Acclimatization to hearing aids

Abstract

Objective: Evidence for a clinically significant effect of acclimatization to hearing aids is mixed. The aim of this study was to test for auditory acclimatization effects in new unilateral and bilateral adult hearing aid users. Hypotheses were i) there would be improvements in aided speech recognition in new hearing aid users, compared to unaided listening and to a control group of experienced hearing aid users, and ii) improvements would correlate with severity of hearing loss, hearing aid use and cognitive capacity.

Design: Speech recognition in noise (SiN) was measured for a 65 and a 75 dB SPL target with the Four Alternative Auditory Feature (FAAF) test. SiN was measured within one week of fitting and retested at 12 weeks post-fitting in new hearing aid users (16 unilateral and 16 bilateral fit). A control group of experienced hearing aid users (n = 17) was tested over a similar time scale. Cognitive capacity (reaction time and working memory) was measured and self-reported change in performance was assessed using the Speech, Spatial and Qualities (SSQ) of Hearing Scale. Hearing aid use was assessed via data logging at the completion of the study.

Results: Mean improvements in speech recognition of up to 4% were observed across conditions and across groups consistent with a general practice effect. On average there was no evidence of auditory acclimatization in the new hearing aid user groups in terms of improvement in aided listening conditions above that observed in unaided recognition or in the control group. There was no correlation between change in aided speech recognition and severity of hearing loss, hearing aid use or cognitive capacity. New users reported
significant improvement over time in aided performance on a self-report questionnaire compared to the control group.

Conclusion: On average, there was no improvement over time in new users’ aided speech recognition relative to unaided recognition or to the control group. This does not support a robust acclimatization effect with non-linear hearing aids. Test-retest variability may obscure small average acclimatization effects; variability was not accounted for by individual differences in severity of hearing loss, hearing aid use or cognitive capacity. New users’ subjective report of increased benefit over time may be reflective of other aspects of adjustment to hearing aid use not examined in this study.
Introduction

Hearing aids selectively amplify acoustic input to compensate for losses of hearing sensitivity. With amplification, previously inaudible speech cues may become audible so that the hearing aid user may experience an immediate benefit from increased speech intelligibility. Performance may improve further over time as the user learns to make optimal use of these newly available acoustic cues. This further improvement, if not due merely to practice or retest effects, represents a form of perceptual learning. This is termed ‘auditory acclimatization’ (Arlinger et al. 1996), and will be referred to simply as ‘acclimatization’ from here on.

Scientific evidence for acclimatization associated with hearing aid use is mixed. In terms of improved speech perception, several studies have reported evidence for acclimatization (Cox et al. 1992; Cox et al. 1996; S. Gatehouse 1992, 1993; Horwitz et al. 1997; Kuk et al. 2003; Munro et al. 2003; Yund et al. 2006), while others have not (Bentler et al. 1993a, 1993b; Humes et al. 2003; Humes et al. 2002; Saunders et al. 1997; Taylor 1993). For a review of these and other studies, see Tuner et al. (1996), Palmer et al. (1998) and Munro (2008). Note that the studies to date have examined acclimatization to linear hearing aids; acclimatization to modern non-linear hearing aids has only been reported by a single study that we are aware of to date (Yund et al. 2006)\(^1\). For those studies that reported acclimatization effects, outcomes were highly variable with some participants showing large improvements over time and others not (Turner et al. 1996). One conclusion that followed from this is that if acclimatization does occur, it may be impacted by a variety of factors such

\(^1\) In his 1998 review of tests of speech recognition published in Scandinavian Audiology, Gatehouse makes reference to a study where he reportedly observed acclimatization to two hearing aids using different types of non-linear processing. However, this study was not described in any detail in this paper and it appears never to have been published.
that it is difficult to predict which users will show a reliable increase in benefit over time.

Furthermore, because of the difficulty in predicting acclimatization in individuals and typically small mean effects at group level, acclimatization effects – if they actually exist at all - were suggested to be clinically insignificant (Turner et al. 1998). Others argued that acclimatization effects do have relevance for research and clinical practice (Mueller et al. 2001; Munro 2008; Palmer et al. 1998), and that the reasons for the inconsistency of the research literature and for inter-individual variability in acclimatization may be ascribed to experimental design factors, unilateral versus bilateral fitting, the type and amount of auditory input, the severity of hearing loss, the amount of hearing aid use and cognitive factors. These are reviewed below.

With respect to experimental design factors, Turner et al (1996) and Palmer et al (1998) identified various aspects of experimental design that varied between acclimatization studies and may have impacted outcomes. Specifically, delays in the initial testing (for example, Taylor, 1993), the inclusion of participants with previous experience with hearing aids (for example, Bentler et al, 1993a), the manner of comparing with a control condition rather than with a separate control group (for example, Gatehouse, 1992; Munro & Lutman, 2003) or lack of control (for example Kuk et al. 2003; Reber and Kompis, 2005), inclusion of participants with a heterogeneous mix of signal processing or fitting schemes (for example, Saunders and Ceinkowski 1997) or relatively mild levels of hearing loss (Palmer et al. 1998) may have affected the outcome.

With regard to the type of auditory input, a key reason that has been suggested to explain the inconsistency of acclimatization in the research literature is that the effect may be limited specifically to those aspects of the stimulus which have been altered by the hearing
aid and which have not been typically experienced in daily life prior to amplification (Munro 2008). Supporting this idea, Gatehouse (1989) found that speech reception in the fitted ear of unilateral hearing aid users was better than in the non-fitted ear for high presentation levels, while it was worse in the fitted ear than in the non-fitted ear for low presentation levels. This suggested that acclimatization may involve selective adjustment to a dynamic range consistent with gain provided by hearing aids. Further indications of the specificity of acclimatization effects to novel inputs come from i) the observation of Gatehouse (1992) that speech recognition improved over the first few months of hearing aid use in new hearing aid users only in listening conditions that mimicked the pattern of amplification provided by the hearing aid, but not in listening conditions with an unfamiliar pattern of amplification or in unamplified listening conditions, and ii) the reports of Munro and Lutman (2003) of greater improvements for aided speech recognition of higher intensity speech stimuli over lower intensity stimuli. Additionally, in comparison to hearing aids, acclimatization effects do seem larger and more reliable with cochlear implants (Tyler et al. 1996), in which there is a more dramatic alteration in auditory input than is experienced with hearing aids. This would be consistent with the requirement for a significant alteration in auditory input in order to observe acclimatization. A necessary condition of acclimatization then is that input is novel and that to detect acclimatization effects, one should use assessments that tap frequencies and intensities that hearing aids make newly available. For example, Saunders and Cienkowski (1997) suggested that a possible reason for the lack of acclimatization effects in their study was that the test materials used were primarily sensitive to low- and mid-frequency information, and so acclimatization effects with high frequency stimuli may have been missed. Despite supportive evidence from the Gatehouse and Munro studies for the specificity of acclimatization effects to newly
experienced frequencies and intensities, not all data are consistent with this notion. Cox et al. (1996) measured aided speech recognition at the time of fitting and 12 weeks post-fitting in 22 unilaterally-fit new hearing aid users. Cox et al. hypothesized that acclimatization would be related to the amount of high frequency gain provided. Small but significant improvements in speech recognition were detected, but these did not correspond to the change in audibility of high frequency sounds. The additional hypothesis that improvements in speech recognition would be greatest for high-frequency speech sounds was also not supported.

With respect to unilateral versus bilateral fitting, there is a tendency for acclimatization effects to be detected in those that focused on unilateral fittings only (Munro 2008). The preponderance of positive findings for studies of unilateral fittings might suggest that acclimatization effects are a consequence of the perceptual asymmetry introduced by unilateral aiding. Some studies have included a mix of unilateral and bilateral users (for example, Cox and Alexander, 1992; Bentler et al. 1993a), although no direct comparison of acclimatization between unilateral and bilateral fitting has been reported to date.

Lastly, in the studies that have shown significant acclimatization effects on average, inter-individual variability in outcome is a notable feature, with some participants showing improvements over time while others do not (Turner et al. 1996). Individual differences between participants have been suggested as accounting for this variability, including the amount of hearing aid use, degree of hearing loss and cognitive capacity (Palmer et al. 1998; Tyler and Summerfield 1996). If acclimatization to hearing aids does depend on regular experience with novel auditory input, it seems necessary that this would require consistent hearing aid use. Most studies have reported the amount of use, although have not
examined use as a moderator of acclimatization. One exception is that of Bentler and colleagues (1993a), who reported no significant difference between ‘part-time’ (<4 hours per day) versus ‘full-time’ (>4 hours per day) self-reported use in improvements in aided speech recognition with a group of experienced hearing aid users who were re-fit with new hearing aids. Note that i) taking an arbitrary cut-point (e.g. 4 hours) to dichotomize the ‘use’ variable is likely to reduce statistical power to detect the impact of use and ii) self-reported hearing aid use may over-estimate actual use (Taubman et al. 1999).

In terms of the degree of hearing loss, Palmer et al (1998) suggested that a possible reason that some studies did not detect acclimatization was that they had included participants with relatively mild losses. Palmer et al suggested that only when high-frequency hearing loss exceeds 40-45 dB HL would there be a significant loss of input from conversational levels of speech, and therefore a greater likelihood of acclimatization following restoration in input. Individual differences in cognitive ability may also explain some inter-individual variability in acclimatization. Acclimatization may involve identification of speech cues within novel patterns of input, associating those cues with established representations of speech and language, and redefining stored representations according to novel patterns of input (Tyler and Summerfield 1996). In this conception, acclimatization is seen as being a cognitively demanding process, and so cognitive ability may be a limiting factor for acclimatization. For example, Ronnberg and colleagues recently proposed that working memory capacity may be a specific skill that mediates speech recognition, whereby working memory supports mapping of unfamiliar or degraded speech input onto established phonological representations (Ronnberg et al. 2008), as may be important for the acclimatization process as suggested by Tyler and Summerfield (1996).
Aims

This study presents two sets of analyses that address two main questions:

1. Existence of acclimatization effects

   • Is there evidence of an acclimatization effect for modern, non-linear digital hearing aids in terms of improved aided speech recognition over time, relative to unaided conditions and to a control group of long-term hearing aid users? If acclimatization effects do occur for intensities and frequencies that are significantly altered by hearing aids and have not typically been experienced in daily life, they should be most evident for higher level stimuli and in those who experience the greatest change in audibility with aiding (Gatehouse 1989, 1992; Munro and Lutman 2003).

   • Are acclimatization effects reflected in new hearing aid users’ self-report of improvement in real-life listening situations?

   • Does acclimatization differ between unilateral and bilateral hearing-aid fittings?

2. Inter-individual variability in acclimatization

   • Is the amount of acclimatization predictable from the severity of hearing loss, hearing aid use or cognitive capacity?

Methods

Participants

A total of 49 experienced and novice hearing aid users were recruited from four local audiology clinics. Ethics approval was obtained from the appropriate bodies and informed
written consent obtained from participants. Minimum sample size was estimated following
the magnitude and variance of changes in benefit over time as reported by Munro and
Lutman (2003), where the FAAF test was also used as the primary outcome measure. A
sample size of n = 15 would be required to provide 80% power to detect a mean change of
4% (SD 5%) using repeated measures t-test with an alpha level of 5% (two-tailed). Minimum
group size was exceeded for all three participant groups to allow for attrition and provide
some leeway in statistical power.

Palmer et al. (1998) suggested that acclimatization effects are more likely to be observed
when hearing loss exceeds 40-45 dB HL. Therefore, in the present study, the selection
criterion was symmetrical, mild-to-moderate, sloping high frequency sensorineural hearing
loss of at least 40 dB at 2−6 kHz. Additional inclusion criteria for new users were a history of
hearing loss of at least 1 year’s duration (based on self report) and no prior hearing aid use.
Additional inclusion criteria for experienced hearing aid users were a minimum of one years’
hearing aid use and self reported daily hearing aid use of at least 6 hours per day. The
assumption was that experienced users would have largely acclimatized to their hearing aids
after at least one years’ use and that any further improvements in aided listening would be
small in comparison to those in the new hearing aid user groups. Exclusion criteria were i)
fluctuating or recent changes in hearing level, ii) asymmetry in air conduction thresholds of
greater than 15 dB at two or more frequencies, iii) an air−bone gap greater than 15 dB at
any test frequency, iv) abnormal middle ear function assessed using oto−admittance
audiometry.

Thirty-two new hearing aid users were recruited, with 16 fit unilaterally and 16 fit
bilaterally. Allocation to unilateral or bilateral fitting was made as participants were
recruited to the study, alternating between unilateral and bilateral fitting. No participant expressed a strong preference for or against unilateral or bilateral aiding, and unilaterally fit participants were able to switch to bilateral aiding (and vice versa) on completion of the study if they wished. Seventeen experienced users were recruited, including 8 bilateral and 9 unilateral hearing aid users.

Mean age of each group was 70 years (SD = 10), 67 years (SD = 11) and 73 years (SD = 6) for the new unilateral, new bilateral and experienced user groups, respectively. Mean hearing loss across 2 to 6 kHz for each group was 51 dB HL (SD = 9), 50 dB HL (SD = 9) and 57 dB HL (SD = 13) for the new unilateral, new bilateral and experienced user groups, respectively. Hearing thresholds were retested on completion of the study. Mean change in threshold across frequencies and groups was 0.2 dB (SD = 5.4 dB). Overall 89% of hearing thresholds changed by 5 dB or less, and 98% changed by 10 dB or less. This is in line with reported test-retest differences for audiometry (Robinson 1991).

Amplification, fitting and electroacoustic measures

New users’ hearing aids were Starkey Radius behind-the-ear (BTE; n = 7) or Destiny completely-in-the-canal (CIC; n = 25) aids. BTE users received standard shell earmolds with a 0.8 mm vent. Both BTE and CIC aids had similar levels of technology. Both had an 8-channel compressor, with gain adjustable in twelve bands, and noise management was activated. Compression in these hearing aids operated with an attack time of 20 ms and release time of 2000 ms (per the ANSI S3.22 standard). Adaptive listening programs were disabled for the duration of the study. New users were fit according to National Acoustic

---

The reason for the mixture of BTE and CIC devices was that CIC devices became available to participants shortly after the commencement of this study. It was found that offering CIC aids as an option increased the participation rate, which is the reason for the greater number of CIC users over BTE users. The choice of BTE or CIC device was made by the participant.
Laboratories Non-linear 1 (NAL-NL1; Dillon et al. 1998) prescription targets based on audiometric thresholds between 0.25 and 8 kHz as measured at baseline, with fit verified with real ear insertion gain (REIG) measurement. REIG was carried out with a Siemens Unity hearing aid analyzer following recommended procedure (British Society of Audiology and the British Academy of Audiology 2008).

All newly fit participants were allowed at least one day’s trial use of the hearing aid, and gain adjusted if the participant felt they could not tolerate the level of gain as prescribed. No further gain adjustments were made during the study period, and participants were not able to adjust hearing aid gain themselves. Following any gain adjustment based on listener preference, REIG was measured. The mean prescription target and the mean user gain for 65 dB input signal level are shown for new users and experienced users in Figure 1. The mean target gain between 1 and 4 kHz (where hearing impairment was greatest) ranged between 10 to 18 dB. The mean user gain was up to 2 dB less than prescription target. Participants may have been conservative with gain settings because they were aware that gain could not be adjusted during the study period. For new hearing aid users, mean aided Speech Intelligibility Index\(^3\) (SII) (ANSI S3.5 1997) for sentences in quiet at 65 dB SPL was 0.79, compared to 0.52 unaided; the hearing aids significantly improved audibility of conversational-level speech.

Maximum output was determined with reference to loudness discomfort level (LDL) obtained during initial assessment and adjusted if the participant reported any undue discomfort to loud sounds presented during the hearing aid fitting. For unilateral fittings,

\(^3\) SII ranges from 0 to 1 and is calculated as the weighted sum of band audibility according to the relative importance of each frequency band to the understanding of speech. The SII calculation assumes that the speech area covers a dynamic range of 30 dB. SII was determined based on overall levels of 65 SPL for sentences in quiet. Aided values were obtained by adding the mean user insertion gain to speech levels.
ear of fitting alternated as participants were recruited to the study, with the exception of a single participant for whom work requirements necessitated fitting in the left ear. Thus, 9 participants were fit in the left ear, 7 in the right. Bilateral hearing aids were fitted at the beginning of the study for the unilateral users, though one hearing aid for each participant was retained in the lab and only used in the test sessions. Newly fit participants were encouraged to allow time to adjust to the input provided by the hearing aids, and to use their hearing aid at least 6 hours per day. Experienced users continued to use their existing hearing aids as usual, and no alterations in hearing aid gain or changes to adaptive listening programs were made during the course of the study. Experienced users’ hearing aids were Danalogic 6070 BTEs (n = 8) and Oticon Spirit 3 BTEs (n = 3), Oticon Synchro CICs (n = 2), Oticon Sprit 2 BTE (n =1), Oticon Sprit Zest BTE (n =2) and Siemens Reflex 1012 (n = 1); all non-linear hearing aids.

Real ear user gain was measured again at the end of the study in order to check stability of amplification over time. Mean difference in REIG ranged from 0.0 (SD = 1.8) to -3.0 (SD = 2.9) dB across frequencies; the mean change in REIG was not statistically significant, based on repeated measures ANOVA. Variation in gain is consistent with short-term test-retest REIG data (Hawkins et al. 1991).

In order for acclimatization to occur, users may require consistent exposure to altered patterns of auditory input via their hearing aids (Palmer et al. 1998). Most previous acclimatization studies relied on self-report of hearing aid use, such as a daily usage diary. However, such self-report measures are known to be unreliable (Taubman et al. 1999), and actual hearing aid use may have been over-estimated. In the present study, the mean number of hours per day of hearing aid use was recorded via hearing aid data logging
facilities, calculated and reported as part of the manufacturer’s hearing aid fitting program. In addition, the percentage of time that the hearing aid was worn and receiving input >50 dB SPL was also recorded to account for users who may wear hearing aids most of the day, but do not obtain significant auditory stimulation due to spending most of their time in relatively quiet home environments. The >50 dB SPL criterion was one of convenience; this is the cut-off point for the manufacturer’s data logging setting. Input below 50 dB SPL is characterized by the manufacturer’s data logging software as ‘quiet’. The assumption was that data logging information based on this 50 dB SPL criterion would provide a reasonable estimate of the proportion of time that hearing aids were providing auditory stimulation.

(Insert Figure 1 here)

Test materials

Four Alternative Auditory Feature test

The Four Alternative Auditory Feature (FAAF) test (Foster et al. 1979, 1987) is a closed-set word identification task, used in previous acclimatization studies (Gatehouse 1992, 1993; Munro and Lutman 2003). The FAAF consists of 20 sets of four binary and minimally paired words based on auditory/phonetic distinctions, resulting in an 80-item test. Either the initial or the final consonant varies within a set. Identification of the target largely depends on high-frequency auditory information, and this ensures that the FAAF is sensitive to putative newly available frequencies from amplification. In administration, the participant hears the phrase “Can you hear the word X clearly?” spoken by an English male speaker and then must select the target “X” from among four rhyming words (for example mail, bail, dale or nail) displayed on a computer monitor. Responses are scored as correct or incorrect, and an
overall percentage correct score is generated along with a breakdown of errors according to confusions of place, manner or voicing.

Target phrases were presented at two intensities; 65 dB SPL (roughly corresponding to ‘conversational’ speech levels) and 75 dB SPL (corresponding to ‘raised’ speech for male speakers (Pearsons et al. 1977)), following Munro and Lutman (2003) who reported greater mean acclimatization effects for more intense speech levels over less intense ones. Targets were presented against a background of steady noise with the same long-term spectrum as the target over the same loudspeaker. Speech and noise levels were calibrated in terms of the overall sound pressure level measured at the reference point; the approximate centre of the participant’s head with the participant absent.

At each test session (baseline and week 12) participants first completed a familiarization run with a target of 65 dB SPL and background noise set at 55 dB SPL. Each ear was then tested separately, the contralateral side being plugged and muffled (following Munro and Lutman’s (2003) methodology). At baseline, the signal to noise ratio (SNR) for 50% correct was determined for each participant and for each listening condition so that testing would occur at an equally challenging level for both aided and unaided listening and for high and low intensity targets, and so that ceiling and floor effects would be avoided. The SNR for 50% correct level on the FAAF was estimated for each participant and for each of the eight listening conditions (aided/unaided, left/right ear, 65/75 dB SPL stimulus intensity level) using an adaptive version of the FAAF. For adaptive testing, the initial noise level was set at 55 dB SPL (i.e. an SNR of +10 dB), with a 4 dB step size reducing to 2 dB steps for the last ten reversals. The final eight reversals were averaged to provide the SNR for 50% criterion performance for each listening condition. Criterion SNR for each listening condition was
then used for testing with the full, 80-item version of the FAAF at baseline and week 12.

Note that because SNR was set independently for each condition, it is not possible to calculate benefit by subtracting aided versus unaided performance, as done for some previous studies. Rather, changes in aided and unaided performance are examined separately. A different random order of testing was used for each test session.

Spatial, Speech and Qualities of Hearing Questionnaire – difference version (SSQ-D)

A questionnaire-based study of acclimatization by Munro and Lutman (2004) concluded that although new users tended to report a general increase in benefit over time, this depended on the time point that was used as a comparison. New users tended to rate performance as generally being better compared to the previous occasion of assessment, though not when compared to initial occasion of assessment. Munro and Lutman suggested that this was evidence of a response bias and recommended caution in interpretation of self-report measures of acclimatization. In order to minimize a retest bias effect such as reported by Munro and Lutman, self report was measured at week 12 only with participants asked to rate changes in performance with reference to baseline using the Spatial, Speech and Qualities of Hearing Scale – Difference version (SSQ-D). The SSQ-D is based on the SSQ (S. Gatehouse et al. 2004), modified by Stuart Gatehouse (W. Noble, personal communication 2009) so that participants are asked to report any change in a specific ability over a particular period of time. The version of the SSQ-D used in the present study was further abridged following Gatehouse and Noble (2004), so that four items from each dimension (Speech, Spatial and Qualities) with the highest correlations with independent measures of hearing handicap are included. Two additional questions from the SSQ that ask directly about listening effort were also included, so that the total number of items in the SSQ-D is
14. At week 12, all participants were given the SSQ-D and asked to report whether for each situation described they felt that there had been any positive or negative change in their ability during the period between baseline and week 12. Participant responses were indicated on a visual analogue scale ranging from -5 “Much Worse”, to zero “Unchanged”, to +5 “Much Better”. Scores for each item were summed to provide an overall index of change, with a possible range of scores between -70 to +70.

Cognitive tests

Two measures of cognitive ability were selected. First, speed of processing was tested, and this was intended as a general measure of cognitive integrity. Associations between speed of processing and other cognitive abilities are so strong that some investigators have concluded that age-related cognitive declines are fundamentally due to declines in processing speed (Salthouse 1996). Second, a recent review suggested that working memory capacity is the cognitive skill which is most consistently associated with speech recognition (Akeroyd 2008). Ronnberg and colleagues propose a model that identified working memory as a key resource for mapping unfamiliar or degraded speech input onto established phonological representations (Ronnberg et al. 2008). Working memory capacity was therefore tested and used as a predictor variable for acclimatization.

Speed-of-processing

Reaction times to visually presented digits were recorded, based on a paradigm used in a large study of cognitive aging (Deary et al. 2005). Use of a visual rather than an auditory reaction time test avoids any possible confound with peripheral hearing loss. Choice reaction times were recorded in response to the numbers 1 to 4 presented on a computer monitor, with participants required to press the corresponding key on a numeric keypad.
There were eight practice trials and 40 test trials. In the test trials, each digit appeared 10 times in randomized order. Correctness of the response and the reaction time in milliseconds were recorded for each test trial. The time interval between a response and presentation of the next digit varied randomly between one and three seconds. Participants were instructed to press the appropriate button ‘as quickly as possible’. Reaction times were averaged to provide an overall reaction time measure in milliseconds.

Working memory

The digits backwards subtest from the Weschler Adult Intelligence Scale-III (Wechsler 1997) was used to estimate working memory capacity. Participants were asked to repeat lists of aurally presented digits of increasing length in reverse order. Responses are scored correct if all the digits are recalled in the correct order. Raw scores were recorded with a maximum score of 14. Participants wore their hearing aids to complete this task in order to minimize the possible effect of hearing loss on performance.

Procedure

All participants completed otoscopy, pure tone audiometry, real ear insertion gain measures and the speech in noise task at baseline. For new users, the time between the day of initial fitting and baseline testing ranged between 0 to 7 days (M = 3, SD = 2 days). Otoscopy, pure tone audiometry, real ear insertion gain measures and the speech in noise task were then repeated after 12 weeks. A re-test interval of approximately 12 weeks was chosen following previous studies that suggested acclimatization effects would be evident by this time (for example Munro & Lutman, 2003). The exact retest interval varied due to the availability of participants for testing. The average time between baseline and week 12 was 93 days (SD =
18, range 70 to 110 days). The impact of time between baseline and week 12 test sessions was tested, and this is reported in the results section below.

Testing was carried out in a sound-treated audiometric booth with a reverberation time of 0.1 seconds across frequencies from 0.125 to 8 kHz, which is slightly more favorable than within an average domestic living room (around 0.3 seconds; Burgess et al. 1984). The participant was seated in the centre of the booth with a loud speaker positioned at ear level directly in front of the participant (at 0° azimuth) at a distance of 1.5 m from the reference test point, which was the approximate centre of the listener’s head. A computer monitor was placed in front of the participant below the level of the speaker to display response options. Stimulus presentation was computer controlled via an external 16-bit sound card at a sample rate of 20 kHz, routed to a Tannoy 607 loud speaker via a Denon amplifier (PMA-250III).

Results

The group mean SNR at baseline is displayed in Table 1. Mean SNR did not differ significantly between New Unilateral and New Bilateral groups. Apart from two conditions (Aided Right 65 dB SPL and Unaided Right 75 dB SPL), experienced users required statistically significantly more favorable mean SNRs compared to new users, based on one-way ANOVA with Fisher’s LSD post-hoc analysis. This is likely due to the older age and generally poorer hearing of the experienced user group.

Establishing the mean SNR for the FAAF test conditions enabled calculation of the SII for each participant in each listening condition. The calculations showed that for new users, amplification on average made less difference in audibility for the higher intensity speech
target than for the lower speech target. Aided SII values, obtained by adding the mean user
insertion gain for 65 dB and 75 dB SPL input levels to the speech and noise levels, were 0.79
aided versus 0.54 unaided for the 65 dB SPL target (a difference of 0.25) and 0.81 aided
versus 0.72 unaided for the 75 dB SPL target (a difference of 0.09).

Mean 80-item FAAF scores at individually set SNR for 50% correct tended to be closer to
60% correct rather than 50% across conditions and ranged between 53.8% and 61.2% across
groups, suggesting that some early practice effect had taken place. However, the aim was to
set SNR at a sufficiently challenging level for each participant and for each listening
condition, and this objective was met. Mean scores for each condition at baseline and week 12 are shown in Table 2.

Figure 2 shows change in FAAF performance for each condition and for each group,
measured by subtracting the baseline aided and unaided FAAF scores from week 12 aided
and unaided FAAF scores for each condition. Positive numbers represent an increase in
accuracy (percent correct). For all groups, small average improvements are evident across
most conditions, and this is consistent with a general practice effect. Variability within each
condition is large. For the new user groups, it was expected that acclimatization effects
would be observed in the form of improvements in the aided conditions (or for New Unilateral users, for the fitted side only) over that of unaided conditions and over that of the Experienced user group. For the New Unilateral group, there was a possible trend for
improvement for the fitted ear, though this was not specific to the aided listening
conditions; unaided performance increased similarly for the 65 dB SPL condition. For the
Bilateral group, there were small improvements on average across aided and unaided conditions for both ears. For the Experienced user group, there was a trend for small improvement in aided over unaided listening conditions in the 75 dB SPL condition.

(Figure 2 here)

The effect of i) Time (baseline/week 12), ii) Group (new unilateral/new bilateral/experienced users), iii) Aiding (aided/unaided), iv) Ear (for new unilateral users; aided/unaided, for new bilateral and experienced users; left/right) and v) Intensity (65/75 dB SPL) and their interaction on percent correct FAAF performance was tested statistically using mixed design ANOVA with Group as between-subjects variable (Table 3). The only significant effect was that of Time, which is consistent with a small positive improvement across groups and listening conditions. There was no significant main effect of Aiding or Intensity, and no significant interaction between Time, Group, Aiding, Ear or Intensity. Errors of recognition according to place, manner and voicing (Error Type) were analyzed similarly to overall accuracy. There was no significant effect of Time, or interaction between Time, Group, Aiding, Ear, Intensity or Error Type (results not shown here). The proportion of different types of errors did not change over time.

(Insert Table 3 here)

Variability in improvements in performance was notable for all three groups. Correlational analyses assessed whether 1) hearing aid use, 2) severity of hearing impairment, 3) cognitive capacity, and 4) degree of stimulus novelty explain the variation. Average daily hearing aid use (hearing aid hours) was 7.5 hours (SD 3.9), and ranged from 1 to 13 hours. Percentage of time that the hearing aid was switched on and receiving input >50 dB SPL was
termed ‘hearing aid active’. The mean for this variable was 63% (SD = 15), and ranged from 33% to 84%. Level of hearing loss (PTA) was indexed by the average PTA threshold across 0.250 kHz to 8 kHz for both ears (see Methods). The retest interval was taken as the number of days between the two test occasions (see the Procedure section). Mean Reaction time was 798 ms (SD = 172, range 532 to 1214), and was log transformed in order to normalize the distribution. Mean working memory score was 5.9 (SD = 1.8, range 3 to 10). Stimulus novelty was quantified as the difference between aided and unaided SII for the 65 dB SPL and 75 dB SPL test stimuli, as described earlier.

In order to provide maximum statistical power to test associations between predictor variables and change in speech perception over time, new unilateral and new bilateral users were pooled to form a single new user group (n=32). This assumes that acclimatization effects are qualitatively similar for unilateral and bilateral fits and that both are constrained in similar ways by the predictor variables. For the bilateral group, left ear performance was used. For new hearing aid users, there was no association between change in aided or unaided speech recognition and retest interval, level of hearing loss, cognitive factors (reaction time and working memory) or hearing aid use variables (amount of use or active use; |r| < 0.30, p > 0.05). Change in SII for a 75 dB SPL stimulus correlated with improvement in the 65 dB SPL Aided condition (r = 0.37, p < 0.05), but not the 75 dB SPL Aided condition (r = 0.20, p > 0.05). Change in SII for a 75 dB SPL stimulus also correlated with the 75 dB SPL Unaided condition (r = 0.37, p > 0.05). Such inconsistent correlations between change in SII and both Aided and Unaided speech recognition may be spurious. Older age was associated with slower reaction time (r = 0.47, p < 0.01) and higher proportion of time spent in quiet environment (hearing aid active; r = 0.26, p < 0.05).
Mean SSQ-D score was 26.2 (SD = 14.5), 20.5 (SD = 14.6) and -0.2 (SD = 11.7) for the New Unilateral, New Bilateral and Experienced user groups respectively, with positive scores associated with an improvement in performance over time. A similar pattern of improvement was seen for all SSQ-D subscales (i.e. spatial, speech, qualities and effort subscales; not reported here). Group SSQ-D scores were statistically significantly different from each other ($F(2,46) = 17.2$, $p = 0.00$). Post hoc analysis (Fishers LSD) revealed that the New Unilateral and New Bilateral user groups did not differ significantly from each other ($p = 0.65$), though the Experienced user group had a significantly lower score than both the new user groups ($p$'s < 0.00); on average both new user groups reported improved performance over time while the experienced user group reported no change. There were no significant correlations between SSQ-D score and change in aided performance on the FAAF for any listening condition (for the new user groups, $r$'s ranged from 0.00 to 0.10, all ns).

**Discussion**

1. **Existence of acclimatization effects**

   There was a small statistically significant improvement across conditions, consistent with a general practice effect. The significant practice effect detected in this study emphasizes the importance of the use of a control group; without such controls, improvements in performance may erroneously be interpreted as acclimatization effects. There was no significant interaction between improvement over time and group, stimulus intensity or ear of aiding. No differences in improvement over time were observed for unilateral versus
bilateral fitting. There was no change in the pattern of errors; given the particular impact of amplification on these phonemes, one might have expected a relative reduction in confusions of place-of-articulation compared to confusions of manner and voicing. There was no evidence of improvement consistent with acclimatization at group level.

There was also no evidence of particular improvements in speech recognition for aided conditions at the higher stimulus intensity level as would be consistent with the hypothesis of ‘stimulus novelty’. This hypothesis was advanced by studies of acclimatization that suggested acclimatization effects are specific to those intensities and frequencies altered by hearing aid amplification, and which had not been typically experienced in daily life prior to amplification (Cox and Alexander 1992; Gatehouse 1989, 1992; Munro and Lutman 2003).

Munro and Lutman (2003) argued that the acclimatization effects observed for high presentation levels in their study was the result of ‘novel’ levels of input. One possible reason for the discrepancy between Munro and Lutman’s (2003) study and the present one is that Munro and Lutman utilized linear aids, as did the majority of previous studies which found acclimatization effects. Linear aids might provide greater stimulus novelty than non-linear aids particularly for higher intensity stimuli because of higher levels of gain provided by linear aids for higher intensity input. For Munro and Lutman’s (2003) study, in the ‘raised’ (69 dB SPL) stimulus condition where acclimatization effects were most evident, the average change in SII (aided-unaided) was 0.22. This compares to a change of 0.25 for the 65 dB SPL stimulus and 0.09 for the 75 dB SPL stimulus in the current study using non-linear devices. The 65 dB SPL stimulus in the present study is closest to the 69 dB SPL stimuli in Munro and Lutman’s (2003) study, and there was a similar mean change in SII with aiding. Lack of ‘stimulus novelty’ does not therefore seem a likely explanation for explaining the lack of
acclimatization observed in the present study compared to Munro and Lutman’s (2003) study. In addition, in the present study stimulus novelty (change in SII with aiding) was inconsistently related to both aided and unaided speech recognition. This pattern of association does not support a reliable effect of stimulus novelty for acclimatization.

An additional possible explanation for the lack of acclimatization effects observed in the present study is that because baseline testing was not done on the day of testing but up to 7 days post-fitting, acclimatization effects may have been underestimated. However, acclimatization has previously been conceptualized as a process that is evident only after at least several weeks of use, and so one would have expected a post-fitting test interval of a matter of a few days not to have greatly impacted outcome.

New users did report large improvements in aided performance on average based on responses to a modified version of the SSQ (SSQ-D). These improvements were statistically significantly greater than those reported by the experienced users, who reported no change on average. There was no significant correlation between self-reported improvement based on the SSQ-D and improvements in aided speech perception based on the FAAF test. However, new hearing aid users appeared to have a subjective impression of improvement in aided performance over the first weeks of hearing aid use. This may reflect bias as, for example, users may have been comparing aided performance to pre-aided performance, rather than describing change in aided performance over time. Alternative possibilities are that new users’ impression of improvement may reflect an aspect of acclimatization not captured by the FAAF test, or to an aspect of adjustment to hearing aid use that is not related to acclimatization, such as increased confidence or familiarity with hearing aid use.

2. Inter-individual variability in acclimatization
On average, no statistically significant acclimatization effects were detected. However, between-participant variability may have obscured a small acclimatization effect. Variability of outcome is a noted feature of previous acclimatization research (Turner et al. 1996), and it has been suggested that individual differences between participants may explain this variability (Palmer et al. 1998; Turner et al. 1996; Tyler and Summerfield 1996). Retest interval, level of hearing loss, cognitive variables and hearing aid use variables were unrelated to change in both aided and unaided speech recognition performance. The variability in change in speech recognition performance observed in this study may thus be related to unknown determinants of acclimatization that were not measured in this study or may merely be a reflection of the test-retest variability of the FAAF test results. Over the 12 week duration of the present study, FAAF scores increased by 1% on average (SD 6%) across conditions, and correlations between baseline and week 12 scores varied between 0.5 and 0.8 (i.e. medium to large-sized correlations, all $p<0.001$) across conditions. These statistics suggest that even over a relatively long retest interval of 12 weeks, FAAF scores are reasonably reliable. Nonetheless, small acclimatization effects may be undetectable within this variability. Test-retest variability may be addressed by establishing the reliability of measures for older, hearing impaired adults over the relatively long time intervals required for acclimatization studies. Note that older adults typically exhibit more variable performance than young adults on a range of measures (MacDonald et al. 2009), so estimates of test-retest reliability based on younger participants’ performance may be unrealistically optimistic.

Conclusion
There was no evidence for acclimatization in the form of improvement in mean aided speech recognition. Test-retest variability may obscure small mean acclimatization effects. This variability was not accounted for by individual differences in the level of hearing loss, cognitive factors or hearing aid use variables. These findings did not support the existence of acclimatization effects for non-linear hearing aids that are large enough to be clinically significant. The absence of an effect in this study suggests that it may not be necessary to take acclimatization into account for speech-in-noise tests in clinical or research settings. Acclimatization effects may be more relevant for linear hearing aids or those with signal processing schemas that result in a more significant alteration in auditory input, such as frequency-transposing hearing aids. In the present study, new hearing aid users did report statistically significant improvement in aided listening compared to a control group of experienced users. This perceived improvement may relate to an aspect of acclimatization not measured in the present study, or to factors associated with beginning hearing aid use other than acclimatization, such as increased confidence or familiarity with hearing aids.

Acknowledgements

Thank you to audiologists Andrea Curran and Aneela Greval for hearing aid fitting and maintenance and Keith Wilbraham for technical assistance. This study was funded by the Starkey Hearing Research Centre.

References


Figure 1. Mean real ear user gain and prescribed gain targets for 65 dB input signal level. Error bars show standard deviation.
Figure 2. Difference in FAAF performance (baseline-week 12) by group. Error bars show standard deviations.
Table 2. Average FAAF score (percent correct) at baseline and week 12 for each participant group. Standard deviations are shown in brackets.

<table>
<thead>
<tr>
<th></th>
<th>Experienced Users</th>
<th></th>
<th></th>
<th>New Bilateral</th>
<th></th>
<th></th>
<th>New Unilateral</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
<td></td>
<td>Week 12</td>
<td>Baseline</td>
<td></td>
<td>Week 12</td>
<td>Baseline</td>
<td></td>
</tr>
<tr>
<td><strong>Unaided</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Right Ear</strong></td>
<td>65 dB</td>
<td>56.0 (7.2)</td>
<td>55.3 (8.3)</td>
<td>57.6 (5.0)</td>
<td>58.3 (6.0)</td>
<td>56.9 (5.0)</td>
<td>56.9 (9.5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>75 dB</td>
<td>56.8 (7.4)</td>
<td>56.6 (7.3)</td>
<td>56.9 (6.6)</td>
<td>58.5 (7.1)</td>
<td>56.9 (6.4)</td>
<td>57.1 (7.6)</td>
<td></td>
</tr>
<tr>
<td><strong>Left Ear</strong></td>
<td>65 dB</td>
<td>53.8 (7.7)</td>
<td>52.5 (10.1)</td>
<td>56.2 (7.4)</td>
<td>58.3 (5.5)</td>
<td>57.1 (7.3)</td>
<td>58.9 (6.6)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>75 dB</td>
<td>57.5 (7.3)</td>
<td>56.2 (8.8)</td>
<td>57.4 (7.7)</td>
<td>59.1 (6.2)</td>
<td>58.9 (7.0)</td>
<td>58.4 (8.3)</td>
<td></td>
</tr>
<tr>
<td><strong>Aided</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Right Ear</strong></td>
<td>65 dB</td>
<td>56.3 (8.3)</td>
<td>58.5 (9.3)</td>
<td>57.2 (7.3)</td>
<td>58.9 (8.8)</td>
<td>61.2 (6.5)</td>
<td>60.6 (8.1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>75 dB</td>
<td>58.8 (7.9)</td>
<td>60.3 (7.1)</td>
<td>53.8 (9.2)</td>
<td>57.4 (8.0)</td>
<td>56.7 (11.1)</td>
<td>57.3 (10.1)</td>
<td></td>
</tr>
<tr>
<td><strong>Left Ear</strong></td>
<td>65 dB</td>
<td>57.7 (7.4)</td>
<td>58.4 (7.8)</td>
<td>54.9 (9.6)</td>
<td>57.3 (7.5)</td>
<td>57.3 (8.7)</td>
<td>59.4 (7.8)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>75 dB</td>
<td>56.6 (6.5)</td>
<td>58.7 (5.6)</td>
<td>56.7 (4.9)</td>
<td>58.4 (6.2)</td>
<td>56.7 (7.5)</td>
<td>59.0 (7.0)</td>
<td></td>
</tr>
</tbody>
</table>
Table 3. Output statistics for mixed design ANOVA

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>Hypothesis df</th>
<th>Error df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>4.4</td>
<td>1</td>
<td>46</td>
<td>0.04</td>
</tr>
<tr>
<td>Time * Group</td>
<td>0.9</td>
<td>2</td>
<td>46</td>
<td>0.40</td>
</tr>
<tr>
<td>Ear</td>
<td>1.0</td>
<td>1</td>
<td>46</td>
<td>0.31</td>
</tr>
<tr>
<td>Ear * Group</td>
<td>1.5</td>
<td>2</td>
<td>46</td>
<td>0.22</td>
</tr>
<tr>
<td>Aid</td>
<td>0.1</td>
<td>1</td>
<td>46</td>
<td>0.69</td>
</tr>
<tr>
<td>Aid * Group</td>
<td>0.3</td>
<td>2</td>
<td>46</td>
<td>0.69</td>
</tr>
<tr>
<td>Level</td>
<td>0.1</td>
<td>1</td>
<td>46</td>
<td>0.81</td>
</tr>
<tr>
<td>Level * Group</td>
<td>0.9</td>
<td>2</td>
<td>46</td>
<td>0.40</td>
</tr>
<tr>
<td>Time * Ear</td>
<td>3.9</td>
<td>1</td>
<td>46</td>
<td>0.05</td>
</tr>
<tr>
<td>Time * Ear * Group</td>
<td>0.7</td>
<td>2</td>
<td>46</td>
<td>0.48</td>
</tr>
<tr>
<td>Time * Aid</td>
<td>0.2</td>
<td>1</td>
<td>46</td>
<td>0.67</td>
</tr>
<tr>
<td>Time * Aid * Group</td>
<td>1.1</td>
<td>2</td>
<td>46</td>
<td>0.33</td>
</tr>
<tr>
<td>Ear * Aid</td>
<td>0.3</td>
<td>1</td>
<td>46</td>
<td>0.58</td>
</tr>
<tr>
<td>Ear * Aid * Group</td>
<td>0.7</td>
<td>2</td>
<td>46</td>
<td>0.47</td>
</tr>
<tr>
<td>Time * Ear * Aid</td>
<td>0.0</td>
<td>1</td>
<td>46</td>
<td>0.82</td>
</tr>
<tr>
<td>Time * Ear * Aid * Group</td>
<td>0.5</td>
<td>2</td>
<td>46</td>
<td>0.57</td>
</tr>
<tr>
<td>Time * Level</td>
<td>0.2</td>
<td>1</td>
<td>46</td>
<td>0.66</td>
</tr>
<tr>
<td>Time * Level * Group</td>
<td>0.1</td>
<td>2</td>
<td>46</td>
<td>0.86</td>
</tr>
<tr>
<td>Ear * Level</td>
<td>1.5</td>
<td>1</td>
<td>46</td>
<td>0.22</td>
</tr>
<tr>
<td>Ear * Level * Group</td>
<td>0.1</td>
<td>2</td>
<td>46</td>
<td>0.86</td>
</tr>
<tr>
<td>Time * Ear * Level</td>
<td>0.6</td>
<td>1</td>
<td>46</td>
<td>0.42</td>
</tr>
<tr>
<td>Time * Ear * Level * Group</td>
<td>0.2</td>
<td>2</td>
<td>46</td>
<td>0.76</td>
</tr>
<tr>
<td>Aid * Level</td>
<td>1.5</td>
<td>1</td>
<td>46</td>
<td>0.23</td>
</tr>
<tr>
<td>Aid * Level * Group</td>
<td>0.4</td>
<td>2</td>
<td>46</td>
<td>0.70</td>
</tr>
<tr>
<td>Time * Aid * Level</td>
<td>0.5</td>
<td>1</td>
<td>46</td>
<td>0.46</td>
</tr>
<tr>
<td>Time * Aid * Level * Group</td>
<td>0.4</td>
<td>2</td>
<td>46</td>
<td>0.66</td>
</tr>
<tr>
<td>Ear * Aid * Level</td>
<td>0.0</td>
<td>1</td>
<td>46</td>
<td>0.98</td>
</tr>
<tr>
<td>Ear * Aid * Level * Group</td>
<td>1.4</td>
<td>2</td>
<td>46</td>
<td>0.25</td>
</tr>
<tr>
<td>Time * Ear * Aid * Level</td>
<td>0.3</td>
<td>1</td>
<td>46</td>
<td>0.61</td>
</tr>
<tr>
<td>Time * Ear * Aid * Level * Group</td>
<td>0.4</td>
<td>2</td>
<td>46</td>
<td>0.67</td>
</tr>
</tbody>
</table>