Comparison of speech intelligibility in quiet and in noise after hearing aid fitting according to a purely prescriptive and a comparative fitting procedure

Mick Metselaar · Bert Maat · Pieta Krijnen · Hans Verschuure · Wouter Dreschler · Louw Feenstra

Received: 23 August 2007 / Accepted: 21 January 2008 / Published online: 2 February 2008
© The Author(s) 2008

Abstract We compared two different types of hearing-aid fitting procedures in a double-blind randomized clinical study. Hearing aid fittings based on a purely prescriptive procedure (the NAL-RP formula) were compared to a comparative fitting procedure based on optimizing speech intelligibility scores. Main outcome measures were improvement of speech intelligibility scores in quiet and in noise. Data were related to the real-ear insertion responses that were measured after fitting. For analysis purposes subgroups were composed according to degree of hearing loss, characterized by unaided speech intelligibility in quiet, previous experience with hearing aids, unilateral or bilateral fittings and type of hearing aid. We found equal improvement of speech intelligibility in quiet, while fitting according to the prescriptive formula resulted in a somewhat better performance as expressed by the speech-to-noise ratio in comparison to the comparative procedure. Both procedures resulted in comparable real-ear insertion responses.

Keywords Fitting hearing aids · Hearing impairment · Speech intelligibility · Speech-in-noise test

Introduction

Within the actual process of hearing aid fitting a selection and evaluation phase can be distinguished [1], although the degree of distinction between these phases will vary according to the fitting procedure. A prescriptive formula initially selects a hearing aid according to some target characteristic, which is usually derived from psychophysical measurements (e.g. pure-tone audiometry or loudness scaling). In a comparative procedure, where the hearing aid is selected by comparison according to some criterion (e.g. speech intelligibility, sound quality), the selection process has a more iterative character. Here, the selection and evaluation phase are more closely linked.

Prescriptive procedures can easily be automated and offer a quick and reproducible method for the initial hearing aid selection. However, although the design of some of these procedures are based on speech intelligibility data, a conscientious implementation of this approach implies that pure-tone thresholds or loudness data directly or indirectly entail all the information required to alleviate hearing impairment, including psychophysical factors like spectral and temporal resolution and ecological factors like lifestyle and acoustics. A comparative procedure principally approximates the primary criterion chosen as close as possible and offers direct clinical evaluation with the hearing aid in place. However, this way of fitting could be expected to be more time-consuming and to be dependent on the knowledge and experience of the hearing aid fitter.

While the number of prescriptive formulae is gradually increasing, little is known about the quality and efficiency...
of this kind of fitting procedures in comparison with a comparative fitting approach which seems to steadily lose popularity. Quality can be defined as the extent to which a fitting procedure will succeed in alleviating limitations in performance, suffered by an individual or a population group and encourage participation in society. Efficiency has to do with the amount of labour, knowledge and money that needs to be invested in an optimum hearing-aid fitting procedure. Comparison of these kinds of fitting procedures in a clinical setting can reveal differences concerning these aspects.

We designed a prospective double-blind randomized clinical trial. Aim of the study was to compare the quality of a prescriptive hearing aid fitting procedure with a comparative method. Prescriptive fittings were carried out exactly according to the NAL-RP method, as this is one of the best-known and extensively documented and validated prescriptive procedures for linear amplification [2–5]. We based the criterion of the comparative procedure on speech intelligibility tests, more or less according to the procedure described by Carhart [6]. Although these measurements are relatively time-consuming, they can provide useful information since improving the intelligibility of speech is one of the main goals in hearing rehabilitation. We defined quality in terms of improvement of speech intelligibility score in quiet and in noise. Insertion responses from the two fitting procedures were compared and related to changes in speech perception data.

Materials and methods

Population

Hearing-impaired patients were primarily recruited from the audiological centers of two university hospitals. Experienced as well as inexperienced hearing aid users were included after informed consent was obtained. Mean pure-tone audiometric thresholds (1, 2 and 4 kHz) had to be at least 35 dB HL at one ear (insurance company criterion for partly reimbursing the expenses in the Netherlands), predominantly or entirely sensorineural. Exclusion criteria were:

- maximum unaided speech score less than 50% at the best ear,
- suspicion of retrocochlear cause of hearing loss,
- Meniere’s disease (active phase),
- (severe) tinnitus,
- significant co-morbidity.

Patients that were included in the study were stratified according to unaided maximum speech intelligibility score in quiet measured at the better ear. Stratification was done in an attempt to achieve a more equally distributed population with respect to the range of hearing impairments. Patients were stratified according to maximum speech intelligibility score at the better ear. Three strata were distinguished: a lower stratum containing speech scores between 50 and 74%, a middle stratum between 75 and 89%, and a high stratum with scores at and beyond 90%.

The number of clients necessary was calculated in a power analysis. Based on a former pilot study (unpublished) we assumed an improvement in aided versus unaided speech intelligibility scores at 65 dB SPL of 15, 11 and 5% in the lower, middle and higher stratum, respectively. For calculation of group-sizes a student t-test was applied with a standard significance (p value < 0.05). When the power is set at 80% the clinical relevant differences between improvement in speech intelligibility for both fitting procedures is 6, 4 and 1.25% in the three strata. These should contain 124, 176, and 180 patients, respectively, requiring a total number of 480 participants.

General procedure

Standard pure-tone audiometry was performed with the Madsen OB-822 clinical audiometer and TDH-39 earphones. Speech audiometry was performed with the same equipment for each ear separately. Lists of 11 phonetically balanced CVC-words [7, 8] were offered at 10 dB intervals. Each consonant or vowel added 3% to the total score.

Hearing aid fittings were carried out according to the NAL-RP formula and the comparative procedure as well. A detailed description of these procedures will be given in the next paragraphs. After inclusion each patient was fitted according to the two hearing aid selection and evaluation procedures in succession, each carried out by a different hearing-aid fitter. Both fitters were not informed about each others results, except for the type of hearing aid prescribed (BTE or ITE) and unilateral or bilateral fitting which was kept the same in both prescriptions in order to keep the procedure masked to the patient. Unilateral as well as bilateral fittings were carried out, depending on the user’s hearing and preference. All hearing aids used in the study had analogue electrical circuits and were adjusted to linear amplification. No digital circuits and/or WDRC compression algorithms were used as clear fitting procedures for these hearing aids and amplification-mode were emerging at the time of the study and would make it impossible to apply the strict design of the project.

The hearing aids selected according to both fitting procedures were specified in a prescription, which also included indications concerning the tone settings and the desired type of earmould and earhook.

One of either prescriptions was randomly selected by an independent person and given to the patient who subsequently consulted the hearing-aid dispenser for conveyance.
of the hearing aid and earmould according to the specifications of the prescription. The patient was unaware of the type of hearing-aid fitting that was selected.

A 12-week period of rehabilitation and experience followed during which the patient was able to get used to the sound and wearing of the aid. Evaluation of the hearing aid was performed once in the middle of this period (after 6 weeks) in order to optimize its setting. At the end of the try-out period (12 weeks), final assessments were performed by a researcher who was not aware of the type of procedure the patient had been fitted with. The measurements consisted of measurements of aided and unaided speech intelligibility in quiet and in noise and real-ear measurements. Several self-report questionnaires had to be completed also at the beginning ($t = 0$), halfway ($t = 6$ weeks) and at the end of the acclimatization period ($t = 12$ weeks).

The blinding ended after having assessed whether the hearing-aid has been satisfactorily fitted or not. This was done according to audiological and client criteria. For approval on audiological grounds, the speech intelligibility in quiet with hearing aid(s) had to be equal or better than the maximum speech score measured in the speech audiogram before fitting. The patient could also indicate whether (s)he was satisfied with the result. NAL-RP prescriptions could then be changed and optimized according to the comparative procedure if necessary. This was required by the Medical-Ethical Committee of the participating hospitals and ensured that patients were provided with at least the same care as when they would not have participated in the study.

NAL-RP fitting

Prescriptive hearing aid fittings were strictly carried out according to the NAL formula [3] with the modification for profound hearing losses [4]. Corrections for an air-bone gap were performed by adding 25% of the difference between the air and bone conduction thresholds to the gain at each specified frequency [9]. Ear canal characteristics were taken into account by measuring the open ear response (“real-ear unaided response”) to correct for the standard coupler response (IEC 126) [10] by the individual real-ear to coupler difference. This procedure has been described by Cornelisse et al. [11].

Hearing aid selection was performed by means of a computer program, that has been exclusively designed for this study. Coupler responses of all hearing aids available in our centers with different settings of tone-controls had been measured in advance on a 2cc coupler [12] using a PortaRem-2000 (RD Rastronics Division, Denmark). These had been stored in the database of the program. The actual selection process consisted of matching the calculated target again of the patient with all coupler responses in the database. The hearing aid that was able to generate a response most similar to the target gain was selected. Correction factors for open ear response and type of earmould were also included in the selection program, the exact type of hearing aid and tone settings and also the specifications of the earmould were prescribed.

After the specified hearing aid had been delivered to the patient, it was adjusted as close as possible to the NAL-RP target real-ear insertion response. This was done by real-ear measurements with the hearing aid and earmould in the ear.

Comparative fitting procedure

The comparative fitting procedure that was used in this study has been described in detail by Verschuure [13]. The aim is to improve speech perception as much as possible; to at least the maximum speech intelligibility found in the (unaided) speech audiogram. Hearing aid selection was therefore performed in a comparative procedure in which evaluation of speech intelligibility in quiet with each of the selected hearing aids in the ear was used as the primary selection criterion. A second criterion was used, based on sound quality judgments by the patient.

After a 6-week period of initial acclimatization to the sound and the wearing of the hearing aid, evaluation and, if necessary, adjustment of hearing aid settings was performed in order to optimize speech intelligibility.

Final assessments were done after a second 6-week period of rehabilitation and experience. Hearing aid fitting was considered finished when both the hearing aid fitter and the patient were satisfied with the result. In case of an unsatisfactory result after 6 or 12 weeks, re-selection of hearing aids took place, which was again followed by fine-tuning and acclimatization.

Outcome measures

The following primary outcome measures were defined:

- improvement of speech intelligibility scores in quiet. Speech intelligibility scores were measured in a free-field condition at 55, 65 and 75 dB SPL, using the recorded NVA lists, each containing 11 CVC-words. These were presented through a loudspeaker at a distance of 1 m from the patient in a sound-treated booth with a reverberation radius of about 1.5 m. Correctly reproduced consonants and vowels (33 for each list) were scored as a percentage score. The aided speech score used for analysis was the highest speech score at one of the levels 55, 65 or 75 dB SPL. Unaided intelligibility was defined as the highest speech score for any sound level measured at the better ear in the speech audiogram.
Results

Population

We were able to include 254 hearing impaired patients in a 3-year period: 92 men (36%) and 162 women. Age ranged from 29 to 95 years with a mean age of 71 years (SD 13.5 years). Mean pure tone audiogram thresholds were 57.5 dB HL and ranged from 30.6 to 102.5 dB HL. Speech reception threshold (SRT) ranged from 11.4 to 94.6 dB with a mean of 53.2 dB. Detailed data are shown in Table 1. The results of all participating centers were comparable.

Speech in quiet

Due to missing data (profound losses; limited amplification at testing) in three of the 184 successful hearing aid fittings, we were able to calculate results for 181 clients. On the whole, speech intelligibility after hearing aid fitting improved to the same extent for both fitting procedures with 6%.

Because maximum possible improvement of the aided speech intelligibility score was determined by the unaided performance, largest improvements were found in the lower stratum: 21% (median), while for the middle and upper stratum 11 and 3% improvement was found. Data are shown in Fig. 1. We found no significant differences between the two fitting procedures. As can be read from Fig. 2, improvements were equal for inexperienced and experienced hearing-aid users in all strata.

Analysis on subgroups (inexperienced and experienced hearing-aid users, unilateral and bilateral fittings, ITE and BTE-fittings) did not show any significant difference between the prescriptive and comparative fitting procedures, except for ITE-fittings, where a significantly larger median improvement in speech intelligibility was found for the prescriptive fitting procedure: 9% compared to 0% for the comparative method. Despite of the small number of ITE fittings, this improvement was significant (p = 0.002; t-test).

Table 1 General features of the study population and numbers for the three strata

| Stratum   | H.a. user | H.a. fitting | Type of h.a. | Sex |
|-----------|-----------|--------------|--------------|-----|
|           | Inexp.    | Exp.         | Unilat.      | Bilat. | BTE | ITE | Male | Female |
| 50–74%    | 7         | 27           | 10           | 24       | 33  | 1   | 20   | 14    |
| 75–89%    | 37        | 42           | 12           | 67       | 71  | 8   | 57   | 22    |
| 90–100%   | 71        | 70           | 36           | 105      | 112 | 25  | 86   | 55    |
| Total     | 115       | 139          | 58           | 196      | 216 | 34  | 163  | 91    |
We concluded that improvement of the intelligibility score of speech in quiet was the same for both NAL-RP fitting procedure and was proportional to the degree of hearing loss characterized by unaided intelligibility of speech in quiet.

Speech in noise

Changes in S/N ratio were calculated for 132 subjects, which is substantially less than the 181 subjects for which we were able to calculate the improvement of the speech scores in quiet. This was due to the fact that speech-in-noise measurements were obtained at a level of 20 dB above the speech reception threshold (SRT). In a number of cases our equipment was not able to deliver the required stimulation level for unaided scores (maximum output level: 100 dB SPL).

When combining all comparative fittings (three strata), no improvement in S/N ratio after hearing aid fitting was found in this group. However, in the prescriptive group, a median improvement of 0.80 dB was found. Although this improvement seemed to be small, it was significantly better when compared the outcome to the comparative fitting group ($p = 0.002$).

The median S/N ratios for the two fitting procedures in all three strata are depicted in Fig. 3. Median improvement turned out to be in favour of the prescriptive procedures in all three strata, although statistical significance could only be proven in the middle stratum ($p = 0.03$).

No significant differences were present in the lower stratum due to the small number of subjects and in the upper stratum due to the small effect in spite of it is relatively large group size. The difference was significant in the middle stratum only ($p = 0.03$) and was in favour of the prescriptive procedure.

No significant differences in improvement of S/N ratios were found in the distinguished subgroups. Data are shown in Fig. 4.

Real-ear insertion response

We found a comparable slope of the real-ear insertion gain measured after fitting according to either procedure. This was also the case for each of the three strata separately (ANOVA; $p > 0.1$). No differences in slope were found in
comparable and significant correlation between the slope of the audiogram and real-ear insertion gain was found in the prescriptive and comparative fitting subgroups (Pearson correlation 0.309; p < 0.001). A scatter plot is depicted in Fig. 5. This was also found in all three strata separately.

We concluded that both fitting procedures are comparable with respect to the slope of the frequency response prescribed for a population of different sloping and varying degrees of hearing losses.

Further investigation was carried out in order to correlate real-ear insertion responses to speech in noise data. This was done in order to search for any relationship between the improvement of the S/N ratio and the amount of high frequency amplification provided. Here, it must be realized that the insertion gains were measured for each fitted ear separately, while S/N ratios were obtained in bilateral conditions. A significant correlation between the improvement of the S/N ratio and the insertion gain slope was not clearly evident (Pearson correlation −0.180; p = 0.05). We only found a significant correlation between the improvement of the S/N ratio and the slope of the audiogram (Pearson correlation −0.278; p < 0.001). A scatter plot is given in Fig. 6. We therefore concluded that patients with high frequency hearing losses (steeply sloping audiograms) tended to benefit most from high-frequency amplification in general, regardless of type of fitting procedure investigated in this study.

Discussion

Comparative evaluations of hearing aid fitting procedures according to a double blind randomized clinical trial are scarce in the literature.
We see that the aim of better speech scores has not been realized to a larger extent in the population fitted according to the comparative procedure than in those fitted according to the NAL-RP formula. Although the starting-points and goals were formulated differently, the outcome appeared to be very similar. It might be surprising that although the comparative procedure aimed to optimize speech intelligibility, the NAL-RP even in a few cases resulted in even slightly better speech scores.

One can argue that the outcome of NAL-RP fittings in a normal clinical setting will be somewhat different from ours, because we may have implemented the way of fitting according to this formula too strictly. Calculated target gains were controlled by measurements of the real-ear insertion gains. Coupler responses that we have measured in order to facilitate computerized selection of hearing aids were obtained from straightforward analogue electrical circuits with linear amplification. Compression was used as little as possible in order to avoid any unpredictable effect of nonlinear amplification. In normal clinical practice some leeway will be given to complaints of patients about too much high-frequency amplification. This may have influenced the outcome of the comparative fitting procedure more than the NAL-RP group. However, we found no clear differences in real ear measurements and must conclude that this aspect would be insignificant.

Regarding the recent and current developments in hearing aid technology, it will be clear that the formula that we have used will not be applicable for the current range of commercially available hearing aids, provided with digital and/or programmable nonlinear circuitry. These hearing aids involve many more features than were considered in this study.

We have deliberately chosen not to implement hearing aids with digital circuits and/or WDRC compression algorithms for a couple of reasons valid at the time of the study:

- The extra value of digital hearing aids had not been proven or was not evident and probably not present in the digital hearing aids at the time.
- Various ways of signal processing (like WDRC) were used in digital hearing aids, which would influence the acoustic response of the aid in a complex and inscrutable way, making a comparison between hearing aids far more complicated if not impossible. Moreover, we had no insight into the detailed working of digital hearing aids and their fitting software.
- Generic fitting procedures for digital/WDRC hearing aids were and are in a state of development.
- The answer to the research question would most probably not be different when hearing aids provided with more sophisticated circuitry were used as the effects would affect both fitting procedures to a similar extent.

We included only 254 patients in the study while the power calculation resulted in a required number of 480 patients. There were some reasons for this. We particularly had an insufficient number of patients in the lower two strata. One reason had to do with the presence of significant co-morbidity in the subgroup of patients with a speech intelligibility between 50 and 75%, which resulted in a higher exclusion rate than anticipated. Another reason was the growing request for digital and/or programmable hearing aids from the potential participants. Extending the inclusion period was not considered a practical option as we expected the second reason to become more relevant and outspoken, particularly in the group of poorly performing users. We therefore had to accept smaller numbers of users participating in the study especially in the lower strata. One has to realize that, where differences between the groups were studied, the significance could not always be proven due to relatively small numbers.

Improvement of speech intelligibility in quiet did not show significant differences between the two fitting procedures. This is in accordance with the results of van Buuren et al. [15] who demonstrated that even the intelligibility of speech in noise in mild to moderate hearing losses appears not to be very critical for the hearing aid gain provided over a wide range of spectra. It was remarkable that although the evaluation of speech intelligibility served as a major criterion during hearing aid selection and fine-tuning in the comparative fitting procedure, no better results in terms of this outcome-measure were achieved. A reason for this may be the limited number of test items, being 33 consonants and vowels for each list of 11 CVC words. This means that performance differences between hearing aids for individuals can only be significant for differences of more than about 10%. This is relatively large in view of the total possible improvement.

The improvement of the S/N ratio measured with the Dutch sentence test after hearing aid fitting is in accordance with the data from Verschuure and van Benthem [16] and van den Heuvel et al. [17] who found a small positive effect of a hearing aid on intelligibility of speech in noise. In our study, this improvement was only found in the NAL-RP subgroup. We were not able to point to a clear reason for this finding. Analysis of (unilateral) real-ear insertion gains did anyhow not reveal significant differences in insertion gain-slope between hearing aid fittings according to the NAL-RP formula and the comparative procedure. It has often been suggested that amplification with high-frequency emphasis should result in a better speech intelligibility in noise in spite of poorer sound quality, but we cannot support this assumption from our results.

From our analysis of insertion gains in relation with audiogram-slope and improvement of S/N ratios, it
appeared that patients with high frequency hearing losses (steeply sloping audiograms) tended to benefit most from high-frequency amplification with respect to improvement of the S/N ratio. This is in accordance with the finding from Lee et al. [18]. From their analysis on a group of patients with high-frequency hearing loss it appeared that speech in noise tests were the most sensitive indication of improved speech recognition after hearing aid fitting.

Conclusions

Our data were obtained from experienced and first-time hearing-aid users with a variety of predominantly sensorineural hearing losses. Analogue hearing aids with linear amplification were prescribed. The conclusions listed below are therefore to apply to comparable populations and hearing aids:

1. Improvement of speech intelligibility in quiet was comparable after hearing aid fitting according to both the comparative and the prescriptive procedure studied.
2. Hearing aid fitting according to the strictly implemented NAL-RP formula resulted in a small improvement of the speech-to-noise ratio. No improvement was found after fitting according to the comparative procedure. The difference in the extent of improvement between the two fitting procedures was significant.
3. Comparable real-ear insertion responses for the range of hearing losses included in the study were found for both fitting procedures. No significant differences in slope of the insertion response were found between hearing-aid fittings according to the NAL-RP formula and the comparative fitting procedure.
4. Patients with high frequency hearing losses (steeply sloping audiograms) tended to benefit most from high-frequency amplification.

Acknowledgments This study has been conducted with financial support by the Dutch CVZ (College voor Zorgverzekeringen). The authors would like to thank all the hearing impaired patients that have participated in this study.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Gatehouse S (1993) Hearing aid evaluation: limitations of present procedures and future requirements. JSLPA Monogr Suppl 1
2. Byrne D, Cotton S (1988) Evaluation of the National Acoustic Laboratories’ new hearing aid selection procedure. J Speech Hear Res 31:178–186
3. Byrne D, Dillon H (1986) The National Acoustic Laboratories’ (NAL) new procedure for selecting the gain and frequency response of a hearing aid. Ear Hear 7:257–265
4. Byrne D, Parkinson A, Newall PH (1990) Hearing aid gain and frequency response requirements for the severely/profoundly hearing impaired. Ear Hear 11:40–49
5. Byrne D, Tonisson W (1976) Selecting the gain of hearing aids for persons with sensorineural hearing impairments. Scand Audiol 5:51–59
6. Carhart R (1946) Tests for selection of hearing aids. Laryngoscope 56:780–794
7. Bosman A (1989) Speech perception by the hearing impaired. Thesis, Utrecht University, The Netherlands
8. Smoorenburg GF (1985) Het spraakaudiogram in relatie tot de gebruikte method. In: Spraakaudiometrie. Uitgever Nederlandse Vereniging voor Audiologie
9. Lybarger SF (1963) Simplified fitting system for hearing aids. Canonsburg, Pennsylvania: Radioear Corp
10. International Electrotechnical Commission (IEC) (1973) IEC reference coupler for the measurements of hearing aids using earphones coupled to the ear by means of ear inserts. Publication 126, 2nd edn. Geneva, Switzerland
11. Cornelisse LE, Seewald RC, Jamieson DG (1995) The input/output formula: a theoretical approach to the fitting of personal amplification devices. J Acoust Soc Am 97:1854–1864
12. International Electrotechnical Commission (IEC) (1983) Hearing aids. Part 7. Measurement of the performance of hearing aids for quality inspection for delivery purposes. Publication 118-7, 1st edn. Geneva, Switzerland
13. Verschuure J (1994) Selektie van het hoortoestel. In: Kapteyn TS, Clemens A, Glazenbourg BE, Joustra J (eds) Slechthorende en hoorstoel. De Leeuw Press, Rijnsburg
14. Plomp R, Mijnpen AM (1979) Improving the reliability of testing the speech reception threshold for sentences. Audiology 18:43–52
15. van Buuren RA, Festen JM, Plomp R (1995) Evaluation of a wide range of amplitude-frequency responses for the hearing impaired. J Speech Hear Res 38:211–221
16. Verschuure J, van Benthem PPG (1992) Effect of hearing aids on speech perception in noisy situations. Audiology 31:205–221
17. van den Heuvel J, Goverts STh, Kapteyn ThS (1997) Evaluation of fitting rules with a programmable hearing aid. Audiology 36:261–278
18. Lee LW, Humes LF, Wilde G (1993) Evaluating performance with high-frequency emphasis amplification. J Am Acad Audiol 4:91–97