Charactering the individual ear by the "Auditory Profile"

Wouter Dreschler[a], Thamar Esch Van[b], Birgitta Larsby[c], Mathias Hallgren[c], Mark Lutman[d], Johannes Lyzenga[e], M Vormann[f] and Birger Kollmeier[g]

[a]AMC, Clinical and Experimental Audiology, 1105 Amsterdam, Netherlands
[b]AMC - Dept. of Clinical and Experimental Audiology, Meibergdreef 9, 1105AZ Amsterdam, Netherlands
[c]Linkoepings Universitet, Hus Origo Campus Valla, 581 83 Linkoeping, Sweden
[d]University of Southampton, University Road, Highfield / ISVR, SO17 iBJ Southampton, UK
[e]Vrije Universiteit Medical Center, Boelelaan 1117, 1081 HV Amsterdam, Netherlands
[f]Hoerzentrum Oldenburg, Hoerzentrum Oldenburg, 26129 Oldenburg, Germany
[g]Universität Oldenburg, Medizinische Physik, Carl-von-Ossietzky Str. 9-11, 26111 Oldenburg, Germany
w.a.dreschler@amc.uva.nl

This paper describes a new approach to auditory diagnostics, which is one of the central themes of the EU-project HEARCOM. For this purpose we defined a so-called "Auditory Profile" that can be assessed for each individual listener using a standardized battery of audiological tests that – in addition to the pure-tone audiogram - focus on loudness perception, frequency resolution, temporal resolution, speech perception, binaural functioning, listening effort, subjective hearing abilities, and cognition. For the sake of testing time only summary tests are included from each of these areas, but the broad approach of characterizing auditory communication problems by means of standardized tests is expected to have an added value above traditional testing in understanding the reasons for poor speech reception. The Auditory Profile may also be relevant in the field of auditory rehabilitation and for design of acoustical environments. The results of an international 5-center study (in 4 countries and in 4 languages) will be presented and the relevance of a broad but well-standardized approach will be discussed.

1 Introduction

The EU project HEARCOM (acronym for Hearing in the Communication Society, see www.hearcom.eu) aims at full participation in the modern communication society by reducing the limitations in auditory communication.

Two of the focus areas of HEARCOM are:

- The identification and characterization of auditory communication limitations
- The development of standardized testing and evaluation procedures

There is still lack of knowledge about the causes of poor speech perception in the individual hearing-impaired person, especially in more complex listening environments with (fluctuating) noise and reverberation. For this reason an "Auditory Profile" has been defined.

The auditory profile (AP) should be applicable as a diagnostic tool in a broad population of subjects with complaints about their performance in (auditory) communication tasks. The diagnostic scope here is not primarily on the underlying impairment, but on auditory disabilities that impact auditory functioning in daily life. After definition, implementation, and verification, the AP may become a standard approach in (specialized) hearing centers and clinics. For this reason, we also consider the potential value of the AP with respect to hearing aid selection and hearing aid fitting.

2 Design of the AP

Consensus within HEARCOM has been reached about a standardized battery of audiological tests that – in addition to the pure-tone audiogram - can be applied to characterize the residual capacities of the hearing-impaired subject in the AP. The auditory profile should include all necessary measures to describe the main characteristics and differences between different hearing impairments. On the other hand, the auditory profile should minimize redundancy between measures. International co-operation allowed comparisons of the audiometric results across countries, even for the speech tests.

The components of the AP should be relevant for auditory communication performance. Usually most emphasis is given to speech perception, but the scope of the auditory profile is broader: the profile should also be related to signal recognition, sound quality, spatial hearing, listening

comfort, listening effort, and adequate processing of sounds. A limited set of tests will never be able to cover all aspects in detail, but the aim is that the AP is broad enough to cover at least the main parameters in these areas.

More specifically, the partners selected seven fields for testing, listed in the first column of Table 1. In each of these fields an inventory of available tests was made and pilot studies were conducted to compare tests.

To be applicable in a clinical environment, also some extra methodological requirements were taken into account:

- Tests should be reliable and reproducible;
- Tests should not exhibit strong learning effects;
- Test procedures should be well described;
- Tests should be applicable in a large variety of hearing impairments.

Field	Test
Loudness perception	Acalos
Frequency resolution and temporal acuity	Combined FT-test
Speech perception	SRT in quiet, stationary and fluctuating noise
Binaural processing	MAA, ILD and BILD
Subjective judgements	Gothenburg Profile Effort scaling speech in noise
Cognitive abilities	Lexical decision making test

Table 1 List of tests included in the AP.

In a consensus meeting appropriate tests have been selected to be included in the preliminary auditory profile, according to the second column of Table 1, based on the pilot studies and the above-mentioned criteria.

One of the most problematic issues is the large number of relevant areas versus the limited testing time available. For the preliminary AP, testing time was constrained to 120 minutes for the complete set of tests (not including standard audiometry). A further reduction of testing time can be realized, based on the results of the preliminary AP. One possibility is a hierarchical structure with limited tests in each of the areas of interest and more detailed tests in areas in which problems appear.

3   Methods used in the AP

Audibility

Pure-tone air- and bone-conduction thresholds are measured at octave frequencies from 250 to 8000 Hz, using standard audiometric procedures with adequate masking of the contra-lateral ear. Air-conduction thresholds are also measured at 3000 and 6000 Hz.

Loudness perception

Loudness perception was measured using ACALOS (Adaptive CAtegorical LOudness Scaling, see [1]), which estimates the loudness growth function on a scale from 0-50, where 50 is "too loud". Measurements were performed using 1/3-octave bands of so called ′low-noise noise′ at 500 and 3000 Hz, and using the broadband speech-shaped ICRA1 (or ICRA1_female) noise. From these measurements most comfortable loudness levels are derived (MCL_{low}, level in dB SPL at 20 categorical loudness units, cu). For all subjects, all the following tests were conducted at equal loudness levels: the MCL_{low} level that will be called MCL in further descriptions. For speech tests and other broadband measurements, the MCL as derived with the speech-shaped noise was used as measurement level (with a maximum of 85 dB) and for narrowband tests MCL as derived with corresponding narrowband noises (with a maximum of 95 dB) were used. For all binaural measurements, MCL of the subject's better ear was used.

Frequency resolution and temporal acuity

The F-T test of Larsby and Arlinger [2] was used to measure spectral and temporal resolution. Masked thresholds of tone pulses in three different noises were measured: octave-band stationary noise, noise with spectral gaps (around signal frequency), and noise with temporal gaps (coinciding with the signals). Thresholds were estimated using a Békésy tracking procedure. Measurements were conducted at 500 and 3000 Hz, in both ears separately. The masking noise is fixed at MCL, and signal level is varied.

Speech perception

Speech perception was measured using Plomp-type [3] sentence tests:

- In quiet, diotically;
- In stationary noise (ICRA-1 or ICRA-1_female, same gender as the speaker), monaurally in both ears [4];
- In fluctuating noise (ICRA-5_250 or ICRA-4_250, same gender as the speaker), monaurally at both ears.

The noise level was fixed, and the speech level was varied. Outcome measure is the speech reception threshold (SRT): the signal to noise ratio (SNR) for 50% correct, except for the quiet condition (the speech level for 50% correct).

Binaural processing

Three tests were conducted involving binaural processing: intelligibility level difference test (ILD), binaural intelligibility level difference test (BILD), and the minimal audible angle test (MAA). As these tests are all conducted via headphones, virtual stimuli are used. This means that all signals were filtered with generic Head-Related Transfer Functions (HRTF) to simulate different directions. ILD and BILD are measured with Hagerman-type sentences [5] with the noise level fixed and varying speech level. These sentences all have a fixed structure, generated from ten names, ten verbs, ten numerals, ten adjectives, and ten objects.

ILD test

For this test, speech recognition thresholds were measured in three conditions with speech-shaped noise:

- S_0N_0: speech and noise both coming from the front (0°) ;
- S_0N_{90}: speech coming from the front (0°) and noise coming from the right side (90°);
- S_0N_{-90}: speech coming from the front (0°) and noise coming from the left side (−90°).

The ILD represents the SRT difference between S_0N_0 and S_0N_{90} or S_0N_{-90} results.

BILD test

To estimate the BILD, two additional, monaural, measurements were conducted:

- S_0N_{90}: speech coming from the front (0°) and noise coming from the right side (90°) with the right ear blocked (so both signals are presented monaurally to the left ear);
- S_0N_{-90}: speech coming from the front (0°) and noise coming from the left side (−90°) with the left ear blocked (both signals presented monaurally to the right ear).

The BILD represents the SRT difference between monaural and binaural S_0N_{90} and S_0N_{-90} results.

MAA test

To test sound localisation ability, a virtual headphone version of the minimal audible angle (MAA) test was used. This test measures the just noticeable difference (JND) in (virtual) horizontal sound direction. Two stimuli were presented consecutively from different directions, symmetrically spaced on different sides of the straight-ahead direction. The order of the sounds (left first or right first) was randomised. The listener's task was to indicate the order of the two sounds. If the two sounds are perceived from different angles the result is the impression of a moving sound. Was the sound going from left to right or from right to left? The sounds were:

- Low-pass noise (filtered at 1500 Hz) to investigate the use of interaural time difference;
- High-pass noise (filtered at 3000 Hz) to investigate the use of interaural level differences;
- Broadband white noise to investigate the interaction between the two difference cues.

Measurements were performed at MCL: MCL at 500 Hz for low-pass noise, MCL at 3000 Hz for high-pass noise, and MCL measured with ICRA1 noise for broadband noise.

Self-report measures

Gothenburg profile

Subjects were asked to fill in the Gothenburg Profile [6] on a PC. This questionnaire measures experienced hearing disability and handicap. It consists of 20 items divided into

four subscales: 'speech perception', 'spatial hearing', 'social interactions' and 'behaviour and reaction'.

Listening effort

Subjects were asked to indicate their experienced effort on a scale while listening binaurally to running speech (fairy tales) in four different conditions: in ICRA-1_female noise at S/N = +5 dB, in ICRA-1_female noise at S/N = -5 dB, in ICRA-4_250 noise at S/N = +5 dB, and in ICRA-4_250 noise at S/N = -5 dB with the noise level fixed in all conditions.

Cognitive abilities

A measure of cognitive abilities was obtained using the Lexical decision-making test [7], which estimates the lexical access of subjects. During the test, items were selected at random from lists of real words and non-words and presented as text on a computer screen. Subjects had to indicate the nature of the presented item (word or non-word) by pressing the corresponding button. Outcome measure of this test is percentage correct divided by average response time.

4   Implementation

All tests have been implemented as headphone tests using a common software platform developed by HörTech in Oldenburg (called OMA, Oldenburg Measurement Applications). The tests on OMA are now available in four languages: English, German, Dutch and Swedish. This applies to all speech and language tests (Speech perception tests, ILD, BILD, and Lexical decision-making test) and to the subjective judgement tests (Effort scaling and the Gothenburg profile). Two types of speech material are available for speech testing: everyday sentences with an open structure (Plomp-type sentences [3]) and artificially composed sentences (Hagerman-type sentences, developed within the HEARCOM project by HörTech, see [8]). It is clear that some differences may occur due to language-specific speech material and testing procedures. Therefore, we collected an extra set of reference data for each language and the results of the reference data were used to calculate corrected results of the speech tests. The same holds for the lexical-decision test. In the results of the multi-center study we will use language-corrected data only. Also a language-specific correction factor was applied to the results of the Gothenburg Profile.

5   Validation in multi-center test

Seventy-three hearing-impaired subjects were invited to participate in this study on a voluntary basis (the target was 15 for each of the participating centres). They are selected from the clinical population according to the following inclusion criteria:

- Age between 18 and 75 years;

- Difference in average pure tone thresholds between the ears < 30 dB;

- No language problems;

- Active and alert and able to perform the tests;

- No complaints of tinnitus.

A control group of 30 normal-hearing subjects (all pure-tone audiogram thresholds better than 20 dB), aged between 18 and 50 years, was included.

The results of this extensive study will be described in detail in a separate paper.

Given the limited space and for the purpose of this paper, it is relevant to have a first glance on the main results.

We investigated the test-retest reliability of the tests by calculating the intra-class correlation coefficient (ICC) for the total group, hearing-impaired listeners and normal-hearing listeners. No clinically relevant learning effects were found: all differences between test and retest values were much smaller than the within-subject standard deviation. We decided to use test values rather than averages of test and retest in further analyses, because this is more clinically relevant. Of course we used the information about effects of test-retest and learning when deciding about including or excluding parameters in the final AP.

5.1   Per-ear variables

First of all, the multi-center study yields a large number of outcome measures per ear. Before we analyzed the relations between per-ear variables, we investigated their distributions to see if they deviated from normal distributions. We did this by performing Kolmogorov-Smirnov tests and by visual inspection. All variables except air-bone gap were distributed approximately normally. We transformed the ABGs using BLOM. Next we performed correlation analyses, partial correlation analyses (partialing out the effects of audibility), regression analyses, and factor analyses for the total group and for the subgroups of hearing-impaired and normally-hearing subjects.

The data provide clear evidence that speech perception is not only determined by audibility. For the hearing-impaired group speech-perception in stationary noise is mainly determined by the hearing loss at 3 kHz and by the frequency resolution at 500 Hz. In fluctuating noise also the temporal resolution at 3 kHz becomes relevant.

The factor analysis for the hearing-impaired group reveals four independent factors, explaining more than 66% of the total variance:

- Factor 1 with high factor loadings for the slope of the loudness curve at 3 kHz and the frequency resolution at 3 kHz (>0.7), representing about 20 % of the variance and interpreted as *high-frequency processing*;

- Factor 2 with high loadings for all MCL-values (>0.7) and the audiometric thresholds at 500 and 3 kHz (>0.6), explaining about 20 % of the variance and interpreted as *audibility*;

- Factor 3 with high factor loadings for the slope of the loudness curves at 500 Hz and for broadband stimuli (>0.7), explaining about 14% of the variance and interpreted as *recruitment*;

- Factor 4 with high factor loadings for the frequency and time resolution at 500 Hz (>0.7), explaining almost 13% of the variance and interpreted as *low-frequency processing*.

5.2 Per-subject variables

Some of the test results cannot be interpreted as per-ear results and we call them 'per subject' variables. These include the binaural speech tests, the MAA-tests, and the lexical decision-making test. We tested normality both by Kolmogorov-Smirnov tests and by visual inspection. We found that all variables except the MAA variables were distributed approximately normally. Consequently, we transformed the MAA variables using BLOM.

In the results of the per-subject variables, most significant correlations were found between parameters derived from the same type of tests. As a result the factor analysis yielded three independent factors, related to:

- *Sound localization* (high factor loadings for all MAA-parameters, explaining almost 30% of the variance);

- *Binaural speech perception* (high factor loadings for SRT in quiet, ILD and BILD, explaining about 20% of the variance);

- *Cognition* (high factor loadings for lexical decision and age, explaining almost 20% of the variance).

5.3 Communication performance

For the subjective tests on communication performance (Gothenburg profile and Listening effort) again the normality of the distributions was verified. We decided to transform the Gothenburg-Profile variables because of their skewed distributions. Distributions of the listening-effort results were approximately normal.

Next, Pearson's correlations were calculated between the communication-performance parameters and the selected 'per-ear' for the better ear and 'per-subject' variables.

It turns out that there is very little correspondence between especially the listening-effort test and the per-ear and 'binaural' tests, except for the speech-perception tests. Listening effort cannot easily be predicted by objective measures. Even correlations between the audiogram or age and listening effort are not significant, except for the pure-tone loss at 3 kHz. Only speech-perception tests are statistically related to perceived effort. It's remarkable that listening effort in fluctuating noise shows more correspondence with other tests than listening effort in continuous noise.

Also the outcomes of the Gothenburg Profile are more strongly related to speech-perception tests than to other per-ear and 'binaural' outcome measures. But for the 'speech-perception' and 'spatial-hearing' subscales of the Gothenburg profile, the low-frequency processing (slope of the loudness curve at 500 Hz and the frequency resolution at 500 Hz) are relevant. The subscales for 'social interactions' and 'behaviour and reactions' are only significantly related to the SRT-results.

6 Relevance of the AP for hearing aid fitting

It should be stressed that the AP described above is primarily focused on the diagnosis of auditory functioning. For the purpose of auditory rehabilitation, some extra tests may be needed in order to select, fit, and evaluate hearing aids.

For application of the AP in rehabilitative audiology, an important question is whether the AP can be used for a classification of HI that has consequences for his/her auditory rehabilitation. Although such an application needs to be investigated in a new validation study, it may be worthwhile to start with some "well-educated" guesses, partly based on clinical experience and partly on the outcomes of the preliminary version of the AP.

The AP may be expected to be able to discriminate between subjects in whom audibility is the main cause of disability and subjects that have more complex disabilities, related to a (severely) reduced dynamic range, reduced frequency resolution, or reduced temporal resolution. In other subjects, problems in binaural integration or problems with cognition may appear.

In our opinion, the AP in its final form can be an important support for hearing aid selection and fitting with respect to the following aspects:

- The Gothenburg Profile may be expected to be a powerful tool to assess the need of a hearing aid in a well-structured way;

- The potential benefit can be predicted from the speech reception tests in noise, not only in stationary, but even more important in fluctuating noise. For counseling purposes, it is important to have valid predictions of the potential benefit. Expectations should be as realistic as possible;

- The choice for one or two hearing aids will be supported by the tests on binaural integration (MAA, ILD, and BILD);

- Options for compression and output limiting can be based on the ACALOS results;

- Noise reduction (including directional microphones) may be considered especially in cases of poor speech perception in noise (from the SRT tests);

- Spectral sharpening may be applied in case of a reduced frequency resolution;

- Fast compression and echo suppression may be used in case of a reduced temporal resolution;

This way the AP can support rehabilitative audiology. In addition, some of the AP tests conducted with headphones can set targets that should be reproduced by subjects with hearing aids in a free-field condition. This applies to the SRT-tests and the tests on binaural functioning (ILD/BILD). By means of an ACALOS test in aided conditions, the effect of amplification on the dynamic range can be evaluated.

7   Further streamlining of the AP

For clinical applications, both in in-depth diagnostics and for application as a diagnostic step in auditory rehabilitation, we feel that the AP should be reduced further. Some of the measures are redundant because they correlate highly with one another. Furthermore, there were some unsatisfactory features of the AP tests that can be improved with minor modifications.

One important issue is the choice of the presentation level. Measurement errors in determining MCL are perpetuated through all the outcome measures through dependence of test outcomes on presentation level. These considerations suggest that basing presentation levels for most tests on MCL reduces the value of the AP, by reducing its ability to demonstrate the dimensions of variation that we wish to investigate. However, the approach of basing stimulus levels on MCL has the advantage of setting a comfortable level for all participants, regardless of hearing impairment. Use of a fixed level for all participants may mean that it is uncomfortably loud for some participants with normal hearing and/or too quiet for some hearing-impaired participants.

In our opinion 60 minutes of testing is the maximum and a stepwise approach could help to reduce testing time even further for individual cases.

8   Discussion and conclusion

International consensus is growing for a broad battery of audiological tests to characterize the residual capacities of the impaired ear. The results of the multi-center study show that the Auditory Profile allows a detailed analysis of auditory disabilities by a very broad diagnosis of auditory deficits. In many subjects problems in auditory communication are not only caused by reduced audibility, but also by a different loudness perception, reduced supra-threshold resolution, reduced binaural cooperation, or problems in cognition. It is worthwhile to assess the strength of contributing factors in individual subjects.

This is work in progress. By factor analyses on the preliminary set of data we found a clustering of test results that indicate that hearing impairment is a multi-dimensional problem. The Auditory Profile is a powerful means to cover different dimensions that are shown to be relevant for auditory problems. The implementation of the tests on a uniform software platform will facilitate clinical application.

The outcomes of the Auditory Profile will help us to understand the causes of the problems and to find the best solutions, either in acoustical requirements (HEARCOM subproject SP2), in signal processing strategies for advanced hearing aids (HEARCOM subproject SP3), or in assistive listening devices (HEARCOM subproject SP4). It is our ambition to set new European standards in Audiology. If the Auditory Profile is able to estimate the problems that individual subjects will encounter in adverse communication situations, this work may stimulate a broad clinical acceptance of such a broad innovative approach to auditory testing.

References

[1] Brand T, Hohmann V (2002) An adaptive procedure for categorical loudness scaling. *J Acoust Soc Am* 112:1597-1604.

[2] Larsby B, Arlinger S (1998) A method for evaluating temporal, spectral and combined temporal-spectral resolution of hearing. *Scand Audio*l 27:3-12.

[3] Plomp R, Mimpen AM (1979a) Improving the reliability of testing the speech reception threshold for sentences. *Audiology* 18:43-52.

[4] Dreschler WA, Verschuure H, Ludvigsen C, Westermann S (2001). ICRA Noises: Artificial Noise Signals with Speech-like Spectral and Temporal Properties for Hearing Instrument Assessment. *Audiology* 2001; 40: 148-157

[5] Hagerman B (1982) Sentences for testing speech intelligibility in noise. *Scand Audiol* 11:79-87.

[6] Ringdahl A, Eriksson-Mangold M, Andersson G (1998) Psychometric evaluation of the Gothenburg Profile for measurement of experienced hearing disability and handicap: applications with new hearing aid candidates and experienced hearing aid users. Br J *Audiology* 32:375-385.

[7] Larsby B, Hällgren M, Lyxell B, Arlinger S (2005) Cognitive performance and perceived effort in speech processing tasks: effects of different noise backgrounds in normals and in hearing-impaired subjects. *International Journal of Audiology*, 44, 131-143.

[8] Wagener K, Brand T, Kollmeier B (2007) International cross-validation of sentence intelligibility tests. *Proceedings EFAS conference Heidelberg*, 2007.