Intelligent processing to optimize the benefits of hearing aids

Boymans, M.

Citation for published version (APA):
CHAPTER 2.

ASSESSMENT OF HEARING AID CANDIDACY
AND HEARING AID BENEFIT
2. Assessment of hearing aid candidacy and hearing aid benefit

In both parts of this thesis we apply audiometric test methods that provide extra information complementary to the pure-tone audiogram in order to facilitate auditory rehabilitation with hearing aids. Therefore, this chapter provides an overview of test methods using speech and non-speech stimuli.

2.1. Psychophysical tests with non-speech stimuli

While the pure-tone audiogram measures the absolute threshold as a function of frequency, other audiometric tests are available that focus on the perception of supra-threshold signals. Part of these tests are relevant for the evaluation of hearing aids and/or for the evaluation of binaural hearing, relevant for the fitting of bilateral hearing aids.

2.1.1. Loudness scaling

Sensory hearing loss affects loudness perception and this can only be measured subjectively. Loudness perception can be measured by means of categorical loudness scaling. It is possible to use different types of noises and different ranges of output levels.

One method is the Würzburger Hörfeld Skalierung (Hellbrück et al., 1985). The scaling of loudness is based on a 50-point scale, ranging from “not heard” to “too loud”. The instruction is to judge loudness at the end of each fragment. Another method has been
proposed by Pascoe (1986). In this method a 10-point scale is used. The results of both methods can be plotted as loudness growth curves; the plot of categorical loudness units (vertical axis) versus presentation level in dB (horizontal axis). The raw data can be fitted by a curve and this curve determines the most comfortable level (“MCL” at 50% of the scale). Also the slope of the loudness growth function can be calculated. MCL is related to the amount of hearing loss and the slope to the amount of recruitment.

Some digital hearing aids provide a form of loudness scaling in their fitting software. Usually, the amount of compression ratio will be adjusted according the results. For normal-hearing people loudness is greater when this is measured binaurally. Loudness summation is rising from 3 dB near the threshold (Dermody, 1975) to about 6-10 dB for higher intensities (Christen, 1980; Haggard, 1982).

2.1.2. Horizontal localization

To assess horizontal localization ability we usually apply a localization experiment with 13 matched loudspeakers, positioned in half a circle in front of the subject (from $-90^\circ$ to $+90^\circ$). The stimuli are usually broadband noise bursts, 300 ms in duration and gated according to a half cosine function to avoid clicks. The hearing-impaired person responses by indicating the number of the box where he/she thought the noise came from.

For the quality of horizontal localization two parameters can be calculated:

- The root mean square value of the stimulus response differences (in degrees). This parameter is used to get information about the absolute values of the faults, weighting large discrepancies between stimulus and response more severe than smaller ones.
The correlation coefficient for the stimulus response patterns. This parameter is used to find out whether response patterns correspond to the spatial ordering of the stimuli, irrespective of the absolute values of the deviations.

Because noise bursts are not realistic, a new localization test has been developed with a random selection of several daily sounds, like dog barking, music etc., presented simultaneously at a fixed intensity level. The different noises were overlapping that way that at every moment at least three noises were present. When the hearing-impaired listener hears a telephone bell, he/she has to indicate from which speaker box the sound came. Now only five boxes are used and the intensity of the telephone bell had a roving level in order to avoid that differences in the output of the loudspeakers would give unwanted cues to the listener about the location of the telephone bell. For this localization test the order of presentations was also randomized, but now resulting in six presentations for each of the five loudspeakers for each measurement (see Chapter 5).

2.1.3. Binaural Masking Level Differences (BMLD)

The auditory system of the human brain can combine signals from the two ears in order to make a better separation between the signals. For a unilaterally presented signal, this results in a better critical signal to noise ratio (S/N), when the noise is presented bilaterally instead of unilaterally. There is also a better critical S/N ratio for a bilateral tone in noise presented bilaterally, when the tone is out of phase instead of in phase.

The amount of noise suppression is called the binaural masking level difference (BMLD), or binaural release from masking, or binaural unmasking or binaural squelch. The BMLD for low frequency sounds is the strongest, about 15 dB. The effect of BMLD for speech is smaller than for low frequency sounds. The BMLD for speech for normal-hearing subjects is 6 - 8 dB (Johansson et al., 2002).
2.1.4. Interaural Time Differences (IATD)

In free field conditions a signal will arrive first at the ear closest to the sound source, and after some time the signal will also arrive at the other ear. The difference between both arrival times is called interaural time difference (IATD). Interaural time differences depend on the direction of the stimuli and the size of the head. There is no interaural time difference when the sound source is located at 0° azimuth, and the IATD is about 0.7 ms for sounds coming from 90° (Kuhn, 1982). Interaural time differences are resulting in interaural phase differences. The chance that the interaural phase difference is zero is higher for high frequencies than for low frequencies. Interaural time differences and the interaural phase differences are used to localize sounds.

IATD can be measured with headphones. The stimulus consists of two noise bursts presented binaural, starting with a short interaural time difference for the first noise burst (Δt), while the interaural time difference is reversed in the second noise burst. For example, the first part of the binaural noise burst is presented first at the right ear and Δt later at the left ear. This causes that one noise burst will be heard at the right side of the head. The leading noise burst dominates according to "the precedence effect" (Gardner, 1968; Moore, 1982; Goverts et al., 2000). The next binaural noise burst will be presented first at the left ear and then at the right ear. Consequently, this noise burst will be heard at the left side of the head. So the two binaural noise bursts give the impression of moving from the right-hand side to the left-hand side. When the interaural time difference is zero, the binaural noise bursts will be heard in the middle of the head. During the test Δt will be varied adaptively in order to find the minimum interaural time difference that causes a moving image in the head. The smaller the value the better the IATD.
2.2. Psychophysical tests with speech stimuli

2.2.1. Intelligibility of single words

Traditionally, speech perception ability is measured with short lists with monosyllabic CVC-words (consonant-vowel-consonant words) (Bosman, 1989). The speech material is presented by headphones at different average speech levels, resulting in the so-called speech audiogram. This test is well standardized, relatively fast and it gives a good impression about the speech intelligibility at different speech levels.

Steeneken et al. (1990), developed a speech test, which is based on existing and fictitious CVC words with a balanced frequency of occurrence for each phoneme, in order to allow an analysis of confusions. This test does not only provide information about the percentage of correctly identified words, but also about the type of confusions between phonemes. These confusions can be related to the acoustical features of the phonemes and allows a qualitative analysis of the intelligibility problems. The CVC-words used in this test are presented in carrier sentences of four words. The subject has to identify always the third word in the sentence. There are only five carrier-sentences, and 51 target CVC-words per list. The carrier sentence is shown on a computer screen and the target CVC-word has to be identified and to be typed into the computer. At the end of the test, a list is shown with the target CVC-words and the answers. This test can also be presented in background noise. The disadvantage of the test is that it is very time-consuming especially when a high number of conditions has to be measured.

The output files can be used to generate confusion matrices and these confusion matrices can be used for multidimensional scaling (INDSCAL analysis, Carroll & Chang, 1970) or for Sequential Information transfer Analysis (SINFA, Wang & Bilger, 1973). SINFA analyses the amount of information transfer for each perceptual phoneme category.
2.2.2. Dichotic discrimination tests

In daily practice the listener, listening to speech, can be distracted by another speech signal present at the same time. To imitate this situation we used a test based on the dichotic discrimination test of Feldmann (1965). In a pilot study we investigated the applicability of the Feldmann test material for the assessment of the benefit of bilateral hearing aids (for details see Boymans & Dreschler, 1993; Dreschler & Boymans, 1994).

Twelve hearing-impaired subjects participated in this experiment with moderate (average loss at 1000, 2000, and 4000 Hz between 40 and 70 dB) and symmetrical (average difference between the ears < 15dB) sensorineural hearing losses. They were recently fitted with two (identical) behind-the-ear hearing aids. We compared the results for the following conditions: right ear provided with a hearing aid, left ear open (condition AD), left ear provided with a hearing aid, right ear open (AS), and bilaterally fitted hearing aids (ADS). The order of conditions was counterbalanced to avoid sequence effects.

In the dichotic speech test two concurrent words (3 syllables) or numbers (4 syllables) were presented exactly simultaneously from $-45^\circ$ and $+45^\circ$ azimuth. The words were recorded from the same speaker. Both words or numbers had to be replicated, if possible. For words the percentages correct replicated syllables for the different sides were calculated. For numbers the correct replicated units and decades were calculated for every side. The realistic aspect of this experiment was that the subjects had to concentrate at both sides simultaneously. In the evaluation of the results of the conditions with an unilateral fitting a distinction was made between the responses at the so-called contra-lateral side (S-contra; words presented at the unaided side) and at the ipsi-lateral side (S-ipsi; words presented at the aided side). The group results of the dichotic discrimination test are presented in Table 2.1.
Assessment of hearing aid candidacy and hearing aid benefit

Table 2.1. Group results for the test on dichotic discrimination. Average values (and st.dev.) are presented for words and numbers separately for the following parameters:

- "unilateral / S-ipsi": the average scores in the unilateral conditions for the speech material from the (unilaterally) aided side.
- "unilateral / S-contra": the average scores in the unilateral conditions for the speech material from the (unilaterally) unaided side.
- "bilateral": the average scores for all speech material presented to the subject wearing two hearing aids.

<table>
<thead>
<tr>
<th></th>
<th>Words</th>
<th>Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Av. Scores</td>
<td>St. dev.</td>
</tr>
<tr>
<td>Unilateral / S-ipsi</td>
<td>33.7 %</td>
<td>17.1</td>
</tr>
<tr>
<td>Unilateral / S-contra</td>
<td>12.0 %</td>
<td>10.5</td>
</tr>
<tr>
<td>Bilateral</td>
<td>31.5 %</td>
<td>16.8</td>
</tr>
</tbody>
</table>

In the unilateral case, ipsi-laterally presented speech material is perceived much better than contra-laterally presented speech material. In the bilateral case, there is only a clear improvement relative to unilateral speech discrimination for the contra-laterally presented speech material. The effect is statistically significant (Wilcoxon, p<0.01), both for words (from 12% to 31.5%) and for numbers (from 32.1% to 57.9%). The results for words and numbers are closely related (correlation coefficient is 0.73).

On average, the perception of ipsi-laterally presented speech information seems to be slightly hampered rather than improved by adding a second hearing aid (and consequently conflicting information). This effect is not in agreement with the results that are usually found in other speech tests, but the effect is only weak (n.s.). The negative trend can be induced by a conflict of attention due to the task to understand both messages.

In this pilot study we found a significant bilateral benefit in dichotic discrimination relative to unilateral conditions with speech at the unaided side. But the effect relative to unilateral
conditions with speech at the aided side was slightly negative. Our results show that dichotic discrimination is much more difficult for words than for numbers. The results suggest that words are too difficult and numbers should be preferred for the dichotic discrimination task.

2.2.3. Intelligibility of sentences

In daily practice speech perception usually concerns the perception of running speech instead of isolated words. Therefore, sentence tests have been developed which can be used to assess objectively the benefits of hearing aids in realistic conditions. The speech reception test (SRT) in noise according to Plomp and Mimpen (1978) is the most well known sentence test used in the Netherlands.

In the SRT-test sentences (spoken with a male or female voice) and noise are presented simultaneously. The noise has a frequency spectrum corresponding to the long-term average spectrum of the speaker and is presented at a constant level (for example 65 dB). The speech level will be varied according to an adaptive up-down procedure following the responses of the subject. The subject repeats the sentences he or she hears. When the sentence cannot be repeated or is not repeated correctly, the next sentence will be presented at a 2 dB higher level each time, until the sentence is repeated completely correctly. Then the next sentence will be presented at a 2 dB lower level, etc., following an adaptive up-down procedure. In total 13 sentences are presented for a single threshold measurement. The average of the last 10 sentences is considered as the SRT-threshold. For normal-hearing subjects, with speech and noise at 0° azimuth, the speech can be presented at about 6 dB below the level of the continuous noise for a 50% correct intelligibility score. Consequently, for normal-hearing subjects the critical S/N ratio is −6 dB (for listening with two ears in the free field). The most important
advantage of this procedure is the high test-retest reliability: standard deviations are in the order of about 1 dB.

In some studies an SRT-test with speech-modulated speech noise is used (Festen & Plomp, 1990) as recorded at the FENAC-CD (Federation of Dutch Audiological Centres). The noise used is speech-noise of a male or a female speaker, modulated according to the modulation spectrum of a single speaker. For normal-hearing subjects the critical S/N ratio is usually 6-10 dB lower in modulated noise than in continuous noise (Duquesnoy, 1983). The reason is that normal-hearing listeners take advantage of the pauses in the background noise. This capacity is affected in hearing-impaired subjects (Festen & Plomp, 1990; Bronkhorst & Plomp, 1992). This results in larger differences between normal-hearing and hearing-impaired listeners in modulated noise than in continuous noise.

For a reliable application of the SRT-test, it is not allowed to use the same sentence lists more than once in the same subjects, because the listener can easily recognize the sentences, even after a long period, and then the test is not reliable any more. Therefore, more speech material with sentences of a male and female voice has been recorded on the VU 98 CD (Versfeld et al., 2000). Again, this CD contains matched background noise signals. Traditionally, the SRT-test is applied with speech and noise from 0° azimuth. But to make the situation more realistic or to incorporate more of the spatial effects that are important in the case of bilateral fitting, the speech and noise sources can be spatially separated. A spatial separation between sound sources usually improves the critical S/N ratio for normal-hearing listeners.

Another speech test is the Oldenburger Satztest (Wagener et al., 1999^a, 1999^b, 1999^c). The speech material consists of a closed set of sentences of five words each. The structure of the sentences is always similar: name-verb-numeral-adjective-object. For each of these five components, 10 words are available. The words can be selected at random and great care is taken to make the transitions between the words as smooth and
natural as possible. The main advantage of this test is that an almost infinite number of different sentences can be constructed and thus this test can be repeated very often. In addition, the test can easily be automated. Of course, the test can also be presented in quiet or in background noise, and if wanted with spatially separated sound sources. At the moment a Dutch/Flemish version of the test is under construction by the Erasmus Medical Centre in Rotterdam, the University of Leuven, and at the AMC in Amsterdam.

2.2.4. Use of the method of adjustment in speech audiometry

As mentioned before, 13 sentences are needed to measure a single critical S/N in the SRT-test. This takes a considerable time, especially when more situations have to be measured. A faster method is the Just Follow Conversation (JFC) test. In contrast to other speech intelligibility tests, this is a subjective speech test. The listener hears sentences in noise and is asked to adjust the speech level by him/her self till he/she could just follow what is being said. The intelligibility of sentences depends on the acoustic features and the redundancy of the sentence. Therefore, it is possible to choose for a closed set of sentences, which will be repeated every time. The listener knows the speech material and can compare the different settings more easily than when the speech material differs every time.

As mentioned above the JFC-test takes less time than the SRT-test. Therefore, more situations with noise from different directions (or different noises), and more hearing aid settings can be tested. When people know the speech material, there is no learning effect, the speech material can be used frequently and the reliability is high. In our test set-up we typically obtain test-retest standard deviations of 1.4 dB. On the other hand, the subjective results are depending on individual criteria and can show large inter-individual differences. The individual criteria are based on speech intelligibility, but could also be based on comfort. The criterion effect can be a problem if individual
measurements have to be compared in absolute terms. For relative measurements (difference measures), this disadvantage is almost absent.

Another implementation of the JFC-test is to use running speech (Neumann et al., 2000). This is less boring, but more difficult to compare different settings in the hearing aid because speech intelligibility is depending on the kind of speech material. The reproducibility of the tests with running speech is usually higher than in our test set-up.

2.2.5. Use of paired comparisons with speech stimuli

In digital hearing aids many settings are possible. But it is not always clear which setting is preferred for each individual. Therefore, subjective judgements are useful additive to more objective information from speech tests. For this purpose, paired comparisons can be used with speech stimuli (Franck et al., 2003). In a paired comparison the subject can make direct comparisons between speech fragments reproduced by a hearing aid in different settings. The subject hears the same sentences for two different hearing aid settings that are to be compared. The sentences can be presented in quiet or in background noise. The subject has to judge which hearing aid setting is preferred, taken into consideration that the hearing aid setting should be used for the whole day. A set of combinations of hearing aid settings can be presented in a tournament-like procedure in order to find the setting that is judged most frequently as the best (i.e. the winner).

2.2.6. Applications of speech stimuli for the evaluation of hearing aid benefit

A lot of evaluation speech tests are possible, but which test do we have to choose? This depends on different factors:
- The kind of information needed.
The relationship with daily communication.

The accuracy of the measurement versus the time requirements.

The degree of difficulty for the subject.

For every study we want an accurate, detailed, easy and fast test, but not everything is possible in the same test. Therefore we have to prioritise.

The first question is: What do we want to measure? Do we need intelligibility scores, speech reception thresholds, phoneme confusion patterns or subjective preferences?

For speech reception thresholds, SRT-tests remain the “golden” standard, but if a large number of conditions have to be compared JFC-tests can be used as a first-order approximation. For comparative measurements between different settings or between different hearing aids, we can start with the measurement of the amount of speech intelligibility before analysing the kind of errors or substitutions made. From an analysis of confusions (as obtained with the test developed by Steeneken) we can gather more qualitative information about the reasons for poor speech perception and/or the effects of signal processing parameters (e.g. attack and release times) on characteristic properties of phoneme identification. For specific aspects of binaural processing, for example the effect of a second hearing aid, the dichotic discrimination test can be used. For subjective measurements, a paired-comparison test can be useful when a direct comparison is needed between different hearing aid settings.

The second question is: How realistic should the test be? In daily practice we speak in sentences, so a sentence test is more realistic than a test with words. However a word-test is easier to analyse phoneme confusions. And with sentences we have to take into account the redundancy that is present in sentences.

Most of the time we have the possibility of lip reading. But for a speech test this gives a lot of bias: we have to separate what is being heard from what is being seen. Therefore, tests without lip reading are generally preferred. In daily practice, often different background noises are present and they are coming from different directions. This is problematic for all hearing-impaired listeners. To make a test realistic, it is useful to
imitate those difficult situations. Background noise could be added in all tests, in principle. In daily practise background noise is often speech, and speech can distract, because there is conflicting information in it. To imitate such a situation, a dichotic discrimination test with words or numbers can be chosen.

The third question is related to the interrelated items of accuracy and time consumption. The more detailed information the more time is needed. For detailed information about the specific difficulties in identifying different speech sounds the CVC-identification test, developed by Steeneken, can be used. The test is time-consuming. So the hearing-impaired listener has to concentrate for quite a long time, especially when more conditions have to be measured. Therefore, we should try to avoid that we measure the concentration of the subject instead of his/her speech perception abilities, especially for older people. As discussed above, the SRT-test is an objective speech test and measures the 50 % point of speech intelligibility. The JFC-test is a more subjective measurement converging to an unknown (individually chosen) criterion, but this test is much faster than the SRT-test. For comparative measurements with a lot of different situations the JFC-test can be considered. A paired comparison test is also a subjective test but the result is only a rank order, and with the JFC-test, more specific judgements are measured.

The last question concerns the degree of difficulty for the subject to conduct the test. This depends on the individual subject, but in general the dichotic discrimination test is the most difficult test. For this test we should take into account the concentration of the subject. Apparently, this test is more difficult when words are used instead of numbers. The SRT-test is not experienced as difficult, but of course the subject needs to concentrate and the tests may not last too long. A paired-comparison test is probably more difficult than a JFC-test. Because with the JFC-test the subject can make his/her own reference better by adjusting the gain of the presented sentences, but a disadvantage is that there is no direct comparison in the JFC-test. When the hearing loss
is too large for speech intelligibility, lip reading can be added to make the test less
difficult. However, not all test material discussed in this section is available with
accompanying video material. There is need for further development of test material
that can be applied audio-visually.

2.3. Subjective evaluation techniques

Differences in hearing aids are not always measurable with objective tests. Besides, the
subjective experiences are important factors. Hearing-impaired people do have different
impairments, experience different amounts of disabilities and feel different degrees of
handicap, wear their hearing aids in different acoustical situations, and experience
different benefits of their hearing aid(s). To map out all those subjective information a
lot of questionnaires have been developed.

2.3.1. Traditional hearing aid questionnaires in the Netherlands

There are a lot of questionnaires in circulation, but only a few of them are validated.
The “Hearing Handicap and Disability Inventory” (van den Brink, 1995) is validated
and focuses on disability and handicap. There is a complete list of 40 questions, and for
brief measurements an abbreviated list with 20 questions has been developed (10
questions about disability and 10 questions about handicap). The hearing-impaired
listeners are asked to answer the questions for common situations with a hearing aid (or
without a hearing aid when this is more usual). For the answers a 4-points scale is used.
Questions are asked for different situations like: a quiet situation, a noisy situation, the
use of telephone, attending a lecture, listening to television, and visiting a shop.
Another validated questionnaire is the Amsterdam Inventory of Auditory Disability and Handicap, which consists of 30 questions (Kramer et al., 1995; Kramer, 1998). The questions are distributed in five basic disability factors: detection of sounds (5 questions), distinction of sounds (8), intelligibility in quiet (5), intelligibility in noise (5), and auditory localization (5). Each question consists of three parts, the first part is about disability at that specific moment, the second part is about the situation in the past, and the last part is about the handicap. Four answer categories were possible. The ‘handicap-question’ is about the extent to which the hearing-impaired subjects are annoyed by the experience of difficulty in hearing in that specific situation and the extent to which they are limited in doing activities. When there is no difficulty in hearing in a specific situation, the hearing-impaired is instructed to skip the handicap part. The questionnaire could be filled in for situations without and/or with a hearing aid.

In Rotterdam a questionnaire was developed (Franck et al., 1999), with questions about the hearing aid in general (sounds, function, frequency of wearing the hearing aid etc.) and about speech intelligibility with the hearing aid in different situations. Situations at home, outside, at work, and at school. The subjects are also asked to fill in how often a situation occurred and how important that situation was for the subject. They are asked to visualise their answer on a visual-analogue-scale. This is a horizontal unmarked line, with end markers such as “good” and “bad” (two extremes). When the subjective rating corresponds to a very good intelligibility he/she has to make a vertical line at the horizontal line near the word “good”. When the subjective judgement is about 50% the vertical line has to be placed in the middle of the line. This questionnaire is not validated, but gives a good impression about the subjective experiences with different programs in hearing aids or with different hearing aids.
2.3.2. Traditional international hearing aid questionnaires

The Abbreviated Profile of Hearing Aid Benefit (APHAB) is a questionnaire that can be used as part of the fitting procedure (Cox et al., 1995). Firstly, questions are asked about the experience with hearing aids, hearing aid use and about the working situation. Then 24 questions (this is a subset of the original PHAB-questions, Cox, 1990) are asked, which refer to four subscales: ease of communication under relatively favourable conditions, communication in reverberant rooms, communication in settings with high background noise levels, and unpleasantness / aversiveness of environmental sounds. Each item is a statement. The hearing-impaired listener is asked to rate the truth of that specific statement on a 7-point scale, for the situation without a hearing aid and with a hearing aid. So differences between both situations can be measured. It is also possible to answer those questions for example with two different hearing aids, or two different settings of the hearing aid in order to determine whether one is significantly superior.

The International Outcome Inventory for Hearing aids (IOI-HA) is developed as a product of an international workshop on Self-Report Outcome measures in Audiological Rehabilitation (Cox et al., 2000). This questionnaire is translated in different languages to facilitate co-operation among researchers in different hearing healthcare settings across national boundaries. The questionnaire consists of only seven questions, with answer possibilities at a five-point scale. One question is about the frequency of hearing aid use, three questions about the residual handicap (factor 2), and three questions about the benefit or satisfaction of the hearing aid (factor 1).

The Glasgow Hearing Aid Benefit Profile (GHABP; Gatehouse 1999) is a questionnaire with eight listening situations. Four situations are pre-specified and four situations are user-specified. Questions are asked about initial disability, handicap, hearing aid use, hearing aid benefit, residual disability and satisfaction for each of these eight conditions. The subjects are asked to answer the questions on a 5-points scale.
2.3.3. Composition of AVETA to evaluate bilateral benefit

For the evaluation of the benefit of bilateral hearing aids there was a co-operative effort of the Free University Amsterdam and our lab to compose a specialised questionnaire from existing questionnaires.

For a retrospective study some general questions were included from other questionnaires and more specific questions were added about the reasons for choosing one or two hearing aids.

A large part of the questionnaire exists of questions about the situations without a hearing aid, with one hearing aid and with two hearing aids. For that purpose parts of the adjusted version of the Amsterdam Inventory Disability and Handicap (AIADH) and the Abbreviated Profile of Hearing Aid Benefit (APHAB) were used. In total 7 categories were composed: detection of sounds (5 questions), speech intelligibility in quiet (5 questions), speech intelligibility in noise (5 questions), directional hearing or localization (5 questions), discrimination or recognition of sounds (1 question), speech intelligibility in reverberation (1 question from the APHAB), and comfort of loud sounds (6 questions from the APHAB). Ten questions from the HHDI were used to get information about handicap.

All seven questions of the new IOI-HA were used to get information about hearing aid use, residual handicap and benefit or satisfaction of the hearing aid. Details about this questionnaire were described by Kramer et al. (2002).

Because the questionnaire of the retrospective study was rather long, we applied – after validation based on the results of the retrospective study - a shortened version in the prospective study.

We still used general questions about the daily situation of the subject and about the reasons for choosing one or two hearing aids. But the selection of questions from the
AIADH and APHAB was decreased from 29 to 18. The question about speech in reverberation was skipped. So, only six categories were left (detection of sounds, speech intelligibility in quiet, speech intelligibility in noise, directional hearing or localization, discrimination or recognition of sounds, and comfort of loud sounds) and for each category three questions were included, selected on the basis of the analyses of the retrospective results.

Ten questions from the HHDI were omitted because there was too much overlap with the handicap part of the IOI-HA. The IOI-HA was included as an integral part. The resulting validated questionnaire is called AVETA (Dutch acronym for Amsterdam Questionnaire for Unilateral or Bilateral Hearing Aid Fittings).
Assessment of hearing aid candidacy and hearing aid benefit