An Examination of Changes in Hearing-Aid Performance and Benefit in the Elderly Over a 3-Year Period of Hearing-Aid Use

This brief report describes the changes in hearing-aid performance and benefit in 9 elderly hearing-aid wearers over a 3-year period following the hearing-aid fitting. Objective measures of hearing-aid performance included three measures of speech recognition: (a) the Nonsense Syllable Test (NST) presented at 65 dB SPL and a +8 dB signal-to-noise ratio (SNR), (b) the Connected Speech Test (CST) presented at 50 dB SPL in quiet, and (c) the CST presented at 65 dB SPL and a +8 dB SNR. Subjective, self-report measures of hearing-aid benefit included the Hearing Aid Performance Inventory (HAPI; B. E. Walden, M. E. Demorest, & E. L. Helper, 1984) and the Hearing Handicap Inventory for the Elderly (HHIE; I. Ventry & B. Weinstein, 1982). Performance and benefit measures were obtained at postfit intervals of 1 month, 6 months, 1 year, 2 years, and 3 years using a standardized measurement protocol. Individual data were evaluated using 95% critical differences established previously for each benefit measure and applied around the scores observed at the 1-month postfit interval. Little evidence was seen for systematic improvement in aided performance or benefit, consistent with that expected from acclimatization, in any participant or for any measure of benefit.

KEY WORDS: hearing aids, elderly, benefit, acclimatization

Over the past decade, there has been keen interest in the time course of changes in performance associated with the use of hearing aids. This time-dependent adaptation phenomenon is popularly termed “acclimatization” of hearing-aid benefit (Cox & Alexander, 1992; Gatehouse, 1992). Recently, Humes, Wilson, Barlow, and Garner (2002) examined changes in hearing-aid benefit in 134 elderly hearing-aid wearers over a 1-year period and in 49 of these same hearing-aid wearers over a 2-year period. Consistent with the findings in a comprehensive review of several smaller scale studies by Turner, Humes, Bentler, and Cox (1996), there was little evidence for acclimatization in the larger scale study by Humes et al. (2002).

Perhaps, however, the acclimatization process requires more than 2 years to take place. This is not inconceivable, because the acclimatization process is essentially self-taught. No formal training or rehabilitation is provided to the wearer. Rather, it is conjectured that the wearer, through prolonged and repeated exposures to amplified sound, learns to better understand amplified speech over time. Given that there is often a period of many years between the first appearance of hearing
problems and the subsequent use of hearing aids in the elderly, it is not unlikely that several years of hearing-aid use would be needed for acclimatization to amplified sound.

This brief report provides the individual data for a variety of hearing-aid benefit and performance measures obtained from 9 elderly hearing-aid wearers followed over a period of 3 years. These data represent a subset of the data from the larger scale study by Humes et al. (2002). The purpose of the presentation of these data is to extend the measurement period for that longitudinal study by a full year to 3 years postfit. Compared to the previously published larger scale study, the present study is smaller both in the number of participants included and in the number of measures of hearing-aid performance and benefit examined. Given the focus on examination of individual data in this study, only those performance and benefit measures for which 95% critical difference values were readily available were included in this report.

**Method**

The materials used and procedures followed are described in detail in the study by Humes et al. (2002). The key features are summarized here.

**Participants**

Of the 49 individuals who returned for the 2-year follow-up measurements in the study by Humes et al. (2002), 15 returned voluntarily for 3-year follow-up measurements. Of these, 2 had missing scores for several outcome measures at various postfit measurement intervals and their data were not evaluated further. In addition, because it was felt to be critical to have some measures of hearing threshold after the initial hearing evaluation (conducted prior to the hearing-aid fitting), only those participants who had complete pure-tone audiograms measured at either the 2-year or 3-year postfit interval were included. Nine of the remaining 13 individuals met this inclusion criterion.

The audiograms for these 9 individuals are presented in Figure 1 for air-conduction thresholds measured about 1 month prior to the hearing-aid fitting (unfilled symbols) and at the most recent postfit interval (filled symbols). The postfit hearing thresholds displayed for the top 3 participants are those obtained at the 3-year interval. For participant S056, whose audiogram is presented in the lower right panel of Figure 1, the postfit thresholds were those obtained at the 4-year postfit interval (which were equivalent to those obtained at the 3-year postfit interval). All other postfit audiograms in Figure 1 were obtained at the 2-year postfit interval. Although many participants revealed slight decreases in hearing sensitivity over the 2–4-year evaluation period, as would be expected with advancing age (Morrell, Gordon-Salant, Pearson, Brant, & Fozard, 1996), only 1 participant (S005) revealed a substantial decline. This change was confined to an ear with initially poorer hearing sensitivity, and the thresholds from the participant’s better-hearing ear did not change appreciably over time. Thus, because better-ear hearing sensitivity in S005 was stable and all unaided and aided speech-recognition measurements reported here were obtained binaurally in a sound-field setting, unaided and aided speech-recognition performance would be expected to remain stable.

**Hearing Aids**

The hearing aids in this study use linear circuits with output-limiting compression and Class D amplifiers. All were full-shell, in-the-ear (ITE) devices and included a telecoil switch on one instrument (determined by the wearer’s preference). Active tone (low-cut only) and output-limiting controls, adjustable select-a-vent venting, and wax guards were included on all devices. The volume-control wheels were marked by the manufacturer with a small white dot at the perimeter to provide a visual reference for its position and adjustment.

The clinician used real-ear insertion gain targets for the National Acoustic Laboratory (NAL-R) prescription formula (Byrne & Dillon, 1986; Byrne, Parkinson, & Newall, 1990) incorporated in the Hearing Aid Selection Program (HASP) fitting software (Version 2.07, distributed by NAL) and adjusted the settings of the controls and vent to achieve the closest match possible to target gain. The test signal was a 60-dB-SPL swept pure-tone signal generated from either Frye 6500 or Audioscan real-ear measurement equipment. Matching criteria were ±10 dB from 250 to 2000 Hz and ±15 dB at 3000 and 4000 Hz. All participants met these fairly broad criteria and most matches were considerably better.

Means and standard errors for the clinician-fit gain at the initial hearing-aid fitting and at the 3-year postfit interval are shown in Figure 2. All speech-recognition measures were obtained with the hearing aids adjusted to the clinician-fit setting, which remained stable in each ear over the 3-year evaluation period. In addition, the mean clinician-fit coupler gain represents a good match to the target coupler gain in each ear. Also plotted in Figure 2 are the as-worn gain settings at the first postfit interval (1-month postfit) and at the final (3-year) postfit interval. As noted previously (Humes, Barlow, Garner, & Wilson, 2000, 2002), as-worn gain is less than the clinician-fit gain. Interestingly, albeit for a smaller sample size, we found a greater shift in as-worn gain between the first and last postfit measurement than
we had observed previously following 1 or 2 years of hearing-aid use. Still, the shift in as-worn gain is on the order of only 3–4 dB and is not a large effect.

In summary, both the clinician-fit gain and as-worn gain are fairly stable over time. Therefore, any differences in observed benefit over time reported in this study are not likely attributable to changes in gain. This is important to document in studies of acclimatization, especially over long time periods (Horwitz & Turner, 1997).

### Procedures

A complete description of procedural details can be found in Humes et al. (2002). Briefly, 2 weeks following the initial fit of the hearing aids, all unaided measures of speech recognition were completed. A total of three unaided speech-recognition scores were obtained, one for each of three test conditions. In each condition, scores were obtained binaurally in the sound field. The
three test conditions considered here were (a) the CUNY Nonsense Syllable Test (NST; Levitt & Resnick, 1978) presented at an overall level of 65 dB SPL and +8 dB speech-to-noise ratio (SNR) in recorded multitalker babble (Kalikow, Stevens, & Elliot, 1977), (b) the Connected Speech Test (CST; Cox, Alexander, Gilmore, & Pusakulich, 1988) presented at an overall level of 50 dB SPL in quiet, and (c) the CST presented at an overall level of 65 dB SPL and +8 dB SNR in the recorded multitalker babble provided with the CST.

All speech materials were commercially available recorded versions. The full 102-item, 11-subtest version of the NST was used. Each score for the CST was based on two consecutive passages, with each passage containing 25 key words for scoring. Different forms of the NST and different passages of the CST were used for each condition. The speech signal for all speech-recognition measurements was presented from a loudspeaker (Radio Shack Optimus 7) located 1 meter in front of the participant at 0° azimuth and elevation. The noise competition was delivered from an identical loudspeaker located 1 meter behind the participant at 180° azimuth and 0° elevation. For the NST, the participant marked the syllable heard on a large-font answer sheet that contained the target stimulus and six to eight alternatives that differed from the target stimulus by only one phoneme. For the CST, the participant was provided with the passage topic and was encouraged to guess if uncertain about what she or he heard. Playback was paused after presentation of each sentence of the passage, the participant orally repeated what had been heard, and the experimenter scored the participant’s response for keywords within the passage. All binaural unaided measures of speech recognition were repeated at the 1-year, 2-year and 3-year postfit intervals.

Approximately 2 weeks later, the participant returned for the 1-month postfit follow-up visit. The hearing aids were examined, evaluated in the testbox and adjusted by the clinician as needed to return their function to the target levels from the initial fitting session. Next, aided speech-recognition measures were obtained for the three test conditions described previously. Test forms (NST) and passages (CST) not used previously were employed in this session. All aided testing was conducted binaurally.

In addition to aided speech-recognition measures, each participant completed, in the examiner’s presence, two pencil-and-paper surveys of subjective benefit. These were the Hearing Aid Performance Inventory (HAPI; Walden, Demorest, & Hepler, 1984) and the Hearing Handicap Inventory for the Elderly (HHIE; Ventry & Weinstein, 1982). These self-report benefit measures were also obtained at the 6-month, 1-year, 2-year, and 3-year follow-up sessions. Other self-report outcome measures also were completed by each participant over the 3-year postfit period (Humes, Wilson, Barlow, Garner, & Amos, 2002), including measures of hearing-aid usage. For the 9 participants in this study, the mean hours of hearing-aid usage per day throughout the first year was 9.8 hr, with a range of 4.5 to 15.7 hr across participants. Numerical estimates of the hours of daily hearing-aid usage were provided by the participants only during the first year of the study. However, additional self-report measures of relative hearing-aid usage (Humes, Garner, Wilson, & Barlow, 2001) were obtained throughout and these measures indicated that usage declined only slightly from the 1-year to the 3-year postfit interval.
Results and Discussion

The individual scores on the NST are presented in Figure 3 for each of the 9 elderly hearing-aid wearers. Each panel displays the speech-recognition score transformed into a “rationalized” arcsine unit (rau; Studebaker, 1985), on the ordinate, as a function of the postfit interval in months, along the abscissa. The 95% critical difference in raus was calculated for a test comprised of 102 items following Studebaker. The shaded region in each panel represents the 95%-critical-difference boundaries constructed around the aided NST score obtained at the 1-month postfit interval. In this figure, as well as in subsequent ones, the filled circles represent aided speech-recognition scores at each postfit interval and unfilled circles represent unaided speech-recognition scores. The filled circles lying outside the shaded boundaries represent aided NST scores that were significantly higher or lower than the aided NST score obtained at the 1-month postfit interval. All told, 5.6% of the aided NST scores at a later postfit interval were significantly better than those at the 1-month postfit interval, whereas 22.2% were significantly worse. Although the focus here is placed on improvements in aided performance over time, it is apparent that unaided performance (unfilled circles) is fairly stable over time as well.

The individual data for the CST, administered in quiet at 50 dB SPL, are shown in Figure 4 in a format similar to Figure 3. Individual scores, in rau, for the Nonsense Syllable Test (NST) at 65 dB SPL and a +8 dB signal-to-noise ratio. Both aided (filled circles) and unaided (unfilled circles) scores are shown at each measurement interval for which the scores were available. The shaded or cross-hatched rectangle in each panel provides the 95% critical-difference boundaries for the aided score obtained at the 1-month postfit interval. For each participant, scores represented by filled circles lying above or below this shaded area are those that were found to be significantly different from the score at the 1-month postfit interval.
identical to that used previously for the NST (see Figure 3). Because Cox et al. (1988) had demonstrated that the binomial distribution failed to describe accurately the measurement variability of the CST, a different approach was taken by Humes et al. (2002) to generate the 95% critical differences for this test. The resulting 95% critical difference for scores based on 50 keywords was found to be 32.2 rau. The shaded region in each panel of Figure 4 spans 32.2 raus above and below the aided CST score obtained at the 1-month postfit interval. Using these boundaries, 11.1% of the subsequent aided CST scores were significantly better than the score at the 1-month postfit interval and 2.8% were worse. It should be noted that, given the high aided scores at the 1-month postfit interval for 4 of the 9 participants (S005, S234, S119, and S056), together with the size of the 95% critical difference (32.2 rau), it was impossible for those participants to show significant improvements over time.

It is also notable that although most aided scores were fairly stable over time, those for S184 and, to a lesser extent S021, underwent large variations.

Figure 5 shows the individual data for all 9 participants for the CST administered at 65 dB SPL with a +8 dB SNR. The same 32.2 rau 95% critical difference used in Figure 4 for the CST has been applied here. A total of 13.9% of the data points at subsequent postfit intervals were worse than those at the 1-month postfit interval and 25% were better. In general, however, those participants who had significantly better performance after the 1-month interval usually returned to nonsignificant differences at later postfit measurement intervals. The results for S119 and S056 are typical of this pattern. Only the data for S545 support a steady improvement in aided (and possibly unaided) performance over time. For 5 of the 9 participants (S213, S071, S005, S021, and S234), however, the combination of a high aided score at the 1-month postfit interval and a 95%
critical difference of 32.2 rau made it impossible for them to show significant improvements in aided performance over time.

It is possible that researchers should not focus exclusively on the aided performance to determine the presence of acclimatization. It may be that acclimatization would lead to a consistently widening gap between aided and unaided scores over time, with not only the aided scores increasing but the unaided scores decreasing over time. As hearing-aid wearers acclimatize to amplified sound, their understanding of unamplified sound might also deteriorate such that the relative benefit, or the difference between aided and unaided performance would increase over time. Review of the individual data in Figures 3, 4, and 5 suggests that the data from one participant, S213, might be consistent with such an interpretation for two of the three test conditions. The aided scores for this participant, however, do not exceed the 95% critical differences for any condition.

Self-report measures of hearing-aid benefit from each of the four scales of the HAPI were evaluated using 95% critical differences from Walden et al. (1984). Specifically, the 95% critical differences were 0.42, 0.37, 0.52, and 0.68 scale unit on the speech-in-noise, speech-in-quiet, speech-without-cues, and non-speech subscales of the HAPI, respectively. The number of individual scores exceeding these critical differences, for each postfit interval relative to the initial measurements at the 1-month postfit interval, was determined for each scale of the HAPI and is shown in Table 1. The scores on the speech-in-noise and speech-without-cues subscales revealed the greatest number of changes, and the majority of these were consistent with a decline in subjective benefit over time. All told, 25 of the 144 possible
paired comparisons, or 17.4%, supported a significant change in HAPI subscale scores following the 1-month postfit interval. Of the 17.4% with significant changes in HAPI subscale scores, 11.1% were declines and 6.3% were improvements. Eight of the 25 significant changes in HAPI subscale scores were associated with 1 participant (S056). The pattern of change in this participant, however, was varied. Three of the four HAPI subscale scores for S056 improved significantly from the 1-month to the 6-month postfit interval followed by significant declines in scores on two of the four HAPI subscales from the 1-month to 2-year postfit intervals and on three of the four HAPI subscales from the 1-month to the 3-year postfit intervals. The factors underlying the fluctuation in HAPI subscale scores for this participant are unknown.

Finally, another self-report measure of subjective benefit compared aided and unaided HHIE scores. The unaided scores were obtained prior to the hearing-aid fitting and the aided scores were obtained at each postfit interval. Because there was only one unaided HHIE score for each participant, fluctuations in benefit over time can be examined by looking at only the aided HHIE scores at each postfit interval. Aided HHIE scores at subsequent postfit intervals were compared to those obtained at the 1-month postfit interval for each participant. Significant differences were established using a 95% critical difference of 19 scale units (Newman & Weinstein, 1989). The bottom row of Table 1 shows the number of significant changes in aided HHIE score over time. Four of the 36 possible paired comparisons, or 11.1%, were found to exceed the 95% critical difference. Among the 11.1% of the aided HHIE scores that had changed significantly, 8.3% were declines and 2.8% were improvements.

In summary, there was little evidence of acclimatization in either objective or subjective measures of hearing-aid performance and benefit for this small sample of 9 elderly hearing-aid wearers followed over a 3-year period. For both objective and subjective types of measurements considered in this study, there were just as many, if not more, significant declines in aided performance over time as there were improvements. This is consistent with the larger scale investigation published previously (Humes et al., 2002) in which 134 and 49 elderly hearing-aid wearers were followed for 1 and 2 years, respectively. The current study extends the postfit timeline to 3 years, over which few significant acclimatization effects have been observed. Perhaps more worrisome, however, is the more frequent observation of significant declines in performance over time, despite stability of hearing thresholds and hearing-aid gain. Still, typically 75%–90% of the between-interval comparisons, pooled across participants, did not differ statistically from those obtained at the 1-month postfit interval regardless of the outcome measure, with the lone exception being the aided speech-recognition scores for the CST at 65 dB SPL and +8 dB SNR. Even for this condition, however, 61% of the between-interval comparisons did not differ significantly when pooled across participants. Thus, change in performance or benefit over time, regardless of direction, appears to be the exception rather than the rule.

Acknowledgment

This research was supported, in part, by Research Grant R01-AG08293 from the National Institute on Aging.

References


Received February 13, 2002
Accepted June 25, 2002
DOI: 10.1044/1092-4388(2003/011)
Contact author: Larry Humes, Department of Speech and Hearing Sciences, Indiana University, 200 South Jordan Avenue, Bloomington, IN 47405-7002.
E-mail: humes@indiana.edu