

## **Auditory tests for characterizing hearing deficits: The BEAR test battery**

Raul Sanchez-Lopez<sup>a,\*</sup>, Silje Grini Nielsen<sup>a</sup>, Mouhamad El-Haj-Ali<sup>b</sup>, Federica Bianchi, Michal Fereczkowski<sup>a,b,c</sup>, Oscar M Cañete<sup>a</sup>, Mengfan Wu<sup>b,c</sup>, Tobias Neher<sup>b,c</sup>, Torsten Dau<sup>a</sup> and Sébastien Santurette<sup>a,d</sup>

<sup>a</sup> *Hearing Systems Section, Department of Health Technology, Technical University of Denmark, Kgs. Lyngby, Denmark.*

<sup>b</sup> *Institute of Clinical Research, University of Southern Denmark, Odense, Denmark.*

<sup>c</sup> *Research Unit for Oto-Rhino-Laryngology, Odense University Hospital, Odense, Denmark*

<sup>d</sup> *Centre for Applied Audiology Research, Oticon A/S, Smørum, Denmark*

\* Corresponding author: Raul Sanchez-Lopez, Hearing Systems Section, Department of Health Technology, Ørstedes Plads 352, R-111, Technical University of Denmark (DTU), 2800 Kgs. Lyngby, Denmark.

Email: [rsalo@dtu.dk](mailto:rsalo@dtu.dk)

## **Auditory tests for characterizing hearing deficits: The BEAR test battery**

**Introduction:** The Better hEARing Rehabilitation (BEAR) project aims to provide a new clinical profiling tool – a test battery – for hearing loss characterization.

Whereas the loss of sensitivity can be efficiently measured using pure-tone audiometry, the assessment of supra-threshold hearing deficits remains a challenge. In contrast to the classical ‘attenuation-distortion’ model, the proposed BEAR approach is based on the hypothesis that the hearing abilities of a given listener can be characterized along two dimensions reflecting independent types of perceptual deficits (distortions). A data-driven approach provided evidence for the existence of different auditory profiles with different degrees of distortions.

**Design:** Eleven tests were included in a test battery, based on their clinical feasibility, time efficiency and related evidence from the literature. The tests were divided into six categories: audibility, speech perception, binaural processing abilities, loudness perception, spectro-temporal modulation sensitivity and spectro-temporal resolution. **Study sample:** Seventy-five listeners with symmetric, mild-to-severe sensorineural hearing loss were selected from a clinical population. **Results:** The analysis of the results showed interrelations among outcomes related to high-frequency processing and outcome measures related to low-frequency processing abilities. **Conclusions:** The results showed the ability of the tests to reveal differences among individuals and their potential use in clinical settings.

200 words

**Keywords:** Hearing loss, supra-threshold auditory deficits, auditory processing, psychoacoustics

## 1 **Introduction**

2 In current clinical practice, hearing loss (HL) is diagnosed mainly on the basis of pure-  
3 tone audiometry (ISO 8253-1, 2010). The audiogram helps differentiate between  
4 conductive and sensorineural hearing losses and can characterize the severity of the  
5 hearing loss from mild to profound. However, the pure-tone audiogram only assesses  
6 the sensitivity to simple sounds, which is not necessarily related to listening abilities at  
7 supra-threshold sound pressure levels (e.g. a person's ability to discriminate speech in  
8 noise).

9 Pure-tone audiometry is often complemented by speech audiometry (ISO 8253-3,  
10 2012), which is a test of word recognition performance in quiet. Although this test can  
11 provide information about supra-threshold deficits (Gelfand, 2009), measurements of  
12 speech understanding in noise have been found more informative (Killion, Niquette,  
13 Gudmundsen, Revit, & Banerjee, 2004; Nilsson, Soli, & Sullivan, 1994). Since  
14 improving speech intelligibility is usually the main goal of successful hearing  
15 rehabilitation, several auditory factors affecting speech intelligibility in noise have been  
16 investigated (e.g. Glasberg & Moore, 1989; Houtgast & Festen, 2008; Strelcyk & Dau,  
17 2009). Audibility (in conditions with fluctuating maskers), frequency selectivity (in  
18 conditions with stationary noise), and temporal processing acuity (in conditions with  
19 speech interferers), have been identified as important factors affecting speech reception  
20 thresholds in noise (e.g. Desloge, Reed, Braida, Perez, & D'Aquila, 2017; Johannesen,  
21 Pérez-González, Kalluri, Blanco, & Lopez-Poveda, 2016; Oxenham & Simonson, 2009;  
22 Rhebergen, Versfeld, & Dreschler, 2006). Thus, a hearing evaluation that goes beyond  
23 pure-tone sensitivity and speech intelligibility in quiet would be expected to provide a  
24 more accurate characterization of a listener's hearing deficits.

25

26 In Denmark, the Better hEARing Rehabilitation (BEAR) project was initiated with the  
27 aim of developing new diagnostic tests and hearing-aid compensation strategies for  
28 audiological practice. Although the assessment of individual hearing deficits can be  
29 complex, new evidence suggests that the perceptual consequences of a hearing loss can  
30 be characterized effectively by two types of hearing deficits, defined as “auditory  
31 distortions” (Sanchez-Lopez, Bianchi, Fereczkowski, Santurette, & Dau, 2018). By  
32 analysing the outcomes of two previous studies (Johannesen et al., 2016; Thorup et al.,  
33 2016) with a data-driven approach, Sanchez-Lopez et al. (2018) identified high-  
34 frequency hearing loss as the main predictor of one of the distortions, whereas the  
35 definition of the second type of distortion was inconclusive. The mixed results obtained  
36 from these analyses were most likely due to differences between the two studies in  
37 terms of hearing loss profiles and outcome measures. Here, a new dataset was therefore  
38 collected based on a heterogeneous group of listeners with audiometric hearing losses  
39 ranging from very mild to severe and with a large range of audiometric profiles. To that  
40 end, the most informative tests resulting from the analysis of Sanchez-Lopez et al.  
41 (2018) were included, together with additional auditory tests that had shown potential  
42 for hearing profiling in other previous studies. The tests included in the current study  
43 are referred to as the BEAR test battery.

44

45 The characterization of hearing deficits beyond the audiogram was considered in several  
46 earlier studies (e.g. Brungart, Sheffield, & Kubli, 2014; Rönnberg et al., 2016;  
47 Santurette & Dau, 2012; Saunders, Field, & Haggard, 1992; Vlaming et al., 2011).  
48 Among them, the HEARCOM project (Vlaming et al., 2011) proposed an extended  
49 hearing profile formed by the results of several behavioural tests. These tests targeted  
50 various auditory domains, such as audibility, loudness perception, speech perception,

51 binaural processing, and spectro-temporal resolution, as well as a test of cognitive  
52 abilities. Importantly, while the auditory domains considered in the BEAR test battery  
53 are similar to the ones considered in the HEARCOM project, the BEAR project aims to  
54 additionally classify the patients in subcategories and to create a link between hearing  
55 capacities and hearing-aid parameter settings.

56

57 The tests included in the BEAR test battery were chosen based on the following criteria:

58 1) There is evidence from the hearing research literature that the considered test is  
59 informative and reliable; 2) The outcomes of the test may be linked to a hearing-aid  
60 fitting strategy; 3) The outcome measures are easy to interpret and to explain to the  
61 patient; 4) The task is reasonably time-efficient or can be suitably modified to meet this  
62 requirement (e.g., by changing the test paradigm or developing an out-of-clinic  
63 solution); 5) The test implementation can be done with equipment available in clinics;  
64 6) The tasks are not too demanding for patients and clinicians; 7) Tests with several  
65 outcome measures are prioritized, and 8) The tests are language independent.

66

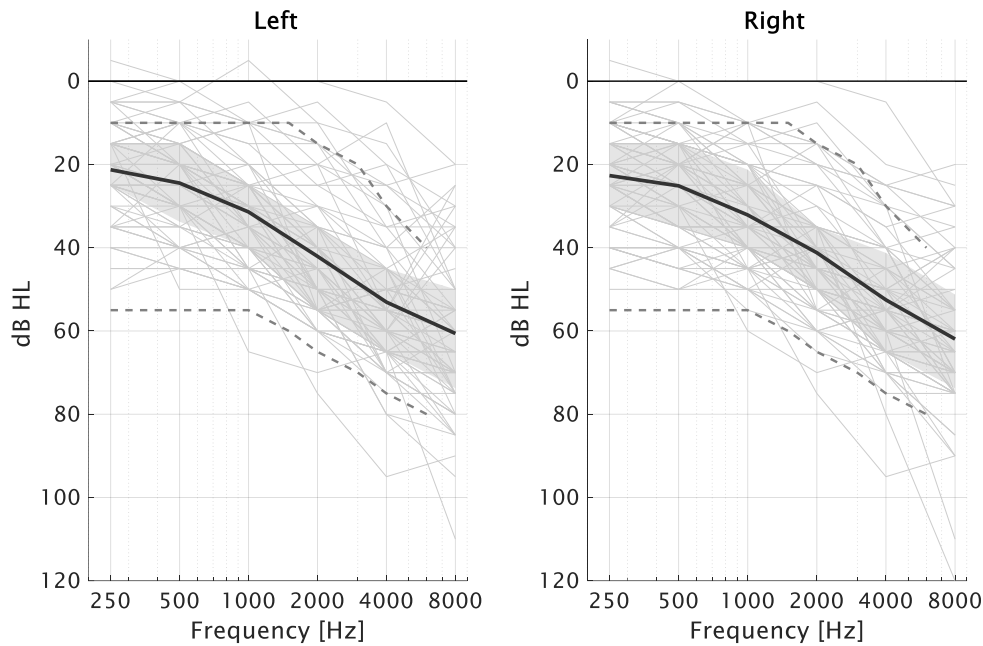
67 The selected test battery included measures of audibility, loudness perception, speech  
68 perception, binaural processing abilities, spectro-temporal modulation (STM) sensitivity  
69 and spectro-temporal resolution. It was implemented and tested in normal-hearing (NH)  
70 and hearing-impaired (HI) listeners. The goals of the study were: 1) To collect reference  
71 data from a representative sample of HI listeners for each of the selected tests, 2) to  
72 analyse the test-retest reliability of these tests, 3) to analyse the relationships between  
73 the different outcome measures, and 4) to propose a version of the test battery that can  
74 be implemented in hearing clinics.

75 **General methods**76 *Participants and general setup*

77 Seventy-five listeners (38 females) participated in the study, who were aged between 59  
78 and 82 years (median: 71 years). Five participants were considered NH ( $PTA \leq 25$  dB  
79 HL). PTA was defined as the pure-tone average between 0.5, 1 and 2 kHz. The HI  
80 group consisted of 70 participants with symmetric sensorineural hearing losses.

81 Symmetric sensorineural hearing loss was defined as an interaural difference (ID)  $\leq 15$   
82 dB HL at frequencies below 8 kHz and ID  $\leq 25$  dB HL at 8 kHz and air-bone gap  $< 10$   
83 dB HL. The pure-tone audiograms of the participants are shown in Figure 1.

84 The participants were recruited from the BEAR database (Wolff et al., 2020) at Odense  
85 University Hospital (OUH), from the patient database at Bispebjerg Hospital (BBH),  
86 and from the database at the Hearing Systems Section at the Technical University of  
87 Denmark (DTU). The tests were performed in a double-walled sound-insulated booth.  
88 Most of the tests were implemented using a modular framework for psychoacoustic  
89 experiments (AFC; Ewert, 2013) and the stimuli were presented through headphones.  
90 The study was approved by the Science-Ethics Committee for the Capital Region of  
91 Denmark H-16036391. All participants gave written informed consent and received  
92 financial compensation for their participation.



93

94 Figure 1: Audiograms of the 75 participants of the study together with the average for  
 95 each ear (dark solid lines) and interquartile ranges (grey areas). The grey dashed lines  
 96 correspond to the standard audiograms N1 and N4 from Bisgaard, Vlaming, &  
 97 Dahlquist (2010).

### 98 *Analysis of test reliability*

99 The test-retest reliability of the test battery was assessed using intraclass correlation  
 100 coefficients (ICC; Koo & Li, 2016) and the standard error of measurement (SEM;  
 101 Stratford & Goldsmith, 1997). Test-retest measurements were performed with a  
 102 subgroup consisting of seven HI and three NH participants for all tests of the test  
 103 battery. The seven HI listeners had bilateral HL with a mean PTA of 31 dB HL. The  
 104 retest session was conducted within four months after the first visit.

### 105 *Overview of the test battery*

106 The proposed tests are divided into six categories. Table 1 shows the tests and the  
 107 corresponding auditory domains. The following sections present all tests individually,  
 108 the experimental method and the summary statistics of the outcome measures presented

109 in Table 2. The dataset is publicly available in a Zenodo repository (Sanchez-Lopez et  
 110 al., 2019). More details about the method can be found in the supplementary material  
 111 and in the data repository.

112 Table 1: List of the tests included in the BEAR test battery and their corresponding  
 113 auditory domains.

<i>Test</i>	<i>Test domain</i>
<i>Pure-tone audiometry:</i>	
<i>Fixed level frequency threshold (eAUD-HF)</i>	<i>Audibility</i>
<i>Word recognition scores in quiet (WRS-4UFC):</i>	<i>Speech perception</i>
<i>Hearing in noise test (HINT)</i>	<i>tests</i>
<i>Maximum frequency for IPD detection (IPD<sub>fmax</sub>)</i>	
<i>Binaural pitch (Bpitch)</i>	<i>Binaural processing</i>
<i>Extended binaural audiometry in noise (eAUD-B)</i>	<i>abilities</i>
<i>Adaptive categorical loudness scaling (ACALOS)</i>	<i>Loudness perception</i>
<i>Fast spectro-temporal modulation sensitivity (fSTM)</i>	<i>Spectro-temporal</i>
	<i>modulation</i>
<i>Extended audiometry in noise:</i>	
<i>Tone in noise detection test (eAUD-N)</i>	<i>Spectro-temporal</i>
<i>Spectral masking release condition (eAUD-S)</i>	<i>resolution</i>
<i>Temporal masking release condition (eAUD-T)</i>	

114

115 Table 2: Summary statistics of the outcome measures of the BEAR test battery for the  
 116 NH and HI group. The results are presented in terms of mean, standard deviation (SD)  
 117 and the 1<sup>st</sup> (Q1) and 3<sup>rd</sup> quantiles (Q3) for the right ear (RE), left ear (LE) or both ears  
 118 (Bin). In the case of frequency-specific examination, the frequency range is either low  
 119 (LF) or high (HF).

<i>Outcome measure</i>	<i>Freq. Range</i>	<i>Ear</i>	<i>NH</i>			<i>HI</i>		
			<i>Mean (SD)</i>	<i>Q1</i>	<i>Q3</i>	<i>Mean (SD)</i>	<i>Q1</i>	<i>Q3</i>
<i>SRT<sub>Q</sub> (dB)</i>		<i>LE</i>	19.9 (7.1)	16.5	19.2	41.5 (13.5)	31.8	50.6
		<i>RE</i>	23.3 (8.9)	17.2	29.0	42.7 (12.6)	33.9	51.1
<i>Max DS (%)</i>		<i>LE</i>	99.2 (1.6)	100.0	100.0	97.2 (4.1)	95.3	100.0
		<i>RE</i>	97.2 (1.8)	95.5	97.6	93.9 (6.4)	92.1	98.4
<i>SRT<sub>N</sub> (dB)</i>		<i>LE</i>	1.0 (0.7)	0.4	1.5	4.1 (3.4)	1.4	6.7
		<i>RE</i>	-0.5 (1.1)	-1.0	0.0	2.6 (3.8)	0.0	4.2
<i>SScore<sup>+4dB</sup> (%)</i>		<i>LE</i>	85.0 (11.7)	85.0	90.0	60.0 (26.6)	40.0	85.0
		<i>RE</i>	91.0 (9.6)	90.0	95.0	62.3 (24.0)	48.7	80.0
<i>MCL (dB HL)</i>	<i>LF</i>	<i>LE</i>	81.5 (14.8)	73.3	84.1	80.6 (8.4)	76.4	85.8
		<i>RE</i>	76.5 (13.2)	70.0	80.0	79.1 (7.9)	74.7	84.1



Outcome measure	Freq. Range	Ear	NH			HI		
			Mean (SD)	Q1	Q3	Mean (SD)	Q1	Q3
	HF	LE	79.0 (17.6)	66.6	90.8	82.7 (12.3)	75.8	90.0
		RE	73.8 (17.2)	65.0	80.0	80.3 (9.9)	74.7	87.5
Slope (CU/dB)	LF	LE	0.35 (0.1)	0.3	0.4	0.45 (0.1)	0.3	0.5
		RE	0.36 (0.1)	0.3	0.4	0.48 (0.2)	0.3	0.5
	HF	LE	0.45 (0.1)	0.3	0.4	0.84 (0.5)	0.5	0.9
		RE	0.41 (0.1)	0.3	0.4	0.81 (0.4)	0.5	0.9
DynR (dB HL)	LF	LE	91.5 (16.8)	78.3	97.5	76.7 (15.8)	64.5	88.3
		RE	91.1 (18.8)	79.1	100.0	73.9 (16.0)	61.6	86.8
	HF	LE	77.6 (18.2)	72.5	85.8	50.8 (15.1)	40.6	60.2
		RE	78.6 (17.9)	67.5	90.8	50.7 (15.5)	38.9	60.4
sSTM -3dB (d')	LF	Bin	2.6 (0.6)	2.4	3.0	1.7 (1.3)	0.4	3.0
	HF		1.6 (0.8)	1.1	2.4	0.6 (1.1)	-0.3	1.4
fSTM (dB)	LF	LE	-7.7 (1.8)	-9.0	-7.6	-2.8 (2.1)	-3.5	-0.8
		RE	-5.1 (3.1)	-7.2	-1.6	-1.6 (1.3)	-2.0	-0.6
	HF	LE	-8.0 (2.0)	-8.6	-6.2	-2.6 (2.4)	-3.8	-0.6
		RE	-5.6 (3.6)	-8.6	-2.1	-1.9 (1.5)	-2.0	-1.0
IPD <sub>fmax</sub> (kHz)		Bin	0.76 (0.26)	0.59	0.98	0.69 (0.27)	0.52	0.88
Bin Pitch 20 (%)		Bin	87.5 (25.0)	87.5	100.0	80.7 (30.9)	70.0	100.0
eAUD-HF		LE	10.9 (1.2)	10.2	11.9	7.57 (2.7)	5.3	10.0
FLFT (kHz)		RE	11.7 (1.1)	10.9	12.5	8.12 (2.3)	6.7	10.2
eAUD-N (dB HL)	LF	LE	70.4 (4.5)	68.0	71.5	71.8 (2.6)	70.2	73.2
		RE	69.2 (4.6)	65.2	72.5	72.0 (2.8)	69.6	74.3
	HF	LE	71.1 (2.5)	69.7	72.7	74.7 (3.4)	72.5	76.1
		RE	70.8 (3.6)	70.5	71.7	74.2 (3.1)	72.0	76.2
TMR eAUD (N-T) (dB)	LF	LE	7.5 (3.4)	6.0	7.5	7.7 (4.0)	6.1	10.1
		RE	5.2 (3.3)	4.0	7.6	8.3 (2.7)	6.5	10.3
	HF	LE	13.0 (0.6)	12.7	13.2	7.9 (5.0)	5.0	11.6
		RE	10.7 (3.1)	9.1	10.2	8.1 (5.2)	5.1	10.7
SMR eAUD (N-S) (dB)	LF	LE	19.3 (3.6)	16.5	21.7	19.6 (17.7)	17.7	23.2
		RE	18.8 (4.6)	17.0	21.2	20.0 (5.2)	16.5	23.8
	HF	LE	26.8 (4.5)	27.5	29.0	19.3 (9.5)	12.1	26.3
		RE	27.2 (3.7)	26.2	29.5	19.5 (9.9)	12.0	26.8
BMR (SON0 – SOSpi) (dB)		Bin	16.5 (4.7)	13.5	17.5	14.7 (4.6)	12.2	17.5

*SRT<sub>Q</sub>*: Speech reception threshold in quiet / *Max DS*: Maximum speech discrimination score. // *SRT<sub>N</sub>*: Speech reception threshold in noise / *Score +4*: Sentence recognition score at +4 dB SNR // *MCL*: Most comfortable level / *Slope*: Slope of the loudness function / *DynR*: Dynamic range // *sSTM*: Sensitivity for detecting a spectro-temporally modulated noise at  $20\log(m) = -3$  dB, where *m* is the modulation depth / *fSTM*: Fast version of the STM test (Bernstein et al., 2016) // *IPD<sub>fmax</sub>*: Frequency threshold for detecting an interaural phase difference of 180°.

<i>Outcome measure</i>	<i>Freq. Range</i>	<i>Ear</i>	<i>NH</i>			<i>HI</i>		
			<i>Mean (SD)</i>	<i>Q1</i>	<i>Q3</i>	<i>Mean (SD)</i>	<i>Q1</i>	<i>Q3</i>
<i>Bin pitch: Binaural pitch detection scores for 20 presentations // eAUD-HF: Fixed-level frequency threshold (FLFT) at 80 dB SPL // eAUD-N: Tone detection in TEN noise // TMR: Temporal masking release // SMR: Spectral masking release // BMR: Binaural masking release.</i>								

120

121 **Speech perception in quiet**122 ***Methods***

123 The word recognition score (WRS-4UFC) test was proposed as a systematic and self-  
 124 administered procedure that allows the estimation of supra-threshold deficits in speech  
 125 perception in quiet. The speech material was the same as the one used for standard  
 126 speech audiometry (Dantale I; Elberling, Ludvigsen, & Lyregaard, 1989) in Danish. The  
 127 self-administered procedure consisted of a 4-interval-unforced-choice paradigm  
 128 (4UFC). After the presentation of each word, the target was placed randomly in one of  
 129 four intervals. The other three words were also taken from the Dantale-I corpus. They  
 130 were chosen based on the lowest Levenshtein phonetic distance (Sanders & Chin, 2009)  
 131 from the target. Four lists of 25 words were presented at 40, 30, 20 and 10 dB above the  
 132 individual PTA, in this order. A logistic function was fitted to the results from each  
 133 individual ear and the speech reception threshold (SRT<sub>Q</sub>) and maximum speech  
 134 discrimination score (Max DS) were estimated using *psignifit 4* software (Schütt,  
 135 Harmeling, Macke, & Wichmann, 2016).

136 ***Results and discussion***

137 The HI listeners' SRT<sub>Q</sub> were, on average, 20 dB higher than the ones of the NH group.  
 138 The interquartile range for the HI group was about 19 dB whereas for the NH group it

139 was 3 dB for the left ear (LE) and 11.8 dB for the right ear (RE). The Max DS for both  
140 groups was close to 100%. However, the HI listeners showed larger variability,  
141 especially in the right ear (SD= 6.42).  
142 In the analysis of the test-retest variability, the WRS-4UFC test showed poor to  
143 moderate reliability especially at low levels (PTA + 10 dB; ICC = 0.25). However, at  
144 the higher presentation levels (i.e. individual PTA + 40 dB) the standard error of the  
145 measurement was only 4% (1 word). Regarding clinical applicability, the WRS-4UFC  
146 needs to be compared to traditional speech audiometry to explore the influence of using  
147 closed- vs. open-set and forced- vs. unforced-choice test procedures on the results.

#### 148 **Speech perception in noise**

149 The Hearing in Noise Test (HINT; Nilsson et al., 1994) is an adaptive sentence  
150 recognition test carried out with speech-shaped noise. The following assumptions are  
151 considered in HINT: 1) Speech materials made of meaningful sentences yield a steep  
152 psychometric function; 2) Stationary noise with the same spectral shape as the average  
153 spectrum of the speech material makes the speech reception threshold in noise ( $SRT_N$ )  
154 less dependent of the spectral characteristics of the speaker's voice. Furthermore, the  
155 signal-to-noise ratio (SNR) between the target and masker is better defined across the  
156 frequency range; 3) The  $SRT_N$  is independent of the absolute noise level as long as the  
157 noise level is above the "internal noise" level. Therefore, it is recommended to present  
158 the noise at least 30 dB above the "internal noise". The internal noise is defined as the  
159 sum of the SRT in quiet of the tested listener and the SRT in noise for NH listeners, for  
160 a given speech material (Reinier Plomp, 1986).

161 **Methods**

162 The Danish HINT was used as in Nielsen & Dau (2011) to obtain the  $SRT_N$ .  
163 Additionally, a 20-sentence list was presented at a fixed signal-to-noise ratio of +4 dB  
164 and scored to obtain a sentence recognition score ( $SScore^{+4dB}$ ). The presentation level of  
165 the noise was set between 65 and 85 dB SPL to ensure that the noise was always  
166 presented 30 dB above the individual PTA. Each ear was tested individually. All  
167 participants were tested using the same list with the same ear. However, for the test-  
168 retest reliability study, the list and ear presented were randomized, only using lists 6-10.

169 **Results and discussion**

170 The  $SRT_N$  for NH listeners were, on average, 2 dB higher than the ones reported  
171 Nielsen and Dau (2011). However, this might be explained by the fact that they used  
172 diotic presentation which can lead to a 1.5 dB improvement (Plomp & Mimpen, 1979).  
173 The results also showed a lower  $SRT_N$  (1.5 dB) and higher  $SScore^{+4dB}$  (4%) for the right  
174 ear in both groups of listeners. According to Nielsen and Dau (2011), there was a  
175 significant main effect of test list. Such differences are seen mainly for lists 1-4, which  
176 were the lists used here. Therefore, the observed interaural difference can be ascribed to  
177 a list effect.

178 The ICC values ( $SRT_N$ : ICC= 0.61;  $SScore^{+4dB}$ : ICC = 0.57) indicated only moderate  
179 reliability of the HINT. The  $SRT_N$  showed an SEM = 1.02 dB, which is below the step  
180 size of the test (2 dB). The  $SScore^{+4dB}$  showed an SEM value of 7.94%, which  
181 corresponds to an error in one of the sentences.

182 The use of speech-in-noise tests can be a useful tool for the characterization of the  
183 listener's hearing deficits that can be performed under different conditions, including  
184 monaural, binaural, unaided and aided stimuli presentations. While here the tests were

185 performed monaurally and unaided, a binaural condition as well as at least one aided  
186 measure (i.e., with hearing aids), could also be included in clinical practice.

### 187 **Loudness perception**

188 Loudness perception can substantially differ between NH and HI listeners and has been  
189 connected to the peripheral non-linearity (e.g. Jürgens, Kollmeier, Brand, & Ewert,  
190 2011). While the growth of loudness shows a non-linear behaviour in a healthy ear, the  
191 results from HI listeners suggest that loudness perception becomes linear when outer-  
192 hair cell (OHC) function is affected (e.g. Moore, 2007). Besides, the possibilities of  
193 characterizing hearing deficits, loudness function can be used for fitting hearing aids  
194 (e.g., Oetting, Hohmann, Appell, Kollmeier, & Ewert, 2018). Adaptive categorical  
195 loudness scaling (ACALOS; Brand & Hohmann, 2002) is the reference method for the  
196 current standard (ISO 16832, 2006) for loudness measurements.

### 197 **Methods**

198 According to the ACALOS method, narrow-band noises were presented sequentially,  
199 and the participant had to judge the perceived loudness using a 13-category scale  
200 ranging from “not heard” to “extremely loud”. The raw results, which correspond to  
201 categorical units (CU) spanned between 0 and 50, were fitted to a model of loudness as  
202 described in Oetting, Brand, & Ewert (2014). The outcome measures of the ACALOS  
203 presented here are the most comfortable level (MCL), the slope of the loudness function  
204 (Slope), and the dynamic range (DynR) defined as the difference between  
205 uncomfortable level (50 CU) and the hearing threshold (0.5 CU). Low-frequency (LF)  
206 average corresponds to frequencies below 1.5 kHz, high-frequency (HF) average  
207 correspond to frequencies above 1.5 kHz

**208 Results and discussion**

209 The average MCL estimate ranged between 73 and 82 dB HL in both groups and for  
210 both frequency ranges. The average slope of the loudness growth was slightly steeper  
211 for the HI listeners in the low-frequency range (0.45 CU/dB HI vs. 0.35 CU/dB NH)  
212 and substantially steeper in the high-frequency range (0.8 vs 0.45). The average  
213 dynamic range was between 80 and 90 dB HL for the NH listeners, and smaller for the  
214 HI listeners, especially at high frequencies (50.8 dB).

215 Regarding the test-retest reliability, ACALOS showed an excellent reliability for  
216 estimating the hearing thresholds (ICC = 0.94; SEM = 4.5 dB), good reliability for  
217 estimating the MCL (ICC = 0.68, SEM = 6.5 dB) and very good reliability for  
218 estimating the slope (ICC = 0.82; SE M = 0.07 CU/dB). Overall, these results supported  
219 the inclusion of ACALOS in a clinical test battery, as it provides several outcomes  
220 (hearing thresholds, growth of loudness, MCL and dynamic range). ACALOS also  
221 showed a high time efficiency (around 10 min. per ear).

**222 Spectro-temporal modulation sensitivity**

223 A speech signal can be decomposed into spectral and temporal modulations. While  
224 speech-in-noise perception assessment leads to some confounds due to the variety of  
225 speech corpora, noise maskers, and test procedures that can all affect the results, the  
226 assessment of the sensitivity of simpler sounds might be of interest for characterizing a  
227 listener's spectro-temporal processing abilities. Bernstein et al. (2013) showed  
228 significant differences between NH and HI listeners for detecting STM in random noise.  
229 These differences corresponded to specific conditions that were also useful for the  
230 prediction of speech-in-noise performance in the same listeners. Lately, the assessment  
231 of STM sensitivity in these specific conditions gained an increasing interest due to its

232 potential for predicting speech intelligibility (Bernstein et al., 2016; Gallun et al., 2018;  
233 Zaar, Simonsen, Bherens, Dau, & Laugesen, 2019) and for assessing cochlear-implant  
234 candidacy (Choi et al., 2016). Here, STM sensitivity was assessed using a new test  
235 paradigm that may be more suitable for a clinical implementation. The test was  
236 performed in two conditions: a low-frequency condition (similar to the one previously  
237 used in Bernstein et al., 2016) and a high-frequency condition (Mehraei, Gallun, Leek,  
238 & Bernstein, 2014).

### 239 ***Methods***

240 The stimuli were similar to those of Bernstein et al. (2016) and Mehraei et al. (2014),  
241 but a different presentation paradigm was employed. A sequence of four noises was  
242 presented in each trial. The first and third stimulus always contained unmodulated  
243 noise, whereas the second and fourth stimuli could be either modulated or unmodulated.  
244 After the sequence was presented, the listener had to respond whether the four sounds  
245 were different ('yes') or the same ('no'). Two procedures involving catch trials were  
246 evaluated. The first test (sSTM -3 dB) was a screening test consisting of 10 stimuli  
247 modulated at -3 dB level and five unmodulated ones presented in random order. The  
248 outcome measure was the listener's sensitivity ( $d'$ ) in the task. The second test (fSTM)  
249 tracked the 80% threshold using the single-interval adjusted matrix (SIAM; Kaernbach,  
250 1990) paradigm.

### 251 ***Results and discussion***

252 The screening STM test shows the sensitivity in terms of  $d'$ , where the maximum value  
253 is  $d' = 3$ , i.e. 10 modulated and 5 unmodulated stimuli correctly detected. In the  
254 hypothetical case when all the catch trials are detected, the lowest  $d'$  value can be -0.3.  
255 The NH listeners showed a high sensitivity in the low-frequency condition ( $d' = 2.6$ )

256 and a somewhat lower sensitivity in the high-frequency condition ( $d' = 1.63$ )  
257 corresponding to 65% correct responses. The HI listeners showed a higher variability  
258 and a lower sensitivity in the low-frequency condition (~70% correct) and substantially  
259 lower sensitivity in the high-frequency condition (0-50% correct responses). The  
260 threshold tracking procedure (fSTM) showed results between -9 and -6 dB in the NH  
261 group, whereas the HI listeners showed thresholds between -3.50 and -0.5 dB. Although  
262 the results of the fSTM low-frequency condition were consistent with Bernstein et al.  
263 (2016), the results in the high-frequency condition showed higher thresholds than the  
264 ones in Mehraei et al. (2014). This can be ascribed to the higher presentation level used  
265 in Mehraei et al. (2014) than in the current test procedure.

266 The fSTM showed an excellent reliability ( $ICC = 0.91$ ;  $SEM = 0.93$  dB) in the LF  
267 condition. However, several HI listeners were not able to complete the procedure for the  
268 HF condition. Overall, the use of the SIAM tracking procedure allowed us to obtain  
269 accurate thresholds, although additional repetitions were required, especially in the HF  
270 condition. This might be because the psychometric function for detecting the stimulus  
271 can be shallower in this condition or because the 100% detection could not be reached  
272 even in the fully-modulated trials. Therefore, a Bayesian procedure being able to  
273 estimate the threshold and slope of the psychometric function, such as the Bayes Fisher  
274 information gain (FIG; Remus & Collins, 2008), might be more suitable for this type of  
275 test.

### 276 **Binaural processing abilities**

277 Binaural hearing is useful for sound localization and the segregation of complex sounds  
278 (Darwin, 1997). Interaural differences in level or timing are processed for spatial  
279 hearing purposes in the auditory system. With hearing loss, the neural signal at the



280 output of the cochlea can be degraded which may lead to reduced binaural abilities  
281 typically connected to temporal fine structure (TFS) processing. Based on a method  
282 estimating the upper-frequency limit for detecting an interaural phase difference (IPD)  
283 of  $180^\circ$  ( $IPD_{f_{max}}$ ; Neher, Laugesen, Søggaard Jensen, & Kragelund, 2011; Ross,  
284 Tremblay, & Picton, 2007; Sébastien Santurette & Dau, 2012b), Füllgrabe, Harland,  
285 Sęk, & Moore, (2017) recently proposed a refined test as a feasible way to evaluate TFS  
286 sensitivity. This paradigm was used in recent research that suggested that  $IPD_{f_{max}}$  might  
287 be related to non-auditory factors (Strelcyk, Zahorik, Shehorn, Patro, & Derleth, 2019)  
288 and affected by factors beyond hearing loss, such as musical training (Bianchi, Carney,  
289 Dau, & Santurette, 2019). Therefore, the  $IPD_{f_{max}}$  might be a task that requires auditory  
290 and non-auditory processing abilities beyond TFS sensitivity.

291 In contrast, binaural pitch detection assesses binaural processing abilities in a different  
292 manner. This test requires the detection of pitch contours embedded in noise, which are  
293 diotically or dichotically evoked. While the diotic condition can be resolved  
294 monaurally, the dichotic condition requires the binaural processing abilities to be  
295 sufficiently intact to detect the contour. Previous studies showed that some listeners  
296 were unable to detect binaural pitch, regardless of the audiometric configuration  
297 (Sanchez-Lopez et al., 2018; Santurette & Dau, 2012). Therefore, it was of interest to  
298 compare the results of these two binaural processing tests.

### 299 ***Methods***

300 The maximum frequency for detecting an IPD of  $180^\circ$  with pure-tones was obtained  
301 using a 2-AFC tracking procedure similar to the one used in Füllgrabe et al. (2017).  
302 The frequency threshold ( $IPD_{f_{max}}$ ) was obtained from the average of two runs.

303 Binaural pitch detection scores were obtained using a clinical implementation of the test  
304 proposed by Santurette & Dau, (2012). A 3-minute sequence of noise was presented  
305 bilaterally. Ten diotic and 10 dichotic pitch contours, embedded in the noise, had to be  
306 detected by the listener. The tones forming the pitch contours were generated by adding  
307 frequency-specific IPDs to the presented noise (Cramer & Huggins, 1958). The  
308 outcome measure of the binaural pitch test was the percentage score averaged across  
309 two repetitions (BP20).

### 310 ***Results and discussion***

311 The listeners in the NH and HI groups showed  $IPD_{fmax}$  thresholds around 700 Hz with a  
312 standard deviation (~270 Hz) and interquartile range (~370 Hz) similarly in both  
313 groups. These results are in line with the ones reported in Füllgrabe & Moore (2017).  
314 The  $IPD_{fmax}$  test showed excellent reliability (ICC = 0.95; SEM = 65.4 Hz), and the  
315 median time needed for two repetitions was 10 minutes. This suggests that  $IPD_{fmax}$  is a  
316 reliable measure of binaural processing abilities that can reveal substantial variability  
317 among both NH and HI listeners, which is valuable for highlighting individual  
318 differences among patients.

319 The overall results from the binaural pitch test for the NH listeners showed >87.5%  
320 correct detection, whereas the HI listeners' results showed a higher variability with an  
321 interquartile range from 70-100%. The test showed excellent reliability (ICC = 0.98;  
322 SEM = 4%). Listeners reported a positive experience due to the test being short and  
323 easy to understand.

### 324 **Extended audiometry in noise (eAUD)**

325 The extended audiometry in noise (eAUD) is a tone detection test intended to assess  
326 different aspects of auditory processing by means of a task similar to pure-tone

327 audiometry. The tone is presented either in noise or in quiet and the listener has to  
328 indicate whether the tone was perceived or not. The aspects of auditory processing  
329 assessed here are 1) high-frequency audibility, 2) spectral and temporal resolution and  
330 3) binaural processing abilities.

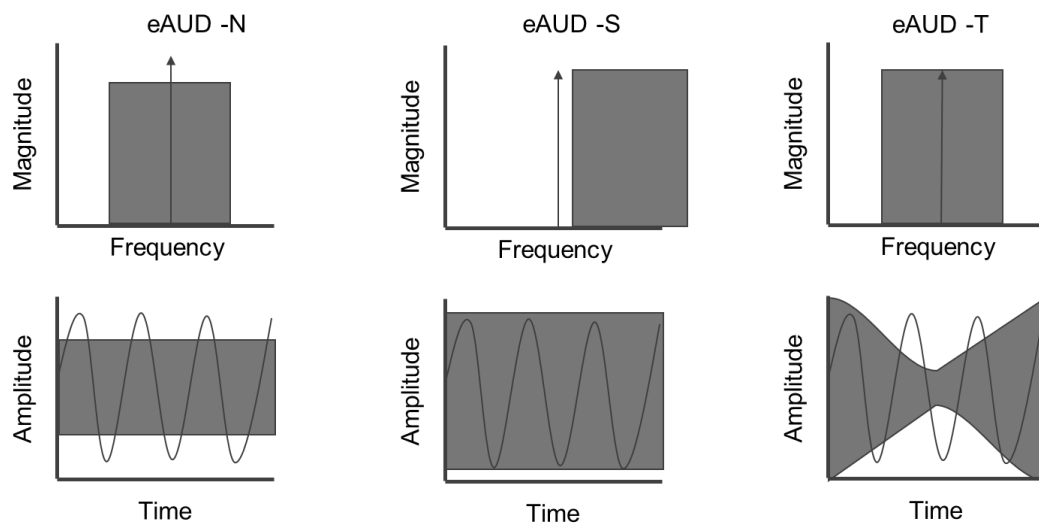
### 331 ***High-frequency audibility***

332 Recently, elevated thresholds at high frequencies (>8 kHz) have been linked to the  
333 concept of “hidden hearing loss” and synaptopathy (Liberman, Epstein, Cleveland,  
334 Wang, & Maison, 2016). However, the measurement of audiometric thresholds above 8  
335 kHz is not part of the current clinical practice. The fixed-level frequency threshold  
336 (FLFT) has been proposed as a quick and efficient alternative to high-frequency  
337 audiometry (Rieke et al., 2017). The test is based on the detection of a tone presented at  
338 a fixed level. The frequency of the tone is varied towards high frequencies and the  
339 maximum audible frequency at the given level is estimated in an adaptive procedure.  
340 Here, a modified version of FLFT was used as the extended audiometry at high  
341 frequencies (eAUD-HF).

### 342 ***Spectro-temporal resolution***

343 Frequency and temporal resolution are aspects of hearing that are fundamental for the  
344 analysis of perceived sounds. While NH listeners exhibit a frequency selectivity on the  
345 order of one third of an octave (from Glasberg & Moore, (1990), HI listeners have  
346 typically broader auditory filters leading to impaired frequency selectivity (e.g. Moore,  
347 2007). Temporal resolution can be characterized by the ability to “listen in the dips”  
348 when the background noise is fluctuating based on the so-called masking release  
349 (Festen & Plomp, 1990). Schorn & Zwicker, (1990) proposed an elaborated technique  
350 for assessing both spectral and temporal resolution using two tests: 1) Psychoacoustical

351 tuning curves and 2) temporal resolution curves. In both cases, the task consists of  
 352 detecting a pure tone that is masked by noise or another tone while the spectral or  
 353 temporal characteristics of the masker are varied. Later, Larsby & Arlinger (1998)  
 354 proposed a similar paradigm, the F-T test, which was successfully tested in HI listeners  
 355 (van Esch & Dreschler, 2011). Here, the spectro-temporal resolution was assessed using  
 356 a new test. This test is a tone-in-noise detection task consisting of three conditions as  
 357 sketched in Figure 2.



358

359 Figure 2: Sketch of the conditions of the spectro-temporal resolution measures of the  
 360 extended audiometry in noise (eAUD). The top panel shows the spectrum of the noise  
 361 and target pure-tone (delta), the bottom panel shows both signals in the time domain.  
 362 Left panel: Tone in noise condition (eAUD-N). Middle panel: Spectral condition  
 363 (eAUD-S). Right panel: Temporal condition (eAUD-T).

364

- 365 1) eAUD-N: The tone is embedded in a 1-octave-wide threshold equalizing noise  
366 (TEN; Moore, 2001). Because of the properties of the TEN, the tone detection  
367 threshold is comparable to the level of the noise in dB HL.
- 368 2) eAUD-S: The tone is embedded in a TEN that has been shifted up in frequency.  
369 In the spectral domain, this yields spectral unmasking of the tone, so the detection  
370 threshold is lower than in eAUD-N.
- 371 3) eAUD-T: The tone is embedded in a temporally-modulated noise with the same  
372 spectral properties as the one in eAUD-N. In the temporal domain, the  
373 modulations of the noise yield temporal unmasking, so the tone can be detected  
374 in the dips.

375 The outcome measures were focused on the temporal and spectral benefits expected in  
376 the eAUD-S and eAUD-T conditions compared to the eAUD-N condition. While in the  
377 noise condition (eAUD-N) the threshold is expected to be approximately at the level of  
378 the noise, in the temporal and spectral conditions the thresholds should be lower  
379 showing temporal masking release (TMR) and spectral masking release (SMR).

### 380 ***Binaural Masking Release***

381 Besides the binaural tests presented previously, another approach for evaluating the  
382 binaural processing abilities is assessing binaural masking release (Durlach, 1963),  
383 which has been used in several studies (Neher, 2017; Strelcyk & Dau, 2009) and  
384 implemented in some commercial audiometers (Brown & Musiek, 2013). In this  
385 paradigm, a tone-in-noise stimulus is presented in two conditions: (1) a diotic condition  
386 where the tone is in phase in the two ears, and (2) a dichotic condition where the tone is  
387 in antiphase in the two ears. The difference between the two yields the benefit for tone  
388 detection due to binaural processing, the so-called binaural masking release (BMR).

**389 Methods**

390 The procedure used here was a yes/no task using a SIAM procedure (Kaernbach,  
391 1990). As in traditional up-down procedures, the target can be presented in a given trial  
392 or not. If the target was detected, the target-presentation level is decreased according to  
393 a given step size; if it was not detected, the level is increased. If the stimulus was not  
394 presented (catch trial) but the listener provided a positive response, the level is  
395 decreased compared to the previous trial.

396 The target stimulus for all the conditions tested here was a warble tone. For each run,  
397 the first two reversals were discarded, and the threshold of each trial was calculated as  
398 the average of the four subsequent reversals. The low-frequency condition (LF)  
399 corresponds to the detection of a 0.5-kHz warble tone, whereas the high-frequency (HF)  
400 condition corresponded to a 2-kHz warble tone. The final threshold was calculated as  
401 the mean threshold of two repetitions. The outcome measures of the eAUD are 1) the  
402 high-frequency threshold (eAUD-HF), 2) the tone-in-noise threshold (eAUD-N), 3) the  
403 SMR, 4) the TMR, and 5) the BMR.

**404 Results and discussion**

405 The maximum frequency threshold for a tone presented at 80 dB SPL (eAUD-HF) was  
406 11 kHz for the NH listeners and 8 kHz for the HI listeners. The HI group showed larger  
407 variability compared to the NH group (interquartile range: 6 kHz vs. 10 kHz). In  
408 contrast, the eAUD-N condition showed a larger variance for the NH group (SD = 4.5  
409 dB HL) at low frequencies. The detection thresholds were in line with previous work  
410 with thresholds close to the noise presentation level (70 dB HL) (Vinay, Hansen, Raen,  
411 & Moore, 2017). The TMR shown by the NH group was larger at high frequencies (10  
412 dB) than at low frequencies (7 dB). The HI group showed, on average, similar TMR

413 only at low frequencies. The SMR shown by the NH listeners was 19 dB for low  
414 frequencies and 26 dB for high frequencies. In contrast, for the HI listeners, the SMR  
415 was 7 dB lower only in the high-frequency condition. The BMR shown by both groups  
416 was around 15 dB, as expected from previous studies (Durlach, 1963).

417 The reliability of the eAUD was moderate for most of the conditions ( $ICC < 0.75$ ). The  
418 eAUD-HF test showed very good reliability ( $ICC = 0.89$ ;  $SEM = 495$  Hz), and the  
419 eAUD-S at low frequencies showed good reliability ( $ICC = 0.85$ ;  $SEM = 1.78$  dB). The  
420 masking release estimates showed good reliability only for the high-frequency  
421 condition. The reason for this might be that masking release is a differential measure,  
422 and the cumulative error is, therefore, higher than that of each individual measure. The  
423 reduced reliability can be explained to some extent by the method used. To have a  
424 similar procedure as in pure-tone audiometry, the parameters of the SIAM tracking  
425 procedure were set accordingly. However, this made the test challenging and the  
426 listeners consistently missed several catch trials. Thus, extra trials were required to  
427 improve measurement accuracy. However, the standard error of the measurement was in  
428 most cases larger than the final step size (2 dB). As in the case of the fSTM, a different  
429 procedure, such as Bayesian adaptive methods, might increase measurement reliability.

### 430 **Exploratory analysis**

431 The collection of tests included in the test battery was intended to explore different and  
432 potentially independent aspects of hearing to obtain an auditory profile with controlled  
433 interrelations among the tests. A factor analysis performed in the HEARCOM study  
434 (Vlaming et al., 2011) based on data from 72 HI subjects revealed auditory dimensions:  
435 1) high-frequency processing, 2) audibility, 3) low-frequency processing and 4)  
436 recruitment. In the current study, the results of the behavioural tests were analysed

437 further in order to explore possible interrelations between the various outcome  
438 measures.

### 439 **Methods**

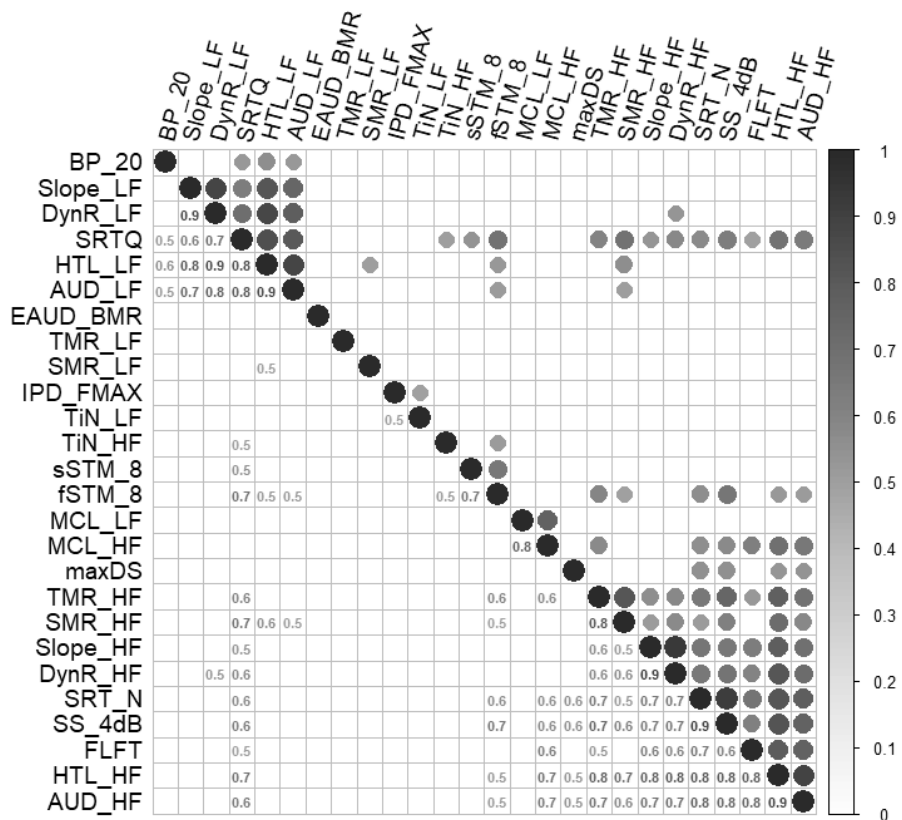
440 First, the data were pre-processed as in Sanchez Lopez et al. (2018) to reduce the  
441 number of variables. The outcome variables of the frequency-specific tests were divided  
442 into LF ( $\leq 1$  kHz) and HF ( $> 1$  kHz) variables. This decision was supported by a  
443 correlation analysis performed on the complete set of outcome variables, where the  
444 outcomes corresponding to 2, 4 and 6 kHz as well as the ones corresponding to 0.25, 0.5  
445 and 1 kHz were highly intercorrelated. For the tests performed monaurally, the mean of  
446 the two ears was taken as the resulting outcome variable. The resulting dataset (BEAR3  
447 dataset) contained 26 variables, divided into six groups corresponding to the six aspects  
448 of auditory processing considered here. The exploratory analysis consisted of a  
449 correlation analysis using Spearman correlations and factor analysis. The factor analysis  
450 was performed using an orthogonal rotation (“varimax”) and the method of maximum  
451 likelihood. The number of components was chosen using parallel analysis, the resulting  
452 number of components was four.

### 453 **Results**

454 Figure 3 shows the results from the correlation analysis performed on the BEAR3  
455 dataset. For convenience, the absolute value of the correlation was used when  
456 visualizing the data to show the strength of the correlation. The circles on the left-hand  
457 side of the figure depict significant correlations ( $p < 0.00001$ ), and the correlation  
458 values are presented on the left-hand side of the figure. Two groups of correlated  
459 variables can be observed. The upper-left corner shows variables related to LF  
460 processing (dynamic range, the slope of the loudness function, and hearing thresholds)



461 and speech intelligibility in quiet. The bottom-right corner shows a larger group of  
 462 correlated variables including HF processing, speech intelligibility in noise, and  
 463 spectro-temporal resolution at high frequencies. The variables that are not significantly  
 464 interrelated are shown in the middle part of Figure 3, including the three variables  
 465 related to binaural processing abilities (IPD<sub>fmax</sub>, BP<sub>20</sub> and BMR) which were not  
 466 significantly correlated to each other. The speech reception threshold in quiet (SRT<sub>Q</sub>)  
 467 and the STM detection were correlated to various variables such as tone-in-noise  
 468 detection, HF spectro-temporal resolution, LF hearing thresholds and speech-in-noise  
 469 perception.



470

471 Figure 3: Correlation plot of the data set BEAR3. The upper part shows the significantly  
 472 correlated variables as coloured circles. The lower panel shows the numeric correlation  
 473 value.

474

475 The four factors resulting from the factor analysis showed 63% of explained cumulative  
 476 variance. The variables with higher loadings ( $> 0.65$ ) for each of the factors are shown  
 477 in Table 3.

478

479

480

481

482

483

484 Table 3: Variables correlated to the four latent orthogonal factors resulting from the  
 485 factor analysis with the method of maximum likelihood (ML). Columns are sorted in  
 486 terms of the variance explained by each factor.

	<i>ML2(19%)</i>	<i>ML1(18%)</i>	<i>ML3(14%)</i>	<i>ML4(12%)</i>
<i>HTL_LF</i>	0.93			
<i>DYNR_LF</i>	-0.90			
<i>AUD_LF</i>	0.82			
<i>SLOPE_LF</i>	0.81			
<i>SRTQ</i>	0.67			
<i>DYNR_HF</i>		-0.93		
<i>SLOPE_HF</i>		0.82		
<i>HTL_HF</i>		0.79		
<i>AUD_HF</i>		0.73		
<i>MCL_HF</i>			0.92	
<i>MCL_LF</i>			0.85	
<i>SRT_N</i>				0.77
<i>SSCORE_4DB</i>				-0.78

487

488 The first factor, in terms of the amount of variance explained (19%), was associated  
 489 with LF loudness perception and speech intelligibility in quiet, whereas the second  
 490 factor (18% of variance explained) was associated with HF loudness perception.

491 Despite loudness perception being associated with the first and second factor, the MCL  
492 was associated, both at high and low frequencies, with the third factor, while the fourth  
493 factor was associated with speech intelligibility in noise.

#### 494 **General discussion**

495 The first goal of the present study was to collect data of a heterogeneous population of  
496 HI listeners, reflecting their hearing abilities in different aspects of auditory processing.  
497 The current study was motivated by the need for a new dataset to refine the data-driven  
498 approach for auditory profiling. The dataset should contain a representative population  
499 of listeners and outcome measures (Sanchez-Lopez et al., 2018) to allow a refined  
500 definition of the two types of auditory distortions and to identify subgroups of listeners  
501 with clinical relevance. To refine the data-driven auditory profiling, the BEAR3 dataset  
502 fulfils all the requirements discussed in Sanchez-Lopez et al. (2018). Other datasets  
503 containing a large number of listeners (Gieseler et al., 2017; Rönnerberg et al., 2016) or  
504 physiological measures (Kamerer, Kopun, Fultz, Neely, & Rasetshwane, 2019) could  
505 also be interesting for complementing the auditory profiling beyond auditory perceptual  
506 measures

#### 507 ***Relationships across different aspects of auditory processing***

508 The proposed test battery considers outcomes divided into six dimensions of auditory  
509 processing. One of the objectives of the study was to investigate the interrelations of  
510 different dimensions and measures. The present analysis showed two interesting  
511 findings. First, the correlation analysis shows two clusters of variables related to either  
512 low- or high-frequency audiometric thresholds. Speech-in-noise perception was  
513 associated with high-frequency sensitivity loss, temporal, and spectral masking release  
514 whereas speech-in-quiet was correlated with both low- and high-frequency hearing loss.

515 Several outcomes were not interrelated, especially the outcomes associated with  
516 binaural processing abilities. Second, factor analysis yielded latent factors related to  
517 low- and high-frequency processing, most comfortable level and speech in noise.  
518 Vlaming et al. (2011) showed four dimensions in the factor analysis of the HEARCOM  
519 project data corresponding to high-and low-frequency spectro-temporal processing,  
520 MCL and recruitment. In contrast, the current study showed that the slopes of the  
521 loudness growth, both at low and high frequencies, were not interrelated and  
522 contributed to the first and second latent factors. Additionally, the speech-in-noise test  
523 performed in HEARCOM was associated with the low-frequency processing, whereas,  
524 in the present study, speech-in-noise dominates the fourth factor and is significantly  
525 correlated with high frequencies. The reason for this discrepancy might be the use of  
526 different types of noise and test procedures in the two studies.

527 Overall, the data of the present study seem to be dominated by the audiometric profiles,  
528 with low- and high-frequency processing reflecting the main sources of variability in  
529 the data. However, binaural processing abilities, loudness perception and speech-in-  
530 noise outcomes showed a greater contribution to the variability of the supra-threshold  
531 measures than spectro-temporal processing outcomes.

### 532 *Towards clinical feasibility of the tests*

533 The test-retest reliability of the test battery was investigated based on the results of a  
534 subset of listeners who participated 2-5 months after the first visit. The analysis was  
535 based on the ICC and the SEM. Some of the tests, such as  $IPD_{fmax}$ , binaural pitch and  
536 FLFT showed good to excellent test-retest reliability with all ICC values above 0.9,  
537 while other tests, such as the extended audiometry in noise and speech intelligibility in  
538 quiet, showed poor reliability.

539 The selected tests were conducted in two sessions and the total time was, on average,  
540 three hours. In realistic clinical setups, a subset of tests with high reliability and a  
541 reasonably low difficulty would need to be prioritized. For a clinical version of the test  
542 battery, other tracking procedures such as Bayesian Functional information (Remus &  
543 Collins, 2008) might be adopted to improve the reliability and time-efficiency in some  
544 tasks such as STM and tone detection in noise. Moreover, if time-efficiency is crucial,  
545 testing some aspects of auditory processing out of the clinic, as other proposed test  
546 batteries for auditory research (Gallun et al., 2018), might be a solution for completing  
547 the patient's hearing profile.

548 A clinical test battery with the subset of tests that showed a good or excellent test-retest  
549 reliability should be evaluated in a large scale study. This should include several aspects  
550 of auditory processing and provide detailed information on the supra-threshold deficits  
551 of the patient. The tests that showed potential for the clinical implementation were  
552 ACALOS, HINT, fSTM, BP and IPD<sub>fmax</sub>. Such a test battery could serve to identify  
553 clinically relevant subset of patients (auditory profiles) that may benefit from specific  
554 types of hearing rehabilitation towards a “stratified approach” (Lonergan et al., 2017)  
555 for audiology practice.

## 556 **Conclusion**

557 The analysis of the data showed that a reduced BEAR test battery has the potential for  
558 clinical implementation, providing relevant and reliable information reflecting several  
559 auditory domains. The proposed test battery showed good reliability, was reasonably  
560 time-efficient and easy to perform. The implementation of a clinical version of the test  
561 battery is publicly available and can be evaluated in future research, e.g. in a larger field  
562 study to further refine the auditory profiling approach. Moreover, the current data will

563 be re-analysed in a continuation study to better define the auditory profiles proposed in  
564 the data-driven approach and the two types of auditory distortions.

#### 565 **Declaration of interest**

566 This work was supported by Innovation Fund Denmark Grand Solutions 5164-00011B  
567 (Better hEARing Rehabilitation project), Oticon, GN Resound, Widex and other partners  
568 (Aalborg University, University of Southern Denmark, the Technical University of  
569 Denmark, Force, Aalborg, Odense and Copenhagen University Hospitals). No potential  
570 conflict of interest was reported by the authors. The data that support the findings of this  
571 study are openly available in Zenodo at <http://doi.org/10.5281/zenodo.3459579>. A  
572 clinical implementation of the test battery is publicly available at  
573 <https://bitbucket.org/hea-dtu/bear-test-battery/src/master/>.

#### 574 **Acknowledgements**

575 We thank the staff from OUH, BBH and HEA, especially JH Schmidt, SS Houmøller, E  
576 Kjærboel, RS Sørensen, and the student helpers from the MSc of Audiology at SDU. The  
577 funding and collaboration of all BEAR partners are sincerely acknowledged. We also  
578 want to show our gratitude to all the participants in the study.

579

#### 580 **References**

- 581 Bernstein, J. G. W., Danielsson, H., Hällgren, M., Stenfelt, S., Rönnberg, J., & Lunner,  
582 T. (2016). Spectrotemporal Modulation Sensitivity as a Predictor of Speech-  
583 Reception Performance in Noise With Hearing Aids. *Trends in Hearing*, 20, 1–17.  
584 <https://doi.org/10.1177/2331216516670387>
- 585 Bernstein, J. G. W., Mehraei, G., Shamma, S., Gallun, F. J., Theodoroff, S. M., & Leek,

- 586 M. R. (2013). Spectrotemporal modulation sensitivity as a predictor of speech  
587 intelligibility for hearing-impaired listeners. *Journal of the American Academy of*  
588 *Audiology*, 24(4), 293–306. <https://doi.org/10.3766/jaaa.24.4.5>
- 589 Bianchi, F., Carney, L. H., Dau, T., & Santurette, S. (2019). Effects of Musical Training  
590 and Hearing Loss on Fundamental Frequency Discrimination and Temporal Fine  
591 Structure Processing: Psychophysics and Modeling. *Journal of the Association for*  
592 *Research in Otolaryngology*, 20(3), 263–277. [https://doi.org/10.1007/s10162-018-](https://doi.org/10.1007/s10162-018-00710-2)  
593 00710-2
- 594 Bisgaard, N., Vlaming, M. S. M. G., & Dahlquist, M. (2010). Standard Audiograms for  
595 the IEC 60118-15 Measurement Procedure. *Trends in Amplification*, 14(2), 113–  
596 120. <https://doi.org/10.1177/1084713810379609>
- 597 Brand, T., & Hohmann, V. (2002). An adaptive procedure for categorical loudness  
598 scaling. *The Journal of the Acoustical Society of America*, 112(4), 1597–1604.  
599 <https://doi.org/10.1121/1.1502902>
- 600 Brown, M., & Musiek, F. (2013). Pathways: The Fundamentals of Masking Level  
601 Differences for Assessing Auditory Function. *The Hearing Journal*, 66(1), 16.  
602 <https://doi.org/10.1097/01.HJ.0000425772.41884.1d>
- 603 Brungart, D. S., Sheffield, B. M., & Kubli, L. R. (2014). Development of a test battery  
604 for evaluating speech perception in complex listening environments. *The Journal*  
605 *of the Acoustical Society of America*, 136(2), 777–790.  
606 <https://doi.org/10.1121/1.4887440>
- 607 Choi, J. E., Hong, S. H., Won, J. H., Park, H.-S., Cho, Y. S., Chung, W.-H., ... Moon, I.  
608 J. (2016). Evaluation of Cochlear Implant Candidates using a Non-linguistic  
609 Spectrotemporal Modulation Detection Test. *Scientific Reports*, 6(April), 35235.  
610 <https://doi.org/10.1038/srep35235>
- 611 Cramer, E. M., & Huggins, W. H. (1958). Creation of Pitch through Binaural  
612 Interaction. *The Journal of the Acoustical Society of America*, 30(5), 413–417.  
613 <https://doi.org/10.1121/1.1909628>

- 614 Darwin, C. J. (1997). Auditory grouping. *Trends in Cognitive Sciences*, *1*(9), 327–333.  
615 [https://doi.org/10.1016/S1364-6613\(97\)01097-8](https://doi.org/10.1016/S1364-6613(97)01097-8)
- 616 Desloge, J. G., Reed, C. M., Braida, L. D., Perez, Z. D., & D'Aquila, L. A. (2017).  
617 Masking release for hearing-impaired listeners: The effect of increased audibility  
618 through reduction of amplitude variability. *The Journal of the Acoustical Society of*  
619 *America*, *141*(6), 4452–4465. <https://doi.org/10.1121/1.4985186>
- 620 Durlach, N. I. (1963). Equalization and Cancellation Theory of Binaural Masking-Level  
621 Differences. *The Journal of the Acoustical Society of America*, *35*(8), 1206–1218.  
622 <https://doi.org/10.1121/1.1918675>
- 623 Elberling, C., Ludvigsen, C., & Lyregaard, P. E. (1989). DANTALE: a new Danish  
624 speech material. *Scandinavian Audiology*, *18*(3), 169–175.  
625 <https://doi.org/10.3109/01050398909070742>
- 626 Ewert, S. (2013). AFC - A modular framework for running psychoacoustic experiments  
627 and computational perception models. *Proceedings of the International Conference*  
628 *on Acoustics AIA-DAGA 2013*, 1326–1329. Retrieved from  
629 [www.aforcedchoice.com](http://www.aforcedchoice.com)
- 630 Festen, J. M., & Plomp, R. (1990). Effects of fluctuating noise and interfering speech on  
631 the speech-reception threshold for impaired and normal hearing. *The Journal of the*  
632 *Acoustical Society of America*, *88*, 1725–1736. <https://doi.org/10.1121/1.400247>
- 633 Füllgrabe, C., Harland, A. J., Şek, A. P., & Moore, B. C. J. (2017). Development of a  
634 method for determining binaural sensitivity to temporal fine structure.  
635 *International Journal of Audiology*, *56*(12), 926–935.  
636 <https://doi.org/10.1080/14992027.2017.1366078>
- 637 Füllgrabe, C., & Moore, B. C. J. (2017). Evaluation of a Method for Determining  
638 Binaural Sensitivity to Temporal Fine Structure (TFS-AF Test) for Older Listeners  
639 With Normal and Impaired Low-Frequency Hearing. *Trends in Hearing*, *21*, 1–14.  
640 <https://doi.org/10.1177/2331216517737230>
- 641 Gallun, F. J., Seitz, A., Eddins, D. A., Molis, M. R., Stavropoulos, T., Jakien, K. M., ...



- 642 Srinivasan, N. (2018). Development and validation of Portable Automated Rapid  
643 Testing (PART) measures for auditory research. *The Journal of the Acoustical*  
644 *Society of America*, 143(3), 050002. <https://doi.org/10.1121/2.0000878>
- 645 Gelfand, S. A. (2009). *Essentials of Audiology* (4th ed., p. 582). 4th ed., p. 582.  
646 <https://doi.org/10.1097/MAO.0b013e3181c99550>
- 647 Gieseler, A., Tahden, M. A. S., Thiel, C. M., Wagener, K. C., Meis, M., & Colonius, H.  
648 (2017). Auditory and Non-Auditory Contributions for Unaided Speech  
649 Recognition in Noise as a Function of Hearing Aid Use. *Frontiers in Psychology*,  
650 8. <https://doi.org/10.3389/fpsyg.2017.00219>
- 651 Glasberg, B. R., & Moore, B. C. (1989). Psychoacoustic abilities of subjects with  
652 unilateral and bilateral cochlear hearing impairments and their relationship to the  
653 ability to understand speech. *Scandinavian Audiology. Supplementum*, 32, 1–25.  
654 Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/2711118>
- 655 Glasberg, B. R., & Moore, B. C. J. (1990). Derivation of auditory filter shapes from  
656 notched-noise data. *Hearing Research*, 47(1–2), 103–138.  
657 [https://doi.org/10.1016/0378-5955\(90\)90170-T](https://doi.org/10.1016/0378-5955(90)90170-T)
- 658 Houtgast, T., & Festen, J. M. (2008). On the auditory and cognitive functions that may  
659 explain an individual's elevation of the speech reception threshold in noise.  
660 *International Journal of Audiology*, 47, 287–295.  
661 <https://doi.org/10.1080/14992020802127109>
- 662 ISO 16832. (2006). *Acoustics - Loudness scaling by means of categories* (Vol. 2006, pp.  
663 1–12). Vol. 2006, pp. 1–12. Retrieved from  
664 [http://www.iso.org/iso/catalogue\\_detail.htm?csnumber=32442](http://www.iso.org/iso/catalogue_detail.htm?csnumber=32442)
- 665 ISO 8253-1. (2010). *Acoustics - Audiometric test methods - Part 1: Pure-tone air and*  
666 *bone conduction audiometry. International Organization for Standardization.*  
667 Retrieved from [http://www.iso.org/iso/catalogue\\_detail.htm?csnumber=43601](http://www.iso.org/iso/catalogue_detail.htm?csnumber=43601)
- 668 ISO 8253-3. (2012). *Acoustics. Audiometric test methods - Part 3: Speech audiometry.*  
669 *International Organization for Standardization.* Retrieved from

- 670 [http://www.iso.org/iso/iso\\_catalogue/catalogue\\_tc/catalogue\\_detail.htm?csnumber](http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber)  
671 =45101
- 672 Johannesen, P. T., Pérez-González, P., Kalluri, S., Blanco, J. L., & Lopez-Poveda, E. A.  
673 (2016). The Influence of Cochlear Mechanical Dysfunction, Temporal Processing  
674 Deficits, and Age on the Intelligibility of Audible Speech in Noise for Hearing-  
675 Impaired Listeners. *Trends in Hearing*, 20, 233121651664105.  
676 <https://doi.org/10.1177/2331216516641055>
- 677 Jürgens, T., Kollmeier, B., Brand, T., & Ewert, S. D. (2011). Assessment of auditory  
678 nonlinearity for listeners with different hearing losses using temporal masking and  
679 categorical loudness scaling. *Hearing Research*, 280(1–2), 177–191.  
680 <https://doi.org/10.1016/j.heares.2011.05.016>
- 681 Kaernbach, C. (1990). A single-interval adjustment-matrix (SIAM) procedure for  
682 unbiased adaptive testing. *The Journal of the Acoustical Society of America*, 88(6),  
683 2645–2655. <https://doi.org/10.1121/1.399985>
- 684 Kamerer, A. M., Kopun, J. G., Fultz, S. E., Neely, S. T., & Rasetshwane, D. M. (2019).  
685 Reliability of Measures Intended to Assess Threshold-Independent Hearing  
686 Disorders. *Ear and Hearing*, 40(6), 1267–1279.  
687 <https://doi.org/10.1097/AUD.0000000000000711>
- 688 Killion, M. C., Niquette, P. a, Gudmundsen, G. I., Revit, L. J., & Banerjee, S. (2004).  
689 Development of a quick speech-in-noise test for measuring signal-to-noise ratio  
690 loss in normal-hearing and hearing-impaired listeners. *The Journal of the*  
691 *Acoustical Society of America*, 116(October 2004), 2395–2405.  
692 <https://doi.org/10.1121/1.2166610>
- 693 Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass  
694 Correlation Coefficients for Reliability Research. *Journal of Chiropractic*  
695 *Medicine*. <https://doi.org/10.1016/j.jcm.2016.02.012>
- 696 Larsby, B., & Arlinger, S. (1998). A method for evaluating temporal, spectral and  
697 combined temporal-spectral resolution of hearing. *Scandinavian Audiology*, 27(1),

- 698 3–12. <https://doi.org/10.1080/010503998419641>
- 699 Liberman, M. C., Epstein, M. J., Cleveland, S. S., Wang, H., & Maison, S. F. (2016).  
700 Toward a Differential Diagnosis of Hidden Hearing Loss in Humans. *PLOS ONE*,  
701 *11*(9), e0162726. <https://doi.org/10.1371/journal.pone.0162726>
- 702 Lonergan, M., Senn, S. J., McNamee, C., Daly, A. K., Sutton, R., Hattersley, A., ...  
703 Pirmohamed, M. (2017). Defining drug response for stratified medicine. *Drug*  
704 *Discovery Today*. <https://doi.org/10.1016/j.drudis.2016.10.016>
- 705 Mehraei, G., Gallun, F. J., Leek, M. R., & Bernstein, J. G. W. (2014). Spectrotemporal  
706 modulation sensitivity for hearing-impaired listeners: Dependence on carrier center  
707 frequency and the relationship to speech intelligibility. *The Journal of the*  
708 *Acoustical Society of America*, *136*(1), 301–316. <https://doi.org/10.1121/1.4881918>
- 709 Moore, B. C. J. (2001). Dead Regions in the Cochlea: Diagnosis, Perceptual  
710 Consequences, and Implications for the Fitting of Hearing Aids. *Trends in*  
711 *Amplification*, *5*(1), 1–34. <https://doi.org/10.1177/108471380100500102>
- 712 Moore, B. C. J. (2007). Cochlear Hearing Loss. In *Cochlear Hearing Loss:*  
713 *Physiological, Psychological and Technical Issues*.  
714 <https://doi.org/10.1002/9780470987889>
- 715 Neher, T. (2017). Characterizing the binaural contribution to speech-in-noise reception  
716 in elderly hearing-impaired listeners. *The Journal of the Acoustical Society of*  
717 *America*. <https://doi.org/10.1121/1.4976327>
- 718 Neher, T., Laugesen, S., Sjøgaard Jensen, N., & Kragelund, L. (2011). Can basic  
719 auditory and cognitive measures predict hearing-impaired listeners' localization  
720 and spatial speech recognition abilities? *The Journal of the Acoustical Society of*  
721 *America*, *130*(3), 1542–1558. <https://doi.org/10.1121/1.3608122>
- 722 Nielsen, J. B., & Dau, T. (2011). The Danish hearing in noise test. *International Journal*  
723 *of Audiology*, *50*(3), 202–208. <https://doi.org/10.3109/14992027.2010.524254>
- 724 Nilsson, M., Soli, S. D., & Sullivan, J. a. (1994). Development of the Hearing In Noise

- 725 Test for the measurement of speech reception thresholds in quiet and in noise. *The*  
726 *Journal of the Acoustical Society of America*, 95(2), 1085–1099.  
727 <https://doi.org/10.1121/1.408469>
- 728 Oetting, D., Brand, T., & Ewert, S. D. (2014). Optimized loudness-function estimation  
729 for categorical loudness scaling data. *Hearing Research*, 316, 16–27.  
730 <https://doi.org/10.1016/j.heares.2014.07.003>
- 731 Oetting, D., Hohmann, V., Appell, J. E., Kollmeier, B., & Ewert, S. D. (2018).  
732 Restoring Perceived Loudness for Listeners With Hearing Loss. *Ear and Hearing*.  
733 <https://doi.org/10.1097/AUD.0000000000000521>
- 734 Oxenham, A. J., & Simonson, A. M. (2009). Masking release for low- and high-pass-  
735 filtered speech in the presence of noise and single-talker interference. *The Journal*  
736 *of the Acoustical Society of America*, 125(1), 457–468.  
737 <https://doi.org/10.1121/1.3021299>
- 738 Plomp, R. (1986). A Signal-to-Noise Ratio Model for the Speech-Reception Threshold  
739 of the Hearing Impaired. *Journal of Speech, Language, and Hearing Research*,  
740 29(2), 146–154. <https://doi.org/10.1044/jshr.2902.146>
- 741 Plomp, R., & Mimpen, A. M. (1979). Improving the Reliability of Testing the Speech  
742 Reception Threshold for Sentences. *International Journal of Audiology*, 18(1), 43–  
743 52. <https://doi.org/10.3109/00206097909072618>
- 744 Remus, J. J., & Collins, L. M. (2008). Comparison of adaptive psychometric procedures  
745 motivated by the Theory of Optimal Experiments: Simulated and experimental  
746 results. *The Journal of the Acoustical Society of America*, 123(1), 315–326.  
747 <https://doi.org/10.1121/1.2816567>
- 748 Rhebergen, K. S., Versfeld, N. J., & Dreschler, W. a. (2006). Extended speech  
749 intelligibility index for the prediction of the speech reception threshold in  
750 fluctuating noise. *The Journal of the Acoustical Society of America*, 120(6), 3988–  
751 3997. <https://doi.org/10.1121/1.2358008>
- 752 Rieke, C. C., Clavier, O. H., Allen, L. V., Anderson, A. P., Brooks, C. A., Fellows, A.

- 753 M., ... Buckey, J. C. (2017). Fixed-Level Frequency Threshold Testing for  
754 Ototoxicity Monitoring. *Ear and Hearing*, 1.  
755 <https://doi.org/10.1097/AUD.0000000000000433>
- 756 Rönnerberg, J., Lunner, T., Ng, E. H. N., Lidestam, B., Zekveld, A. A., Sörqvist, P., ...  
757 Stenfelt, S. (2016). Hearing impairment, cognition and speech understanding:  
758 exploratory factor analyses of a comprehensive test battery for a group of hearing  
759 aid users, the n200 study. *International Journal of Audiology*, 55(11), 623–642.  
760 <https://doi.org/10.1080/14992027.2016.1219775>
- 761 Ross, B., Tremblay, K. L., & Picton, T. W. (2007). Physiological detection of interaural  
762 phase differences. *The Journal of the Acoustical Society of America*, 121(2), 1017–  
763 1027. <https://doi.org/10.1121/1.2404915>
- 764 Sanchez-Lopez, R., Bianchi, F., Fereczkowski, M., Santurette, S., & Dau, T. (2018).  
765 Data-Driven Approach for Auditory Profiling and Characterization of Individual  
766 Hearing Loss. *Trends in Hearing*, 22(1), 233121651880740.  
767 <https://doi.org/10.1177/2331216518807400>
- 768 Sanchez-Lopez, R., Nielsen, S. G., El-Haj-Ali, M., Bianchi, F., Fereczkowski, M.,  
769 Cañete, O., ... Santurette, S. (2019). Data from “Auditory tests for characterizing  
770 hearing deficits: The BEAR test battery.” <https://doi.org/10.5281/zenodo.3459579>
- 771 Sanders, N. C., & Chin, S. B. (2009). Phonological Distance Measures\*. *Journal of*  
772 *Quantitative Linguistics*, 16(1), 96–114.  
773 <https://doi.org/10.1080/09296170802514138>
- 774 Santurette, S., & Dau, T. (2012). Relating binaural pitch perception to the individual  
775 listener’s auditory profile. *The Journal of the Acoustical Society of America*,  
776 131(4), 2968–2986. <https://doi.org/10.1121/1.3689554>
- 777 Saunders, G. H., Field, D. L., & Haggard, M. P. (1992). A clinical test battery for  
778 obscure auditory dysfunction (OAD): Development, selection and use of tests.  
779 *British Journal of Audiology*, 26(1), 33–42.  
780 <https://doi.org/10.3109/03005369209077869>

- 781 Schorn, K., & Zwicker, E. (1990). Frequency selectivity and temporal resolution in  
782 patients with various inner ear disorders. *Audiology : Official Organ of the*  
783 *International Society of Audiology*, 29(1), 8–20.  
784 <https://doi.org/10.3109/00206099009081641>
- 785 Schütt, H. H., Harmeling, S., Macke, J. H., & Wichmann, F. A. (2016). Painfree and  
786 accurate Bayesian estimation of psychometric functions for (potentially)  
787 overdispersed data. *Vision Research*, 122, 105–123.  
788 <https://doi.org/10.1016/j.visres.2016.02.002>
- 789 Stratford, P. W., & Goldsmith, C. H. (1997). Use of the Standard Error as a Reliability  
790 Index of Interest : An Applied Example Using Elbow Flexor Strength Data.  
791 *Physical Therapy*, 77(7), 745–750. <https://doi.org/10.1093/ptj/77.7.745>
- 792 Strelcyk, O., & Dau, T. (2009). Relations between frequency selectivity, temporal fine-  
793 structure processing, and speech reception in impaired hearing. *The Journal of the*  
794 *Acoustical Society of America*, 125, 3328–3345. <https://doi.org/10.1121/1.3097469>
- 795 Strelcyk, O., Zahorik, P., Shehorn, J., Patro, C., & Derleth, R. P. (2019). Sensitivity to  
796 Interaural Phase in Older Hearing-Impaired Listeners Correlates With Nonauditory  
797 Trail Making Scores and With a Spatial Auditory Task of Unrelated Peripheral  
798 Origin. *Trends in Hearing*, 23, 233121651986449.  
799 <https://doi.org/10.1177/2331216519864499>
- 800 Thorup, N., Santurette, S., Jørgensen, S., Kjærboel, E., Dau, T., & Friis, M. (2016).  
801 Auditory profiling and hearing-aid satisfaction in hearing-aid candidates. *Danish*  
802 *Medical Journal*, 63(10). Retrieved from  
803 <http://www.ncbi.nlm.nih.gov/pubmed/27697129>
- 804 van Esch, T. E. M., & Dreschler, W. A. (2011). Measuring spectral and temporal  
805 resolution simultaneously: a comparison between two tests. *International Journal*  
806 *of Audiology*, 50(7), 477–490. <https://doi.org/10.3109/14992027.2011.572083>
- 807 Vinay, Hansen, A. S., Raen, Ø., & Moore, B. C. J. (2017). Reference thresholds for the  
808 TEN(HL) test for people with normal hearing. *International Journal of Audiology*,

- 809           56(9), 672–676. <https://doi.org/10.1080/14992027.2017.1307531>
- 810   Vlaming, M. S. M. G., Kollmeier, B., Dreschler, W. A., Martin, R., Wouters, J., Grover,  
811    B., ... Houtgast, T. (2011). HearCom: Hearing in the Communication Society.  
812    *Acta Acustica United with Acustica*, 97(2), 175–192.  
813    <https://doi.org/10.3813/AAA.918397>
- 814   Wolff, A., Houmøller, S. S., Hougaard, D., Gaihede, M., Hammershøi, D., & Schmidt,  
815    J. (2020). Health-related Quality of Life in Hearing Impaired Danish Adults before  
816    and after Hearing Aid Rehabilitation. *International Journal of Audiology*, (u. rev.).
- 817   Zaar, J., Simonsen, L. B., Bherens, T., Dau, T., & Laugesen, S. (2019). Investigating the  
818    relationship between spectro-temporal modulation detection, aided speech  
819    perception, and noise reduction preference. *Proceedings of ISAAR 2019: Auditory*  
820    *Learning in Biological and Artificial Systems. Vol 7*. Nyborg (Denmark).
- 
- 821  
822