The effect of speech presentation level on measurement of auditory acclimatization to amplified speech[a]

Kevin J. Munro[b] and Mark E. Lutman
Institute of Sound and Vibration Research, University of Southampton, Southampton SO17 1BJ, United Kingdom

(Received 3 May 2002; revised 30 March 2003; accepted 7 April 2003)

A systematic improvement in auditory performance over time, following a change in the acoustic information available to the listener (that cannot be attributed to task, procedural or training effects) is known as auditory acclimatization. However, there is conflicting evidence concerning the existence of auditory acclimatization; some studies show an improvement in performance over time while other studies show no change. In an attempt to resolve this conflict, speech recognition abilities of 16 subjects with bilateral sensorineural hearing impairments were measured over a 12-week period following provision of a monaural hearing instrument for the first time. The not-fitted ear was used as the control. Three presentation levels were used representing quiet, normal, and raised speech. The results confirm the presence of acclimatization. In addition, the results show that acclimatization is evident at the higher presentation levels but not at the lowest. © *2003 Acoustical Society of America.* [DOI: 10.1121/1.1577556]

PACS numbers: 43.71.Ky, 43.71.Pc, 43.66.Ts [CWT]

I. INTRODUCTION

With the provision of amplification, the hearing-impaired 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 auditory neural representation. These changes may immediately confer greater intelligibility. However, provision of amplification may result in additional improvement in performance over time, possibly as a result of reorganization within the auditory cortex. This improvement over time is known as auditory acclimatization and has been defined by Arlinger *et al.* (1996) as ‘‘A systematic improvement in auditory performance over time, following a change in the acoustic information available to the listener. It involves an improvement in performance that cannot be attributed to task, procedural or training effects.’’ Evidence for auditory acclimatization after prolonged use of acoustic hearing aids is mixed; some authors state the acclimatization effects are small or nonexistent (Turner and Bentler, 1998). For a review, see the articles by Turner *et al.* (1996) and Palmer *et al.* (1998).

The most frequently cited acoustic hearing instrument studies providing evidence of auditory acclimatization include Cox and Alexander (1992), Cox *et al.* (1996), Gatehouse (1992, 1993), and Horwitz and Turner (1997). Cox and Alexander (1992) studied eight new hearing instrument users and four experienced users fitted with new instruments. Speech recognition was measured using the Connected Speech Test (CST) at the time of fitting and 10 weeks later. No significant change in mean benefit (aided minus unaided performance) was measured in noisy or reverberant listening conditions. However, in a low-noise background there was a statistically significant increase in mean benefit over time of 5%–6%. There was no control group, so it is not clear if the improvement was due to practice from repeated use of the same test material. In a further study, Cox *et al.* (1996) included experienced hearing instrument users as a control. The experimental group of 22 new hearing instrument users was fitted monaurally and tested at the time of fitting and 12 weeks later. The mean benefit on CST increased from 4% at the time of fitting to 8% 12 weeks later; there was no improvement in the control group. The change in benefit was due to an increase in aided performance with no change in unaided performance.

In 1992, Gatehouse used the Four-Alternative Auditory Feature (FAAF) test to track performance over a 12-week postfitting period in four new hearing instrument users fitted monaurally. Benefit increased from 5% at the time of fitting to greater than 15% at 12 weeks postfitting; the improvement commenced at around 6 weeks postfitting and was due to both an increase in the aided condition and a decrease in the unaided condition. No improvement in benefit was observed in the not-fitted (control) ear of these monaurally aided subjects. In a subsequent experiment, Gatehouse (1993) refitted 36 experienced hearing instrument users with a new instrument that provided greater high-frequency amplification than their previous one. Aided performance was measured on the FAAF test at 0, 8, and 16 weeks after the new fitting. Mean scores with the old and new hearing instrument were similar initially but increased significantly by 2.3% at 8 weeks and 4.4% at 16 weeks.

Horwitz and Turner (1997) compared 13 new hearing instrument users fitted monaurally with 13 experienced hearing instrument users as the controls. Speech recognition was measured over 18 weeks using the Nonsense Syllable Test (NST) with the hearing instrument at a fixed initial gain set-

---

[a] Portions of this work were presented at the British Society of Audiology Experimental Short Papers Meeting, Oxford, UK, September 2001.

[b] Now at: Human Communication and Deafness Group, School of Education, University of Manchester, Manchester M13 9PL, UK. Electronic mail: kevin.j.munro@man.ac.uk

ting and at the subject's daily adjusted gain setting. There was a gradual increase in mean benefit from around 6% at the time of fitting to around 14% at 18 weeks for the fixed gain setting in the new users. This was due to increases in aided performance rather than reductions in the unaided performance. A similar finding was reported for the user-adjusted gain setting. There was no improvement in the control group.

By contrast, several studies have failed to demonstrate auditory acclimatization, including Bentler *et al.* (1993), Humes *et al.* (1996, 2002), Munro and Lutman (2000), Saunders and Cienkowski (1997), Surr *et al.* (1998), and Taylor (1993). There are at least three possible explanations for this. First, the subjects may have had little opportunity to improve over time due to mild hearing loss (Taylor, 1993), previous experience of a hearing instrument (Bentler *et al.*, 1993), or limited use (Taylor, 1993; Bentler *et al.*, 1993). Second, Robinson and Summerfield (1996) suggest that the negative findings may be due to the specificity of perceptual learning: the test methods employed may not have picked up changes that had occurred. Third, findings may relate to the presentation level of the test material used to demonstrate acclimatization. The main focus of the present study is this third possibility.

Munro and Lutman (2000) used very similar methodology to Gatehouse (1992), but did not show an improvement in benefit over time. One difference between the two studies is that Munro and Lutman used a presentation level of 58 dB SPL, whereas Gatehouse used a presentation level of 65 dB SPL. This difference in presentation level is reflected in a difference in the mean initial benefit score between the studies (14% for Munro and Lutman but only 5% for Gatehouse); at higher presentation levels the unaided speech is already audible so the difference between the aided and unaided score is small. The dependence of acclimatization on presentation level can be rationalized as follows. 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 speech test materials are presented at 55 dB, they are amplified to 75 dB and hence there is little difference compared to that experienced before amplification. However, if material is presented at 70 dB and amplified to 90 dB, this reaches a level not previously experienced for normal speech. The organism must adapt and reorganize to discriminate and utilize the speech cues that are now coded into an unfamiliar part of the neural representation of sound. It is proposed that this process of reorganization underlies the phenomenon of auditory acclimatization. In the foregoing example, reorganization would be required to discriminate speech presented at 70 dB but not at 55 dB.

The aim of this study is to test the hypothesis that auditory acclimatization is only revealed when testing involves using speech materials at higher presentation levels. This is achieved by measuring aided and unaided speech recognition performance in newly aided subjects at three different speech presentation levels repeatedly over 12-weeks postfitting. The null hypothesis is that there are no differential changes in benefit over time as a function of presentation level. The alternative hypothesis is that improvements are greater at the higher presentation levels and negligible at the lowest level. In order to remove hearing instrument gain setting as a confounding variable, subjects are fitted with linear amplification with the gain control disabled for the duration of the study.

II. METHODS

A. Subjects

Sixteen subjects were recruited (10 male, 6 female) with a mean age of 70 years (s.d. ± 5.5). The sample size was calculated for paired data with a mean difference of 4% and a standard deviation of the difference of 5%. Fifteen subjects were required for a statistical power of 80% at a two-tailed significance level of 5% using Student's *t*-test, but 16 were recruited to allow for attrition. No subjects withdrew from the study and no subjects missed any test session.

Subjects were all first-time hearing instrument users recruited from the local hospital audiology service. They complained 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 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; however, they were naive to the changes expected over time. The study was approved by local hospital and university ethical committees.

Table I summarizes hearing thresholds and uncomfortable loudness levels, showing symmetrical, high-frequency sensorineural hearing impairments with reduced dynamic range. The mean change in hearing threshold level between the beginning and end of the study [initial minus final measurement] was 0 dB (s.d. ± 6 dB). Over 80% of hearing thresholds changed by less than 5 dB at each audiometric frequency and 90% changed by less than 10 dB; this is consistent with audiometric test–retest differences reported in the literature (for example, Robinson, 1991).

Each subject received standard shell earmolds to which a parallel 0.8-mm vent was added. This is standard clinical practice to equalize pressure and avoid a build-up of moisture in the ear canal: it is likely to have a negligible effect on the low-frequency gain of the amplified sound path. The same model of hearing instrument was used for all subjects: Phonak Sono-Forte 331X-L PiCS which is a miniature, high-gain, digitally programmable BTE hearing instrument. It has a three-way audio filter allowing considerable flexibility when tailoring the frequency response. It can also store three programs that can only be accessed using a hand control. The subjects were not issued the hand control and were therefore unable to change programs during the study. The fittings were all monaural, which allowed the not-fitted ear to be used as a control. The not-fitted control ear underwent the

TABLE I. Summary of audiometric data for the subjects. The table includes the mean air conduction, not-masked bone conduction hearing threshold levels, and mean uncomfortable loudness levels. One standard deviation is given in brackets ($n=16$).

Frequency (kHz)	0.25	0.5	0.75	1	1.5	2	3	4	6
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)	

same test protocol as the fitted ear using the subjects' own hearing instrument. The only difference between test and control ears was wearing of the instrument between sessions. It was important to have a control condition because increases in performance might otherwise be explained by practice effects due to repeated exposure to the test material. However, since both ears had identical opportunities for practice with the test materials, taking the difference between scores on the two ears gave a robust measure of acclimatization. Some subjects expressed a clear preference for ear to be fitted: the fitted ear in the remaining subjects was selected at random. Six subjects were fitted in the right ear and ten subjects were fitted in the left ear. The earmold for the ear to be the control ear was retained in the laboratory and not given to the subject.

The hearing instrument response was tailored to the NAL-RM target values (Byrne and Dillon, 1986) using the real-ear insertion gain (REIG) protocol on the Rastronics Portarem 2000 probe-tube measurement system. This involved measuring the real-ear unaided response and subtracting it from the real-ear aided response for each subject. Table II shows the mean target values (±1 s.d.) for the fitted ear along with the best match to target. The mean target gain from 1–4 kHz (where the hearing impairment is greatest) is 22–23 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 7 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). Since the gain control was disabled for the duration of the study, special care was taken to avoid exceeding uncomfortable loudness levels. The maximum output was set close to uncomfortable loudness level and adjusted if the subject reported any undue discomfort. The mean real-ear saturation response was set around 104 dB SPL using mainly peak-clipping output limiting. These settings were entered into memory one of the hearing instruments since 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 postfitting. The results are shown in Table III. 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 and 1.5 dB, respectively). While it is possible that there were real changes in REIG, the most likely explanation 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.

B. Test materials

The speech recognition test used was a digitized version of the Four-Alternative Auditory Feature (FAAF) test, as used in acclimatization studies of Gatehouse (1992, 1993). The FAAF test is a forced-choice word recognition task based on the rhyme test principle, described by Foster and Haggard (1979, 1987). It consists of 20 sets of four minimally paired words, each based on two binary auditory/

TABLE II. 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 (kHz)	0.25	0.5	0.75	1	1.5	2	3	4
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)

TABLE III. Mean change in gain (dB) at the user setting, relative to the time of fitting, for the fitted ear. One standard deviation is given in brackets.

Frequency (kHz)	0.25	0.5	0.75	1	1.5	2	3	4
				Real-ear insertion gain				
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)
				2-cc coupler gain				
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)

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 is presented acoustically in the carrier phrase (spoken by a male speaker) and the subject’s task is to select the correct word from the choice of four shown on a touch-screen monitor. The frequency spectrum of the key words is shown in Fig. 1 in comparison with the frequency response of conversational speech (from ANSI, 1997). The FAAF key words show slightly more emphasis at the high frequencies compared to the conversational speech.

The FAAF materials were replayed from a standard 16-bit computer sound card, at a sample rate of 20 kHz, routed via a Grason-Stadler GSI 61 clinical audiometer to a Fostex 6301B loudspeaker. They were presented against a background of steady noise that was filtered to give a similar long-term spectrum to key words and delivered by the same loudspeaker. The spectrum of this noise is also shown in Fig. 1. The loudspeaker was located at an azimuth of 0° and a distance of 1.5 m from the subject. Speech and noise levels were defined in terms of the overall SPL measured at the reference point, defined as the position occupied by the center of the subject’s head, with the subject absent (ISO 8253-2, 1998).

Pearsons *et al.* (1977) measured the mean speech levels in a variety of settings and reported “casual,” “normal,” and “raised” speech for male speakers of 56, 61, and 68 dB SPL, respectively, with female speakers approximately 2 dB lower. Therefore, the FAAF test was presented at overall levels of 55, 62, and 69 dB SPL to approximate quiet, normal, and raised speech.

The mean Speech Intelligibility Index (SII; ANSI, 1997) values for the aided and unaided condition were used to provide a guide to changes in audibility for each presentation level and are shown in Table IV. The equivalent speech spectrum level of the FAAF test was based on overall levels of 55, 62, and 69 dB SPL for the key words. The SII assumes that the speech area covers a dynamic range of 30 dB. The SII is the weighted sum of band audibility where the weight ranges from zero to 1 and represents the relative importance of each frequency band to the understanding of speech. For recognizing nonsense syllables, the band around 2 kHz is most important; for recognizing sentences, the band around 0.5 kHz is most important. The band importance function for the FAAF test has not been determined but is likely to be similar to that for nonsense syllables. The SII includes a level distortion factor (LDF) to account for the deterioration in speech performance that occurs when overall level exceeds 73 dB SPL. The equivalent noise spectrum of the FAAF-shaped noise spectrum was determined for each subject (based on the SNR). The aided values were obtained by

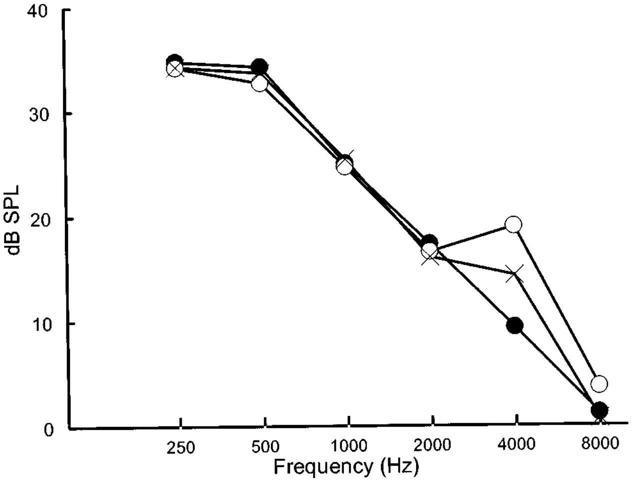

FIG. 1. Octave band levels for conversational speech at 62.35 dB SPL. FAAF words, open circles; FAAF-shaped noise, crosses; conversational speech from ANSI S3.7 (1997), filled circles.

TABLE IV. Mean Speech Intelligibility Index for the fitted ear. The values in parentheses incorporate the desensitization correction from Pavlovic *et al.* (1986). The equivalent speech spectrum level of the FAAF test was based on overall SPLs of 55, 62, and 69 dB for the key words. The equivalent noise spectrum of the FAAF-shaped noise spectrum was determined for each subject (based on the SNR). The aided values were 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.

	Overall
Unaided	
55 dB SPL	0.08(0.06)
62 dB SPL	0.12(0.09)
69 dB SPL	0.20(0.13)
Aided	
55 dB SPL	0.25(0.15)
62 dB SPL	0.39(0.21)
69 dB SPL	0.42(0.22)
Change (aided−unaided)	
55 dB SPL	0.17(0.09)
62 dB SPL	0.27(0.12)
69 dB SPL	0.22(0.09)

adding the user insertion gain to the speech and noise levels. [The values in parentheses in Table IV incorporate the desensitization correction from Pavlovic *et al.* (1986) to account for the poorer performance of hearing impaired subjects compared to normal-hearing subjects for a given level of audibility.] The low aided SII value at 55 dB SPL indicates that part of the speech envelope remains inaudible even after amplification. The small improvement after aiding at a presentation level of 69 dB SPL indicates that much 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 condition.

Each subject was tested at a fixed signal-to-noise ratio (SNR) across all test conditions. A pilot study indicated that a single individually tailored SNR would give a score in the range 40%–80% across all test conditions which should be sufficiently high to maintain subject motivation but avoid ceiling effects. The SNR was determined for each subject using the adaptive FAAF test strategy described by Lutman and Clark (1986), which targets 71% correct with a speech presentation level of 62 dB SPL. SNR was defined as the difference in decibels between the SPL of the key words and the SPL of the noise. Subjects then practiced the test at all three presentation levels, while the investigator adjusted the SNR to compensate for any improvement in performance with practice; this normally involved increasing the noise by 1–3 dB over the course of several practice sessions. Performance was then measured throughout the study at this fixed SNR for all speech presentation levels. The SNR ranged from 0 to -5 dB across subjects with a median value of -2 dB. Subject performance was assessed as the number of words correctly identified and was expressed as percent correct. A random numbers table was employed to generate different orders of the 80 items for the multiple presentations involved.

Testing was carried out in a quasifree sound field [ISO 8253-2 (1998)] with reverberation times <0.3 s when measured in octave intervals from 0.125 to 4 kHz. The ambient noise levels measured at the reference test point (with the test equipment switched on but subject absent) was sufficiently low to have allowed 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 (1998)].

C. Procedures

Subjects attended several sessions (usually 4–6) each lasting approximately 2 h before obtaining their hearing instrument. This allowed tailoring of the hearing instrument frequency response and determination of 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 baseline FAAF tests were performed within 7 days of fitting; testing could not commence until the fixed user gain had been established. Data for the initial test session is reported as occurring at zero days postfitting; however, for some subjects this may have taken place as late as 7 days postfitting. Testing was repeated at 6 and 12 weeks postfitting. The order of testing was balanced across subjects and test sessions.

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 (4, 4–8, 8–12, >12 h). Subjects were required to wear the hearing instrument for at least 6–8 h per day. The number of subjects who reported using their hearing instrument more than 8 h per day was never less than 13 (82%) during any 1-week period.

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 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, intermodulation distortion was assessed using a composite speech-weighted input signal. None of the hearing instruments tested failed the electroacoustic tests and none was operating with observable intermodulation distortion for the settings and input levels used in the present study. One subject reported a nonfunctioning device between test sessions and a replacement model was fitted the following day. The mean input–output function of the hearing instrument was measured 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 output was linear at input levels below 85 dB SPL. As the maximum combined level of speech and noise used in the study was 72 dB SPL, there was approximately 13 dB of headroom and hence hearing instruments were not likely to be operating in significant saturation for any of the speech input levels used.

The first FAAF list of each session was used as a practice run before commencing formal testing. Two FAAF lists (i.e., 160 words) were then used for every test condition: this involved two word lists presented at each level (55, 62, and 69 dB SPL) in the aided and unaided condition for each ear. The subject's own hearing aid was used when measuring aided performance in both the test ear and the control ear. All testing was performed monaurally with the nontest ear plugged and muffed. Benefit scores were obtained by subtracting the unaided score from the aided score. The complete FAAF testing took 100–140 min, including a 20-min break. After the last test session was completed, pure-tone audiometry was repeated to check for any change in hearing sensitivity.

The data were inspected before analysis to confirm it was appropriate to use parametric statistics. Statistical analysis of the mean benefit, aided and unaided scores, consisted primarily of repeated-measures analysis of variance (ANOVA), using SPSS version 10.0, to determine if performance changed significantly over the course of the study. The three factors entered into the ANOVA were ear (test and control), presentation level (55, 62, and 69 dB SPL) and postfitting time (0, 6, and 12 weeks). The degrees of freedom were modified using the Greenhouse–Geisser correction

TABLE V. Mean score (in percent) as a function of postfitting time. One standard deviation is given in parentheses ($n=16$).

Postfitting time (weeks)	0	6	12	Mean
Benefit scores				
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)
Aided scores				
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)
Unaided scores				
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)

when there was a statistically significant deviation from sphericity on Mauchly's test.

III. RESULTS

A. Change in benefit score over time

The mean FAAF benefit (±1 s.d.), as a function of postfitting time, is given in Table V. The mean benefit scores in the fitted ear with speech presented at 55, 62, and 69 dB SPL at the time of fitting were 19%, 8%, and 0%, respectively. Similar results were obtained in the not-fitted control ear.

The study was concerned primarily with changes in benefit, relative to the time of fitting (Fig. 2). The scores increased for all presentation levels in the fitted ear, with increasing presentation level. By 12 weeks, the mean scores had increased by 2.0%, 3.2%, and 5.9% across the three presentation levels. The corresponding increases in the control ear were all less than 1%.

The mean benefit scores were analyzed using a three-factor (time [3], ear [2], and SPL [3]) repeated-measures ANOVA (see Table VI). As expected, there was a statistically significant difference in mean benefit score with presentation level. The ANOVA orthogonal polynomial breakdown revealed a significant linear component [$F(1,15)=36.7$; $p<0.01$] showing that the slope of the curves in Fig. 2(a) increased with presentation level. The main effects of ear and postfitting time (averaged across other factors) were not statistically significant; however, there was a statistically significant interaction between ear and presentation level. Again, only the linear component of the polynomial breakdown was significant [$F(1,15)=16.2$; $p<0.01$]. This interaction indicates that the fitted and control ears differ over time; the fitted ear showed increasing scores while the control ear showed little change. This was confirmed by investigating each ear separately using a two-factor (time [3] and SPL [3]) repeated measures ANOVA; time was statistically significant in the fitted ear but not the control ear (see Table VII).

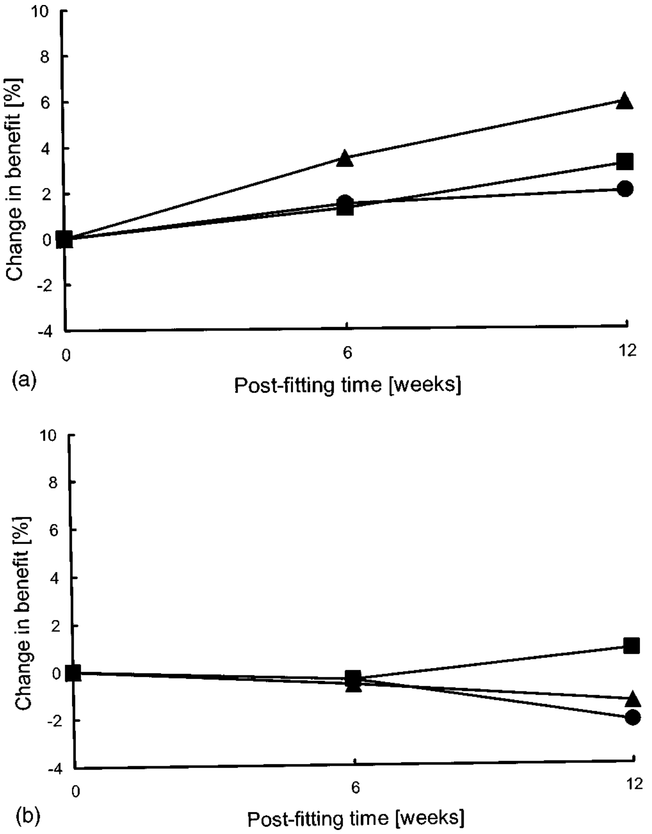

FIG. 2. Mean change in FAAF benefit (aided minus unaided), relative to the time of fitting, as a function of postfitting time with speech at 55 (circles), 62 (squares), and 69 dB SPL (triangles). The fitted ear is shown in (a) and the control ear in (b).

Although postfitting time was not significant in the three-factor ANOVA, one-factor (time [3]) repeated-measures ANOVAs performed separately for each presenta-

TABLE VI. 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 VII. Summary of simple main effects when the benefit scores from each ear were 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.

tion level showed no statistically significant change in benefit at 55 dB [$F(2,30)=0.5$, $p>0.05$], but there was a statistically significant increase in benefit at 69 dB [$F(2,30)=5.6$, $p<0.01$]. The improvement in benefit in the fitted ear over time was not simply due to skewing by a small number of subjects; the change in benefit at the highest presentation level ranged from −4% to +19%, with eight (50%) subjects showing an increase in excess of the mean improvement of +5.9%.

B. Aided performance over time

The mean aided and unaided scores (±1 s.d.) from which the benefit scores are calculated are also given in Table V. The initial aided scores were around 72% in both ears. The changes in score are shown in Fig. 3 for the fitted ear and the control ear. The aided scores increased over time in the fitted ear for all presentation levels (by 1.0%, 2.6%, and 3.6% at 55, 62, and 69 dB SPL, respectively). Changes in the control ear were all negative and typically around −2.0%.

The mean aided scores were analyzed using a three-factor (time [3], ear [2], and SPL [3]) repeated-measures ANOVA (see Table VIII). The main effect of ear was statistically significant; the mean score (averaged across other factors) in the fitted ear was 73.9% compared with 72.2% in the control ear. The linear component of the ANOVA polynomial breakdown was significant [$F(1,15)=6.0$; $p=0.03$]. In addition, there was a statistically significant interaction between ear and postfitting time [$F(2.30)=15.5$; $p<0.01$]. This interaction indicates that the fitted and control ears differ over time; the fitted ear showed increasing scores while the control ear showed slightly decreasing scores. Again, only the linear component of the polynomial breakdown was significant [$F(1,15)=27.0$; $p<0.01$]. Two-factor (time [3] and SPL [3]) repeated-measures ANOVA performed on each ear separately (see Table IX) revealed postfitting time to be statistically significant for each ear, but there was no interaction with presentation level. The linear component of the polynomial breakdown was significant for both ears [fitted ear, $F(1,15)=13.4$; $p<0.01$; control ear, $F(1,15)=7.4$, $p=0.02$].

Consistent with the benefit data, when ANOVAs were performed separately for each, presentation level showed no statistically significant change at 55 dB [$F(2,30)=1.2, p>0.05$] but there was a statistically significant increase in benefit at 69 dB [$F(2,30)=6.5, p<0.01$]. The improvement in aided performance in the fitted ear was not simply due to skewing by a small number of subjects; the change in aided performance at the highest presentation level ranged from −4% to +14% with nine (56%) subjects showing an increase in excess of the mean improvement of +3.6%.

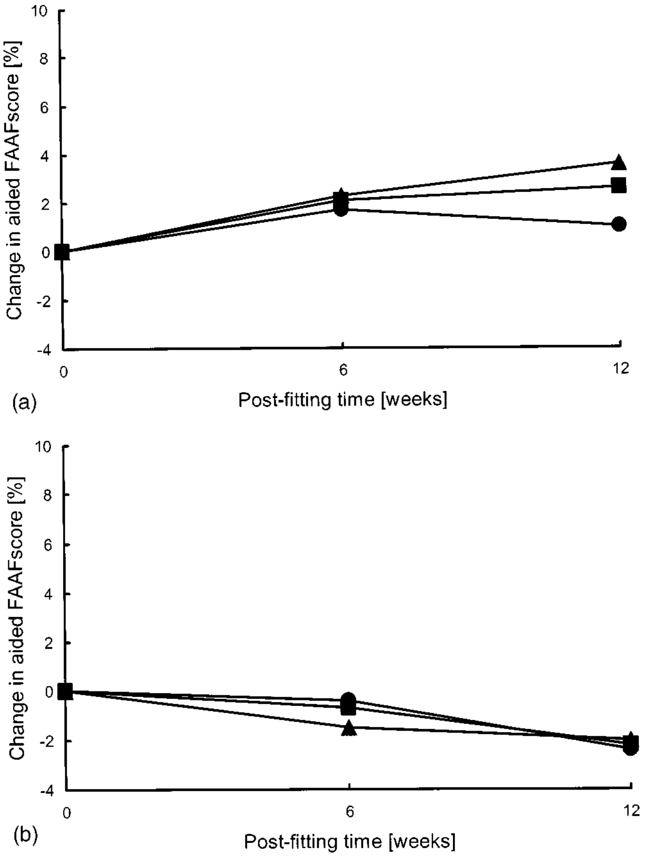

FIG. 3. Mean change in aided FAAF score (%) relative to fitting, at a speech presentation level of 55 (circles), 62 (squares), and 69 dB SPL (triangles). The fitted ear is shown in (a) and the control ear in (b).

C. Unaided performance over time

The mean unaided scores (±1 s.d.), as a function of postfitting time, are also given in Table V. The initial performance was typically around 55% with speech at 55 dB SPL but nearer 71% with speech at 69 dB SPL. The change in

TABLE VIII. Summary of repeated-measures analysis of variance on the aided scores [$n=16$]. The three within-subject factors were time [3], ear [2], 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 IX. 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 X. 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

score relative to the time of fitting is shown in Fig. 4 for both ears. There was a trend for scores to decrease slightly over time for all presentation levels in both ears, but this was not significant (see Table X).

IV. DISCUSSION

A. Speech recognition performance at the time of fitting

The magnitudes of the derived benefit scores (aided minus unaided) are greater for the lower presentation levels. The unaided scores are greater at higher presentation levels whereas the aided scores are similar across levels; this has the effect of reduced benefit score at the higher level. This is expected since performance in the unaided condition is mediated mainly by audibility, whereas in the aided condition most speech is audible and performance is limited equally for all levels by the competing FAAF-shaped noise. The mean benefit scores at the time of fitting were comparable in both ears, consistent with the symmetrical hearing loss.

Mean aided performance with speech presented at 69 dB SPL was 1%–2% lower than at the two lower presentation levels. The decrement was small and not statistically significant. Some studies indicate that performance remains constant at high speech presentation levels (for example, Duquesnoy and Plomp, 1983), while others indicate that performance is reduced (for example, Larson *et al.*, 2000; Studebaker *et al.*, 1999). This latter "roll-over" effect is presumably related to increased upward spread of masking resulting in vowels masking consonants.

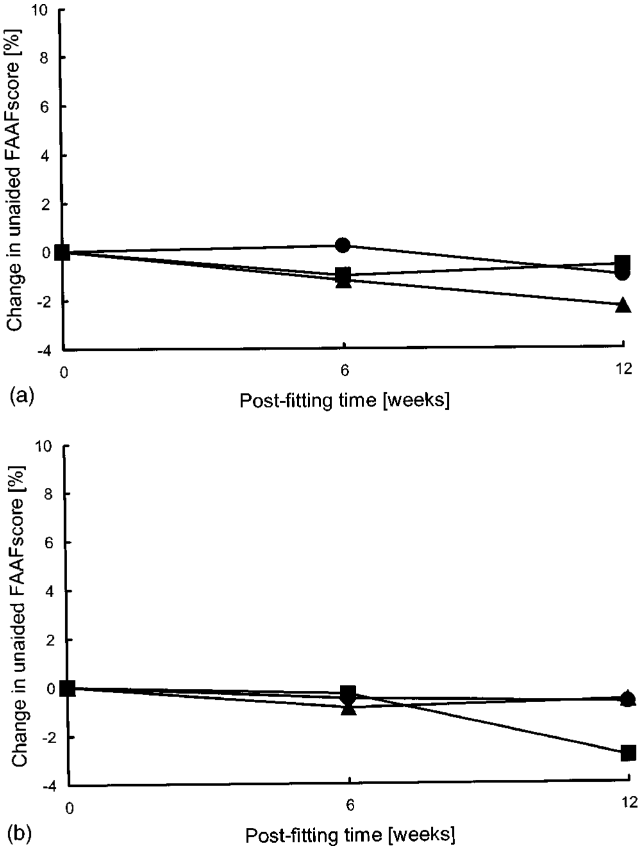

FIG. 4. Mean change in unaided FAAF score (%) relative to fitting, at a speech presentation level of 55 (circles), 62 (squares), and 69 dB SPL (triangles). The fitted ear is shown in (a) and the control ear in (b).

B. Change in benefit over time

Following 12 weeks of hearing instrument use there was a systematic increase in benefit of around 4% (averaged across presentation level) 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. The significant increase in benefit over time replicates the findings of studies discussed earlier that demonstrate acclimatization.

The only previous study that used the not-fitted ear as a control was Gatehouse (1992); this also showed no increase in the control condition. This finding is consistent with the specific nature of perceptual learning discussed by Robinson and Summerfield (1996). Studies where human subjects are trained on frequency discrimination (Demany, 1985; Irvine *et al.*, 2000) temporal discrimination (Wright *et al.*, 1997), and sound localization (Wright and Fitzgerald, 2001) tasks show 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 the fitted ear and may not transfer to the control ear.

The magnitude of the acclimatization 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 postfitting. Gatehouse reported a mean unaided score of around 71% and a benefit score of around 4% at the time of fitting.

Similar scores were obtained in the present study with a presentation level of 69 dB SPL. At this presentation level, the change in mean benefit in the present study was 5.9%. A possible explanation for the apparent discrepancy is the difference in the REIG provided in the two studies; Gatehouse provided approximately 12 dB more gain than the present study. This means that there was a larger difference in level between the amplified and unamplified speech. This may mean that more reorganization was required by the subjects in the Gatehouse study, resulting in greater improvement over time.

The implications for clinical practice are unclear. Several authors (for example, Gatehouse, 1997, 1998; Humes *et al.*, 2001) have demonstrated that word recognition scores, in isolation, give an incomplete representation of hearing aid benefit. In addition, the increase in mean aided performance for average speech was only +2.6%. The study cannot speak to longer time scales because there is no evidence of performance reaching a plateau by 12 weeks postfitting. This finding is consistent with other studies conducted over a similar time scale (Gatehouse, 1992; Horwitz and Turner, 1997).

C. Changes in aided and unaided scores over time

The change in benefit score in the fitted ear was due to both an increase in aided performance and a decrease in unaided performance, although only the former reached statistical significance. There are mixed findings in the acclimatization literature regarding changes in the unaided condition. The studies by Cox *et al.* (1996) and Horwitz and Turner (1997) do not show any reduction in the unaided scores over a 12–18-week postfitting period (although a subgroup of subjects from the Cox *et al.* study was reviewed 9 months postfitting and the results showed a dramatic reduction in unaided performance). The clearest demonstration of a reduction in performance in the unaided condition (approximately 8%) comes from Gatehouse (1992). This discrepancy may also be related to the greater amount of REIG provided by Gatehouse. If the fitted ear becomes accustomed to hearing speech at a high amplified level, then over time it will perform worse at the less familiar level of unaided speech.

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, 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 deprived of adequate stimulation relative to the fitted ear. This finding is consistent with retrospective reports of late onset auditory deprivation (see the review by Neuman, 1996), although it occurred 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 postfitting time period of 12 weeks. It may be that resources within the auditory system are diverted to subserve the fitted ear at the expense of the other ear.

D. The influence of presentation level

The results show greater increases in benefit over time at the higher presentation levels. This is supported by the one-factor (time) repeated-measures ANOVAs performed separately for each presentation level. 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 variability means that the statistical power was only around 40%: this resulted in a material risk of a type II error. Approximately double the number of subjects would be 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 postfitting time period, a wider range of presentation levels, or if subjects were provided with greater REIG.

A summary of an experiment that measured the acclimatization effects with linear and nonlinear signal processing was included in an article by Gatehouse (1998). One finding was that the acclimatization effect with linear hearing instruments differed with presentation level: acclimatization was shown 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.

Several previous studies have failed to show an acclimatization effect despite using a presentation level similar to the highest level used in the present study. Bentler *et al.* (1993) used a speech presentation level between 50–60 dB HL. When transformed from hearing level to sound-pressure level, the presentation level is approximately 65–70 dB. However, there are potentially important methodological differences between the two studies that could explain the difference in outcome. For example, Bentler *et al.* included subjects with a relatively mild hearing impairment, some were existing hearing aid users, the frequency response failed to meet NAL-RM targets above 2 kHz, and the gain control was not fixed across time. Despite these differences, and the lack of statistically significant improvements in mean performance over time, several of the performance measures (low predictability SPIN sentences and NST in noise) showed an increase of around 4% at 1–3 months postfitting; this is comparable with the present study.

There are a number of additional studies that demonstrate a level-dependent effect after hearing instrument use. Gatehouse (1989) measured the SNR required for an unaided performance of 50% 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 favorable 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 score 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, these differences in score can be converted into differences in SNR using the average FAAF performance intensity function obtained from hearing-impaired subjects as 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 ears of approximately 2 dB. Thus, the magnitude of change expressed in terms of SNR is similar to that reported by Gatehouse.

In 1995, Robinson and Gatehouse (1995) reported the results of a retrospective study that showed that the fitted ear of subjects aided monaurally was better able to discriminate intensity than the not-fitted ear at levels above 95 dB SPL; in contrast, the not-fitted ear was better able to discriminate intensity than the fitted ear at lower levels. In 1996, Robinson and Gatehouse (1996) reported a prospective study 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 perform better at high levels but worse at lower 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 louder in the fitted ear than in the control ear over the range of intensities studied. 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. 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 become accustomed.

Before fitting, the ears of the subjects used in the present study would be expected to be performing most efficiently over the range 55 to 69 dB SPL since this is assumed to cover the range of levels experienced in real life. After fitting, this range will be shifted to a higher level at the eardrum; however, there is an overlap between the range of levels experienced when aided and unaided (raised speech before aiding is similar in level to quiet speech after aiding). Hence, only the level of aided speech at 69 dB SPL is ‘‘new’’ to the subjects and it is for this reason that a period of acclimatization is required for this condition.

An alternative explanation for the maximum acclimatization effect occurring at the highest presentation level is 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. Thus, the maximum acclimatization effect occurred when listening in the presence of the highest noise levels. It is possible that the subject required time to become able to extract the speech signal from the high-level background noise. For this to be correct, the subject would have to be exposed to high levels of noise in every day life. The FAAF-shaped noise was used in the study to limit the speech recognition scores at the time of fitting: it was not selected to be representative of everyday listening conditions. Although the noise levels experienced by the subjects in their daily 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 more favorable than the median of -2 dB used in the present study. Thus, it appears unlikely that there would have been sufficient opportunity (in every day life) for subjects to learn to extract a signal at the negative SNRs used in the present study. In addition, this does not explain why performance decreased over time in the fitted ear when the hearing instrument was removed. It should be possible to design an experiment that involves speech material being presented with and without masking noise (but avoiding ceiling effects). If performance were to increase under both conditions, then acclimatization cannot be due solely to the ability to extract speech from noise.

It is interesting to speculate on the physiological mechanism responsible for acclimatization. It is possible that acclimatization may entail reallocation of resources to areas of the auditory system representing high-level stimuli. This is consistent with the progressive improvements in intensity discrimination reported by Gatehouse and Robinson (1996).

With the onset of a hearing impairment, one might envisage that high-threshold regions within the auditory system may become underused because the sound levels seldom reach sufficient levels to activate them. This might lead to the lower threshold regions invading the underused areas due to lack of competition from high stimulus levels. After provision of amplification, the auditory system has to reorganize or reacclimatize to the range of high signal levels. The effect may be that the auditory system increases its representation of 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 reorganization. This may explain why Gatehouse reported a larger acclimatization effect in subjects who were fitted with hearing instruments that provided greater gain than in other studies. Reorganization in the intensity domain is consistent with the remapping of the relationship between intensity and loudness reported by Gatehouse and Robinson (1996); the greatest capacity for discriminating differences in intensity shifts from lower to higher intensities as subjects gain amplification experience.

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

These processes take time to occur and can explain the phenomenon of auditory acclimatization. Because the reallocation in the intensity domain involves the upper end of the intensity range, auditory acclimatization is most readily demonstrated using speech at relatively high presentation levels. The present study demonstrates clear acclimatization for speech at 69 dB but minimal acclimatization for speech at 55 dB, consistent with the above explanation.

V. CONCLUSIONS

The results of this study confirm the presence of auditory acclimatization predominantly when using speech in noise at a high presentation level. This suggests that studies of auditory acclimatization must use a level of amplified speech signal that is higher than commonly experienced in everyday life, prior to aiding. Further experiments are desirable to confirm this effect across a wider range of presentation levels.

ACKNOWLEDGMENTS

This work was supported by a research studentship from the NHS Research and Development Directorate in the UK.

ANSI (**1989**). ANSI S3.6 (1989), ''Specification for Audiometers'' (American National Standards Institute, Inc., New York).

ANSI (**1997**). ANSI S3.5 (1997), ''American National Standard Methods for the Calculation of the Speech Intelligibility Index'' (American National Standards Institute, Inc., New York).

Arlinger, S., Gatehouse, S., Bentler, R. A., Byrne, D., Cox, R. M., Dirks, D. D., Humes, L., Neuman, A., Ponton, K., Robinson, K., Silman, S., Summerfield, A. Q., Turner, C. W., Tyler, R. S., and Willott, J. F. (**1996**). ''Report of the Eriksholm workshop on auditory deprivation and acclimatization,'' Ear Hear. **17**, 87S–98S.

Bentler, R. A., Neibuhr, D. P., Getta, J. P., and Anderson, C. V. (**1993**). ''Longitudinal study of hearing aid effectiveness. I. Objective measures,'' J. Speech Hear. Res. **36**, 808–819.

Byrne, D., and Dillon, H. (**1986**). ''The National Acoustics Laboratories (NAL) new procedure for selecting the gain and frequency response of a hearing aid,'' Ear Hear. **7**, 257–265.

Cox, R. M., and Alexander, G. C. (**1992**). ''Maturation of hearing aid benefit: Objective and subjective measurements,'' Ear Hear. **13**, 1311–1341.

Cox, R. M., Alexander, G. C., Taylor, I. M., and Gray, G. A. (**1996**). ''Benefit acclimatization in elderly hearing aid users,'' J. Am. Acad. Audiol **7**, 428–441.

Demany, L. (**1985**). ''Perceptual learning in frequency discrimination,'' J. Acoust. Soc. Am. **78**, 1118–1120.

Duquesnoy, A. J., and Plomp, R. (**1983**). ''The effect of a hearing aid on the speech reception threshold of hearing-impaired listeners in quiet and in noise,'' J. Acoust. Soc. Am. **73**, 2166–2173.

Foster, J. R., and Haggard, M. P. (**1979**). ''FAAF. An effective analytical test of speech perception,'' Proc. Inst. Acous. **182**, 9–12.

Foster, J. R., and Haggard, M. P. (**1987**). ''The four alternative auditory feature test (FAAF): Linguistic and psychometric properties of the material with normative data in noise,'' Br. J. Audiol. **21**, 165–174.

Gatehouse, S. (**1989**). ''Apparent auditory deprivation effects of late onset: The role of presentation level,'' J. Acoust. Soc. Am. **86**, 2103–2106.

Gatehouse, S. (**1992**). ''The time course and magnitude of perceptual acclimatization to frequency responses: Evidence from monaural fitting of hearing aids,'' J. Acoust. Soc. Am. **92**, 1258–1268.

Gatehouse, S. (**1993**). ''Role of perceptual acclimatization in the selection of frequency responses for hearing aids,'' J. Am. Acad. Audiol **4**, 296–306.

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

Gatehouse, S. (**1998**). ''Speech tests as measures of outcome,'' Scand. Audiol. Suppl. **27**, 54–60.

Gatehouse, S., and Robinson, K. (**1996**). ''Acclimatization to monaural hearing aid fitting effects on loudness functions and preliminary evidence for parallel electrophysiological and behavioural effects,'' in *Psychoacoustics, Speech and Hearing aids*, edited by B. Kollmeier (World Scientific, Singapore), pp. 319–330.

Hawkins, D., Alvarez, E., and Houlihan, J. (**1991**). ''Reliability of three types of probe tube microphone measurements,'' Hear. Instru. **42**, 14–16.

Horwitz, A. R., and Turner, C. W. (**1997**). ''The time course of hearing aid benefit,'' Ear Hear. **18**, 1–11.

Humes, L. E., Barlow, N. N., Garner, C. B., and Wilson, D. L. (**2000**). ''Prescribed clinician-fit versus as-worn coupler gain in a group of elderly hearing-aid users,'' J. Speech Lang. Hear. Res. **43**, 879–892.

Humes, L. E., Garner, C. B., Wilson, D. L., and Barlow, N. N. (**2001**). ''Hearing-aid outcome measures following one month of hearing aid use by the elderly,'' J. Speech Lang. Hear. Res. **44**, 469–486.

Humes, L. E., Halling, D., and Coughlin, M. (**1996**). ''Reliability and stability of various hearing-aid outcome measures in a group of elderly hearing-aid wearers,'' J. Speech Hear. Res. **39**, 923–935.

Humes, L. E., Wilson, D. L., Barlow, N. N., and Garner, C. (**2002**). ''Changes in hearing-aid benefit following one, two, or three years of hearing aid use by the elderly,'' J. Speech Hear. Res. (in press).

ISO 8253-2. (**1998**). *Acoustics—Audiometric test methods. Part 2: Sound Field Audiometry with Pure Tone and Narrow-band Test Signals* (International Organization for Standardization, Geneva, Switzerland).

Irvine, D. R. F., Martin, R. L., Klimkeit, E., and Smith, R. (**2000**). ''Specificity of perceptual learning in a frequency discrimination task,'' J. Acoust. Soc. Am. **108**, 2964–2968.

Larson, V. D., Williams, D. W., Henderson, W. G., Leuthke, L. E., Beck, L. B., Noffinger, D., Wilson, R. H., Dobie, R. A., Haskell, G. B., Bratt, G. W., Shanks, J. E., Stemmoachowicz, P., Studebaker, G. A., Boysen, A. E., Donahue, A., Canalis, R., Fausti, S. A., and Rappaport, B. Z. (**2000**). ''Efficacy of three commonly used hearing aid circuits,'' J. Am. Med. Assoc. **11**, 1806–1813.

Lutman, M. E., and Clark, J. (**1986**). ''Speech identification under simulated hearing-aid frequency response characteristics in relation to sensitivity, frequency resolution, and temporal resolution,'' J. Acoust. Soc. Am. **80**, 1030–1040.

Munro, K. J., and Lutman, M. E. (**2000**). ''Auditory acclimatization to amplified speech,'' Poster presentation at IHCON 2000, Lake Tahoe.

Neuman, A. C. (**1996**). ''Late-onset auditory deprivation: A review of past research and an assessment of future research needs,'' Ear Hear. **17**, 3S–13S.

Palmer, C. V., Nelson, C. T., and Lindley, G. A. (**1998**). ''The functionally and physiologically plastic adult auditory system,'' J. Acoust. Soc. Am. **103**, 1705–1721.

Pavlovic, C. V., Studebaker, G. A., and Sherbecoe, R. L. (**1986**). ''An articulation index based procedure for predicting the speech recognition performance of hearing-impaired individuals,'' J. Acoust. Soc. Am. **80**, 50–57.

Pearsons, K. S., Bennett, R. L., and Fidells, S. (**1977**). ''Speech levels in various noise environments'' (Report No. EPA-600/1-77-025). (U.S. Environmental Protection Agency, Washington, D.C.).

Robinson, D. W. (**1991**). ''Long-term repeatability of the pure-tone hearing threshold and its relation to noise exposure,'' Br. J. Audiol. **25**, 219–235.

Robinson, K., and Gatehouse, S. (**1995**). ''Changes in intensity discrimination following monaural long-term use of a hearing aid,'' J. Acoust. Soc. Am. **97**, 1183–1190.

Robinson, K., and Gatehouse, S. (**1996**). ''The time course of effects on intensity discrimination following monaural fitting of hearing aids,'' J. Acoust. Soc. Am. **99**, 1255–1258.

Robinson, K., and Summerfield, A. Q. (**1996**). ''Adult auditory learning and training,'' Ear Hear. **17**, 51S–65S.

Saunders, G. H., and Cienkowski, K. M. (**1997**). ''Acclimatization to hearing aids,'' Ear Hear. **18**, 129–139.

Shields, P. W., and Campbell, D. R. (**2001**). ''Intelligibility, subjective ratings and completion time scores using the FAAF test with hearing-impaired subjects and noisy reverberant environments,'' Br. J. Audiol. **35**, 237–245.

Studebaker, G. A., Sherbecoe, R. L., McDaniel, D. M., and Gwaltney, C. A. (**1999**). ''Monosyllabic word recognition at higher-than-normal speech and noise levels,'' J. Acoust. Soc. Am. **105**, 2431–2444.

Surr, R. K., Cord, M. T., and Walden, B. E. (**1998**). ‘‘Long-term versus short-term hearing aid benefit,’’ J. Am. Acad. Audiol. **9**, 165–171.

Taylor, K. S. (**1993**). ‘‘Self-perceived and audiometric evaluations of hearing aid benefit in the elderly,’’ Ear Hear. **14**, 390–394.

Turner, C. W., and Bentler, R. A. (**1998**). ‘‘Does hearing aid benefit increase over time?’’ Letter to Editor. J. Acoust. Soc. Am. **104**, 3673–3674.

Turner, C. W., Humes, L. E., Bentler, R. A., and Cox, R. M. (**1996**). ‘‘A review of past research on changes in hearing aid benefit over time,’’ Ear Hear. **17**, 14S–28S.

Wright, B. A., Buonomano, D. V., Mahnvcke, H. W., and Merzenich, M. M. (**1997**). ‘‘Learning and generalization of auditory temporal-interval discrimination in humans,’’ J. Neurosci. **17**, 3956–3963.

Wright, B. A., and Fitzgerald, M. B. (**2001**). ‘‘Different patterns of human discrimination learning for two interaural cues to sound-source localization,’’ Proc. Nat. Acad. Sci. U.S.A. **98**, 12307–12312.