Audiologist-Driven Versus Patient-Driven Fine Tuning of Hearing Instruments

Trends in Amplification
16(1) 49–58
© The Author(s) 2012
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1084713811424884
<http://tia.sagepub.com>

Monique Boymans, PhD¹ and Wouter A. Dreschler, PhD¹

Abstract

Two methods of fine tuning the initial settings of hearing aids were compared: An audiologist-driven approach—using real ear measurements and a patient-driven fine-tuning approach—using feedback from real-life situations. The patient-driven fine tuning was conducted by employing the Amplifit[®] II system using audiovideo clips. The audiologist-driven fine tuning was based on the NAL-NLI prescription rule. Both settings were compared using the same hearing aids in two 6-week trial periods following a randomized blinded cross-over design. After each trial period, the settings were evaluated by insertion-gain measurements. Performance was evaluated by speech tests in quiet, in noise, and in time-reversed speech, presented at 0° and with spatially separated sound sources. Subjective results were evaluated using extensive questionnaires and audiovisual video clips. A total of 73 participants were included. On average, higher gain values were found for the audiologist-driven settings than for the patient-driven settings, especially at 1000 and 2000 Hz. Better objective performance was obtained for the audiologist-driven settings for speech perception in quiet and in time-reversed speech. This was supported by better scores on a number of subjective judgments and in the subjective ratings of video clips. The perception of loud sounds scored higher than when patient-driven, but the overall preference was in favor of the audiologist-driven settings for 67% of the participants.

Keywords

hearing aid fitting, aided speech intelligibility, amplification, listening comfort

Introduction

There is no unequivocal way to fit modern hearing aids. Different approaches have been developed and several are currently applied in clinical practice.

In modern hearing aids, proprietary fitting rules developed by the manufacturer are usually based on theoretical concepts and on population studies. Those proprietary rules are in many cases not well documented in peer-reviewed articles (Russ, 2001; Svard, Spens, Back, Ahlner, & Barrenasm, 2005). The “first-fit” of hearing aids according to proprietary fitting rules is generally based on audiometric data (usually the pure-tone audiogram only) and on the most important acoustic environments for the individual hearing aid user.

Killion and Gudmundsen (2005) did an evidenced-based review on the correlation between prefitting measures and hearing aid satisfaction based on self-report measures. In none of the reviewed studies significant correlations were found between traditional unaided prefitting speech measures and hearing aid satisfaction, except in a study by Walden and Walden (2004) for speech perception in noise. Wesselkamp, Margolf-Hackl, and Kiessling (2001) and Pastoors, Gebhart, and Kiessling (2001) investigated the effect of a fitting strategy based on loudness scaling. They found no evidence that loudness scaling leads to better fitting results. Cox, Alexander, and Gray (1999, 2007) found that personality is associated

with self-reported outcome data. However, separate personality data at the prefitting stage proved to be minimally useful in prediction of long-term fitting outcomes.

After the “first-fit,” the hearing aid user can start a trial period of some weeks in his or her own acoustical environment. After this trial, the hearing aid user can recall his or her positive and negative experiences, and the audiologist or hearing aid dispenser often uses a complaint-driven fine-tuning procedure to reduce the negative experiences encountered in the trial period (Jenstad, Van Tasell, & Ewert, C, 2003; Kuk, 1999; Nelson, 2001). Although the starting point of such a procedure is relatively well defined (the “first-fit” of the manufacturers) the individual fine-tuning process is not. However, given the fact that there is a large interindividual variation in the acoustical situations and in the perception across patients (even for patients with a similar pure-tone audiogram) the fine-tuning process is crucial but very time-consuming. The acoustic environments of the

¹Academic Medical Center Amsterdam, Department of Clinical and Experimental Audiology, Amsterdam, The Netherlands

Corresponding Author:

Monique Boymans, Department of Clinical and Experimental Audiology, Academic Medical Center, Meibergdreef 9, 1105 AZ Amsterdam, Netherlands
Email: M.Boymans@amc.nl

hearing aid users are not defined well enough to be able to fine tune the hearing aids effectively (Kiessling, 2001). Furthermore, a complaint-driven approach needs effective complaints to work well, and some hearing aid users find it very difficult to formulate the differences in sound quality.

Given the complications mentioned above, a realistic approach is to use a patient-driven fine-tuning procedure. For this purpose, well-structured sets of background noises can be applied to improve the “first-fit” parameters interactively during the initial fitting session in the clinic or in the shop of the hearing aid dispenser (Robinson, Russ, & Siu, 2002). Punch, Robb, and Shovels (1994) found in their study that generalization of the results to all everyday listening situations is not possible. However, they suggested that the predictive value of clinical-based preference judgments, when used in hearing aid fitting, might be substantially improved by using stimulus materials simulating the characteristics of real-world environments.

In this study, the Amplifit® II system is applied using audiovideo clips recorded on DVD, representing specific listening tasks in background noises to simulate specific sound environments. For each hearing aid user, a subset of audiovideo clips is selected that matches best with his or her individual listening tasks and listening environments. The hearing aid user is asked to answer closed questions about speech intelligibility, sound dynamics, and sound quality and based on these responses, the hearing aid settings can be changed directly by the hearing aid dispenser. The potential advantage is a better tailoring of the hearing aid settings toward specific listening conditions that are relevant for the individual user. The focus is on subjective aspects of speech intelligibility and listening comfort, but it is not known how effective this method is for finding the optimal balance between the sometimes contrasting demands of listening comfort and speech intelligibility.

The second approach of hearing aid fine tuning is based on real ear measurements with well-defined and widely accepted prescription rules like NAL-NL1 (Byrne, Dillon, Ching, Katsch, & Keidser, 2001) and DSL-i/o (Scollie et al., 2005) that have proven to optimize speech intelligibility for large groups. These fitting rules are generic, meaning that they are generally applicable and not restricted to the use for hearing aids of one specific brand. This approach can be characterized as audiologist-driven fine tuning. To compensate for individual acoustic effects in the ear canal or in the ear mould that might cause the aid frequency gain characteristics to deviate from the required responses, real ear measurements can be conducted. The influence of subjective listening comfort is minimized, and the main focus is on restoration of speech intelligibility (Ching, Dillon, Katsch, & Byrne, 2001; Keidser & Grant, 2001; Mueller, 2005).

It is not clear which method of fine-tuning produces a more positive fitting outcome. This study is focused on two questions with respect to the fine tuning of a first-fit hearing aid setting:

Research Question 1: Are there measureable differences in measured gain levels for patient-driven fine tuning and audiologist-driven fine tuning, and what are the effects of these differences on both objective performance (speech perception) and subjective judgment (quality judgments and questionnaires)?

Research Question 2: Are there measureable differences between subgroups that benefit more from an audiologist-driven fine-tuning approach?

Method

Participants

To be less dependent on personal fitting philosophies and fitting expertizes, this study was conducted as a multicenter study in four audiological centers in different parts of the Netherlands.¹ For this project each audiological center cooperated with two or three selected hearing aid dispensers in its region.

Hearing aid users were invited to participate in this study on a voluntary basis after they had decided to start a trial period with bilateral hearing aids. The inclusion criteria required that their hearing losses be relatively symmetrical (interaural differences for each frequency lower than 30 dB). The participants could be new or experienced users, so new hearing aid users visited the clinic for their first hearing aids; experienced hearing aid users came for replacement of their old hearing aids. They had to speak Dutch, have good vision, and be physically able to complete speech intelligibility tests. They were asked to sign an informed consent before participating in the study. Seventy-three participants were included in this study.

The hearing aid users received a financial compensation for their travel costs and an additional €75 for the three extra visits that were necessary to complete the experimental part of the hearing aid fitting and fine tuning.

In this multicenter study, 73 participants (42 men and 31 women) were included, with an average age of 65 years (range: 43-80). The majority of the hearing losses were sensorineural (92%), 8% were mixed hearing losses (bone conduction > 10 dB [HL] and the air bone gap > 10 dB for at least one ear). The average hearing loss (500, 1000, 2000, 4000 Hz) was 43.5 dB ($\pm$ 10.5 dB) for the right ears and 44.2 dB ($\pm$ 10.5 dB) for the left ears (for details, see Table 1). Fifty participants represented a first fitting (68%) and half of the participants had an open fitting (n = 36). In total, 47 different hearing aid models were prescribed, representing seven different brands.

Experimental Procedures

The focus of this study was on fine tuning, and therefore the first-fit (including the selection of the brand, style, and model of the hearing aids as well as the type of ear mould)

Table 1. The Average Audiogram and Standard Deviations of Right and Left Ears

Frequency (Hz)	Right ear	Left ear
250	32 dB (± 16)	30 dB (± 15)
500	32 dB (± 14)	32 dB (± 13)
1000	37 dB (± 15)	36 dB (± 15)
2000	46 dB (± 15)	48 dB (± 15)
4000	60 dB (± 15)	61 dB (± 16)

was implemented by the hearing aid dispenser, without restrictions. The dispenser made ear impressions and produced appropriate ear mould(s) (or housings for “in the ear” hearing aids), if necessary.

The Amplifit® II system was used for *patient-driven fine tuning* (PD approach) of the hearing aids (with the “first fit” of the manufacturer as starting point). This system inventoried the subject’s difficulties in auditory communication using an individually tailored selection of video clips, which were chosen to simulate specific real-world conditions relevant for the subject’s auditory communication situation.

First the subject selected videos that characterized his or her individual communication situations as closely as possible with respect to six aspects of hearing: speech in quiet, speech in noise, speech in reverberation, the perception of soft sounds, the perception of loud sounds, and the quality of sounds. The individually selected subset of videos aims to represent situations the participants encounter in everyday listening as good as possible but in a well-standardized way. For example, this could be a subset of videos like a discussion in a quiet room or a discussion in an office with more background noises or just sounds of a bird whistling. The videos of these individually tailored selections were reproduced in freefield by a surround-sound system with five external boxes and calibrated sound levels to reproduce real-life presentation levels (each system was calibrated with a calibration signal). After each video clip with a length of about 20 to 30 s, the participants were given questions about speech intelligibility in quiet, in noise, and in reverberation, loudness of low-intensity and high-intensity sounds, and sound quality with a closed set of answers.² Each answer was correlated with a 5-point score (5 being the best rating). This procedure resulted in an Amplifit profile, containing average scores for each of the six above-mentioned parameters.

Before the (new) hearing aids were fitted, an initial Amplifit profile was measured. The initial test (prefitting) was without hearing aids for the new hearing aid users and with their old hearing aids for the experienced hearing aid users. After the fitting of the hearing aids, the same video clips were judged in the aided condition with the new hearing aids fitted according to the first-fit settings, usually as suggested by the proprietary fitting software. The hearing aid dispenser used the well-structured responses of the Amplifit® II system as a tool to assess the quality of the hearing aid

settings. If the result was not yet satisfying, settings were optimized according to the fine-tuning suggestions available in most hearing aid fitting software programs and/or based on the experience of the hearing aid dispenser. After one or more fine-tuning sessions, the same video clips were judged again by the hearing aid user, until he or she was content with the fine-tuned hearing aid settings for the selected target conditions.

The initial Amplifit profile and the final profile with the new hearing aids (for new and experienced hearing aid users) fitted by the hearing aid dispenser was conducted again at the audiological centers. Insertion-gain measurements provided objective data to objectify the effect of the ear moulds of the new hearing aids. The settings of the hearing aids obtained by patient-driven fine tuning were documented.

The audiologist-driven fine tuning (AD approach) at the audiological centers was based on the NAL-NL1 target (Byrne et al., 2001), carried out under control of real ear measurements derived from broadband noises presented at 65 and 80 dB SPL. After switching off advanced functionalities such as noise reduction and directionality, the frequency gain curves were changed by fine tuning manually to match, as closely as possible, the curves prescribed by the NAL-NL1 prescription rule. Where applicable, advanced functionality was restored with the aim of keeping the PD and AD settings of the hearing aid similar except for the frequency gain characteristics at different levels. The same applied to the use of the comfort programs. The settings of the hearing aids obtained by the audiologist-driven fine tuning were also documented.

After the fitting and fine-tuning procedures, two consecutive field trials were started. Both field trials consisted of 6 weeks each. The order of field trials (PD followed by AD or AD followed by PD) was determined by a cross-over design. The participants were not informed about the order of the trials (single-blinded design).

Evaluation Method

Real ear measurements. To characterize the gain differences between the patient-driven fine tuning and the audiologist-driven fine tuning real ear insertion gain (reig) curves were measured with continuous broadband speech noises presented at input levels of 65 dB and 80 dB (SPL) with Madsen Aurical real ear equipment. Before the measurements, the directivity and noise reduction options were switched off temporarily in the quiet setting. For open fits, the substitution method was used to avoid interference of the test signal and the feedback-reduction signal in the reference microphone (Dyrlund, Ostergaard, Haastrup, & Lantz, 2005). The measurements were conducted before and after each field trial. In this way, it was possible to verify that the particular setting was preserved in the hearing aid during the field trial.

SRT test. To evaluate the effects of the hearing aid settings on speech intelligibility, Speech Reception Threshold (SRT)

tests were conducted in quiet, in continuous speech noise, and in time-reversed speech noise using male and female voices (Festen & Plomp, 1990; Plomp & Mimpen, 1978). Time-reversed speech was used to simulate as much as possible the daily-life situation without being disruptive due to informational masking. Both the speech and the interfering signals came from the front, from the same loudspeaker and the interfering signals were presented at a fixed level of 65 dB(A). The participant was located in the direct field at least 75 cm from the boxes, in a sound isolated room. The speech material for the SRT test was taken from VU98 sentences³ (Versfeld, Daalder, Festen, & Houtgas, 2000), recorded from a male and a female talker. In quiet, only the female voice was used. So in total, 5 tests were conducted with stimuli in front of the participant.

To simulate relevant daily-life situations, spatially separated sound sources were also used. The SRT test was administered with a spatial separation between the target speech and the noise from two loudspeakers positioned at -45° and $+45^\circ$. The “noise” used was time-reversed speech of the opposite gender to facilitate the distinction between both speakers.

With speech from the right-hand side, the interfering signal came from the left-hand side and vice versa. The order was randomized, and the participant did not know beforehand from which side the speech would be presented. The interfering signal was presented at a fixed level of 65 dB(A), whereas the level of the target speech was varied adaptively. The tests were conducted after each trial period with the hearing aid settings corresponding to the preceding trial period. If hearing aids had more than one program with manual switching, the tests were performed in the “quiet” program for the quiet conditions and in the “noise” program for the conditions with interfering noise.

Video clips. After each field trial, the previous individually composed set of video clips was used again but this time only for evaluating purposes and not for fine tuning. Again the participants had to rate subjectively speech intelligibility in quiet, in noise, and in reverberation and they were asked to judge soft and loud sounds, as well as the quality of sounds of specific video clips. If hearing aids had more than one program with manual switching, the video clips were performed in the “quiet” program for quiet situations and in the “noise” program for noisy situations.

Questionnaires. To retrieve information about the subjective experiences with both hearing aid settings during the 6-week trial periods, the hearing aid users were asked to complete the Speech, Spatial, and Qualities of Hearing Scale (SSQ; Gatehouse & Noble, 2004) during each field trial. This questionnaire has been designed to provide an inventory of the remaining problems experienced in a wide variety of realistic listening situation in everyday life. A visual analogue scale was used with a range from 0 to 10, with only labels at 0 (for example “not at all”) and at 10 (for example “perfectly”). The outcome values were related to speech intelligibility, spatial hearing, and quality of sounds.

A direct comparison between both hearing aid settings was performed by means of a shortened version of the AVETA (Amsterdamse Vragenlijst voor Eén en Tweezijdige Aanpassingen—A validated Dutch Questionnaire for unilateral and bilateral hearing aid fittings) questionnaire (Boymans, 2003), a questionnaire that was partly based on existing questionnaires such as the Amsterdam Inventory of Auditory Disability and Handicap (AIADH; Kramer, Kapteyn, Festen, & Tobi, 1995) and the Abbreviated Profile of Hearing Aid Benefit (APHAB; Cox & Alexander, 1995). Based on 18 questions, outcome values were obtained for detection, discrimination, speech in quiet, speech in noise, localization, and aversion to loud sounds. Each category was represented by three questions. In a few cases some answers were missing. If the majority of answers (2 or 3 out of each set of 3) were available, the average scores for each participant and each category were calculated. If only 1 out of 3 answers was available, the end result was regarded as missing.

All scales ranged from 1 to 4. The AIADH and APHAB questions were posed for the conditions prefitting, with the hearing aid according to the first hearing aid trial and with the hearing aid according to the second trial. The randomization scheme determined whether these answers were related to the PD or AD settings.

At the end of both field trials, the hearing aid users were asked about their overall preference for one of the hearing aid settings, just by asking, “Which setting do you prefer, the setting of the first trial period or the setting of the second trial period.” This preferred setting was then programmed into their hearing aids for continuing use.

Data Analyses

The results were analyzed in two sections. The first section focuses on the differences between AD and PD settings (Research Question 1) and describes the differences for the objective outcome measures (gain and speech perception), the subjective outcome measures (video clips and SSQ questionnaire) and the outcome measures of the direct comparison (AVETA questionnaire). Paired *t* tests were used to determine the significance of the differences between the results with the AD and PD settings.

The second section focuses on the identification of subject groups that benefit more from one or the other fine-tuning approach (Research Question 2). For this purpose, we used analyses of variance (repeated measures). The results show the effects of the main parameters (as within subject factors) for the different subgroups. The subgroups were distinguished based on differences in age, hearing aid experience (as between subject factors), type of ear mould (open or closed), and gain differences.

In the analysis of the subjective results, missing values were encountered. The results presented are averages of all responses available. Paired *t* tests included only pairs without missing values. The level of significance used in the analyses was defined as $p < .05$.

Figure 1. The average target curve according the NAL-NL1 at 65 dB and the average REIG values for the patient-driven and audiologist-driven settings averaged across all subjects and across left and right ears.

Results

Real Ear Measurements

Real ear measurements were conducted for the patient-driven and audiologist-driven settings with a broadband speech-shaped noise at 65 dB SPL. The average NAL-NL1 target curves for left and right ears ($n = 146$) and the insertion gain values of the PD and AD settings are represented in Figure 1.

On average, the curve of the AD settings was higher than the curve of the PD settings. The average gain difference (500, 1000, 2000, 4000 Hz) was 3.2 dB ($p < .001$). The gain differences were slightly less for lower frequencies (a significant difference of 2.6 dB at 500 and 1000 Hz; $p < .001$) than for higher frequencies (a significant difference of 3.7 dB at 2000 and 4000 Hz; $p < .001$).

The curve of the AD settings was 2 dB lower than the target curves for 500, 1000, and 2000 Hz. At 250 Hz the difference was 5 dB, and at 3,000 and 4000 Hz the differences were 4 and 9 dB, respectively. This was due to the fact that NAL-NL1 could not always be reached without feedback problems, whereas the ear mould could not be changed to keep it constant for both fine-tuning approaches.

The average NAL-NL1 target curves for left and right ears with a broadband speech-shaped noise at 80 dB SPL and the average insertion gain values of the patient-driven and the audiologist-driven settings are represented in Figure 2.

The differences between the target curve and the AD curve were relatively small. On average, the curve of the AD settings was again higher than the curve of the PD settings. However, the distances between the curves were smaller than for the 65-dB input level. The average gain difference (500, 1000, 2000, 4000 Hz) was 1.4 dB ($p < .001$). Again, the gain differences were slightly less for lower frequencies (a significant difference of 1.1 dB at 500 and 1000 Hz; $p < .001$) than for higher frequencies (a significant difference of 1.6 dB at 2000 and 4000 Hz; $p < .001$). The differences between the target curve and the average results of the PD settings were 4.5 at 250 Hz and 3 and 9 dB at 3000 and 4000 Hz, respectively.

Figure 2. The average target curve according the NAL-NL1 at 80 dB and the average REIG values for the patient-driven and audiologist-driven settings averaged across all subjects and across left and right ears.

Table 2. The Results of the SRT Test in Quiet, in Different Noises, and With Spatially Separated Sound Sources Averaged for -45° and 45° , for PD and AD

	Patient-driven settings	Audiologist-driven settings	p
Speech and noise at 0°			
SRT in quiet	43.5 dB (± 5.7)	40.7 dB (± 5.6)	$p < .001$
SRT in continuous noise	0.8 dB (± 2.5)	0.9 dB (± 2.4)	ns
SRT in the reversed speech	-1.1 dB (± 3.3)	-1.7 dB (± 3.7)	$p < .05$
Speech and noise spatially separated			
average -45° and 45°	-7.2 dB (± 3.8)	-8.2 dB (± 4.4)	$p < .001$

Note: The significance of the differences is indicated in the last column.

Objective Performance for Speech Perception

Table 2 presents the results of the SRT test in quiet, in continuous noise, and in time-reversed speech from the front (0° azimuth), for the PD and AD settings, respectively. The degree of significance is shown in the last column.

The first row represents the average presentation level in quiet in which 50% of the sentences were repeated correctly. The second and third row represent the signal to noise ratios (SNRs) averaged for the target sentences spoken with a male and a female voice (lower SNRs indicate a more favorable result).

In quiet, the average thresholds for the AD settings were significantly better than for the PD settings ($p < .001$). The difference in continuous noise was not significant but in time-reversed speech again a significantly better SNR was found for the AD settings ($p < .05$).

The last row in Table 2 shows the average results for the SRT test with spatially separated sound sources consisting of a target speech signal masked by an interfering (time-reversed) speech signal of the other gender at the other side.

Table 3. The Benefit From Temporal Gaps (Difference Continuous Noise Vs. Time Reversed Speech), From Spatial Separation (Difference Between 0° Vs. -45° and 45°), and the Benefit From the Gender of the Speaker

	Patient-driven settings	Audiologist-driven settings	<i>p</i>
Benefit from temporal gaps	1.9 dB (±2.2)	2.6 dB (±2.3)	<i>p</i> < .01
Benefit from spatial separation	6.1 dB (±2.2)	6.5 dB (±2.7)	<i>ns</i>
Benefit from female speaker	2.2 dB (±2.0)	2.4 dB (±2.1)	<i>ns</i>

Note: The significance of the differences between PD and AD is indicated in the last column.

Figure 3. The average subjective ratings for the PD and AD settings based on judgments of the video clips for the situations: speech in quiet, in background noise, in reverberation, soft and loud sounds, and the quality of sounds. The significance of the differences between the PD and AD settings is indicated by asterisks: **p* < .05.

The SNRs with the AD settings were significantly better than with the PD settings (*p* < .001).

The benefit from temporal gaps (difference between the results in continuous noise and a fluctuating background as used in time-reversed speech), from spatial separation (difference between the results at 0° vs. -45° and 45°), and the benefit from the female speaker (difference between the female vs. male speaker in continuous noise and in time-reversed speech at 0°) are represented in Table 3.

There was significantly more benefit of the temporal gaps with the AD settings than with the PD settings.

Video clips. The average subjective judgments of the video clips are shown in Figure 3. Higher bars represent better results. For speech in quiet, speech in background noise, and for the quality of sounds, the subjective scores with the AD settings were significantly higher than with the PD settings (*p* < .05).

Table 4. Average Scores of the SSQ Questionnaire for PD and AD, the Overall Scores for the Subscales Speech, Space, and Quality, and the Scores of Single Questions With a Statistically Significant Difference Between PD and AD

	Patient-driven settings	Audiologist-driven settings	<i>p</i>
Overall score			
Total speech	5.4 dB (±1.6)	5.7 dB (±1.4)	<i>ns</i>
Total space	6.0 dB (±1.6)	6.2 dB (±1.5)	<i>ns</i>
Total quality	6.5 dB (±1.4)	6.5 dB (±1.3)	<i>ns</i>
Specific questions			
Talking with 1 person in quiet	7.1 dB (±1.7)	7.6 dB (±1.5)	<i>p</i> < .01
Conversation in echoic environment	5.3 dB (±2.1)	5.9 dB (±2.1)	<i>p</i> < .05
Naturalness own voice	6.6 dB (±2.4)	6.0 dB (±2.3)	<i>p</i> < .01
Understand car passenger	5.6 dB (±2.2)	6.0 dB (±2.1)	<i>p</i> < .05

Note: The significance of the differences between the PD and AD is indicated in the last column. The significance of the effects disappeared after Bonferroni correction.

However, the PD settings yielded better scores for subjective judgments of loud sounds (*p* < .05).

Questionnaires

SSQ. The averaged scores on the overall subscales speech, space, and quality, of the SSQ for the PD and AD settings are shown in the first four rows of Table 4. The average scores did not differ significantly for PD and AD settings. In the individual questions, some significant differences were found (shown in the last four rows).

For the speech-related question “Talking with one person in a quiet room” and “Having a conversation in an echoic environment” and for the quality-related question “When you are a passenger, can you easily hear what the driver is saying?” the AD settings scored significantly higher than the PD settings (*p* < .01, *p* < .05, and *p* < .05 respectively). For the quality-related question “Does your own voice sound natural to you?” the PD settings scored significantly higher than the AD settings (*p* < .05).

However, after a Bonferroni correction, the differences based on specific questions all lost their significance.

AVETA. The subjective scores on the six subscales used in the AVETA are shown in Figure 4. The bars represent the average results (with standard deviations) for participants providing valid responses for the conditions prefitting and postfitting with PD and AD settings, respectively. For all scales, higher bars indicate a more positive result (note that the scale of aversiveness was inverted to a scale of “comfort for loud sounds”).

In the total group, the scores postfitting were significantly higher than prefitting (without hearing aids or with the old hearing aids) for detection of sounds (*p* < .001), discrimination

Figure 4. The average results of the subjective judgments, before fitting and with the new hearing aids in PD and AD settings, for all 6 categories of the AVETA questionnaire. The significance of the differences between the PD and AD settings is indicated by asterisks: * $p < .05$.

of sounds ($p < .01$), speech in quiet, and speech in noise ($p < .001$). However, for the comfort of loud sounds, the conditions with hearing aids were significantly poorer than the prefitting situation ($p < .001$). No significant differences were found for localization of sounds.

For this study, the comparisons between the PD and AD settings are especially relevant. As indicated with asterisks in Figure 4, the results for the AD settings were significantly better scores for the categories detection ($p < .05$), speech in noise ($p < .05$), and localization ($p < .05$).

Overall preferences. After completing both trial periods, 67% of the participants preferred the audiologist-driven setting, 15% did not have a preference, and 18% preferred the patient-driven setting.

Analysis of the for Subgroups

The effect of the main parameters described above were analyzed for different subgroups based on differences in age, open or closed fits, and hearing aid experience.

Subgroups according to age. Three age groups were distinguished: Younger than 60 years old ($n = 15$), between 60 and 70 years ($n = 35$), and 70 years or older ($n = 23$). No significant group interactions were found for the outcome parameters and age.

Subgroups according to previous experience. New hearing aid users ($n = 50$) visited the clinic for their first hearing aids; experienced hearing aid users ($n = 23$) came for replacement of their old hearing aids. There was a significant interaction between the effects of fine-tuning method and hearing aid experience. The effect that more gain was found for the audiologist-driven settings was larger for the subgroup without previous hearing aid experience than for the experienced hearing aid group ($p < .05$). As expected the difference in SRT results between the AD settings and PD settings for the

subgroup without hearing aid experience was significantly larger than for the experienced subgroup ($p < .05$). No significant group interactions were found for the other outcome parameters.

Subgroups according to type of hearing aid. The numbers of open and closed fittings were about equal ($n = 36$ and 37 , respectively). For the closed fitted group half of the participants were new users ($n = 19$) and half of the participants were experienced users ($n = 18$). However, in the open fitted group, most participants were new users ($n = 31$) and only 5 participants were experienced users. For the real ear measurements, there were no significant interactions between age or the type of ear mould (open or closed). For the SRT-test, a significant group interaction was found between the open and closed fittings ($p < .05$). The effect that audiologist-driven settings yielded better speech intelligibility in background noise than patient-driven settings was larger for the group with open fits than for the group with closed fits. No significant group interactions were found for the subjective ratings of the video clips, the SSQ questionnaire, and the AVETA questionnaire.

Subgroups according to gain differences. The results presented above indicate that gain is an important factor for the differences between the PD setting and the AD setting. To investigate the effects of gain, the results of subgroups with larger gain differences ($AD - PD > 3$ dB) were compared with the results of subgroups with minor differences in gain ($AD - PD \leq 3$ dB) for three frequency areas (broadband: 500, 1000, 2000, 4000 Hz; low frequencies: 500, 1000 Hz; high frequencies: 2000, 4000 Hz).

Only one significant difference was found for the SRT test. The subgroup $AD - PD > 3$ dB for broadband frequencies showed significantly more benefit ($p < .05$) from the AD settings for speech in quiet (3.5 dB; $n = 36$) than the group $AD - PD \leq 3$ dB (1.9 dB; $n = 37$; see Table 5).

In the AVETA, only one significant group effect was found for subgroup $AD - PD > 3$ dB for broadband frequencies: the comfort of loud sounds was better for the PD settings.

For the Amplifit videos and the SSQ questionnaires, no significant differences were found for the different subgroups.

Discussion

This study investigated the effects of different fine-tuning approaches on a variety of outcome data, both objective and subjective. Clear differences were found between a patient-driven and an audiologist-driven fine-tuning approach when the shaping of frequency gain was modified, whereas other features such as noise reduction and directionality were identical.

This is in contrast with a study of Hornsby and Mueller (2008), who found no difference between the setting according to NAL-NL1, and the setting after adjusting the hearing aid gain to optimize listening comfort and speech clarity by the participants themselves. The contrast between that study

Table 5. Difference Values for SRT Outcome Measures for Subgroups Composed on the Basis of a 3-dB Difference in Real Ear Measurements

Benefit AD-PD SRT-test	.5, 1, 2, 4 kHz		.5, 1 kHz		2, 4 kHz	
	AD-PD ≤ 3 dB	AD-PD > 3 dB	AD-PD ≤ 3 dB	AD-PD > 3 dB	AD-PD ≤ 3 dB	AD-PD > 3 dB
Number	37	36	40	33	33	40
Quiet	1.88	3.52	2.22	3.25	1.94	3.31
Continuous noise	-0.04	-0.30	-0.08	-0.26	0.29	-0.54
Time reversed speech	0.54	1.21	0.59	1.22	0.86	0.89

Note: In subgroup "AD-PD ≤ 3 dB" gain differences were smaller than 3 dB; in the other subgroup the gain for AD is more than 3 dB higher than for PD. These differences were analyzed for broadband frequencies (columns 2 and 3), for low frequencies (columns 4 and 5), and for high frequencies (columns 6 and 7). $p < 0.05$.

and ours may be related to differences in the starting point before the fine-tuning process (Dreschler, Keidser, Convery, & Dillon, 2008; Hornsby & Mueller, 2008; Mueller, Hornsby, & Weber, 2008). In our study, the starting point of the interactive fitting was the first-fit rule of the different hearing aids.

Souza, Yueh, Sarubbi, and Loois, (2000) found no relationship between audibility and patient ratings of communication. However, in our study, better speech intelligibility results in quiet were found for the audiologist-driven setting. This is likely due to more overall gain relative to the PD setting.

The patient-driven fine tuning can be considered as a more comfort-driven approach, whereas the audiologist-driven fine tuning can be considered as a more intelligibility driven approach. The initial hypothesis was that the patient-driven setting should score higher on comfort or quality aspects than the audiologist-driven settings based on the NAL-NL1 rule. This effect was only demonstrated for the comfort of loud sounds with video clips and for one question of the SSQ. During video clips, the comfort of loud sounds was higher for the patient-driven settings than for the audiologist-driven settings. The PD settings yielded better scores for subjective judgments of loud sounds. This suggests that subjective judgments of loudness were important for PD settings. However, this effect was not statistically significant in the direct comparison questionnaire. In the SSQ questionnaire, the question about the naturalness of their own voice scored higher with the patient-driven settings than with the audiologist-driven settings, but this significance disappeared after Bonferroni correction. So there is only weak evidence that the hypothesis on listening comfort is correct.

In contrast with the expectations, the quality of sounds during the video clips scored higher for the audiologist-driven settings than for the patient-driven settings, and again this was not confirmed on a statistically significant level by the subscale for sound quality in the SSQ questionnaire.

The second hypothesis was that the audiologist-driven fine tuning should result in better speech intelligibility scores because NAL-NL1 is designed for optimizing speech intelligibility. The data confirmed that speech intelligibility scores were significantly better for the audiologist-driven settings

than for the patient-driven settings, both for the SRT tests and for the AVETA questionnaire. However, for the subscale Speech in the SSQ questionnaire, no significant improvement was found for the audiologist-driven approach. The responses to the SSQ questionnaire produced only little differences between the AD and PD settings. One possible explanation for this is that many hearing impaired participants experienced difficulties in completing this questionnaire because not all situations were replicated during both trial periods. As a consequence, participants may have been uncertain as to how to respond to some questions.

The video clips and the insertion gain measurements were both used as tools for fine tuning and as tools for evaluation. Although this approach may have some methodological drawbacks due to lack of independence of fitting and evaluation, the advantage is that both fine-tuning approaches can be evaluated according to the same methods that have been used for the fine tuning. The method used to optimize the intelligibility-driven approach is also used to evaluate the comfort-driven approach and vice versa.

For the videos, we do not expect learning effects, because the hearing aid users were already familiar with the video clips. (The clips had been viewed two or more times with the hearing aid dispenser and two times at the audiological center.) If some learning effects were still present, these would have been compensated for by the randomization of the order of the hearing aid settings.

The larger difference in gain between both settings (more gain for the audiologist-driven fine tuning) for the new users versus the experienced hearing aid users may be explained by the fact that experienced hearing aid users were more used to the gain and the sound of the hearing aid than the new users. From clinical experience, it is well known that new users tend to be reluctant to accept higher gain values, partly because they are used to soft sounds and soft speech and partly because they do not know what to expect. In literature, there is some evidence that new hearing aid users prefer less gain than experienced users (Marriage, Moore, & Alcántara, 2004) or show more gain adaptation (Keidser, O'Brien, Carter, McLelland, & Yeend, 2008). Cox and Alexander (1992) found that experienced users obtained more benefit

than new users. However, they found similar time-related changes in benefit during the first 10 weeks of hearing aid use.

In the patient-driven approach of our study, the participants had the opportunity to keep the gain values low if that was their preference. The results of the SRT tests could be explained by the difference in the real ear measurements. There was more gain for the AD than for the PD settings especially at 1000 and 2000 Hz. So in a quiet environment, the input level could be set lower for the AD settings than the PD settings to achieve 50% correct speech recognition. In continuous noise, there was no difference between the PD and AD settings because the level of speech and noise were both raised. For the situation in time-reversed speech, the levels of both speech and noise were raised; however, a higher input signal (in this case for the AD settings) produced better speech intelligibility in the temporal gaps.

In this study, comfort of loud sounds measured with video clips was rated higher for the PD settings than for the AD settings. Despite the fact that the AD settings produced better speech intelligibility in quiet and in noise and also scored better quality, special attention should be given to the comfort of loud sounds.

The results from this study suggest that a patient-driven approach may lead to suboptimal speech perception. Therefore, an objective verification by real ear measurements is important. However, the general acceptance and comfort of hearing may profit from an interactive procedure using realistic daily-life situations as simulated in this study by video clips. A way to handle these possibly conflicting requirements is to start with a first fit based on insertion gain measurements and a generic fitting rule and then continue with a fine-tuning process like Amplifit® II system. Although this study did not provide hard data, it may be assumed that the patient-driven approach with videos will be more cost-effective than a patient-driven approach with trial and error during successive trial periods. The presence of complaints during trial periods may increase the number of negative trials because of a disappointing result. Such an audio-visual system could also be used for counseling. Saunders, Lewis, and Forsline (2009) found small but significant effects on expectations due to counseling (with or without such a system).

Conclusions

The audiologist-driven approach yielded on average a higher gain than the patient-driven approach, especially at 1000 and 2000 Hz. The objective performance for speech perception in quiet and in time-reversed speech was better with the audiologist-driven settings than with the patient-driven settings. The subjective judgments about speech perception in quiet, in noise and the quality of sounds, with video clips, favored the audiologist-driven settings. The subjective judgments of the comfort of loud sounds favored the patient-driven settings. The subjective judgments about detection, speech in noise, and localization with the AVETA question-

naire were more positive for the audiologist-driven settings. There is a trend that the new users (having more open fits) report greater benefit from the audiologist-driven approach than from the patient-driven approach. After both trial periods, most hearing-impaired participants preferred the audiologist-driven settings (67%) over the patient-driven hearing aid settings (18%), 15% did not have a preference.

Acknowledgments

This study was conducted by four Audiological Centers, co-operating in the Foundation PACT (Platform for Audiological Clinical Testing), funded by Amplifon (Italy) and Beter Horen (the Netherlands). The authors would like to thank the colleagues from the Audiological Centers in Eindhoven, Hoensbroek, and Tilburg and the hearing aid dispensers of Beter Horen in the area of Amsterdam, Eindhoven, Hoensbroek, and Tilburg. They also would like to thank Johannes Borgstein for correcting the English text. In addition, the authors are grateful to László Korösy (AMC) for his indispensable technical support in this project.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: from Amplifon.

Notes

1. The participating centers were all members of PACT (Platform of Audiological Clinical Testing), a foundation for multicenter clinical studies.
2. Some examples of the questions are "How much did you understand the speech of person . . ." (questions related to speech perception); "How did you perceive/hear . . ." (questions related to the loudness of soft and loud sounds and the quality of sounds).
3. For example, "The children were home alone" (in Dutch).

References

- Boymans, M. (2003). *Intelligent processing to optimize the benefits of hearing aids* (Doctoral dissertation). University of Amsterdam, Amsterdam, Netherlands.
- Byrne, D., Dillon, H., Ching, T. Y., Katsch, R., & Keidser, G. (2001). NAL-NL1 procedure for fitting nonlinear hearing aids: Characteristics and comparisons with other procedures. *Journal of the American Academy of Audiology*, 12(1), 37-51.
- Ching, T. Y., Dillon, H., Katsch, R., & Byrne, D. (2001). Maximizing effective audibility in hearing aid fitting. *Ear and Hearing*, 22, 212-224.
- Cox, R. M., & Alexander, G. C. (1992). Maturation of hearing aid benefit: Objective and subjective measurements. *Ear and Hearing*, 13, 131-141.
- Cox, R. M., & Alexander, G. C. (1995). The abbreviated profile of hearing aid benefit. *Ear and Hearing*, 16, 176-183.

- Cox, R. M., Alexander, G. C., & Gray G. (1999). Personality and the subjective assessment of hearing aids. *Journal of the American Academy of Audiology*, 10(1), 1-13.
- Cox, R. M., Alexander, G. C., & Gray G. A. (2007). Personality, hearing problems, and amplification characteristics: Contributions to self-report hearing aid outcomes. *Ear and Hearing*, 28(2), 141-162.
- Dreschler, W. A., Keidser, G., Convery, E., & Dillon, H. (2008). Client-based adjustments of hearing aid gain: The effect of different control configurations. *Ear and Hearing*, 29, 214-227.
- Dyrlund, O., Ostergaard, S., Hastrup, A., & Lantz, J. (2005). REIG measurements on open fittings using digital feedback suppression. *Proceedings of the 21th Danavox Symposium on Hearing aid fitting, Denmark*, 1, (167-172).
- Festen, J. M., & Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech reception threshold for impaired and normal hearing. *Journal of the Acoustical Society of America*, 88, 1725-1736.
- Gatehouse, S., & Noble, W. (2004). The Speech, Spatial and Qualities of Hearing Scale (SSQ). *International Journal of Audiology*, 43, 85-99.
- Hornsby, B. W., & Mueller, H. G. (2008). User preference and reliability of bilateral hearing aid gain adjustments. *Journal of the American Academy of Audiology*, 19(2), 158-170.
- Jenstad, L. M., Van Tasell, D. J., & Ewert, C. (2003). Hearing aid troubleshooting based on patients' descriptions. *Journal of the American Academy of Audiology*, 14(7), 347-360.
- Keidser, G., & Grant, F. (2001). Comparing loudness normalization (IHAF) with speech intelligibility maximalization (NAL-NL1) when implemented in a two-channel device. *Ear and Hearing*, 22, 501-515.
- Keidser, G., O'Brien, A., Carter, L., McLelland, M., & Yeend, I. (2008). Variation in preferred gain with experience for hearing-aid users. *International Journal of Audiology*, 47, 621-635.
- Kiessling, J. (2001). Hearing aid fitting procedures—State-of-the-art and current issues. *Scandinavian Audiology Supplement*, 52, 7-9.
- Killion, M. C., & Gudmundsen, G. I. (2005). Fitting hearing aids using clinical prefitting speech measures: An evidence-based review. *Journal of the American Academy of Audiology*, 16, 439-447.
- Kramer, S. E., Kapteyn, T. S., Festen, J. M., & Tobi, H. (1995). Factors in subjective hearing disability. *Audiology*, 34, 311-320.
- Kuk, F. K. (1999). How flow charts can help you troubleshoot hearing aid problems. *Hearing Journal*, 52, 46-52.
- Marriage, J., Moore, B. C., & Alcántara, J. I. (2004). Comparison of three procedures for initial fitting of compression hearing aids III. Inexperienced versus experienced users. *International Journal of Audiology*, 43(4), 198-210.
- Mueller, G. (2005). Fitting hearing aids to adults using prescriptive methods: An evidence-based review of effectiveness. *Journal of the American Academy of Audiology*, 16, 448-460.
- Mueller, H. G., Hornsby, B. W., & Weber, J. E. (2008). Using trainable hearing aids to examine real-world preferred gain. *Journal of the American Academy of Audiology*, 19, 758-773.
- Nelson, J. A. (2001). Finetuning multichannel compression hearing instruments. *Hearing Review*, 58, 30-35.
- Pastoors, A. D., Gebhart, T. M., & Kiessling, J. (2001). A fitting strategy for digital hearing aids based on loudness and sound quality. *Scandinavian Audiology*, 52, 60-64.
- Plomp, R., & Mimpfen, A. M. (1978). Improving the reliability of testing the speech reception threshold for sentences. *Audiology*, 18, 43-52.
- Punch, J. L., Robb, R., & Shovels, A. H. (1994). Aided listener preferences in laboratory versus real-world environments. *Ear and Hearing*, 15(1), 50-61.
- Robinson, J., Russ, D., & Siu, B. (2002). Effect of a multimedia fitting tool on final gain settings. *Hearing Review*, 9(3), 42-45.
- Russ, D. (2001). Beltone's adaptive fitting algorithm: Combining loudness normalization and loudness equalization to achieve target gain. *Audiology Online*. Retrieved from www.audiologyonline.com
- Saunders, G. H., Lewis, M. S., & Forsline, A. (2009). Expectations, prefitting counseling, and hearing aid outcome. *Journal of the American Academy of Audiology*, 20(5), 320-334.
- Scollie, S., Seewald, R., Cornelisse, L., Moodie, S., Bagatto, M., Launagaray, D., . . . Pumford, J. (2005). The desired sensation level multistage input/output algorithm. *Trends in Amplification*, 9(4), 159-197.
- Souza, P. E., Yueh, B., Sarubbi, M., & Loovis, C. F. (2000). Fitting hearing aids with the Articulation Index: Impact on hearing aid effectiveness. *Journal of Rehabilitation Research and Development*, 37, 473-481.
- Svard, I., Spens, K. E., Back, L., Ahlner, B. H., & Barrenasm, M. L. (2005). The benefit method: Fitting hearing aids in noise. *Noise & Health*, 7(29), 12-23.
- Versfeld, N. J., Daalder, L., Festen, J. M., & Houtgast, T. (2000). Method for the selection of sentence materials for efficient measurement of the speech reception threshold. *Journal of the Acoustical Society of America*, 107, 1671-1684.
- Walden, T. C., & Walden, B. E. (2004). Predicting success with hearing aids in everyday living. *Journal of the American Academy of Audiology*, 15, 342-352.
- Wesselkamp, M., Margolf-Hackl, S., & Kiessling, J. (2001). Comparison of two digital hearing instrument fitting strategies. *Scandinavian Audiology*, 52, 73-75.