified GHADP when subjects were asked to compare performance to the time of fitting. This finding is different from the previous experiment when the subjects were asked to compare performance to the previous visit. It is not known which, if any, of these approaches is correct. The use of self-report is likely to be of limited benefit in studies investigating auditory acclimatisation since changes could be due to factors such as increased familiarity with the hearing instrument controls or errors in retrospective reporting. There is no evidence for acclimatisation on paired comparison tests but this may be because the difference between the two frequency responses was small.

CHAPTER FIVE

PERFORMANCE AS A FUNCTION OF PRESENTATION LEVEL

5.0 Introduction to experiment three

Experiments one and two failed to demonstrate a sustained acclimatisation effect. In both experiments the FAAF test was presented at an overall level of 60 dB SPL. Mean performance increased from 56% in the unaided condition[1] to 70% in the aided condition [i.e., a mean benefit of 14%]. Studies that have shown an acclimatisation effect have reported a much smaller initial benefit score, typically less than 7%. A small benefit score suggests that the hearing instrument was providing little gain or, more likely, speech was presented at a relatively high SPL before aiding. It is not straightforward to compare the speech presentation level directly across studies because of differences in calibration procedures and materials. For example, Horwitz and Turner [1997] presented speech at an overall level of 70 dB SPL whereas Gatehouse [1992] presented speech at 65 dB SPL defined for a 1 kHz calibration tone that had a SPL equal to the mean of the peaks of the test words.

Gatehouse [1989] has reported level-dependent differences in SNR between the two ears of subjects fitted monaurally: the fitted ear required a more favourable SNR at low presentation levels compared to the not-fitted control ear while the converse was true for higher presentation levels. In addition, Robinson and Gatehouse [1995, 1996] have reported level-dependent effects for intensity discrimination in monaurally aided subjects: the fitted ear had better discrimination at high presentation levels but the control ear was better at low presentation levels. Thus, there is evidence that auditory acclimatisation is dependent on presentation level.

No studies have reported speech recognition performance, over time, as a function of presentation level. The aim of this experiment was to measure aided and unaided speech recognition scores over a range of presentation levels that could realistically be expected to occur in everyday life. The model in Figure 4.25 shows a range of vocal

---

[1] Unaided performance was only measured in the first experiment.

efforts to represent those experienced by a hypothetical subject in everyday life. Values of 55, 65 and 75 dB SPL were used for illustrative purposes to represent quiet, average and raised speech. The model assumes that there is no change in these levels once the subject has been provided with amplification i.e., it does not account for the possible reduction in vocal effort required when communicating with the subject after aiding. It also does not attempt to account for changes in frequency spectra with different vocal efforts before and after amplification. The explanation for the decreased benefit with increasing vocal effort was given in section 4.8.2. The important point is that amplified quiet speech is raised to a level that approximates the level of raised speech before aiding whereas amplified raised speech will be received at a higher [and less familiar] level than previously experienced before aiding. Thus, the change in overall listening level with amplification will be greater for raised speech than for quiet speech and hence, the opportunity to learn will be greatest for raised speech relative to lower speech levels. The null hypothesis was that there would be no differential change in benefit, over time, as a function of presentation level. The alternative hypothesis was that improvements over time would be greatest at higher presentation levels where initial benefit scores would be small. Since we were interested in changes in performance with presentation level, it was important that these were not confounded with changes in hearing instrument gain setting. Therefore, it was decided to only use subjects who agreed to have the gain control disabled for the duration of the study. The subjects were tested over a 12-week post-fitting period. This was shorter than the previous experiments but consistent with the duration used in published studies [for example, Gatehouse, 1992a; Cox *et al.*, 1996] and sufficient to show differences in acclimatisation with presentation level. The subjects completed all test conditions around the time of fitting and again at 6 and 12 weeks post-fitting. Testing at six weeks post-fitting was included as this allowed regular contact with the subject. It also enabled a more detailed measure of the time course of any changes in performance.

5.1 Additional methodology

Sixteen subjects were recruited [10 male, 6 female] with a mean age of 70 years [± 5.5]. Six subjects were fitted in the right ear and 10 subjects were fitted in the left ear.
No subjects withdrew from the study and no subjects missed any test session. Although experiment two matched the hearing instrument response to DSL target values,

experiment three used the NAL-RM target values. While the NAL-RM method is the most widely used prescription method in clinical practice [Martin *et al.*, 1998] and in previous acclimatisation studies, the main justification for this change was that for a given hearing impairment, the overall loudness of NAL-RM target values is less than for the DSL target and the specific loudness pattern is flatter. This should increase user comfort and reduce the need to make adjustments to the gain control [Moore and Glasberg, 1998]. The NAL-RM targets provide less amplification than DSL at the high frequencies and this should increase the headroom and reduce the risk of distortion from the hearing instrument with higher input levels.

Table 5.1. Mean real-ear insertion gain [dB] at NAL-RM target, best match to target and at the preferred user setting, as a function of frequency, in the test ear. One standard deviation is given in brackets.

Frequency [Hz]	250	500	750	1000	1500	2000	3000	4000
Target [dB]	3 [3.4]	10 [6.6]	16 [5.7]	22 [5.1]	23 [4.2]	22 [3.2]	23 [2.5]	24 [2.7]
Match [dB]	3 [2.8]	8 [7.5]	16 [4.7]	21 [4.6]	24 [5.1]	24 [4.5]	23 [2.5]	22 [4.2]
User [dB]	1 [1.9]	3 [5.2]	12 [5.2]	17 [4.6]	20 [6.5]	20 [6.1]	19 [2.5]	17 [4.7]

The hearing instrument response was tailored to the NAL-RM REIG target values using the REIG protocol on the Rastronics Portarem 2000 probe-tube measurement system. This involved measuring the REUR and the REAR from each subject. Figure 5.1 shows the mean target values for the fitted ear along with the best match to target. The numerical data for the mean and standard deviations are given in Table 5.1. The mean target gain between 1-4 kHz [where the hearing impairment is greatest] is 23 dB. There is very good agreement between the target values and the best match: the mean difference is typically less than 1-2 dB. Also shown is the user-gain setting that was fixed for the duration of the study. The subjects were free to adjust the gain control of the hearing instrument during the first few days after fitting before returning to the laboratory to have this disabled at their preferred user-gain setting. All adjustments were made within seven days of fitting. The mean user-gain was 4 dB below the target gain from 1-4 kHz. It is possible that subjects may have been conservative in selecting their preferred gain because they were aware that this would be fixed for the duration of the study. However, other studies have reported a similar reduction in user gain compared to NAL-RM target values [Humes *et al.*, 2000]. The maximum output was set close to the subject's uncomfortable loudness level and adjusted if the subject reported any undue discomfort. These settings were entered into memory one of the

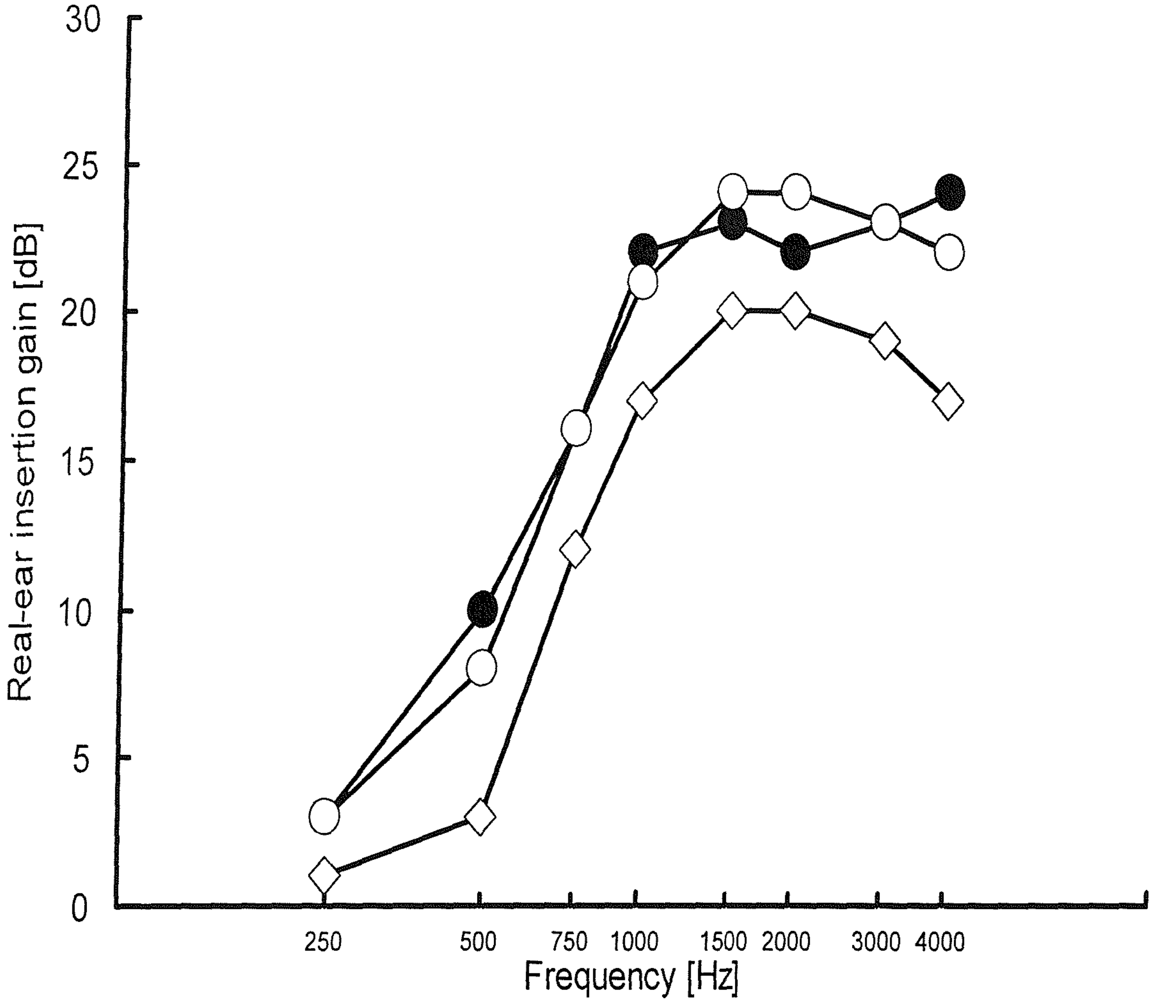


Figure 5.1 Mean real-ear insertion gain for the fitted ear. The NAL-RM target gain, the best match to NAL-RM target and the actual user gain are shown by the open circles, closed circles and open diamonds respectively.

hearing instrument as this was the default memory used by the subject in their home environment.

The REIG and the HA2 2-cc coupler gain were measured at the time of fitting and again at 6 and 12 weeks post-fitting. The results are shown in Table 5.2 and 5.3 respectively. The mean change in gain was less than 1 dB. The standard deviation was higher for insertion gain than for 2-cc coupler gain [typically 2.5 dB and 1.5 dB respectively]. While it is possible that there were real changes in REIG, the most likely reason is that probe-tube measurements are inherently more variable than coupler measurements. There are no studies available that report long-term test-retest variability for REIG; however, the standard deviations in the present study are typical of those reported for short-term test-retest data by Hawkins *et al.* [1991]. Thus, each subject was provided with relatively constant gain.

Table 5.2. Mean change in real-ear insertion gain [dB] at the user setting, relative to the time of fitting, for the fitted ear. One standard deviation is given in brackets.

Frequency [Hz]	250	500	750	1000	1500	2000	3000	4000
Week 6 [dB]	0 [1.7]	0 [2.7]	0 [1.8]	0 [2.8]	0 [2.0]	0 [2.7]	0 [2.4]	1 [1.7]
Week 12 [dB]	0 [2.3]	0 [2.1]	1 [1.7]	0 [2.2]	1 [1.5]	0 [3.3]	–1 [2.7]	2 [3.6]

Table 5.3. Mean change in 2cc-coupler gain [dB] at user setting, relative to the time of fitting, in the fitted ear. One standard deviation is given in brackets.

Frequency [Hz]	250	500	750	1000	1500	2000	3000	4000
Week 6	0 [1.5]	0 [1.4]	–1 [1.2]	0 [1.2]	–1 [1.6]	–1 [1.4]	–1 [1.6]	–1 [1.5]
Week 12	1 [1.0]	2 [1.6]	0 [1.1]	0 [1.5]	0 [1.3]	–1 [1.6]	–1 [1.0]	–1 [0.9]

The relationship between hearing threshold level and speech before and after amplification [with an overall SPL in the undisturbed sound field of 65 dB] is illustrated in Figure 5.2. The auditory dynamic range is shown as filled circles [hearing threshold, lower circles; uncomfortable loudness level, upper circles]. The long-term average spectrum of speech [LTASS] from Cox and Moore [1988] is shown as squares [unamplified, open squares; amplified, filled squares]. The spectrum of the FAAF test is similar to the LTASS [as shown in Appendix A]. The peaks and troughs of the LTASS have been omitted for simplicity. The real ear saturation response of the hearing instrument was measured with the gain control set to maximum and a swept

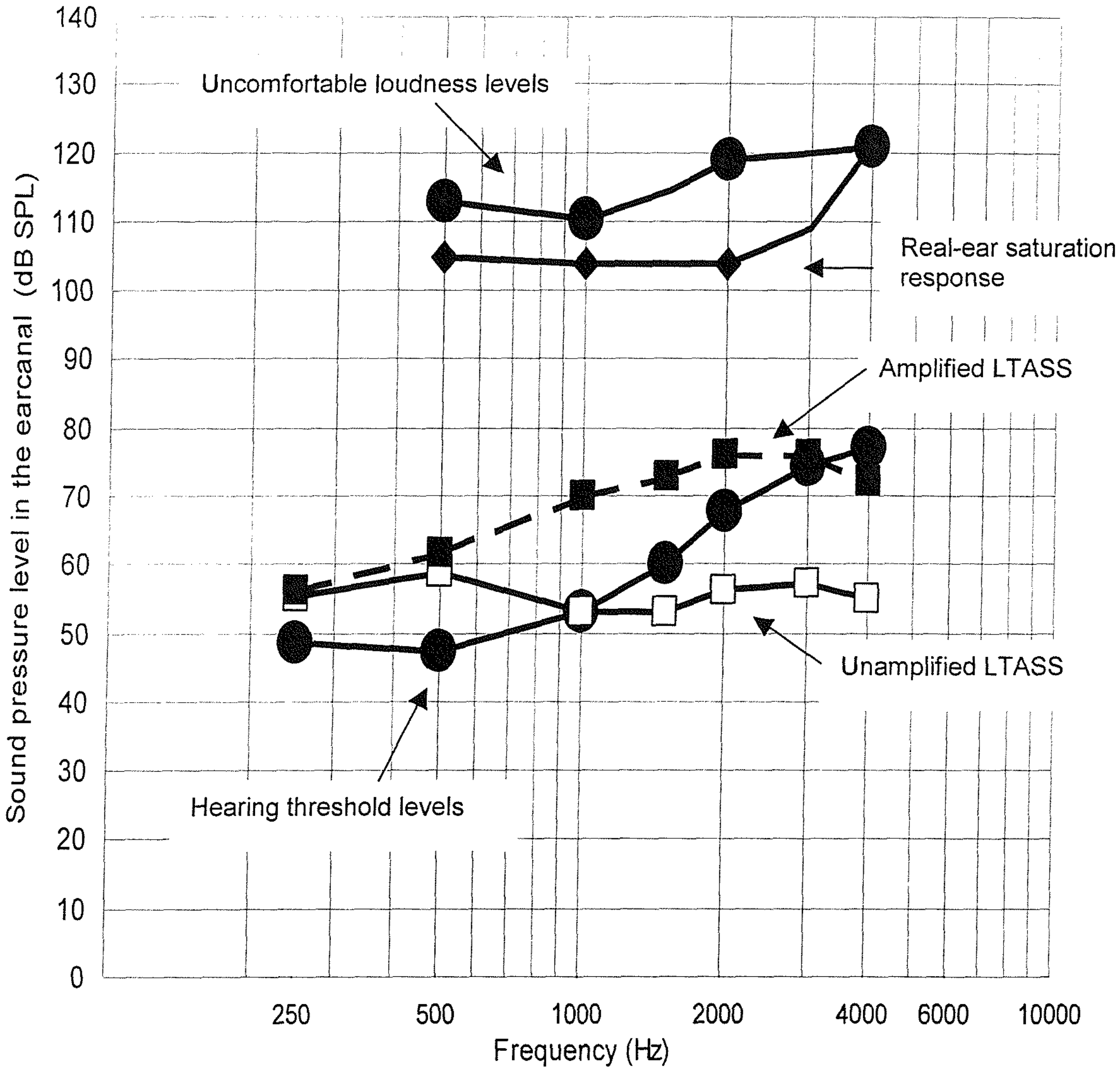


Figure 5.2. The mean sound pressure level of hearing threshold (lower filled circles) and uncomfortable loudness level (upper filled circles) plotted as a function of frequency along with long term average spectrum of speech before (open squares) and after (filled squares) amplification and the real ear saturation response (filled diamonds) of the hearing instrument. All measurements are plotted in SPL in the ear canal.

pure tone input of 90 dB SPL. All variables were converted to ear canal SPL using average acoustic transformations. The audiometric data were converted from dB HL to ear canal dB SPL in a two stage process; first, the data were converted to 6-cc coupler SPL by adding the RETSPLs obtained from ISO 389-1 and second, converted from coupler to real-ear SPL by adding the 6-cc to eardrum transfer function from Bentler and Pavlovic [1989]. The 1/3 octave band levels of the unaided LTASS were converted to ear canal SPL by adding the real ear unaided response from Shaw and Vailancourt [1985]. The aided LTASS was obtained by adding the REIG to the unaided LTASS. Finally, the saturation sound pressure level of the hearing instrument was converted to ear canal SPL by adding the RECD transform from Munro and Hatton [2000]. The figure shows that mean LTASS was audible at frequencies up to around 1 kHz before amplification: after amplification it was audible up to 3 kHz.

Since the linear gain control was disabled for the duration of the study, extra care was taken to avoid exceeding the subjects' uncomfortable loudness levels. The RESR was set around 104 dB SPL. The NAL target value for maximum output [based on the average hearing threshold level at 0.5, 1 and 2 kHz] is 102 dB in a 2-cc coupler [Dillon and Storey, 1998]. In the average adult ear canal this is likely to be nearer 107 dB SPL. The RESR measurements were made using a pure tone signal as this indicates the maximum output the hearing instrument is capable of producing when all of the power is concentrated into one narrow frequency region at a time. With a broadband speech-like signal, the total power has to be divided among all the signal frequencies and this means that the power available for any particular frequency region will be less than for the pure tone measurement. For this reason, the maximum output may be considerably lower than shown in Figure 5.2. In order to determine the saturation level, the output of the hearing instrument was measured using a broadband input signal. Figure 5.3 shows the mean input-output function of the hearing instrument using FAAF-shaped noise. The hearing instrument was positioned at the reference test position in the sound field and the sound pressure level was measured in an HA2 2-cc coupler. The curve is linear at input levels below 85 dB SPL. The maximum speech level used in the present study was 69 dB SPL but when presented in a similar noise level, the total becomes 72 dB SPL. This means that there was approximately 13 dB of headroom before the response becomes non-linear at the highest presentation level used in this study. This confirms

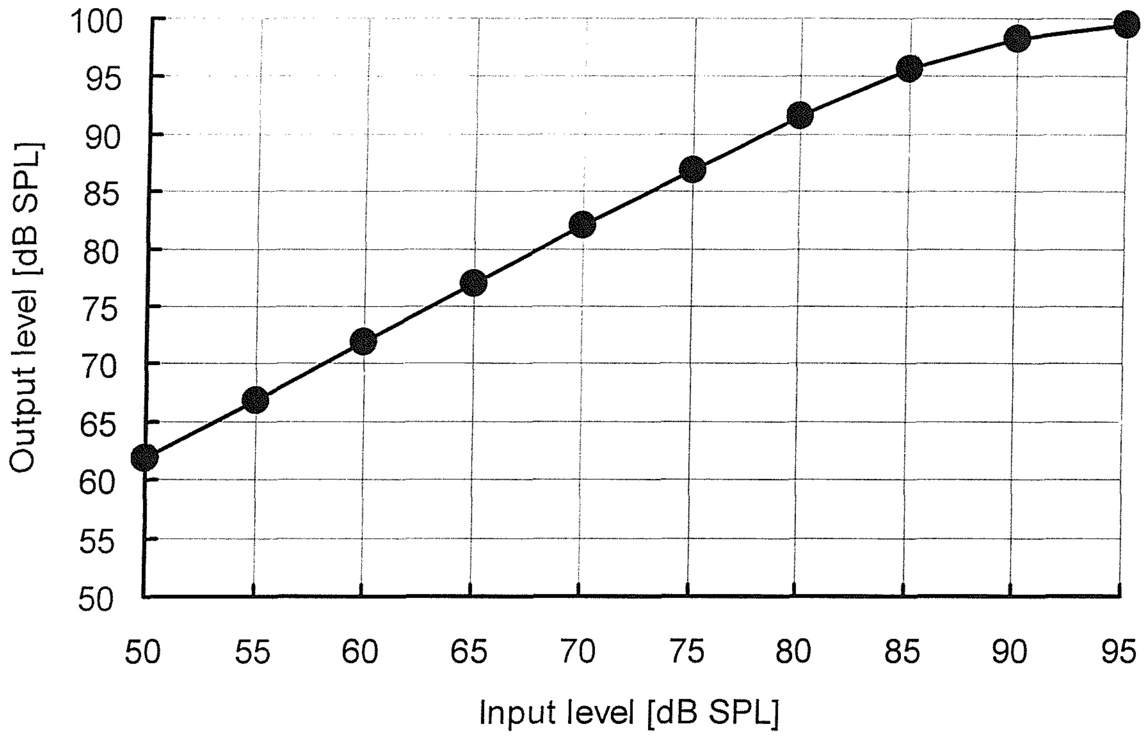


Figure 5.3. The mean input/output function for the hearing instrument settings used in the present study. The input signal was FAAF-shaped noise. The hearing instrument was positioned at the reference test position in the sound field and the sound pressure level was measured in an HA2 2-cc coupler.

that the hearing instruments were not operating in saturation for any of the speech input levels used in the study.

The FAAF test [with visual feedback disabled] was presented at levels that could realistically be expected to occur in typical everyday listening situations. In 1998, Olsen summarised the work of Pearsons *et al.* [1977] who measured the mean speech levels in a variety of settings. The average levels of what was referred to as 'casual', 'normal' and 'raised' speech for a male speaker were reported as 56, 61 and 68 dB SPL respectively. The levels obtained with female speakers were approximately 2 dB lower. They also reported the average level of speech measured at a conversational distance inside and outside the home as 58 dB SPL and 68 dB SPL respectively [no value was given for casual speech]. The SII standard assumes an overall level of 62 dB SPL for normal speech and 68 dB SPL for raised speech [no value is given for quiet speech]. Walden [1997] suggests that hearing instrument clinical trials should be performed with speech at 50, 60 and 70 dB[A]. Recently, a multi-centre trial reported by Larson *et al.* [2000] used a speech presentation level of 52, 62 and 74 dB SPL to represent quiet, average and loud speech. Gatehouse [personal communication] measured the daily noise dose of a number of subjects. The mean level was 62 dB SPL with an inter-quartile range of 55 to 69 dB SPL. Although the measurements were not restricted to speech alone, the values are consistent with the previous studies showing speech within the range of 55-70 dB SPL. Based on the above studies, the overall [rms] levels that were used to approximate quiet, normal and raised speech were 55, 62 and 69 dB SPL respectively.

Filtered steady noise was used to limit the initial performance somewhere below the maximum possible score [i.e., to avoid a ceiling effect] so that it would be possible to measure improvements over time. A pilot study using the not-fitted ear of ten monaurally aided subjects [typically age 70 years], was undertaken to determine if a single SNR could be used for all test conditions. These individuals had participated in the earlier experiments so they were very experienced at performing the FAAF test. Around five hours of testing, spread over two visits, was carried out on each subject. The test conditions were quasi-random and there was typically only time for one list of 80 items per condition [after an initial practice list]. Data were collected over a range of presentation levels and SNRs. For the speech levels used in the present study the mean values for the aided and unaided conditions were in the range 55-75% with a SNR of –3 dB. The lowest unaided score for any individual was 40% and the highest aided score was 80%. This range of scores should be sufficiently high to maintain subject motivation but avoid ceiling effects. Thus, it was decided to test each subject at a fixed SNR across all test conditions. The SNR would be determined for each subject but this was expected to be around –3 dB.

Once the subjects had completed the practice sessions, they were fitted with the hearing instrument and FAAF testing was performed within seven days of fitting. This was different from the earlier experiments when performance was measured on the day of fitting: testing could not commence in the present study until the fixed user-gain had been established. Testing was repeated at 6 and 12 weeks post-fitting. Some previous studies have suggested that subjects react differently on their last visit perhaps because of differences in motivation and attention. In order to avoid this, subjects were informed that testing would also take place on week 13 [although, unknown to the subject, this was usually only post-study audiometry etc].

Testing was completed aided or unaided on each ear before changing condition as this reduced the number of occasions the subject's concentration was disturbed. It also reduced the number of occasions the subject's head and hearing instrument was moved. The order of testing was balanced across subjects and test sessions.

Table 5.4. Self-reported daily use of the hearing instrument [hours] as a function of post-fitting time. Each subject was issued with a diary that required them to indicate the hours of daily use of the hearing instrument. The categories available to the subject were: less than 4 hours, 4-8 hours, 8-12 hours or greater than 12 hours. The daily hours of use for each subject was taken as the most frequently ticked category in each week [n=16].

Post-fitting time [weeks]	1	2	3	4	5	6	7	8	9	10	11	12
< 4 hours	0	0	0	0	0	0	0	0	0	0	0	0
4-8 hours	0	0	2	2	3	3	0	0	0	1	1	0
8-12 hours	11	10	7	7	6	6	9	7	9	7	8	8
> 12 hours	5	6	7	7	7	7	7	9	7	8	7	8
Missing	0	0	0	0	0	0	0	0	0	0	0	0

5.2 Results

The mean self-reported daily use of the hearing instrument as a function of post-fitting time is shown in Table 5.4. The number of subjects who reported using their hearing instrument more than 8 hours per day was never less than 13 [82%] during any one-week period.

Table 5.5. Mean Speech Intelligibility Index for the fitted ear. The values in brackets incorporate the desensitisation correction from Pavlovic *et al.* [1986]. The equivalent speech spectrum level of the FAAF test was based on an overall SPL of 55, 62 and 69 dB. The equivalent noise spectrum of the FAAF shaped noise spectrum was determined for each subject [based on the SNR]. The aided condition was obtained by adding the user insertion gain to the speech and noise levels. The band importance function for nonsense syllables was used in the calculation.

	0.25	0.5	1	2	4	8	Overall
			Unaided				
55 dB SPL	0.02[0.01]	0.05[0.04]	0.01[0.00]	0.00[0.00]	0.00[0.00]	0.00[0.00]	0.08[0.06]
62 dB SPL	0.02[0.01]	0.05[0.04]	0.05[0.03]	0.00[0.00]	0.00[0.00]	0.00[0.0]	0.12[0.09]
69 dB SPL	0.02[0.01]	0.05[0.04]	0.08[0.05]	0.03[0.01]	0.02[0.01]	0.00[0.00]	0.20[0.13]
			Aided				
55 dB SPL	0.02[0.01]	0.06[0.04]	0.09[0.05]	0.09[0.05]	0.00[0.00]	0.00[0.00]	0.25[0.15]
62 dB SPL	0.02[0.01]	0.05[0.04]	0.08[0.05]	0.14[0.07]	0.10[0.04]	0.00[0.0]	0.39[0.21]
69 dB SPL	0.02[0.01]	0.05[0.04]	0.09[0.05]	0.13[0.06]	0.13[0.05]	0.00[0.0]	0.42[0.22]
			Change [aided – unaided]				
55 dB SPL	0.00[0.00]	0.01[0.00]	0.08[0.05]	0.09[0.05]	0.00[0.00]	0.00[0.00]	0.17[0.09]
62 dB SPL	0.00[0.00]	0.00[0.00]	0.03[0.02]	0.14[0.07]	0.10[0.04]	0.00[0.00]	0.27[0.12]
69 dB SPL	0.00[0.00]	0.00[0.00]	0.01[0.00]	0.10[0.05]	0.11[0.04]	0.00[0.00]	0.22[0.09]

The mean SII values for the aided and unaided condition are shown in Table 5.5. These incorporate the desensitisation correction from Pavlovic *et al.* [1986]. The SII was used to provide a guide to changes in audibility for each presentation level before and after amplification. The equivalent speech spectrum level was based on the average of the scoring FAAF words at 55, 62 and 69 dB SPL. The equivalent spectrum of the FAAF-shaped noise was determined for each individual subject [based on the SNR used during testing]. The aided condition was obtained by adding the insertion gain values [at the user-gain setting] to the speech and noise levels. The band importance function for nonsense syllables was used in the calculation. The mean unaided SII values for speech presentation levels of 55, 62 and 69 dB SPL were 0.08, 0.12 and 0.20 respectively. The corresponding values when aided were 0.25, 0.39 and 0.42. Therefore, the improvements in SII after aiding were 0.17, 0.27 and 0.22 respectively. The improvement was due primarily to the improved audibility in the 2-4 kHz region. The lower aided SII value at 55 dB SPL indicates that part of the speech envelope remains inaudible even after amplification. The smaller improvement after aiding at a presentation level of 69 dB SPL indicates that a larger proportion of the speech signal was already audible before aiding. The SII values confirm that performance should improve as the presentation level increases, especially for the unaided conditions.

5.2.1 Change in benefit score over time

Figure 5.4 shows mean FAAF benefit [aided minus unaided] as a function of post-fitting time in the fitted ear for speech presented at 55, 62 and 69 dB SPL [circles, squares and triangles respectively]. The results for the not-fitted control ear are shown in Figure 5.5. The numerical data for both ears are given in Table 5.6. The mean benefit scores in the fitted ear at the time of fitting were 19, 8 and 0% with speech presented at 55, 62 and 69 dB SPL respectively. Similar results were obtained in the not-fitted control ear.

Table 5.6. Mean benefit score [in percent] as a function of post-fitting time. One standard deviation is given in brackets [n=16].

Post-fitting time [weeks]	0	6	12	Mean
Fitted ear				
55 dB SPL	19.0 [8.5]	20.6 [11.0]	21.1 [10.4]	20.2 [9.9]
62 dB SPL	8.0 [7.8]	9.3 [9.8]	11.2 [9.3]	9.5 [8.9]
69 dB SPL	0.0 [7.9]	3.5 [6.2]	5.9 [8.3]	3.1 [7.8]
Mean	9.0 [11.2]	11.1 [11.5]	12.7 [11.1]	10.9 [11.3]
Control ear				
55 dB SPL	17.2 [11.2]	16.8 [11.9]	15.0 [13.0]	16.3 [11.9]
62 dB SPL	7.7 [9.2]	7.3 [10.5]	8.4 [10.1]	7.8 [9.7]
69 dB SPL	1.3 [6.5]	0.7 [11.4]	0.0 [8.2]	0.7 [8.8]
Mean	8.7 [11.1]	8.3 [12.9]	7.8 [12.1]	8.3 [12.0]

The change in benefit, relative to the time of fitting, is seen more easily in Figures 5.6 and 5.7 [numerical data in Table 5.7]. The scores increase over time for all presentation levels in the fitted ear and the effect increases as the presentation level increases. There are increases of 2.0, 3.2 and 5.9% at presentation levels of 55, 62 and 69 dB SPL respectively. The corresponding increase in the not-fitted control ear was always less than 1%.

A summary of the repeated-measures analysis of variance [ANOVA] is shown in Table XX, Appendix E. The three factors that were treated as repeated-measures were

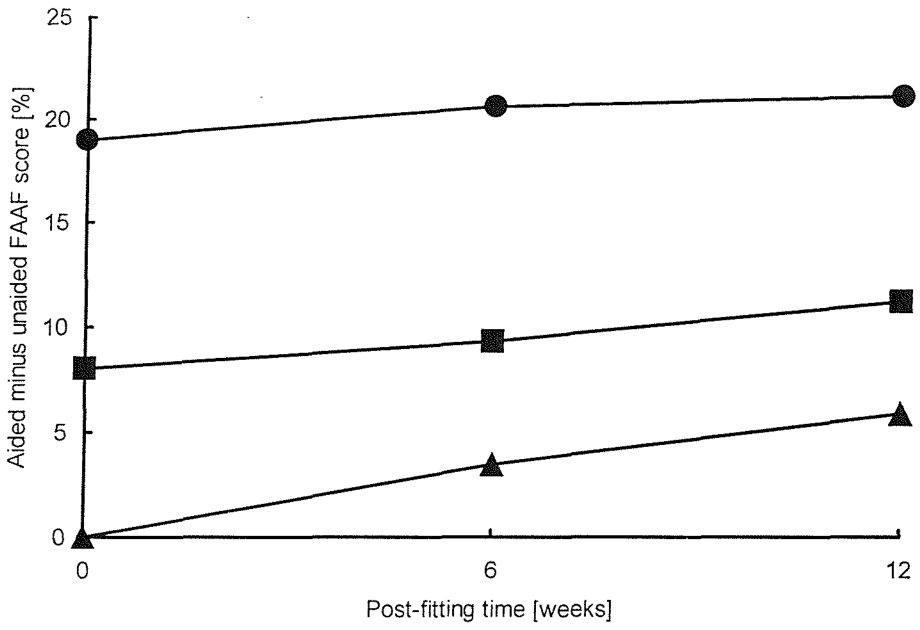


Figure 5.4. Mean FAAF benefit [aided minus unaided] in the fitted ear as a function of post-fitting time with speech at 55 [circles], 62 [squares] and 69 dB SPL [triangles].

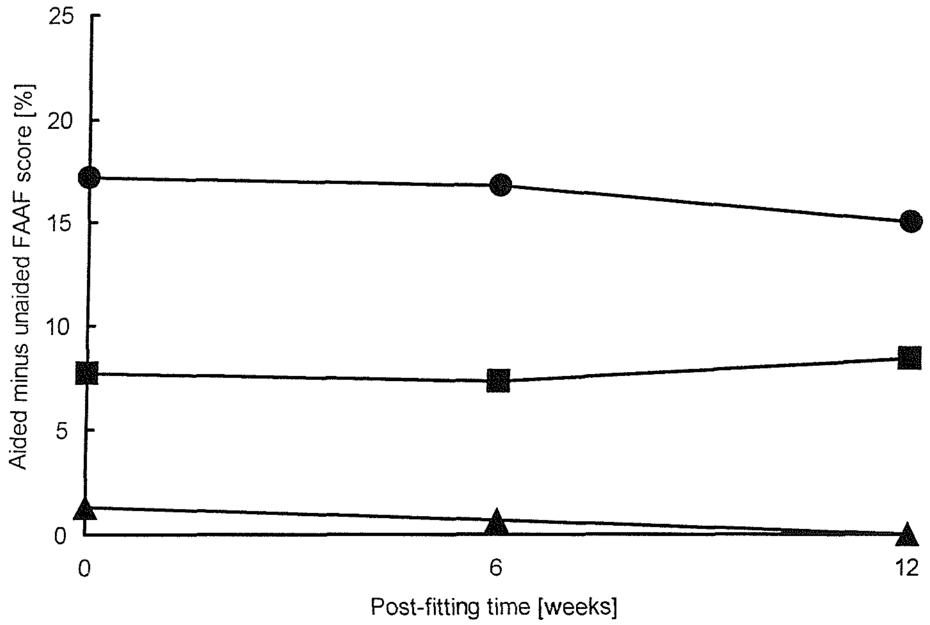


Figure 5.5. Mean FAAF benefit [aided minus unaided] in the control ear as a function of post-fitting time with speech at 55 [circles], 62 [squares] and 69 dB SPL [triangles].

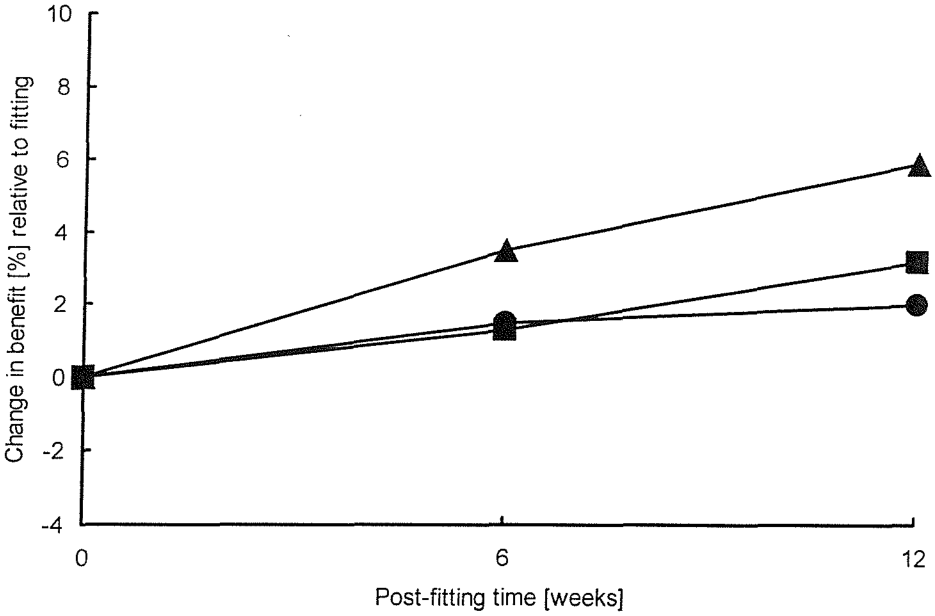


Figure 5.6. Mean change in benefit in the fitted ear relative to the time of fitting at a speech presentation level of 55 [circles], 62 [squares] and 69 dB SPL [triangles].

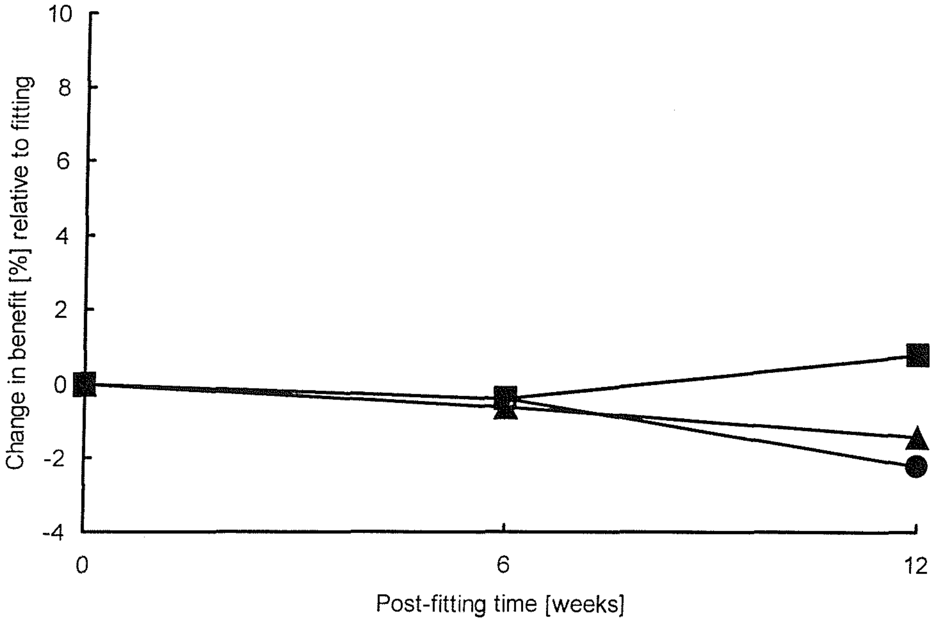


Figure 5.7. Mean change in benefit in the control ear relative to the time of fitting at a speech presentation level of 55 [circles], 62 [squares] and 69 dB SPL [triangles].

ear [test and control], presentation level [55, 62 and 69 dB SPL] and post-fitting time [0, 6 and 12 weeks]. As expected, there was a statistically significant difference with presentation level. The orthogonal polynomial breakdown revealed that only the linear component was significant [F(1,15)=36.7; p<0.01]. The effects of ear and post-fitting time were not statistically significant; however, there was a statistically significant interaction between these two factors. Again, it was only the linear component of the polynomial breakdown that was significant [F(1,15)=16.2; p<0.01]. This interaction indicates that the ears were differentially affected by time; the fitted ear increases while the control ear shows little change. This was confirmed by investigating each ear separately [see Table XXI, Appendix E]; time was statistically significant in the fitted ear but not the control ear.

Table 5.7. Mean change in benefit score [in percent] relative to the time of fitting. One standard deviation is given in brackets [n=16].

Post-fitting time [weeks]	6	12	Mean
Fitted ear			
55 dB SPL	1.5 [7.2]	2.0 [8.9]	1.8 [8.0]
62 dB SPL	1.3 [5.7]	3.2 [6.3]	2.3 [6.0]
69 dB SPL	3.5 [8.0]	5.9 [6.4]	4.7 [7.2]
Mean	2.1 [6.9]	3.7 [7.4]	2.9 [7.2]
Control ear			
55 dB SPL	−0.4 [3.5]	−2.2 [6.3]	−1.3 [5.1]
62 dB SPL	−0.4 [5.8]	0.8 [7.1]	0.2 [8.0]
69 dB SPL	−0.6 [9.6]	−1.4 [6.4]	−1.0 [8.0]
Mean	−0.5 [6.0]	−1.0 [6.6]	−0.7 [6.6]

An increase in benefit could be due to either an increase in aided performance or a decrease in unaided performance [or some combination of the two]. To investigate this further, the aided and unaided scores were examined separately and the results are reported in the following sections.

5.2.2 Aided performance over time

Figure 5.8 shows the mean aided scores, as a function of post-fitting time in the fitted ear at a speech presentation level of 55, 62 and 69 dB SPL [circles, squares and

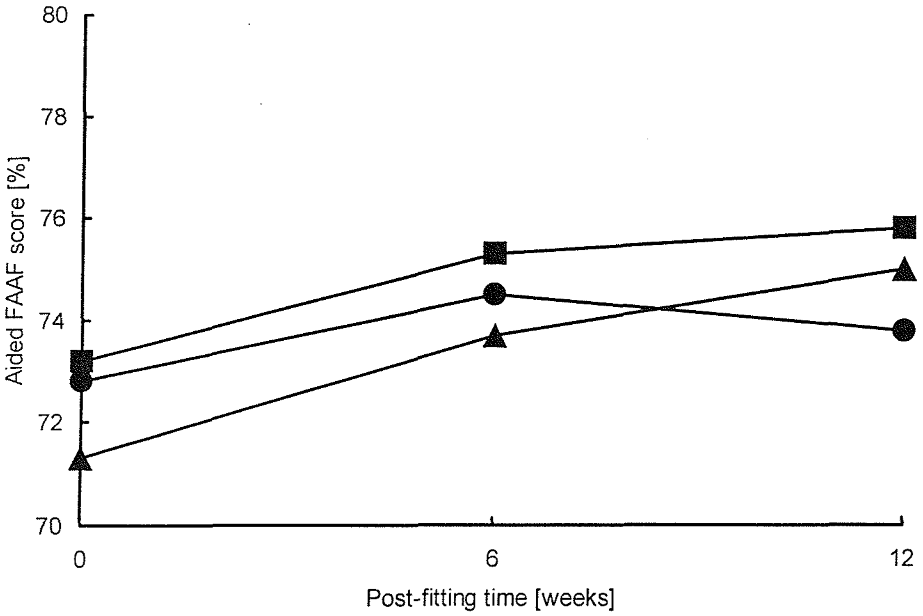


Figure 5.8. **Mean aided FAAF score [%] in the fitted ear as a function of post-fitting time with speech at 55 [circles], 62 [squares] and 69 dB SPL [triangles].**

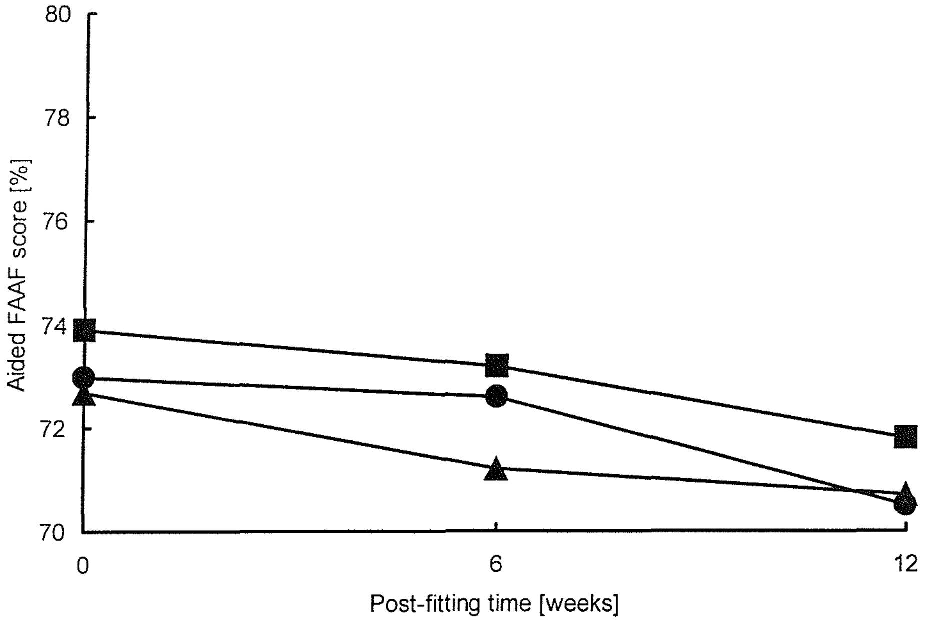


Figure 5.9. **Mean aided FAAF score [%] in the control ear as a function of post-fitting time with speech at 55 [circles], 62 [squares] and 69 dB SPL [triangles].**

triangles respectively]. The results from the control ear are shown in Figure 5.9. The numerical data are given in Table 5.8. The initial scores are around 72% in both ears although, interestingly, the lowest score at the time of fitting was obtained with the highest speech [and noise] presentation level. The change in score, relative to the time of fitting, is more easily seen in Figures 5.10 and 5.11 [numerical data in Table 5.9]. The aided scores increased over time in the fitted ear for all presentation levels although the effect increases with presentation level. There is an increase of 1.0, 2.6 and 3.6% at a presentation level of 55, 62 and 69 dB SPL respectively. The corresponding change in the not-fitted control ear was always negative and typically around –2.0%.

Table 5.8 Mean aided FAAF score [in percent] as a function of post-fitting time. One standard deviation is given in brackets [n=16].

Post-fitting time [weeks]	0	6	12	Mean
	Fitted ear			
55 dB SPL	72.8 [4.5]	74.5 [5.1]	73.8 [5.0]	73.7 [4.9]
62 dB SPL	73.2 [3.8]	75.3 [5.3]	75.8 [4.9]	74.8 [4.7]
69 dB SPL	71.3 [4.9]	73.7 [4.0]	75.0 [3.7]	73.3 [4.4]
Mean	72.5 [4.4]	74.5 [4.8]	74.9 [4.6]	73.9 [4.9]
	Control ear			
55 dB SPL	73.0 [6.2]	72.6 [4.8]	70.5 [5.6]	72.0 [5.6]
62 dB SPL	73.9 [6.0]	73.2 [6.0]	71.8 [6.2]	73.0 [6.0]
69 dB SPL	72.7 [5.8]	71.2 [7.7]	70.7 [7.2]	71.5 [6.8]
Mean	73.2 [5.9]	72.3 [6.2]	71.0 [6.2]	72.2 [6.1]

Table 5.9. Mean change in aided FAAF score [in percent] relative to the time of fitting. One standard deviation is given in brackets [n=16].

Post-fitting time [weeks]	6	12	Mean
	Fitted ear		
55 dB SPL	1.7 [4.2]	1.0 [3.9]	1.3 [4.1]
62 dB SPL	2.1 [4.1]	2.6 [3.9]	2.3 [3.9]
69 dB SPL	2.3 [4.9]	3.6 [4.1]	3.0 [4.5]
Mean	2.0 [4.3]	2.4 [4.0]	2.2 [4.2]
	Control ear		
55 dB SPL	-0.4 [3.5]	-2.4 [3.3]	-1.4 [3.5]
62 dB SPL	-0.7 [3.6]	-2.2 [4.0]	-1.4 [3.8]
69 dB SPL	-1.5 [5.1]	-2.0 [4.9]	-1.7 [4.9]
Mean	-0.9 [4.1]	-2.2 [4.0]	-1.5 [4.1]

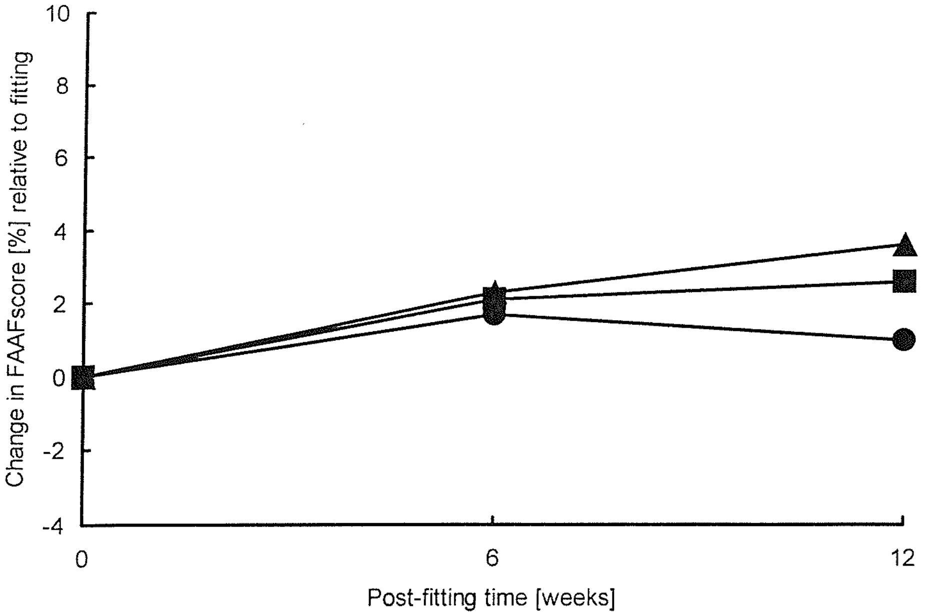


Figure 5.10. Mean change in aided FAAF score [%] in the fitted ear, relative to fitting at a speech presentation level of 55 [circles], 62 [squares] and 69 dB SPL [triangles].

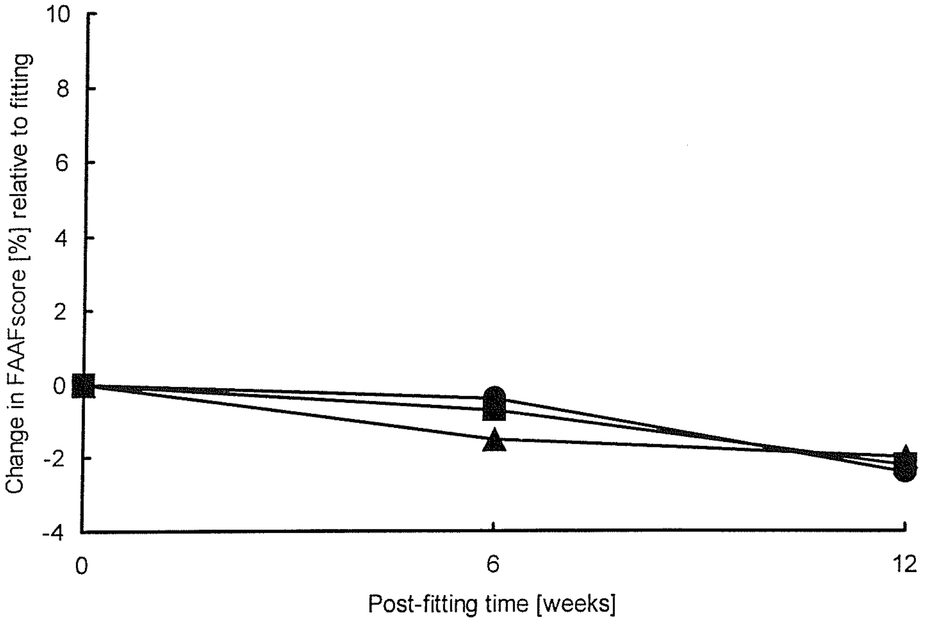


Figure 5.11. Mean change in aided FAAF score [%] in the control ear relative to the time of fitting at a speech presentation level of 55 [circles], 62 [squares] and 69 dB SPL [triangles].

A summary of the repeated-measures ANOVA is shown in Table XXII, Appendix E. The three factors that were treated as repeated-measures were ear [test and control], presentation level [55, 62 and 69 dB SPL] and post-fitting time [0, 6 and 12 weeks]. The effect of ear was statistically significant; the mean score in the fitted ear was 73.9% compared with 72.2% in the control ear. Only the linear component of the polynomial breakdown was significant [F(1,15)=6.0; p=0.03]. In addition, there was a statistically significant interaction between ear and post-fitting time. This interaction indicates that the ears were differentially affected by time; the difference between ears was only 0.7% at the start of the study but had reached 3.9% by the end of the study. Again, only the linear component of the polynomial breakdown was significant [F(1,15)=27.0; p<0.01]. An ANOVA was performed on each ear separately [see Table XXIII, Appendix E]. The factor of post-fitting time was statistically significant for both ears but there was no interaction with presentation level. Only the linear component of the polynomial breakdown was significant [fitted ear F(1,15)=13.4,p<0.01; control ear F(1,15)=7.4, p=0.02].

5.2.3 Unaided performance over time

Figure 5.12 shows the mean unaided score, as a function of post-fitting time, in the fitted ear for speech presented at 55, 62 and 69 dB SPL [circles, squares and triangles respectively]. The results from the control ear are shown in Figure 5.13. The numerical data are given in Table 5.10. The initial performance was related to presentation level; it was typically around 55% with speech at 55 dB SPL but nearer 71% with speech at 69 dB SPL. The change in score relative to the time of fitting is more easily seen in Figures 5.14 and 5.15 [numerical data in Table 5.11]. There is a trend towards the scores decreasing over time for all presentation levels and in both ears.

A summary of the repeated-measures ANOVA is shown in Table XXIV, Appendix E. The three factors that were treated as repeated-measures were ear [test and control], SPL [55, 62 and 69 dB SPL] and post-fitting time [0, 6 and 12 weeks]. The only statistically significant effect was the difference due to presentation level. Importantly, there was no statistically significant change over time for either ear.

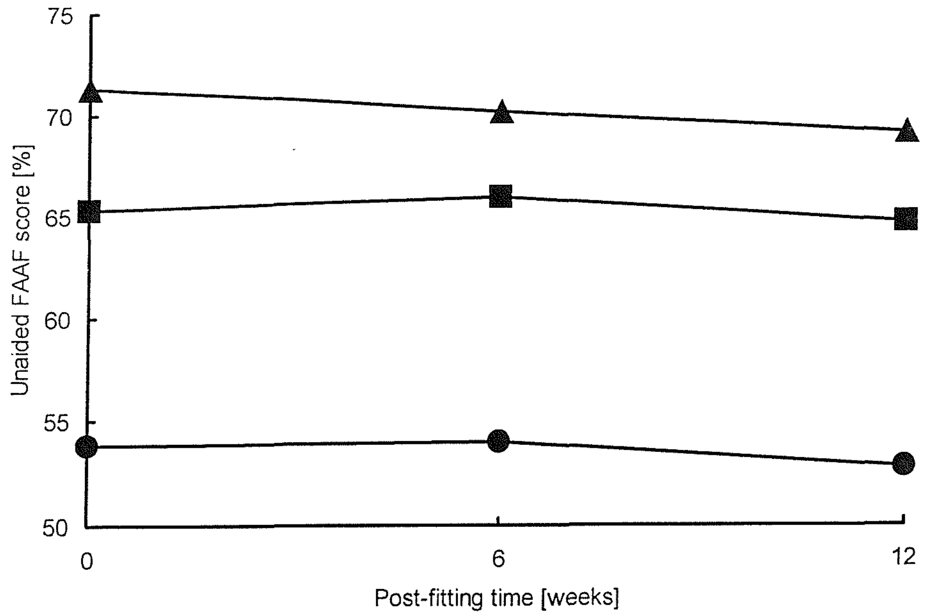


Figure 5.12. Mean unaided FAAF score [%] in the fitted ear as a function of post-fitting time at a speech presentation level of 55 [circles], 62 [squares] and 69 dB SPL [triangles].

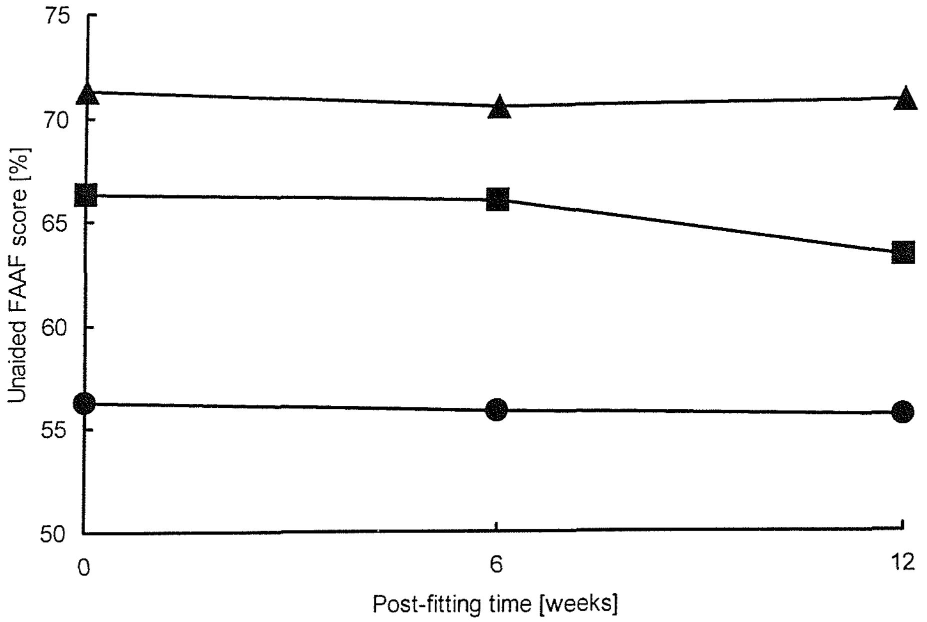


Figure 5.13. Mean unaided FAAF score [%] in the control ear as a function of post-fitting time at a speech presentation level of 55 [circles], 62 [squares] and 69 dB SPL [triangles].

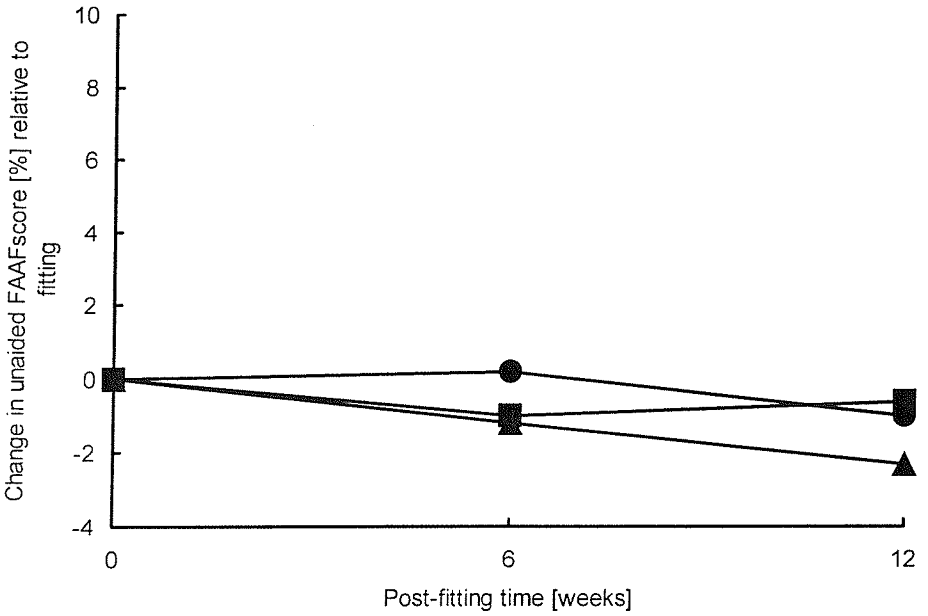


Figure 5.14. Mean change in unaided FAAF score [%] in the fitted ear, relative to fitting at a speech presentation level of 55 [circles], 62 [squares] and 69 dB SPL [triangles].

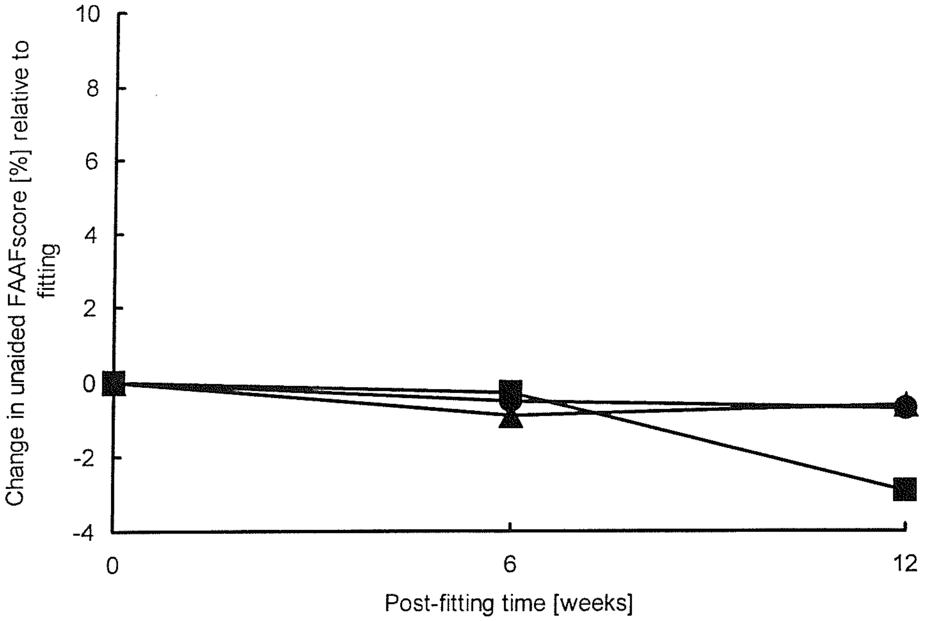


Figure 5.15. Mean change in unaided FAAF score [%] in the control ear, relative to fitting at week 0 at a speech presentation level of 55 [circles], 62 [squares] and 69 dB SPL [triangles].

Table 5.10. Mean unaided FAAF score [in percent] as a function of post-fitting time. One standard deviation is given in brackets [n=16].

Post-fitting time [weeks]	0	6	12	Mean
	Fitted ear			
55 dB SPL	53.8 [7.9]	54.0 [11.7]	52.8 [9.4]	53.5 [9.6]
62 dB SPL	65.3 [6.2]	66.0 [7.6]	64.7 [7.5]	65.3 [7.0]
69 dB SPL	71.3 [4.6]	70.2 [4.7]	69.1 [6.3]	70.2 [5.2]
Mean	63.5 [9.6]	63.4 [10.8]	62.2 [10.3]	63.0 [10.2]
	Control ear			
55 dB SPL	56.3 [10.3]	55.8 [10.1]	55.6 [11.6]	55.9 [10.4]
62 dB SPL	66.3 [6.7]	66.0 [7.5]	63.3 [7.5]	65.2 [7.2]
69 dB SPL	71.3 [5.4]	70.5 [7.7]	70.8 [5.7]	70.9 [6.2]
Mean	64.6 [9.9]	64.1 [10.4]	63.2 [10.5]	64.0 [10.2]

Table 5.11. Mean change in unaided FAAF score [in percent] relative to the time of fitting. One standard deviation is given in brackets [n=16].

Post-fitting time [weeks]	6	12	Mean
	Fitted ear		
55 dB SPL	0.2 [7.9]	−1.0 [7.4]	−0.4 [7.5]
62 dB SPL	−1.0 [8.0]	−0.6 [4.5]	−0.8 [6.4]
69 dB SPL	−1.2 [4.9]	−2.3 [4.8]	−1.7 [4.8]
Mean	−0.7 [6.9]	−1.3 [5.6]	−1.0 [6.3]
	Control ear		
55 dB SPL	−0.5 [4.7]	−0.7 [6.1]	−0.6 [5.6]
62 dB SPL	−0.3 [4.4]	−2.9 [7.1]	−1.6 [5.9]
69 dB SPL	−0.9 [5.8]	−0.6 [4.0]	−0.7 [4.9]
Mean	−0.5 [4.9]	−1.4 [5.9]	−1.0 [5.4]

5.2.4 Summary of statistical analysis

The results of the statistical analysis on the mean FAAF test scores can be summarised as follows:

Performance at the time of fitting

- There was a statistically significant decrease in mean benefit score as the speech presentation level increased [19, 8 and 0% at 55, 62 and 69 dB SPL respectively].

- This decrease in benefit was due to a statistically significant increase in the mean unaided score as the speech presentation level increased [53.8, 65.3 and 71.3% at 55, 62 and 69 dB SPL respectively] with no statistically significant change in the mean aided scores.

Change in performance over time

- There was a statistically significant increase in mean benefit over time in the fitted ear [+2.9%] but not the control ear [–0.7%].
- In the fitted ear, there was a trend towards a larger increase in mean benefit as presentation level increased [1.8, 2.3 and 4.7% at 55, 62 and 69 dB SPL respectively] but this was not statistically significant.
- The increase in benefit over time in the fitted ear was primarily due to a statistically significant increase in mean aided score [+2.2%] but there was also a trend of performance decreasing in the unaided condition [–1.0%].
- In the fitted ear, there was a trend towards a larger increase in mean aided score as presentation level increased [1.3, 2.3 and 3.0% at 55, 62 and 69 dB SPL respectively] and a larger decrease in the mean unaided scores [–0.4, –0.8 and –1.7% at 55, 62 and 69 dB SPL respectively] but these were not statistically significant.
- In the control ear, there was a reduction in the mean aided and unaided scores [–1.5% and –1.0% respectively] but only the aided condition was statistically significant.

5.3 Discussion

Experiment three examined changes in FAAF scores, as a function of presentation level, over a post-fitting time period of 12 weeks. Sixteen subjects were fitted monaurally at a fixed-gain setting for the duration of the study. The not-fitted ear was used as the control condition.

5.3.1 Performance at the time of fitting

The magnitude of the derived benefit score, at the time of fitting, decreased as the presentation level increased. This was expected since performance in the unaided condition is mediated by audibility whereas performance in the aided condition is limited by the competing FAAF-shaped noise. As a result, the unaided scores increase with presentation level whereas the aided scores are similar across levels; this has the effect of reducing the benefit score.

Reference data for the FAAF test are available from Foster and Haggard [1987]. These are plotted on Figure 5.16 along with the data from the present study. The filled circles correspond to the mean data from the 16 hearing impaired subjects in the present study while the open circles correspond to the mean data for 32 normal hearing subjects [64 ears] from Foster and Haggard. The Foster and Haggard data have been shifted to the left by 2 dB to account for differences in the calibration of the noise between the two studies [Foster, personal communication]. In order to allow a direct comparison between the studies, the mean performance in the present study has been plotted as a function of SNR. This involved converting an SII value to a SNR. Essentially, the peaks and troughs of conversational speech are assumed to cover a range of +12 to −18 dB. At a SNR of −12 dB, the SII for conversational speech is zero since the +12 dB peaks of speech are inaudible. The SII then increases by 0.1 for every +3 dB improvement in SNR until it reaches a maximum value of 1.0 at a SNR +18 dB [when even the −18 dB troughs of speech will be audible to a normal hearing listener]. The same approach has been taken by others. For example, Killion [1985] has used this approach for the performance functions shown in Figure 15 of ANSI S3.5 [1969]. The aided performance scores at a presentation level of 62 and 69 dB SPL are almost superimposed since the SII values are similar for these conditions.

More recent data from Shields and Campbell [2001] have also been plotted on Figure 5.16. The open triangles correspond to nine normal hearing individuals while the filled triangles correspond to 15 hearing impaired subjects. The data from the hearing impaired subjects have been shifted 5 dB to the left to account for the typical difference in SNR caused by the sensorineural hearing impairment [Moore, 1998]. There is some similarity across all of the studies although unaided performance in the present study is better than for all of the published studies. The most striking feature is the rather abrupt flattening of the function for the aided conditions. The data from the published studies indicate a slope of approximately 6% per dB in the 40 to 85% region

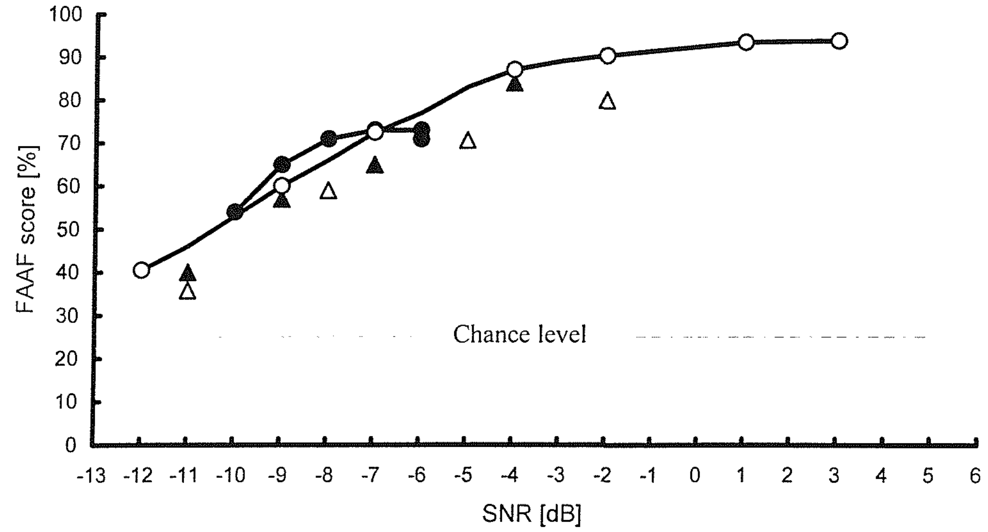


Figure 5.16. The mean FAAF score, in percent, as a function of SNR. The filled circles show the mean scores obtained from the present study, at the time of fitting [see text for the explanation of conversion to SNR]. The open circles show the reference data for normal hearing subjects from Foster and Haggard [1987]. This has been shifted to the left by 2 dB to account for differences in calibration of the noise level. The triangles show the reference data from Shields and Campbell [2001]. The open triangles are from normal subjects and the filled triangles from hearing impaired subjects.

of the average performance function of normal hearing subjects while the slope of the performance function for hearing impaired subjects is nearer 3.5% per dB. Using these slopes, it may have been anticipated that the performance in the present study would be some 7 to 12% higher at the highest aided presentation levels. This point is discussed more fully in section 5.3.3.

The relationship between all of these studies should be interpreted cautiously since they are influenced by a number of factors such as differences in calibration procedures. Also, since the data from the present study were converted to SNR from SII, the performance function will be influenced by the SII values. The SII calculations were based on band importance functions for nonsense syllables [since importance functions are not available for the FAAF test material]. The proficiency factor [i.e., a measure of individual differences that are not explained by audibility such as experience or familiarity with the FAAF test] was assumed to be one [i.e., the equivalent of a skilled normal hearing subject]. The proficiency factor, in particular, may be important because the subjects were highly experienced at performing the FAAF test: performance during pre-fitting practice sessions was clearly seen to improve in all subjects. Although this improvement was not recorded in a systematic manner, it was normal for criterion performance of 71% to be obtained with a poorer SNR [often by around 2-3 dB] after approximately 12 hours of practice. This latter point may explain why the performance curve is shifted to the left of the Shields and Campbell data. We also applied an average desensitisation correction from Pavlovic *et al.* [1986] to our SII values to account for the decreased frequency resolution and time resolving ability associated with a sensorineural hearing impairment. This involved multiplying the uncorrected SII in each octave band by a hearing threshold level dependent factor.

The mean performance at 55 dB SPL was similar to those obtained in experiments one and two. This was to be expected since the subjects had similar hearing thresholds and were tested in the same laboratory with the same equipment. There was a 1-2% decrease in mean aided performance with speech presented at 69 dB SPL compared to the two lower presentation levels. The reduction is small and not statistically significant. Some studies indicate that performance remains constant at high speech presentation levels while others indicate that performance begins to decrease. An example of the latter is the multi-centre trial by Larson *et al.* [2000]. In their study, subjects were tested aided and unaided at 52, 62 and 74 dB SPL. The signal-to-babble ratio [SBR] was fixed across presentation level at a mean nominal value of 0 dB [the actual SBR was probably nearer +7 dB]. Performance decreased from around 70% at 52 and 62 dB SPL to nearer 55% at 74 dB SPL. This fall in performance is more marked than the present study but the presentation level was 74 dB SPL compared to 69 dB SPL. There were other differences between the two studies that may be relevant including differences in the characteristics of the masking noise [babble versus speech-shaped noise]. Studebaker *et al.* [1999] have shown that performance deteriorates at high speech and noise levels in normal subjects [and hearing impaired subjects when aided]. They measured performance in a group of 12 normally hearing subjects with high levels of speech and noise and showed that performance was around 14% lower than expected based on SII values. This 'roll-over' is presumably related to upward spread of masking which results in vowels masking consonants. The results from Studebaker *et al.* suggest that this is not simply due to an increase in the level of the speech signal [i.e., the speech 'level distortion factor' incorporated into the SII]; instead, there is an interaction between speech and noise: the negative effect of adding noise is greater when the speech level is high.

The mean benefit scores [and the aided and unaided scores] at the time of fitting are comparable in both ears. This is consistent with the symmetrical findings on pure tone audiometry. Although no comprehensive psychoacoustic measurements were undertaken on the subjects [for example, there were no tests of frequency and temporal resolution], there is no reason to suppose that there was any systematic asymmetry of auditory function that might have affected the benefit resulting from the provision of amplification.

5.3.2 Change in benefit over time

Following 12 weeks of hearing instrument use, there was a systematic increase in benefit of around 4% in the fitted ear. There was no change in the control ear. This means that the increase in benefit cannot be explained in terms of simple practice effects. Thus, the results of this study, showing a significant increase in benefit over time, replicate the findings of Gatehouse [1992], Cox and Alexander [1992], Cox *et al.*

[1996] and Horwitz and Turner [1997]. The only previous study that has used the not-fitted ear as the control condition was Gatehouse [1992]. This also showed no increase in the control condition. The lack of increase in performance in the control ear means that acclimatisation does not transfer to the not-fitted ear. There are other studies that have demonstrated that changes over time do not transfer to the not-fitted ear. These include the intensity discrimination studies of Robinson and Gatehouse [1995a,b]. This is consistent with the very specific nature of perceptual learning discussed in section 2.2. For example, Recanzone *et al.* [1993] reported that owl monkeys showed an improvement in frequency discrimination with training but this was specific to the frequency of the stimulus used for training. Another example is the study of Karni and Sagi [1991] who showed that improvements in the discrimination of visual texture in one eye did not transfer to the other eye. Thus, the results are consistent with a process of perceptual learning.

The magnitude of the acclimatisation effect observed in the present study is consistent with most of the previous studies. The one exception is the study by Gatehouse [1992] that reported a 14% increase in benefit at 12 weeks post-fitting. Gatehouse [1992] calibrated FAAF key words [in peak SPL] and speech-shaped noise [using the 'A' frequency weighting] in a 6-cc coupler. This was different from the calibration method used in the present study so it is difficult to make direct comparisons between presentation levels. However, he reported a mean unaided score of around 71% and a benefit score of around 4% at the time of fitting. A similar level of performance was obtained in the present study with a presentation level of 69 dB SPL[2]. At this presentation level, the change in mean benefit in the present study was 5.9% compared to the 14% reported by Gatehouse. A possible explanation for this difference is the amount of REIG provided by the hearing instrument; Gatehouse provided approximately 12 dB more gain. The larger difference in SPL between aided and unaided speech in the Gatehouse study may mean that more learning was required by the subjects compared to the present study.

---

[2] Although mean unaided performance was similar, Gatehouse reported slightly better performance when aided. This suggests that Gatehouse used a slightly lower presentation level with a more favourable SNR. This would give the same magnitude of audibility unaided but the aided score would be higher because of the lower noise level.

There is no indication that the increase in performance was reaching an asymptote when the study terminated at 12 weeks post-fitting. Benefit was still increasing in the Gatehouse study when it terminated at 12 weeks post-fitting and in the Horwitz and Turner study when it terminated at 18 weeks post-fitting. It is possible that further increases in performance would have been measured in the present study if a longer post-fitting time period had been used.

5.3.3 Changes in aided and unaided scores over time

The change in benefit score in the fitted ear was due to an increase in aided performance and a decrease in unaided performance, although only the former reached statistical significance. Mean aided scores increased by 1.0, 2.6 and 3.6% at a presentation level of 55, 62 and 69 dB SPL respectively. In experiment two, the mean aided score increased by around 1.5 dB at a presentation level of 60 dB SPL [see footnote two].

The unaided scores showed a mean reduction of around 1%. There are mixed findings in the acclimatisation literature regarding changes in the unaided condition. The studies of Cox *et al.* [1996] and Horwitz and Turner [1997] do not show any reduction in the unaided scores over the 12-18 week post-fitting period [although, a subgroup of subjects from the Cox *et al.* study were reviewed nine months post-fitting and the results showed a dramatic reduction in unaided performance]. The clearest demonstration of a reduction in performance in the unaided condition, of approximately 8%, comes from Gatehouse [1992]. This conflict may be explained, at least in part, in terms of the change brought about by providing amplification. The amount of REIG [1-4 kHz] was around 12[3], 19, 19 and 31 dB in the studies by Horwitz and Turner, Cox *et al.*, the present study and Gatehouse respectively. The studies providing most amplification are also the studies that show a reduction in unaided performance. Essentially, if the fitted ear becomes experienced at hearing speech at a relatively high sensation level then, over time, it will perform poorly at the lower, less familiar, level associated with unaided speech.

---

[3] Horwitz and Turner provided a 2-cc coupler gain at 2 kHz of ca. 12 dB. Since the CORFIG transform is approximately 0 dB at 2 kHz [Bentler and Pavlovic, 1989] the mean REIG will be around 12 dB.

In addition to the changes observed in the fitted ear, there was a progressive decrease in performance over time in the control ear. By 12 weeks post-fitting, the mean aided scores showed a statistically significant decrease of 2.2% while the unaided scores showed a trend towards reduced performance of 1.4%. This suggests that the control ear was essentially deprived of adequate stimulation relative to the fitted ear. In experiment two, the aided score in the control ear decreased by around 1.6 % at 12 weeks post-fitting. These findings are consistent with the retrospective reports of late onset auditory deprivation reviewed in section 2.3 although it occurs over a much shorter time scale in the present study. However, a similar finding was reported by Gatehouse [1992] who showed a statistically significant reduction in performance in three out of six test conditions in the control ear over a post-fitting time period of 12 weeks.

In summary, it appears that the more gain that is provided by the hearing instrument [and hence the greater the difference in sensation level between aided and unaided speech] the larger will be the change over time in the aided and unaided conditions in both the fitted ear and the control ear.

5.3.4 The influence of presentation level

There is a trend towards increasing benefit over time as the presentation level increases. In support of this trend, one-factor [time] repeated-measures ANOVA were performed on each presentation level separately. These showed that there was no statistically significant change in benefit [or aided performance] over time for a presentation level of 55 dB SPL [$p>0.05$] but there was a statistically significant increase in benefit [and aided performance] over time at a presentation level of 69 dB SPL [$p<0.01$]. The lack of statistically significant interaction between presentation level and time occurred because the statistical analysis was underpowered. The mean difference between the increase in benefit at 55 and 69 dB SPL was 3.8% but the standard deviation was 8.3%. The higher than expected inter-subject variability means that the statistical power was only around 40%: this resulted in a material risk of a Type II error. In the first two experiments, the aided performance score of each subject, at the time of fitting, was close to 71%. In experiment three, no attempt was made to set aided performance at a criterion level so scores varied across subjects. It is possible that the rate and/or magnitude of improvement may differ based on the initial performance and this may account for the higher variability in the present study. Approximately double the number of subjects would have been required to increase the statistical power to 80%. Alternatively, the effect size may have increased if the design had been modified to include a longer post-fitting time period, a wider range of presentation levels or if subjects were provided with greater REIG.

At an international workshop in Sweden, 1997, Gatehouse gave an invited presentation on the topic of 'Speech tests as measures of outcome'. The presentation, which has now appeared in a journal supplement, included a short summary of the results of an experiment to investigate acclimatisation effects to non-linear processing [Gatehouse, 1998]. One reported finding was that acclimatisation effects with linear hearing instruments differed across presentation level: acclimatisation was present [on FAAF; Gatehouse, personal communication] at a presentation level of 65 and 75 dB SPL but not at 55 dB SPL. While the study has yet to be published in detail, this finding is consistent with the level dependent effects observed in the present experiment. No acclimatisation studies have been published that report the effect of different speech presentation levels. However, there have been studies that show an intensity-dependent acclimatisation effect. Gatehouse [1989] measured SNR for 50% performance on FAAF in 24 monaurally aided subjects and showed an interaction between presentation level [65-90 dB SPL] and ear. Criterion performance was obtained with a less favourable SNR when the fitted ear was tested at high levels compared to the not-fitted ear while the reverse occurred at lower presentation levels. This intensity dependence was taken as evidence that an ear performs most efficiently at its familiar listening level. The difference in SNR between ears at the highest presentation level of 90 dB SPL was 2.9 dB. The difference in aided performance between the two ears in the present study, at a presentation level of 69 dB SPL, was 5.6% [test ear +3.6, control ear –2.0]. To allow a comparison with the change in SNR from the Gatehouse study, the difference in performance can be converted into a difference in SNR using the average FAAF performance intensity function obtained from the hearing impaired subjects reported by Shields and Campbell [2001]. At a performance level of 70%, a change of 3% occurs for every 1 dB change in SNR. Therefore, the difference between ears of 5.6% corresponds to a difference in SNR between ears of approximately 2 dB.

Thus, the magnitude of change expressed in terms of SNR is similar, if a little lower, than Gatehouse. There are several possible explanations for the smaller difference in predicted SNR. Firstly, hearing instrument use was limited to 12 weeks in the present study whereas the mean duration of hearing instrument use in the Gatehouse study was 4.8 years. The results in the present study may have approached that reported by Gatehouse if the post-fitting duration had been longer. Secondly, the input-output function in Figure 5.3 for FAAF-shaped noise shows that the aided presentation level was approximately 12 dB above the unaided presentation level. This means that an input SPL of 69 dB would result in an output SPL of 82 dB SPL. The 2.9 dB reported by Gatehouse occurred at a SPL of 95 dB. The difference in SNR between ears was 2 dB when averaged over 80 and 85 dB SPL. Therefore, there is some similarity between the two studies.

In 1996, Robinson and Gatehouse reported an experiment that measured the ability to discriminate changes in intensity in five new users of monaural hearing instruments. After 12-18 weeks of hearing instrument use, the subjects displayed level- and frequency-dependent changes in intensity discrimination for the fitted ear but not the control ear. Specifically, the fitted ear was shown to be superior at high presentation levels but inferior at lower presentation levels. These changes only occurred at the frequencies that received material benefit from amplification.

Gatehouse and Robinson [1996] subsequently reported an experiment that showed changes in loudness function in four subjects who were fitted monaurally. Specifically, sounds were rated less loud in the fitted ear than in the control ear over the range of intensities studies. Again, the changes were confined to the high frequencies where the hearing instrument provided benefit. In the same study, the authors also reported a single case study where the changes in intensity discrimination were accompanied by changes in amplitude of the N1-P2 complex of the slow vertex cortical evoked response. At a high intensity level the amplitude was greater in the fitted ear but smaller at lower intensity levels. These two studies show that improved intensity discrimination at a higher SPL in the fitted ear appears to be associated with a steeper loudness growth function, whereas at lower levels the reduced intensity discrimination ability in the fitted ear appears to be associated with a less steep loudness function.

Gatehouse has subsequently presented unpublished data from five subjects showing a close coupling between changes in audibility, intensity discrimination and speech recognition performance. Specifically, this retrospective study demonstrated that the improvement in intensity discrimination and speech recognition were limited to the high frequencies where amplification had improved audibility. Taken collectively, these studies show a level-dependent effect and this has generally been explained in terms of the ear performing most efficiently at the presentation level to which it has typically been exposed.

The present study supports a level-dependent effect but the pattern of results is not entirely consistent with the notion that this is due to improved audibility. Since the SII is primarily a measure of audibility, acclimatisation should occur at the presentation level where amplification resulted in the greatest change in SII. The SII values, summarised in Table 5.5, show that the greatest change occurred at a presentation level of 62 dB SPL. However, at the time of fitting, the greatest change in performance occurred at a presentation level of 55 dB SPL. It is possible that aided performance at 62 dB SPL [and 69 dB SPL] was reduced due to upward spread of masking caused by the high speech and masker level.

Despite the discrepancy between the SII and the measured performance, the largest change over time occurred at the highest presentation level. As a further check on the contribution that audibility made to the subject's performance, the data were split into two groups according to the magnitude of the acclimatisation effect. This allowed a comparison of the eight subjects who demonstrated the largest change with the remaining eight subjects who showed the smallest change. The results are shown in Table 5.12. The mean change in benefit for each group was 10.8 and 1.0 and the mean change in aided performance was 6.7 and 0.6. The difference between groups of 9.85 and 6.13 for benefit and aided performance respectively, was statistically significant on independent samples *t*-tests [$p<0.01$]. Table 5.12 also shows the change in audibility [aided minus unaided SII values] for each group. The difference in audibility between groups is small and not statistically significant on independent samples *t*-tests [$p>0.05$]. Thus, the difference in performance between the two groups is not supported by the hypothesis that this is a direct result of changes in audibility.

Table 5.12. Mean change in performance of the fitted ear at 12 weeks relative to the time of fitting at a presentation level of 69 dB SPL. The data for the 16 subjects has been split into two groups according to the magnitude of change. Group A is the eight subjects who showed the largest changes while Group B is the eight subjects who showed the smallest change. Also shown is the mean change in SII after aiding. One SD is given in brackets.

	Change in performance		Change in SII	
	Group A	Group B	Group A	Group B
Benefit	+10.8 [4.9]	+1.0 [3.3]	0.17 [0.09]	0.30 [0.17]
Aided	+6.7 [3.1]	+0.6 [2.4]	0.19 [0.10]	0.29 [0.18]

An explanation for the maximum acclimatisation effect occurring at the highest presentation level may be related to the subject's ability to extract speech from background noise. Since the SNR was fixed across presentation levels, the level of noise increased along with the speech presentation level. It is possible that the subject required time to learn to extract the speech signal from the background noise. There is anecdotal evidence that the ability of factory workers to understand speech in noise improves over time.

The FAAF-shaped noise was used to limit the speech recognition scores at the time of fitting. It was not selected to be representative of everyday listening situations. Although the noise levels experienced by the subjects in their every day environments were not measured, Pearsons *et al.* [1977] reported the SNRs measured for conversational speech in and around urban homes to be of the order of +5 to +9 dB. This is a more favourable SNR than the –2 dB used in the present study. It is also likely that hearing impaired subjects would develop strategies to improve adverse SNRs. Such strategies may involve moving nearer the speaker, increasing the level of radio and television, or avoiding difficult listening situations altogether. Thus, it appears unlikely that there would have been sufficient opportunity to learn to extract a signal at the negative SNRs used in the present study. In addition, it does not explain why performance decreased over time in the fitted ear when the hearing instrument was removed. However, more recently, Gatehouse [personal communication] has shown that the SNR is less favourable at high listening levels than at lower listening levels [presumably because the level of speech has to be raised in noisy environments]. Since experiment three used a relatively poor SNR, it is possible that it was only at the high presentation level that it resembled a familiar listening condition. This might explain why acclimatisation was most marked for the highest presentation levels.

It may be possible to test the above explanation by designing a study that involves the test material being presented with and without masking noise. If performance increases under both conditions, then acclimatisation cannot be due solely to the ability to extract speech from noise. In order to avoid a ceiling effect with the FAAF test in quiet, it may be possible to use selective filtering. An example would be to use a band pass filter around 1.5-3 kHz since this would restrict performance to changes that occur in the amplified frequency region. Alternatively, a more difficult speech test, such as a confusion test with many more alternatives could be used since performance is highly dependent on the constraints placed upon the message being communicated. The greater the constraints, the lower the performance for a given level of audibility. The effect of using nonsense syllables such as the CVC test would be to shift the performance intensity function to the right [i.e., poorer performance for a given SNR].

Table 5.13. Conceptualisation of the range of SPL to which the ear performs most efficiently before and after provision of amplification. Before amplification, speech is presented over a range of SPL from 55-70 dB SPL. After provision of a linear hearing instrument, the ear will learn to perform most eficiently at 70-85 dB SPL

Input SPL	40	45	50	55	60	65	70	75	80	85	90	95	100
Pre-amplifiation				■	■	■	■						
Post-amplification							■	■	■	■			

An alternative explanation is that acclimatisation occurred as a direct consequence of the high presentation level of speech and is unrelated to the SNR. This is illustrated in Table 5.13. The SPL of speech before and after amplification is shown as the filled boxes. Before amplification, the SPL of speech is typically in the range of 55-70 dB SPL. Thus, the ear will function most efficiently over this range. After amplification, the range of speech will be raised, for example, by 15 dB so speech will now be 70-85 dB SPL. A period of time may be required for the auditory system to acclimatise to the new range of SPL. Thus, over time, performance will increase over the new range of SPL but decrease at the unaided range. In the above example, there may be little change when speech at 55 dB SPL is amplified because, in terms of overall SPL, this is no different from the higher levels of speech before amplification.

Although the physiological mechanism responsible for this change is not yet understood, it may be similar to the neural plasticity changes observed in the frequency domain. At the level of the auditory nerve, it has been proposed that intensity coding is achieved by the activation of neurones with different threshold and dynamic range characteristics. The sensitivity of neurones to intensity is typically depicted by means of a spike-rate intensity function where the spike discharge rate of a single neurone is plotted against stimulus intensity. It provides a quantitative measure of the threshold and the intensity dynamic range of the neurone for that stimulus. Figure 5.17 shows examples of different rate-level functions from the cat [Sachs and Abbas, 1974]. For neurones with a high rate of spontaneous activity and low intensity threshold, the discharge rate increases with level in a sigmoid fashion, reaching a maximum discharge rate about 20 dB above threshold. The neurones with a low spontaneous discharge rate and high intensity threshold do not fully saturate although, at very high intensity levels, the rate of discharge is greatly reduced. The dynamic range of these neurones can be nearer 40 dB. Thus, there are neurones that encompass the dynamic range of the ear. While neurones within the CANS also show sigmoid and non-saturating rate-intensity functions, there are also neurones with non-monotonic functions. For these neurones, the discharge rates increase with intensity until they reach a maximum and then decrease at higher intensities. An example of non-monotonic functions from neurones in the auditory cortex of the monkey was reported by Pfingst and O'Connor [1981] and is reproduced in Figure 5.18. The intensity that elicits the maximum firing rate is usually referred to as the neurones 'best' SPL. The best SPL for these non-monotonic neurones in the cortex may be as low as 15 dB SPL or as high as 106 dB SPL.

Not only is there clear evidence that neurones respond to different intensities but there are studies that show a topographical map for intensity within the CANS. Heil *et al.* [1992] showed that, in the auditory cortex of the cat, neurones with a best SPL at high intensity levels alternate with neurones that have a best SPL at low intensity levels. This pattern of alternating high and low threshold neurones occurred in different cats although there was considerable variation in the physical location of high and low threshold neurones as well as the overall range of intensities that could be detected. Schreiner *et al.* [1992] have also shown that neurones are not randomly distributed

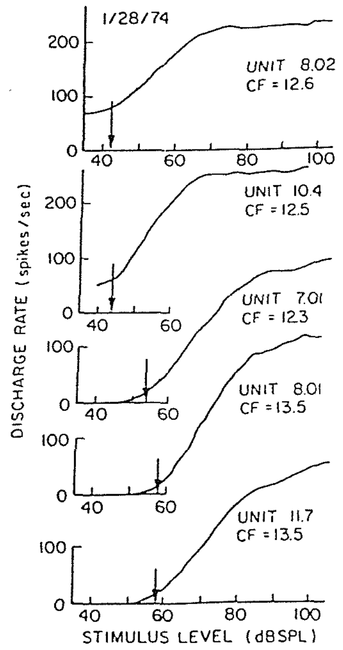


Figure 5.17. Rate-level functions [at CF] for five cochlear nerve fibres from a single cat. The fibres have CFs of 12.3-13.5 kHz. The arrows indicate the mean rate thresholds. Reproduced from Sachs and Abbas [1974] with permission.

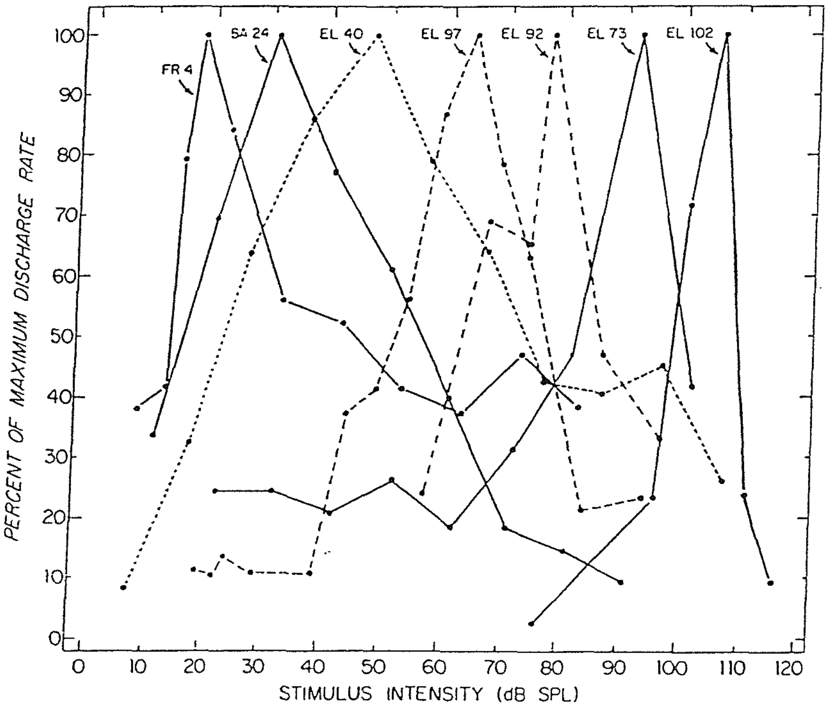


Figure 5.18. Non-monotonic rate-level functions from seven cortical neurones of the unanaethetised monkey performing an auditory task. Reproduced from Pfingst and O'Connor [1981] with permission.

across the auditory cortex but are arranged in clusters that respond maximally to high and low intensity levels. In a further study, Heil *et al.* [1994] investigated the spatial representation in the topographical map for changes in intensity level. An increase in intensity of a low level signal resulted in an increase in activity in the spatial areas where low threshold neurones were present. An increase in intensity of a high level signal resulted in an overall change in the spatial distribution of activity without any overall change in activity. This is probably because increased activity in some neurones is offset by reduced activity in other, non-monotonic, neurones. Thus, intensity could be represented in the cortex of the cat by both the neuronal discharge rate in small groups of cells and by the spatial distribution of activity. A schematic representation of the changes that occur with increasing stimulus intensity is illustrated in Figure 5.19. The topographic map of alternating high and low SPL neurones along an iso-frequency contour is represented by three high SPL neurones and two low SPL neurones [denoted H and L respectively]. At low intensity levels, activity occurs in low SPL neurones [diagram A]. Small increases in intensity result in increased activity in these areas [diagram B]. At higher intensity levels, activity can also be detected in high threshold neurones and this results in activity over a wider area [diagram C]. At even higher intensity levels, the distribution of activity changes although the overall level of activity remains relatively constant [diagram D]. The transition between low and high threshold neurones has been shown as abrupt changes although in real life there is a more gradual transition. Thus, there is evidence that neurones are tuned to a specific intensity level.

It is known that there is plasticity within the CANS of mature animals and studies reporting changes in the frequency domain were reviewed in Chapter Two. It is not clear where, in the auditory system, these changes take place although Harrison [2001] suggests that, in the mature animal, this is likely to be confined to levels above the midbrain. Little is known about the mechanism that allows the cortex to reorganise to behaviourally important stimuli while ignoring irrelevant stimuli although there is growing evidence that this decision takes place in the basal nucleus that is situated in the forebrain. For example, Kilgard and Merzenich [1998] and Weinberger [1998] have shown that when an auditory stimulus is paired with electrical stimulation to the basal nucleus, there is substantially more reorganisation of the auditory cortex than

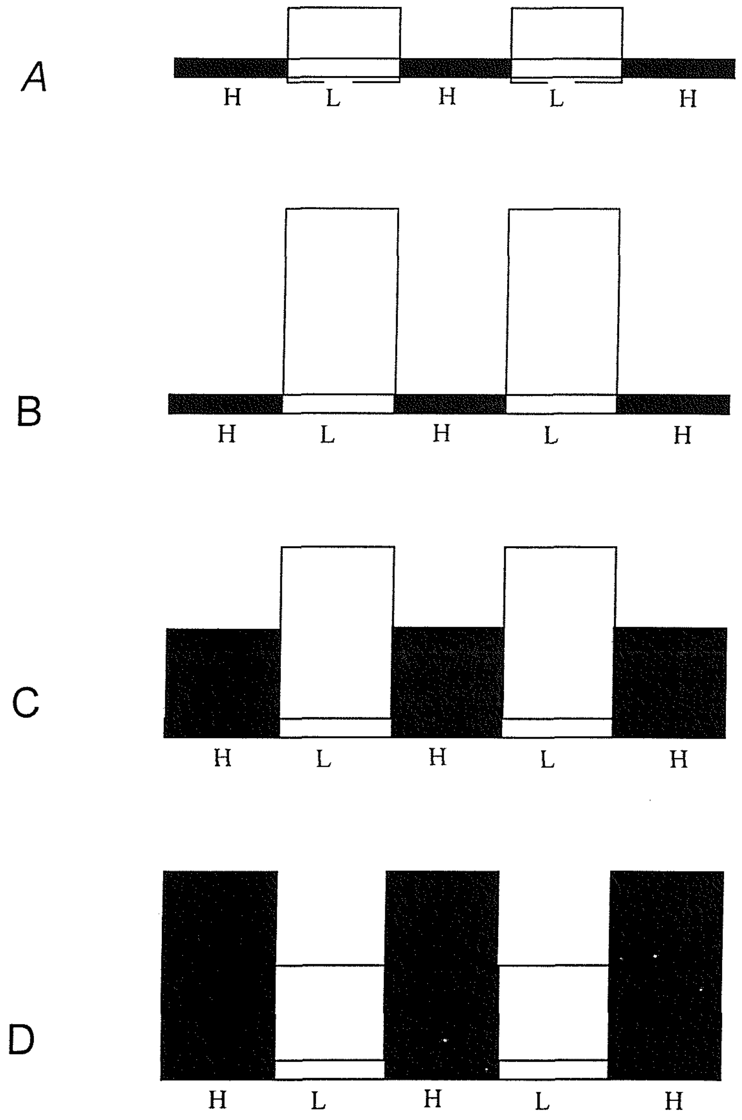


Figure 5.19. A schematic representation of changes in the normal auditory cortex with increasing stimulus intensity. The topographic organisation of neurones in the auditory cortex are shown in a periodic pattern. This is illustrated by the alternating high and low threshold neurones [denotes H and L]. At low intensity levels [diagram A] activity can be detected in neurones with a low threshold. As the intensity increases the activity in these low threshold neurones increases. This is shown by the increasing height of the boxes in diagram B. At higher intensity levels activity can be detected in both the low threshold and high threshold neurones. This is represented by the black boxes in diagram C. At even higher intensity levels the overall activity within the auditory cortex remains relatively constant but the spatial pattern changes. This occurs because there are non-monotonic low threshold neurones that decrease in activity as the intensity levels. This is represented in diagram D by the reverse in height of the high and low threshold neurone activity.

providing the auditory stimulus alone. Thus, electrical stimulation of the basal nucleus has the same effect as presenting a behaviourally important stimulus [such as the food reward used by Recanzone *et al.* [1993] reported in section 2.1. Kilgard and Merzenich have also cited studies that show that learning can be impaired if the basal nucleus is damaged. This demonstrates that signals from the basal nucleus influence plasticity in the auditory cortex.

Although no study has reported reorganisation of the topographical map for intensity, it is possible that this is the physiological mechanism for the acclimatisation effect observed in the present study. With the provision of amplification and a high-level input, subjects have to listen at higher levels than previously experienced. This may require a change to existing, but previously less effective, synapses and neurones. Alternatively, there may be structural changes that take place that may involve, for example, synapses being modified in size or there may be local sprouting of axons and dendrites. With the onset of a hearing impairment, one might envisage that high-threshold regions of the cortex may become underused because the sound levels never reach sufficient levels to activate them. This might lead to the lower threshold regions invading the unused cortex due to lack of competition from high stimulus levels. After provision of amplification, the cortex has to reorganise or re-acclimatise to the range of high signal levels. The effect is that the auditory system somehow increases its 'representation' to high levels of speech, at the expense of the previously lower unaided levels, in order to make maximum use of the newly amplified speech signal. This suggests that larger changes in listening level will result in greater reorganisation. This may explain why Gatehouse reported a larger acclimatisation effect in subjects who were fitted with hearing instruments that provided greater gain than in other studies. Reorganisation of intensity is consistent with the re-mapping of the relationship between intensity and loudness reported by Gatehouse and Robinson [1996]; the capacity for discriminating differences in intensity shift from lower to higher intensities as subjects gain amplification experience.

In summary, it is inferred that, after provision of a hearing instrument, the auditory cortex reallocates its resources in the intensity domain, in addition to changes in the frequency domain. This results in an increase in cortical representation to the behaviourally important speech sounds that are now presented at a higher intensity than previously with a corresponding reduction in representation at lower intensities.

5.3.5 Predicting individuals who will show changes over time

In keeping with previous acclimatisation studies, there is a wide range of variability in performance across subjects. This is a common finding in studies of perceptual learning; in a recent article, Ahissar [2001] reports that the rate of learning and the specificity of learning differ greatly across subjects. It is of interest to determine if the subjects showing acclimatisation could be predicted from some of the variables measured in the study. The variables of interest are summarised in Table 5.14. The subjects were split into two equal groups based on the amount of increase in benefit that occurred over time in the fitted ear using a presentation level of 69 dB SPL. Group A showed the smallest increase while group B showed the largest increase. The mean difference in age between groups was small and there was no difference in the ratio of male and female subjects within each group. The mean hearing threshold at 2-4 kHz was 3 dB better in the group showing most change over time. This probably explains why there was slightly less insertion gain in this group [and hence the smaller difference on SII]. The output SPL that caused the hearing instrument to saturate was similar for both groups. There was no difference in the self-reported number of hours of daily use of the hearing instrument. The SNR used for FAAF testing was 1 dB poorer in the group with the slightly better hearing thresholds. This probably also explains why the mean aided FAAF score was slightly lower in this group and why the initial benefit score was also poorer. On inspection of the data, one individual [subject 12] was responsible for negative benefit observed in Group B. When this subject was removed from the analysis the mean unaided and aided scores become 71 and 72% respectively [with negligible change to the other variables]. None of the differences between groups were statistically significant on independent samples *t*-tests [$p>0.05$]. A similar outcome was obtained when the subjects were categorised according to changes in aided performance. This was not surprising since this made little difference to the subjects allocated to each group. Thus, for the small sample size of 16 new hearing instrument users, no clear associations were observed. However, a larger group of subjects would be required to more appropriately assess these relationships.

Table 5.14 Summary of mean difference in variables with the subjects split into two groups according in the change in benefit at 12 weeks post-fitting. Subjects showing the smallest increase are in Group A [n=8] and subjects showing the largest increase are in Group B [n=8]. The values for air conduction threshold, insertion gain and SII were all calculated as the average at 2-4 kHz. None of the differences were significant on independent sample *t*-test [p>0.05].

	Group A (small)	Group B (large)	Difference
Age [years]	69.6	70.6	1.0
Sex [female:male]	3:5	3:5	0
AC [dB HL]	60	57	3
Hearing instrument			
Fitted ear [R:L]	5:3	5:3	0
Insertion gain [dB]	20	17	3
Change in SII [aided minus unaided]	0.18	0.13	0.05
Output SPL before non-linear	84	86	2
Self-reported daily use >8 hrs	8	8	0
Speech recognition testing			
SNR [dB]	–1	–2	1
Unaided FAAF [%]	71	72	1
Aided FAAF [%]	73	70	3
FAAF benefit	+2	–2	4

Two previous studies have investigated the relationship between the amount of high frequency insertion gain and changes in benefit over time. Cox *et al.* [1996] used three hearing instruments that provided small differences in high frequency insertion gain. There were differences in the change in benefit scores on the CST but this did not appear to be related to differences in high frequency insertion gain. Horwitz and Turner [1997] showed no correlation when comparing the amount of 4 kHz insertion gain with the change in benefit measured by the NST. Like the present study, both studies used small numbers of subjects [22 and 13 respectively] and were not designed to specifically investigate the relationship between insertion gain and changes in benefit. The statistical power in the Cox study was only 44%. Therefore, a more comprehensive study investigating this relationship has yet to be published.

In a recent review article, Kricos [2000] summarised a number of audiological and non-audiological variables that interact and influence the eventual outcome of auditory rehabilitation. The non-audiological factors included race/ethnicity, gender, age and personality. It is also likely that there are psychological variables that impact on a subject's ability to learn. Gantz *et al.* [1993] derived a predictive index of accuracy of word recognition in sentences from a multiple regression analysis. The prediction was based on a combination of measures of peripheral neural survival [duration of deafness and hearing sensitivity], motivation [Krantz Health Opinion Survey and CPHI], and cognition [Visual Monitoring Task and measures of Lip-reading]. Tait *et al.* [2000] have shown that up to 25% of the variance in speech identification performance of young children who receive a cochlear implant may be predicted from characteristics that are inherent to the child before implantation. The characteristics are represented by the demonstration of autonomy in preverbal communication interactions.

5.3.6 Interpretation of individual versus group data

It is not clear what the practical significance of the acclimatisation effect is in terms of the benefit experienced by the subject or of the reduction of disability and handicap they may experience. The mean benefit score at the time of fitting at a presentation level of 69 dB SPL was 0% but this increased to 5.9% after 12 weeks of hearing instrument use. This means that the subsequent benefit far exceeds the initial benefit score. However, for most subjects, the main purpose of the hearing instrument is to make speech audible and this is less likely to be a problem at these relatively high presentation levels where the acclimatisation effect is most apparent. It is possible that a subject who is frequently in an environment where speech is at a relatively high presentation level may benefit more than a subject who is in an environment with a relatively quiet presentation level. However, Gatehouse [1997] has demonstrated that the benefit perceived by the subject in their normal listening environment is best described using a combination of speech presentation levels and signal-to-noise ratios. He also demonstrated that the maximum correlation between reported benefit occurred for the response time [not recognition score] of the speech task that is related to ease of listening. Thus, there is concern about the extent to which laboratory test conditions can provide reliable information concerning the listening environments experienced by the subject.

The most extensive study of hearing aid outcomes amongst elderly people in terms of the combination of the number of subjects and the number of outcome measures was Humes *et al.* [2001]. They used 26 outcome measures in 173 elderly hearing aid users after one month of hearing aid use. They identified seven independent dimensions of outcome including self-reported benefit and satisfaction, benefit on performance tests, sound quality rating and use of hearing aid. Thus, a more complete picture of hearing aid outcome requires measurement of more than the subject's recognition score on the FAAF test.

The changes reported in the present study were obtained with monosyllabic words presented in steady noise. This may not be related to the ability to understand everyday speech in a simple manner. Boothroyd and Nittrouer [1988] have described the complex interactions between recognition of individual phonemes, words and whole sentences. For example, a moderate score on a monosyllabic word recognition test may result in an excellent score on a sentence test. It is also known that sentence material designed to simulate 'real world' speech signals do not depend as strongly on high frequency information as does the FAAF test. Studies have yet to be undertaken that investigate the implications of auditory acclimatisation in every day listening situations. This may include a comparison of changes in performance between the FAAF test and BKB sentences at high presentation levels.

CHAPTER SIX

CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusions

A comprehensive literature review revealed that some studies show an acclimatisation effect while other studies have reported no change. This conflict can be understood if auditory acclimatisation is viewed in terms of perceptual learning: reorganisation of the CNS to a change in behaviourally important stimuli can lead to an improvement in performance. Thus, acclimatisation will not be manifest if subjects had little opportunity to assimilate new skills required to perform the task used for its measurement; for example, if they have only a mild hearing impairment or if they have previous experience of amplification. Furthermore, acclimatisation will not be manifest if the task does not include the new skills that have been assimilated. The conflict in the literature may be due to variations in the tasks used to measure acclimatisation and in the opportunities to develop new skills.

A number of factors conducive to perceptual learning were taken into account when designing the experiments. A recognised prescription target was used to ensure that there was a change in audibility after aiding. Hearing instruments were worn for at least 6-8 hours per day by more than 80% of subject. A homogeneous group of subjects with a hearing impairment consistent with previous successful acclimatisation studies was used since the generality of acclimatisation for different degrees and configurations of hearing impairment is not known. In addition, subjects were fitted with the same model of linear hearing instrument since it is not known if acclimatisation effects are the same for different amplification and processing strategies. Performance was measured at fixed gain and user gain using a speech recognition test that is sensitive to changes in high frequency audibility was used. Despite controlling for the factors listed above, there was no evidence of acclimatisation in the first two experiments. This suggests that there must be additional factors not yet considered. Studies showing acclimatization have tended to be those where there was a small initial benefit score whereas many of the studies that have not demonstrated acclimatisation show sizeable initial benefit. Small benefit scores suggest that the speech test materials were presented at a relatively high sound pressure level; at high levels the unaided speech is already audible so the difference between the aided and unaided score is small.

The results of the third experiment confirm that presentation level is important when measuring acclimatisation. Acclimatisation was present at high presentation levels but not lower ones: this may explain the conflict reported in the literature. Acclimatisation does not appear to be measurable at the moderate presentation levels such as that used to represent conversation speech [but may have been measurable if amplification had been greater]. When amplified, these levels may not extend the functional auditory dynamic range beyond that present before amplification: presentation level to use appears to be a high level that ensures aided speech extends materially above the range of speech experienced before aiding. However, care has to be taken not to use excessively high levels since acclimatisation may not be measurable at levels exceeding everyday listening levels. Experiment three does not support the commonly held notion that acclimatisation is related to learning to make use of speech cues that were previously inaudible and are reintroduced by aiding: if this were the case, learning would have been expected to be greatest for lower presentation levels where the change in audibility is most evident. The fact that acclimatisation is related to the higher end of the intensity range suggests some re-allocation of neural resources to areas of the CNS stimulated by high-level stimuli. It is hypothesised that, after provision of a hearing instrument, the auditory cortex reallocates resources in the intensity domain, in addition to changes in the frequency domain. This results in an increased in cortical representation for the behaviourally important speech sounds that are presented at a higher intensity than previously with a corresponding reduction in representation for lower intensities. These processes take time to occur and can explain the phenomenon of auditory acclimatisation. Because the reallocation in the intensity domain involves the upper end of the intensity range, auditory acclimatisation is most readily demonstrated using speech at relatively high presentation levels. The present study demonstrates clear acclimatisation for speech at 69 dB but minimal acclimatisation for speech at 55 dB, consistent with the above explanation. The findings from experiment three indicate that future studies on auditory acclimatisation should include speech at relatively high presentation levels.

The lack of acclimatisation in the control ear is consistent with the auditory training studies showing some degree of specificity to the training stimulus. Acclimatisation can be viewed as perceptual learning that has occurred without formal training. Representation of the auditory environment can be considered perceptually in terms of space having [at least] three dimensions: intensity, frequency and laterality. The mapping of the physical parameters [intensity, frequency, side] onto the corresponding perceptual dimensions is not fixed but may change due to plastic processes within the CNS. Plastic changes occur for commonly occurring and important stimuli such as everyday speech used for communication and social/emotional purposes. They occur slowly in response to sustained change in effective stimulation [e.g., hearing impairment, amplification]. Hearing impairment leads to mapping favouring relatively low effective intensities and a preponderance of low-frequency information within the available perceptual space. If hearing impairment is asymmetrical, there may also be changed mapping to favour the better hearing side. When a hearing instrument is used consistently, further re-mapping occurs that tends to reverse those effects: a greater part of the perceptual space is given back to the higher effective intensities of amplified speech, to the preferentially amplified higher frequencies and in favour of the aided ear [for unilateral instrument use]. The acclimatisation effects shown for higher intensity speech and the lack of acclimatisation shown for lower intensity speech in the present and other studies can be understood within this framework.

It is not possible to conclude the time course of acclimatisation other than to report that is longer than 12 weeks. It is possible that the rate, as well as the magnitude, of change may vary with presentation level: the change may be slower for presentation levels that occur relatively infrequently in everyday life because there is less opportunity for perceptual learning to occur.

Subjects were asked to compare present level of benefit and satisfaction with two retrospective times: at the time of fitting and on the previous visit three weeks earlier. Compared to the time of fitting, there was no further reported increase in benefit and satisfaction; however, there was an increase in both reported benefit and satisfaction when compared to the previous visit. These results are paradoxical and highlight the difficulty associated with retrospective reporting. The use of self-report questionnaires for measuring acclimatisation should be treated with caution. A paired comparison technique did not show any preference for the fitted frequency response over time. This may be related to the presentation level used for the quality judgements or because the difference between the two responses was not sufficiently large.

The literature review identified five conditions [summarised in Table 1.1] that were thought to be necessary for an effect of acclimatisation to be measured. The description of the first two conditions can now be modified as a result of the knowledge gained from the present series of experiments; these conditions concern the choice of subject and the definition of what is meant by providing 'new information'. A modified list of the conditions that are required for auditory acclimatisation to be measured is given in Table 6.1.

The modified conditions can be explained by referring to Figures 6.1-6.3. Figure 6.1 shows the threshold of hearing [in dB SPL at the eardrum] in one ear of a hypothetical subject with a moderate-to-severe high frequency sensorineural hearing impairment. Also shown [as dashed lines] are the average long-term levels for quiet and raised speech. These dashed lines represent the upper and lower boundary for the 'typical' range of speech levels experienced by the subject before amplification [for simplicity, each level of speech is represented by a single line and the change in frequency response associated with different vocal efforts has been ignored]. For this hypothetical subject, quiet speech is audible from 0.25 to 0.75 kHz whereas raised speech is audible across most of the frequency range. Figures 6.2 and 6.3 show the effect of amplifying quiet and raised speech respectively. For simplicity, the figures assume linear amplification. Figure 6.2 shows the effect of high frequency amplification of quiet speech. The audibility of quiet speech now extends from 0.25 to 3 kHz. This increase in audibility will result in an improvement in performance on a speech recognition test immediately after fitting. Despite the increase in audibility of quiet speech after fitting, the newly amplified speech signal still falls within the same range of levels experienced before fitting. This scenario will be referred to as Model A. Figure 6.3 shows the effect of amplifying raised speech. Raised speech continues to be audible after aiding; for this reason, the immediate improvement in performance at the time of fitting is expected to be less than that observed with quiet speech. This scenario will be referred to as Model B.

Table 6.1 Modified list of conditions required for the measurement of auditory acclimatisation. This list is similar to the one shown in Table 1.1 except for revisions to the first two points.

1. hearing-impaired and possibly normal hearing subjects
2. re-mapping of the relationship between intensity and loudness
3. outcome measure should be sensitive to the changes that occur as a result of acclimatisation
4. outcome measure should be reliable enough to show the effect size and be free from floor/ceiling effects
5. there should be a control condition so that improvements due to wearing the hearing instrument can be differentiated from other improvements

In real-life, the new information provided by the hearing instrument depends on the range of speech levels experienced before fitting as well as the gain of the hearing instrument. As a result, the situation may be more complex than indicated in Figures 6.1-6.3; there may be a requirement for a combination of both Model A [i.e., change in audibility] and Model B [i.e., new listening level] to occur before acclimatisation is measured. In experiment three, amplification resulted in an increase in audibility, as calculated using the SII, for all speech presentation levels. Acclimatisation was greatest for the highest presentation level where the change in audibility was smallest. However, the largest change in listening level, relative to the range of levels experienced before aiding will occur for the highest speech presentation level. This requires the greatest extent of re-mapping [see Figure 1.1]. This finding suggests that presenting speech at a higher level than previously experienced routinely before aiding is necessary for the measurement of acclimatisation.

It is not clear if the new presentation level needs to be accompanied by a change in audibility for performance to improve over time. For example, if the amplification characteristics of the hearing instrument were changed so that the low frequencies [that are audible before fitting] were amplified, then there may still be an improvement in performance over time because of the new and unfamiliar listening level at these frequencies. For the same reason, it is possible that acclimatisation may occur in normal hearing subjects if there is a change in presentation level. This is because the information that is 'new' to the subject is the change in level of the speech signal rather than a change in audibility as illustrated in Figure 1.1.

The theoretical and empirical analyses reported in this thesis add in the following ways to our understanding of the conditions under which auditory acclimatisation can be measured. First, it may be incorrect to assume that the subject needs to be hearing impaired for measurement of acclimatisation. Second, amplifying speech to a new, less familiar listening level is required for measurement of acclimatisation. Finally, it is not clear if this change needs to be accompanied with a change in audibility.

6.2 Recommendations

It would be useful to extend the final experiment by including a wider range of presentation levels and to use a longer post-fitting time. The rate as well as the magnitude of improvement may differ with presentation level. It is suggested that acclimatisation will not occur for presentation levels that are higher than those experienced commonly in everyday life since there will neither have been much opportunity for perceptual learning to occur nor any importance attached to such stimuli. A longer post-fitting time period is required to investigate the rate and full time course of acclimatisation.

It would be helpful to measure acclimatisation to speech in quiet to determine if the effect is due to signal presentation level or the ability to extract speech from noise. If performance increases in both quiet and in noise then acclimatisation cannot be due solely to the ability to extract speech from noise. In order to avoid a ceiling effect with the FAAF test in quiet, it may be possible to use selective filtering. An example would be to use a band pass filter around 1.5-3 kHz since this would restrict performance to changes that occur in the amplified frequency region. Other possibilities include speeded speech or minimal pairs that are acoustically very similar.

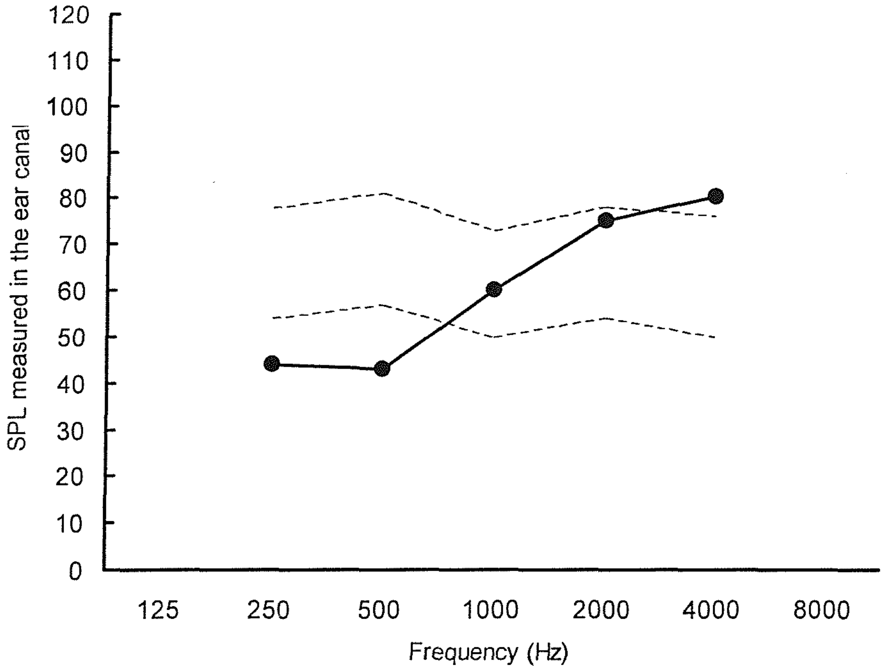


Figure 6.1. The relationship between the hearing thresholds of a hypothetical subject with a high frequency hearing impairment [filled circles] and unaided speech [dashed lines]. The lower dashed line represents the long-term average levels for quiet speech while the upper dashed line represents the long-term average levels for raised speech.

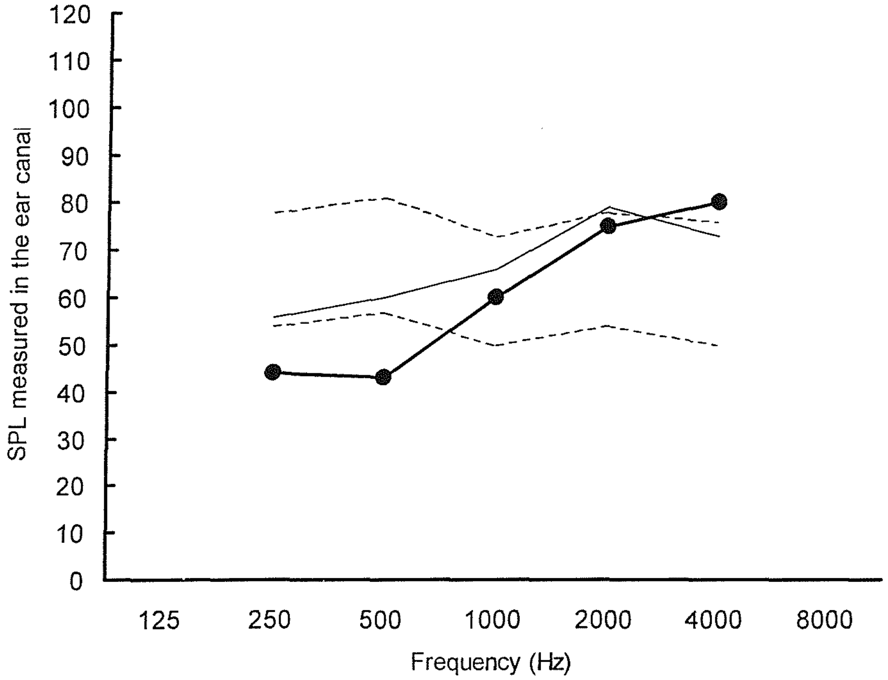


Figure 6.2. Same as Fig 1 but the effect of high frequency amplification on quiet speech has been added as a solid line. This scenario is referred to as model A.

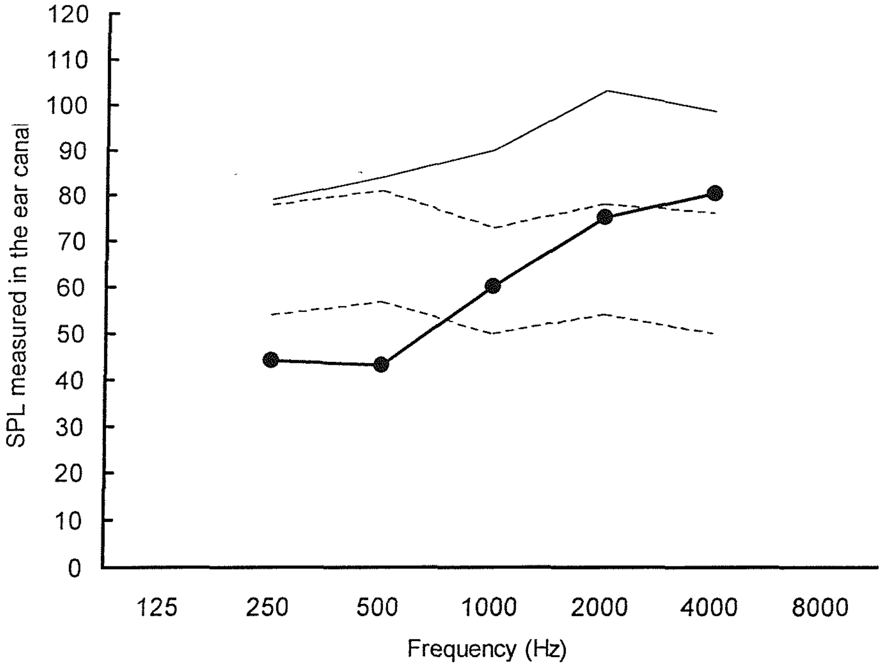


Figure 6.3. Same as Fig 1 but the effect of high frequency amplification on raised speech has been added as a solid line. This scenario is referred to as model B.

Most studies that have demonstrated acclimatisation have included participants fitted monaurally. It is not known if acclimatisation occurs as a result of asymmetrical aiding which disrupts binaural processing. The conceptual framework outlined above suggests that a preponderance of the neural resources will be reallocated towards the fitted ear. If this is correct, then acclimatisation may be less for binaural fitting where the resources are presumably shared between both ears equally. This hypothesis requires testing by using monaurally and binaurally fitted subjects and testing them both monaurally and binaurally. In studies where the participants are fitted monaurally, the not-fitted ear is occluded during testing. This asymmetry may enhance the acclimatisation effect perhaps by reducing the initial benefit score. This can be investigated by having the non-test ear occluded and unoccluded during monaural testing.

It would be informative to investigate the relationship between changes in psychoacoustic abilities, such as intensity discrimination, and acclimatisation to speech. The relationship between underlying psychoacoustic abilities and improvements in speech recognition may provide evidence of the underlying basic mechanisms responsible for acclimatisation. It would also be interesting to investigate the relationship between behavioural, electrophysiological and imaging techniques. This may identify the anatomical location for reorganisation and help to understand the anatomical and physiological changes underlying changes in performance. Imaging techniques such as PET and fMRI can provide estimates of the extent and location of activation produced by auditory stimulation. Since they provide a reliable method for assessing longitudinal changes in cortical activity, longitudinal studies may provide direct physiological evidence of a change in magnitude, locus or distribution of activity after hearing instrument fitting. Little use has so far been made of these imaging techniques to study acclimatisation.

It would be useful to determine what variables contribute to the magnitude and time course of acclimatisation. Variables may include age, cognitive abilities, acoustic environment, personality profile and audiometric/psychoacoustic measures. It may be possible to use these variables to predict the subjects who will show an acclimatisation effect. It is not known if acclimatisation varies with degree of impairment or, indeed, if it is present with both sensorineural and conductive hearing impairment. It is also not known if acclimatisation can be accelerated, for example, with training in speech perception tasks. Little work has been done to investigate the benefit from different processing strategies after allowing for acclimatisation. This information is important in establishing the clinical significance of the effect.

REFERENCES

Ahissar M. Perceptual training: A tool for both modifying the brain and exploring it. Proc Nat Acad Sc 2001; 98: 11842-3.

ANSI [1969]. ANSI S3.5-1969. American National Standard methods for the calculation of the Speech Intelligibility Index. New York: American National Standards Institute, Inc.

ANSI [R1997]. ANSI S3.5-1997. American National Standard methods for the calculation of the Speech Intelligibility Index. New York: American National Standards Institute, Inc.

Arlinger S, Billermark E. One year follow-up of users of a digital hearing aid. Brit J Audiol 1999, 33: 223-32.

Arlinger S, Billermark E, Oberg M, Lunner T, Hellgren J. Clinical trial of a digital hearing aid. Scand Audiol 1998, 27: 51-61.

Arlinger S, Gatehouse S, Bentler RA, Byrne D, Cox RM, Dirks DD, Humes L, Neuman A, Ponton K, Robinson K, Silman S, Summerfield AQ, Turner CW, Tyler RS, Willott JF. Report of the Eriksholm workshop on auditory deprivation and acclimatisation. Ear Hear 1996; 17: 87S-98S.

Balfour PB, Hawkins DB. A comparison of sound quality judgements for monaural and binaural hearing aid processed stimuli. Ear Hear 1992; 5: 331-9.

Bauer RW, Matusza JL, Blackmer RF. Noise localization after unilateral attenuation. J Acoust Soc Am 1966; 40: 441-4.

Bench J, Kowal A, Bamford J. The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. Brit J Audiol 1979; 13: 108-12.

Bentler R, Holte L, Turner C. An update on the acclimatisation issue. The Hearing Journal 1999, 52: 44-8.

Bentler RA, Neibuhr DP, Getta JP, Anderson CV. Longitudinal study of hearing aid effectiveness I: Objective measures. J Speech Hear Res 1993a; 36: 808-19.

Bentler RA, Neibuhr DP, Getta JP, Anderson CV. Longitudinal study of hearing aid effectiveness II: Subjective measures. J Speech Hearing Res 1993b; 36: 820-31.

Bentler RA, Pavlovic CV. Transfer functions and correction factors used in hearing aid evaluation and research. Ear Hear 1989; 10: 58-63.

Bilger RC, Neutzel JM, Rabinowitz WM, Rzeczkowski C. Standardization of a test of speech perception in noise. J Speech Hear Res 1984; 27: 32-48.

Boothroyd A. Evaluation of speech production in the hearing impaired: Some benefits of force-choice testing. J Speech Hear Res 1985; 28:185-96.

Boothroyd A, Nittrouer S. Mathematical treatment of context effects in phoneme and word recognition. J Acoust Soc Am 1988; 84: 101-14.

Brooks DN. Use of post-aural aids by National Health Service patients. Brit J Audiol 1981; 15: 79-86.

Butler RA. An analysis of the monaural displacement of sound in space. Perception and Psychophysics 1987; 41: 1-7.

Byrne D, Dillon H. The National Acoustics Laboratories (NAL) new procedure for selecting the gain and frequency response of a hearing aid. Ear Hear 1986; 7: 257-65.

Byrne D, Dirks D. Effects of acclimatization and deprivation on non-speech auditory abilities. Ear Hear 1996; 17: 29S-37S.

Cacace AT, Lovely TJ, Mcfarland DJ, Parnes SM, Winter DF. Anomalous cross-modal plasticity following posterior fossa surgery: some speculations on gaze-evoked tinnitus. Hear Res 1994; 81: 22-32.

Ching TYC, Newall P, Wigney D. Comparison of severely and profoundly hearing-impaired children's amplification preferences with the NAL-RP and the DSL 3.0 prescriptions. Scand Audiol 1997; 26: 219-22.

Ching TYC, Hill M, Birtles G, Beecham L. Clinical use of paired comparisons to evaluate hearing aid fitting of severely/profoundly hearing impaired children. The Australian and New Zealand J Audiol 1999; 21: 51-63.

Cox RM. The MSU hearing instrument prescription procedure version 3.0. Hearing Instruments 1988; 39: 6-10.

Cox RM, Alexander GC. Maturation of hearing aid benefit: Objective and subjective measurements. Ear Hear 1992; 13: 131-41.

Cox RM, Alexander GC, Gilmour C, Pusakulich KM. Use of the Connected Speech Test (CST) with hearing impaired listeners. Ear Hear 1988; 9: 198-207.

Cox RM, Alexander GC, Rivera IM. Accuracy of audiometric test room simulations of three real-world listening environments. J Acoust Soc Am 1991a; 90:764-72.

Cox RM, Alexander GC, Taylor IM, Gray GA. Benefit acclimatization in elderly hearing aid users. J Am Acad Audiol 1996; 7: 428-41.

Cox RM, Gilmour C, Alexander GC. Comparison of two questionnaires for patient-assessed hearing aid benefit. J Am Acad Audiol 1991b; 2: 134-45.

Cox RM, Moore JN. Composite speech spectrum for hearing aid prescriptions. J Speech Hear Res 1988; 31: 102-7.

Cullington H. Tinnitus evoked by finger movement: Brain plasticity after peripheral deafferentation. Neurology 2001; 56: 978.

Demany L. Perceptual learning in frequency discrimination. J Acoust Soc Am 1985; 78: 1118-20.

Dillon H. A procedure for subjective quality rating of hearing aids. NAL report No. 100, Camberra, Australia: Government Publishing Service, 1984.

Dillon H, Storey L. The National Acoustic Laboratories' procedure for selecting the saturation sound pressure level of hearing aids: theoretical derivation. Ear Hear 1998; 19: 255-66.

Dorman MK, Dankowskie K, McCandless G, Parkin JL, Smith L. Longitudinal changes in word recognition by patients who use the Ineraid cochlear implant. Ear Hear 1990, 455-9.

Dorman MF, Loizou PC. Changes in speech intelligibility as a function of time and signal processing strategy for an Ineraid patient fitted with continuous interleaved sampling (CIS) processors. Ear Hear 1997, 18: 147-5.

Eimas PD, Siqueland ER, Jusczyk PW, Vigorito J. Speech Percepton in infants. Science 1971; 171: 303-6.

Finney EM, Cobb S, Hickok G, Dobkins KR. Visual stimuli activate auditory cortex of the deaf. Nat Neurosci 2001; 12: 1171-3.

Finney FM, Dobkins KR. Visual contrast sensitivity in deaf versus hearing populations: exploring the perceptual consequencies of auditory deprivation and experience with a visual language. Cog Brain Res 2001; 11: 171-83.

Florentine F. Relation between lateralization and loudness in asymmetrical hearing loss. J Am Audiol Soc 1976; 1: 243-51.

Foster JR, Haggard MP. FAAF. An effective analytical test of speech perception. Proceedings of the Institute of Acoustics 1979; 182: 9-12.

Foster JR, Haggard MP. The four alternative auditory feature test [FAAF]: Linguistic and psychometric properties of the material with normative data in noise. Brit J Audiol 1987; 21: 165-74.

Foust KO, Gengel RW. Speech discrimination by sensorineural hearing-impaired persons using a transposer hearing aid. Scan Audiol 1973; 2:161-70.

Fryauf-Bertschy H, Tyler RS, Kelsay DM, Gantz BJ. Performance over time of congenitally deaf and postlingually deafened children using a multichannel cochlear implant. J Speech Hear Res 1992, 35: 913-20.

Gantz BJ, Woodworth GG, Knutson JF, Abbas PJ. Multivariate predictors of success with cochlear implants. Adv Otology Rhinology Laryngology 1993; 102; 909-16.

Gatehouse S. Apparent auditory deprivation effects of late onset: The role of presentation level. J Acoust Soc Am 1989; 86: 2103-6.

Gatehouse S. The time course and magnitude of perceptual acclimatisation to frequency responses: Evidence from monaural fitting of hearing aids. J Acoust Soc Am 1992a; 92: 1258-68.

Gatehouse S. The evaluation of the Sentence verification test for the assessment of hearing aid benefit. 1992b, Unpublished data.

Gatehouse S. Role of perceptual acclimatization in the selection of frequency responses for hearing aids. J Am Acad Audiol 1993; 4: 296-306.

Gatehouse S. Speech tests as measures of outcome. Sand Audiol 1998; 27(suppl.): 54-60.

Gatehouse S. Glasgow Hearing aid Benefit profile: Derivation and validation of a client-centred outcome measure for hearing aid services. J Am Acad Audiol 1999; 10: 80-103.

Gatehouse S. Outcome measures for the evaluation of adult hearing aid fittings and services. Scientific and Technical Report to the Department of Health, London, England, 1997.

Gatehouse S, Robinson K. Acclimatisation to monaural hearing aid fitting- effects on loudness functions and preliminary evidence for parallel electrophysiological and behavioural effects. In B. Kollmeier (Ed.), Psychoacoustics, speech and hearing aids. Singapore: World Scientific. 319-30, 1996.

Gelfand SA, Silman S, Ross L. Long-term effects of monaural, binaural and no amplification in subjects with bilateral hearing loss. Scand Audiol 1987; 16: 201-7.

Giolas TG, Owens E, Lamb SH, Schubert ED. Hearing Performance Inventory. J Speech Hear Disorders 1979; 44: 169-95.

Goldstone RL. Perceptual learning. Annu Rev Psychol 1998, 49: 585-612.

Harrison RV. Age-related tonotopic map plasticity in the central auditory pathways. Scand Audiol 2001; 30: Suppl 53: 8-14.

Hawkins D, Alvarez E, Houlihan J. Reliability of three types of probe tube microphone measurements. Hearing Instruments 1991; 42: 14-6.

Heil P, Rajan R, Irvine DRF. Sensitivity of neurons in cat primary auditory cortex to tones and frequency-modulated stimuli. III. Topographical representation along the 'isofrequency' dimension. Hear Res 1992; 63: 135-56.

Heil P, Rajan R, Irvine DRF. Topographic representation of tone intensity along the isofrequency axis of cat primary auditory cortex. Hear Res 1994; 76: 188-202.

Holte L. Acclimatisation to hearing aids. Presented at the Iowa Hearing Aid Conference. Iowa City, IA, 1997.

Horwitz AR, Turner CW. The time course of hearing aid benefit. Ear Hear 1997; 18: 1-11.

Humes LE, Barlow NN, Garner CB, Wilson DL. Prescribed clinician-fit versus as-worn coupler gain in a group of elderly hearing-aid users. J Speech Lang Hear Res 2000; 43: 879-92.

Humes LE, Garner CB, Wilson DL, Barlow NN. Hearing-aid outcome measures following one month of hearing aid use by the elderly. J Speech Lang Hear Res 2001; 44: 469-86.

Humes LE, Halling D, Coughlin M. Reliability and stability of various hearing-aid outcome measures in a group of elderly hearing-aid wearers. J Speech Hear Res 1996; 39: 923-35.

Humes LE, Wilson DL, Barlow NN, Garner C. Changes in hearing-aid benefit following one, two or three year of hearing aid use by the elderly. J Speech Hear Res [In press].

ISO 389-1. Reference zero for the calibration of audiometric equipment – Part 1: Reference equivalent threshold sound pressure levels for pure tones and supra-aural earphones. Geneva: International Organization for Standardization, 1991.

ISO 8253-2. Acoustics – Audiometric test methods. Part 2: Sound field audiometry with pure tone and narrow-band test signals. Geneva: International Organization for Standardization, 1998.

Irvine DRF, Martin RL, Klimkeit E, Smith R. Specificity of perceptual learning in a frequency discrimination task. J Acoust Soc Am 2000; 108: 2964-8.

Jastreboff PJ. Phantom auditory perception [tinnitus]: Mechanisms of generation and perception. Neurosci Res 1990, 8: 221-54.

Kalikow DN, Stevens KN, Elliot LL. Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. J Acoust Soc Am 1977; 61: 1337-51.

Kaltenbach JA, Czaja JM, Kaplan CR. Changes in the tonotopic map of the dorsal cochlear nucleus following induction of cochlear lesions by exposure to intense sound. Hear Res 1992; 59: 213-23.

Karni A, Sagi D. Where practice makes perfect in texture discrimination: Evidence for primary visual cortex plasticity. Proc Nat Acad Sc, USA 1991; 88: 4966-70.

Kass JH. The reorganisation of sensory and motor maps in adult mammals. In: Gazzaniga MS [ed] The cognitive neurosciences. Cambridge, MA: MIT press 1995, 51-71.

Kiessling J, Steffens T. Comparison of a programmable 3-channel compression hearing system with single-channel AGC instruments. Scand Audiol 1993; Suppl. 38: 65-74.

Kilgard MP, Merzenich MM. Cortical map reorganisation enables by nucleus basalis activity. Science 1998; 279: 1714-8.

Killion M. The noise problem: There's hope. Hear Instr 1985; 36: 26-32.

Kraus N, McGee T, Carrell T, King C, Tremblay K, Nicol T. Central auditory system plasticity associated with speech discrimination training. J Cogn Neurosci 1995; 7: 25-32.

Kricos PB. The influence of nonaudiological variables on audiological rehabilitation outcomes. Ear Hear 2000; 21: 7S-14S.

Kuk FK. Acclimatization? Possible evidence seen in people with severe-to-profound hearing losses. Presentation at International Conference on Hearing Aid Research, Lake Tahoe, 2000.

Larson VD, Williams DW, Henderson WG, Luethke LE, Beck LB, Noffinger D, Wilson RH, Dobie RA, Haskell GB, Bratt GW, Shanks JE, Stemmoachowicz P, Studebaker GA, Boysen AE, Donahue A, Canalis R, Fausti SA, Rappaport BZ. Efficacy of three commonly used hearing aid circuits. J Am Med Ass 2000; 11: 1806-13.

Levitt H. Transformed up-down methods in psychoacoustics. J Acoust Soc Am 1971; 49: 467-77.

Lively SE, Logan JS, Pisoni DB. Training Japanese listeners to identify English /r/ and /l/. The role of phonetic environment and talker variability in learning new perceptual categories. J Acoust Soc Am 1993, 94: 1242-55.

Lockwood AH, Burkard RF, Salvi RJ Positron emission tomographic studies [PET] of gaze evked tinnitus. Association for Research in Otolaryngology, Feb 1999 [abstract]

Lockwood AH, Wack DS, Burkard RF, Coad ML, Reyes SA, Arnold SA, Salvi RJ. The functional anatomy of gaze-evoked tinnitus and sustained lateral gaze. Neurology 2001; 56: 472-80.

Lutman ME, Clark J. Speech identification under simulated hearing-aid frequency response characteristics in relation to sensitivity, frequency resolution and temporal resolution. J Acoust Soc Am 1986; 80: 1030-40.

Lutman ME, Payne E. Comparison of alternative compression strategies in new and experienced hearing-aid users. ISVR Contract Report No. 99/31. Institute of Sound and Vibration Research, University of Southampton 1999.

Malinoff RL, Weinstein BE. Changes in self-assessment of hearing handicap over the first year of hearing aid use by older adults. J Ac Rehab Audiol 1989; 22: 54-60.

Markides A. Binaural hearing aids. London: Academic Press 1977.

Martin FN, Champlin CA, Chambers JA. Seventh survey of audiometric practices in the United States. J Am Acad Audiol 1998; 9: 95-104.

McDermott HJ, Lech M, Kornblum MS, Irvine DRF. Loudness perception and frequency discrimnation in subjects with steeply sloping hearing loss; possible correlates of neural plasticity. J Acoust Soc Am 1998; 104: 2314-25.

McGarr NS. The intelligibility of deaf speech to experienced and inexperienced listeners. J Speech Hear res 1983; 26: 451-8.

Melzack R. Phantom limbs and the concept of a neuromatrix. Trends Neurosci 1990, 13: 88-92.

Merenich MM, Kass JH, Wall JT, Surr M, Nelson RJ, Felleman DJ. Progression of change following median nerve section in the cortical representation of the hand in areas 3b and 1 in adult owl and squirrel monkeys. Neurosci 1983, 10: 639-65.

Merzenich MM, Nelson RJ, Stryker MP, Cynader MS, Schoppman A, Zook JM. Somatosensory cortical map changes following digit amputation in adult monkeys. J Comp Neuro 1984, 224: 591-605.

Mogilner A, Grossman JAI, Ribory U, Joliet M, Volkmann J, Rapaport D, Bensley RW, Llinas RR. Somatosensory cortical plasticity in adult humans revealed my magnetoenchephalography. Proc Nat Acad Sci 1993, 90: 3593-97.

Moore BCJ. Cochlear hearing loss. London: Whurr, 1998.

Moore BCJ, Glasberg BR. Use of a loudness model for hearing-aid fitting. I. Linear hearing aids. Brit J Audiol 1998; 32: 317-35.

Most T, Weisel A, Lev-Matezky A. Speech intelligibility and the evaluation of personal qualities by experienced and inexperienced listeners. The Volta Review 1997; 98: 181-90.

Mulrow CD, Tuley MR, Aguilar C. Sustained benefits of hearing aids. J Speech Hear Res 1992: 35, 1402-5.

Munro KJ, Hatton N. Customized acoustic transform functions and their accuracy at predicting real-ear hearing aid performance. Ear Hear 2000; 21: 59-69.

Neuman AC. Balachandran R, Compton CL, Fine PS, Levitt H. Acclimatisation to hearing aids. Presented at: The second biennial hearing aid research and development conference, NIDCD and DVA, Bethesda, Maryland, 1997.

Nilsson M, Soli SD, Sullivan JA. Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. J Acoust Soc Am 1994; 95: 1085-99.

Noble W, Byrne D. A comparison of different hearing aid systems for sound localization in the horizontal and vertical planes. Brit J Audiol 1990; 24; 335-42.

Noble W, Byrne D. Auditory localization under conditions of unilateral fitting of different hearing aid systems. Brit J Audiol 1991; 25: 237-50.

Neuman AC. Late-onset auditory deprivation: a review of past research and an assessment of future research needs. Ear Hear 1996; 17: 3S-13S.

Olsen SO, Rasmussen AN, Nielson LH, Borgkvist BV. Loudness perception is influenced by long-term hearing aid use. Audiology 1999; 38: 202-5.

Olsen WO. Average speech levels and spectra in various speaking/listening conditions: A summary of the Pearsons, Bennett, & Fidell (1977) report. Am J Audiol 1998; 21: 21-5.

Ovegard A, Lundberg G, Hagerman B, Gabrielsson A, Bengtsson M, Brandstrom U. Sound quality judgement during acclimatization of hearing aid. Scan Audiol 1997; 26: 43-51.

Palmer CV, Nelson CT, Lindley GA. The functionally and physiologically plastic adult auditory system. J Acoust Soc Am 1998; 103: 1705-21.

Palmeri TJ, Goldinger SD, Pisoni DB. Episodic encoding of voice attributes and recognition memory for spoken words. J Exp Psychol: Learn Mem Cogn 1993, 19: 309-28.

Pavlovic CV, Studebaker GA, Sherbecoe RL. An articulation index based procedure for predicting the speech recognition performance of hearing-impaired individuals. J Acoust Soc Am 1986; 80: 50-7.

Paykel ES. Methodological aspects of life events research. J Psychosomatic Res 1983; 27: 341-52.

Pearsons KS, Bennett RL, Fidells S. Speech levels in various noise environments (Report No. EPA-600/1-77-025). Washington, DC: U.S. Environmental protection Agency, 1977.

Pfingst BE, O'Connor TA. Characteristics of neurons in auditory cortex of monkeys performing a simple auditory task. J Neurophysiol 1981; 45: 16-31.

Philibert B, Collet L, Veuillet E. Evidence of functional plasticity induced by auditory rehabilitation in binaurally fitted listeners. I: perceptive data. Presentation at Fifth European Federation of Audiology Societies, Bordeaux 2001a.

Philibert B, Collet L, Veuillet E. Evidence of functional plasticity induced by auditory rehabilitation in binaurally fitted listeners. II: Electrophysiological data. Presentation at Fifth European Federation of Audiology Societies, Bordeaux 2001b.

Quittner AL, Smith LB, Osberger MJ, Mitchell TV, Katz DB. The impact of audition on the development of visual attention. Psychol Sci 1994, 5: 347-53.

Recanzone GH, Merzenich MM, Dinse HR. Expansion of the cortical representation of a specific skin field in primary somatosensory cortex by intracortical microstimulation. Cerebral Cortex 1992, 2: 181-96.

Recanzone GH, Schreiner CE, Merzenich MM. Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. J Neurosci 1993; 13: 87-103.

Robertson D, Irvine DRF. Plasticity of frequency organisation in auditory cortex of guinea pigs with partial unilateral deafness. J Comp Neurol 1989; 282: 456-71.

Robinson DW. Long-term repeatability of the pure-tone hearing threshold and its relation to noise exposure. Brit J Audiol 1991; 25: 219-35.

Robinson K, Gatehouse S. Changes in intensity discrimination following monaural long-term use of a hearing aid. J Acoust Soc Am 1995; 97: 1183-90.

Robinson K, Gatehouse S. The time course of effects on intensity discrimination following monaural fitting of hearing aids. J Acoust Soc Am 1996; 99: 1255-8.

Robinson K, Summerfield AQ. Adult auditory learning and training. Ear Hear 1996; 17: 51S-65S.

Sachs MB, Abbas PJ. Rate versus level functions for auditory nerve fibres in cats: Tone burst stimuli. J Acoust Soc Am 1974; 56: 1835-47.

Sagi D, Tanne D. Perceptual learning: learning to see. Curr Opin Nuerobiol 1994, 4: 195-9.

Saunders GH, Cienkowski KM. Acclimatization to hearing aids. Ear Hear 1997; 18: 129-39.

Schreiner CE, Mendelson JR, Sutter ML. Functional topography of cat primary auditory cortex: distribution of integrated excitation. Exp Brain Res 1992; 92: 105-22.

Schum DJ. Test-retest reliability of a shortened version of the Hearing Aid Performance Inventory. J Am Acad Audiol 1993; 4: 18-21.

Schwaber MK, Garraghty PE, Kaas JH. Neuroplasticity of the adult primate auditory cortex following cochlear hearing loss. Am J Otol 1993; 14: 252-8.

Seewald RC, Ramji KV, Sinclair ST, Moodie KS, Jamieson DG. A computer-assisted implementation of the desired sensation level method for electroacoustic selection and fitting in children. User's manual. University of Ontario, 1993.

Shaw EAG, Vaillancourt MM. Transformation of sound-pressure level from the free field to the eardrum presented in numerical form. J Acoust Soc Am 1985, 78: 1120-3.

Shields PW, Campbell DR. Intelligibility, subjective ratings and completion time scores using the FAAF test with hearing-impaired subjects and noisy reverberant environments. Brit J Audiol 2001; 35: 237-45.

Silman S, Gelfand SA, Silverman CA. Late-onset auditory deprivation: Effects of monaural versus binaural hearing aids. J Acoust Soc Am 1984; 76: 1357-61.

Silman S, Silverman CA, Emmer MB, Gelfand SA. Effects of prolonged lack of amplification on speech-recognition performance: Preliminary findings. J Rehab Res Develop 1993; 30: 326-32.

Sireteanu R, Rettenbach R. Perceptual learning in visual search: fast, enduring but non-specific. Vis Res 1995, 35: 2037-43.

Spivak LG, Waltzman SB. Performance of cochlear implant patients as a function of time. J Speech Hear Res 1990, 33: 511-9.

Studebaker GA. A rationalized arcsine transform. J Speech Hear Res 1985; 28: 455-62.

Studebaker GA, Sherbecoe RL, McDaniel DM, Gwaltney CA. Monosyllabic word recognition at higher-than-normal speech and noise levels. J Acoust Soc Am 1999; 105: 2431-44.

Summerfield AQ, Marshall DH. Outcomes from adult implantation II: Performance. In: Cochlear implantation in the UK 1990-1994, London, HMSO 1996, 62-90.

Surr RK, Cord MT, Walden BE. Long-term versus short-term hearing aid benefit. J Am Acad Audiolo 1998; 9: 165-71.

Tait M, Lutman ME, Robinson K. Preimplant measures of preverbal communication behavior as predictors of cochlear implant outcomes in children. Ear Hear 2000, 21:18-24.

Taylor KS. Self-perceived and audiometric evaluations of hearing aid benefit in the elderly. Ear Hear 1993; 14: 390-4.

Theodoridis GC, Schoeny ZA. Procedural learning effects in speech perception tests. Audiology 1990; 29: 228-39.

Thompson G, Lassman F. Listener preference for selective vs flat amplification for a high-frequency hearing-loss population. J Speech Hear Res 1970; 13: 670-2.

Turner CW, Bentler RA. Does hearing aid benefit increase over time? Letter to editor. J Acoust Soc Am 1998; 104: 3673-4.

Turner CW, Humes LE, Bentler RA and Cox RM. A review of past research on changes in hearing aid benefit over time. Ear Hear 1996;17: 14S-28S.

Tyler RS, Summerfield AQ. Cochlear implantation: Relationships with research on auditory deprivation and acclimatization. Ear Hear 1996; 17: 38S-50S.

Tyler RS, Parkinson AJ, Woodworth GG, Lowder MW, Gantz BJ. Performance over time of adult patients using the Ineraid or Nucleus cochlear implant. J Acoust Soc Am 1997, 102: 508-22.

Vasama J, Makela JP. Auditory pathway plasticity in adult humans after unilateral idiopathic sudden sensorineural hearing loss. Hear Res 1995; 87: 132-40.

Ventry IM, Weinstein BE. The hearing aid inventory for adults: A new tool. Ear Hear 1982; 3: 128-34.

Walden BE. Toward a model clinical-trials protocol for substantiating hearing aid user-benefit claims. Am J Audiol 1997; 6: 13-24.

Watson C, Knudsen V. Selective amplification in hearing aids. J Acoust Soc Am 1940; 2: 406-19.

Weinberger NM. Learning induced changes of auditory receptive fields. Curr Opin Neurobiol. 1993, 3: 570-77.

Willot J. Changes in frequency representation in the auditory system of mice with age-related hearing impairment. Brain Res 1984; 309: 159-62.

Willott JF. Effects of ageing, hearing loss and anatomical location on thresholds of inferior colliculus neurons in C57BL/6 and CBA mice. J Neurophysiol 1986; 56: 391-408.

Willott JF, Parham K, Hunter KP. Response properties of inferior colliculus neurons in middle-age C57BL/6J mice with presbyacusis. Hear Res 1988; 37: 15-38.

Willott JF, Aitkin LM, McFadden SL. Plasticity of auditory cortex associated with sensorineural hearing loss in adult C57BL/6J mice. J Comp Neurol 1993; 329: 402-11.

Willott JF. Auditory system plasticity in the adult C57BL/6J mouse. In: Auditory system plasticity and regeneration. Ed RJ Salvi, D Henderson, F Fiorino, V Colletto. New York: Thieme Medical, 1996a, 297-316.

Willot JF. Physiological plasticity in the auditory system and its possible relevance to hearing aid use, deprivation effects, and acclimatisation. Ear Hear 1996b, 17: 38S-50S.

Wilmington D, Gray L, Jahrsdoerfer R. Binaural processing after corrected unilateral conductive hearing loss. Hear Res 1994; 74: 99-114.

Wood SA, Lutman ME. Evaluation of the relative benefits of typical NHS amplification and advanced digital devices. ISVR Contract Report No. 01/02. Institute of Sound and Vibration Research, University of Southampton 2001.

Wright BA, Buonomano DV, Mahnvcke HW, Merzenich MM. Learning and generalisation of auditory temporal-interval discrimination in humans. J Neurosci 1997; 17: 3956-63.

Wright BA, Fitzgerald MB. Different patterns of human discrimination learning for two interaural cues to sound-source localisation. Proc Natl Acad Sci 2001; 98: 12307-12.

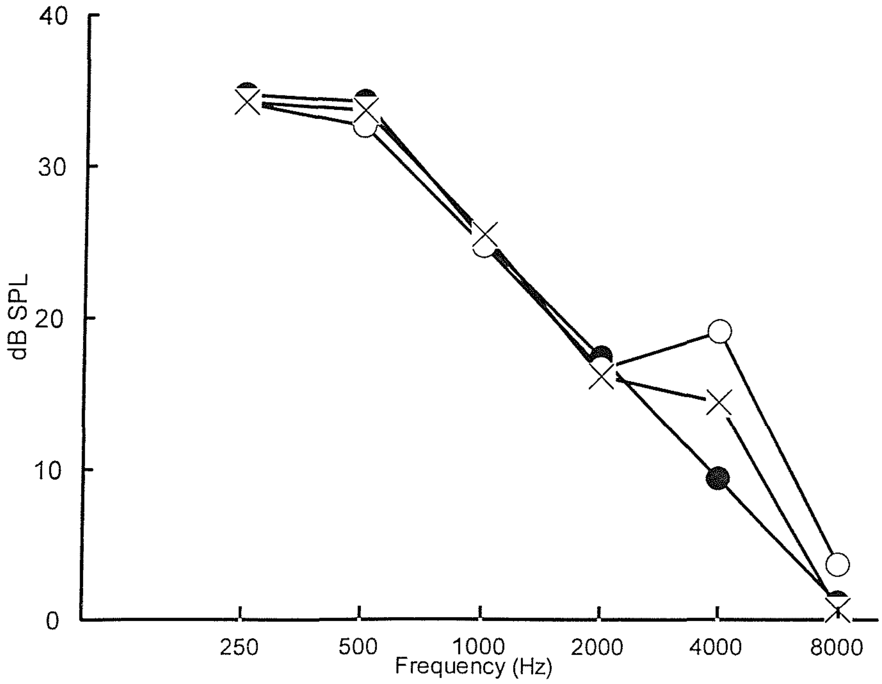
Yund EW, Buckles KM. Discrimination of multichannel compressed speech in noise: Long-term learning in hearing-impaired subjects. Ear Hear 1995; 16: 417-27.

APPENDIX A

The Four Alternative Auditory Feature [FAAF] test

Eighty key words presented in four alternative forced-choice format and delivered in the carrier phrase ‘Can you hear *XXXX* clearly?’

nan	*than*	*van*	*man*
some	*sud*	*sun*	*sub*
ridge	*rich*	*rids*	*ritz*
coast	*post*	*boast*	*ghost*
din	*pin*	*tin*	*bin*
rode	*rose*	*robe*	*rove*
milks	*milk*	*mick*	*mix*
cab	*gab*	*dab*	*tab*
wet	*bet*	*get*	*yet*
teen	*sheen*	*seen*	*keen*
mash	*match*	*mats*	*mass*
cop	*cot*	*cod*	*cob*
rib	*rick*	*rig*	*rip*
port	*fought*	*thought*	*taught*
lad	*lands*	*lads*	*land*
bag	*back*	*bat*	*bad*
dale	*nail*	*mail*	*bail*
veal	*seal*	*zeal*	*feel*
how	*high*	*ham*	*hang*
bang	*bag*	*bad*	*ban*



Octave band levels for conversational speech at 62.35 dB SPL. FAAF words, open circles; FAAF-shaped noise, crosses; conversational speech from ANSI S3.7 [1997], filled circles.

APPENDIX B

The BKB sentences

Sentence List 1
01 the CLOWN had a FUNNY FACE.
02 the CAR ENGINE's RUNNING.
03 SHE CUT with her KNIFE.
04 CHILDREN LIKE STRAWBERRIES.
05 the HOUSE had NINE ROOMS.
06 THEY're BUYING some BREAD.
07 the GREEN TOMATOES are SMALL.
08 HE PLAYED with his TRAIN.
09 the POSTMAN SHUT the GATE.
10 THEY're LOOKING AT the CLOCK.
11 the BAG BUMPS on the GROUND.
12 the BOY DID a HANDSTAND.
13 a CAT SITS ON the BED.
14 the LORRY CARRIED FRUIT.
15 the RAIN CAME DOWN.
16 the ICE CREAM was PINK.

Sentence List 2
01 the LADDER's NEAR the DOOR.
02 THEY had a LOVELY DAY.
03 the BALL WENT INTO the GOAL.
04 the OLD GLOVES are DIRTY.
05 HE CUT his FINGER.
06 the THIN DOG was HUNGRY.
07 the BOY KNEW the GAME.
08 SNOW FALLS at CHRISTMAS.
09 SHE's TAKING her COAT.
10 the POLICE CHASED the CAR.
11 a MOUSE RAN DOWN the HOLE.
12 the LADY's MAKING a TOY.
13 some STICKS were UNDER the TREE.
14 the LITTLE BABY SLEEPS.
15 THEY're WATCHING the TRAIN.
16 the SCHOOL FINISHED EARLY.

Sentence List 3
01 the GLASS BOWL BROKE.
02 the DOG PLAYED with a STICK.
03 the KETTLE's QUITE HOT.
04 the FARMER KEEPS a BULL.
05 THEY SAY some SILLY THINGS.
06 the LADY WORE a COAT.
07 the CHILDREN are WALKING HOME.
08 HE NEEDED his HOLIDAY.
09 the MILK CAME in a BOTTLE.
10 the MAN CLEANED his SHOES.
11 THEY ATE the LEMON JELLY.
12 the BOY's RUNNING AWAY.
13 FATHER LOOKED at the BOOK.
14 SHE DRINKS from her CUP.
15 the ROOM's GETTING COLD.
16 a GIRL KICKED the TABLE.

Sentence List 4
01 the WIFE HELPED her HUSBAND.
02 the MACHINE was QUITE NOISY.
03 the OLD MAN WORRIES.
04 a BOY RAN down the PATH.
05 the HOUSE had a NICE GARDEN.
06 SHE SPOKE TO her SON.
07 THEY're CROSSING the STREET.
08 LEMONS GROW on TREES.
09 HE FOUND his BROTHER.
10 some ANIMALS SLEEP ON STRAW.
11 the JAM JAR was FULL.
12 THEY're KNEELING DOWN.
13 the GIRL LOST her DOLL.
14 the COOK's MAKING a CAKE.
15 the CHILD GRABS the TOY.
16 the MUD STUCK on his SHOE.

Sentence List 5
01 the BATH TOWEL was WET.
02 the MATCHES LIE on the SHELF.
03 THEY're RUNNING PAST the HOUSE.
04 the TRAIN had a BAD CRASH.
05 the KITCHEN SINK's EMPTY.
06 a BOY FELL from the WINDOW.
07 SHE USED her SPOON.
08 the PARK's NEAR the ROAD.
09 the COOK CUT some ONIONS.
10 the DOG MADE an ANGRY NOISE.
11 HE's WASHING his FACE.
12 SOMEBODY TOOK the MONEY.
13 the LIGHT WENT OUT.
14 THEY WANTED some POTATOES.
15 the NAUGHTY GIRL's SHOUTING.
16 the COLD MILK's in a JUG.

Sentence List 6
01 the PAINT DRIPPED on the GROUND.
02 the MOTHER STIRS the TEA.
03 THEY LAUGHED at his STORY.
04 MEN WEAR LONG TROUSERS.
05 the SMALL BOY was ASLEEP.
06 the LADY GOES TO the SHOP.
07 the SUN MELTED the SNOW.
08 the FATHER's COMING HOME.
09 SHE had her POCKET MONEY.
10 the LORRY DROVE up the ROAD.
11 HE's BRINGING his RAINCOAT.
12 a SHARP KNIFE's DANGEROUS.
13 THEY TOOK some FOOD.
14 the CLEVER GIRLS are READING.
15 the BROOM STOOD in the CORNER.
16 the WOMAN TIDIED her HOUSE.

Sentence List 7

01 the CHILDREN DROPPED the BAG.
02 the DOG CAME BACK.
03 the FLOOR LOOKED CLEAN.
04 SHE FOUND her PURSE.
05 the FRUIT LIES on the GROUND.
06 MOTHER FETCHES a SAUCEPAN.
07 THEY WASHED in COLD WATER.
08 the YOUNG PEOPLE are DANCING.
09 the BUS WENT EARLY.
10 THEY had TWO EMPTY BOTTLES.
11 a BALL's BOUNCING ALONG.
12 the FATHER FORGOT the BREAD.
13 the GIRL has a PICTURE BOOK.
14 the ORANGE was QUITE SWEET.
15 HE's HOLDING his NOSE.
16 the NEW ROAD's on the MAP.

Sentence List 8

01 the BOY FORGOT his BOOK.
02 a FRIEND CAME for LUNCH.
03 the MATCH BOXES are EMPTY.
04 HE CLIMBED his LADDER.
05 the FAMILY BOUGHT a HOUSE.
06 the JUG STOOD on the SHELF.
07 the BALL BROKE the WINDOW.
08 THEY're SHOPPING for CHEESE.
09 the POND WATER's DIRTY.
10 THEY HEARD a FUNNY NOISE.
11 POLICE are CLEARING the ROAD.
12 the BUS STOPPED SUDDENLY.
13 SHE WRITES to her BROTHER.
14 the FOOTBALLER LOST a BOOT.
15 the THREE GIRLS are LISTENING.
16 the COAT LIES ON a CHAIR.

Sentence List 9

01 the BOOK TELLS a STORY.
02 the YOUNG BOY LEFT HOME.
03 THEY're CLIMBING the TREE.
04 SHE STOOD near her WINDOW.
05 the TABLE has THREE LEGS.
06 a LETTER FELL on the MAT.
07 the FIVE MEN are WORKING.
08 HE LISTENS TO his FATHER.
09 the SHOES were VERY DIRTY.
10 THEY WENT on HOLIDAY.
11 BABY BROKE his MUG.
12 the LADY PACKED her BAG.
13 the DINNER PLATE's HOT.
14 the TRAIN's MOVING FAST.
15 the CHILD DRANK some MILK.
16 the CAR HIT a WALL.

Sentence List 10

01 a TEA TOWEL's by the SINK.
02 the CLEANER USED a BROOM.
03 SHE LOOKED IN her MIRROR.
04 the GOOD BOY's HELPING.
05 THEY FOLLOWED the PATH.
06 the KITCHEN CLOCK was WRONG.
07 the DOG JUMPED ON the CHAIR.
08 SOMEONE's CROSSING the ROAD.
09 the POSTMAN BRINGS a LETTER.
10 THEY're CYCLING ALONG.
11 HE BROKE his LEG.
12 the MILK was by the FRONT DOOR.
13 the SHIRTS HANG in the CUPBOARD.
14 the GROUND was TOO HARD.
15 the BUCKETS HOLD WATER.
16 the CHICKEN LAID some EGGS.

Sentence List 11

01 the SWEET SHOP was EMPTY.
02 the DOGS GO for a WALK.
03 SHE's WASHING her DRESS.
04 the LADY STAYED for TEA.
05 the DRIVER WAITS by the CORNER.
06 THEY FINISHED the DINNER.
07 the POLICEMAN KNOWS the WAY.
08 the LITTLE GIRL was HAPPY.
09 HE WORE his YELLOW SHIRT.
10 THEY're COMING for CHRISTMAS.
11 the COW GAVE some MILK.
12 the BOY GOT INTO BED.
13 the TWO FARMERS are TALKING.
14 MOTHER PICKED some FLOWERS.
15 a FISH LAY on the PLATE.
16 the FATHER WRITES a LETTER.

Sentence List 12

01 the FOOD COST a LOT.
02 the GIRL's WASHING her HAIR.
03 the FRONT GARDEN was PRETTY.
04 HE LOST his HAT.
05 the TAPS are ABOVE the SINK.
06 FATHER PAID AT the GATE.
07 SHE's WAITING for her BUS.
08 the BREAD VAN's COMING.
09 THEY had some COLD MEAT.
10 the FOOTBALL GAME's OVER.
11 THEY CARRY some SHOPPING BAGS.
12 the CHILDREN HELP the MILKMAN.
13 the PICTURE CAME from a BOOK.
14 the RICE PUDDING was READY.
15 the BOY had a TOY DRAGON.
16 a TREE FELL on the HOUSE.

Sentence List 13

01 the FRUIT CAME in a BOX.
02 the HUSBAND BRINGS some FLOWERS.
03 THEY're PLAYING in the PARK.
04 SHE ARGUED with her SISTER.
05 a MAN TOLD the POLICE.
06 POTATOES GROW in the GROUND.
07 HE's CLEANING his CAR.
08 the MOUSE FOUND the CHEESE.
09 THEY WAITED for ONE HOUR.
10 the BIG DOG was DANGEROUS.
11 the STRAWBERRY JAM was SWEET.
12 the PLANT HANGS ABOVE the DOOR.
13 the CHILDREN are ALL EATING.
14 the BOY has BLACK HAIR.
15 the MOTHER HEARD her BABY.
16 the LORRY CLIMBED the HILL.

Sentence List 14

01 the ANGRY MAN SHOUTED.
02 the DOG SLEEPS in a BASKET.
03 THEY're DRINKING TEA.
04 MOTHER OPENS the DRAWER.
05 an OLD WOMAN was at HOME.
06 HE DROPPED his MONEY.
07 THEY BROKE ALL the EGGS.
08 the KITCHEN WINDOW was CLEAN.
09 the GIRL PLAYS with the BABY.
10 the BIG FISH GOT AWAY.
11 SHE's HELPING her FRIEND.
12 the CHILDREN WASHED the PLATES.
13 the POSTMAN COMES EARLY.
14 the SIGN SHOWED the WAY.
15 the GRASS is GETTING LONG.
16 the MATCH FELL on the FLOOR.

Sentence List 15

01 a MAN's TURNING the TAP.
02 the FIRE was VERY HOT.
03 HE's SUCKING his THUMB.
04 the SHOP CLOSED for LUNCH.
05 the DRIVER STARTS the ENGINE.
06 the BOY HURRIED to SCHOOL.
07 some NICE PEOPLE are COMING.
08 SHE BUMPED her HEAD.
09 THEY MET SOME FRIENDS.
10 FLOWERS GROW in the GARDEN.
11 the TINY BABY was PRETTY.
12 the DAUGHTER LAID the TABLE.
13 THEY WALKED ACROSS the GRASS.
14 the MOTHER TIED the STRING.
15 the TRAIN STOPS at the STATION.
16 the PUPPY PLAYS with a BALL.

Sentence List 16

01 the CHILDREN WAVE at the TRAIN.
02 MOTHER CUT the CHRISTMAS CAKE.
03 HE CLOSED his EYES.
04 the RAINCOAT's VERY WET.
05 a LADY BUYS some BUTTER.
06 THEY CALLED an AMBULANCE.
07 SHE's PAYING for her BREAD.
08 the POLICEMAN FOUND a DOG.
09 some MEN SHAVE in the MORNING.
10 the DRIVER LOST his WAY.
11 THEY STARED at the PICTURE.
12 the CAT DRANK from a SAUCER.
13 the OVEN DOOR was OPEN.
14 the CAR's GOING TOO FAST.
15 the SILLY BOY's HIDING.
16 the PAINTER USED a BRUSH.

Sentence List 17

01 the APPLE PIE's COOKING.
02 HE DRINKS from his MUG.
03 the SKY was VERY BLUE.
04 THEY KNOCKED on the WINDOW.
05 the BIG BOY KICKED the BALL.
06 PEOPLE are GOING HOME.
07 the BABY WANTS his BOTTLE.
08 the LADY SAT on her CHAIR.
09 THEY had some JAM PUDDING.
10 the SCISSORS are QUITE SHARP.
11 SHE's CALLING her DAUGHTER.
12 some BROWN LEAVES FELL off the TREE.
13 the MILKMAN CARRIED the CREAM.
14 a GIRL RAN ALONG.
15 the MOTHER READS a PAPER.
16 the DOG CHASED the CAT.

Sentence List 18

01 the CAKE SHOP's OPENING.
02 THEY LIKE ORANGE MARMALADE.
03 the MOTHER SHUT the WINDOW.
04 HE's SKATING WITH his FRIEND.
05 the CHEESE PIE was GOOD.
06 RAIN FALLS from CLOUDS.
07 SHE TALKED to her DOLL.
08 THEY PAINTED the WALL.
09 the TOWEL DROPPED on the FLOOR.
10 the DOG's EATING some MEAT.
11 a BOY BROKE the FENCE.
12 the YELLOW PEARS were LOVELY.
13 the POLICE HELP the DRIVER.
14 the SNOW LAY on the ROOF.
15 the LADY WASHED the SHIRT.
16 the CUP HANGS on a HOOK.

Sentence List 19

01 the FAMILY LIKE FISH.
02 SUGAR's VERY SWEET.
03 the BABY LAY on a RUG.
04 the WASHING MACHINE BROKE.
05 THEY're CLEARING the TABLE.
06 the CLEANER SWEPT the FLOOR.
07 a GROCER SELLS BUTTER.
08 the BATH WATER was WARM.
09 HE's REACHING for his SPOON.
10 SHE HURT her HAND.
11 the MILKMAN DRIVES a SMALL VAN.
12 the BOY SLIPPED ON the STAIRS.
13 THEY're STAYING for SUPPER.
14 the GIRL HELD a MIRROR.
15 the CUP STOOD on a SAUCER.
16 the COWS WENT to MARKET.

Sentence List 20

01 the BOY GOT into TROUBLE.
02 THEY're GOING OUT.
03 the FOOTBALL HIT the GOALPOST.
04 HE PAID his BILL.
05 the TEACLOTH's QUITE WET.
06 a CAT JUMPED OFF the FENCE.
07 the BABY has BLUE EYES.
08 THEY SAT on a WOODEN BENCH.
09 MOTHER MADE some CURTAINS.
10 the OVEN's TOO HOT.
11 the GIRL CAUGHT a COLD.
12 the RAINCOAT's HANGING UP.
13 SHE BRUSHED her HAIR.
14 the TWO CHILDREN are LAUGHING.
15 the MAN TIED his SCARF.
16 the FLOWER STANDS in a POT.

Sentence List 21

01 the PEPPER POT was EMPTY.
02 the DOG DRANK from a BOWL.
03 a GIRL CAME into the ROOM.
04 THEY're PUSHING an OLD CAR.
05 the CAT CAUGHT a MOUSE.
06 the ROAD GOES UP a HILL.
07 SHE MADE her BED.
08 BANANAS are YELLOW FRUIT.
09 the COW LIES on the GRASS.
10 the EGG CUPS are on the TABLE.
11 HE FRIGHTENED his SISTER.
12 the CRICKET TEAM's PLAYING.
13 the FATHER PICKED some PEARS.
14 the KETTLE BOILED QUICKLY.
15 the MAN's PAINTING a SIGN.
16 THEY LOST some MONEY.

APPENDIX C

Glasgow Hearing Aid Difference Profile

Does this situation happen in your life?
0 ___ No 1 ___ Yes

LISTENING TO THE TELEVISION WITH OTHER FAMILY OR FRIENDS WHEN THE VOLUME IS ADJUSTED TO SUIT OTHER PEOPLE

With your current hearing aid, how much difficulty do you have in this situation?	In this situation what proportion of the time do you wear your current hearing aid?	In this situation, with your new hearing aid, how much difficulty do you now have?	In this situation, what proportion of the time do you wear your new hearing aid?	In this situation, how much more does your new hearing aid help compared to your previous one?	For this situation, how much more satisfied are you with your new aid than with your previous one?
0__N/A 1__No difficulty 2__Only slight difficulty 3__Moderate difficulty 4__Great difficulty 5__Cannot manage at all	0__N/A 1__Never/Not at all 2__About ¼ of the time 3__About ½ of the time 4__About ¾ of the time 5__All the time	0__N/A 1__No difficulty 2__Only slight difficulty 3__Moderate difficulty 4__Great difficulty 5__Cannot manage at all	0__N/A 1__Never/Not at all 2__About ¼ of the time 3__About ½ of the time 4__About ¾ of the time 5__All the time	0__N/A 1__New aid much worse 2__New aid worse 3__New aid the same 4__New aid better 5__New aid much better	0__N/A 1__Much less satisfied 2__Less satisfied 3__Equally satisfied 4__More satisfied 5__Much more satisfied

Does this situation happen in your life?
0 ___ No 1 ___ Yes

HAVING A CONVERSATION WITH ONE OTHER PERSON WHEN THERE IS NO BACKGROUND NOISE

With your current hearing aid, how much difficulty do you have in this situation?	In this situation what proportion of the time do you wear your current hearing aid?	In this situation, with your new hearing aid, how much difficulty do you now have?	In this situation, what proportion of the time do you wear your new hearing aid?	In this situation, how much more does your new hearing aid help compared to your previous one?	For this situation, how much more satisfied are you with your new aid than with your previous one?
0__N/A 1__No difficulty 2__Only slight difficulty 3__Moderate difficulty 4__Great difficulty 5__Cannot manage at all	0__N/A 1__Never/Not at all 2__About ¼ of the time 3__About ½ of the time 4__About ¾ of the time 5__All the time	0__N/A 1__No difficulty 2__Only slight difficulty 3__Moderate difficulty 4__Great difficulty 5__Cannot manage at all	0__N/A 1__Never/Not at all 2__About ¼ of the time 3__About ½ of the time 4__About ¾ of the time 5__All the time	0__N/A 1__New aid much worse 2__New aid worse 3__New aid the same 4__New aid better 5__New aid much better	0__N/A 1__Much less satisfied 2__Less satisfied 3__Equally satisfied 4__More satisfied 5__Much more satisfied

Does this situation happen in your life?
0 ___No 1 ___ Yes

CARRYING ON A CONVERSATION IN A BUSY STREET OR SHOP

With your current hearing aid, how much difficulty do you have in this situation?	In this situation what proportion of the time do you wear your current hearing aid?	In this situation, with your new hearing aid, how much difficulty do you now have?	In this situation, what proportion of the time do you wear your new hearing aid?	In this situation, how much more does your new hearing aid help compared to your previous one?	For this situation, how much more satisfied are you with your new aid than with your previous one?
0__N/A 1__No difficulty 2__Only slight difficulty 3__Moderate difficulty 4__Great difficulty 5__Cannot manage at all	0__N/A 1__Never/Not at all 2__About ¼ of the time 3__About ½ of the time 4__About ¾ of the time 5__All the time	0__N/A 1__No difficulty 2__Only slight difficulty 3__Moderate difficulty 4__Great difficulty 5__Cannot manage at all	0__N/A 1__Never/Not at all 2__About ¼ of the time 3__About ½ of the time 4__About ¾ of the time 5__All the time	0__N/A 1__New aid much worse 2__New aid worse 3__New aid the same 4__New aid better 5__New aid much better	0__N/A 1__Much less satisfied 2__Less satisfied 3__Equally satisfied 4__More satisfied 5__Much more satisfied

Does this situation happen in your life?
0 ___ No 1 ___ Yes

HAVING A CONVERSATION WITH SEVERAL PEOPLE IN A GROUP

With your current hearing aid, how much difficulty do you have in this situation?	In this situation what proportion of the time do you wear your current hearing aid?	In this situation, with your new hearing aid, how much difficulty do you now have?	In this situation, what proportion of the time do you wear your new hearing aid?	In this situation, how much more does your new hearing aid help compared to your previous one?	For this situation, how much more satisfied are you with your new aid than with your previous one?
0__N/A 1__No difficulty 2__Only slight difficulty 3__Moderate difficulty 4__Great difficulty 5__Cannot manage at all	0__N/A 1__Never/Not at all 2__About ¼ of the time 3__About ½ of the time 4__About ¾ of the time 5__All the time	0__N/A 1__No difficulty 2__Only slight difficulty 3__Moderate difficulty 4__Great difficulty 5__Cannot manage at all	0__N/A 1__Never/Not at all 2__About ¼ of the time 3__About ½ of the time 4__About ¾ of the time 5__All the time	0__N/A 1__New aid much worse 2__New aid worse 3__New aid the same 4__New aid better 5__New aid much better	0__N/A 1__Much less satisfied 2__Less satisfied 3__Equally satisfied 4__More satisfied 5__Much more satisfied

We have dealt with some of the situations which in our experience can lead to difficulty with hearing. What we would now like you to do is to nominate up to four new situations in which it is important for you as an individual to be able to hear as well as possible.

With your current hearing aid, how much difficulty do you have in this situation?	In this situation what proportion of the time do you wear your current hearing aid?	In this situation, with your new hearing aid, how much difficulty do you now have?	In this situation, what proportion of the time do you wear your new hearing aid?	In this situation, how much more does your new hearing aid help compared to your previous one?	For this situation, how much more satisfied are you with your new hearing aid than with your previous one?
0_N/A 1_No difficulty 2_Only slight difficulty 3_Moderate difficulty 4_Great difficulty 5_Cannot manage at all	0_N/A 1_Never/Not at all 2_About ¼ of the time 3_About ½ of the time 4_About ¾ of the time 5_All the time	0_N/A 1_No difficulty 2_Only slight difficulty 3_Moderate difficulty 4_Great difficulty 5_Cannot manage at all	0_N/A 1_Never/Not at all 2_About ¼ of the time 3_About ½ of the time 4_About ¾ of the time 5_All the time	0_N/A 1_New aid much worse 2_New aid a little worse 3_New aid the same 4_New aid a little better 5_New aid much better	0_N/A 1_Much less satisfied 2_Less satisfied 3_Equally satisfied 4_More satisfied 5_Much more satisfied

With your current hearing aid, how much difficulty do you have in this situation?	In this situation what proportion of the time do you wear your current hearing aid?	In this situation, with your new hearing aid, how much difficulty do you now have?	In this situation, what proportion of the time do you wear your new hearing aid?	In this situation, how much more does your new hearing aid help compared to your previous one?	For this situation, how much more satisfied are you with your new hearing aid than with your previous one?
0_N/A 1_No difficulty 2_Only slight difficulty 3_Moderate difficulty 4_Great difficulty 5_Cannot manage at all	0_N/A 1_Never/Not at all 2_About ¼ of the time 3_About ½ of the time 4_About ¾ of the time 5_All the time	0_N/A 1_No difficulty 2_Only slight difficulty 3_Moderate difficulty 4_Great difficulty 5_Cannot manage at all	0_N/A 1_Never/Not at all 2_About ¼ of the time 3_About ½ of the time 4_About ¾ of the time 5_All the time	0_N/A 1_New aid much worse 2_New aid a little worse 3_New aid the same 4_New aid a little better 5_New aid much better	0_N/A 1_Much less satisfied 2_Less satisfied 3_Equally satisfied 4_More satisfied 5_Much more satisfied

With your current hearing aid, how much difficulty do you have in this situation?	In this situation what proportion of the time do you wear your current hearing aid?	In this situation, with your new hearing aid, how much difficulty do you now have?	In this situation, what proportion of the time do you wear your new hearing aid?	In this situation, how much more does your new hearing aid help compared to your previous one?	For this situation, how much more satisfied are you with your new hearing aid than with your previous one?
0_N/A 1_No difficulty 2_Only slight difficulty 3_Moderate difficulty 4_Great difficulty 5_Cannot manage at all	0_N/A 1_Never/Not at all 2_About ¼ of the time 3_About ½ of the time 4_About ¾ of the time 5_All the time	0_N/A 1_No difficulty 2_Only slight difficulty 3_Moderate difficulty 4_Great difficulty 5_Cannot manage at all	0_N/A 1_Never/Not at all 2_About ¼ of the time 3_About ½ of the time 4_About ¾ of the time 5_All the time	0_N/A 1_New aid much worse 2_New aid a little worse 3_New aid the same 4_New aid a little better 5_New aid much better	0_N/A 1_Much less satisfied 2_Less satisfied 3_Equally satisfied 4_More satisfied 5_Much more satisfied

With your current hearing aid, how much difficulty do you have in this situation?	In this situation what proportion of the time do you wear your current hearing aid?	In this situation, with your new hearing aid, how much difficulty do you now have?	In this situation, what proportion of the time do you wear your new hearing aid?	In this situation, how much more does your new hearing aid help compared to your previous one?	For this situation, how much more satisfied are you with your new hearing aid than with your previous one?
0_N/A 1_No difficulty 2_Only slight difficulty 3_Moderate difficulty 4_Great difficulty 5_Cannot manage at all	0_N/A 1_Never/Not at all 2_About ¼ of the time 3_About ½ of the time 4_About ¾ of the time 5_All the time	0_N/A 1_No difficulty 2_Only slight difficulty 3_Moderate difficulty 4_Great difficulty 5_Cannot manage at all	0_N/A 1_Never/Not at all 2_About ¼ of the time 3_About ½ of the time 4_About ¾ of the time 5_All the time	0_N/A 1_New aid much worse 2_New aid a little worse 3_New aid the same 4_New aid a little better 5_New aid much better	0_N/A 1_Much less satisfied 2_Less satisfied 3_Equally satisfied 4_More satisfied 5_Much more satisfied

APPENDIX D

RAW DATA

EXPERIMENT ONE: FAAF

FAAF

FITTED: Unaided

id	**0**	**3**	**6**	**9**	**12**	**15**	**18**	**21**	**24**
1	65	65.5	65.5	66	68.5	75	69.5	73.5	67.5
2	76	76	71	78.5	78	72	73.5	71	75
3	57	59	61	59.5	66	67	63	59	70
4	55.5 *	*	*		52	56	63	61	60.5
5	57.5	53.5	52	48.5	59.5	53.5	47.5 *		54.5
6	66.5	49.5	44.5	46	50.5	59	55.5	59.5	53
7	62	62	55.5	62	63.5	62.5	66.5	61	61
8	55	44.5	55	52	59.5	58.5 *		56	56
9	46.5	46.5	44 *		47.5	47.5	50	46.5	50.5
10	48.5	59	50	54.5	57.5	62.5	58 *		62
11	64.5	68	64.5	66	69.5	77	64	64.5	57.5
12	49	47	50	53.5	46	57.5	54 *		56
13	66	67.5	76	73.5	73.5	73.5	72	76	73.5
14	34.5	51.5	47	54.5	58	59	59.5	55	47.5
15	48	42.5	45.5	43.5	45.5	41.5	45	41.5	39.5
16	49	52	53	57.5	52	52.5	53	51	57.5
	56.3	**56.3**	**55.6**	**58.3**	**59.2**	**60.9**	**59.6**	**59.7**	**58.8**

FITTED: Aided fixed

id	**0**	**3**	**6**	**9**	**12**	**15**	**18**	**21**	**24**
1	66	67	66	60	60.5	67.5	71.5	78.5	76.5
2	73.5	74.5	69.5	72.5	72	75.5	70.5	68	72
3	72.5	74	68.5	69	69	73	74	66	68
4	73 *	*	*		74	72.5	75.5	72.5	75
5	67.5	71.5	70.5	66	74.5	74	79.5 *		76.5
6	70.5	78	78.5	82.5	76	80	82	78	78.5
7	71	72	67.5	70.5	71	71.5	66	70	70
8	65.5	66.5	70.5	67	63.5	71.5	67.5	66	78
9	70	71	74.5 *		77	70.5	69	64	71.5
10	71.5	75.5	75	79	82	81.5	84 *		82.5
11	70.5	72.5	72.5	78	80	86.5	83	81.5	83
12	69	68.5	67.5	75.5	77	75.5	62.5 *		68.5
13	68	69	75	73	75.5	74.5	78	75.5	73.5
14	71.5	71.5	75.5	71	70.5	69.5	77	79	78.5
15	73	66.5	67.5	68.5	66	69	70.5	73	69.5
16	72	71.5	73.5	73.5	73	73.5	82.5	79.5	82
	70.3	**71.3**	**71.5**	**71.9**	**72.6**	**74.1**	**74.6**	**73.2**	**75.2**

FITTED: Aided user

id	**0**	**3**	**6**	**9**	**12**	**15**	**18**	**21**	**24**
1	65	65.5	66	71	73.5	70	73	74.5	74.5
2	74	74.5	81	77.5	79.5	80	77	79.5	71
3	61	72	68.5	72.5	73	67.5	61	69.5	72.5
4	63.5 *	*	*		68.5	74.5	72	66	53.5
5	68.5	70	75.5	78	74.5	79	77 *		79.5
6	75.5	84.5	79.5	78	84	80.5	76.5	82	82.5
7	68	69	77	71	66.5	64.5	71	72	72

8	70	61	68	63	63.5	67	69.5	68.5	73
9	68	69	75 *		73	65.5	66	67.5	68.5
10	72	72	80	80.5	80.5	79.5	82 *		83
11	70	74.5	79.5	81.5	78.5	86	80	89.5	82.5
12	76.5	74	72.5	69.5	73	74	70.5 *		65.5
13	76	77	78.5	84	77	69.5	79	70.5	76.5
14	73	67.5	79.5	75	77	70	73	75.5	73.5
15	71	71.5	73.5	69.5	71.5	68	73	69	76
16	76	71.5	74.5		76	73.5	79.5	78	77
	70.5	**71.6**	**75.2**	**74.7**	**74.3**	**73.1**	**73.8**	**74.0**	**73.8**

FAAF

CONTROL: Unaided

id	0	3	6	9	12	15	18	21	24
1	61.5	63	70	63.5	53	72.5	66	67.5	67.5
2	65.5	67	70.5	63	70.5	61	68	68.5	69.5
3	70.5	72.5	72.5	73	73.5	77.5	75	69.5	68.5
4	48 *	*	*		44	45.5	58	47.5	46
5	57.5	54	57.5	59.5	51	56	58.5 *		57
6	58	61	62.5	66	65.5	67	62.5	66	67.5
7	47	45	55.5	50	58.5	62	50.5	59	51.5
8	71.5	75	72	68.5	68	68	79	80.5	70.5
9	52.5	54	52 *		51	54	56	53.5	57
10	54	57	44.5	53.5	59	53	53.5 *		62
11	50	57	62	51	61	77	53	62.5	56
12	42	50.5	51.5	52.5	54	67.5	64 *		37.5
13	63	62	66.5	62	62.5	64.5	63.5	66	61
14	61.5	55.5	60.5	58.5	62.5	54	53.5	64.5	58.5
15	40	46.5	49.5	45.5	44.5	48.5	45	39.5	35.5
16	55.5	55	53	56	55	57	50.5	63	47.5
	56.1	**58.3**	**60.0**	**58.8**	**58.3**	**61.6**	**59.8**	**62.1**	**57.1**

CONTROL: Aided

id	0	3	6	9	12	15	18	21	24
1	63	63.5	70.5	69.5	72	72	69	67.5	74.5
2	68.5	69	73	68	68.5	72.5	74.5	69.5	69.5
3	68	72.5	72.5	72	76.5	73.5	73.5	80	71.5
4	73.5 *	*	*		63.5	72	72.5	77	78
5	66	65.5	74.5	69.5	71.5	71	78 *		69
6	65.5	62	58.5	66	64	88.5	75.5	79	74.5
7	70.5	69	63		75	65	62	65.5	64.5
8	72	73.5	72.5	75	70.5	78	78.5	76.5	76
9	57.5	58	77 *		61	64.5	70	64	70.5
10	69.5	72.5	73	69.5	77	74.5	80 *		77
11	71	72.5	72.5	73	69.5	78.5	76.5	77	83
12	64	68.5	68.5	72	71.5	71.5	73.5 *		70.5
13	72	72	75	67	64.5	71.5	76.5	76	74
14	68	66.5	74	71	75.5	71	74.5	79	75.5
15	67	64	69.5	69.5	67	72	67.5	67	68
16	69	68	71.5	65.5	71.5	76.5	73	74.5	74.5
	67.8	**67.8**	**71.0**	**69.8**	**69.9**	**73.3**	**73.4**	**73.3**	**73.2**

EXPERIMENT ONE: BKB Sentences

BKB

FITTED: Unaided

id	0	6	12	18	24
1 *		70.5	77.3	74	90.5
2	62	69.5	62.7	85.5	69.5
3	42.5	41	22	25	32
4	40.5	35	45	54.5	41
5	50 *		72.5	67	87.5
6	56.5	66	72.5	61	61.5
7	34.5	31.5	42	64	59
8	59	44	73.5	71	76.5
9	39	53	48	43.5	62.5
10	38.7	42	72.5	51.5	58.5
11	10	7.5	9	21.5	15
12	36.5	25	37	65	60
13	35.5	35	41.3	28	17.5
14	45	43	52.5	43	51.5
15	49.5	53	50	44	53.5
	42.8	**44.0**	**51.9**	**53.2**	**55.7**
	13.0	**17.6**	**20.2**	**18.9**	**22.2**

CONTROL: Uniaded

id	0	6	12	18	24
1 *		62.5	83	72.5	59.5
2	40.5	34.5	20	21.5	21.5
3	47.5	75	39.5	32	60
4	38	53	44	52 *	
5	38 *		51.5	44.5	47.5
6	38	34.5	41.5	48.5	51
7	28.5	47.5	70	44	59
8	46	29.5	52	60.5	43
9	53	56	64	57.5	75.5
10	56.7	65.5	74.5	79	74.5
11	18.5	9	26.5	18	17.5
12	74	45	61.5	57.5	47.5
13	44	31	35	39.5	20
14	43.5	35	24.5	47.5	51
15	51	59.5	43.5	72	48.5
	44.1	**45.5**	**48.7**	**49.8**	**48.3**
	13.1	**17.7**	**18.9**	**17.7**	**18.2**

FITTED: Aided

id	0	6	12	18	24
1 *		69.5	97	85.5	96
2	85	59.5	75.3	86	82.5
3	73	68.5	72	60	69
4 .		64.5	80.5	82.5	80.5
5	72.5 *		86	90	85.5
6	76.5	90.5	95.5	99	98
7	74	72 *		88.5	83.5
8	47.3	65.5	90.5	80	58
9	64	74	77	87	85
10	71	92	95	97	93.5
11	69	50	72	67.5	63
12	64.5	80	81	76.5	82.5
13	68.5	60.5	58	61	44
14	74.5	69.5	59.5	51.5	65
15	69	81.5	83.5	77	77.5
	69.9	**71.3**	**80.2**	**79.3**	**77.6**
	8.7	**11.7**	**12.3**	**13.9**	**15.1**

CONTROL: Aided

id	0	6	12	18	24
1 *		71 *		85.5	77
2	74.5	43	46.7	73.5	71
3	33	62.7	69	68	52
4	54.5	62	79.5	76.5	69.5
5	64 *		67	80.5	80
6	68.5	72	75.5	77.5	83
7	62.5	60	78	71	87.5
8	77	64	86	63	73
9	73	79	84	84	94
10	62	74	86.5	90	90
11	69.5	62	77.5	81.5	84.5
12	68	55.5	72	77.5	70.5
13	72.5	63.5	61	60.5	47.5
14	63	56	60.5	47.5	58
15	68	77	73.5	89	77.5
	65.0	**64.4**	**72.6**	**75.0**	**74.3**
	11.0	**9.6**	**11.1**	**11.6**	**13.5**

EXPERIMENT ONE: Modified GHADP

	Disability	Handicap
id	**0**	**0**
1	70.8	83.3
2	41.67	0.67
3	60	55
4	43.75	31.25
5	30	25
6	55	70
7	57.14	60.71
8	46.4	37.5
9	55	50
10	56.3	43.75
11	37.5	50
12	50	41.7
13	37.5	37.5
14	40	35
15	68.75	87.5
16	50	40
	50.0	**42.7**

Residual Disability

id	**3**	**6**	**9**	**12**	**15**	**18**	**21**	**24**
1	31.25	10	10	6.25	37.5	41.7	25	25
2	8.33	0	0	0	0	0	0	6.25
3	0	16.67 .		12.5	8.33	0	0	1
4	31.25	50	33.33	37.5	41.7	31.3	33.3	33.3
5	25	25	31.25	25	31.25	31.25 *		18.75
6 *	*	*		45	50	43.75	25	25
7	15	20.83	15	15	10	15	16.67	12.5
8	37.5	37.5	8.33	8.33	16.67	0 *		12.5
9	0	0	16.67	8.33	8.33 *		16.67	8.33
10	25	16.7	8.3	18.75	25	16.7	18.8	20
11	0	0	0	0	0	0 *		0
12	16.67	8.33	0	0	0	0	0	0
13	25	31.25	18.75	25	31.25	20	25	12.5
14	15	15	10	15	10	0	5	5
15	16.67	0	6.25	0	18.8	25	12.5	18.75
16	0 *		0	6.25	0	0	50	50
.	**16.7**	**15.8**	**9.2**	**10.4**	**13.3**	**15.0**	**16.7**	**12.5**

Use

id	**3**	**6**	**9**	**12**	**15**	**18**	**21**	**24**
1	100	100	100	100	87.5	100	100	100
2	75	87.5	65	33.3	100	100	100	100
3	58.33	91.67 .		100	100	91.67	100	100
4	75	93.75	100	87.5	83.3	93.75	100	83.3
5	58.33	83.33	68.75	56.25	100	68.75 *		100
6 *&	*	*		65	81.25	100	100	66.67
7	75	91.67	95	90	90	90	91.67	75
8	87.5	62.5	91.67	66.67	58.33	58.33 *		62.5
9	91.67	87.5	83.33	83.33	75 *		66.67	75
10	66.67	100	100	100	100	100	100	100
11	100	93.75	87.5	87.5	93.75	93.75 *		93.25
12	91.67	100	100	100	100	100	100	100

13	100	100	100	100	100	100	100	100
14	100	100	100	100	100	100	100	80
15	100	100	87.5	100	87.5	91.67	93.75	87.5
16	60 *		75	68.75	68.8	75	62.5	56.25
	87.5	***93.8***	***93.3***	***88.8***	***91.9***	***93.8***	***100.0***	***90.4***

Additional Benefit

id	**3**	**6**	**9**	**12**	**15**	**18**	**21**	**24**
1	62.5	85	55	50	25	100	50	50
2 .		50	50	50	50	50	50	50
3	50	80 .		91.67	58.33	75	50	50
4	50	68.75	75	56.25	50	50	58.3	50
5	66.75	58.33	68.75	56.25	50	75 *		50
6 *	*	*	*		75	75	50	66.67
7	68.75	68.75	55	60	75	50	65	56.25
8	50	50	50	50	50	50 *		50
9	50	50	50	50	50 *		50	50
10	66.67	66.67	58.33	75	50	50	75	50
11	100	81.25	81.25	56.25	50	50 *		50
12	100	50	62.5	50	50	50	50	50
13	75	75	66.67	75	75	66.67	56.25	50
14	93.75	75	60	50	40	60	75	50
15	100	100	56.25	62.5	25	25	87.5	56.25
16	50 *		50	50	50	50	50	50
	66.7	**68.8**	**57.3**	**56.3**	**50.0**	**50.0**	**50.0**	**50.0**

Additional Satisfaction

id	**3**	**6**	**9**	**12**	**15**	**18**	**21**	**34**
1	75	90	50	50	25	37.5	50	25
2 .		50	50	50	50	50	50	50
3	50	83.33 .		91.67	75	75	50	50
4	50	56.25	58.33	56.25	50	50	50	50
5	25	58.33	68.75	50	50	75 *		50
6 *	*	*	*		56.25	62.5	50	66.67
7	45	56.25	50	55	70	50	85	62.5
8	50	62.5	50	50	50	50 *		50
9	50	50	50	50	25 *		50	50
10	83.33	58.33	50	50	50	50	50	50
11	100	100	75	56.25	50	50 *		50
12	75	50	50	75	50	50	50	50
13	75	58.33	75	75	75	58.33	62.5	50
14	75	50	70	50	55	50	75	50
15	83.33	75	50	56.25	43.85	25	75	75
16	65 *		50	50	50	50	50	50
	70.0	**58.3**	**50.0**	**50.0**	**50.0**	**50.0**	**50.0**	**50.0**

EXPERIMENT TWO: FAAF

FITTED EAR: Fixed gain

id	0	4	8	12	16	20	24 *
1	73	76.5	75.5	73	75	74	74.5
2	66	69	65	70.5	66.5	69.5	61
3	68.5	72.5	66.5	71.5	66.5	70.5	68.5
4	70	73	70	66.5	70	70	64.5
5	70	73	75	77	76.5	77	79
6	68.5	65.5 *		72	74	68	75.5
7	74	78	75	72.5 *		75	83.5
8	73.5	68.5	75	80.5	77.5	72.5	76
9	69.5	79	76	79.5	78	81.5	72
10	72.5	73	77	80	76.5	79	76.5
11	69.5	74.5	74.5	74.5	73	68.5	70
12	72	75.5	76	79	78	71.5	73
13	71.5	73.5	76	74.5	72.5	73	75
14	73	68	61.5	65.5	67	67.5	67.5
15	71	71.5 *		70	73.5	72	72
16	70	71 *		67.5	75	76.5	67.5
	70.8	**72.6**	**72.5**	**73.4**	**73.3**	**72.9**	**72.3**

FITTED EAR: User gian

id	0	4	8	12	16	20	24
1	73	69.5	73.5	73	75	77	75.5
2	67	67.5	66.5	68.5	59	74	67
3	71.5	67.5	62.5	61.5	72	63.5	58.5
4 *		69.5	65	66.5	72	72.5	62.5
5	69	73 *		74	79	76.5	74
6	70.5	69.5 *		72	72	74.5	78.5
7	74.5	62.5	79.5	76.5 *		77.5	78.5
8	70.5	75	75.5	77	73	75	77.5
9	69	76.5	71	76.5	74	78.5	72
10	72	71.5	77	76.5	76.5	79	79
11	70.5	74.5	73	72	75.5	68.5	70
12	70.5	73	78.5	80.5	82	78.5	79
13	73	66.5	83.5	76	76.5	73	75
14	71.5	68	72	67	67	67.5	67.5
15	72	71.5	73	73	74.5	75	73
16	78.5	77.5 *		74	75	77.5	72.5
8.5	**71.5**	**70.8**	**73.1**	**72.8**	**73.5**	**74.3**	**72.5**

FITTED EAR: Simulated fixed gain

id	0	4	8	12	16	20	24
1	71	75.5	73.5	67.5	75	74.5	77
2	66.5	65.5	66.5	67	64.5	66.5	67
3	66.5	62.5	64.5	63.5	67	72	58.5
4 *		71.5	71.5	66.5	64.5	70.5	64
5	73	74 *		76	78	80	75
6	68.5	65 *		73.5	68.5	69.5	72.5
7	76.5	79.5	73	72.5 *		79	77
8	70.5	78	69.5	82.5	79.5	74.5	77
9	74	75.5	70	80.5	73	83	72
10	79.5	73.5	77	80	76.5	79	78.5
11	68	74.5	77	76.5	73	68.5	70
12	69	78	77	80	78.5	73	82
13	73	77	76.5	67.5	76.5	73.5	75
14	68.5	68	67	69.5	67	67.5	67.5
15	76	83.5 *		76.5	75	74.5	75
16	75.5	75 *		75.5	78	73.5	66
	71.7	**73.5**	**71.9**	**73.4**	**73.0**	**73.7**	**72.1**

FITTED EAR: Simulated user gain

id	0	4	8	12	16	20	24
1	75	74.5	71.5	67.5	75	77.5	71
2	69.5	71	61.5	68.5	59.5	67.5	73
3	71	60.5	55	62.5	52.5	59	60.5
4 *		73	70	62.5	71.5	65.5	70
5	64.5	70 *		70.5	71	75.5	74.5
6	67	65.5 *		74	74	72	77.5
7	75.5	73.5	78	68.5 *		78.5	77.5
8	67	74	75	74.5	76.5	69.5	74.5
9	69	73	77	80	79	80	72
10	73.5	68	77	79	76.5	73.5	79.5
11	72.5	74.5	74	77	73.5	68.5	70
12	74.5	77.5	74	71.5	71.5	74.5	78
13	78.5	68	73.5	76.5	72	77	75
14	74	68	66	69.5	67	67.5	67.5
15	67.5	72 *		67.5	74.5	73	72
16	67.5	70 *		66.5	70.5	76.5	69.5
8.5	**71.1**	**70.8**	**71.0**	**71.0**	**71.0**	**72.2**	**72.6**

CONTROL EAR

id	0	4	8	12	16	20	24
1	69.5	70.5	72	69	68 *		73
2	69.5	72.5	68.5	61.5	63	65	68.5
3	69.5	70	69.5	70	69	74	67
4	70	71	72.5	72	66	74	69
5	71	71	68	68	69	75.5	69.5
6	70	75.5 *		73	71.5	69.5	70.5
7	75.5	70	65.5	76.5 *		72.5	71
8	71.5	72.5	59.5	66.5	67.5	71.5	77
9	67	71.5	73	65.5	72.5	73	71.5
10	70.5	66	62	67.5	68.5	65	69
11	74	68.5	67.5	70	71	72.5	76.5
12	68.5	70.5	69	69	72.5	76	73.5
13	71.5	67.5	64	69	69.5	68	74

14	71.5	69	64	68.5	73.5	71.5	68
15	67.5	70 *		69.5	70	65.5	68
16	73	66 *		68.5	79	72	73
	70.6	**70.1**	**67.3**	**69.0**	**70.0**	**71.0**	**71.2**

EXPERIMENT TWO: Modified GHADP

	Disability	**Handicap**
1	63	67
2	60	70
3	44	50
4	60	45
5	63	58
6	63	67
7	67	75
8	69	75
9	66	72
10	58	58
11	54	42
12	81	75
13	75	33
14	46	43
15	31	25
16	30	25
	61.5	**58.0**

Residual Disability

id	**4**	**8**	**12**	**16**	**20**	**24**
1	5.0	5.0	0.0 *		10.0	5.0
2	30.0	35.0	25.0	20.0	15.0	20.0
3	0.0	0.0	0.0	0.0	0.0	0.0
4	10.0	5.0	6.3	10.0	10.0 *	
5	20.0 *		0.0	0.0	0.0	5.0
6	0.0 *		7.5	0.0	25.0	16.7
7	45.0	33.3	40.0 *		60.0	50.0
8	50.0	44.0	25.0	31.3	31.3	31.3
9	11.0	8.3	10.8	21.4	28.6	28.6
10	12.5	25.0	29.2	29.0	25.0	33.3
11	20.0 *		30.0	25.0	25.0	25.0
12	18.8	12.5	31.3	31.3	12.5	12.5
13	0.0	63.0	0.0	0.0	0.0 *	
14	20.0	20.8	20.8	16.7	16.7	16.7
15	0.0 *		18.8	31.3	25.0 *	
16	10.0 *		12.5	6.3	12.5	16.7
	11.8	**20.8**	**15.6**	**18.3**	**15.8**	**16.7**

Use of hearing aid

id	**4**	**8**	**12**	**16**	**20**	**24**
1	90.0	95.0	100.0 *		90.0	95.0
2	65.0	85.0	80.0	80.0	80.0	85.0
3	100.0	100.0	100.0	100.0	100.0	100.0
4	75.0	90.0	100.0	45.0	75.0 *	
5	80.0 *		87.5	75.0	90.0	95.0
6	100.0 *		100.0	100.0	100.0	100.0
7	100.0	100.0	100.0 *		100.0	100.0
8	100.0	100.0	100.0	100.0	100.0	100.0
9	100.0	100.0	100.0	100.0	100.0	100.0
10	100.0	100.0	100.0	70.8	100.0	100.0
11	70.0 *		100.0	100.0	100.0	100.0
12	100.0	100.0	100.0	100.0	100.0	100.0
13	100.0	100.0	100.0	100.0	100.0 *	
14	37.5	62.5	87.5	62.5	62.5	62.5
15	50.0 *		68.8	43.8	56.3 *	
16	70.0 *		100.0	62.5	75.0	100.0
	95.0	**100.0**	**100.0**	**90.0**	**100.0**	**100.0**

Additional Benefit

id	**4**	**8**	**12**	**16**	**20**	**24**
1	95.0	100.0	100.0 *		100.0	100.0
2	75.0	100.0	85.0	50.0	90.0	90.0
3	92.0	100.0	100.0	100.0	100.0	100.0
4	70.0	80.0	68.8	50.0	50.0 *	
5	81.3 *		100.0	100.0	100.0	95.0
6	50.0 *		50.0	50.0	50.0	50.0

Additional Satisfaction

id	**4**	**8**	**12**	**16**	**20**	**24**
1	75.0	95.0	75.0 *		75.0	100.0
2	60.0	100.0	80.0	50.0	75.0	75.0
3	100.0	100.0	100.0	100.0	100.0	100.0
4	75.0	55.0	62.5	50.0	50.0 *	
5	68.8 *		93.8	100.0	100.0	95.0
6	50.0 *		50.0	50.0	50.0	50.0

7	85.0	75.0	60.0 *		75.0	75.0	7	60.0	60.0	55.0 *		50.0	50.0
8	58.0	56.0	56.3	87.5	87.5	87.5	8	58.0	50.0	62.5	62.5	87.5	81.3
9	70.0	50.0	50.0	50.0	50.0	50.0	9	50.0	50.0	50.0	50.0	50.0	50.0
10	87.5	58.3	66.7	75.0	100.0	62.5	10	70.8	66.7	58.3	75.0	75.0	100.0
11	70.0 *		75.0	75.0	75.0	75.0	11	60.0 *		75.0	75.0	75.0	75.0
12	100.0	100.0	100.0	87.5	100.0	100.0	12	87.5	100.0	100.0	100.0	100.0	100.0
13	83.0	50.0	87.5	100.0	100.0 *		13	75.0	50.0	87.5	100.0	100.0 *	
14	75.0	75.0	50.0	50.0	62.5	62.5	14	60.0	70.8	58.3	50.0	45.8	45.9
15	37.5 *		43.8	56.3	75.0 *		15	37.5 *		43.8	56.3	75.0 *	
16	66.7 *		68.8	50.0	56.3	50.0	16	50.0 *		56.3	50.0	50.0	50.0
	75.0	**75.0**	**68.8**	**65.6**	**81.3**	**75.0**	**8.5**	**60.0**	**66.7**	**62.5**	**59.4**	**75.0**	**75.0**

EXPERIMENT TWO: Paired Comparisons

clarity in quiet

	0	**4**	**8**	**12**	**16**	**20**	**24**
1	0.5	-0.2 *	*		0.5 *	*	
2	-1.5	0.5	-0.5	0.2	0.5	1.2	0.5
3	0	2.5	-0.2	-0.5	-0.2	-0.2	2.5
4	-0.5	0.5	0.5	0.5	0.5	0.5	0.5
5	-1.5	0.2 *		-0.5	0.2	-0.2	-1.5
6	0	0.5 *		0.2	0.5	1.2	0.2
7	0	-0.2	-0.5	0.5 *		-0.2	-0.2
8	-0.2	0.5	1.2	0.8	0.8	0.5	0.8
9	-0.8	2.5	0.2	0.5	-0.5	-0.2	-0.5
10	-1.5	-2.5	-2.5	-1.8	1.5	-0.8	1.8
11	0.2	0.5	0	0.5	0.5	1.5	0.5
12	-1.2	-1.5	1.5	1.2	-1.2	0.5	-0.5
13	-0.8	-0.5	-1.5	-0.2	-0.5	-0.2	-0.5
14	1.8	0.8	-0.5	-0.2	-0.5	-0.5	0.5
15	0.8	0.5 *		-1.2	-1.2	1.5 ***	
16	-1.5	0.5 *		1.2	1.5	0.5	-0.5
mean	**-0.4**	**0.3**	**-0.2**	**0.1**	**0.2**	**0.3**	**0.3**

clarity in noise

	0	**4**	**8**	**12**	**16**	**20**	**24**
1	0.2	-1.2 *	*		1.2 *	*	
2	-1.5	0	-0.5	-0.2	0.5	1.2	1.5
3	-0.2	1.8	-0.5	-0.5	0.2	0.2	-0.8
4	0.5	0.5	0.8	0.8	1.5	1.2	0.5
5	-1.5	-0.5 *		-0.2	0.2	-0.2	-0.5
6	-0.5	-0.2 *		-0.2	-0.2	0.2	-0.2
7	0.5	0.2	-0.2	0.5 *		0.5	0.2
8	0.5	0.8	1.5	0.5	0.8	0.5	-0.5
9	-0.2	-0.5	-0.2	0.5	-0.5	-0.5	-0.5
10	-0.5	0.8	-2.5	-1.5	0.8	-0.5	2.5
11	0.2	0.2	-2	1.2	0.5	-0.8	0.83
12	-1.5	0.2	1.8	-0.8	-1.2	1.5	2.2
13	0.8	0.5	-0.5	-1.5	1.2	-0.5	0.8
14	-0.2	0.8	-0.8	-0.5	0.8	1.5	0.5
15	0.6	0.8 *		-2.2	-0.2	1.2 ***	
16	1.5	2.2 *		1.5	1.2	1.5	0.5
mean	**-0.1**	**0.4**	**-0.3**	**-0.2**	**0.5**	**0.5**	**0.5**

clarity in babble

	0	4	8	12	16	20	24
1	1.5	1.5 *	*		0.8 *	*	
2	-1.5	1.2	-0.2	0.8	0.5	1.2	1.5
3	-0.5	2.5	0.2	0.2	-0.2	0.5	-0.8
4	0.5	0.2	0.8	0.5	0.8	1.2	0.5
5	-1.5	0.8 *		1.8	0.2	0.2	0.5
6	-0.5	-0.5 *		0.2	0.2	1.5	-0.5
7	0.2	0.5	0.8	0.8 *		0.5	0.2
8	0.2	0.8	0.8	0.8	0.2	0.5	0.2
9	-0.2	-1.5	0.2	-0.2	0	-0.5	-0.5
10	2.5	2.5	-2.5	0.2	2.5	0.8	2.5
11	0.5	0.8	0	1.5	1.5	1.5	0.5
12	1.5	0.8	2.5	0.8	-0.2	1.5	2.2
13	1.2	0.5	1	-0.8	1.5	0.8	0.5
14	1.5	1.8	-0.8	0.8	-0.5	1.2	0.5
15	1.5	1.2 *		-1	-0.2	0.5 ***	
16	1.5	1.5 *		1.5	0.8	0.5	0.5
mean	**0.5**	**0.9**	**0.3**	**0.5**	**0.5**	**0.8**	**0.6**

comfort in quiet

	0	4	8	12	16	20	24
1	-0.5	-1.2 *	*		-0.2 *	*	
2	-1.5	-0.8	-0.05	0.5	0.5	-0.5	-1.5
3	-0.5	0.5	-0.5	0.2	-0.2	-0.2	-2.5
4	0.2	0.5	-0.8	0.5	-0.8	-0.5	-0.5
5	-1.5	-1.5 *		-1.5	-0.5	0.2	-1.5
6	0	-0.5 *		-0.8	-0.5	-0.5	-0.8
7	-0.8	-0.5	-1.2	0.5 *		0.8	0.5
8	-1.5	-1.2	0.5	-0.8	0.8	-0.8	-0.5
9	-0.5	-0.5	0.2	-0.2	-0.5	-0.5	-0.5
10	-2.5	-2.5	-1.5	-2.5	-2.5	-2.5	-1.2
11	0.54	0.2	-0.5	-0.2	0.5	0.5	-0.5
12	2.5	1.8	-2.5	-1.2	0.5	0.5	1.2
13	1.2	-1.5	0.2	-0.5	-0.5	-1.2	-1.5
14	-0.5	-1.8	0.2	0.5	-0.5	0.5	-0.5
15	0.2	-0.5 *		-0.5	0.8	-1 ***	
16	1.5	0.2 *		-1.2	-1.5	-0.5	1.5
mean	**-0.2**	**-0.6**	**-0.5**	**-0.5**	**-0.3**	**-0.4**	**-0.6**

comfort in noise

	0	4	8	12	16	20	24
1	0	-1.5 *	*		-0.5 *	*	
2	-1.5	-0.2	-0.5	0.5	0.5	1.2	1.5
3	-0.5	0.5	-0.5	-0.5	0.5	0.2	-2.5
4	0.5	0.2	-0.5	1.2	0.5	1.2	-0.5
5	-1.5	-1.5 *		0.5	-0.2	0.2	-0.5
6	-0.5	0.5 *		1.5	-0.5	-1.5	-0.5
7	-0.5	-0.5	-0.5	0.5 *		-0.8	0.5
8	-1.5	-1.5	1.2	-1.2	0.8	-0.5	-0.5
9	-0.5	-0.5	0.2	-0.2	-0.5	-2.5	-0.5
10	-2.5	-2.5	-0.5	-2.5	-2.5	-0.5	-1.2
11	0.2	-0.8	-1.5	-0.2	0.5	-0.5	-0.5
12	-0.8	2.2	-2.5	-1.2	0.5	-0.5	2.2
13	0.5	0.5	-0.5	-0.8	0.5	-0.5	-0.5
14	-1.5	-2.2	-0.5	0.5	-0.5	0.5	-0.5

15	-0.8	-0.8 *		-1.5	0.2	-1.5 ***	
16	1.5	0 *		-1.5	-1.5	-0.5	1.5
mean	**-0.6**	**-0.5**	**-0.6**	**-0.3**	**-0.1**	**-0.4**	**-0.1**

comfort in babble

	0	**4**	**8**	**12**	**16**	**20**	**24**
1	0	-1.2 *	*		0.5 *	*	
2	-1.5	-0.8	-0.5	0.5	0.5	1.2	0.5
3	-0.5	0.5	-0.5	0.2	-0.5	0.5	-2.5
4	0.5	0.5	1.2	0.5	1.2	0.5	-0.5
5	-1.5	-1.5 *		1.8	-0.2	0.2	-0.5
6	-0.5	-0.5 *		-1.5	-0.5	-1.5	-0.8
7	0	-0.5	-0.5	0.5 *		1.5	0.5
8	-0.8	-0.8	0.8	-0.8	0.5	-0.5	-0.5
9	-0.5	-0.5	-0.2	-0.2	-0.5	-0.5	-0.5
10	-2.5	-2.5	-0.5	-2.5	-0.8	1.2	-1.5
11	0.5	0.2	-0.5	-0.5	-0.5	-0.5	-0.5
12	-2.2	-0.8	-1.8	-1.8	0.5	1.5	2.5
13	1.5	1.5	-0.5	-0.8	0.2	0.8	-1.8
14	-1.2	-1.5	-1.2	0.5	-0.5	0.5	-0.5
15	-0.5	-0.5 *		-0.8	0.5	-1.5 ***	
16	1.5	0 *		-1.2	-0.5	-0.5	-1.5
mean	**-0.5**	**-0.5**	**-0.4**	**-0.4**	**0.0**	**0.2**	**-0.5**

overall in quiet

	0	**4**	**8**	**12**	**16**	**20**	**24**
1 *		-0.5 *	*		-0.8 *	*	
2	-1.5	0.8	-0.2	0.2	0.5	-0.2	0.5
3	0	2.5	-0.5	-0.2	-0.2	0.2	-2.5
4	-0.2	0.5	0.5	0.5	0.2	0.5	0.5
5	-1.5	-1.5 *		-1.5	0.2	0.2	-0.5
6	0	0.5 *		-0.8	-0.2	-0.5	-0.2
7	-0.5	-0.5	-0.5	-0.5 *		1.5	0.5
8	-0.5	-0.5	-0.2	-0.8	0.5	-0.5	-0.5
9	-0.5	-0.5	-0.2	-0.5	-0.5	-0.5	-0.5
10	-2.5	-2.5	-2.2	-2.5	-0.8	-2.5	-2.5
11	0.5	0.2 *		1.2	0.5	1.2	0.5
12	-0.8	-0.2	-0.2	-1.2	0.5	0.5	-1.2
13	-0.8	-0.5	-1.5	-0.5	-0.2	-1.5	-0.5
14	0.2	-1.5	0.5	0.5	-0.5	0.5	-0.5
15	0.2	0.2 *		-1.5	0.8	0.5 ***	
16	1.5	-0.5 *		-0.8	-1.5	0.5	0.5
mean	**-0.4**	**-0.3**	**-0.5**	**-0.6**	**0.0**	**0.0**	**-0.5**

overall in noise

	0	**4**	**8**	**12**	**16**	**20**	**24**
1 *		-0.5 *	*		0.2 *	*	
2	-1.8	1.5	-0.2	0.2	0.5	0.5	0.2
3	0.2	2.5	-0.5	-0.2	-0.2	0.5	-2.5
4	0.5	0.5	0.2	0.5	1.2	0.8	-0.5
5	-1.5	-1.5 *		-1.5	-0.2	0.2	-0.5
6	0	-1.5 *		-1.2	-0.5	-2.5	-0.5
7	0.2	-0.2	-0.5	-0.2 *		1.5	0.5
8	-0.8	-0.5	-0.5	-1.2	-0.2	-0.8	-0.5
9	-0.5	-1.5	-0.2	-0.5	-0.5	-0.5	-0.5
10	-2.5	-2.5	-1.2	-1.8	-2.5	-1.8	-1.2

11	0.2	-1.5 *		0.5	1.5	1.5	-0.5
12	-0.5	1.8	-1.2	-2.5	0.2	2	1.5
13	0.8	-0.2	0	-1.5	-0.2	0.5	-0.2
14	-0.8	-0.8	0.2	0.5	-1	0.5	0.5
15	-0.8	-0.5 *		-1.5	-1.5	-0.5 ***	
16	2	0.5 *		-0.8	-1.5	0.5	0.5
	-0.4	**-0.3**	**-0.4**	**-0.7**	**-0.3**	**0.2**	**-0.3**

overall in babble

	0	**4**	**8**	**12**	**16**	**20**	**24**
1 *		-0.5 *	*		0.2 *	*	
2	-1.8	1.5	-0.2	0.5	0.5	0.2	0.5
3	0	2.5	0.2	-0.2	0.2	0.5	-0.8
4	0.5	0.5	1.2	0.5	1.2	0.8	0.5
5	-1.5	-1.5 *		1.5	0.2	0.5	0.5
6	0	-0.5 *		-0.2	-0.5	-2.5	-0.2
7	0.2	0.5	0.5	0.2 *		1.5	0.5
8	-0.8	0.2	-0.2	-0.5	0.5	-0.5	-0.5
9	-0.5	-0.5	0.2	-0.5	-0.5	-0.5	-0.5
10	-2.5	-2.5	-0.8	-1.8	-2.5	0.8	-1.8
11	0.2	0.2 *		0.2	1.5	-0.2	-0.5
12	1.2	1.8	-0.8	0.8	1.2	2.5	2.5
13	0.5	0.5	0.5	-1.5	0.2	0.2	1.5
14	-0.8	-0.2	0.2	0.5	0.5	0.5	0.5
15	-0.2	0.5 *		-0.8	-0.5	-0.5 ***	
16	2	-0.5 *		-0.5	-1.5	0.5	0.5
mean	**-0.2**	**0.1**	**0.1**	**-0.1**	**0.0**	**0.3**	**0.2**

EXPERIMENT THREE: FAAF

FITTED EAR: Unaided

	55 dB SPL			62 dB SPL			69 dB SPL		
id	**0**	**6**	**12**	**0**	**6**	**12**	**0**	**6**	**12**
1	60	52	52	71.5	73	70	73	75	74
2	44	45	41	65	56.5	56	71	72	58
3	50	47.5	51	60	64.5	61	68	69	64.5
4	64	65	66.5	67.5	69	71	68	66	65
5	56	69	57	69	72	64	68	68.5	66
6	51	38.5	44.5	58	60.5	51	61	62	59.5
7	52	53.5	45	66	63.5	60	72	67.5	73
8	50	51.5	54.5	57	54	58	68	69.5	66
9	50	39	43	52.5	52	54	71	72.5	71.5
10	36.5	29.5	39	63	59	63	77.5	65	75.5
11	56	59	43	60	73.5	66	72	71	74
12	66.5	59	65	73.5	74	73	77.5	66	75.5
13	53	67	63	70.5	69.5	69	75	71	75
14	53.5	59	65.5	71.5	71.5	72	67	70.5	69.5
15	67	69.5	56	70.5	68	70	75.5	75	61
16	51.5	59.5	58.5	68.5	75.5	76.5	77	82	77.5
	53.8	**54.0**	**52.8**	**65.3**	**66.0**	**64.7**	**71.3**	**70.2**	**69.1**

FITTED EAR: Aided

	55 dB SPL			62 dB SPL			69 dB SPL		
id	**0**	**6**	**12**	**0**	**6**	**12**	**0**	**6**	**12**
1	75	74	71.5	74.5	74	75	68.5	73.5	71
2	73	75.5	71	71	76.5	83	75	75.5	76
3	70	68.5	74	73	74.5	76	76	74.5	77
4	74	71	77	72	71	76	72	70	77
5	73	79.5	72.5	72	69.5	71.5	71	71	75.5
6	62.5	61.5	64	72.5	73	71.5	73.5	79	79.5
7	71	73	70	72	70	74	73.5	72.5	75
8	74.5	71.5	78	78	78.5	79	77	74	76
9	74.5	78	75	75	80	75	72	77	78
10	76.5	76.5	76.5	72.5	75.5	76.5	57.5	72	71.5
11	80.5	77	81.5	78	86.5	78.5	74.5	77.5	78.5
12	79.5	84	81	81.5	81	85	73	75	80
13	66	77.5	71	72.5	80	79	75	73.5	71
14	74.5	74	66	69.5	64.5	64.5	66.5	62.5	66
15	71	78.5	80	72.5	76.5	78	70.5	79	75
16	70	72.5	72.5	65	73.5	71	66	72	72.5
	72.8	**74.5**	**73.8**	**73.2**	**75.3**	**75.8**	**71.3**	**73.7**	**75.0**

CONTROL EAR: Unaided

	55 dB SPL			62 dB SPL			69 dB SPL		
id	**0**	**6**	**12**	**0**	**6**	**12**	**0**	**6**	**12**
1	30	29	30	58	53.5	47	73	73	76
2	46	49	40	61.5	57	55	69.5	64	71
3	57	64	71	73.5	71.5	71	79	77	80
4	65	63	67	66	61.5	65	69	61	69.5
5	58.5	52.5	51	66	67.5	63	64	56.5	64
6	48	55.5	48.5	60	53	53.5	63	58	59.5
7	53.5	52	49	61	62	64	77.5	71.5	73.5
8	64.5	62	63	71	73	68	80.5	80.5	81
9	63	55	56	52	57	60	71	67	69
10	43.5	40.5	44	62.5	66.5	70	75	80	75
11	54	58.5	56	67	74.5	68	69.5	81	72
12	63	58	66	75	74	73	71	68	71.5
13	66	67	66	67	72.5	69	76	71	65
14	57	61.5	64	72.5	71.5	60	70	72.5	71
15	62	56	50	75	69	56	71.5	76	66
16	69	69	67.5	72	71.5	71	62	70.5	68.5
8.5	**56.3**	**55.8**	**55.6**	**66.3**	**66.0**	**63.3**	**71.3**	**70.5**	**70.8**

CONTROL EAR: Aided

	55 dB SPL			62 dB SPL			69 dB SPL		
id	**0**	**6**	**12**	**0**	**6**	**12**	**0**	**6**	**12**
1	73.5	75	74	80	76	81	80	79	82.5
2	75.5	73	69	74	72.5	70	75	71	79
3	75	70	71	73	69	73	70	73	71
4	72	70	73	72	73	74	68	74.5	74
5	70	71.5	70	69	71	68	70.5	74	68
6	55	62	59	70	73	72	71	74.5	72.5
7	69.5	70.5	66	74	72	74.5	69	71	70
8	79	78	76	79	74	76.5	84	77.5	76
9	79.5	76.5	78	77.5	84	76	76	77.5	75
10	69	73.5	68	66	70.5	70	66.5	63	63
11	79	79.5	77.5	86	81.5	75.5	81.5	79	76
12	77.5	76	71	82.5	81	82	77	78	71
13	69	72	65	73	70.5	66	73	66	65
14	69	65	64	67	63	59	65.5	55	56
15	78.5	79	79	75	78	69	71	71	74
16	76.5	70	68	65	62.5	62	65	55	58.5
8.5	**73.0**	**72.6**	**70.5**	**73.9**	**73.2**	**71.8**	**72.7**	**71.2**	**70.7**

APPENDIX E

Summary of statistical analysis

Table I. Summary of the repeated-measures analysis of variance on the benefit scores at the fixed-gain setting. The two factors that were treated as repeated-measures were ear [test and control] and post-fitting time. [n=15]

Factor	df	F	Significance
Ear	1, 14	2.5	0.13
Time	8, 112	1.5	0.16
Ear*Time	8, 112	0.5	0.88

Table II. Summary of the repeated-measures analysis of variance on the benefit scores at the user-gain setting. The two factors that were treated as repeated-measures were ear [test and control] and post-fitting time. [n=15]

Factor	df	F	Significance
Ear	1, 14	5.6	0.03
Time	8, 112	1.9	0.07
Ear*Time	8, 112	1.6	0.14

Table III. Summary of paired *t*-tests between the benefit scores in the fitted and control ear for each post-fitting test session. The user-gain scores for the fitted ear were used in the comparison. The Bonferroni correction for nine comparisons results in a criterion for 5% significance overall of 0.006.

Post-fitting time [Weeks]	n	Mean difference	*t*	df	Significance
0	16	2.5	−0.8	15	0.46
3	15	5.8	−1.8	14	0.09
6	15	8.6	-3.2	14	0.01
9	12	6.5	-2.0	11	0.07
12	16	3.6	−1.2	15	0.23
15	16	0.4	−0.3	15	0.80
18	15	0.2	0.1	14	0.92
21	13	3.2	−1.0	12	0.35
24	16	1.1	0.3	15	0.76

Table IV. Summary of the repeated-measures analysis of variance on the unaided data. The two factors that were treated as repeated-measures were ear [fitted and control] and post-fitting time. One subject was excluded because they missed three consecutive test sessions. The data was smoothed [i.e., linear interpolation was used for four subjects who missed a single test session]. A total of 15 subjects were used in the analysis.

Factor	df	F	Significance
Ear	1, 14	0.4	0.55
Time	8, 112	4.2	0.01
Ear*Time	8, 112	1.1	0.38

Table V. Summary of the repeated-measures analysis of variance on the aided data. The two factors that were treated as repeated-measures were ear and post-fitting time. One subject was excluded because they missed three consecutive test sessions. The data was smoothed [i.e., linear interpolation was used for four subjects who missed a single test session]. A total of 15 subjects were used in the analysis.

Factor	df	F	Significance
	Aided at fixed-gain		
Ear	1, 14	3.8	0.07
Time	8, 112	7.0	0.01
Ear*Time	8, 112	0.7	0.68
	Aided at user-gain		
Ear	1, 14	6.2	0.03
Time	8, 112	6.3	0.01
Ear*Time	8, 112	2.1	0.04

Table VI. Summary of the repeated-measures analysis of variance on the simple main effects. One subject was excluded because they missed 3 consecutive test sessions. The data were smoothed [i.e., linear interpolation was used for 4 subjects who missed a single test session]. A total of 15 subjects were used in the analysis.

Factor	df	F	Significance
	Unaided		
Fitted ear	8, 112	2.5	0.02
Control ear	8, 112	2.9	0.01
	Aided at fixed-gain		
Fitted ear	8, 112	3.0	0.01
Control ear	8, 112	5.6	0.01
	Aided at user-gain		
Fitted ear	8, 112	3.0	0.01

Table VII. Summary of paired *t*-tests between the mean scores in the fitted and the control ear at each post-fitting test session. The user-gain results were used in the comparison. The Bonferroni correction for nine comparisons results in a significance value of 0.006.

Post-fitting time [Weeks]	n	Mean difference	t	df	Significance
0	16	2.7	1.8	15	0.10
3	15	3.8	2.0	14	0.07
6	15	4.2	2.3	14	0.04
9	13	4.6	2.2	12	0.05
12	16	4.4	2.4	15	0.03
15	16	0.2	0.2	15	0.88
18	16	0.3	0.2	15	0.83
21	13	0.7	0.3	12	0.74
24	16	0.7	0.3	15	0.75

Table VIII. Summary of the repeated-measures analysis of variance on the BKB benefit scores. The two factors that were treated as repeated-measures were ear and post-fitting time. Two subjects missed the first or last visit and their results were excluded from the analysis. Four subjects did not attend every test session so the data were smoothed using linear interpolation.

Factor	df	F	Significance
Ear	1, 12	0.2	0.76
Time	4, 48	0.2	0.94
Ear*time	4, 48	1.6	0.18

Table IX. Summary of the repeated-measures analysis of variance on the mean unaided BKB scores. The two factors that were treated as repeated-measures were ear and post-fitting time. Four subjects did not attend every test session so the data were smoothed using linear interpolation. However, three subjects missed the first or last visit and their results were excluded from the analysis.

Factor	df	F	Significance
Ear	1, 12	0.3	0.61
Time	4, 48	2.1	0.10
Ear*time	4, 48	0.8	0.51

Table X. Summary of the repeated-measures analysis of variance on the mean aided BKB scores. The two factors that were treated as repeated-measures were ear and post-fitting time. Four subjects did not attend every test session so the data were smoothed using linear interpolation. However, three subjects missed the first or last visit and their results were excluded from the analysis.

Factor	df	F	Significance
Ear	1, 12	2.5	0.14
Time	4, 48	3.7	0.01
Ear*time	4, 48	0.8	0.55

Table XI. Summary of the repeated-measures analysis of variance on the simple main effects of BKB aided scores over time. Thirteen subjects were used in the analysis of each ear.

Factor	df	F	Significance
Test ear	4,48	2.0	0.13
Control ear	4, 48	3.6	0.01

Table XII. Summary of the one-factor repeated measure analysis of variance on the post-fitting sub-scales of the Glasgow Hearing Aid Difference Profile. The four sub-scales were residual disability, use of hearing aid, additional benefit since previous visit and additional satisfaction since previous visit. One subject was excluded from the analysis because he did not complete the questionnaire on three successive occasions. Linear interpolation was used to smooth the remaining data [n=15]. With the exception of the additional satisfaction sub-scale, the sub-scales were statistically significant on Mauchly's test of sphericity. The degrees of freedom on these sub-scales were reduced for univariate tests using the Greenhouse-Geisser correction.

	df	F	Significance
Residual disability	2.8, 39.7	0.6	0.59
Use of hearing aid	2.7, 37.4	1.1	0.38
Additional benefit	3.6, 50.4	3.9	0.01
Additional satisfaction	7, 98	2.6	0.02

Table XIII. Summary of repeated-measures ANOVA. Post-fitting time was treated as a repeated measure and the procedure yields an orthogonal polynomial breakdown. The table contains the F value and associated significance level for the over-all effect of post-fitting time with and without feedback, together with the linear, quadratic and cubic components [n=16].

Factor	Linear	Quadratic	Cubic	Overall
No feedback*	F[1,15]=4.1; NS	F[1,15]=1.7; NS	F[1,15]=7.0; p=0.02	F[5,75]=2.5; p=0.04

* Not significant when week eight was removed [F[4,60]=1.3; p=0.27]

Table XIV. Summary of repeated-measures ANOVA on the mean aided FAAF recognition scores as a function of post-fitting time for different hearing instrument gain conditions. The two factors that were treated as repeated-measures were post-fitting time [7] and gain condition [3; fixed gain in each ear and also user gain in test ear]. The data were smoothed [i.e., linear interpolation was used for five subjects who missed a single test session]. The degrees of freedom have been adjusted using the Greenhouse-Geisser correction for the interactions since these were statistically significant on Mauchly's test of sphericity [n=16].

Factor	df	F	Significance
Time	6, 90	2.0	0.08
Gain condition	2, 30	7.7	<0.01
Time*Gain condition*	5.69, 180	1.8	0.11

* When sphericity was assumed the interaction was borderline significant [F[12,180]=1.8; p=0.05]

Table XV . Repeat of the previous analysis but with the not-fitted control condition removed from the analysis. The degrees of freedom have been adjusted using the Greenhouse-Geisser correction for the interactions since these were statistically significant on Mauchly's test of sphericity [n=16].

Factor	df	F	Significance
Time	3.2, 48.0	1.7	0.17
Gain condition	1, 15	0.2	0.64
Time*Gain condition	3.5, 51.9	1.4	0.24

Table XVI. Summary of repeated-measures ANOVA on the mean aided FAAF recognition scores as a function of post-fitting time for the fixed and simulated fixed-gain conditions. The two factors that were treated as repeated-measures were post-fitting time [7] and gain condition [2]. The data were smoothed [i.e., linear interpolation was used for five subjects who missed a single test session], n=16.

Factor	df	F	Significance
Time	6, 90	1.4	0.22
Gain condition	1,15	10.8	0.42
Time*Gain condition	6, 90	0.3	0.92

Table XVII. Summary of repeated-measures ANOVA on the mean aided FAAF recognition scores as a function of post-fitting time for the user and simulated user-gain conditions. The two factors that were treated as repeated-measures were post-fitting time [7] and gain condition [2]. The data were smoothed i.e., linear interpolation was used for five subjects who missed a single test session [n=16].

Factor	df	F	Significance
Time	6, 90	1.0	0.41
Gain condition	1, 15	5.8	0.03
Time*Gain condition	3.9, 57.8	1.4	0.26

Table XVIII. Summary of one-factor repeated-measures ANOVA on the post-fitting subscales of the Glasgow Hearing Aid Difference Profile. The four sub-scales were: residual disability, use of hearing aid, additional benefit and additional satisfaction. Linear interpolation was used to smooth the data [n=16]. All of the subscales were statistically significant on Mauchly's test of sphericity. The degrees of freedom on these subscales were reduced using the Greenhouse-Geisser correction.

	df	F	Significance
Residual Disability	2.27, 33.97	0.4	0.71
Use of Hearing Aid	2.65, 39.79	5.2	0.01
Additional Benefit	2.45, 36.67	0.9	0.48
Additional Satisfaction	2.53, 37.89	1.4	0.26

Table XIX. Summary of repeated-measures analysis of variance on the preference judgements for the two frequency responses. The two factors which were treated as repeated measures were background [quiet, noise and babble] and post-fitting time. The data from five subjects were smoothed using linear interpolation to account for a missing test session. The degrees of freedom have been adjusted using the Greenhouse-Geisser correction for the interactions since these were statistically significant on Mauchly's test of sphericity.

Factor	df	F	Significance
	Clarity		
Background	2, 28	5.9	0.01
Time	6, 84	1.5	0.19
Background*Time	5.1, 71.4	0.9	0.47
	Comfort		
Background	2, 28	0.5	0.59
Time	6, 84	0.7	0.68
Background*Time	5.6, 78.2	1.49	0.23
	Overall Preference		
Background	2, 28	5.8	0.01
Time	6, 84	1.4	0.24
Background*Time	5.3, 73.8	0.7	0.61

Table XX. Summary of repeated-measures analysis of variance on the benefit scores [n=16]. The three within-subject factors were time [3], ear [2] and SPL [3].

Factor	df	F	Significance
Time	2,30	0.9	0.42
Ear	1,15	1.9	0.18
SPL	2,30	31.5	<0.01
Time*Ear	2,30	6.4	<0.01
Time*SPL	4,60	1.1	0.37
Ear*SPL	2,30	0.8	0.46
Time*Ear*SPL	4,60	0.7	0.59

Table XXI. Summary of simple main effects when each ear was tested separately [n=16]. The two within-subject factors were time [3] and SPL [3].

Factor	df	F	Significance
	Fitted ear		
Time	2,30	4.1	0.03
SPL [a]	1.4,30	32.1	<0.01
Time*SPL	4,60	0.93	0.45
	Control ear		
Time	2,30	0.35	0.71
SPL [a]	1.3, 20.2	21.9	<0.01
Time*SPL	4,60	0.84	0.50

[a] *Mauchly test significant; degrees of freedom adjusted using Greenhouse-Geisser correction*

Table XXII. Summary of repeated-measures analysis of variance on the aided scores [n=16]. The three within-subject factors were time [3], ear [2] and SPL [3].

Factor	df	F	Significance
Time	2,30	0.6	0.55
Ear	1,15	6.0	0.03
SPL [a]	1.3,19.5	1.1	0.33
Time*Ear	2,30	15.5	<0.01
Time*SPL	4,60	1.3	0.30
Ear*SPL	2,30	0.0	0.99
Time*Ear*SPL	4,60	0.4	0.80

[a] *Mauchly test significant; degrees of freedom adjusted using Greenhouse-Geisser correction*

Table XXIII. Summary of simple main effects on the aided scores when each ear was tested separately [n=16]. The two within-subject factors were time [3] and SPL [3].

Factor	df	F	Significance
	Test ear		
Time	2,30	7.4	<0.01
SPL [a]	1.4,21.2	1.0	0.36
Time*SPL	4,60	1.1	0.38
	Control ear		
Time	2,30	4.9	0.02
SPL [a]	1.3,30	0.8	0.43
Time*SPL	4,60	0.5	0.75

[a] *Mauchly test significant; degrees of freedom adjusted using Greenhouse-Geisser correction*

Table XXIV. Summary of repeated-measures analysis of variance on the unaided scores [n=16]. The three within-subject factors were time [3], ear [2] and SPL [3].

Factor	df	F	Significance
Time	2,30	1.4	0.26
Ear	1,15	0.4	0.56
SPL	2,30	42.0	<0.01
Time*Ear	2,30	0.1	0.89
Time*SPL	4,60	0.6	0.67
Ear*SPL	2,30	0.8	0.45
Time*Ear*SPL	4,60	0.6	0.66

APPENDIX F

Speech Intelligibility Index [SII]

The articulation index [AI] was first proposed as a method to relate audibility to speech intelligibility by French and Steinberg [1947]. ANSI [1969] was concerned with methods for the calculation of AI. This has since been revised [ANSI, R1997] and is now known as the Speech Intelligibility Index [SII]. The SII is the proportion of the total speech information that is received by the listener's ear and is measured in the range of zero to one; when none of the speech signal is audible, the value is zero: when all of the speech signal is audible, the value is one. The calculation of SII is obtained by dividing the speech signal into several frequency bands, each weighted according to the theoretical contribution of that band to speech intelligibility. ANSI [R1997] outlines procedures for calculating SII with different frequency bands [critical band, ⅓ octave and octave bands].

The SII can be represented by the following equation:

$$\mathrm{SII} = \sum I_i \times A_i$$

The SII is the weighted sum of band audibility A_i. The weight, I_i, ranges from zero to one, and represents the relative importance of the frequency band to the understanding of speech. For recognising nonsense syllables, the band around 2 kHz is most important. For recognising sentences, the band around 0.5 kHz is most important. The frequency band importance function for the FAAF test has not been determined but is likely to be similar to that for nonsense syllables.

The Ai value is represented by the equation:

$$A_i = K_i \times L_i$$

K_i is the proportion of speech that is audible, and L_i is the level distortion factor [LDF]. Both K_i and L_i are given a value within the range zero to one. The speech area is assumed to have a dynamic range of 30 dB. When the speech signal is ≥ 30 dB above hearing threshold, the frequency band is making its maximal contribution to intelligibility and is give a value of one. When the speech signal is ≤30 dB below hearing threshold, the frequency band is making no contribution to the intelligibility and is given the value zero. L_i is a new addition to the SII calculation method compared to the formula in ANSI 1969. The LDF is based on the observation that speech performance deteriorates at high sound pressure levels in normally hearing subjects, presumably from distortions caused by, for example, upward spread of masking. The LDF allows for a reduced contribution to speech intelligibility from a maximum of one when the overall SPL exceeds 73 dB SPL. When a hearing-impaired subject is fitted with a hearing instrument, they receive speech at a level where even the normal-hearing subject has distortion. The LDF is used together with audibility to account for the degraded speech performance observed among hearing-impaired subjects.

For any particular type of speech material, the SII can be related to speech performance by a transfer function. This makes it possible to predict speech scores from audibility. Performance using sentences is higher than for nonsense syllables for a given SII and the transfer function is steeper [because of the redundancy in sentence material]. No transfer function has been published for the FAAF material although the low redundancy means that it is likely to be somewhat similar to the nonsense syllables transfer function.

A worked example of how to calculate the SII for a hearing impaired individual listening to conversational speech at an overall level of 62.35 dB SPL at a SNR of +3 dB is summarised in Table A. The SII of 0.61 was obtained by summing the values given in column [m].

Table A. Calculation of SII for a hypothetical subject using the octave band procedure

[a]	[b]	[c]	[d]	[e]	[f]	[g]	[h]	[j]	[j]	[k]	[l]	[m]
band	frequency	threshold	speech	noise	internal noise	disturbance	standard speech	L_i	K_i	A	I_i	SII
1	250	30	34.75	31.75	-3.90	31.75	34.75	1	0.6	0.6	0.0437	0.03622
2	500	30	34.27	31.27	-9.70	31.27	34.27	1	0.6	0.6	0.1294	0.07764
3	1000	40	25.01	22.01	-12.50	22.01	25.01	1	0.6	0.6	0.2025	0.1215
4	2000	50	17.32	14.32	-17.70	14.32	17.32	1	0.6	0.6	0.3117	0.18702
5	4000	60	9.33	6.33	-25.90	6.33	9.33	1	0.6	0.6	0.2576	0.15456
6	8000	70	1.13	-1.87	-7.10	-1.87	1.13	1	0.6	0.6	0.0551	0.03306

[a] frequency band number

[b] centre frequency of octave band [Hz]

[c] hearing threshold level [dB HL]

[d] speech spectrum level [dB SPL] with overall level of 62.35 dB

[e] masking noise spectrum level [dB SPL] for SNR of + 3 dB

[f] reference internal noise spectrum [dB SPL; from Table IV, ANSI R1997]

[g] equivalent disturbance level is the greater of column [e] and [f]

[h] standard speech spectrum with normal effort [dB SPL; from Table IV, ANSI R1997]

[i] level distortion factor which is calculated as 1 – ([c] - [h] / 160)

[j] the proportion of speech that is audible which is calculated as ([d] – [g] + 15) / 30

[k] the band audibility function is the product of [i] and [j].

[l] band importance function [for nonsense syllables] from Table IV, ANSI R1997.

[m] this is the product of columns [k] and [l]

Many studies have shown that the performance of hearing-impaired subjects is often poorer than normal hearing subjects for a given level of audibility. This reduction in performance can be included into the SII calculations by multiplying column [m] with a 'speech desensitisation factor' [SDF] such as the one reported by Pavlovic *et al.* [1986] and reproduced in numerical format in Table B.

Table B. The speech desensitisation factor [Pavlovic *et al.*, 1986]

Threshold [dB HL]	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95
SDF	0.936	0.873	0.809	0.746	0.682	0.619	0.555	0.492	0.428	0.365	0.301	0.238	0.174	0.111	0.047	0

References for Speech Intelligibility Index

ANSI [1969]. ANSI S3.5-1969. American National Standard methods for the calculation of the Speech Intelligibility Index. New York: American National Standards Institute, Inc.

ANSI [R1997]. ANSI S3.5-1997. American National Standard methods for the calculation of the Speech Intelligibility Index. New York: American National Standards Institute, Inc

French NR, Steinberg JC. Factors governing the intelligibility of speech sounds. J Acoust Soc Am 1947; 19: 90-119.

Pavlovic CV, Studebaker GA, Sherbecoe RL. An articulation index based procedure for predicting the speech recognition performance of hearing-impaired individuals. J Acoust Soc Am 1986; 80: 507.