


Relations Between the Intelligibility of Speech in Noise and Psychophysical Measures of Hearing Measured in Four Languages Using the Auditory Profile Test Battery

Trends in Hearing
2015, Vol. 19: 1–12
© The Author(s) 2015
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/2331216515618902
tia.sagepub.com


T. E. M. Van Esch¹ and W. A. Dreschler¹

Abstract

The aim of the present study was to determine the relations between the intelligibility of speech in noise and measures of auditory resolution, loudness recruitment, and cognitive function. The analyses were based on data published earlier as part of the presentation of the Auditory Profile, a test battery implemented in four languages. Tests of the intelligibility of speech, resolution, loudness recruitment, and lexical decision making were measured using headphones in five centers: in Germany, the Netherlands, Sweden, and the United Kingdom. Correlations and stepwise linear regression models were calculated. In sum, 72 hearing-impaired listeners aged 22 to 91 years with a broad range of hearing losses were included in the study. Several significant correlations were found with the intelligibility of speech in noise. Stepwise linear regression analyses showed that pure-tone average, age, spectral and temporal resolution, and loudness recruitment were significant predictors of the intelligibility of speech in fluctuating noise. Complex interrelationships between auditory factors and the intelligibility of speech in noise were revealed using the Auditory Profile data set in four languages. After taking into account the effects of pure-tone average and age, spectral and temporal resolution and loudness recruitment had an added value in the prediction of variation among listeners with respect to the intelligibility of speech in noise. The results of the lexical decision making test were not related to the intelligibility of speech in noise, in the population studied.

Keywords

speech intelligibility, hearing impaired, clinical tests, multicenter study, audiological diagnosis, auditory profile

Introduction

As the result of an international research effort, an Auditory Profile test battery has been developed (Van Esch et al., 2013). This test battery characterizes hearing in terms of seven domains: loudness perception, spectral and temporal resolution, the intelligibility of speech 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 multicenter 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 of 30 normal-hearing subjects and investigated the clinical applicability and usability of each test in 72 hearing-impaired subjects. They concluded that the

individual tests showed a good test–retest reliability, could be performed well by naïve subjects, and the tests were relatively fast. The tests are therefore potentially applicable in clinical audiology in four different languages (Dutch, English, German, and Swedish). For normal-hearing listeners, results of all tests were comparable across centers after baseline correction to account for differences between test materials. However, for hearing-impaired listeners, differences between test materials had to be taken into account when interpreting the

¹Academic Medical Centre, Amsterdam, The Netherlands

Corresponding author:

W. A. Dreschler, Clinical and Experimental Audiology, ENT Department D2-240, Amsterdam Medical Centre, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands.
Email: w.a.dreschler@amc.uva.nl



results of the language-dependent tests (i.e., the tests that have different materials in each language). As a first validation step, the results of the Auditory Profile were compared 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 several tests from the Auditory Profile and the intelligibility of speech in noise. This serves as a second step of the validation of the Auditory Profile test battery and provides insight into the causes of reduced speech understanding in noise. The test results will indicate the added value of extra tests in the test battery and it 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, the intelligibility of speech 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 recognized that hearing-impaired listeners have poorer intelligibility of speech in noise than normal-hearing listeners, and that the difference increases in fluctuating noise, as hearing-impaired listeners have reduced ability to take advantage of the gaps in noise (e.g., De Laat & Plomp, 1983; Festen & Plomp, 1990; Versfeld & Dreschler, 2002). The reduced fluctuating-masker benefit in hearing-impaired subjects can partly be attributed to the fact that adaptive procedures converge to different signal to noise ratios for normal-hearing and hearing-impaired subjects (Bernstein & Grant, 2009) but also suprathreshold processing plays a role (e.g., Bacon, Opie, & Montoya, 1998; Eisenberg, Dirks, & Bell, 1995; Oxenham & Kreft, 2014; 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 intelligibility of speech 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 (Dreschler & Plomp, 1985; Festen & Plomp, 1983) and loudness recruitment (Brand & Hohmann, 2001, 2002; Villchur, 1974). These aspects of impairment have been shown to be related to the intelligibility of speech in noise. Relations between spectral and temporal resolution and the intelligibility of speech in noise were investigated by many authors (e.g., Dreschler & Leeuw, 1990; George, Festen, & Houtgast, 2006; George et al., 2007; Glasberg & Moore, 1989; Noordhoek, Houtgast, & Festen, 2001; Patterson, Nimmo-Smith, Weber, & Milroy, 1982; Van Esch & Dreschler, 2011). Villchur (1974) and Moore and

Glasberg (1993) demonstrated the influence of loudness recruitment on the intelligibility of speech in noise. Besides that, significant relations between SRT scores and outcomes of several cognitive abilities have been shown (e.g., George et al., 2007; Humes, 2002; Kramer, Zekveld, & Houtgast, 2009; Larsby, Hällgren, & Lyxell, 2012; for a review see Akeroyd, 2008). Several studies have also investigated the influence of both auditory and cognitive factors on speech recognition in noise. Houtgast and Festen (2008) gave an overview of several laboratory studies that related a subset of measures of hearing thresholds, spectral resolution, temporal resolution, intensity difference limen, age, and cognitive abilities to the SRT in noise. Their overall conclusion was that typically 70% of the variance in speech recognition data could be explained by these factors. However, one has to keep in mind that percentages of explained variance are strongly affected by the precision of the tests relative to the range of SRTs in noise, and thus by the range of hearing impairment included in a study. Moreover, Houtgast and Festen (2008) found that the audiogram and age were generally good predictors across the different studies.

Many of the aforementioned studies investigated the influence of only one factor, included only a small number of listeners, and used time-consuming experimental methods that are not applicable in clinical audiology. Despite the fact that loudness recruitment, spectral and temporal resolution, and cognitive abilities, all have been shown to play a significant role in the prediction of the intelligibility of speech in noise, it is not clear to what extent their information is mutually exclusive and of value in addition to the pure-tone audiogram and age. It can be difficult to unravel causal relationships from the correlations shown in such observational studies, due to covariation between the explanatory variables. The Auditory Profile data set offers the possibility of examining the relationships between the intelligibility of speech in noise and a broad set of test results, including loudness recruitment, spectral and temporal resolution, and cognitive abilities. The advantages of the Auditory profile study are that standardized tests were used and that the sample of listeners involved was relatively large.

In the Auditory Profile test battery, *loudness recruitment* is defined as the slope of the lower part of the loudness curve as measured with the Adaptive CAtegorical LOudness Scaling (ACALOS) test, developed by Brand and Hohmann (2001). Van Esch et al. (2013) showed that the ACALOS results from the Auditory Profile test battery corresponded very well (i.e., within 5dB) with data presented in ISO 16832 (2006) and with data from Brand and Hohmann (2001). *Spectral and temporal resolution* is measured in a combined test (the F/T test [test of frequency and temporal resolution]). Van Esch

et al. (2013) found good agreement between the spectral and temporal resolution results from the Auditory Profile and previously published results for hearing-impaired listeners. Moreover, Van Esch and Dreschler (2011) showed significant correlations between results from the F/T test and results from conventional spectral and temporal resolution tests.

Cognitive ability is measured in the Auditory Profile using a lexical decision-making test. Van Esch et al. (2013) reported that their results for normal-hearing Swedish subjects agreed very well with the results presented by Hällgren, Larsby, Lyxell, and Arlinger (2001), but that the results for the Swedish hearing-impaired subjects from Van Esch et al. (2013) showed better performance than those presented by Hällgren et al. (2001). Age is a possible explanation, because the hearing-impaired listeners in this study were slightly younger than in the study by Hällgren et al.

Finally, the intelligibility of speech in noise is measured using short meaningful sentences in the language corresponding to each center. Van Esch et al. (2013) showed that the results for the normal-hearing subjects in stationary noise from the Auditory Profile were quantitatively in line with the results obtained in previous studies and also that the effects of fluctuating noise and hearing loss agreed well with the literature.

A potential drawback of the Auditory Profile data set is the fact that some test center effects were found in the data of the hearing-impaired subjects from the five centers in four different countries. Van Esch et al. (2013) found that the results for the hearing-impaired subjects from the different centers differed significantly for the lexical decision test and for the intelligibility of speech in noise, even after the use of a language correction factor. Consequently, even with language-specific corrections, we have not been able to produce test materials that are completely language independent. This may hamper conclusions from multicenter studies across languages. This study will find out how critical this is in case of the relationships between SRT and the other tests in the Auditory Profile. This will determine the applicability of the Auditory Profile in future multicenter research.

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

1. As a further step in the validation of the Auditory Profile, we examined whether the well-known correlations between the intelligibility of speech, loudness recruitment, spectral and temporal resolution, and cognitive abilities were reproduced in the results from the Auditory Profile in four languages.

2. We investigated the additional value (relevance) of the Auditory Profile for the intelligibility of speech in noise, by identifying predictors of the intelligibility of speech in noise that are not available in the traditional approach. We applied a stepwise linear regression and included the traditionally known variables (pure-tone audiogram and age) in the first step and the other variables in successive steps. It should be recognized that this approach tends to overestimate the importance of audiogram and age, as some of the variation in the intelligibility of speech scores that is intrinsically caused by other psychophysical measures may be attributed to the audiogram or to age, due to covariation of the psychophysical measures with hearing threshold level.

Methods

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

Test Set-Up

The tests were implemented on the Oldenburg Measurement Application, 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 with free-field compensation. Experiments took place in sound-insulated booths. Written instructions were translated in the four languages (Dutch, English, German, and Swedish) and used in all centers, complemented with oral explanations when needed.

Centers and Listeners

The five participating centers were audiological centers 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). The first four centers included 15 hearing-impaired subjects, and UK-ISVR included 12 hearing-impaired subjects, summing up to a group of 72 subjects.

All centers were approved by their local research ethics committees for 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. The 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 the left and right ears are shown in Figure 1 for the listeners with sensorineural losses, defined as having an air-bone gap (ABG) of less than 10 dB averaged over thresholds at 0.5 and 1 kHz ($n = 58$), and the listeners with conductive components ($n = 14$) separately. The majority of the listeners had symmetric hearing losses, but there were 13 listeners in the group with an asymmetry of 10 dB or more (averaged over 0.5, 1, 2, and 4 kHz). Details of the listeners and conductive and asymmetric hearing losses are described by Van Esch et al. (2013).

Protocol

The Auditory Profile comprised tests on loudness perception (ACALOS), intelligibility for speech in noise (SRT), spectral and temporal resolution (F/T test), and cognitive abilities (Lexical decision-making test).²

The tests were conducted in test and retest in two sessions on separate days (1–3 weeks apart), except for the pure-tone audiogram which was measured only once. For the other tests, the means of test and retest values were used as the pooled measures in the present analyses. The tests did not show clinically relevant learning effects (Van Esch et al., 2013).

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

All auditory tests were conducted at similar subjective loudness levels. Individual loudness levels were obtained from the ACALOS test and presentation levels were limited to maxima of 95 and 85 dB sound pressure level for narrowband and broadband signals, respectively (see later for details).

In a reference group of 80 normal-hearing listeners (20 per languages), language-validation studies 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 average scores of the normal-hearing listeners for each language. The language-dependent tests comprised the SRT tests and the lexical decision-making test. In the present article, only corrected data (i.e., data after subtraction of reference values) are presented.

Test Procedures

ACALOS test. Loudness perception 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 50-point scale, based on which the stimulus level was adaptively varied. We used three types of stimuli: broadband speech-shaped noise (ICRA, see Dreschler, Verschuure, Ludvigsen, & Westermann, 2001) and narrow-band noises (with 1/3-octave bandwidth at 0.5 and 3 kHz). For each stimulus type, individual loudness growth curves were fit. From these curves

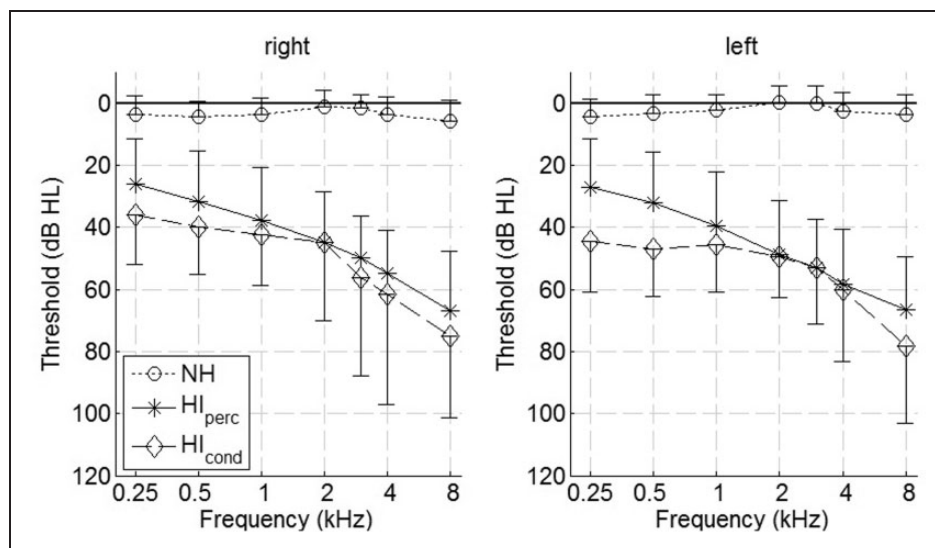


Figure 1. Average air-conduction hearing thresholds of left and right ears (mean \pm SD) of normal-hearing listeners (circles, dotted lines, $n = 30$) and hearing-impaired listeners with purely perceptive hearing loss (asterisks, solid lines, $n = 58$) or conductive components (diamonds, dashed lines, $n = 14$). Figure adapted from Van Esch et al. (2013).

most comfortable levels (MCL) were calculated as the levels corresponding to a perceived loudness of 20 categorical units, 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 loudness recruitment.

F/T test. Spectral and temporal resolution were measured at 0.5 and 3 kHz using a combined test, as described by Larsby and Arlinger (1999) with the modifications suggested by Van Esch and Dreschler (2011). The test measured 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 $\frac{1}{2}$ -octave spectral gap, centered at the test-tone frequency, was cut in the masking noise and 50 ms temporal gaps (the center of which coincided with the center of the test-tone pulse) were introduced in the $\frac{1}{2}$ -octave band noise to allow the assessment of temporal resolution.

SRT test. The intelligibility of speech performance in noise was measured using short meaningful sentences in the language corresponding to each center: the Dutch centers (NL-AMC and NL-VUMC) used the Versfeld sentences (Versfeld, Daalder, Festen, & Houtgast, 2000), DE-HZO used the Göttinger sentences (Brand & Kollmeier, 2002; Kollmeier & Wesselkamp, 1997), UK-ISVR used the BKB sentences (Bench, Kowal, & Bamford, 1979), and SE-LINK used the Swedish HINT sentences (Hällgren, Larsby, & Arlinger, 2006). Tests were conducted according to the local standards, as described in the works cited earlier. 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 the stationary noise. For the fluctuating noise, the ICRA5_250 or ICRA4_250 was used, which are male- and female-weighted versions of fluctuating speech-shaped noises (see Wagener, Brand, & Kollmeier, 2006). In both conditions, the speech level was adaptively varied to track the level (the SRT) at which 50% of the sentences was repeated correctly. The noise level was fixed at the individual MCL for broadband noise, with a maximum of 85 dB sound pressure level, and the outcome measure was the SRT expressed as signal to noise ratio in dB.

Lexical decision-making test. A measure of cognitive ability was obtained using the lexical decision-making test (Hällgren et al., 2001), which measured speed and accuracy of lexical access of subjects. This test was originally

developed in Sweden (Hällgren et al., 2001) and was translated into Dutch, English, and German in the HearCom (Hearing in the Communication Society) project (www.hearcom.eu). The task was to discriminate words from nonwords 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 correspond to better performance).

Statistical Analysis

Normality. Normality of the outcome measures was tested by Van Esch et al. (2013) by visual inspection and the Shapiro-Wilk and Kolmogorov-Smirnov tests. All outcome measures that we used in the present analyses were distributed (approximately) normally.

Linear regression analyses and inclusion of audiogram thresholds. Multiple linear regression analyses were performed in SPSS, on data for the total group of 72 listeners. SRTs for speech in stationary and fluctuating noise were used as dependent variables. The regression models involved stepwise inclusion of possible predictors (inclusion: $p < .05$ and exclusion: $p > .10$). As we wanted to investigate the relevance of the Auditory Profile tests in addition to the pure-tone audiogram and age, we first included the audiogram measures and age, before including all other measures. 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 (see Durbin & Watson, 1950, 1951).

Results

Relations between the SRT data, both in stationary and in fluctuating noise (SRT_{stat} and SRT_{fluct}) and the audiogram, age, loudness recruitment, spectral and temporal resolution, and cognitive abilities were analyzed. Audiogram measures were the pure-tone average (PTA, 0.5, 1, 2, 4 kHz), the slope of the audiogram (difference between thresholds at 0.5 and 4 kHz), and the ABG (average of 0.5 and 1 kHz). Measures of loudness recruitment were the slopes from the lower level parts of the ACALOS curves measured at 0.5 and 3 kHz (SL500 and SL3k). The spectral resolution at 0.5 and 3 kHz (F500 and F3k) and temporal resolution at 0.5 and 3 kHz (T500 and T3k) were the test results from the F/T test, expressed as RoM (more negative values indicate better performance). Lexical decision-making results were included as a measure of cognitive ability.

Table 1. Pearson Correlation Coefficients in the Group of 72 Hearing-Impaired Subjects Between SRT, Age, Audiogram, Loudness Recruitment, Spectral and Temporal Resolution, and Lexical Decision Making Results.

	Age	Audiogram			ACALOS		F/T test				Lexical decision making
		PTA	Slope	ABG	SL500	SL3k	F500	T500	F3k	T3k	
SRT _{stat}	0.29	0.46**	0.28	0.06	0.15	0.40**	0.40**	0.20	0.27	0.35*	0.10
SRT _{fluct}	0.34*	0.67**	0.14	0.03	0.36**	0.45**	0.56**	0.32	0.19	0.40**	0.06

Note. The following measures are included: SRT in stationary and fluctuating noise (SRT_{stat} and SRT_{fluct}), audiogram: PTA, slope, and ABG; loudness recruitment at 0.5 and 3 kHz (SL500 and SL3k); spectral and temporal resolution at 0.5 and 3 kHz (F500, T500, F3k, and T3k); and lexical decision-making results. For all auditory measures, results of analyses of right-ear data are displayed. Significant correlations after Bonferroni correction for multiple testing at the $p < .01$ and $p < .05$ level are marked ** and *, respectively. PTA = pure-tone average; ABG = air-bone gap; SRT = speech reception threshold; ACALOS = Adaptive CAtegorical LOudness Scaling; F/T = test of frequency and temporal resolution.

Table 2. Results of Stepwise Linear Regression Analyses of Right-Ear Data.

Dependent	Model	Predictors	R	Adjusted R ²
SRT _{stat}	Block 1	PTA, slope	0.57	0.30
	Blocks 1 and 2	PTA, slope, F500	0.63	0.36
SRT _{fluct}	Block 1	PTA, age	0.70	0.47
	Blocks 1 and 2	PTA, age, T3k, F500, SL3k	0.81	0.63

For two dependents (SRT_{stat} and SRT_{fluct}), the significant predictors from both blocks of the stepwise linear regression models are shown, as well as R and adjusted R^2 values of these models. All shown models have a significance $p < .001$. PTA = pure-tone average; SRT = speech reception threshold; F500 = spectral resolution at 0.5 kHz, T3k = temporal resolution at 3 kHz, and SL3k = loudness recruitment at 3 kHz.

Correlations

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

Note that the scales of all measures were arranged such that poorer performance corresponded to more positive values; hence, positive correlation coefficients were expected. Correlations for right-ear data are shown in Table 1. Left-ear data showed very similar results and are not shown here.

It can be seen that age and the PTA were correlated significantly with SRTs for speech in stationary and fluctuating noise. Most results from the ACALOS test and half of the results of the F/T test correlated significantly with the SRT results, and some of the correlations were about equally strong as those between the audiogram and SRT. No significant correlations between the cognitive test and SRT results were found.

Stepwise Linear Regression Analyses

We applied stepwise linear regression analyses to build models for the prediction of the SRT data for speech, both in stationary and in fluctuating noise. Analyses were conducted in two successive blocks. In the first block, age and audiogram measures were included in a stepwise procedure (the best predictor first, then adding

the second best, etc.) to find the significant predictors among those variables (PTA, slope, ABG, and age). In the second block, the other parameters from the Auditory Profile described earlier (spectral and temporal resolution, loudness recruitment, and lexical decision making) were included as explanatory factors, again in a stepwise procedure.

Analyses on right-ear data. Results of the analyses for right-ear data are displayed in Table 2. (The regression models were verified on the left-ear data – see later.) For predictions of SRTs for speech in stationary and in fluctuating noise, results were shown for the models with only age and audiogram variables (Block 1) and for the models with audiogram variables and other predictors (Blocks 1 and 2). For each model, the included variables, values of R and adjusted R^2 were calculated. Adjusted R^2 values gave the percentages of variance explained by the models, with a correction for the degrees of freedom. Variables were listed in the order that they were included by the stepwise procedures, so in order of decreasing R^2 values.

For stationary noise, both PTA and audiogram slope were significant predictors from the first block and explained 30% of the variance. Age and the ABG were not significant. When including the second block of variables, frequency resolution at 0.5 kHz improved the prediction of SRT for speech in stationary noise significantly (R^2 change: $p < .05$), while the other predictors

were not found to be significant. The regression model for the SRT for speech in stationary noise was

$$\text{SRT}_{\text{stat}} = 3.1 + 0.078 \times \text{PTA} + 0.065 \times \text{slope} \\ + 0.386 \times \text{F500}$$

This model explained 36% of the variance in the right-ear results of SRT for speech in stationary noise. Please note that SRT_{stat} refers to corrected SRT scores (see Methods section and Van Esch et al., 2013). This implied that SRT_{stat} for normal-hearing listeners was expected to be zero, on average, and higher values correspond to poorer intelligibility of speech. The regression model showed that, for the test group of hearing-impaired listeners, SRT_{stat} was on average 3.1 dB higher than the reference for normal-hearing listeners. Furthermore, SRT_{stat} increased 0.78 dB for every 10 dB of hearing loss (PTA) and 0.65 dB for every 10 dB/oct increment in audiogram slope. This underlined the importance of high-frequency hearing for speech intelligibility in stationary noise. Finally, listeners with poorer frequency resolution at 0.5 kHz had poorer SRT_{stat} (0.39 dB increment per dB RoM for spectral resolution).

For fluctuating noise, PTA and age were significant in the first block; together they explained 47% of the variance. Slope of the audiogram and ABG did not contribute significantly. Temporal resolution at 3 kHz, spectral resolution at 0.5 kHz, and loudness recruitment at 3 kHz improved the prediction significantly (R^2 change: $p < .001$), while lexical decision-making did not contribute significantly to the regression model. The regression equation for SRT for speech in fluctuating noise (corrected scores for) was

$$\text{SRT}_{\text{fluct}} = 2.43 + 0.125 \times \text{PTA} + 0.075 \times \text{age} + 0.309 \\ \times \text{T3k} + 0.442 \times \text{F500} + 4.32 \times \text{SL3k}$$

and explained 63% of the variance in SRT for speech in fluctuating noise. This equation showed that, for the test group of hearing-impaired listeners, $\text{SRT}_{\text{fluct}}$ was on average 2.4 dB higher than the reference for normal-hearing listeners, with additional increments proportional to their PTA, age, loss of spectral resolution at 0.5 kHz, loss of temporal resolution at 3 kHz, and loudness recruitment at 3 kHz.

Test-center effects. Van Esch et al. (2013) reported significant test-center effects in the SRT (and lexical decision-making) results. To investigate whether these differences were associated with a significant effect of center on the regression analyses, the linear regression models were evaluated separately for the different centers. The number of subjects were relatively low for a formal analysis, but scatter plots of predicted versus measured SRT scores by center (not shown) were visually inspected and

showed considerable overlap and very similar results for subjects in the different centers. Indeed, no significant effect of center was found in the residuals from the linear regression analyses (one-way ANOVA, $F(4,62) = 2.40$, $p = .06$ for SRT_{stat} , and $F(4,61) = 1.97$, $p = 0.111$ for $\text{SRT}_{\text{fluct}}$). Therefore, we concluded that similar regression models apply for predicting the SRTs for speech in noise at the different centers.

Generalizability and verification using 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 to the 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 (0.63 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 showed that the predictive power of the regression models was in the same order of magnitude for the left-ear data as for the right-ear data from which they were derived.

Discussion

In the present article, we examined the relevance of assessing loudness recruitment, spectral and temporal resolution, and lexical decision making for the prediction of the intelligibility of speech in stationary and fluctuating noise, all measured with the Auditory Profile test battery in four languages.

Loudness recruitment. Significant correlations between measures of loudness recruitment and SRTs ($p < .01$, see Table 1) were found. This is in agreement with a previous study that showed the influence of loudness recruitment on the intelligibility of speech in noise (Dreschler & Plomp, 1985). The results are also in line with two studies in which the effect of loudness recruitment on the intelligibility of speech in noise was simulated (Moore & Glasberg, 1993; Villchur, 1974).

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

Presumably, the small additional effect of loudness recruitment in the regression analyses was caused by

covariance between loudness recruitment and the audiogram ($r = .66$ and $r = .51$ at 0.5 and 3 kHz, respectively, both $p < .01$). As the correlations between the audiogram and SRT were of the same order of magnitude as those between loudness recruitment and SRT, the stepwise linear regression analysis model had to select either the hearing thresholds or loudness recruitment, while it is likely that both factors contributed to the variation. However, in the present study, we searched for additional predictive power of the Auditory Profile tests for the intelligibility of speech in noise, after accounting for the pure-tone audiogram and age. As loudness recruitment hardly improves the prediction of SRT, we conclude that loudness recruitment, although significantly correlated with the intelligibility of speech in noise, is only of minor importance for the prediction of the intelligibility of speech in fluctuating noise, once the audiogram has been taken into account. It was not possible to learn from our data whether the correlation between the audiogram and SRT was actually a consequence of threshold elevation, or a consequence of loudness recruitment, leading to apparent correlations with the audiogram due to the mutual correlation between the audiogram and loudness recruitment. However, presentation of the speech materials at similar loudness levels for all listeners should have reduced the influence of audibility, suggesting that threshold elevation itself may not have been the key factor.

Spectral and temporal resolution. In the present study, we found several significant correlations between spectral and temporal resolution and the intelligibility of speech in noise (see Table 1). This indicated that the previously reported relationship between spectral and temporal resolution and the intelligibility of speech in noise (e.g., Dreschler & Plomp, 1985; George et al., 2007; Noordhoek et al., 2001) also existed in the results from the Auditory Profile in four languages.

In the stepwise linear regression analyses, both spectral and temporal resolution were significant predictors of the 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 intelligibility of speech. In addition to the audiogram (PTA and slope), spectral resolution explained 6% of the variance in SRT. In fluctuating noise, temporal resolution at 3 kHz and spectral resolution at 0.5 kHz were significant predictors. Again, poorer resolution, both spectral and temporal, corresponded to poorer intelligibility of speech in noise. Altogether we conclude that spectral and temporal resolution were both important additional factors for the prediction of the intelligibility of speech in noise, after the audiogram has been taken into account.

The same considerations as described for loudness recruitment apply to measures of spectral and temporal

resolution. As the speech materials were presented at similar loudness levels for all listeners, the influence of audibility was reduced. The results suggested that audibility itself may not have been the key factor and part of the variation in the intelligibility of speech attributed to the audiogram in the regression models was actually mediated via reduced spectral and temporal resolution.

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 the intelligibility of speech in noise.

Larsby et al. (2012) examined the relation between results from the lexical decision-making test and the intelligibility of speech in noise for forty Swedish hearing-impaired listeners. They found significant correlations, even after partialling out the effects of PTA (0.5, 1, 2, 4 kHz). Their finding was in line with Akeroyd (2008), who concluded that there was a link between several different cognitive abilities and the intelligibility of speech in noise. However, the relation between cognitive abilities and the intelligibility of speech in noise that Larsby et al. (2012) found was not present for our group of listeners from different countries, neither in continuous noise nor in fluctuating noise. A possible explanation is that language-dependent factors may compromise the relationship between lexical decision making and the intelligibility for speech in noise. However, when analyzed in subgroups for each language, no significant correlations between the results for the lexical decision-making test and the intelligibility of speech in noise were found. To avoid testing in small subgroups, an analysis of variance (an univariate general linear model) was conducted on the SRT results with language as factor and the results of the lexical decision making test as covariate. The main effect *language* was significant, but the main effect *lexical decision* and the interaction effect between *language* and *lexical decision* were not. Finally, an explanation could be the limited range of cognitive abilities in our study sample: The language-corrected outcome measure in %correct divided by response time in ms ranged from -0.03 up to 0.08 , and there was considerable overlap between the results of hearing-impaired subjects and the results of the normal-hearing subjects in the reference group.

Test center effect. Van Esch et al. (2013) reported significant center effects for the SRT (and lexical decision making) results. However, this does not necessarily imply that there is also a significant center effect in the prediction of the intelligibility of speech in noise. In the present study, no significant center effect was found in the residuals from the linear regression analyses.

Nevertheless, center effects in the data may have contaminated or weakened the relationships between tests.

Predictions of the intelligibility of speech in noise in the literature. Predictive factors (both auditory and cognitive) for the intelligibility of speech in noise have been examined in several previous studies, see Houtgast and Festen (2008) for a review. The present study differs from the studies described by Houtgast and Festen in several aspects. Most importantly, the measurements in the present study were conducted in different languages and at different centers. Moreover, the tests from the Auditory Profile that were used in the present study, were, in general, faster to administer than the procedures that were used in most of the previous laboratory studies. A third essential difference concerned the use of equal subjective loudness levels in the present study. This new approach has both advantages and disadvantages, as discussed by Van Esch and Dreschler (2011) and Van Esch et al. (2013). It is important to realize that the use of equal subjective loudness levels will presumably lead to a smaller range of the results for the intelligibility of speech in noise, as the effect of audibility is reduced. As mentioned earlier, a smaller range of the results leads to less variation to be explained by the regression models. Finally, our analysis method differed from that of the previous studies. We included predictors in the regression models in two blocks, contrary to previous studies in which regression analyses were performed. We included the audiogram (and age) in the equations before the other measures were included. This may have overemphasized the role of the audiogram in the regression models, but it is a fair way to determine the real added value of the other test parameters provided by the Auditory Profile.

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

A quantitative comparison of our results and the studies listed by Houtgast and Festen (2008) showed 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 fixed loudness levels that might have caused a smaller range of scores for the

intelligibility of speech in noise, and our different way of building regression models. We expressed the variance explained by our models as adjusted R^2 values: R^2 values that have been corrected for the available degrees of freedom. These values are 36% for the SRT for speech in stationary noise and 63% for the SRT for speech in fluctuating noise (see Table 2). These values tended to be somewhat lower than in the studies reviewed by Houtgast and Festen (2008). But the main outcomes were in line with these studies.

It can be concluded that—especially for speech perception in stationary noise—a considerable degree of variance could not be explained by the test parameters used as predictors, despite the diversity and expected relevance of the tests included in the Auditory Profile. An interesting finding is that the SRT for speech in fluctuating noise could be predicted better than the SRT for speech in stationary noise. The results of this study suggested that temporal resolution is one of the factors responsible for the higher percentages of variance explained for SRTs for speech in fluctuating noise. But the results obtained in Block 1 of the analyses indicated that the percentage of explained variance was also higher for SRT for speech in fluctuating noise for the predictions based on audiogram and age only. This suggested that threshold levels in the gaps may have played an important role in speech intelligibility in fluctuating noise as well.

Additional value of the Auditory Profile for the intelligibility of speech in noise. A major aim of the present study was to investigate the value of the Auditory Profile in addition to the pure-tone audiogram and age. Rather than seeking to understand the underlying causal factors, we followed a pragmatic approach to investigate the value of additional testing in a clinical environment. The Auditory Profile related to differences among hearing-impaired listeners, such as those attending audiology clinics for diagnostic assessment and advice regarding rehabilitative options. Giving the relatively low predictive power of our regression models, one might argue that the use of all audiogram data would have given better predictions of the intelligibility of speech in noise. We therefore did additional comparative analyses by building regression models with only audiogram and age as independent variables. All audiogram thresholds (at 6-octave frequencies from 0.25 to 8 kHz) and age were entered into predictive models for SRT_{stat} and SRT_{fluct} . We found that the regression models on audiogram and age explained 37% and 52% of the variance in SRT_{stat} and SRT_{fluct} , respectively (adjusted R^2 values). Comparing these values to the adjusted R^2 values from the models with the Auditory Profile factors described earlier (36% for SRT_{stat} and 63% SRT_{fluct}), we concluded that the added value of the Auditory Profile

was mainly found for prediction of the intelligibility of speech in fluctuating noise. For the prediction of the intelligibility of speech in stationary noise, the added value seems to be questionable, especially when considering the extra time that is needed to perform the Auditory Profile tests.

However, from the analyses of the hearing thresholds per frequency, we learned that the use of hearing losses at specific frequencies resulted in higher percentages of explained variance than the use of *overall* measures as PTA and audiometric slope, even if the number of factors is kept constant. For SRT for speech in stationary noise, a set of only two predictors using the hearing loss at 4 kHz (HL4k) and frequency resolution at 500 Hz (F500) resulted in R values of 0.66 (explaining 42% of the variance). The regression model was

$$\text{SRT}_{\text{stat}} = 2.1 + 0.105 \times \text{HL4k} + 0.373 \times \text{F500}$$

Also for SRT for speech in fluctuating noise, HL4k and F500 proved to be the most important predictors. A simple regression model resulted in R -values of 0.77 (explaining 58% of the variance). The regression model was

$$\text{SRT}_{\text{fluct}} = 9.5 + 0.152 \times \text{HL4k} + 0.827 \times \text{F500}$$

Both analyses underline the relative importance of the hearing loss at 4 kHz for the prediction of SRT for speech in stationary as well as fluctuating noises and show that the test of frequency resolution at 500 kHz, as included in the Auditory Profile, contributes essential information, in addition to the data that is already available in the pure-tone audiogram.

In a number of cases Auditory Profile parameters have shown to be more effective in predicting the intelligibility of speech in stationary noise, than the audiogram thresholds alone. When the clinician needs extra information, but if there is only limited time for extra testing, Auditory Profile tests can be used selectively in order to contribute additional information for diagnostic purposes and for hearing aid selection and fitting. For clinical use, the test of frequency resolution at 500 Hz (F500) seems to be worthwhile, especially in complex cases.

Conclusions

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

On the basis of correlation analyses, we concluded that previously published relationships between loudness recruitment and the intelligibility of speech in noise, and

between spectral and temporal resolution and the intelligibility of speech in noise, can be reproduced using the Auditory Profile data set in four languages. No significant correlation between lexical decision making and the intelligibility of speech in noise was found. This may be specific to the newly developed lexical decision making test that was used and to the significant center effects in the data. It may also reflect the limited range of outcome values in the sample.

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

Acknowledgements

The authors like to thank the colleagues from the labs and clinics that were involved in this study: Mark Lutman and Sheetal Athalye (Institute for Sound and Vibration Research, Southampton, United Kingdom, UK), Matthias Vormann and Birger Kollmeier (HörTech gGmbH, Oldenburg, Germany), Johannes Lyzenga and Tammo Houtgast (Free University Hospital, Amsterdam, The Netherlands), and Matthias Hällgren and Brigitta Larsby (Linköping University, Linköping, Sweden). The authors also thank Daniel Berg for technical support and implementation of the tests on the platform (Oldenburg Measurement Application), Kirsten Wagener for her work on consolidation of the speech tests, and Jan Wouters for his useful comments on an earlier version of this manuscript. Finally, the authors thank the subjects for their participation.

Author Note

Parts of this work have been presented at the *International Hearing Aid Research Conference: Evaluation of the 'Auditory Profile' Test Battery in an International Multicenter Study*. Data from the present study were published by Van Esch et al. (2013). Examples of words and nonwords from the lexical decision-making test in the four languages are available from the HearCom website (www.hearcom.eu). 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 his/her sole risk and

liability. The authors alone are responsible for the content and writing of the paper.

Declaration of conflicting interests

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

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Supported by grants from the European Union FP6, Project 0004171 HearCom.

Notes

1. NL-AMC: no. 05/127 # 05.17.0934, dated August 3, 2005; HZO-DE: "Klinische Tests zur Bestimmung individueller Hördefizite und Kommunikationsfähigkeiten," dated November 15, 2006; ISVR-UK: 791, dated February 13, 2007; SE-LINK: M83-06; VUMC-NL: MEC05/12 – 2006/171, dated November 2, 2006.
2. The Auditory Profile also included a measurement of the intelligibility of speech 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.

References

- Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *International Journal of Audiology*, *47*, S53–S71.
- Bacon, S. P., Opie, J. M., & Montoya, D. Y. (1998). The effects of hearing loss and noise masking on the masking release for speech in temporally complex backgrounds. *Journal of Speech Language and Hearing Research*, *41*, 549–563.
- Bench, J., Kowal, A., & Bamford, J. (1979). The BKB (Bamford-Kowal-Bench) sentence lists for partially hearing children. *British Journal of Audiology*, *13*, 108–112.
- Bernstein, J. G. W., & Grant, K. W. (2009). Auditory and auditory-visual intelligibility of speech in fluctuating maskers for normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America*, *125*, 3358–3372.
- Brand, T., & Hohmann, V. (2001). Effect of hearing loss, centre frequency, and bandwidth on the shape of loudness functions in categorical loudness scaling. *Audiology*, *40*, 92–103.
- Brand, T., & Hohmann, H. (2002). An adaptive procedure for categorical loudness scaling. *Journal of the Acoustical Society of America*, *112*, 1597–1604.
- Brand, T., & Kollmeier, B. (2002). Efficient adaptive procedures for threshold and concurrent slope estimates for psychophysics and speech intelligibility tests. *Journal of the Acoustical Society of America*, *111*, 2801–2810.
- De Laat, J. A. P. M., & Plomp, R. (1983). The reception threshold of interrupted speech for hearing-impaired listeners. In R. Klinke, & R. Hartman (Eds.), *The reception threshold of interrupted speech for hearing-impaired listeners. Hearing—Physiological bases and psychophysics*, 359–363.
- Dreschler, W. A., & Leeuw, A. R. (1990). Speech reception in reverberation related to temporal resolution. *Journal of Speech and Hearing Research*, *33*, 181–187.
- Dreschler, W. A., & Plomp, R. (1985). Relations between psychophysical data and speech perception for hearing-impaired subjects. II. *Journal of the Acoustical Society of America*, *78*, 1261–1270.
- Dreschler, W. A., Verschuure, H., Ludvigsen, C., & Westermann, S. (2001). ICRA noises: Artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment. International Collegium for Rehabilitative Audiology. *Audiology*, *40*, 148–157.
- Durbin, J., & Watson, G. S. (1950). Testing for serial correlation in least squares regression, I. *Biometrika*, *37*(3–4), 409–428.
- Durbin, J., & Watson, G. S. (1951). Testing for serial correlation in least squares regression, II. *Biometrika*, *38*(1–2), 159–179.
- Eisenberg, L. S., Dirks, D. D., & Bell, T. S. (1995). Speech recognition in amplitude-modulated noise of listeners with normal and listeners with impaired hearing. *Journal of Speech and Hearing Research*, *38*, 222–233.
- Festen, J. M., & Plomp, R. (1983). Relations between auditory functions in impaired hearing. *Journal of the Acoustical Society of America*, *73*, 652–662.
- Festen, J. M., & Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *Journal of the Acoustical Society of America*, *88*, 1725–1736.
- George, E. L., Festen, J. M., & Houtgast, T. (2006). Factors affecting masking release for speech in modulated noise for normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America*, *120*, 2295–2311.
- George, E. L., Zekveld, A. A., Kramer, S. E., Goverts, S. T., Festen, J. M., Houtgast, T. (2007). Auditory and nonauditory factors affecting speech reception in noise by older listeners. *Journal of the Acoustical Society of America*, *121*, 2362–2375.
- Glasberg, B. R., & Moore, B. C. (1989). Psychoacoustic abilities of subjects with unilateral and bilateral cochlear hearing impairments and their relationship to the ability to understand speech. *Scandinavian Audiology (Supplementum)*, *32*, 1–25.
- Hällgren, M., Larsby, B., & Arlinger, S. (2006). A Swedish version of the hearing in noise test (HINT) for measurement of speech recognition. *International Journal of Audiology*, *45*, 227–237.
- Hällgren, M., Larsby, B., Lyxell, B., & Arlinger, S. (2001). Evaluation of a cognitive test battery in young and elderly normal-hearing and hearing-impaired persons. *Journal of the American Academy of Audiology*, *12*, 357–370.
- Houtgast, T., & Festen, J. M. (2008). On the auditory and cognitive functions that may explain an individual's elevation of the speech reception threshold in noise. *International Journal of Audiology*, *47*, 287–295.

- Humes, L. E. (2002). Factors underlying the speech-recognition performance of elderly hearing-aid wearers. *Journal of the Acoustical Society of America*, *112*, 1112–1132.
- ISO 389-1. (1998). *Acoustics—Reference zero for the calibration of audiometric equipment—Part 1: Reference equivalent threshold sound pressure levels for pure tones and supra-aural earphones*. Geneva, Switzerland: International Organization for Standardization.
- ISO 16832. (2006). *Acoustics—Loudness scaling by means of categories*. Geneva, Switzerland: International Organization for Standardization.
- Kollmeier, B., & Wesselkamp, M. (1997). Development and evaluation of a German sentence test for objective and subjective speech intelligibility assessment. *Journal of the Acoustical Society of America*, *102*, 2412–2421.
- Kramer, S. E., Zekveld, A. A., & Houtgast, T. (2009). Measuring cognitive factors in speech comprehension: The value of using the text reception threshold test as a visual equivalent of the SRT test. *Scandinavian Journal of Psychology*, *50*, 507–515.
- Larsby, B., & Arlinger, S. (1999). Auditory temporal and spectral resolution in normal and impaired hearing. *Journal of the American Academy of Audiology*, *10*, 198–210.
- Larsby, M., Hllgren, M., & Lyxell, B. (2012). The role of working memory capacity and speed of lexical access in speech recognition in noise. In T. Dau, M. L. Jepsen, J. C. Dalsgaard, & T. Poulsen (Eds.), *Speech Perception and Auditory Disorders (494 pp)* (pp. 95–102), ISBN 978-87-990013-3-0.
- Moore, B. C. J., & Glasberg, B. R. (1993). Simulation of the effects of loudness recruitment and threshold elevation on the intelligibility of speech in quiet and in a background of speech. *Journal of the Acoustical Society of America*, *94*, 2050–2062.
- Noordhoek, I. M., Houtgast, T., & Festen, J. M. (2001). Relations between intelligibility of narrow-band speech and auditory functions, both in the 1-kHz frequency region. *Journal of the Acoustical Society of America*, *109*, 1197–1212.
- Oxenham, A. J., & Krefl, H. A. (2014). Speech perception in tones and noise via cochlear implants reveals influence of spectral resolution on temporal processing. *Trends in Hearing*, *18*, 1–14.
- Patterson, R. D., Nimmo-Smith, I., Weber, D. L., & Milroy, R. (1982). The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold. *Journal of the Acoustical Society of America*, *72*, 1788–1803.
- Plomp, R. (1986). A signal-to-noise ratio model for the speech-reception threshold of the hearing impaired. *Journal of Speech and Hearing Research*, *29*, 146–154.
- Summers, V., & Molis, M. R. (2004). Speech recognition in fluctuating and continuous maskers: Effects of hearing loss and presentation level. *Journal of Speech Language and Hearing Research*, *47*, 245–256.
- ter Keurs, M., Festen, J. M., & Plomp, R. (1993). Limited resolution of spectral contrast and hearing loss for speech in noise. *Journal of the Acoustical Society of America*, *94*, 1307–1314.
- Van, Esch, T. E., & Dreschler, W. A. (2011). Measuring spectral and temporal resolution simultaneously: A comparison between two tests. *International Journal of Audiology*, *50*, 477–490.
- Van, Esch, T. E., Kollmeier, B., Vormann, M., Lyzenga, J., Houtgast, T., & Hällgren, M., et al. (2013). Evaluation of the preliminary auditory profile test battery in an international multi-centre study. *International Journal of Audiology*, *52*, 305–321.
- Versfeld, N. J., Daalder, L., Festen, J. M., & Houtgast, T. (2000). Method for the selection of sentence materials for efficient measurement of the speech reception threshold. *Journal of the Acoustical Society of America*, *107*, 1671–1684.
- Versfeld, N. J., & Dreschler, W. A. (2002). The relationship between the intelligibility of time-compressed speech and speech in noise in young and elderly listeners. *Journal of the Acoustical Society of America*, *111*, 401–408.
- Villchur, E. (1974). Simulation of the effect of recruitment on loudness relationships in speech. *Journal of the Acoustical Society of America*, *56*, 1601–1611.
- Wagener, K. C., Brand, T., & Kollmeier, B. (2006). The role of silent intervals for sentence intelligibility in fluctuating noise in hearing-impaired listeners. *International Journal of Audiology*, *45*, 26–33.