Clinical measurement of various aspects of hearing impairment and their relation to auditory functioning: the development of an Auditory Profile

van Esch, Thamar

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Relations between speech perception in noise and psychophysical measures of hearing measured in four languages using the preliminary Auditory Profile test battery

Thamar E.M. van Esch¹, Mark E. Lutman², Matthias Vormann², Johannes Lyzenga³, Mathias Hällgren⁴, Birgitta Larsby⁴, Sheetal P. Athalye⁵, Tammo Houtgast³ Birger Kollmeier², Wouter A. Dreschler¹

¹ Academic Medical Centre, Amsterdam, The Netherlands (NL-AMC)  
² Hörzentrum Oldenburg GmbH, Oldenburg, Germany (DE-HZO)  
³ Free University Hospital, Amsterdam, The Netherlands (NL-VUMC)  
⁴ Linköping University, Linköping, Sweden, (SE-LINK)  
⁵ Institute for Sound and Vibration Research, Southampton, United Kingdom (UK_ISVR)

under revision
Abstract

Objective: The aim of the present study was to determine the relations between speech perception in noise and measures of auditory resolution, recruitment and cognitive function. All were measured using the preliminary Auditory Profile, a test battery implemented in four languages.

Design: Tests of speech perception, resolution, recruitment, and cognitive function were measured using headphones in five centres: in Germany, the Netherlands, Sweden, and the United Kingdom. Correlations and stepwise linear regression models were calculated in the hearing-impaired listeners.

Study sample: Thirty normally hearing listeners aged 19–39 years and 72 hearing-impaired listeners aged 22–91 years with a broad range of hearing losses.

Results: Several significant correlations were found with speech perception in noise. Stepwise linear regression analyses showed that spectral and temporal resolution and recruitment were significant predictors for speech perception in noise.

Conclusions: Complex inter-relationships between auditory factors and speech perception in noise are revealed using the preliminary Auditory Profile data set in four languages. Spectral and temporal resolution and recruitment predict variation among listeners in speech perception in noise, in addition to the variation explained by audiogram and age. The measure of cognitive function that was used, is not related to speech perception in noise, in the population studied.
4.1 Introduction

An international multi-centre effort was recently undertaken to develop an Auditory Profile test battery (Van Esch et al, 2013). This test battery characterizes hearing in terms of seven domains: loudness perception, spectral and temporal resolution, speech perception in quiet and in noise, spatial hearing, cognitive abilities, listening effort, and self-reported disability and handicap. The preliminary test battery was evaluated in an international multi-centre study with over 100 listeners from four countries.

Van Esch et al (2013) have described the composition and evaluation of this test battery. They presented reference data, investigated the clinical applicability and usability of each individual test, and compared their results to results from previous studies. They concluded that the individual tests are potentially applicable clinically in four different languages (Dutch, English, German, and Swedish). For normally hearing listeners (NH), all tests were comparable across centres after baseline correction to account for differences between test materials. For hearing-impaired listeners (HI) however, differences between test materials have to be taken into account when interpreting the results of the language-dependent tests (i.e. the tests that have different materials in each language). Furthermore, they found that the results from the preliminary Auditory Profile show generally good test-retest reliability. As a first validation step, they compared the results from the preliminary Auditory Profile with previously published data. For most tests, good agreement was found (see Van Esch et al, 2013).

In the present study, we investigate the relationship between speech perception in noise and several other tests from the preliminary Auditory Profile. This serves as second step of the validation of the Auditory Profile test battery and provides insight into the causes of reduced speech understanding in noise. It also allows the potential usefulness (relevance) of the test battery in a clinical setting to be evaluated.

Reduced speech understanding in noisy situations, especially in fluctuating noise, is a very common complaint among hearing-impaired listeners. In clinical audiology, speech perception in noise is often measured as the speech reception threshold (SRT, Plomp, 1986), which is the speech level required to achieve a 50% correct score. It is widely recognised that HI listeners have poorer speech perception in noise than NH listeners, and that
these differences increase in fluctuating noise, as HI listeners have reduced ability to take advantage of the gaps in noise (e.g. De Laat & Plomp, 1983, Festen & Plomp, 1990, Versfeld & Dreschler, 2002). It has been shown that the reduced speech perception in (fluctuating) noise cannot be explained entirely by audiometric differences and hence there have to be other causes (e.g. Eisenberg et al, 1995; Bacon et al, 1998; Summers & Molis, 2004). Understanding these causes is of great importance as it may help understanding the nature of the underlying hearing loss. Moreover, understanding the reasons for reduced speech perception is of clinical importance too, because it may possibly help selecting appropriate rehabilitation strategies.

The main consequences of cochlear hearing loss, apart from threshold elevation, are reduced spectral and temporal resolution (Festen & Plomp, 1983; Dreschler & Plomp, 1985) and loudness recruitment (Villchur, 1974; Brand & Hohmann, 2001, 2002). These aspects of impairment have been shown to relate to speech perception in noise. Relations between spectral and temporal resolution and speech perception in noise have been investigated by many authors (e.g. Patterson et al, 1982; Glasberg & Moore, 1989; Dreschler & Leeuw, 1990; Noordhoek et al, 2001; George et al, 2006, 2007; Van Esch & Dreschler, 2011). Villchur (1974) and Moore & Glasberg (1993) demonstrated the influence of recruitment on speech perception in noise. Besides that, significant relations between SRT scores and outcomes of several cognitive abilities have been shown (e.g. Humes, 2007), Kramer et al (2009), and Larsby et al (2012); see Akeroyd (2008) for an overview). Several studies also investigated the influence of both auditory and cognitive factors on speech perception in noise. Houtgast & Festen (2008) give an overview of studies that related a subset of measures of hearing thresholds, spectral resolution, temporal resolution, intensity difference limen, age, and cognitive abilities to the speech reception threshold in noise. Their overall conclusion is that typically 70% of the variance in speech perception data can be explained by these factors. However, one can argue that the result range of speech reception thresholds in noise, and thus the range of hearing impairment included in a study, strongly affects the percentage of variance that theoretically can be explained by a certain model. Moreover, Houtgast & Festen (2008) found that the audiogram and age are generally good predictors across the different studies.
Many of the above-mentioned studies investigate the influence of only one factor, include only a small number of listeners and/or use time-consuming experimental methods. Although the fact that recruitment, spectral and temporal resolution, and cognitive abilities, all have shown to play a significant role in the prediction of speech perception in noise, it is not clear to what extent their information is mutually exclusive and of added value in addition to the pure-tone audiogram and age. It can be difficult to unravel causal relationships from the correlation shown in such observational studies, due to co-variation between the explanatory variables.

The preliminary Auditory Profile data set offers further possibility to examine the relationships between speech perception in noise and recruitment, spectral and temporal resolution, and cognitive abilities. It has the advantage that, clinically applicable, standardised tests were used and the sample of listeners involved was relatively large. In the preliminary Auditory Profile test battery, recruitment is determined as the slope of the lower part of the loudness curve as measured with the Acalos test that was developed by Brand & Hohmann (2001). Van Esch et al (2013) showed that the Acalos results from the preliminary Auditory Profile test battery corresponded very well (i.e., within 5 dB) with data presented in the ISO standard (ISO 16832, 2006) and with data from Brand & Hohmann (2001). Furthermore, spectral and temporal resolution are measured in a combined test (the F&T test). Van Esch et al (2013) found good agreement for HI listeners between the spectral and temporal resolution results from the preliminary Auditory Profile and previously published results. Moreover, Van Esch & Dreschler (2011) showed significant correlations between results from the F&T test and results from conventional spectral and temporal resolution tests. Cognitive abilities are measured in the preliminary Auditory Profile using a lexical decision-making test. Van Esch et al (2013) reported that their Swedish NH results agree very well with the results presented by Hälgren et al (2001), but that the Swedish HI results from Van Esch et al (2013) are better than those presented by Hälgren et al (2001). Finally, speech perception performance in noise is measured using short meaningful sentences in the language corresponding to each centre. Van Esch et al (2013) showed that the NH results in stationary noise from the preliminary
Auditory Profile were quantitatively in line with previously studies results, and that also the effects of fluctuating noise and hearing loss agreed with the literature.

If known relationships between SRT and the other tests hold in the preliminary Auditory Profile data set, this will largely increase the possible applicability of the Auditory Profile in future research projects. However, a potential drawback of the preliminary Auditory Profile data set is the fact that some test centre effects were found in the HI data from the five centres in four different countries. Van Esch et al (2013) found that for the lexical decision test and for speech perception in noise, HI results from the different centres differed significantly from each other, even after subtraction of a language correction factor.

The aim of the present study was to examine the relationships between the psychophysical measures in the preliminary Auditory Profile (recruitment, spectral and temporal resolution, and cognitive abilities for speech perception in noise) and speech perception in noise. To that end, relationships between test results are investigated in two steps.

- First, as a further step in the validation of the preliminary Auditory Profile, we examined whether the well-known correlations between speech perception, recruitment, spectral and temporal resolution, and cognitive abilities can be reproduced in the results from the preliminary Auditory Profile in four languages.
- Secondly, we investigated the additional value (relevance) of the preliminary Auditory Profile for speech perception in noise, by identifying predictors of speech perception in noise, after accounting for the pure-tone audiogram and age. This will be done by means of stepwise linear regression. It should be recognised that this approach tends to over-estimate the importance of the audiogram, since some of the variation in speech perception scores that is intrinsically caused by other psychophysical measures may be attributed to the audiogram, due to co-variation of the psychophysical measures with hearing threshold level.
4.2 Methods

Materials and methods of the experiments are described in detail by Van Esch et al (2013). Here, a shortened version of the general methods as well as brief descriptions of the tests to be examined in this paper is presented.

4.2.1 Test set-up

The tests were implemented on the Oldenburg Measurement Application (OMA), which is a combined software and hardware test platform. Tests ran on a PC and sounds were played via an RME soundcard (type Fireface 800, DIGI96/8 PAD or HDSP 9632) and fed through an amplifier to Sennheiser HDA 200 headphones. Experiments took place in sound-insulated booths. Written instructions were translated in the four languages (Dutch, English, German and Swedish) and used in all centres, complemented with oral explanations when needed.

4.2.2 Centres and listeners

The five participating centres were audiological centres from Academic Medical Centre, Amsterdam, The Netherlands (NL-AMC), Hörzentrum Oldenburg GmbH, Oldenburg, Germany (DE-HZO), Free University Hospital, Amsterdam, The Netherlands (NL-VUMC), Linköping University, Linköping, Sweden, (SE-LINK), and the Institute for Sound and Vibration Research, Southampton, United Kingdom (UK-ISVR) (Van Esch et al, 2013). Table 4.1 shows the distribution of NH and HI listeners over the different centres.

All centres were approved by their local research ethics committees for

<table>
<thead>
<tr>
<th>#</th>
<th>NL-AMC</th>
<th>DE-HZO</th>
<th>UK-ISVR</th>
<th>SE-LINK</th>
<th>NL-VUMC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH</td>
<td>5</td>
<td>5</td>
<td>10</td>
<td>5</td>
<td>5</td>
<td>30</td>
</tr>
<tr>
<td>HI</td>
<td>15</td>
<td>15</td>
<td>12</td>
<td>15</td>
<td>15</td>
<td>72</td>
</tr>
<tr>
<td>Total</td>
<td>20</td>
<td>20</td>
<td>22</td>
<td>20</td>
<td>20</td>
<td>102</td>
</tr>
</tbody>
</table>
the conduct of the study\(^1\), in accordance with the Declaration of Helsinki, and all listeners gave written informed consent to participate in the study. NH listeners were aged 19 to 39 years (mean: 26) and HI listeners were aged 22 to 91 years (mean: 63).

Pure-tone audiometry was conducted prior to the test session using a clinical audiometer calibrated according to ISO 389-1 (1998). Mean air-conduction audiograms of left and right ears are shown in Figure 4.1 for NH listeners (n=30), HI listeners with sensorineural losses, defined as having an air-bone gap of less than 10 dB averaged over thresholds at 0.5 and 1 kHz (n=58), and HI listeners with conductive components (n=14) separately.

The majority of listeners had symmetric hearing losses, but there were

13 listeners in the HI group with an asymmetry of 10 dB or more (averaged over 0.5, 1, 2, and 4 kHz). Details about the listeners and conductive and asymmetric hearing losses are described by Van Esch et al (2013). Subsequent analyses are performed on the total group of HI listeners.

4.2 Methods

The tests that are used in the present study are listed in Table 4.2.

The tests were conducted in test and retest in two sessions on separate days (1–3 weeks apart). The listed tests did show minimal learning effects (Van Esch et al, 2013). Therefore, means of test and retest values are used as the pooled measures in the present analyses. The exception is the pure-tone audiogram which was only measured once.

The auditory tests were conducted unaided via headphones, monaurally on each ear separately.

All auditory tests were conducted at similar subjective loudness levels to obtain similar audibility for all listeners. Individual loudness levels were obtained from the Acalos test and were limited to maxima of 95 and 85 dB SPL for narrowband and broadband signals, respectively (see below for details).

Language-validation studies with NH listeners were conducted in separate independent experiments for the language-dependent tests (see Van Esch et al, 2013). The results were used to correct for test-material effects by presenting all outcome measures relative to reference values, based on the

---

Table 4.2: Testing fields and selected tests of the preliminary Auditory Profile that are used in the present analyses.

<table>
<thead>
<tr>
<th>Testing field</th>
<th>Selected test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loudness perception</td>
<td>Acalos test</td>
</tr>
<tr>
<td>Speech perception in noise</td>
<td>SRTs</td>
</tr>
<tr>
<td>Spectral and temporal resolution</td>
<td>FT test</td>
</tr>
<tr>
<td>Cognitive abilities</td>
<td>Lexical decision-making test</td>
</tr>
</tbody>
</table>

---

The preliminary Auditory Profile also included a measurement of speech perception in quiet, a questionnaire on disability and handicap, spatial hearing tests, and a listening-effort test (see Van Esch et al, 2013). These tests are not included here, because their results will not be used in the present analysis.
average scores of NH listeners for each language. The language-dependent tests comprised the SRT tests and the lexical decision-making test. In the present paper, only corrected data (i.e. data after subtraction of reference values) are presented.

4.2.4 Test procedures

Acalos test
Loudness scaling was measured using an adaptive, categorical procedure: the Acalos test (Adaptive, CAtegorical LOudness Scaling) as described by Brand and Hohmann (2002). Listeners judged loudness on a scale, based on which the stimulus level was adaptively varied. For three stimulus types (broadband noise and narrow-band noises at 0.5 and 3 kHz), individual loudness growth curves were fit. From these curves most-comfortable levels (MCL) were calculated as the level corresponding to the perceived loudness of 20 categorical units (cu), on a scale from 0 to 50. These MCLs were used as presentation levels in subsequent tests. Slopes of the lower-level parts of the curves were used as measures of the degree of recruitment.

F&T test
Spectral and temporal resolution were measured at 0.5 and 3 kHz using a combined test, as described by Larsby & Arlinger (1999), with the modifications that were suggested by Van Esch & Dreschler (2011). The test measures spectral and temporal resolution simultaneously by assessing the release of masking (RoM) of pulsed test tones in different masking noises. Masked thresholds were recorded using a Békésy tracking procedure with noise level fixed at MCL. To assess spectral resolution, a ½-octave spectral gap, centred at the test-tone frequency, was cut in the masking noise and 50 ms temporal gaps (the centre of which coincided with the centre of the test-tone pulse) were introduced in the octave-band noise to allow the assessment of temporal resolution.

SRT test
Speech perception performance in noise was measured using short meaningful sentences in the language corresponding to each centre: in the Dutch centres (NL- AMC and NL-VUMC) the Versfeld sentences (Versfeld et al, 2000) were used, DE-HZO used the Göttinger sentences
Methods

(Kollmeier & Wesselkamp, 1997; Brand & Kollmeier 2002), UK-ISVR the BKB sentences (Bench et al, 1979) and SE-LINK the Swedish HINT sentences (Hällgren et al, 2006). Tests were conducted according to the local standards, as described in the papers cited above. Two conditions were tested: speech in stationary noise and speech in fluctuating noise. Depending on the speaker of the local test, the male or female version of the ICRA1 noise, a universal speech-shaped noise (see Dreschler et al, 2001) was used for a stationary noise. For a fluctuating noise the ICRA5_250 or ICRA4_250 was used, which are male- and female-weighted versions of fluctuating speech-shaped noises (see Wagener et al, 2006a). In both conditions, the speech level was adaptively varied to assess the speech level at which 50% of the sentences were repeated correctly: the speech reception threshold (SRT). The noise level was fixed at the individual MCL for broadband noise, with a maximum of 85 dB SPL, and the outcome measure was the SRT expressed as signal to noise ratio (SNR) in dB.

Lexical decision-making test
A measure of cognitive abilities was obtained using the lexical decision-making test (Hällgren et al, 2001), which measures speed and accuracy of lexical access of subjects. This test was originally developed in Sweden (Hällgren et al, 2001) and was translated to Dutch, English, and German in the HearCom project (www.hearcom.eu). The subject’s task was to discriminate words from non-words that were presented as text on a computer screen, by pressing predefined response buttons. Both accuracy and speed of performance were assessed. The outcome measure of this test was the value of percentage correct divided by response time (in ms), multiplied by −1 (making lower values corresponding to better performance).

4.2.5 Statistical analysis

Normality
Normality of the outcome measures was tested in Van Esch et al (2013) by visual inspection and the Shapiro-Wilk and Kolmogorov-Smirnov tests. In the HI group, all outcome measures that are used in the present analyses were distributed (approximately) normally, meaning that for all tests the majority of outcomes were distributed normally according the abovementioned tests.
Linear regression analyses and inclusion of audiogram thresholds
Multiple linear regression analyses were performed in SPSS, on the group of HI listeners. SRTs in stationary and fluctuating noise were used as dependent variables. The regression models involved stepwise inclusion of possible predictors (inclusion: $p<0.05$ and exclusion: $p>0.10$). As we wanted to investigate the relevance of the preliminary Auditory Profile tests in addition to the pure-tone audiogram and age, we first included the audiogram measures and age in one block, before including all other measures in a second block. For every analysis, the distribution of residuals was checked for approximate normality, a plot of residuals versus predicted values was checked for linearity and homoscedasticity, and autocorrelation of the residuals was tested using the Durbin-Watson statistic.

4.3 Results

Relations between the SRT data, both in stationary and in fluctuating noise (‘SRT\textsubscript{stat}’ and ‘SRT\textsubscript{fluct}’) and the audiogram, age, recruitment, spectral and temporal resolution, and cognitive abilities were analysed. Audiogram measures that were used are the pure-tone average (0.5, 1, 2, 4 kHz, PTA), the slope of the audiogram (difference between thresholds at 0.5 and 4 kHz), and the air-bone gap (average of 0.5 and 1 kHz, ABG). Measures of recruitment were the slopes from the lower-level parts of the Acalos curves measured at 0.5 and 3 kHz (SL\textsubscript{500} and SL\textsubscript{3k}). Spectral resolution at 0.5 and 3 kHz (F\textsubscript{500} and F\textsubscript{3k}) and temporal resolution at 0.5 and 3 kHz (T\textsubscript{500} and T\textsubscript{3k}) from the F&T test were used and expressed in inverse release of masking (i.e. more negative values indicate better performance). Lexical decision-making results were included as a measure of cognitive abilities.

4.3.1 Correlations

Pearson correlation coefficients were calculated in the HI group, between SRTs, age, audiogram, recruitment, spectral and temporal resolution, and lexical decision-making results.

Note that the scales of all measures were arranged such that poorer per-
Table 4.3: Pearson correlation coefficients in the HI group between SRT, age, audiogram, recruitment, spectral and temporal resolution, and lexical decision-making (LDM) results. Significant correlations at the $p<0.01$ and $p<0.05$ level are marked ** and * respectively.

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>Audiogram</th>
<th>Acalos</th>
<th>FT test</th>
<th>LDM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PTA</td>
<td>slope</td>
<td>ABG</td>
<td>SL500</td>
<td>SL3k</td>
</tr>
<tr>
<td><strong>SRT_{stat}</strong></td>
<td>0.29*</td>
<td>0.46**</td>
<td>0.28*</td>
<td>0.06</td>
<td>0.15</td>
</tr>
<tr>
<td><strong>SSRT_{fluct}</strong></td>
<td>0.34**</td>
<td>0.67**</td>
<td>0.14</td>
<td>0.03</td>
<td>0.36**</td>
</tr>
</tbody>
</table>

Performance corresponded to more positive values; hence positive correlation coefficients were expected. Correlations between right-ear data are shown in Table 4.3, left-ear data showed very similar results and are not shown here.

It can be seen that age and the PTA correlated significantly with SRTs in stationary and fluctuating noise. The slope of the audiogram correlated with SRT_{stat}, but not with SRT_{fluct}. Also most results from the Acalos test and the F&T test correlated significantly with the SRT results, and correlations were of the same order of magnitude as those between the audiogram and SRT, or smaller. No significant correlations between the cognitive test (lexical decision-making test) and SRT results were found.

4.3.2 Stepwise linear regression analyses

Stepwise linear regression analyses were performed on the SRT data, both in stationary and in fluctuating noise. Analyses were conducted in two blocks. In the first block, age and audiogram measures were included in a stepwise procedure to find the significant predictors among those variables (PTA, slope, ABG, and age). The second block involved stepwise inclusion of all possible predictors from the preliminary Auditory Profile: spec-

---

3 The following measures are included: SRT in stationary and fluctuating noise (SRT_{stat} and SRT_{fluct}), audiogram: PTA, slope, and ABG; recruitment at 0.5 and 3 kHz (SL500 and SL3k); spectral and temporal resolution at 0.5 and 3 kHz (F500, T500, F3k, T3k); and lexical decision-making results. For all auditory measures results of analyses of right-ear data are displayed.
### Table 4.4: Results of stepwise linear regression analyses of right-ear data.

For two dependents (SRT\textsubscript{stat} and SRT\textsubscript{fluct}) the significant predictors from both blocks of the stepwise linear regression models are shown, as well as R, and adjusted R\textsuperscript{2} values of these models. All shown models have a significance p<0.001.

<table>
<thead>
<tr>
<th>Dependent</th>
<th>Model</th>
<th>Predictors</th>
<th>R</th>
<th>Adjusted R\textsuperscript{2}</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRT\textsubscript{stat}</td>
<td>block 1</td>
<td>PTA, slope</td>
<td>0.56</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td>blocks 1 &amp; 2</td>
<td>PTA, slope, F500</td>
<td>0.60</td>
<td>0.33</td>
</tr>
<tr>
<td>SRT\textsubscript{fluct}</td>
<td>block 1</td>
<td>PTA, age</td>
<td>0.70</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>blocks 1 &amp; 2</td>
<td>PTA, age, T3k, F500, SL3k</td>
<td>0.81</td>
<td>0.63</td>
</tr>
</tbody>
</table>

Central and temporal resolution (F500, F3k, T500, T3k), recruitment (SL500, SL3k), and lexical decision-making.

**Analyses on right-ear data**

Results of these analyses for right-ear data are displayed in Table 4.4. (The regression models are verified later on the left-ear data – see below.) For predictions of SRTs in stationary and in fluctuating noise, results for both the models with only age and audiogram variables (block 1) as well as the models with audiogram variables and other predictors (blocks 1 and 2) are shown. For each model the included variables, values of R and adjusted R\textsuperscript{2} are presented. Adjusted R\textsuperscript{2} values give the percentage of explained variance by the model, corrected for the available degrees of freedom. Variables are printed in the order that they were included by the stepwise procedures, so in order of decreasing R\textsuperscript{2} values.

For stationary noise, both PTA and audiogram slope were significant predictors from the first block and explained 29% of the variance. Age and the ABG proved not to be significant. When including the second block of variables, temporal resolution at 0.5 kHz improved the prediction of SRT in stationary noise significantly (R\textsuperscript{2} change: p<0.05), while the other possible predictors were not found to be significant factors for speech perception in stationary noise. The regression model for SRT in stationary noise for HI listeners was:

\[
\text{SRT}_{\text{stat}} = 2.60 + 0.77 \text{PTA/10} + 0.52 \text{slope/10} + 0.28 \text{F500}
\]

This model explained 33% of the variance in right-ear results of SRT in
stationary noise. Please note that SRT\textsubscript{stat} refers to corrected SRT scores (see Methods section and Van Esch et al, 2013), so for NH listeners SRT\textsubscript{stat} was zero, on average, and higher values corresponded to poorer speech perception. The regression formula indicates that SRT\textsubscript{stat} increased 0.77 dB for every 10 dB of hearing loss (PTA) and 0.52 dB for every 10 dB/oct increment in audiogram slope. Furthermore, listeners with poorer temporal resolution at 0.5 kHz had poorer SRT\textsubscript{stat} (0.28 dB increment per dB release of masking (RoM) for spectral resolution).

For fluctuating noise, PTA and age were significant in the first block; together they explained 47% of the variance. Slope from the audiogram and ABG did not prove to contribute significantly. Temporal resolution at 3 kHz, spectral resolution at 0.5 kHz, and recruitment at 3 kHz improved the prediction significantly ($R^2$ change: $p<0.001$), while lexical decision-making did not contribute significantly to the regression model. The regression equation for (corrected scores for) SRT in fluctuating noise for HI listeners was:

$$SRT_{\text{fluct}} = 2.43 + 0.13 \frac{\text{PTA}}{10} + 0.75 \frac{\text{age}}{10} + 0.31 \text{T3k} + 0.44 \text{F500} + 4.32 \text{SL3k}$$

and explained 63% of the variance in SRT in fluctuating noise. This equation shows that the present HI group, SRT\textsubscript{fluct} was on average 2.43 dB higher than the NH reference, with additional increments proportional to their PTA, age, loss of spectral resolution at 0.5 kHz and loss of temporal resolution and recruitment at 3 kHz.

Test-centre effects
Van Esch et al reported significant test centre effects in the SRT (and lexical-decision making) results (see Van Esch et al, 2013). In order to investigate whether these differences are also associated with a significant effect of centre on the regression analyses, the linear regression models were evaluated separately for the different centres. Scatter plots of predicted versus measured SRT scores by centre (not shown) were visually inspected and showed overlapping and very similar results for the different centres. Indeed, no significant effect of centre was found in the residuals from the linear regression analyses (oneway ANOVA, $F(4,62)=2.40$, $p=0.06$ for SRT\textsubscript{stat}, and
F(4,61)=1.97, p=0.111 for SRT\textsubscript{fluct}). Therefore, we concluded that similar regression models apply for predicting the SRTs in noise at the different centres.

**Generizability and verification on left-ear data**

To test the validity of the models beyond the right-ear data on which they were based, the models as described by the above equations were applied on left-ear data. In these analyses, somewhat higher R values than those for the right-ear data were found: R=0.85 for stationary noise (0.60 for right-ear data) and R= 0.91 (0.81 for right-ear data) for fluctuating noise (for the models with both blocks included). This shows that the predictive power of the regression models was a little higher for the left-ear data set than for the right-ear data set from which they were derived.

### 4.4 Discussion

In the present paper, we examined the relevance of assessing recruitment, spectral and temporal resolution, and cognitive abilities for speech perception in noise, all measured with the preliminary Auditory Profile test battery in four languages. We first tested whether known correlations between speech perception in noise and recruitment, spectral and temporal resolution, and cognitive abilities can be reproduced in the data set from the preliminary Auditory Profile in four different languages, as a further validation of the preliminary Auditory Profile. Next, we performed stepwise linear regression analyses to examine the additional value of the preliminary Auditory Profile for speech perception and to find the most important variables for predicting the SRTs, in addition to the pure-tone hearing thresholds and age. Below, we address both issues sequentially for recruitment, spectral and temporal resolution, and cognitive abilities. Next, we discuss the centre effect in our data set and we will compare our results to previous research on predictive factors for speech perception in noise. Finally, the additional value of the preliminary Auditory Profile for speech perception in noise, in addition the audiogram and age, is discussed.
4.4.1 Discussion | 117

Recruitment

In the present study, significant correlations between measures of recruitment and SRTs (p<0.01, see Table 4.3) were found. This is in agreement with a previous study that showed the influence of recruitment on speech perception in noise (Dreschler & Plomp, 1985). The results are also in line with two studies in which the effect of loudness recruitment on speech perception in noise was simulated (Villchur, 1974 and Moore & Glasberg, 1993).

According to the results of the stepwise linear regression analysis, recruitment was not a significant predictor for SRT in stationary noise. For the SRT in fluctuating noise, recruitment at 3 kHz was a significant predictor. In other words, more recruitment than expected from the audiogram was associated with poorer speech reception than expected from the audiogram. However, recruitment was the last factor that was included in the model for SRT in fluctuating noise and the effect was small, leading to an increment in $R^2$ of only 3%.

Presumably, the small additional effect of recruitment in the regression analyses is caused by covariance between recruitment and the audiogram ($r=0.66$ and $r=0.51$ at 0.5 and 3 kHz respectively, both $p<0.01$). As the correlations between the audiogram and SRT were also of the same order of magnitude as those between recruitment and SRT, the stepwise linear regression analysis model had to select either the hearing thresholds or recruitment, while in reality perhaps both factors co-produced the variation. However, in the present study we search for additional predictive power of the preliminary Auditory Profile tests for speech perception in noise, after accounting for the pure-tone audiogram and age. As recruitment hardly improves the prediction of SRT, we conclude that recruitment, although significantly correlated with speech perception in noise, is of only minor importance for the prediction of speech in fluctuating noise, once the audiogram has been taken into account. It is not possible from our data to know whether the correlation between the audiogram and SRT is actually a consequence of threshold elevation, or actually a consequence of recruitment, leading to apparent correlation with the audiogram due to the mutual correlation between the audiogram and recruitment. However, presentation of the speech materials at similar loudness levels for all listeners
must have minimised the influence of audibility, suggesting that threshold elevation itself may not be the key factor.

4.4.2 Spectral and temporal resolution

In the present study, we found several significant correlations between spectral and temporal resolution and speech perception in noise (see Table 4.3). This indicates that the previously reported relationship between spectral and temporal resolution and speech perception in noise (e.g. Dreschler & Plomp, 1985; Noordhoek et al, 2001; and George et al, 2007) also exists in the results from the preliminary Auditory Profile in four languages.

In the stepwise linear regression analyses, both spectral and temporal resolution were found to be significant additional predictors of SRT in addition to the audiogram and age. In stationary noise, spectral resolution at 0.5 kHz was included, with poorer resolution related to poorer speech perception. In addition to the audiogram (PTA and slope) spectral resolution explained 4% of the variance in SRT. In fluctuating noise, temporal resolution at 3 kHz and spectral resolution at 0.5 kHz proved to be significant predictors. Again, poorer resolution, both spectral and temporal, corresponded to poorer speech perception in noise. Together, these two predictors explained 13% of the variance. Altogether we conclude that spectral and temporal resolution are both important additional factors for the prediction of speech perception in noise, after the audiogram has been taken into account.

The same considerations apply to measures of spectral and temporal resolution as already described for recruitment. As the speech materials were presented at similar loudness levels for all listeners, the influence of audibility was minimised. This suggests that threshold elevation itself may not be the key factor and part of the variation in speech perception attributed to the audiogram in the regression models is actually mediated via reduced spectral and temporal resolution.
4.4.3 Lexical decision-making results

In the present study, no significant correlations were found between the lexical decision-making results and SRT. Likewise, in the linear regression analyses, the lexical decision-making test did not contribute significantly to the predictions of speech perception in noise.

Larsby et al (2012) examined the relation between results from the lexical decision-making test and speech perception in noise in forty Swedish HI listeners. They found significant correlations, even after correction for the PTA(0.5, 1, 2, 4 kHz). Their finding is in line with Akeroyd (2008), who concluded that there is a link between several different cognitive abilities and speech perception in noise. However, the relation between cognitive abilities and speech perception in noise that Larsby et al (2012) found was not present in our group of listeners from different countries. Possible causes are the fact that this test was newly developed in German, English, and Dutch, or the significant centre effects in the corrected results of both the lexical decision-making test and the speech perception in noise results (see Van Esch et al, 2013). However, also when the present results were analysed in sub-groups per language, no significant correlations between the lexical decision-making test and speech perception in noise were found. A potential explanation is the smaller number of listeners per language in the present study (12–15 HI listeners per centre, versus 40 in the study by Larsby et al). Another explanation may be a limited range of cognitive ability in our study sample.

4.4.4 Test centre effect

Van Esch et al reported significant centre effects in the SRT (and lexical-decision making) results (see Van Esch et al, 2013). However, this does not necessarily imply that there is also a significant centre effect in the prediction of speech perception in noise. In the present study, no significant centre effect was found in the residuals from the linear regression analyses. Nevertheless, centre effects in the data may still have contaminated or weakened the relationships between tests.
Predictions of speech perception in noise in the literature

Predictive factors (both auditory and cognitive) for speech perception in noise have been examined in several previous studies: see Houtgast & Festen (2008) for an overview. The present study differs from the studies in Houtgast & Festen (2008) in several aspects. Most importantly, the measurements in the present study are conducted in different languages and different centres, in contrast to the previous studies. Moreover, the tests from the preliminary Auditory Profile that were used in the present study, are, in general, faster than the more extensive research procedures that were used in most of the previous studies. A third essential difference concerns the use of equal subjective loudness levels in the present study. This new approach has both advantages and disadvantages as is discussed by Van Esch & Dreschler (2011) and Van Esch et al (2013). Here, it is important to realise that the use of equal subjective loudness levels will presumably lead to a smaller range of speech perception in noise results, as the effect of audibility is minimized. As mentioned above, a smaller range of speech perception in noise results induces a smaller variation to be explained by the regression models. Finally, also our analysis method differs from that of the previous studies. We included predictors in the regression models in two blocks, contrary to the previous studies in which regression analyses were performed. We first included the audiogram (and age) in the equations before the other measures were included. This may have overdrawn the role of the audiogram in the regression models.

Despite these substantial differences, our results agree qualitatively with the results from earlier studies. We found that the audiogram, age, and spectral and temporal resolution are the most important predictors for speech perception in noise, which is in line with most of the previous studies listed by Houtgast & Festen (2008, see Table 1). Noordhoek et al (2001), Dreschler & Plomp (1985), and Glasberg and Moore (1989) also reported both spectral and temporal resolution as significant predictors for speech perception in noise. Some others found that only spectral resolution (Festen & Plomp, 1983; Ter Keurs et al, 1993) or only temporal resolution (George et al, 2006 & 2007) predicted speech perception in noise significantly.

A quantitative comparison of our results and the studies listed by Houtgast & Festen (2008) shows that the predictive power of our models is lower.
than that of the previous studies. We realize that a direct comparison is not possible, due to language differences, the faster (potentially less accurate) tests, the different loudness levels that might have caused a smaller result range for speech perception in noise, and our different way of building regression models. We express the variance explained by our models in adjusted $R^2$ values: $R^2$ values that are corrected for the available degrees of freedom. These values are 33% for the SRT in stationary noise and 63% for the SRT in fluctuating noise (see Table 4.4). Nevertheless, these values are lower than found in Houtgast & Festen (2008). They reported uncorrected $R^2$ values of 62% and 85% for stationary and fluctuating noise, respectively (see Table 4.4).

4.4.6 Additional value of the preliminary Auditory Profile for speech perception in noise

A major aim of the present study was to investigate the additional value of the preliminary Auditory Profile, in addition to the pure-tone audiogram and age. This is a pragmatic issue, rather than seeking to understand the underlying causal factors. The Auditory Profile relates to differences among HI listeners, such as those attending audiology clinics for diagnostic assessment and advice regarding rehabilitative options. Seeing the relatively low predictive power of our regression models, one might argue that using all audiogram data might give better predictions of speech perception in noise. We therefore did additional comparative analyses by building regression models with only audiogram and age as independent variables. All audiogram thresholds (at six octave frequencies from 0.25 to 8 kHz) and age were entered in predictive models for $SRT_{stat}$ and $SRT_{fluct}$. We found that these regression models explained 34% and 50% of the variance in respectively $SRT_{stat}$ and $SRT_{fluct}$ (adjusted $R^2$ values). Comparing these values to the adjusted $R^2$ values from the models with the Auditory Profile factors described above (33% for $SRT_{stat}$ and 63% $SRT_{fluct}$) we conclude that certainly for the prediction of speech perception in fluctuating noise, the preliminary Auditory Profile had considerable added value. For the prediction of speech perception in stationary noise the added value is questionable, especially when considering the extra time that is needed to perform the Auditory Profile tests. However, the fact that a similar percentage of explained vari-
ance can be realized with a lower number of predictors in the model with Auditory Profile variables (three predictors) (or five when also counting the underlying audiogram thresholds for PTA and slope) than in the model with only audiogram and age (seven predictors) suggests that the Auditory Profile parameters are more effective in predicting speech perception in stationary noise, than the audiogram thresholds.

4.5 Conclusions

In the present study we examined the relevance of recruitment, spectral and temporal resolution, and cognitive abilities for speech perception in noise, all measured with the preliminary Auditory Profile test battery in four languages.

Based on correlation analyses, we conclude that previously published relationships between recruitment and speech perception in noise, and between spectral and temporal resolution and speech perception in noise, can be reproduced using the preliminary Auditory Profile data set in four languages. No significant correlation between cognitive abilities and speech perception in noise was found in our study sample. This may be specific to the newly developed lexical-decision making test that was used and to the significant centre effects in the data. It may also reflect the limited range of cognitive ability in the sample.

According to stepwise linear regression analyses, spectral and temporal resolution are the most important factors for the prediction of speech perception in noise, after accounting for the pure-tone audiogram and age. Recruitment, although significantly correlated with speech perception in noise, did not add much information to the predictions for reduced speech perception in noise based on the hearing thresholds themselves. The lexical decision-making test did not contribute significantly to the linear regression models. A comparison of these results with models predicting speech perception in noise based on all audiogram thresholds and age, showed that the preliminary Auditory Profile adds explanatory power for speech perception in fluctuating noise, in addition to the audiogram and age, at least for the range of hearing abilities included in our sample.