Amplification and Aural Rehabilitation

Reliability and Intersubject Variability of the Real Ear Unaided Response

Michael Valente, PhD; Maureen Valente, MA; Joel Goebel, MD

The Washington University School of Medicine [M. V,. J. G.] Department of Otolaryngology and Department of Communication Disorders [Ma. V.], St. Louis University, St. Louis, Missouri

ABSTRACT

Intratester test-retest reliability of the real ear unaided response (REUR) was determined on 49 ears using the Frye 6500 real ear analyzer. Results revealed mean differences of less than 1 dB for repeat measurements at seven test frequencies between 250 and 4000 Hz. The average peak resonant frequency of the repeated measure was within 16 Hz of the initial measure. In addition, the intersubject variability of the amplitude of REUR was quite large. A range of 7 dB was found at 250 to 500 Hz with the range expanding to 15 to 20 dB at 2000 to 4000 Hz. Also, the peak resonant frequency varied between 2100-4800 Hz. These results are discussed in terms of those dispensers who use the REUR to "custom" order hearing aids (Ear Hear 12 3: 216-220).

EACH DAY, HEARING aids are selected having electroacoustic characteristics felt to be appropriate for a given hearing loss. Recently, Martin and Morris (1989) reported that selection of these characteristics is usually based upon a target of the desired real ear insertion response (REIR) for discrete frequencies recommended by Berger, Hagberg and Rane (1977), Byrne and Dillon (1986), Libby (1985; 1986), or McCandless and Lyregaard (1983). However, the selection of the characteristics may also be based upon a target of the desired real ear aided response (REAR) for discrete frequencies recommended by Cox (1988), Pascoe (1975), Seewald, Ross and Spiro (1985) or Skinner (1980).

Recently, several investigators (Mueller, 1989; Upfold & Bryne, 1988; Valente, Valente & Vass, 1990a; 1990b) have suggested that the real ear unaided response (REUR) of the individual at discrete frequencies should be included in the hearing aid selection process in order to accurately determine the required electroacoustic characteristics necessary to achieve desired REIR. This suggestion is based upon the belief that the natural resonance of the ear canal is eliminated when an earmold or hearing aid is inserted in the ear canal and therefore, significant deviations of the individual REUR from the average REUR may result in difficulty in achieving desired REIR. For example, if the amplitude of the REUR is greater than average, a reduction in the measured REIR (i.e., insertion loss) may occur within that frequency region, and greater coupler gain may be required to obtain desired REIR. On the other hand, if the amplitude of the REUR is less than average, measured REIR may be greater than desired and less coupler gain may be required. Either effect could result in undesirable peaks and troughs in the measured REIR if corrections are not implemented.

In determining individual deviations of the REUR from average, some investigators have used the free-field to eardrum transformation data reported by Shaw (1974) and Shaw and Vaillancourt (1985). In addition, software included in some probe tube units contains the Shaw data as a reference for corrections for the individual REUR. However, the data reported by Shaw is an average of a compilation of 12 investigations in which the REUR was obtained in a manner which, in many respects, is significantly different from the way the REUR is measured with many probe tube units. For example, Shaw used the center of the head in an unobstructed free-field as the reference after the REUR was measured with a probe microphone in the ear canal. The use of the center of the head as the reference results in the inclusion of head diffraction and body baffle effects in the measured REUR. On the other hand, many probe units use an "at the ear" or "under the ear" location for the reference microphone position. The use of this reference point for "equalizing" or "leveling" the test condition excludes head diffraction and body baffle effects from the REUR measure. Exclusion of these effects can result in the measured REUR being different from the Shaw (1974) data from −0.5 to 4 dB (Kuhn, 1979). Bentler (1989), using a Rastronics CCI 10/3 (under the ear reference microphone) at 0° azimuth, reported the REUR in children above the age of 2 yr was reasonably close to the findings of Shaw. However, the mean REUR revealed slightly less gain in the lower and upper frequency regions,

0196/0202/91/1203-0216$03.00/0 • EAR AND HEARING

which she attributes, in part, to the calibration method used.

In addition to measuring the REUR for the purpose of "custom" ordering hearing aids at discrete frequencies, several investigators (Kruger, 1987; Upfold & Bryne, 1988) have suggested that greater user benefit or acceptance may occur if the peak gain of the REAR matched the peak frequency of the REUR to obtain a "transparent" fit. Currently, several hearing aid manufacturers provide a coupler response which mimics the average REUR. This response should be effective in compensating for the loss of the ear canal resonance if the individual REUR has the same peak amplitude and frequency as the average REUR. However, large intersubject variability of these REUR parameters has been reported in the literature (Bentler, 1989; Upfold & Bryne, 1988). This suggests that the effectiveness of this circuit design may be of limited use if a listener has an expected REUR which differs significantly from average. It is expected that technological advances in the near future will result in the ability to select peak coupler gain to match the individual REUR measured in the hearing aid selection process. In addition, these same technological advances may soon allow the dispenser to actively shift the peak frequency during REAR probe measures to match the peak frequency of the individual REUR.

Due to an increased interest in the use of the REUR in the hearing aid selection process, it would be beneficial to determine the test-retest reliability of the REUR for one commercially available unit (Frye 6500). More dispensers seem to be using the REUR as a correction factor to determine the coupler response necessary to achieve desired REIR. Also, dispensers may consider using the REUR as a reference of where to adjust or select the peak frequency of a hearing aid. If the REUR measurement is unreliable (i.e., significant differences are present between measures) then using the REUR to customize hearing aids or for matching the peak frequency may not be appropriate. On the other hand, it would be comforting to know that the REUR is a reliable measure for those who choose to use this measure in the hearing aid selection process.

METHODS

Subjects

Forty-nine ears of 25 subjects (mean age = 64.4 yr; S.D. = 13.7 yr) were included in this study. Tympanograms were within normal limits for pressure (daPa) and amplitude (ml) and no ear had a history of otologic surgery. Excessive cerumen was not present during otoscopic observation before measuring the REUR.

REUR Measures

Subjects sat in a double-walled sound suite directly facing a loudspeaker (Radio Shack Minimus, 3.5 in.) connected to a Frye 6500 real ear analyzer. The loudspeaker was placed at ear level 12 in. from the subject as recommended by the manufacturer.

The reference and probe microphones were placed on the test ear and held in place via a Velcro headband (reference microphone) and "earhanger" (probe microphone). The front of the reference microphone was placed over the apex of the pinna and directly faced the loudspeaker. The soft silicon probe tube from the probe microphone was marked 20 mm from the tip. The probe was placed in the ear canal so that the red mark was slightly medial to the orifice of the ear canal and taped in place to prevent movement. Assuming the average adult ear canal is 25 mm in length, the end of the probe tip was estimated to be within approximately 4 mm from the eardrum of the average subject.

It should be noted that this procedure places the probe, for all subjects, at the same distance from the orifice of the ear canal. However, due to intersubject differences in actual canal length, the distance from the probe to the eardrum probably varied among subjects. Gilman and Dirks (1986) and Chan and Geisler (1990) report that measured SPL at probe positions as far as 12 mm from the eardrum may be as much as 4 dB less at higher frequencies relative to a probe position close to the eardrum. As noted by Bentler (1989) "wide intersubject variability of resonance amplitude may be related, in part, to the small, although significant, difference in probe-to-eardrum distance differences among subjects (p. 286)."

With microphones in place, the frequency response of the loudspeaker, as measured by the reference microphone on the subject's head, was leveled according to manufacturer instructions before each measurement of the REUR. The leveling process is rather quick, using a 2.5-sec burst of a flat spectrum composite signal. Finally, both microphones were calibrated daily according to the manufacturer's instructions.

The following measures were obtained using speech-weighted composite noise presented at an overall level of 70 dB SPL with a duration of 1 to 2 sec. First, an REUR was generated for each ear while the subject focused on an orange dot placed above the cone of the loudspeaker when the signal was introduced. This graphic response was then converted to numeric data by the microprocessor of the Frye 6500 for future statistical analysis. From this printed output, the peak resonant frequency was defined as the frequency (in Hz) corresponding to the greatest numeric value of probe microphone SPL versus reference microphone SPL. The peak amplitude was defined as the amplitude (in dB) recorded at the peak frequency. This same procedure was repeated by the same examiner approximately 1.5 to 2 weeks later.

RESULTS AND DISCUSSION

Test-Retest Reliability of the REUR at Discrete Frequencies

The mean REURs for the initial and repeat measures are reported in Table 1 for each of the seven discrete frequencies (250, 500, 1000, 1500, 2000, 3000 and 4000 Hz). For 250 Hz, the amplitude was determined by averaging the amplitude at 200 and 300 Hz. Also provided are the SD, ranges, SE of the mean, *t*-tests and Pearson product correlation coefficients at each test frequency and for each measure. As can be seen, the difference between the means for each measure is less than 1 dB at all test frequencies. Although these differences were rather small, a two-tailed *t*-test of paired comparisons revealed the differences at 500, 1000 and 3000 Hz were significant ($p < 0.01$). Pearson product

Table 1. Mean REUR (dB) for the initial and repeated measure for 49 ears at the seven discrete test frequencies. Also reported is the standard deviation, range, standard error of the mean across subjects, mean difference between the two measures, Pearson product correlation coefficients and the *t*-test of paired comparisons at each frequency.

	Frequency (Hz)						
Condition	250	500	1000	1500	2000	3000	4000
Mean REUR (1)	3.4	4.5	4.2	5.8	13.2	14.7	10.4
SD	1.2	0.9	1.7	2.6	3.1	3.9	3.9
Range	1–7	2–7	1–8	0–11	5–19	4–23	3–19
SE	0.2	0.1	0.2	0.4	0.4	0.6	0.6
Mean REUR (2)	3.7	4.9	4.9	5.9	13.6	15.6	11.2
SD	0.9	1.0	1.9	2.4	3.2	3.6	4.2
Range	1–7	3–8	2–9	0–11	7–21	8–23	3–22
SE	0.1	0.1	0.3	0.3	0.5	0.5	0.6
Difference	0.3	0.4	0.7	0.1	0.4	0.9	0.8
Grand mean	3.6	4.7	4.6	5.8	13.4	15.2	10.8
Correlation	0.55*	0.30**	0.32**	0.70*	0.78*	0.81*	0.65*
t-test	−1.79	−2.71*	−2.52*	−0.29	−1.15	−2.76*	−1.59

Note: (*) $p < 0.01$; (**) $p < 0.05$.

Table 2. Mean peak amplitude (dB) and peak resonant frequency (Hz) for the initial and repeated measure of the REUR for 49 ears. Also reported is the standard deviation, range, standard error of the mean across subjects, mean difference between the two measures, Pearson product correlation coefficients and the *t*-test of paired comparisons at each frequency.

Condition	Peak Amplitude (dB)	Resonant Frequency (Hz)
Mean REUR (1)	18.1	2593
SD	3.3	463.4
Range	11–24	2100–4800
SE	0.5	69.0
Mean REUR (2)	18.9	2577
SD	3.3	326.8
Range	11–25	2100–3500
SE	0.5	48.7
Difference	0.8	16.0
Grand mean	18.5	2585
Correlation	0.77*	0.46*
t-test	−2.62*	0.25

Note: (*) $p < 0.01$.

correlations were poorer at the lower frequencies, but improved at the higher test frequencies.

Table 2 discloses the peak frequency and the corresponding peak amplitude for the initial and repeat measure. Also shown are the SD, range, SE of the mean, *t*-test and Pearson product correlation coefficients at each test frequency. Results reveal a mean amplitude of 18.5 dB and a mean peak frequency of 2585 Hz. Mean test-retest differences were 0.8 dB for peak amplitude and 16 Hz for the peak frequency. Although not shown, the repeated measure of the peak frequency was equal to the initial measure in 34% of the cases. In addition, the repeated measure of the peak frequency differed from the initial measure by 100 Hz in 23% of the cases and by 200 Hz in an additional 25% of the cases. In all, the repeated measure of the peak frequency was within 300 Hz of the initial measure in 93% of the measures. This is important because it indicates that the measured peak frequency for the same listener is stable over time and, therefore, the dispenser can order, with a fair degree of confidence, a hearing aid whose coupler response matches the peak frequency of the measured REUR to obtain a desired transparent fit.

Again, although the mean difference between measures of the peak amplitude were less than 1 dB, this difference was found to be significant ($p < 0.01$) using the two-tailed *t*-test of paired comparisons. Pearson product correlations between the initial and repeat measures of the peak amplitude and peak resonant frequency were significant ($p < 0.01$).

These findings are in good agreement with previous studies. Upfold and Bryne (1988), using a Rastronics CCI 10/3 under unknown measuring conditions, revealed a mean peak amplitude of 18 dB and a peak frequency of 2968 Hz. However, the method used by Upfold and Bryne to determine peak frequency was quite different from the procedure used in this study. They defined the peak frequency as the midpoint between the upper and lower frequencies which were 6 dB down from the frequency having the highest value. Bentler (1989), also using the Rastronics CCI 10/3 under free-field conditions, reported a mean peak amplitude of 18.9 dB and mean peak frequency of 2849 Hz in children ranging in age from 40 to 164 mo. Kruger (1987), using a noncommercially available system under diffuse-field conditions, reported a mean peak frequency of approximately 2700 Hz for children who were older than 2 yr. As reported by Kuhn (1979) and Shaw (1980), the amplitude and the peak frequency of the measured REUR can be slightly lower for diffuse measures when compared to free-field measures.

Often, mean differences between measures do not accurately reflect the variability present between repeated measures. Table 3 presents the average SD of the absolute test-retest differences in the REUR as well as the 95% confidence interval at each frequency. At all frequencies the average SD of the test-retest differ-

Table 3. Average standard deviation of the absolute differences between the initial and repeated measure of the REUR as well as the 95% confidence interval for the 49 ears.

Frequency (Hz)	SD of Difference	95% Confidence Interval (dB)
250	1.04	2.04
500	1.21	2.37
1000	2.10	4.12
1500	1.94	3.80
2000	2.11	4.13
3000	2.33	4.56
4000	3.41	6.68
Mean	2.02	3.95

ences was less than 3 dB (with the exception of 4000 Hz) and the grand mean was 2.02 dB. In addition, the 95% confidence interval was as small as 2 dB at 250 Hz to as great as nearly 7 dB at 4000 Hz. Figure 1 illustrates the test-retest variability of the REUR. For example, at 250 Hz, no differences were revealed between the initial and repeat measure in nearly 37% of the cases. Differences between measures at this frequency did not exceed 2 dB in 100% of the cases. By comparison, at 4000 Hz, no differences between measures were present in only 12.2% of the cases and differences did not exceed 5 dB in 91.8% of the cases. In addition, test-retest differences did not exceed ±3 dB in over 80% of the comparisons between 250 and 3000 Hz and nearly 90% of the comparisons did not exceed ±4 dB between 250 and 3000 Hz.

Intersubject Variability

As has been described in the past (Bentler, 1989; Kruger, 1987; Upfold & Byrne, 1988), the intersubject variability in the amplitude of the REUR can be rather large. The increased intersubject variability is especially true as frequency increases. This is reflected in the data of Tables 1 and 2 by the presence of larger SD and ranges across subjects for the initial and repeat measure of the REUR as frequency increases. Figure 2 illustrates the range of the lowest and highest amplitude of the REUR measured at each of the seven test frequencies as well as the same data for the peak amplitude and peak frequency for the 49 ears. As can be seen, the range was as small as 7 dB at 250 to 500 Hz to as great as 20 dB at 3000 to 4000 Hz; the range was 15 dB for the peak amplitude. Also, as indicated in Table 2, the range of the peak frequency was between 2100 and 4800 Hz. Only one subject revealed a peak frequency as high as 4800 Hz. If that subject were removed from this study, the upper range of the peak frequency would have been 3500 Hz. These results generally agree with Bentler (1989), who reported a range of 17 dB in the peak amplitude and a range of 1774 to 4039 Hz in the peak frequency and with Upfold and Bryne (1988), who reported a range of 13 dB in the peak amplitude. In addition, these ranges are in general agreement with Chan and Geisler (1990), who report a range of 6 to 25 dB in REUR amplitude at various test frequencies at the distance of the probe increases from the eardrum from 2 to 12 mm. As mentioned earlier, the probe was always 20 mm from the orifice of the ear canal, but there may have been small, but significant differences in the distance of the probe from the eardrum across the 49 ears. Intersubject differences in the distance of the probe from the eardrum, as well as intersubject differences in head diffraction and body baffle effects may account, in part, for the large intersubject variation revealed in the amplitude of the REUR.

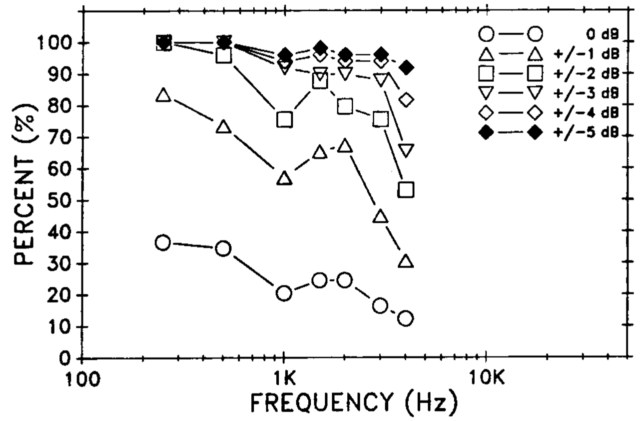

Figure 1. Difference (in dB) between initial and repeat measurements of the REUR at seven test frequencies (reported as percent).

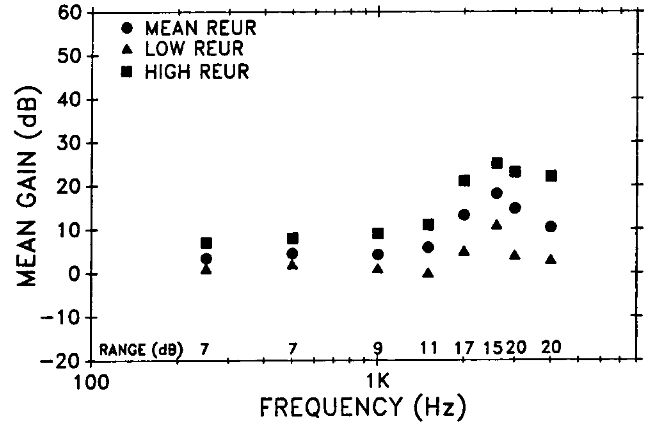

Figure 2. Range of the REUR measured in 49 ears.

SUMMARY

The results of this study suggest that the test-retest reliability of the REUR, when measured with the Frye 6500, is rather good with mean differences not exceeding 1 dB in amplitude and 16 Hz in the peak frequency. In addition, it is reassuring to know that the peak frequency was within 300 Hz of the initial measure for the same subject in 93% of the cases. This suggests that users of the Frye 6500 who utilize the REUR to customize ordering of hearing aids for a patient are using a rather reliable measure in which to make the necessary corrections for significant deviations in the REUR from average. The wide intersubject variability in both the amplitude and peak frequency of the REUR once again focuses on the importance of obtaining this measure on the individual instead of using values from averaged group data.

This large intersubject variability places some question on the universal acceptance of hearing aid circuits mimicking the average REUR. Upfold and Bryne (1988) expressed this concern in regards to the design of uniform attenuation earplugs as well as hearing aid circuits mimicking the average REUR. In this study, although the average test-retest difference of the peak frequency of the REUR was 16 Hz, the range of the peak frequency was 2100 to 3500 Hz (4800 in one ear). If a dispenser assumes the patient has an average REUR and orders a circuit mimicking the average REUR to obtain a transparent fit, there is the possibility that the actual REUR of the patient may be considerably different than the average REUR. This difference could result in the measured REIR being significantly higher or lower than was anticipated when compared to the desired REIR. For example, in this study only 12% of the measured peak frequencies occurred in the frequency region (2700 to 2800) where many manufacturers place the peak gain to mimic the REUR. This concern will be reduced when dispensers have the technology available to actively shift the peak gain of the measured REAR to match the measured REUR.

Finally, the results of this study are relevant only to the Frye 6500 under the conditions specified under "Methods." Additional studies are needed to determine the test-retest reliability of the REUR and other real ear measures for additional commercially available units as well as the various signals (pure tones, complex noise, clicks, speech-weighted noise) and azimuths (0° and 45°) typically utilized in clinics around the world.

REFERENCES

Bentler R. External ear resonance characteristics in children. J Speech Hear Disord 1989;54:264–268.

Berger K, Hagberg N, and Rane R. Prescription of Hearing Aids. Kent: Herald Publishing House, 1977.

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

Chan J and Geisler C. Estimation of eardrum acoustic pressure and ear canal length from remote points in the canal. J Acoust Soc Am 1990;87:1237–1247.

Cox R. The MSU hearing instrument prescription procedure. Hear Instrum 1988;39:6,8,10.

Gilman S and Dirks D. Acoustics of ear canal measurement of eardrum SPL in simulators. J Acoust Soc Am 1986;80:783–793.

Kruger B. An update of the external ear resonance in infants and young children. Ear Hear 1987;8:333–336.

Kuhn G. The pressure transformation from a diffuse sound field to the external ear and to the body and head surface. J Acoust Soc Am 1979;65:991–1000.

Libby E. State-of-the-art of hearing aid selection procedures. Hear Instrum 1985;36:30–38,62.

Libby ER. The ⅓-⅔ insertion gain hearing aid selection guide. Hear Instrum 1986;37:27–28.

Martin F and Morris L. Current audiologic practices in the United States. Hear J 1989;42:25–33,36–44.

McCandless G and Lyregaard P. Prescription of gain/output (POGO) for hearing aids. Hear Instrum 1983;34:16–21.

Mueller HG. Individualizing the ordering of custom hearing aids. Hear Instrum 1989;40:18,20,22.

Pascoe D. Frequency responses of hearing aids and their effects on the speech perception of hearing-impaired subjects. Ann Otol Rhinol Laryngol 1975;84(Suppl 23):1–40.

Seewald R, Ross R, and Spiro M. Selecting amplification characteristics for young hearing-impaired children. Ear Hear 1985;6:48–53.

Shaw AG. The acoustics of the external ear. In Studebaker GA and Hochberg I, Eds. Acoustical Factors Affecting Hearing Aid Performance. Baltimore: University Park Press, 1980:109–125.

Shaw AG. Transformation of sound pressure level from the free field to the eardrum in the horizontal plane. J Acoust Soc Amer 1974;56:1848–1861.

Shaw AG and Vaillancourt MM. Transformation of sound-pressure level from the free field to the eardrum presented in numerical form. J Acoust Soc Amer 1985;78:1120–1123.

Skinner MW. Speech intelligibility in noise-induced hearing loss: Effects of high-frequency compensation. J Acoust Soc Amer 1980;67:306–317.

Upfold G and Byrne D. Variability of earcanal resonance and its implications for the design of hearing aids and earplugs. Aust J Audiol 1988;10:97–102.

Valente M, Valente M, and Vass W. (in press). Use of real-ear measures to select the gain and output of hearing aids. Semin Hear 1990a.

Valente M, Valente M, and Vass W. Selecting an appropriate matrix for ITE and ITC hearing aids. Hear Instrum 1990b;41:20–22,23–24.

Acknowledgments: The authors would like to thank Larry Revit, M.A., Ruth Bentler, Ph.D., and an unknown reviewer for their comments and suggestions, which have resulted in a stronger paper. We would also like to thank Deborah Smith, Ph.D., from Biostatistics at Washington University School of Medicine for her help in the statistical analysis used in this paper.

Address reprint requests to Michael Valente, Ph.D., Washington University School of Medicine, Department of Otolaryngology, Division of Audiology, 517 S. Euclid Ave., St. Louis, MO 63110.

Received May 7, 1990; accepted January 25, 1991.