The “Auditory Profile”: Proposal from the European HEARCOM project

Wouter A. Dreschler[1], Thamar van Esch[1], Birgitta Larsby[2], Mathias Hällgren[2], Mark E. Lutman[3], Johannes Lyzenga[4], Matthias Vormann[5], Birger Kollmeier[5]

[1] *Academic Medical Center, Amsterdam, The Netherlands*
[2] *Linköping University, Linköping, Sweden,*
[3] *Institute for Sound and Vibration Research, Southampton, United Kingdom*
[4] *VU University Hospital, Amsterdam, The Netherlands*
[5] *HörTech GmbH, Oldenburg, Germany*

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 acuity, 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 test 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.

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 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 auditory profile may become a standard approach in (specialized) hearing centres and clinics.

Auditory signal processing in hearing-impaired listeners. 1st International Symposium on Auditory and Audiological Research (ISAAR 2007). T. Dau, J. M. Buchholz, J. M. Harte, T. U. Christiansen (Eds.).
ISBN: 87-990013-1-4. Print: Centertryk A/S.

THE DESIGN OF THE PRELIMINARY AUDITORY PROFILE

Consensus within HEARCOM has been reached about a standardised battery of audiological tests that – in addition to the pure-tone audiogram - can be applied to characterise the residual capacities of the hearing-impaired subject in the Auditory Profile. 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 will allow the comparison of the audiometric results across countries, even for the speech tests.

The components of the auditory profile 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 auditory profile is broad enough to cover at least the main parameters in these areas.

More specifically, the partners selected the following seven fields for testing:

- Audibility and loudness perception
- Frequency resolution and temporal acuity
- Speech perception in noise
- Binaural processing
- Subjective judgments and listening effort
- Cognitive abilities

In each of these fields an inventory of available tests was made and in a consensus meeting appropriate tests have been selected to be included in the preliminary auditory profile.

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

- The tests should be reliable and reproducible
- The tests should not exhibit strong learning effects
- The tests should be efficient because of limited testing time
- The test procedures should be well described
- The test should be applicable in a large variety of hearing impairments

One of the most problematic issues is the large number of relevant areas (see above) versus the limited testing time available. For the preliminary Auditory Profile, testing time was constrained to 90 minutes for the complete set of tests (not including standard audiometry). A further reduction of testing time should be realized, based on the results of the preliminary auditory profile. 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. The preliminary auditory profile now contains the tests listed in Table 1.

It should be stressed that the auditory profile described above is primarily focussed 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.

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 and effort scaling for speech in noise
Cognitive abilities	Lexical decision making test

Table 1: List of tests included in the auditory profile.

METHODS USED IN THE AUDITORY PROFILE

Audibility

Pure-tone thresholds are measured using a standard audiometer. Air-conduction thresholds are measured at 250, 500, 1000, 2000, 3000, 4000, 6000, and 8000 Hz and bone-conduction thresholds at 250, 500, 1000, and 3000 Hz with adequate masking of the contra-lateral ear.

Loudness perception

Loudness perception was measured using ACALOS (Adaptive CAtegorical LOudness Scaling, Brand and Hohmann, 2002), 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 noise (depending on the gender of the speaker in the speech tests, see below, a male- or female-weighted version of this noise was used). 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 measurement, 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, the MCL of the subject's better ear was used.

Frequency resolution and temporal acuity

The F-T test of Larsby and Arlinger (1998) 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 (1979) sentence tests:

- In quiet, diotically
- In stationary noise (ICRA-1, male- or female-weighted version, same gender as the speaker), monaurally in both ears
- 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 recognition 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 (1982) 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°)
- $S0N_{-90}$: 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 speech-shaped noise.

Self-report measures

Gothenburg profile

Subjects were asked to fill in the Gothenburg Profile (Ringdahl, 1998) on a PC. This questionnaire measures experienced hearing disability and handicap. It consists of 20 items divided into four subscales: 'speech percpetion', '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 noise at S/N = +5 dB, in ICRA-1 noise at S/N = -5 dB, in ICRA-5 noise at S/N = +5 dB, and in ICRA-5 noise at S/N = -5 dB with the noise level fixed in all conditions. A male- or female-weighted version of the ICRA-1 noise was used, depending on the gender of the speaker.

Cognitive abilities

A measure of cognitive abilities was obtained using the Lexical decision-making test (Larsby, 2005), 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.

IMPLEMENTATION AND CROSS-LANGUAGE VALIDATION

Besides the pure-tone audiogram, all tests have been implemented as headphone tests, using a common software platform developed by HörTech in Oldenburg (called OMA, Oldenburg Measurement Applications). This will facilitate the dissemination of the tests to the target groups described above. 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).

For speech testing two types of speech material are available: everyday sentences with an open structure (Plomp-type sentences) and artificially composed sentences (Hagerman-type sentences). The Hagerman-type sentences in English, German and Dutch have been developed within the HEARCOM project, with a large involvement of the partners of HörTech (see Wagener, 1999). 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 (see Wagener et al., 2007). The same holds for the lexical-decision test. In the results of the multi-centre study we will use language-corrected data only.

VERIFICATION IN AN INTERNATIONAL MULTI-CENTRE STUDY

Seventy-five hearing-impaired subjects were invited to participate in this study on a voluntary basis (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.
- Maximum difference in PTA between the two ears of 30 dB
- No language problems.
- Active and alert and able to perform the tests.
- No complaints of tinnitus.

In addition, 40 normal-hearing subjects (all pure-tone audiogram thresholds better than 20 dB), aged between 18 and 50 years, were included.

This paper presents the preliminary results obtained in 56 hearing-impaired subjects (HI) and 27 normal-hearing participants (NH). A description of the main results is given in Whisker-Box plots, to show inter-individual variations and the extreme values. In each of the following figures, the results of the NH group are plotted at the left-hand side and the results of the HI-group at the right-hand side. For the tests that were measured per ear, results of the better ear of each subject (as defined by PTA) are presented.

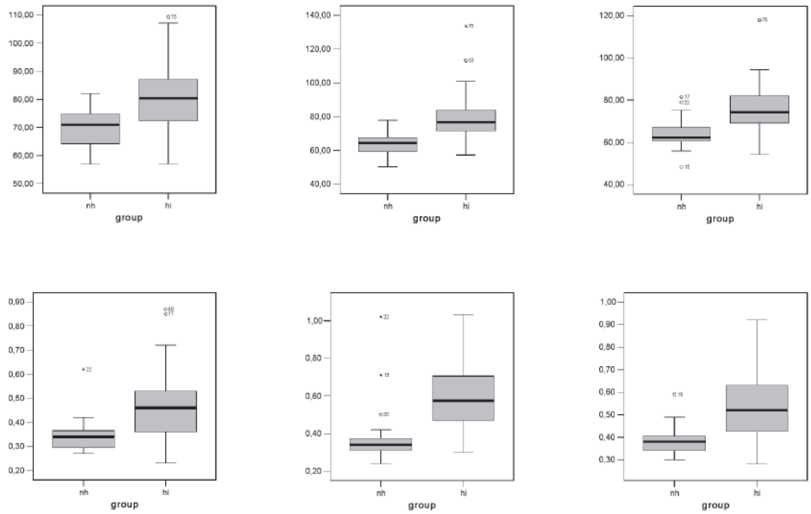

Fig. 1: ACALOS results for normal-hearing and hearing-impaired listeners separately. The upper row show most comfortable loudness levels (dB SPL) and the bottom row shows slopes of lower parts the fitted curves (cu/dB). Results from narrowband noises are shown in the left (500 Hz) and middle (3000 Hz) panels and results or broadband noises are displayed in the right column. Please note that axis scales differ across graphs.

Audibility and loudness perception

In the hearing-impaired listeners (HI) the median audiometric threshold at the better ear is at 30 dB HL 500 Hz (range 5 to 65 dB HL) and 45 dB HL at 3 kHz (range 20 to 85 dB HL).

Loudness scaling results (ACALOS) are shown in figure 1 for the better ear of each subject. It can be seen that in general, HI have higher MCLs and steeper slopes than normal-hearing listeners (NH). Moreover, there is more spread in the HI data than in the NH data.

Spectral resolution and temporal acuity

Figure 2 shows the results of the FT test for NH and HI listeners for the better ears. On average, NH listeners have much better spectral and temporal resolution (more negative release of masking values) than HI listeners. Differences are most pronounced in the resolutions at 3000 Hz. There is another important difference between the two frequencies: at 500 Hz, spread among HI is much larger than among NH, whereas at 3 kHz the spread among HI is very small, especially in temporal resolution. This suggests that perhaps the temporal resolution at 3 kHz was too hard for HI listeners, causing all HI subjects responding the same (poor resolution, no release of masking).

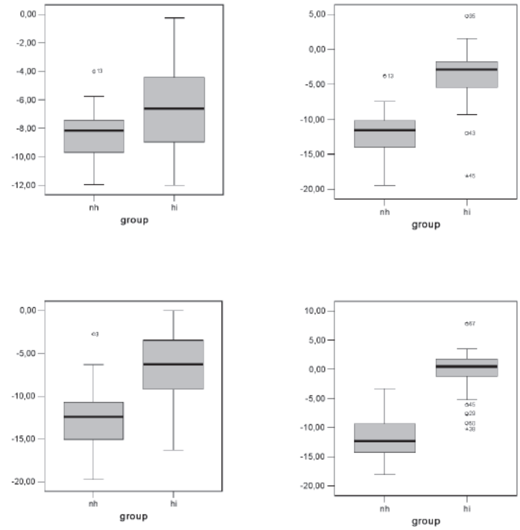

Fig. 2: FT-test results. Vertical axes represent release of masking (dB) in noise with spectral (top panels) or temporal (bottom panels) gaps (more negative values refer to better release of masking). Results for 500 Hz are shown in the left panels and for 3000 Hz in the right panels.

Speech perception

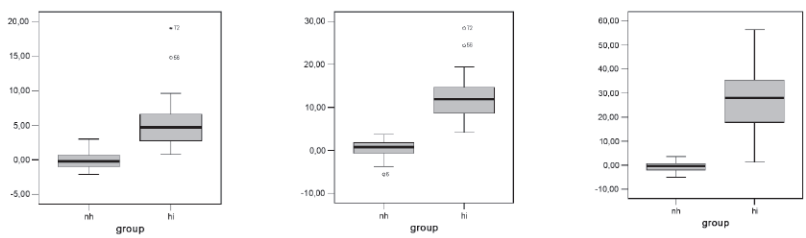

Fig. 3: SRT in stationary (left panel) and fluctuating noise (centre panel) and in quiet (right panel). SRTs in noise are displayed in corrected signal-to-noise ratios (reference values are subtracted from original data) and SRT in quiet represents the corrected speech level.

Figure 3 shows corrected SRT results in quiet (binaural) and in stationary and fluctuating noise (monaural, better ears). SRTs are corrected by subtracting average NH SRTs from a previous reference study, to compensate for language or test effects. As expected, there is very little spread in the NH data, and NH data are centred on 0 dB because of the correction. Among HI, there is much more spread, and overall, they perform much worse than NH (higher SRTs). Additionally, differences between NH and HI listeners are larger in fluctuating noise than in stationary noise.

Binaural processing

Figure 4 shows results of the MAA test for normal-hearing and hearing-impaired listeners, for three different stimulus types (low-pass noise, high-pass noise, and broadband noise). For each condition, outcome measure of this test is the mean MAA of two measurements on one session day. In general, NH perform better than HI (smaller MAAs) and spread among NH is smaller than among HI.

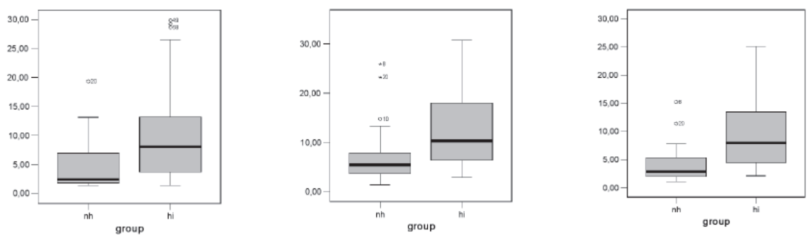

Fig. 4: MAA results for normal-hearing and hearing-impaired listeners. The three panels show MAAs of low-pass (left), high-pass (centre), and broadband (right) noise. Vertical axes show minimum audible angles in degrees. Please note that the scales of vertical axes are not the same in all three graphs.

It can be seen that for normal-hearing listeners average MAA for broadband noise is around 3o, which is a little more than the 1o that is generally found for real MAA measurements (measured with loudspeakers and extensively trained listeners). This difference will probably be caused by the fact that in the current experiment generic HRTFs are used, that are, of course, not as good as a subject’s own HRTFs. It has to be mentioned that the low-pass condition was considered rather difficult, and that some HI subjects were not able to perform the test with the low-pass noise. We do not have an explanation for this phenomenon yet.

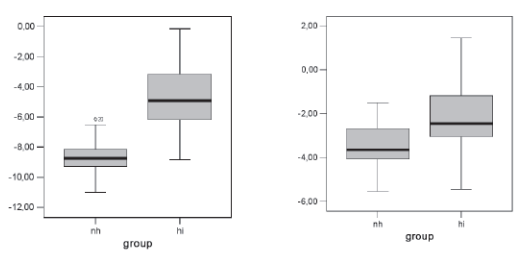

Fig. 5: ILD (left panel) and BILD (right panel) with noise at the side of the poorer ear. Vertical axes show release of masking for the more favourable condition (more negative values refer to better binaural hearing).

Results of the binaural processing tests ILD and BILD are shown in figure 5. In both tests, more negative values refer to better binaural processing.

Hearing-impaired listeners have less benefit from spatial separation (higher values at ILD) and binaural hearing (higher values at BILD) than normal-hearing listeners. The ILD effect in NH is rather large, there is a considerable difference between NH and HI performance, and relatively little spread in the NH data. For BILD, the effect size for NH is quite small, which makes the spread relatively large, and the difference between NH and HI smaller.

Self-report measures

Listening effort results for both normal-hearing and hearing-impaired listeners are presented in figure 6. It is remarkable that in this test, in contrast to most previously described results, there is large spread in the normal-hearing data, and also normal-hearing listeners need quite some effort to understand the speech in the more difficult situations (SNR=−5). In the most difficult condition (stationary noise at SNR=−5), there is substantial overlap between NH and HI results. However, this is possibly due to the fact that the listening effort scale is subjective. NH and HI groups may respond differently to this subjective scale because the latter group has become used to having difficulties and has lower expectations.

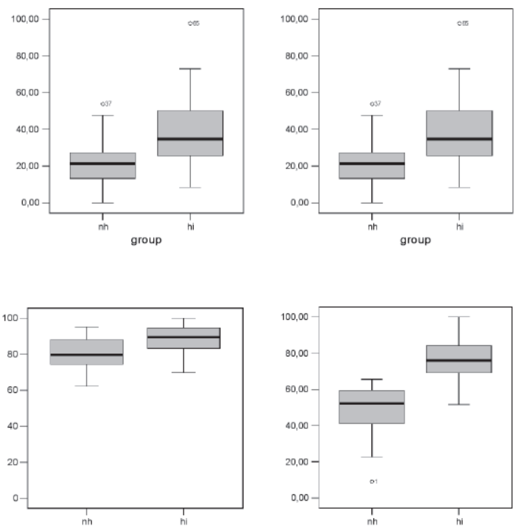

Fig. 6: Perceived listening effort (higher values indicate more effort) by NH and HI listeners when listening to speech in stationary noise (left panels) or fluctuating noise (right panels), at SNR=+5 (top panels) or SNR=−5 (bottom panels). All scales (of vertical axes) are equal.

Figure 7 shows results of normal-hearing and hearing-impaired listeners on the Gothenburg Profile subscales: speech perception, spatial hearing, social interactions, and behaviour and reaction. On all four subscales there is very little spread in the normal-hearing data, and much more spread (and higher scores, so more problems) in the hearing-impaired data. Scores are mean values of answers on five questions on each topic.

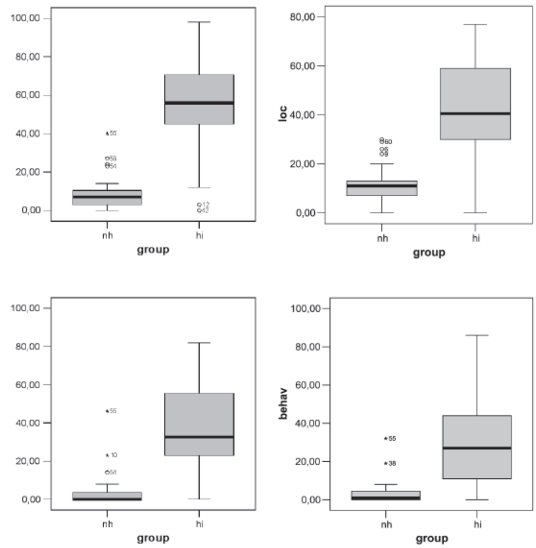

Fig. 7: Gothenburg Profile results. The panels present scores (more negative scores refer to better hearing/less problems) of the four subscales of the questionnaire: speech perception (upper left), spatial hearing (upper right), social interactions (lower left) and behaviour and reaction (lower right).

Cognitive abilities

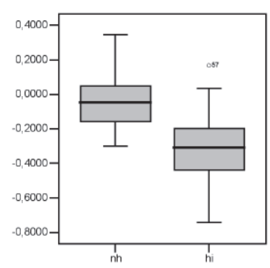

Fig. 8: Lexical decision making results. Vertical axis represents (%correct)/(response time), so higher values refer to better performance. Results are corrected by reference values from a pilot study.

Figure 8 shows results of the cognitive test (lexical decision-making test). On the vertical axis, percentage correct divided by response time is shown, corrected by reference values from a reference study. Although NH subjects are expected to be clustered around 0 because of the correction, the median is slightly below 0. This indicates that listeners from the present study perform a little worse than subjects from the reference study. A possible explanation for this effect is that in the reference study the lexical-decision test was the only test to be done, while in the current study it was just part of an extensive battery of tests, so subjects might have been more tired or concentrated less in the present study. Although the task is not auditory, HI perform worse than NH on this test. This effect may have been biased by the age difference between the two groups.

DISCUSSION AND CONCLUSIONS

International consensus is growing for a broad battery of audiological tests to characterise the residual capacities of the impaired ear. The results of the multi-centre 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 the 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 analyse this multi-dimensionality. 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.

ACKNOWLEDGMENTS

This work was supported by grants from the EU FP6, Project 004171 HEARCOM. The information in this document is provided as is and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability.

REFERENCES

Brand, T., and Hohmann, V. (**2002**). “An adaptive procedure for categorical loudness scaling,” J. Acoust Soc. Am. **112**, 1597-1604.

Hagerman, B. (**1982**). ”Sentences for testing speech intelligibility in noise,” Scand. Audiol. **11**, 79-87.

Larsby, B., and Arlinger, S. (**1998**). “A method for evaluating temporal, spectral and combined temporal-spectral resolution of hearing,” Scand. Audiol. **27**, 3-12.

Larsby, B., Hällgren, M., Lyxell, B., and 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.

Plomp, R., and Mimpen, A. M. (**1979a**). “Improving the reliability of testing the speech reception threshold for sentences,” Audiology **18**, 43-52.

Ringdahl, A., Eriksson-Mangold, M., and 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. Audiol. **32**, 375-385.

Wagener, K., Brand, T., and Kollmeier, B. (**1999a**). ”Development and evaluation of a German sentence test part I: Design of the Oldenburg sentence test,” Z. Audiol. **38**, 4-15.

Wagener, K., Brand, T., and Kollmeier, B. (**1999b**). “Development and evaluation of a German sentence test part II: Optimization of the Oldenburg sentence test,” Z. Audiol. **38**:44-56.

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