Robust data-driven auditory profiling for precision audiology

Raul Sanchez Lopez[1], Michal Fereczkowski[1,2,3], Tobias Neher[2,3], Sébastien Santurette[1,4], and Torsten Dau[1]

[1] Hearing Systems Section, Department of Health Technology, Technical University of Denmark, Kgs. Lyngby, Denmark

[2] Institute of Clinical Research, Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark

[3] Research Unit for Oto-Rhino-Laryngology, Odense University Hospital, Odense, Denmark

[4] Centre for Applied Audiology Research, Oticon A/S, Smørum, Denmark

Abstract

Data-driven profiling allows uncovering complex hidden structures in a dataset and has been used as a diagnostic tool in various fields of work. In audiology, the clinical characterization of hearing deficits for hearing-aid fitting is typically based on the pure-tone audiogram only. Implicitly, this relies on the assumption that the audiogram can predict a listener's supra-threshold hearing abilities. Sanchez-Lopez *et al.* [Trends in hearing vol. 22 (2018)] hypothesized that the hearing deficits of a given listener, both at hearing threshold and at supra-threshold sound levels, result from two independent types of "auditory distortions". The authors performed a data-driven analysis of two large datasets with results from numerous tests, which led to the identification of four distinct "auditory profiles". However, the definition of the two types of distortion was challenged by differences between the two datasets in terms of the selected tests and type of listeners included in the datasets. Here, a new dataset was generated with the aim of overcoming those limitations. A heterogeneous group of listeners (N = 75) was tested using measures of speech intelligibility, loudness perception, binaural processing abilities and spectro-temporal resolution. The subsequent data analysis allowed refining the auditory profiles proposed by Sanchez-Lopez *et al.* (2018). Besides, a robust iterative data-driven method is proposed here to reduce the influence of the individual data in the definition of the auditory profiles. The updated auditory profiles may provide a useful basis for improved hearing rehabilitation, e.g. through profile-based hearing-aid fitting.

Keywords: Audiology, hearing deficits, precision medicine, data-driven analysis.

INTRODUCTION

Nowadays, “profiling” has gained broad attention as a tool for typifying groups of observations (e.g. users, recordings or patients) that follow similar patterns. Data-driven profiling allows uncovering complex structures that are “hidden” in the data. It has been used as a diagnostic tool in various fields, such as functional imaging (Krohne et al., 2019), psychology (Gerlach et al., 2018) or logopedics (Sharma et al., 2019). The idea of using computational data analysis that applies principles of the knowledge discovery from databases (KDD; Frawley et al., 1992) has recently gained attention in the field of audiology, for example in connection to hearing-aid features (Lansbergen & Dreschler, 2020; Mellor et al., 2018). As in stratified medicine (Trusheim et al., 2007), which pursues the identification of subgroups of patients (phenotypes) for the purpose of implementing more targeted treatments, it is of interest to identify subgroups of hearing-impaired (HI) listeners who might benefit from targeted hearing-aid fittings. As such, data-driven auditory profiling could help identify groups of listeners that are characterized by specific hearing disabilities, as a basis for precision audiology.

Hearing devices are the usual treatment for a hearing loss. Hearing-aid fitting mainly consists of the adjustment of amplification parameters to compensate for audibility loss and impaired loudness perception. Advanced hearing-aid signal processing features such as beamforming and noise reduction are typically not individually adjusted in this process, even though they could, in principle, be targeted towards the compensation of supra-threshold hearing deficits (Kiessling, 2001; Neher et al., 2016). However, the characterization of individual supra-threshold hearing deficits can be complex and requires

hearing assessments beyond standard audiometry. The definition of supra-threshold auditory deficits is commonly based on Plomp's (Plomp, 1978) "attenuation-distortion" model. Plomp suggested that hearing deficits affecting speech intelligibility comprise an "attenuation" and a "distortion" component. Whereas the attenuation component is assumed to affect speech intelligibility only in quiet, the distortion component is assumed to do so also in noise, yielding elevated speech reception thresholds in both cases. Kollmeier & Kiessling (2018) extended Plomp's approach and suggested a model that includes an attenuation component (affecting pure-tone sensitivity), a distortion component (affecting speech intelligibility in noise), and a neural component (affecting binaural processing abilities). Their model assumes that a sensorineural hearing loss is characterized by several factors: an "audibility loss", a "compression loss", a "central loss" and a "binaural loss". In general, the modelling approaches by Plomp (1978) and Kollmeier and Kiessling (2018) are rather conceptual and do not pinpoint specific underlying impairment factors nor specific measures to quantify these types of losses.

There have been some attempts to stratify HI listeners based on the shapes of their audiograms. Several classification schemes have been proposed in earlier studies, some of which were based on data-driven approaches (Bisgaard et al., 2010; Chang et al., 2019). Based on results from some human temporal bone studies, Schuknecht & Gacek (1993) proposed four different types of age-related hearing loss: Sensory presbycusis, neural presbycusis, metabolic presbycusis and mechanical or conductive presbycusis. Sensory presbycusis was related to alterations in the Organ of Corti and typically associated with basilar membrane compression loss, reduced frequency selectivity and elevated audiometric thresholds. This type of age-related hearing loss was considered to reflect the

loss of outer hair cells (OHC; Ahroon et al., 1993) and/or inner hair cells (IHC; Lobarinas, Salvi, & Ding, 2013), and was characterized by sloping audiograms. Neural presbycusis was related to a substantial loss of nerve fibers in the spiral ganglion. This type of presbyacusis was characterized by a progressive loss of speech discrimination performance, even though the audiometric thresholds remained unchanged over the same time period. Metabolic presbycusis was related to the atrophy of the stria vascularis that affects the transduction of the sensory cells because of a decreased endo-cochlear potential (EP). This type of impairment was associated with flat audiograms and did not affect speech discrimination (Pauler et al., 1986). Finally, conductive presbycusis corresponded to a gently sloping hearing loss at high frequencies, not reflecting morphological alterations in the sensory cells or stria vascularis but yielding elevated thresholds. However, recent results obtained with new techniques developed for histopathological analysis suggested that OHC dysfunction might have been underestimated in the case of conductive presbycusis and for some of the other types of age-related hearing loss (P. Wu et al., 2020).

Animal studies, where selective damage to the sensory cells or a change of the EP was induced, have allowed a consistent definition of the metabolic and sensory types of impairments in terms of audibility loss (Ahroon et al., 1993; Lobarinas et al., 2013; Mills et al., 2006). Dubno, Eckert, Lee, Matthews, & Schmiedt (2013) proposed a classification into sensory and metabolic audiometric phenotypes based on an approach that combined findings from animal models, expert medical advice and data-driven techniques. The main goal of their study was to analyze a large database of audiograms of HI individuals, and to identify connections between the findings from the animal studies with induced hearing losses and those based on human data. Whereas Schuknecht & Gacek (1993) characterized

the metabolic and sensory types of presbyacusis in terms of physiological impairments observed in humans, Dubno et al. (2013) proposed a phenotypical classification of the audiograms of HI listeners. Dubno et al.'s (2013) classification was thus solely based on the shape of the pure-tone audiogram. While this may help predict the possible origin of a listener's audibility loss, the resultant audiometric phenotypes cannot inform about any supra-threshold auditory processing deficits. The perceptual consequences of sensory or metabolic presbyacusis beyond audibility loss have not yet been studied.

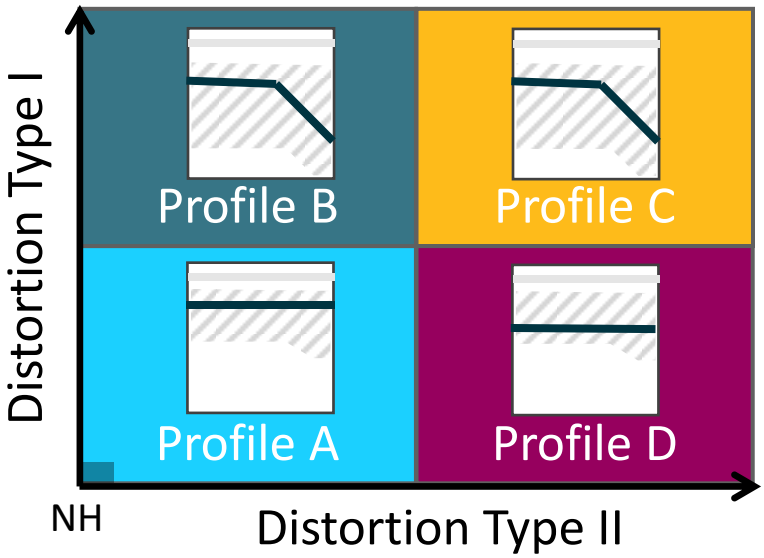

Figure 1: Sketch of the hypothesis of Sanchez-Lopez *et al.* (2018). The hearing deficits of a given listener can be described as a combination of two independent perceptual distortions. In a two-dimensional space, there would be four subgroups of listeners (Profiles A-D), which exhibit different *degrees* of the two distortion types.

Recently, Sanchez Lopez, Bianchi, Fereczkowski, Santurette, & Dau (2018) proposed a method for auditory profiling that was tested and verified by analyzing two datasets from

previous experimental studies (Johannesen et al., 2016; Thorup et al., 2016). Their method was tailored to the hypothesis that a listener's hearing deficits can be characterized by two independent types of “auditory distortion”, type-I and type-II, as illustrated in Figure 1. In this two-dimensional space, a normal-hearing (NH) listener would be placed at the origin whereas other listeners, with hearing losses that differ in the degree of the two types of distortions, would be placed at different positions along the two dimensions. Each type of distortion would then be defined by specific deficits observed in behavioral tasks that covary together and define a given auditory profile. While Profile C represents a high degree of both types of distortion, profiles B and D reflect hearing deficits dominated by one of the two distortions. Profile A, the group with a low degree of distortions, represents near-normal hearing and thus only mild hearing deficits.

In Sanchez Lopez et al. (2018) it was hypothesized that distortion type-I covaries with a loss of audibility, whereas distortion type-II was assumed to be unrelated to audibility. However, the results of the analysis of two different datasets did not support this hypothesis. In fact, the analysis of the two datasets showed that distortion type-I was connected to high-frequency hearing loss and reduced speech intelligibility. Regarding distortion type-II, the analysis of one of the datasets (Thorup et al., 2016) provided a link to reduced binaural processing abilities, whereas the analysis of the other dataset (Johannesen et al., 2016) was linked to low-frequency hearing loss. These mixed results were attributed to differences between the two datasets in terms of the selection of the listeners and chosen behavioral tests. It was concluded that there was a need for a new dataset that included a larger variability of impairment factors across listeners to better characterize the listeners’ auditory distortions and, thus, the auditory profiles. Furthermore,

it was suggested that the chosen tests should investigate several aspects of auditory processing while at the same time being clinically feasible.

A new dataset was therefore generated with the aim of overcoming these limitations. Seventy-five listeners were tested in a clinical environment. The behavioural tasks included measures of audibility, loudness perception, binaural processing abilities, speech perception, spectro-temporal modulation sensitivity and spectro-temporal resolution (Sanchez-Lopez et al., 2020). They included several measures that can be connected to previous approaches such as the attenuation-distortion model (regarding speech perception measures) and the neural component (regarding binaural processing abilities). The outcome measures related to binaural processing abilities were not found to be correlated to each other, suggesting that they might correspond to different aspects of auditory processing (Sanchez-Lopez et al., 2020). Therefore, it was of interest to further investigate the connections between outcome measures and the two distortion types in a data-driven approach.

The analysis of the new dataset was performed with a refined version of the data-driven method provided in Sanchez-Lopez et al. (2018). Importantly, the current study did not aim to disentangle the effects of audibility and supra-threshold deficits but to identify four robust listener subpopulations based on the data-driven analysis of the new dataset. The outcomes of the analysis were discussed in relation to previous classification approaches as well as in terms of implications towards profile-based rehabilitation strategies. Moreover, a decision tree consisting of the auditory measures that best classified the listeners into the four profiles was generated.

zsanchez_et_al_2020_RobustProfiling_medRxiv

METHOD

The development of the data-driven method for auditory profiling was based on two premises: 1) the identifaction of relevant outcome measures that tap into two independent sources of variation, and 2) the identification of extreme exemplars that can serve as "prototypes" of different subgroups of listeners.

Description of the dataset

Seventy-five listeners participated in the study. Seventy of the listeners presented various degrees and shapes of symmetrical, sensorineural hearing losses, while five showed near-normal audiometric thresholds. The listeners were recruited from the clinical databases at Odense University Hospital (OUH), Odense, Denmark and Bispebjerg Hospital (BBH), Copenhagen, Denmark. All listeners completed the "BEAR test battery" (Sanchez-Lopez et al., 2020). This test battery consists of a total of 10 psychoacoustic tests. The tests are divided into six aspects of auditory processing: audibility, speech perception, loudness perception, binaural processing abilities, spectro-temporal modulation sensitivity and spectro-temporal resolution. All participants gave written informed consent and received financial compensation for their participation (approved by the Science-Ethics Committee for the Capital Region of Denmark H-16036391).

The tests were carried out in a double-walled booth (at BBH and DTU) or in a small anechoic chamber (at OUH). The stimuli were presented via headphones (Sennheiser HDA200). For each of the tests, the outcome measures of interest were extracted from the raw results. For example, the speech reception threshold (SRT) in quiet was estimated from

the word discrimination scores obtained at different speech levels. When the tests contained frequency-specific measures, the results were grouped into low-frequency (≤1 kHz) and high-frequency (>1 kHz) averages. This decision was motivated by previous studies (Bernstein et al., 2016; Sanchez-Lopez et al., 2018). In the case of monaural measures, the mean values across ears were used. The data were cleaned following the principles of the KDD, to remove outliers or unreliable data before the analysis. For example, some of the listeners performed the speech-in-noise test at lower levels than the level recommended for the hearing-in-noise measurements (Plomp, 1986). Since speech-in-noise perception is of great interest in the present analysis, unreliable measurements of speech reception thresholds in noise (SRT_N) and sentence recognition scores ($SScore^{4dB}$) were considered as missing data.

In the next step, the data were normalized between the 25th and 75th percentiles, such that the 25th percentile corresponded to a value of -0.5 and the 75th to a value of 0.5. In total, 26 variables were selected from the outcome measures, as shown in Table I. The resulting dataset (‘BEAR3’) is publicly available in a Zenodo repository (doi: 10.5281/zenodo.3459579; Sanchez-Lopez et al., 2019)

zsanchez_et_al_2020_RobustProfiling_medRxiv

Table 1: Description of the tests, dimensions and outcome measures contained in the BEAR3 dataset (Sanchez-Lopez et al., 2019). For each test, a reference is included. The tests are divided by categories, and the outcome variables are presented in the right column.

Test name	***Reference***	***Category***	***Outcome variable***
Pure-tone audiometry	*ISO 8253-1, (2010)*	*Audibility (AUD)*	AUD_x
Fixed level frequency threshold	*Rieke et al., (2017)*		*FLFT*
Adaptive categorical loudness scaling Loudness	*Brand & Hohmann, (2002)*		HTL_x,
		Loudness perception (LOUD)	MCL_x $DynR_x$, $Slope_x$
Word recognition scores Speech	*ISO 8253-3, (2012)*	*Speech in quiet (SiQ)*	SRT_Q, *Max DS*
Hearing in noise test	*Nielsen & Dau, (2011)*	*Speech in noise (SiN)*	SRT_N, $SScore^{4dB}$
Maximum frequency for IPD detection Binaural	*Füllgrabe, Harland, Sęk, & Moore, (2017)*	*Binaural processing abilities (BIN)*	IPD_{fmax}
Binaural pitch processing	*Santurette & Dau, (2012)*		*BP20*
Extended binaural audiometry in noise abilities	*Durlach, (1963)*		*BMR*
Fast spectro-temporal modulation Spectro-sensitivity temporal	*Bernstein et al., (2016)*	*Spectro-temporal modulation sensitivity (STM)*	$sSTM_8$, $fSTM_8$
Extended audiometry in noise	*Moore, Huss, Vickers, Glasberg, & Alcántara, (2000)*	*Spectro-temporal resolution (STR)*	TiN_x
	Schorn & Zwicker, (1990; van Esch et		SMR_x, TMR_x

zsanchez_et_al_2020_RobustProfiling_medRxiv

Test name	***Reference***	***Category***	***Outcome variable***
	al., (2013)		
AUD_x: Pure-tone average at low (x=LF; f ≤ 1kHz) or high (x=HF; f > 1kHz) frequencies. // ACALOS outcome variabless are averaged for low (x=LF; f ≤ 1kHz) and high (x=HF; f > 1kHz) frequencies. // Extended audiometry outcome measures were measured at 0.5 kHz (x=LF) and at 2 kHz (x=HF).			

There were six variables related to audibility and loudness perception: 1) pure-tone average at low frequencies (AUD_{LF}; f ≤ 1kHz) and at high frequencies (AUD_{HF}; f > 1kHz); 2) fixed-level frequency threshold (FLFT) measured at 80 dB sound pressure level (SPL); 3) hearing threshold levels (HTL) estimated from the loudness function, averaged for low (HTL_{LF}) and high (HTL_{HF}) frequencies; 4) most comfortable level (MCL) estimated from the loudness function, averaged for low (MCL_{LF}) and high (MCL_{HF}) frequencies; 5) dynamic range (DynR) estimated as the difference between the uncomfortable level and HTL, estimated from the loudness function for low ($DynR_{LF}$) and high ($DynR_{HF}$) frequencies; and 6) slope of the loudness function at low ($Slope_{LF}$) and high ($Slope_{HF}$) frequencies. For the outcome measures estimated from the loudness function, the low-frequency average corresponded to the center frequencies 0.25, 0.5 and 1 kHz and the high-frequency average corresponded to the center frequencies 2, 4 and 6 kHz.

There were four variables related to speech perception: 1) speech reception threshold in quiet (SRT_Q); 2) maximum word recognition score (Max DS); 3) speech reception threshold in noise (SRT_N); and 4) sentence recognition score at +4 dB SNR (SS^{4dB}).

There were three variables related to binaural processing abilities: maximum frequency for detecting an interaural phase difference of 180° (IPD_{fmax}); binaural pitch detection

zsanchez_et_al_2020_RobustProfiling_medRxiv

performance, estimated as the percent correct of the dichotic presentations (BP_{20}); and binaural masking release (BMR). BMR was estimated as the difference between the threshold in the diotic tone-in-noise detection condition (N_0S_0) and the threshold in the dichotic tone-in-noise detection condition where the tone was out of phase between the ears (N_0S_π). The frequency of the tone presented in the two conditions was 0.5 kHz.

The spectro-temporal processing variables included: 1) spectro-temporal modulation sensitivity at +3 dB modulation depth ($sSTM_8$) and 2) the "fast" spectro-temporal modulation detection threshold ($fSTM_8$); 3) the tone-in-noise detection threshold at 500 Hz (TiN_{LF}) and at 2 kHz (TiN_{HF}); 4) the spectral masking release (SMR) estimated as the difference between the tone-in-noise detection threshold (TiN) and the corresponding threshold with the noise shifted towards high frequencies ($f_{c,noise} = 1.1\ f_{tone}$); and 5) the temporal masking release (TMR) estimated as the difference between the TiN masked threshold and the corresponding threshold with the tone presented in temporally-modulated noise (modulation frequency, $f_m = 4$ Hz).

Stages of the data-driven method

As in Sanchez-Lopez et al. (2018), the data-driven analysis used here was based on unsupervised learning and was divided into three main steps illustrated in the top panel of Figure 2:

I. Dimensionality reduction: Based on principal component analysis (PCA), a subset of variables that were highly correlated with the first two principal components, PC1 and PC2, was kept for the following steps (II and III). The subset could consist

of 3, 4 or 5 variables per PC. Hence, up to 10 variables could be kept for the next step. The to-be-kept variables were chosen in an iterative process using a leave-one-out cross-validation. In each iteration, one variable was removed according to the variance explained by the remaining variables i.e., the subset of variables that explained the largest amount of variance was kept and the left-out variable was discarded. Additionally, since the use of several intercorrelated variables in PCA can bias the results, highly correlated variables were removed. If two variables resulting from step I were highly correlated (Pearson's correlation coefficient, $r > 0.85$), one of them was dropped and this step was repeated.

II. Archetypal analysis: The data were decomposed into two matrices – the 'test matrix', which contained the extreme patterns of the data (archetypes) and the 'subject matrix', which contained the weights for each archetype. A given subject was then represented as a convex combination of the archetypes (Cutler & Breiman, 1994). The specific method used here was similar to the one proposed in Mørup & Hansen (2012). The analysis was limited to four archetypes to improve the interpretability of the results on the scope of the hypothesis.

III. Profile identification: The subject matrix was used to estimate the distance between observations and the four archetypes. Each listener (subject) was then assigned to an auditory profile group based on their weights in the subject matrix. The sum of weights for each listener was always 1. Listeners with a weight above 0.51 for one of the four archetypes were identified as belonging to that auditory profile (Ragozini et al., 2017). Otherwise, they were left "unidentified" ('U').

zsanchez_et_al_2020_RobustProfiling_medRxiv

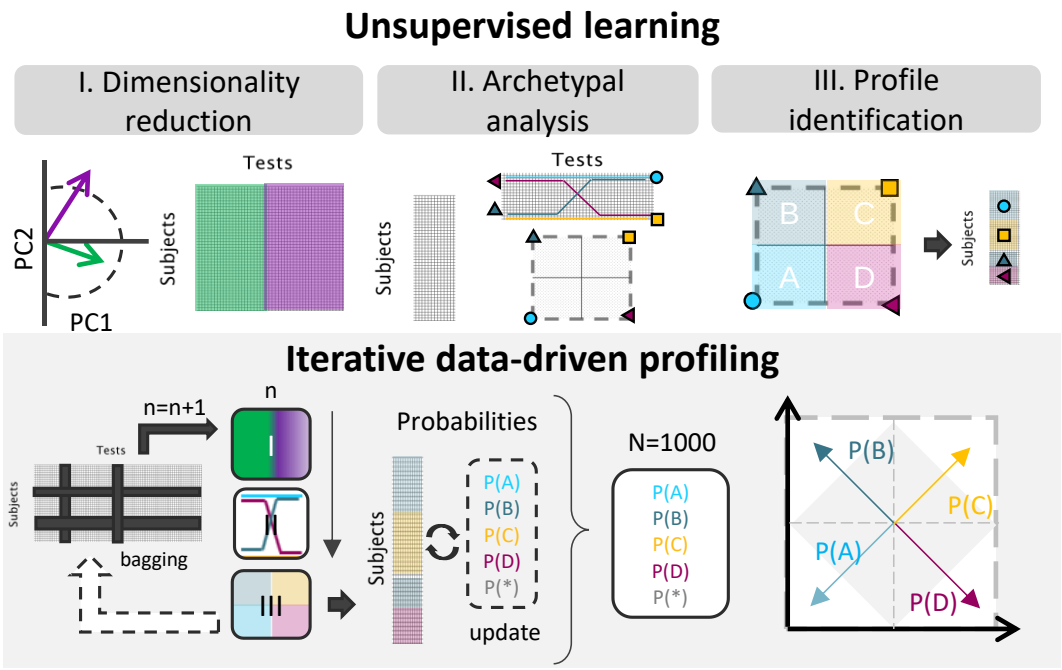

Figure 2: Sketch of the refined data-driven method for auditory profiling. Top panel: The unsupervised learning stages of Sanchez-Lopez et al. (2018): I) dimensionality reduction; II) archetypal analysis; III) profile identification. Bottom panel: In each iteration, a subset of the dataset was processed using dimensionality reduction, archetypal analysis and profile identification. The profile identification stage was two-fold: 1) In each of the iterations, the profiles were identified based on the archetypal analysis. 2) After 1000 iterations, the probability was calculated based on the prevalence of each observation and the number of identifications as each of the profiles. Listeners with higher probabilities of belonging to an auditory profile were placed close to the corners in the square representations and the ones with lower probability ($P < 0.5$) were located inside the grey square.

Iterative data-driven profiling

The robust data-driven auditory profiling method aimed to improve the previous method proposed in Sanchez-Lopez et al. (2018) by reducing the influence of the data on the

definition of the auditory profiles. In any data-driven analysis, and especially in unsupervised learning, individual data points can highly influence the results and lead to misinterpretations. Resample techniques such as bagging are commonly used for supervised learning. Moreover, bagging can improve cluster analysis, making the results less sensitive to the type and number of variables (Dudoit & Fridlyand, 2003). The three unsupervised learning steps were repeated 1000 times, as illustrated in the bottom panel of Figure 2. Before each repetition, the full dataset was decimated randomly in terms of subjects and tests in each iteration. The analysis was performed with only 83% of the data (69 out of 75 listeners and 24 out of 26 variables) in each repetition. In the case of missing data, an algorithm based on spring metaphor (D'Errico, 2020) was used to predict those data points. Furthermore, in step I (dimensionality reduction), the number of selected variables (6, 8 or 10) was also randomly selected in each iteration to further randomize the procedure. Steps II (archetypal analysis) and III (profile identification) yielded a pre-classification of the subjects contained in the subset of the data corresponding to each iteration. The probability of each listener of being identified as a given auditory profile depended on the number of times a given listener was "out-of-bag" in individual repetitions and the profile identification result from step III. In each iteration, the profile probabilities [P(A), P(B), P(C) or P(D)] and the probability of being unidentified [P(U)] were updated.

After 1000 repetitions, the listeners were divided into four subgroups based on the computed probabilities. If a given listener showed a probability above 0.5 of belonging to any of the auditory profiles, the listener was assigned to that profile. However, if the highest probability was below 0.5, the criterion for being included in one of the four clusters was that the difference between the two highest probabilities had to be above 0.1 to be

considered significant. The projection of the probabilities on a two-dimensional space was done by considering four vectors, one for each profile probability, pointing towards each of the corners in a squared representation, as depicted in the right-bottom panel of Figure 2. Graphically, the listeners belonging to an auditory profile were then placed close to the corners.

Distortion estimation from the square representation

The final output of the refined data-driven method was the probability, *P*, of being identified as belonging to an auditory profile (A-D). Regarding the square representation or convex hull, which resembled the hypothesis shown in Figure 1, the probabilities of belonging to an auditory profile were depicted as vectors with the origin at the center of the square and oriented towards each of the four corners (Figure 2). Assuming that *P*(*B*) and *P*(*C*) are proportional to auditory distortion type-I (AD_I) and that assuming that *P*(*C*) and *P*(*D*) are proportional to auditory distortion type-II (AD_{II}), this yields:

$$AD_I = \frac{1}{2}(1 + P(B \cup C) - P(A \cup D)), \quad (1)$$

$$AD_{II} = \frac{1}{2}\big(1 + P(C \cup D) - P(A \cup B)\big). \quad (2)$$

Each listener was placed in the two-dimensional space along with the two estimated distortions. In addition, prototypes reflecting extreme exemplars, equivalent to the archetypes yielded by the archetypal analysis, were estimated by averaging the results of the five listeners with the highest probabilities of belonging to a given auditory profile (A-D).

zsanchez_et_al_2020_RobustProfiling_medRxiv

The relations between the AD_I and AD_{II} with the variables considered in the study were investigated using stepwise linear regression models. The variables included in the model fitting were the outcome variables resulting from the supra-threshold tests, except for AUD_{LF} and AUD_{HF}. Listeners with a high probability of not being identified as any of the four profiles ($P(U) > 0.5$) were discarded. The criterion for adding a variable as a predictor of one of the distortions was an improvement of the adjusted R^2 by more than 0.01.

Decision trees

A decision tree was fitted to the entire dataset following the splitting criterion of weighted impurity (Breiman et al., 2017). Since it was of interest to obtain a decision tree with outcome measures beyond audiometry, the variables from the pure-tone audiometry were excluded from this analysis. The resulting decision tree was pruned to only have three levels and a maximum of seven binary splits. Because of the missing data, the decision tree was surrogated, i.e., it ignored the missing data to facilitate its interpretability.

RESULTS

zsanchez_et_al_2020_RobustProfiling_medRxiv

Data-driven auditory profiling

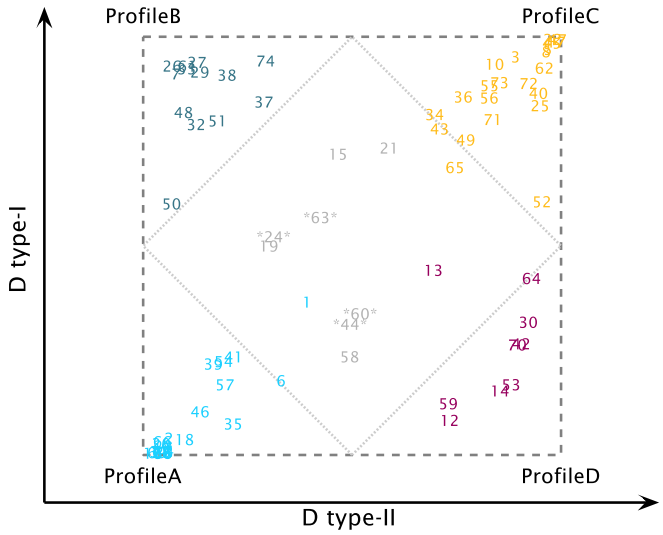

Figure 3: Square representation of the auditory profiles. The listeners are placed in the square representation based on their probability of belonging to one of the subgroups. The inner rhombus delimits the area where the listeners showed $P < 0.50$ of belonging to any subgroup. The listeners labeled as ** showed $P(U) > 0.5$.

Figure 3 shows the results of the analysis where each listener is located in the two-dimensional space according to their degree of type-I and type-II distortion. The degree of distortion was calculated based on the probability of belonging to any of the four auditory profiles. Listeners located close to a corner exhibited a high probability of belonging to a corresponding profile. Unidentified listeners are placed in-between the four quadrants and are marked with grey. Profile A (n = 24) and Profile C (n = 22) represented the most populated groups. The five normal-hearing listeners were placed at the bottom-lefthand corner in Profile A. Profiles B (n = 13) and Profile D (n = 9) represented smaller subgroups.

The probabilities indicated in the Profile B listeners ($0.58 < P(B) < 0.89$) and in the Profile D listeners ($0.36 < P(D) < 0.73$) were, on average, lower than for the Profile A listeners ($0.30 < P(A) < 0.97$) and the Profile C listeners ($0.43 < P(C) < 0.98$). Four listeners showed a high probability of being unidentified ($0.33 < P(U) < 0.77$), and four other listeners were "inconclusive" as reflected in similar probabilites of belonging to two profiles with both $P < 0.3$. The five listeners showing the largest probabilities of belonging to one of the auditory profiles (excluding the normal-hearing listeners) were considered to represent the prototypes shown in Figure 4.

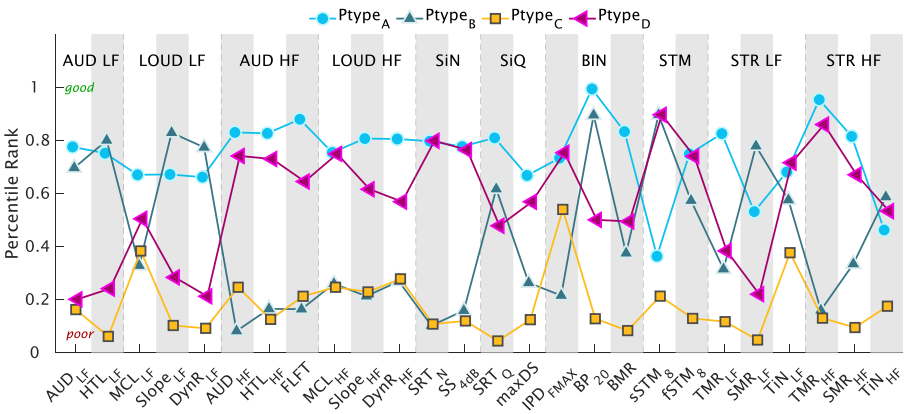

Figure 4. Prototypes (Ptype): Percentile rank across variables corresponding to the extreme exemplars of the different patterns found in the data. The ranks are shown for the 26 outcomes corresponding to the different aspects of auditory processing. AUD: Audibility, LOUD: Loudness, SiN: Speech-in-noise perception, SiQ: Speech-in-quiet, BIN: Binaural processing abilities. STM: Spectro-temporal modulation sensitivity, STP: Spectro-temporal processing abilities, divided into temporal and spectral masking release as well as tone in noise detection. Subgroups of measures with frequency-specific outcomes were divided into low (LF) and high (HF) frequencies.

zsanchez_et_al_2020_RobustProfiling_medRxiv

The prototypes show archetypal patterns in the data associated with the performance obtained by the four different groups. A higher percentile rank corresponds to a higher percentile of the overall data distribution and thus to a "good" performance. Each point in Figure 4 corresponds to the mean of the listeners forming the corresponding prototype. Likewise, a low percentile rank corresponds to a "poorer" performance. Prototype A (blue circles in Figure 4) showed a good performance in most of the outcome measures. However, the outcomes of the tests related to $sSTM_8$ and TiN_{LF} were below the 50th percentile. Prototype C (yellow squares) showed the poorest performance for most outcome measures, with only MCL_{LF} and IPD_{fmax} above the 30th percentile. Prototype B (dark-green upwards-pointing triangles), with a high degree of distortion type-I and a low degree of distortion type-II, showed a good performance for the outcome measures obtained at low frequencies and for BP_{20}, whereas performance was poor for the outcomes obtained at high frequencies, IPD_{fmax} and for the speech-in-noise perception tests (SiN). In contrast, prototype D (magenta left-pointing triangles), with a high-degree of distortion type-II and a low degree of distortion type-I, showed a good performance in terms of SiN and IPD_{fmax}, and a relatively good performance (above the 60th percentile) for most outcomes measures obtained at high frequencies, whereas the performance was poor for outcome measures obtained at low frequencies, especially in terms of loudness (LOUD), TMR_{LF} and SMR_{LF}. The prototypes showed opposite results for the profiles located in opposite corners of Figure 3 (A vs. C and B vs. D).

zsanchez_et_al_2020_RobustProfiling_medRxiv

Relations between auditory distortion types and outcome measures

Table 2: Stepwise regression analysis of auditory distortion (AD) type-I and type-II. The priority was established based on the accumulated adjusted $R^2 > 0.01$. Columns show the predictor name, the estimate, standard deviation (SE), *t*-value and probability of a significant contribution (*p*).

Priority	*Predictor*	*Estimate*	*SE*	*t*	*p*	*Adj R^2*
Model AD type-I						
n/a	*(Intercept)*	*250.0*	*109.0*	*2.3*	*<0.05*	*-*
1	*HTL_{HF}*	*-9.7*	*2.5*	*-3.9*	*<0.0001*	*0.79*
2	*TMR_{HF}*	*-1.6*	*0.6*	*-2.6*	*<0.05*	*0.82*
3	*TiN_{LF}*	*-3.5*	*1.5*	*-2.3*	*<0.05*	*0.83*
4	*$HTL_{HF}:TiN_{LF}$*	*0.2*	*0.1*	*4.6*	*<0.0001*	*0.87*
5	*TMR_{LF}*	*-1.8*	*0.6*	*-2.8*	*<0.01*	*0.88*
Model AD type-II						
n/a	*(Intercept)*	*-18.7*	*3.9*	*-4.7*	*<0.0001*	*-*
1	*HTL_{LF}*	*2.4*	*0.14*	*17.4*	*<0.0001*	*0.84*

The relations between the two types of distortions and outcome measures were studied using stepwise regression analysis (Table 2). Distortion type-I was found to be associated with elevated hearing thresholds at high frequencies, a reduced temporal masking release and increased tone-in-noise detection thresholds at low frequencies. Furthermore, distortion type-I was significantly correlated with SRT_N ($r = 0.76$; $p < 0.0001$), even when the effects of audibility were partialled out ($r = 0.33$; $p < 0.01$). In contrast, the correlations found between distortion type-I and speech recognition in quiet ($r = 0.71$; $p < 0.0001$) were not significant when partialling out audibility ($r = 0.15$; $p > 0.1$). Distortion type-II was only associated with hearing thresholds at low frequencies. The restrictive criterion (increased of $R^2 > 0.01$) did not include other variables in the model. However, distortion type-II was significantly correlated were with the slope of the loudness function ($r = 0.72$;

zsanchez_et_al_2020_RobustProfiling_medRxiv

$p < 0.0001$) and with the amount of spectral masking release at low frequencies ($r = 0.61$; $p > 0.0001$). In addition, distortion type-II was correlated with SRT_Q ($r = 0.83$; $p < 0.0001$) but not with SRT_N ($r = 0.21$; $p > 0.05$). However, the correlation between SRT_Q and distortion type-II was weaker when controlling for the effects of audibility ($r = 0.30$; $p < 0.05$). Moreover, the majority of the auditory outcomes were not significantly correlated with distortion type-II when hearing thresholds were partialled out, except for TMR_{HF} ($r = 0.35$; $p < 0.01$).

The outcome measures related to binaural processing abilities (Figure 4) provided unexpected results. Indeed, the prototypes showed opposite trends for IPD_{fmax} and BP_{20}, which could indicate that they reflect different auditory distortions. Distortion type-I was significantly correlated with both IPD_{fmax} and BP_{20}, but only BP_{20} remained significant after controlling for audibility ($r = -0.36$; $p < 0.01$). In contrast, distortion type-II was only correlated with BP_{20} ($r = -0.58$; $p < 0.0001$) before partialling out the effects of audibility ($r = -0.1$; $p = 0.5$). Besides, IPD_{fmax} was neither correlated with any of the two distortion types when controlling for audibility nor with any of the other BIN outcome measures ($r << 0.1$; $p > 0.15$). Instead, IPD_{fmax} was highly correlated with the tone-in-noise detection threshold at low frequencies ($r = -0.53$; $p < 0.0001$) – one of the main predictors of distortion type-I – even when audibility was partialled out ($r = -0.56$; $p < 0.0001$).

zsanchez_et_al_2020_RobustProfiling_medRxiv

Decision tree for the identified auditory profiles

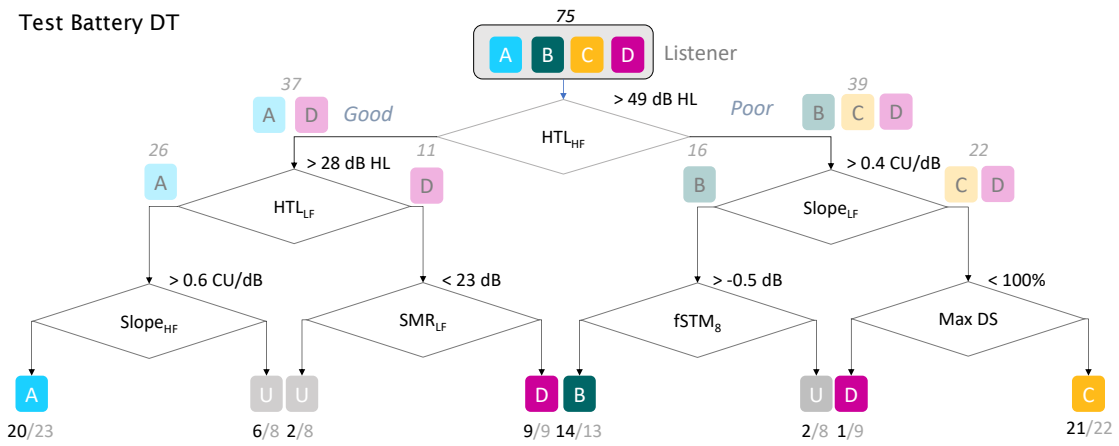

Figure 5: Decision tree fitted to the BEAR3 dataset using the auditory profiles as the output. For each binary split, the right branch corresponds to a "poor" result and the left branch to a "good" result. In each binary split, the number of listeners assigned to each branch are shown together with the most likely outputs. The classes (A-D) are together with the number of listeners belonging to that class and the number of identified listeners for a given profile.

Figure 5 shows the decision tree fitted to the BEAR3 dataset using the identified auditory profiles as well as the unidentified listeners. The decision tree has three levels. The first level corresponds to high-frequency hearing loss as estimated using ACALOS, which splits the listeners into two branches: Profiles A and D (HTL_{HF} < 49 dB HL) are separated from Profiles B and C (HTL_{HF} > 49 dB HL), together with one listener from Profile D. Thus, this first level makes a classification based on the degree of distortion type-I. The second level corresponds to outcomes measured at low frequencies and estimated using the loudness functions, which divide the listeners according to their degree of distortion type-II. Profile D (HTL_{LF} > 28 dB HL) and Profile C ($Slope_{LF}$ > 0.4 CU/dB and maxDS <

100%). The third level makes use of outcomes related to loudness, spectro-temporal modulation and spectral masking release for classifing the unidentified listeners.

DISCUSSION

Two types of distortion to characterize individual hearing loss

The term "distortion" in hearing science has typically been associated with elevated SRT_N, as reflected in Plomp's SRT model (Plomp, 1978). Here, we introduced the term "auditory distortions" to describe the perceptual consequences of sensory hearing impairment, including (but not limited to) loss of sensitivity. The two types of perceptual distortions considered here should thus be considered as consequences, and not sources of, sensory impairments. An interesting aspect of our data-driven profiling method is that the auditory distortions reflect two fairly independent dimensions of perceptual deficits associated with sensorineural hearing impairments.

Although Plomp's attenuation and distortion components are often assumed to be independent, some impairment mechanisms may, in fact, affect both speech-in-noise perception and audiometric thresholds, especially at high frequencies (Moore, 2016), which is consistent with distortion type I. Schädler, Hülsmeier, Warzybok, & Kollmeier (2020) attempted to model supra-threshold auditory deficits that are independent of audibility loss. Their results suggested that reduced speech intelligibility represents an auditory perceptual deficit that may be associated with reduced tone-in-noise detection which is in agreement with the results from the current study. However, as demonstrated here, speech-in-noise perception can also be affected by deficits that covary with the

audiometric thresholds (distortion type-I), which should not be underestimated, especially when the high-frequency hearing loss exceeds 50 dB HL (Profiles B and C), as depicted in Figure 6.

Regarding the 'neural component' associated with reduced binaural processing abilities (Kollmeier & Kiessling, 2018), the BIN measures considered in the present study provided contradictiory results in connection to the proposed auditory profiles. Even though IPD_{fmax} represents a test that has been proposed to reveal binaural disabilities related to the disruption of temporal fine structure (TFS) coding (Füllgrabe et al., 2017), a recent study linked the detection of interaural phase differences to outcomes from cognitive tests (Strelcyk et al., 2019). This suggests that IPD_{fmax} might not reflect a purely auditory process but might also depend on top-down processes such as processing speed. Since IPD_{fmax} and TiN_{LF} were strongly correlated, the two tasks might be affected by either cognitive or auditory processes, which should be investigated further.

The two types of auditory distortions shown here were consistent with Plomp's (1978) approach. However, the two auditory distortion types presented here are, in fact, the result of a data-driven analysis of a large multi-dimensional dataset rather than the conceptual interpretation of speech intelligibility deficits. Distortion type-I may then be considered as a "speech intelligibility related distortion" and distortion type-II as a "loudness perception related distortion". Nevertheless, the listeners with higher degrees of the two types of distortions showed perceptual deficits with respect to spectro-temporal processing and binaural processing abilities thus reflecting deficits that are beyond a simple combination of loudness and speech-intelligibility deficits.

zsanchez_et_al_2020_RobustProfiling_medRxiv

Auditory profiles and hearing-loss phenotypes

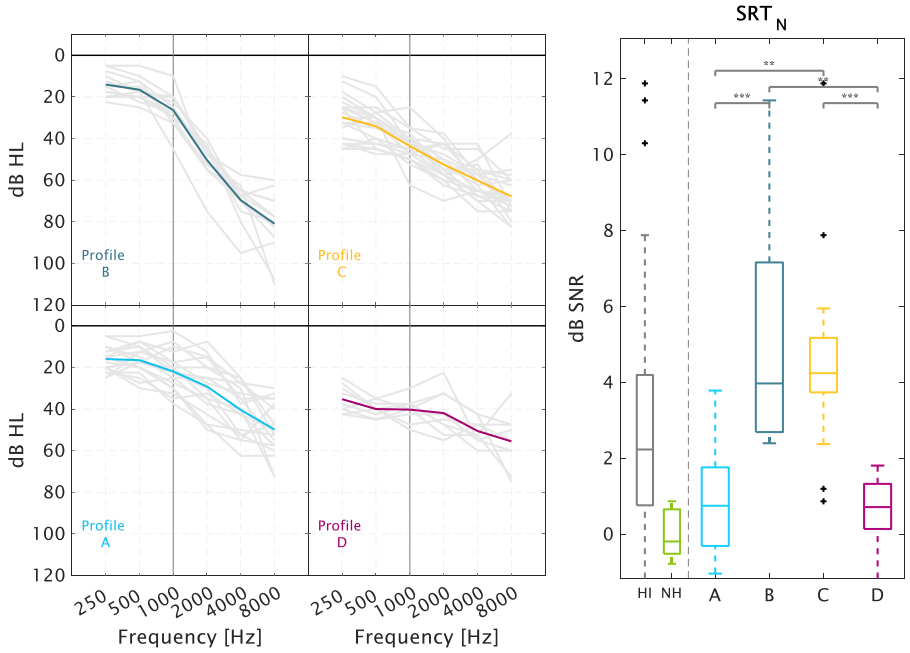

Figure 6: Audiometric thresholds of the four auditory profiles and speech intelligibility in noise. Left panel: The averaged audiometric thresholds of each profile are shown together with the individual audiograms. Right panel: Speech reception thresholds in noise (SRT_N), with boxplots of the HI and NH data (left) and the four auditory profiles (right). The multicomparison analysis revealed significant differences between the groups (*** $p < 0.0001$, ** $p < 0.001$).

Figure 6 shows the audiometric thresholds corresponding to the four robust auditory profiles. Profile A corresponds to a mild, gently sloping high-frequency hearing loss; Profile B corresponds to a steeply sloping high-frequency hearing loss; Profile C corresponds to a low-frequency hearing loss between 30 and 50 dB HL and above 50 dB HL at high frequencies; and Profile D corresponds to a fairly flat hearing loss with audiometric thresholds between 30 and 50 dB HL. Interestingly, these four audiometric configurations look similar to the audiometric phenotypes of Dubno et al. (2013), which

are based on Schuknecht's metabolic and sensory types of presbyacusis (Schuknecht & Gacek, 1993). The main difference between the two approaches is that the audiometric thresholds shown here correspond to four subgroups of HI listeners, which are the result of a data-driven analysis involving several auditory measures and not only the audiometric thresholds.

In previous studies, metabolic hearing loss (MHL) yielded flat elevated audiometric thresholds, but did not affect speech intelligibility in noise (Pauler et al., 1986), which is consistent with the results of the present study for Profile D listeners. In a MHL, the atrophy of the stria vascularis produces a reduction of the EP in the scala media (Schmiedt et al., 2002). The EP loss mainly affects the electromotility properties of the OHCs, i.e. the cochlear amplifier. Therefore, metabolic hearing loss can be considered as a cochlear gain loss that impairs OHC function across the entire basilar membrane. This, in turn, affects the hearing thresholds and is associated with a reduced frequency selectivity (Henry et al., 2019). In the present study, Profile D was characterized by an abnormal loudness function, particularly at low frequencies, and a significantly reduced spectral masking release, although speech-in-noise intelligibility and binaural TFS sensitivity were near-normal. However, one needs to bear in mind that the results observed for the listeners in Profile D might also be compatible with other types of impairments. Sensory hearing loss (SHL) is typically associated with OHC dysfunction, which yields elevated thresholds at more specific frequency regions, a loss of cochlear compression and reduced frequency selectivity (Ahroon et al., 1993). However, audiometric thresholds above about 50 dB HL at high frequencies cannot be attributed only to OHC due to the limited amount of gain induced by the OHC motion, which implies additional IHC loss or a loss of nerve fibers

zsanchez_et_al_2020_RobustProfiling_medRxiv

(Hamernik et al., 1989; Stebbins et al., 1979). Therefore, listeners classified as Profile B or Profile C (i.e. with a higher degree of distortion type-I and a high-frequency hearing loss) may exhibit a certain amount of IHC dysfunction that might produce substantial supra-threshold deficits. Animal studies have shown that audiometric thresholds seem to be insensitive to IHC losses of up to about 80% (Lobarinas et al., 2013). This suggests that HTL > 50 dB HL might indicate the presence of hearing deficits that may distort the internal representation, not only in terms of frequency tuning but also in terms of a disruption of temporal coding due to the lack of sensory cells (Moore, 2001; Stebbins et al., 1979).

Profile B's audiometric thresholds are characterized by a sloping hearing loss with normal values below 1 kHz. However, Profile B exhibited the poorest performance in the IPD_{fmax} test, which cannot be explained by an audibility loss. Neural presbyacusis is characterized by a loss of nerve fibers in the spiral ganglion that is not reflected in the audiogram. Furthermore, primary neural neurodegeneration, recently termed cochlear synaptopathy (Kujawa & Liberman, 2009; P. Z. Wu et al., 2019) or deafferentiation (Lopez-Poveda, 2014), might be reflected in the results of some of the supra-threshold auditory tasks used here. However, the perceptual consequences of primary neural degeneration are still unclear due to the difficulty of assessing auditory nerve fibers loss in living humans (Bramhall et al., 2019). This makes it difficult to link the effects of deafferentation to the reduced binaural processing abilities observed in listeners in Profile B and Profile C.

As suggested by Dubno et al. (2013), the audiometric phenotype characterized by a severe hearing loss (similar to the one corresponding to Profile C) might be ascribed to a combination of MHL and SHL. In the present study, Profile C listeners performed similarly

zsanchez_et_al_2020_RobustProfiling_medRxiv

to Profile B listeners in supra-thresholds tasks related to distortion type-I (e.g. SRT_N and TMR_{HF}) and also similarly to Profile D listeners in tasks related to distortion type-II (e.g. loudness perception). In contrast, Profile C listeners also showed poorer performance in tests such as binaural pitch detection, tone-in-noise detection and spectro-temporal modulation sensitivity, which is not consistent with the idea of a simple superposition of the other profiles. As mentioned above, these deficits observed in Profile C listeners might be a consequence of auditory impairments that are unrelated to the loss of sensitivity, such as deafferentation, which can be aggravated by the presence of a MHL and SHL. However, Bernstein et al. (2016) found spectro-temporal modulation sensitivity to be a good predictor of aided speech perception only in the cases of a moderate high-frequency hearing loss. They suggested that cognitive factors might be involved in the decreased speech intelligibility performance when the high-frequency hearing loss is >50 dB HL. Therefore, Profile C listeners might be affected by both auditory and non-auditory factors that worsen their performance in some demanding tasks.

Stratification in hearing research and hearing rehabilitation

In the present study, the two principal components of the dataset seemed to be dominated by the listeners' low- and high-frequency hearing thresholds. Therefore, another supra-threshold hearing deficit might be hidden in the four auditory profiles that could explain the individual differences across listeners belonging to the same profile. To explore these "additional deficits" not covered by the present approach, a stratification of the listeners might be necessary. Lőcsei et al. (2016) investigated the influence of TFS on speech perception for different interferers. In their study, the HI listeners were divided into groups

based on the degree of hearing loss at high frequencies. In their study, stratification of the listeners into two subgroups helped reduce the potential effect of audibility on speech intelligibility. In another study, Papakonstantinou, Strelcyk, & Dau, (2011) studied the correlation of different perceptual and physiological measures with speech intelligibility in stationary noise. In their study, all the listeners had a steeply sloping high-frequency hearing loss consistent with Profile B. Both studies (Lőcsei et al., 2016; Papakonstantinou et al., 2011) included measures of frequency discrimination thresholds and speech-in-noise perception. However, Papakonstantinou et al. (2011) tested a larger group of HI listeners with fairly similar audiograms in only one speech condition that led to a highly significant correlation between frequency discrimination and speech intelligibility in stationary noise. This suggests that the stratification of the listeners and the investigation of certain phenomena in separated auditory profiles might reveal new knowledge about hearing impairments that are not generalized to the entire population of HI listeners.

Other approaches have attempted to identify why listeners with similar audiograms present substantial differences in suprathreshold performance. Recently, Souza, Gallun, & Wright (2020) showed how older HI listeners vary in terms of their ability to use specific cues (either spectral or temporal cues) for speech identification. Their results showed a so-called "profile cue" that characterizes the listener's abilities in terms of spectro-temporal processing. Some listeners utilized temporal envelope cues and showed good temporal discrimination abilities, whereas other listeners relied on spectral cues and were able to discriminate spectral modifications in a speech signal. The "profile cue" was associated with the spectral discrimination task; however, this test seemed to be influenced by the audiometric thresholds. Since their participants presented audiograms similar to the ones

observed in Profile B, C and D (Figure 6), it is possible that the categorization of the listeners based on auditory profiling could help identify "profile cues" in connection to the supra-threshold auditory deficits observed in each of the four auditory profiles reported here.

Hearing-aid users often show a large variability in terms of benefit and preference to specific forms of hearing-aid processing (Neher et al., 2016; Picou et al., 2015; Smeds et al., 2006; Souza et al., 2019). In some studies, the HI listeners were stratified based on their audiograms (Gatehouse et al., 2006; Keidser et al., 1995; Keidser & Grant, 2001; Larson et al., 2002). However, the existing hearing-aid fitting rules do not make use of supra-threshold auditory measures that might help tune the large parameter space of modern hearing technology. In fact, the HA parameters are still adjusted based on the audiogram and empirical findings that provide some fine-tuning according to the HA user experience or the gender of the patient (Keidser et al., 2012). The four auditory profiles presented here showed significant differences in supra-threshold measures related to two independent dimensions, a "speech intelligibility related distortion" and a "loudness perception related distortion". Therefore, auditory profiling allows stratification of the listeners beyond the audiogram, which may help optimize hearing-aid parameters for a given patient using existing HA technology. Recently, it has been suggested that different advanced signal processing strategies should be used to compensate for different cochlear pathologies (Henry et al., 2019). Since the four auditory profiles showed interesting similarities to the sensory and metabolic phenotypes (Dubno et al., 2013), new forms of signal processing, dedicated to overcoming the hearing deficits in the two identified dimensions, might be developed and evaluated towards a profile-based compensation strategy.

zsanchez_et_al_2020_RobustProfiling_medRxiv

Limitations of the data-driven auditory profiling approach

The data-driven method for auditory profiling presented here provides new knowledge about hearing loss characterization. Regarding the previous data-driven auditory profiling (Sanchez-Lopez et al., 2018), the present results are in better agreement with the analysis performed on the Johannesen et al. (2016) dataset than on the Thorup et al. (2016) dataset. Indeed, the decision tree, obtained from the analysis of the Johannesen et al. (2016) data showed that Profiles D and C exhibited a loss of cochlear nonlinearity at low frequencies and Profiles B and C exhibited a high-frequency hearing loss, which is similar to the first two levels of the decision tree presented here (see Figure 5), based on the analysis of the BEAR3 data set (Sanchez-Lopez et al., 2020). This suggests that the use of data from a representative sample of different degrees of hearing loss (e.g. in Johannesen et al., 2016) and a normal-hearing reference (e.g. in Thorup et al., 2016) is crucial for robust profile-based hearing-loss characterization.

The refined definition of the auditory profiles reflected the main sources of hearing deficits in a relatively large and heterogeneous population of HI listeners. However, this group only contained older adults (>60 years) with symmetric sensorineural hearing losses. An extension of the auditory profiling method proposed here might contain an even more heterogeneous group, which might require different data-analysis techniques for proper analysis and interpretation (Hinrich et al., 2016; Luo et al., 2017). The insights from the current method could then be applied mainly to a population of mild-to-severe age-related hearing losses and to some extent to other types of non-syndromic hearing losses, e.g.

noise-induced hearing loss, but cannot be generalized to the whole variability of existing auditory pathologies.

Conclusion

Using a data-driven approach, four auditory profiles (A-B-C-D) were identified that showed significant differences in terms of several supra-threshold auditory tasks. The listeners' hearing deficits could be characterized by two independent types of auditory distortion, a "speech intelligibility-related distortion" affecting listeners with audiometric thresholds >50 dB HL at high frequencies, and a "loudness perception-related distortion" affecting listeners with audiometric thresholds >30 dB HL at low frequencies. The four profiles showed similarities to the audiometric phenotypes proposed by Dubno et al. (2013), suggesting that Profile B may be resulting from a sensory loss and Profile D may be resulting from a metabolic loss. Profile C may reflect a combination of a sensory and metabolic loss, or a different type of hearing loss that results in substantially poorer supra-threshold auditory processing performance. The current approach shows promise for uncovering auditory phenomena that are usually hidden by, or entangled with, audibility loss. Moreover, the use of auditory profiling for hearing-aid fitting, where the chosen solutions are based not only on the loss of sensitivity but also other hearing deficits, may provide greater benefit to the user and thus enable precision audiology.

Acknowledgements

We thank F Bianchi, SG Nielsen, M El-haj-Ali, OM Cañete, and M Wu for their contribution to the design of the BEAR test battery, and work in the data collection. Also, we appreciate the valuable input from JB Nielsen, J Harte, W. Whitmer, G Naylor and O. Strelcyk among others. We would also like to thank G Encina-Llamas and the anonymous reviewers who suggested some improvements in an earlier version of this manuscript.

Declaration of Conflicting Interests

This work was supported by Innovation Fund Denmark Grand Solutions 5164-00011B (Better hEAring Rehabilitation project) Oticon, GN Hearing, WSAudiology and other partners (Aalborg University, University of Southern Denmark, the Technical University of Denmark, Force, Aalborg, Odense and Copenhagen University Hospitals). The funding and collaboration of all partners are sincerely acknowledged. The authors declare that there is no conflict of interest.

References

Ahroon, W. A., Davis, R. I., & Hamemik, R. P. (1993). The Role of Tuning Curve Variables and Threshold Measures in the Estimation of Sensory Cell Loss. *International Journal of Audiology*, *32*(4), 244–259. https://doi.org/10.3109/00206099309072940

Bernstein, J. G. W., Danielsson, H., Hällgren, M., Stenfelt, S., Rönnberg, J., & Lunner, T. (2016). Spectrotemporal Modulation Sensitivity as a Predictor of Speech-Reception

Performance in Noise With Hearing Aids. *Trends in Hearing*, *20*, 1–17. https://doi.org/10.1177/2331216516670387

Bisgaard, N., Vlaming, M. S. M. G. M. G., & Dahlquist, M. (2010). Standard Audiograms for the IEC 60118-15 Measurement Procedure. *Trends in Amplification*, *14*(2), 113–120. https://doi.org/10.1177/1084713810379609

Bramhall, N., Beach, E. F., Epp, B., Le Prell, C. G., Lopez-Poveda, E. A., Plack, C. J., Schaette, R., Verhulst, S., & Canlon, B. (2019). The search for noise-induced cochlear synaptopathy in humans: Mission impossible? *Hearing Research*, *377*, 88–103. https://doi.org/10.1016/j.heares.2019.02.016

Brand, T., & Hohmann, V. (2002). An adaptive procedure for categorical loudness scaling. *The Journal of the Acoustical Society of America*, *112*(4), 1597–1604. https://doi.org/10.1121/1.1502902

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (2017). Classification and regression trees. In *Classification and Regression Trees*. https://doi.org/10.1201/9781315139470

Chang, Y.-S., Yoon, S. H., Kim, J. R., Baek, S.-Y., Cho, Y. S., Hong, S. H., Kim, S., & Moon, I. J. (2019). Standard Audiograms for Koreans Derived through Hierarchical Clustering Using Data from the Korean National Health and Nutrition Examination Survey 2009–2012. *Scientific Reports*, *9*(1), 3675. https://doi.org/10.1038/s41598-019-40300-7

Cutler, A., & Breiman, L. (1994). Archetypal Analysis. *Technometrics*, *36*(4), 338–347. https://doi.org/10.1080/00401706.1994.10485840

D'Errico, J. (2020). *inpaint_nans*. MATLAB Central File Exchange. https://www.mathworks.com/matlabcentral/fileexchange/4551-inpaint_nans

Dubno, J. R., Eckert, M. A., Lee, F.-S., Matthews, L. J., & Schmiedt, R. A. (2013). Classifying Human Audiometric Phenotypes of Age-Related Hearing Loss from

Animal Models. *Journal of the Association for Research in Otolaryngology*, *14*(5), 687–701. https://doi.org/10.1007/s10162-013-0396-x

Dudoit, S., & Fridlyand, J. (2003). Bagging to improve the accuracy of a clustering procedure. *Bioinformatics*, *19*(9), 1090–1099. https://doi.org/10.1093/bioinformatics/btg038

Durlach, N. I. (1963). Equalization and Cancellation Theory of Binaural Masking-Level Differences. *The Journal of the Acoustical Society of America*, *35*(8), 1206–1218. https://doi.org/10.1121/1.1918675

Frawley, W. J., Piatetsky-Shapiro, G., & Matheus, C. J. (1992). Knowledge Discovery in Databases: An Overview. *AI Magazine*, *13*(3), 57. https://doi.org/10.1609/aimag.v13i3.1011

Füllgrabe, C., Harland, A. J., Sęk, A. P., & Moore, B. C. J. (2017). Development of a method for determining binaural sensitivity to temporal fine structure. *International Journal of Audiology*, *56*(12), 926–935. https://doi.org/10.1080/14992027.2017.1366078

Gatehouse, S., Naylor, G., & Elberling, C. (2006). Linear and nonlinear hearing aid fittings – 1. Patterns of benefit. *International Journal of Audiology*, *45*(3), 130–152. https://doi.org/10.1080/14992020500429518

Gerlach, M., Farb, B., Revelle, W., & Nunes Amaral, L. A. (2018). A robust data-driven approach identifies four personality types across four large data sets. *Nature Human Behaviour*, *2*(10), 735–742. https://doi.org/10.1038/s41562-018-0419-z

Hamernik, R. P., Patterson, J. H., Turrentine, G. A., & Ahroon, W. A. (1989). The quantitative relation between sensory cell loss and hearing thresholds. *Hearing Research*, *38*(3), 199–211. https://doi.org/10.1016/0378-5955(89)90065-8

Henry, K. S., Sayles, M., Hickox, A. E., & Heinz, M. G. (2019). Divergent Auditory Nerve Encoding Deficits Between Two Common Etiologies of Sensorineural Hearing Loss.

The Journal of Neuroscience, *39*(35), 6879–6887. https://doi.org/10.1523/JNEUROSCI.0038-19.2019

Hinrich, J. L., Bardenfleth, S. E., Roge, R. E., Churchill, N. W., Madsen, K. H., & Morup, M. (2016). Archetypal Analysis for Modeling Multisubject fMRI Data. *IEEE Journal of Selected Topics in Signal Processing*, *10*(7), 1160–1171. https://doi.org/10.1109/JSTSP.2016.2595103

ISO 8253-1. (2010). Acoustics - Audiometric test methods - Part 1: Pure-tone air and bone conduction audiometry. *International Organization for Standardization*. http://www.iso.org/iso/catalogue_detail.htm?csnumber=43601

ISO 8253-3. (2012). Acoustics. Audiometric test methods - Part 3: Speech audiometry. *International Organization for Standardization*. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=45101

Johannesen, P. T., Pérez-González, P., Kalluri, S., Blanco, J. L., & Lopez-Poveda, E. A. (2016). The Influence of Cochlear Mechanical Dysfunction, Temporal Processing Deficits, and Age on the Intelligibility of Audible Speech in Noise for Hearing-Impaired Listeners. *Trends in Hearing*, *20*, 233121651664105. https://doi.org/10.1177/2331216516641055

Keidser, G., Dillon, H., & Byrne, D. (1995). Candidates for Multiple Frequency Response Characteristics. *Ear and Hearing*, *16*(6), 562–574. https://doi.org/10.1097/00003446-199512000-00003

Keidser, G., Dillon, H., Carter, L., & O'Brien, A. (2012). NAL-NL2 Empirical Adjustments. *Trends in Amplification*, *16*(4), 211–223. https://doi.org/10.1177/1084713812468511

Keidser, G., & Grant, F. (2001). Comparing Loudness Normalization (IHAFF) with Speech Intelligibility Maximization (NAL-NL1) when Implemented in a Two-

zsanchez_et_al_2020_RobustProfiling_medRxiv

Channel Device. *Ear and Hearing*, *22*(6), 501–515. https://doi.org/10.1097/00003446-200112000-00006

Kiessling, J. (2001). Hearing aid fitting procedures - state-of-the-art and current issues. *Scandinavian Audiology*, *30*(1), 57–59. https://doi.org/10.1080/010503901300007074

Kollmeier, B., & Kiessling, J. (2018). Functionality of hearing aids: state-of-the-art and future model-based solutions. *International Journal of Audiology*, *57*(sup3), S3–S28. https://doi.org/10.1080/14992027.2016.1256504

Krohne, L. G., Wang, Y., Hinrich, J. L., Moerup, M., Chan, R. C. K., & Madsen, K. H. (2019). Classification of social anhedonia using temporal and spatial network features from a social cognition fMRI task. *Human Brain Mapping*, *40*(17), 4965–4981. https://doi.org/10.1002/hbm.24751

Kujawa, S. G., & Liberman, M. C. (2009). Adding Insult to Injury: Cochlear Nerve Degeneration after "Temporary" Noise-Induced Hearing Loss. *Journal of Neuroscience*, *29*(45), 14077–14085. https://doi.org/10.1523/JNEUROSCI.2845-09.2009

Lansbergen, S., & Dreschler, W. A. (2020). Classification of Hearing Aids Into Feature Profiles Using Hierarchical Latent Class Analysis Applied to a Large Dataset of Hearing Aids. *Ear & Hearing*, *[in press]*. https://doi.org/10.1097/AUD.0000000000000877

Larson, V. D., Williams, D. W., Henderson, W. G., Luethke, L. E., Beck, L. B., Noffsinger, D., Bratt, G. W., Dobie, R. A., Fausti, S. A., Haskell, G. B., Rappaport, B. Z., Shanks, J. E., Wilson, R. H., VD, L., DW, W., WG, H., LE, L., LB, B., Noffsinger, D., … Wilson, R. H. (2002). A Multi-Center, Double Blind Clinical Trial Comparing Benefit from Three Commonly Used Hearing Aid Circuits*. *Ear and Hearing*, *23*(4), 269–276. https://doi.org/10.1097/00003446-200208000-00001

zsanchez_et_al_2020_RobustProfiling_medRxiv

Lobarinas, E., Salvi, R., & Ding, D. (2013). Insensitivity of the audiogram to carboplatin induced inner hair cell loss in chinchillas. *Hearing Research*, *302*, 113–120. https://doi.org/10.1016/j.heares.2013.03.012

Lőcsei, G., Pedersen, J. H., Laugesen, S., Santurette, S., Dau, T., & MacDonald, E. N. (2016). Temporal Fine-Structure Coding and Lateralized Speech Perception in Normal-Hearing and Hearing-Impaired Listeners. *Trends in Hearing*, *20*(0), 233121651666096. https://doi.org/10.1177/2331216516660962

Lopez-Poveda, E. A. (2014). Why do I hear but not understand? Stochastic undersampling as a model of degraded neural encoding of speech. *Frontiers in Neuroscience*, *8*(October), 1–7. https://doi.org/10.3389/fnins.2014.00348

Luo, Y., Ahmad, F. S., & Shah, S. J. (2017). Tensor Factorization for Precision Medicine in Heart Failure with Preserved Ejection Fraction. *Journal of Cardiovascular Translational Research*, *10*(3), 305–312. https://doi.org/10.1007/s12265-016-9727-8

Mellor, J. C., Stone, M. A., & Keane, J. (2018). Application of Data Mining to "Big Data" Acquired in Audiology: Principles and Potential. *Trends in Hearing*, *22*, 233121651877681. https://doi.org/10.1177/2331216518776817

Mills, J., Schmiedt, R., Schulte, B., & Dubno, J. (2006). Age-Related Hearing Loss: A Loss of Voltage, Not Hair Cells. *Seminars in Hearing*, *27*(4), 228–236. https://doi.org/10.1055/s-2006-954849

Moore, B. C. J. (2001). Dead Regions in the Cochlea: Diagnosis, Perceptual Consequences, and Implications for the Fitting of Hearing Aids. *Trends in Amplification*, *5*(1), 1–34. https://doi.org/10.1177/108471380100500102

Moore, B. C. J. (2016). A review of the perceptual effects of hearing loss for frequencies above 3 kHz. *International Journal of Audiology*, *55*(July), 707–714. https://doi.org/10.1080/14992027.2016.1204565

Moore, B. C. J., Huss, M., Vickers, D. A., Glasberg, B. R., & Alcántara, J. I. (2000). A test

zsanchez_et_al_2020_RobustProfiling_medRxiv

for the diagnosis of dead regions in the cochlea. *British Journal of Audiology*, *34*(4), 205–224.

Mørup, M., & Hansen, L. K. (2012). Archetypal analysis for machine learning and data mining. *Neurocomputing*, *80*, 54–63. https://doi.org/10.1016/j.neucom.2011.06.033

Neher, T., Wagener, K. C., & Fischer, R.-L. (2016). Directional Processing and Noise Reduction in Hearing Aids: Individual and Situational Influences on Preferred Setting. *Journal of the American Academy of Audiology*, *27*(8), 628–646. https://doi.org/10.3766/jaaa.15062

Nielsen, J. B., & Dau, T. (2011). The Danish hearing in noise test. *International Journal of Audiology*, *50*(3), 202–208. https://doi.org/10.3109/14992027.2010.524254

Papakonstantinou, A., Strelcyk, O., & Dau, T. (2011). Relations between perceptual measures of temporal processing, auditory-evoked brainstem responses and speech intelligibility in noise. *Hearing Research*, *280*(1–2), 30–37. https://doi.org/10.1016/j.heares.2011.02.005

Pauler, M., Schuknecht, H. F., & Thornton, A. R. (1986). Correlative studies of cochlear neuronal loss with speech discrimination and pure-tone thresholds. *Archives of Oto-Rhino-Laryngology*, *243*(3), 200–206. https://doi.org/10.1007/BF00470622

Picou, E. M., Marcrum, S. C., & Ricketts, T. A. (2015). Evaluation of the effects of nonlinear frequency compression on speech recognition and sound quality for adults with mild to moderate hearing loss. *International Journal of Audiology*, *54*(3), 162–169. https://doi.org/10.3109/14992027.2014.961662

Plomp, R. (1978). Auditory handicap of hearing impairment and the limited benefit of hearing aids. *The Journal of the Acoustical Society of America*, *63*(2), 533–549. https://doi.org/10.1121/1.381753

Plomp, R. (1986). A Signal-to-Noise Ratio Model for the Speech-Reception Threshold of the Hearing Impaired. *Journal of Speech, Language, and Hearing Research*, *29*(2),

146–154. https://doi.org/10.1044/jshr.2902.146

Ragozini, G., Palumbo, F., & D'Esposito, M. R. (2017). Archetypal analysis for data-driven prototype identification. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, *10*(1), 6–20. https://doi.org/10.1002/sam.11325

Rieke, C. C., Clavier, O. H., Allen, L. V., Anderson, A. P., Brooks, C. A., Fellows, A. M., Brungart, D. S., & Buckey, J. C. (2017). Fixed-Level Frequency Threshold Testing for Ototoxicity Monitoring. *Ear and Hearing*, *38*(6), e369–e375. https://doi.org/10.1097/AUD.0000000000000433

Sanchez-Lopez, R., Bianchi, F., Fereczkowski, M., Santurette, S., & Dau, T. (2018). Data-Driven Approach for Auditory Profiling and Characterization of Individual Hearing Loss. *Trends in Hearing*, *22*(1), 233121651880740. https://doi.org/10.1177/2331216518807400

Sanchez-Lopez, R., Nielsen, S. G., El-Haj-Ali, M., Bianchi, F., Fereckzowski, M., Cañete, O. M., Wu, M., Neher, T., Dau, T., & Santurette, S. (2020). Auditory tests for characterizing hearing deficits: The BEAR test battery. *MedRxiv*. https://doi.org/10.1101/2020.02.17.20021949

Sanchez-Lopez, R., Nielsen, S. G., El-Haj-Ali, M., Bianchi, F., Fereckzowski, M., Cañete, O., Wu, M., Neher, T., Dau, T., & Santurette, S. (2019). *Data from "Auditory tests for characterizing hearing deficits: The BEAR test battery"* (v1.0). Zenodo; Zenodo. https://doi.org/10.5281/zenodo.3459579

Santurette, S., & Dau, T. (2012). Relating binaural pitch perception to the individual listener's auditory profile. *The Journal of the Acoustical Society of America*, *131*(4), 2968–2986. https://doi.org/10.1121/1.3689554

Schädler, M. R., Hülsmeier, D., Warzybok, A., & Kollmeier, B. (2020). Individual aided speech recognition performance and benefit predictions for listeners with impaired hearing. *Trends in Hearing, [in press]*.

Schmiedt, R. A., Lang, H., Okamura, H. O., & Schulte, B. A. (2002). Effects of furosemide applied chronically to the round window: A model of metabolic presbyacusis. *Journal of Neuroscience*, *22*(21), 9643–9650. https://doi.org/10.1523/jneurosci.22-21-09643.2002

Schorn, K., & Zwicker, E. (1990). Frequency selectivity and temporal resolution in patients with various inner ear disorders. *Audiology : Official Organ of the International Society of Audiology*, *29*(1), 8–20. https://doi.org/10.3109/00206099009081641

Schuknecht, H. F., & Gacek, M. R. (1993). Cochlear Pathology in Presbycusis. *Annals of Otology, Rhinology & Laryngology*, *102*(1_suppl), 1–16. https://doi.org/10.1177/00034894931020S101

Sharma, M., Purdy, S. C., & Humburg, P. (2019). Cluster Analyses Reveals Subgroups of Children With Suspected Auditory Processing Disorders. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02481

Smeds, K., Keidser, G., Zakis, J., Dillon, H., Leijon, A., Grant, F., Convery, E., & Brew, C. (2006). Preferred overall loudness. II: Listening through hearing aids in field and laboratory tests. *International Journal of Audiology*, *45*(1), 12–25. https://doi.org/10.1080/14992020500190177

Souza, P., Arehart, K., Schoof, T., Anderson, M., Strori, D., & Balmert, L. (2019). Understanding Variability in Individual Response to Hearing Aid Signal Processing in Wearable Hearing Aids. *Ear and Hearing*, *40*(6), 1280–1292. https://doi.org/10.1097/AUD.0000000000000717

Souza, P., Gallun, F., & Wright, R. (2020). Contributions to Speech-Cue Weighting in Older Adults With Impaired Hearing. *Journal of Speech, Language, and Hearing Research*, *63*(1), 334–344. https://doi.org/10.1044/2019_JSLHR-19-00176

Stebbins, W. C., Hawkins, J. E., Johnsson, L.-G., & Moody, D. B. (1979). Hearing thresholds with outer and inner hair cell loss. *American Journal of Otolaryngology*,

1(1), 15–27. https://doi.org/10.1016/S0196-0709(79)80004-6

Strelcyk, O., Zahorik, P., Shehorn, J., Patro, C., & Derleth, R. P. (2019). Sensitivity to Interaural Phase in Older Hearing-Impaired Listeners Correlates With Nonauditory Trail Making Scores and With a Spatial Auditory Task of Unrelated Peripheral Origin. *Trends in Hearing*, *23*, 233121651986449. https://doi.org/10.1177/2331216519864499

Thorup, N., Santurette, S., Jørgensen, S., Kjærbøl, E., Dau, T., & Friis, M. (2016). Auditory profiling and hearing-aid satisfaction in hearing-aid candidates. *Danish Medical Journal*, *63*(10), 1–5. http://www.ncbi.nlm.nih.gov/pubmed/27697129

Trusheim, M. R., Berndt, E. R., & Douglas, F. L. (2007). Stratified medicine: strategic and economic implications of combining drugs and clinical biomarkers. *Nature Reviews Drug Discovery*, *6*(4), 287–293. https://doi.org/10.1038/nrd2251

van Esch, T. E. M., Kollmeier, B., Vormann, M., Lyzenga, J., Houtgast, T., Hällgren, M., Larsby, B., Athalye, S. P., Lutman, M. E., & Dreschler, W. A. (2013). Evaluation of the preliminary auditory profile test battery in an international multi-centre study. *International Journal of Audiology*, *52*(5), 305–321. https://doi.org/10.3109/14992027.2012.759665

Wu, P., Wen, W., O'Malley, J. T., & Liberman, M. C. (2020). Assessing fractional hair cell survival in archival human temporal bones. *The Laryngoscope*, *130*(2), 487–495. https://doi.org/10.1002/lary.27991

Wu, P. Z., Liberman, L. D., Bennett, K., de Gruttola, V., O'Malley, J. T., & Liberman, M. C. (2019). Primary Neural Degeneration in the Human Cochlea: Evidence for Hidden Hearing Loss in the Aging Ear. *Neuroscience*, *407*, 8–20. https://doi.org/10.1016/j.neuroscience.2018.07.053

zsanchez_et_al_2020_RobustProfiling_medRxiv