

Data-Driven Approach for Auditory Profiling and Characterization of Individual Hearing Loss

Trends in Hearing
Volume 22: 1–12
© The Author(s) 2018
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/2331216518807400
journals.sagepub.com/home/tia



Raul Sanchez Lopez¹ , Federica Bianchi¹, Michal Fereczkowski¹, Sébastien Santurette^{1,2}, and Torsten Dau¹

Abstract

Pure-tone audiometry still represents the main measure to characterize individual hearing loss and the basis for hearing-aid fitting. However, the perceptual consequences of hearing loss are typically associated not only with a loss of sensitivity but also with a loss of clarity that is not captured by the audiogram. A detailed characterization of a hearing loss may be complex and needs to be simplified to efficiently explore the specific compensation needs of the individual listener. Here, it is hypothesized that any listener's hearing profile can be characterized along two dimensions of distortion: Type I and Type II. While Type I can be linked to factors affecting audibility, Type II reflects non-audibility-related distortions. To test this hypothesis, the individual performance data from two previous studies were reanalyzed using an unsupervised-learning technique to identify extreme patterns in the data, thus forming the basis for different auditory profiles. Next, a decision tree was determined to classify the listeners into one of the profiles. The analysis provides evidence for the existence of four profiles in the data. The most significant predictors for profile identification were related to binaural processing, auditory nonlinearity, and speech-in-noise perception. This approach could be valuable for analyzing other data sets to select the most relevant tests for auditory profiling and propose more efficient hearing-deficit compensation strategies.

Keywords

auditory profile, data-driven, hearing loss, suprathreshold auditory perception

Date received: 9 February 2018; revised: 12 September 2018; accepted: 17 September 2018

Introduction

Currently, the pure-tone audiogram is the main tool used to characterize the degree of hearing loss and for hearing-aid fitting. However, the perceptual consequences of hearing loss are typically associated not only with a loss of sensitivity, as reflected by the audiogram, but also with a loss of clarity that is not captured by the audiogram (e.g., Killion & Niquette, 2000). This loss of clarity may be associated with distortions in the auditory processing of suprathreshold sounds. Although amplification can effectively compensate for the loss of sensitivity, suprathreshold distortions may require more advanced signal processing to overcome the loss of clarity and improve speech intelligibility, particularly in complex acoustic conditions (e.g., Kollmeier & Kiessling, 2016; Plomp, 1978). Plomp (1978) suggested that a hearing loss can be divided into two components: an attenuation component and a distortion component.

When a pure attenuation loss, also referred to as audibility loss or sensitivity loss, is compensated for by amplification, the speech reception threshold in stationary noise (SRT_N) is similar to that of a normal-hearing (NH) listener. In contrast, when a distortion component is present, SRT_N remains elevated despite amplification.

Several studies have attempted to shed light on the potential mechanisms underlying suprathreshold distortions (e.g., Glasberg & Moore, 1989; Houtgast & Festen, 2008; Strelcyk & Dau, 2009; Summers, Makashay,

¹Hearing Systems Group, Department of Electrical Engineering, Technical University of Denmark, Kongens Lyngby, Denmark

²Department of Otorhinolaryngology, Head and Neck Surgery and Audiology, Rigshospitalet, Copenhagen, Denmark

Corresponding author:

Raul Sanchez Lopez, Hearing Systems Group, Ørstedes Plads 352, Kongens Lyngby 2800, Denmark.

Email: rsalo@elektro.dtu.dk



Theodoroff, & Leek, 2013; Johannesen, Pérez-González, Kalluri, Blanco, & Lopez-Poveda, 2016). In these studies, different psychoacoustic tests were considered in listeners with various degrees of hearing loss in an attempt to explain the variance observed in the listeners' speech-in-noise intelligibility performance. It was suggested that, beyond pure-tone audiometry, an elevated SRT_N could be associated with outcome measures related to spectral or temporal processing deficits. The suprathreshold distortions relevant for speech intelligibility in noise may thus reflect inaccuracies in the coding and representation of spectral or temporal stimulus features in the auditory system. To achieve the optimal compensation strategy for the individual hearing-impaired (HI) listener, a characterization of the listener's hearing deficits in terms of audibility loss as well as clarity loss seems essential.

Large-scale studies have attempted to establish a new hearing profile based on test batteries involving suprathreshold outcome measures in addition to pure-tone audiometry. As a part of the European project HEARCOM (van Esch & Dreschler 2015; van Esch et al., 2013; Vlaming et al., 2011), new screening tests were proposed, as well as a test battery designed for assessing the specific hearing deficits of the patients. The factor analysis performed in a study with 72 HI subjects revealed that the test outcomes could be grouped in four dimensions: audibility, high-frequency processing, low-frequency processing, and recruitment (Vlaming et al., 2011). Rönnberg et al. (2016) also used a factor analysis to explore the relations between hearing, cognition, and speech-in-noise intelligibility in a large-scale study with 200 listeners. Even though the sensitivity loss was still a dominant factor, the new test battery provided information about suprathreshold processing using new outcome measures. However, these additional outcome measures were highly cross-correlated, which complicated the factor analysis. Although the earlier studies explored the relative importance of diverse factors in the individual subject, the interpretation of an individual hearing profile became more complex, particularly because the clinical tests were highly interrelated.

Other studies have suggested strategies for classifying HI listeners based on a characterization of their hearing deficits. In one approach, Lopez-Poveda (2014) reviewed the mechanisms associated with hearing loss and their perceptual consequences for speech. In a two-dimensional space, the hearing loss was considered to represent the sum of an outer hair cell (OHC) loss and an inner hair cell (IHC) loss (Lopez-Poveda & Johannesen, 2012). The importance of this distinction is related to the way these mechanisms affect speech. Although the OHC loss has been associated with audibility loss and reduced frequency selectivity, the IHC loss may yield a loss of clarity and temporal processing deficits (Killion & Niquette,

2000). However, since OHC and IHC loss can only be estimated by indirect outcome measures (Jürgens, Kollmeier, Brand, & Ewert, 2011; Lopez-Poveda & Johannesen, 2012), and since pure-tone audiometry only reflects the mixed effects of OHC and IHC loss (Moore, Vickers, Plack, & Oxenham, 1999), this approach seems limited in terms of an individual hearing-loss characterization in a clinical setting. Another approach was presented by Dubno, Eckert, Lee, Matthews, and Schmiedt (2013), who suggested four audiometric phenotypes to account for age-related hearing loss. The phenotypes were proposed based on animal models with either a metabolic or a sensory impairment. Using a large database of audiograms from older humans, the corresponding human exemplars of the four audiometric phenotypes were identified by an expert researcher. Finally, a classifier trained on these exemplars was used to classify the remaining audiograms into the audiometric phenotypes. Although the audiometric phenotypes can be linked to the underlying mechanism of the hearing loss, a limitation of this approach is that it is fully based on the information provided by the audiogram. Hence, suprathreshold distortions are not, or only partly, reflected in this classification.

Inspired by the studies of Lopez-Poveda (2014) and Dubno et al. (2013), a two-dimensional approach was also considered in this study. However, in contrast to these previous approaches, the classification of the listeners was mainly based on perceptual outcome measures, rather than physiological indicators of hearing loss. Although the physiological indicators, such as OHC and IHC loss, cannot be assessed directly in humans, it is likely that their corresponding perceptual distortions can be quantified by means of psychoacoustic tests. The aims of this study were (a) to achieve a new hearing-loss characterization strategy that takes suprathreshold hearing performance into account and is based on functional tests reflecting auditory perception and (b) to propose a new statistical analysis protocol that can be used to reanalyze existing data sets to improve and optimize such a characterization.

It was hypothesized here that any listener's hearing can be characterized along two independent dimensions: Distortion Type I and Distortion Type II, as indicated in Figure 1. Distortion Type I was hypothesized to reflect deficits that have been found to covary with a loss of audibility, such as a loss of frequency selectivity and of cochlear compression (Moore et al., 1999). Distortion Type II was hypothesized to reflect deficits that typically do not covary with audibility loss and may be related to inaccuracies in terms of temporal coding according to the conclusions from other studies (Strelcyk & Dau, 2009; Summers et al., 2013; Johannesen et al., 2016). The two dimensions can be roughly defined as audibility-related and non-audibility-related distortions. In this

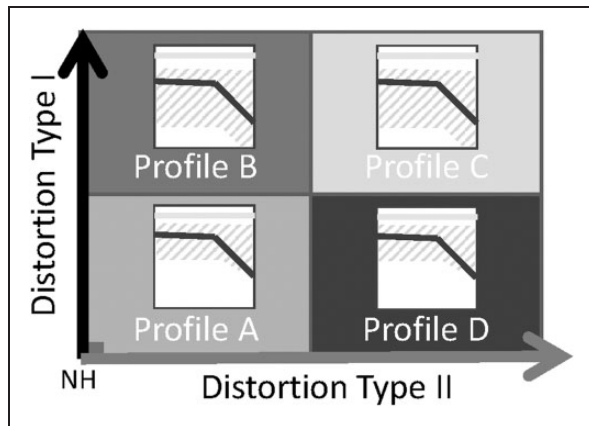


Figure 1. Sketch of the hypothesis. Hearing deficits arise from two independent types of distortions. Distortion Type I: distortions that accompany the loss of sensitivity. Distortion Type II: distortions that do not covary with sensitivity loss. Profile A: low distortion for both types. Profile B: high Type I distortion and low Type II distortion. Profile C: High distortion for both types. Profile D: low Type I distortion and high Type II distortion. NH = normal hearing.

two-dimensional space, NH listeners are placed in the bottom-left corner and defined as not exhibiting any type of distortion. Then, four profiles may thus be identified, depending on the extent to which each type of distortion is present in the individual listener (Figure 1).

To test this hypothesis, a new data-driven statistical method is proposed here and used to reanalyze two existing data sets and exploit the individual differences of HI listeners in terms of perceptual outcome measures. In line with the hypothesis, the method divides different perceptual measures into two independent dimensions. Next, the method identifies patterns in the data, hence the analysis is considered data-driven. The approach is similar to the one used to identify the four audiometric phenotypes in Dubno et al. (2013) but considers additional outcome measures beyond audiometry for the classification of the listeners into the four auditory profiles. The proposed statistical analysis is based on an archetypal analysis (Cutler & Breiman, 1994), an unsupervised learning method that is particularly useful for identifying patterns in data, and has been suggested for prototyping and benchmarking purposes (Ragozini, Palumbo, & D'Esposito, 2017). The main advantage of using unsupervised learning in terms of auditory profiling is that the analysis involves the performance of the listener in different tests, in contrast to correlations between single tests or regression analyses (e.g., Glasberg & Moore, 1989; Houtgast & Festen, 2008; Summers et al., 2013) which explore relations between various hearing disabilities in a HI population, rather than in an individual listener. The novel method was evaluated by reanalyzing the data from two previous studies presented in Thorup et al. (2016) and Johannesen et al.

(2016). In both studies, an extensive auditory test battery was proposed and tested in HI listeners, with the aim to better characterize hearing deficits. While the analysis of those studies focused on finding correlates of speech intelligibility in noise and hearing-aid benefit, the goal here was to further define the two hypothesized distortion types and identify which outcome measures are most relevant to classify listeners into the four suggested auditory profiles.

Method

The data-driven approach was conducted in two stages (Figure 2). First, unsupervised learning was used to identify the trends in the data that could be used to estimate the amount of each distortion type in individual listeners and thus categorize the listeners into different profiles. The second stage consisted of supervised learning. Once the subjects were assigned to a profile, the data were analyzed again to find the best classification structure that could predict the identified profile.

Unsupervised Learning

Unsupervised learning aims to identify patterns occurring in the data, where the output is unknown and the statistical properties of the whole data set are explored (Cutler & Breiman, 1994). In contrast to a regression analysis, unsupervised learning does not aim to predict a specific output, for example, speech intelligibility. In the present approach, the identified auditory profiles were eventually inferred using various unsupervised learning techniques. First, a list of outcome measures obtained from different tests in the reanalyzed study of interest was selected as the input to the unsupervised learning stage.

As two types of distortions to characterize the individual hearing loss were assumed, a principal component analysis (PCA) was run as the first step of the data analysis to reduce the dimensionality of the data. The reduction was done by keeping the variables that were strongly correlated to the first principal component (PC1) or the second principal component (PC2). These variables were used in the further analysis rather than using directly the PCs (Figure 2(I)). Therefore, a dimensionality reduction algorithm was implemented as follows: The optimal subset of variables that suggested a strong relation with each of the two principal components was chosen using a leave-one-out cross-validation in an iterative PCA. In each iteration, a single variable was left out of the subset, and the variance explained by the two principal components was recalculated for the remaining set of variables. If the variance increased, the outcome measure that was left out in this iteration was discarded. This process was repeated until either the variance explained was higher than 90% or the number of variables was

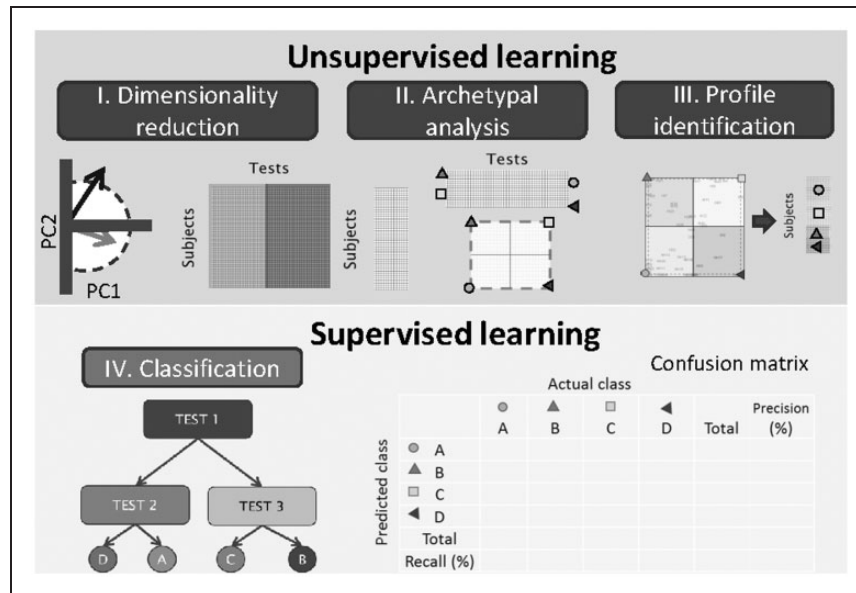


Figure 2. Sketch of the method considered in this study. The upper panel shows the unsupervised learning techniques applied to the whole data set. The bottom panel shows the supervised learning method, which uses the original data as the input and the identified profiles from the archetypal analysis as the output. PC = principal component.

lower than eight (four in each dimension). This reduction in number of variables ensures that the variables are balanced with regard to both distortion types and the chosen variables are connected with the hypothesis in an unsupervised process.

An archetypal analysis (Cutler & Breiman, 1994) was then performed on the output of the dimensionality reduction stage (Figure 2(II)). This technique combines the characteristics of matrix factorization and cluster analysis. In this study, an algorithm similar to the one described in Mørup and Hansen (2012) was used. The aim of this analysis was to identify extreme patterns in the data (*archetypes*). As a result, the listeners were no longer defined by the performance in each of the tests but by their similarity to the extreme exemplars present in the data, i.e., the archetypes.

Based on the archetypal analysis, the subjects were placed in a simplex plot (square visualization) to perform profile identification (Figure 2(III)). In such a plot, the archetypes are located at each corner, and the listeners are placed in the two-dimensional space according to the distance to each archetype. In the present analysis, it was assumed that the subjects placed close to an archetype would belong to the same cluster. Consequently, each subject was labeled according to the nearest archetype.

Supervised Learning

Once the profiles were identified, supervised learning could be performed. The purpose of this stage was to explore the accuracy of a classification scheme that

makes use of only few variables (here, outcome measures from different tests of auditory function). The joint probability density of the data set and the output (i.e., the identified profiles) could then be used to select the most relevant tests for the classification of the subjects into the four auditory profiles.

Decision trees were used to classify each individual listener (Figure 2(IV)). Here, each relevant outcome measure was used in the nodes forming a logical expression and dividing the observations accordingly. As a decision tree needs to be trained with a subset of the data and a known output, the identified auditory profiles (Figure 2(III)) were used as the response variable and a five-fold cross-validation was used to train the classifier. In the cross-validation procedure, the data were randomly divided into five segments. Four segments were used to train the classifier and the remaining one was used for testing. This was done iteratively 10 times. The decision tree which provided the minimum test error was used as the “optimal classifier.” In addition, the optimal decision tree was *pruned* to only have three nodes. This ensured that an efficient classification of the listeners based on only three tests could be considered in future clinical protocols.

Description of the Data Sets

In this study, the data from Thorup et al. (2016) (Study 1) and Johannesen et al. (2016) (Study 2) were reanalyzed with the unsupervised and supervised learning techniques described earlier.

The data set from Study 1 contained 59 listeners, among which 26 listeners had NH thresholds, 29 listeners were hearing impaired, and 4 had been previously identified as suffering from obscure dysfunction, that is, with NH thresholds but self-reports of hearing difficulties. The total number of variables (outcome measures from the different tests) considered in the analysis was 27. The variables used in the analysis were as follows (for details, see Thorup et al., 2016):

- Audiometric thresholds at low (HL_{LF}) and high frequencies (HL_{HF}).
- Spectral (MR_{spec}) and temporal (MR_{temp}) resolution at low and high frequencies.
- Binaural temporal fine structure (TFS) processing measured by interaural phase difference (IPD) frequency thresholds.
- Speech recognition thresholds in stationary (SRT_N) and fluctuating (SRT_{ISTS}) noise.
- Reading-span test of working memory.

The results of additional tests, not reported in Thorup et al. (2016) but collected in the same listeners, were also included in the present analysis:

- Binaural pitch test, using a procedure adapted from Santurette and Dau (2012) measuring detection of pitch contours presented either diotically or dichotically. The variables used here were Bp_{dicho} (percent correct for dichotic presentations only) and Bp_{total} (percent correct for the total number of presentations, i.e., diotic and dichotic).
- Speech reception threshold in quiet (SRT_Q) and discrimination scores (DS) using the Dantale I (Elberling, Ludvigsen, & Lyregaard, 1989) speech material.
- Most comfortable level (MCL) and lower slope of the growth of loudness at low and high frequencies, determined from adaptive categorical loudness scaling (ACALOS, Brand & Hohmann, 2002)

The data set from Johannesen et al. (2016) (Study 2) contained 67 HI listeners. The total number of variables considered in the analysis was 11:

- Audiometric thresholds at low (HL_{LF}) and high frequencies (HL_{HF}).
- Aided speech recognition thresholds in stationary noise ($HINT_{SSN}$) and reversed two-talker masker ($HINT_{R2TM}$).
- Frequency modulation detection threshold (FMDT).
- Basilar membrane compression (BM Comp) and OHC and IHC loss estimated from the results of a temporal masking curve experiment (Nelson, Schroder, & Wojtczak, 2001). These three variables

were each divided into a high-frequency and a low-frequency estimate.

Preprocessing of the Data Sets

For both data sets, the performance in each outcome measure was normalized such that the 25th percentile equaled -0.5 and the 75th percentile equaled $+0.5$. To more easily compare the tests, a *good* performance thus always corresponded to a positive number and a *poor* performance corresponded to a negative number.

The tests that corresponded to measures taken at different frequencies, for example, pure-tone audiometry, were reduced to the mean at low frequencies (≤ 1 kHz) and at high frequencies (> 1 kHz).¹ In addition, when the tests were performed in more than one ear, the average between the two ears was used as the outcome measure.² Listeners that did not complete more than three of the considered tests were excluded from the analysis. Furthermore, an artificial observation with an optimal performance (+1) in all tests was created, which served as an ideal NH reference that did not exhibit any type of distortion. This observation was always the Archetype A, located in the origin of coordinates of the hypothesis stated in Figure 1. The preprocessing was performed identically for both data sets.

Results

The two studies were analyzed using an identical method. For convenience, results corresponding to the reanalysis of the data from Thorup et al. (2016) are referred to using the Subindex 1, for example, Profile A_1 . The results from the reanalysis of the Johannesen et al. (2016) data are referred to using Subindex 2, for example, Profile A_2 . For general mentions of an auditory profile, no subindex is added. The whole data set was reduced to the variables that were strongly correlated to Dimension I (PC1) or Dimension II (PC2), as summarized in Table 1.

For Study 1, the dimensionality reduction revealed that the performance in binaural tests was largely independent of hearing thresholds, suggesting that Dimension II may be related to binaural processing abilities and Dimension I to audibility at low and high frequencies. The PCA could explain 80.3% of the variance in the performance for different hearing tests with only two components, with 63.0% explained by PC1 and 17.3% by PC2.

For Study 2, Dimension II was more dominated by low-frequency processing abilities and Dimension I by high-frequency processing abilities. The PCA could explain 67.8% of the variance in the performance for

Table 1. Results From the Dimensionality Reduction of the Two Data Sets.

Study I: Thorup et al. (2016)				Study II: Johannesen et al. (2016)			
Variable	Test	PC1	PC2	Variable	Test	PC1	PC2
HL _{LF}	Hearing loss at low frequencies	0.45	-0.03	HL _{HF}	Hearing loss at high frequencies	0.51	0.08
HL _{HF}	Hearing loss at high frequencies	0.41	-0.22	OHC loss _{HF}	Outer hair cell loss estimated at high frequencies	0.55	0.05
SRT _Q	Speech reception threshold (SRT) in quiet	0.46	-0.01	IHC loss _{HF}	Inner hair cell loss estimated at high frequencies	0.37	-0.03
SRT _{ISTS}	SRT in noise using international speech test signal	0.47	-0.17	BM Comp _{HF}	Basilar membrane compression at high frequencies	0.51	-0.01
DS	Word discrimination score	0.33	-0.24	HL _{LF}	Hearing loss at low frequencies	-0.00	0.62
MCL _{LF}	Most comfortable level at low frequencies	0.14	0.46	FMDT	Frequency modulation discrimination threshold	-0.02	0.42
Bp _{dicho}	BP dichotic condition	0.20	0.53	OHC loss _{LF}	Outer hair cell loss estimated at low frequencies	0.03	0.45
Bp _{tot}	BP diotic + dichotic	0.16	0.61	IHC loss _{LF}	Inner hair cell loss estimated at low frequencies	-0.11	0.45

Note. The table includes variables strongly correlated to PC1 (Distortion Type I, top four rows) and PC2 (Distortion Type II, bottom four rows) and their correlation coefficient obtained from the loadings of the principal component analysis.

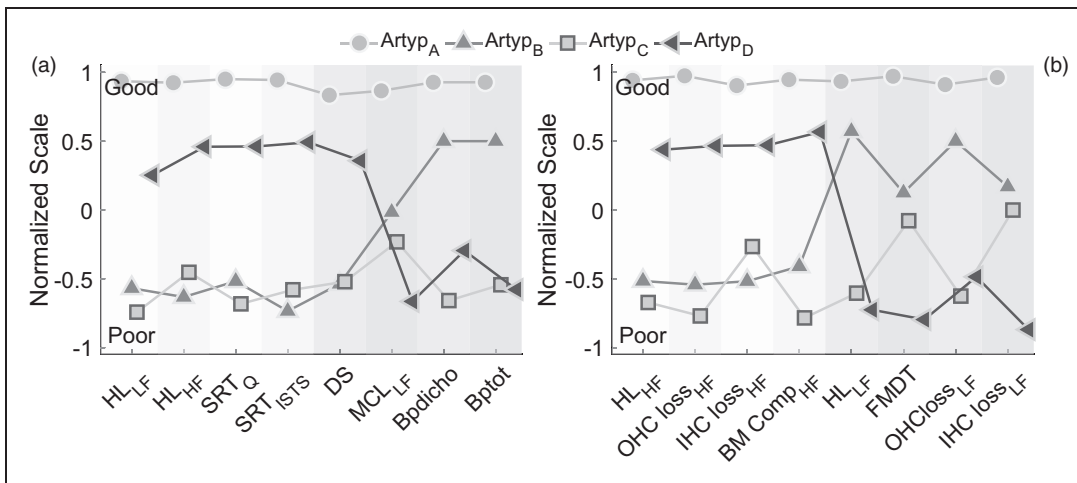


Figure 3. Archetypes (Artp): Extreme exemplars of the different patterns found in the data. (a) Normalized performance of each of the four archetypes from Study I. (b) The same for Study 2. The variables are divided according to Table 1. The first four variables correspond to Distortion Type I and the remaining four to Distortion Type II. HL_{LF} = hearing loss at low frequencies; HL_{HF} = hearing loss at high frequencies; SRT_Q = speech reception threshold (SRT) in quiet; SRT_{ISTS} = SRT in noise using international speech test signal; DS = word discrimination score; MCL_{LF} = most comfortable level at low frequencies; Bp_{dicho} = BP dichotic condition; Bp_{tot} = BP diotic + dichotic; OHC loss_{HF} = outer hair cell loss estimated at high frequencies; IHC loss_{HF} = inner hair cell loss estimated at high frequencies; BM Comp_{HF} = basilar membrane compression at high frequencies; FMDT = frequency modulation discrimination threshold; OHC loss_{LF} = outer hair cell loss estimated at low frequencies; IHC loss_{LF} = inner hair cell loss estimated at low frequencies.

the behavioral tasks with the chosen variables, with 37.2% explained by PC1 and 30.6% by PC2.

The archetypal analysis was used to identify four archetypes using the variables from Table 1. As shown in Figure 3, in both studies, Profile A (Archetype A) exhibited the best performance in both dimensions and

Profile C the worst. Profile B showed poor performance only in Dimension I, while Profile D showed poor performance only in Dimension II.

Figure 3(a) illustrates the four archetypes from Study 1. The performance in the tests related to Distortion Type I was clearly good for archetypes A₁ and D₁ and

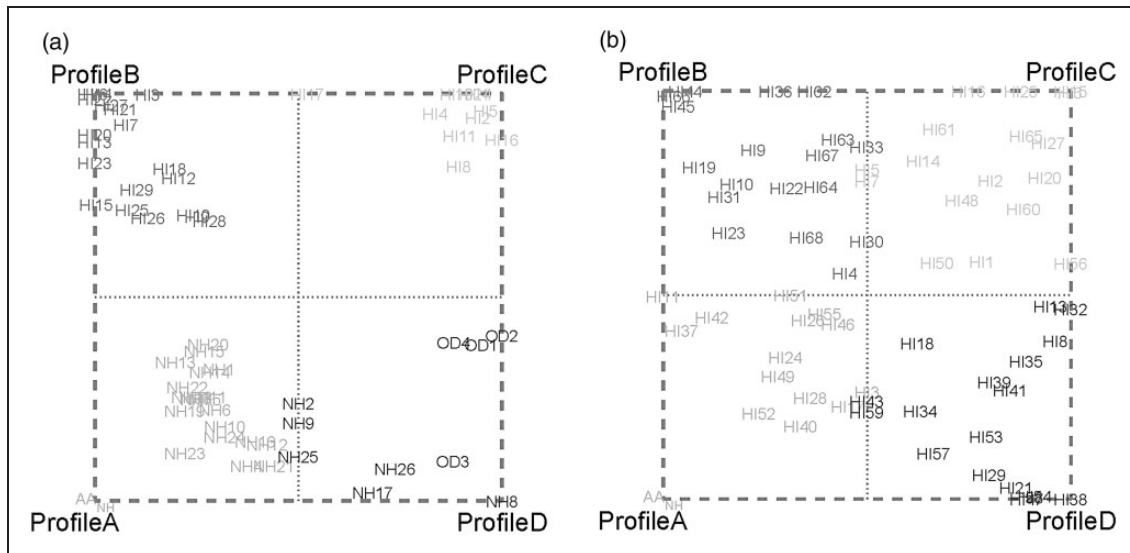


Figure 4. Simplex plots for (a) Study 1 and (b) Study 2. Representation of the listeners in a two-dimensional space. The four archetypes are located at the corners, and the remaining observations are placed in the simplex plot depending on their similarity with the archetypes. HI = hearing impaired; NH, normal hearing; OD = obscure dysfunction.

poor for B_1 and C_1 , in line with the hypothesis of this study. However, the performance in the tests corresponding to Distortion Type II was less consistent. Archetypes A_1 and B_1 , with an expected low degree of Distortion Type II, exhibited good performance in the binaural tests. Archetypes C_1 and D_1 showed poor performance in $B_{p_{dicho}}$, $B_{p_{tot}}$, and MCL_{LF} , but not in DS. This is due to the fact that DS was correlated to both principal components. As described in the “Method” section, the number of variables per dimension was set to four. Hence, DS should not be considered as a representative variable of Distortion Type II.

Figure 3(b) depicts the four archetypes from Study 2. The performance in the tests related to Distortion Type I was clearly good for Archetypes A_2 and D_2 and poor for B_2 and C_2 , in line with the hypothesis of this study. In addition, the performance in the tests related to Distortion Type II was clearly better for Archetypes A_2 and B_2 than for D_2 and C_2 , also in line with the hypothesis of the existence of four auditory profiles along two independent types of distortion.

Based on the archetypes presented in Figure 3, each listener was assigned to the auditory profile defined by the nearest archetype. Results from Study 1 are depicted in Figure 4(a)). The simplex representation shows how the listeners could be divided into clear clusters in the two-dimensional space. In the case of Study 2 (Figure 4(b)), the listeners were more spread out across the two-dimensional space and no clear groups could be identified. It should be noted that, in this case, Archetype A_2 , labeled as AA_{NH} , corresponded to the artificial observation with a good performance in all tests. It is located in the bottom-left corner in the simplex plot and

is far from the rest of observations because, in contrast to Study 1, the data set did not contain any data from NH listeners.

Figure 5 depicts the results of the supervised learning analysis. Decision trees were obtained using the raw data as an input and the identified auditory profiles as the output. In Study 1, the classification tree based on HL_{HF} and binaural pitch showed a very high sensitivity (95% true positives). In Study 2, the classification was based not only on the audibility loss at high and low frequencies but also on the estimate of OHC loss at low frequencies. The sensitivity of this classifier was slightly lower (91%). Although HL_{HF} is in the first node of both classifiers, the amount of hearing loss at high frequencies required to divide the listeners along the Distortion Type I dimension (i.e., Subgroups A–D and B–C) was lower for Study 1 than for Study 2. This is mainly due to the differences in the distribution of the hearing thresholds among the participants of each study. The differences observed between the two classification trees are further discussed in the following section.

Discussion

Two Types of Distortion to Characterize Individual Hearing Loss

This study proposed a new data-driven statistical analysis protocol, which was applied to two existing data sets. The goal was to determine the nature of the two main independent dimensions for individual hearing loss characterization. Based on existing literature findings, it was hypothesized that one dimension (Distortion Type I)

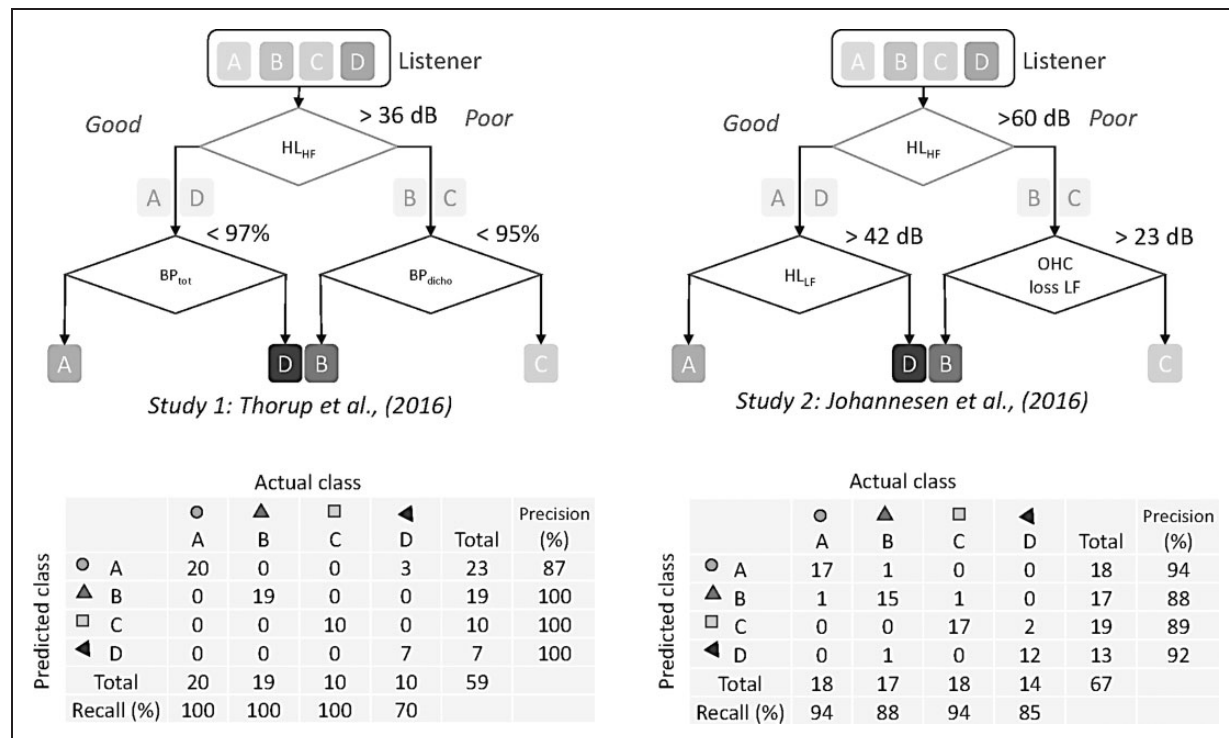


Figure 5. Decision trees and confusion matrices of the classifiers from the analysis of both data sets. For each study, the resulting classifier has three nodes. The right branch corresponds to a poor performance and the left branch to a good performance. The accuracy of the classifier is shown in the form of a confusion matrix where the correspondence of the actual classes and predicted classes are evaluated. HL_{HF} = hearing loss at high frequencies; Bp_{dicho} = BP dichotic condition; Bp_{tot} = BP diotic + dichotic; HL_{LF} = hearing loss at low frequencies; OHC loss LF = outer hair cell loss estimated at low frequencies.

would reflect audibility-related distortions, while the other dimension would reflect non-audibility-related distortions (Distortion Type II). The analysis performed on the two data sets provided different results, which need to be interpreted taking the differences between the two studies into account.

The analysis of the data set in Study 1, with a population of both near-NH and HI listeners, revealed that binaural processing tests were highly sensitive for the classification of the listeners and a main contributor to the Distortion Type II. In that study, the listeners presented a mild low-frequency hearing loss (24 dB $HL \pm 6$ dB) and a higher degree of high-frequency hearing loss (55 dB $HL \pm 6$ dB). As shown in Figure 5, the HI listeners were classified into Profile B or C according to their high-frequency hearing loss and were divided along the Distortion Type II dimension according to their binaural processing abilities. The analysis of the data set in Study 2, with only HI listeners, suggested that Distortion Type I was also related to high-frequency processing, while Distortion Type II was related to low-frequency processing. In Study 2, the listeners presented a higher degree of low-frequency hearing loss (37 dB $HL \pm 12$ dB) compared to Study 1 and a similar degree of high-frequency hearing loss (58 dB $HL \pm 12$ dB) but with a larger

variance. Although the listeners of Study 2 were distributed across the four profiles, they were not clearly divided into clusters as in Study 1 (Figure 4). This suggests that, although the hearing loss at low versus high frequencies may in this case be considered as a good indicator of Distortion Type I versus Type II, the corresponding auditory profiles were less clearly separated than in Study 1. This is probably due to the lack of NH or near-NH listeners in Study 2. Studies 1 and 2 also differed in terms of test batteries. In fact, Study 2 did not consider any test of binaural TFS processing, which may partly account for the difference in the reduced variance explained in the analysis of Study 2 compared to Study 1.

Although Studies 1 and 2 differed both in terms of listeners and test batteries, the analysis performed here revealed that in both cases Distortion Type I was dominated by high-frequency hearing loss. This was observed also in previous studies in which the sensitivity loss, particularly at high frequencies, was the main predictor of the differences among listeners (e.g., Vlaming et al., 2011). The loss of sensitivity at high frequencies can be ascribed to a loss of sensory cells, specifically OHC loss, which yields a loss of cochlear compression and a reduced frequency selectivity (Moore et al., 1999).

Other relevant dimensions suggested in previous studies were related to TFS processing (Rönnerberg et al., 2016) and low-frequency processing (Vlaming et al., 2011). In agreement with this, the analysis of Study 1 pointed toward measures of binaural TFS processing abilities (IPD detection frequency limit and binaural pitch test) for the second dimension, measures that may be correlated to FMDTs (Strelcyk & Dau, 2009), a measure assumed to involve monaural TFS processing abilities. In contrast, Study 2 contained tests that estimated OHC and IHC loss as well as BM compression. As shown in Table 1, HL_{HF} and $BM\ Comp_{HF}$ were strongly correlated to PC1, and HL_{LF} was strongly correlated to PC2 together with FMDTs. This suggests that, while HL_{HF} can be ascribed to a compression loss, HL_{LF} is most likely related to temporal coding deficits, as reflected by FMDTs. Despite the different outcome measures used in the two studies, the analysis of both studies is consistent with Distortion Type II being related to TFS processing. In summary, the outcomes of this study support the hypothesis that Distortion Type I may be more related to functional measures of spectral auditory processing deficits and Distortion Type II may be more related to functional measures of temporal auditory processing deficits.

Auditory Profiling and the Audibility-Distortion Model

In this study, it was assumed that there are two independent types of distortion that affect the overall listening experience and functional performance of the listener. Although it was hypothesized, based on earlier literature findings, that Distortion Type I involved deficits that covaried with the loss of sensitivity, audibility itself was not a priori considered as a fully separate dimension as in previous approaches. According to the proposal of this study, the two types of distortions are, ideally, fully independent. In Plomp's model, besides the attenuation component, a distortion component related to the suprathreshold deficits was proposed to account for the elevated SRTs in speech-in-noise intelligibility tests (Plomp, 1994). However, Humes and Lee (1994) argued that the distortion component can, in fact, appear as a consequence of a nonoptimal compensation of the spectral configuration of the audibility loss and not because of additional and independent suprathreshold deficits. They also stated that the effective compensation of the attenuation component should be performed prior to further investigation of the origin of the suprathreshold distortions. Humes (2007) reviewed previous studies of aided speech intelligibility and concluded that the main factors that explained the individual differences in speech intelligibility in older adults were audibility and cognitive factors. In the analysis presented here, both reanalyzed studies included

hearing threshold and speech intelligibility outcomes and Study 1 included a cognitive test of working memory. As audibility and cognitive factors are known to indirectly influence the performance in some of the other functional tests used in the analysis (e.g., Humes, 2007), it was decided to not treat them as independent dimensions, as this would have biased the analysis, and the aim was to take advantage of a data-driven statistical method to neutrally define the assumed two independent dimensions. Although audibility was clearly reflected as a contributor to the two distortion types in the present analysis, cognition did not emerge as a key variable. This is consistent with Lopez-Poveda et al. (2017), where it was found that working memory was only weakly related to outcome measures of hearing-aid benefit. However, as only one cognitive test was included in the present analysis, the findings do not allow for a clear conclusion about the role of cognition. Applying the present statistical method to test batteries that include more extensive cognitive measures might help clarify this aspect.

In contrast to this study, Kollmeier and Kiessling (2016) explained the factors contributing to hearing loss by three components: an attenuation component that produces a loss of sensitivity due to OHC and IHC loss, a distortion component associated with a reduced frequency selectivity, and a neural component related to degradations presented in the neural representation of the stimulus and associated to binaural processing deficits. The three components were not assumed to be independent such that the loss of sensitivity (the attenuation component) could covary with reduced frequency selectivity (the distortion component) and with IHC loss (the neural component). Despite the important difference of the assumption of an independent attenuation component between the two approaches, the present findings do reconcile rather well with Kollmeier and Kiessling's (2016) approach. Although Distortion Type I was found to be related to compression loss and elevated speech-in-noise recognition thresholds, Distortion Type II was associated to temporal and binaural processing deficits. Distortion Type I in this study can thus be compared to the distortion component from Kollmeier and Kiessling (2016) and Distortion Type II to their neural component. The two approaches thus share some similarities, except for the assumption of independence of the two distortion components in the current approach versus the assumption of an additional attenuation component in Kollmeier and Kiessling (2016).

Evaluation of the Data-Driven Method

The method used in this study was designed based on the hypothesis that the listeners could be divided into four auditory profiles according to the results from their perceptual outcome measures. First, two independent types

of distortions were assumed to characterize the individual hearing deficits of the listeners. Second, the extreme exemplars, that is, the archetypes, contained in the data were identified and the listeners were defined according to their similarity to the nearest exemplar. Third, the outcome measures that were most relevant for the classification of the listeners were identified. Other methods, such as linear regression, make use of the outcome measures to predict the performance in specific tests. This is typically done to explore the effects of different outcome measures on speech intelligibility. The novelty of this method lies in the fact that the characterization of the hearing deficits was carried out by analyzing the whole data set with the goal of achieving an individual hearing loss characterization. As this is a data-driven method, the results are highly influenced by the data included in the analysis. Therefore, one should be cautious when interpreting the results.

The method considered only two principal dimensions for explaining the data. The number of variables was reduced to have only four tests in each dimension. This decision makes the archetypes strongly connected to the hypothesis and keeps the number of variables per dimension balanced. However, if fewer than four variables are representative of one of the dimensions, the current algorithm also considers variables that can be correlated to both principal components. In this study, this was the case for only one variable of Study 1 (DS), which did not yield significant changes in the analysis. This limitation can be solved by imposing the assumption of orthogonality in the selection of the variables instead of using cross-validation. In this case, all the variables that are considered to belong to both dimensions are initially discarded. However, the explained variance might be lower and the number of the representative variables might change when using that method instead of an iterative cross-validation.

The archetypes, representing extreme exemplars, were used here as prototypes of the auditory profiles. Therefore, the rest of the listeners were assumed to belong to the same category as the nearest *archetype*. This has two main disadvantages. First, if outliers are present in the data, these will be most likely used as *archetypes*. Second, subgroups of listeners that are not well represented by any of the four auditory profiles are not considered here. In contrast, in Dubno et al. (2013), the identification of the exemplars corresponding to each audiometric phenotypes was done by an experienced researcher. That method is, however, not feasible for large data sets and may also be prone to judgement bias from the researchers. The use of unsupervised learning provides a solution to this potential problem. To better define the auditory profiles, alternative clustering as well as other advanced pattern recognition techniques may also be explored instead of an archetypal analysis

for profile identification and benchmarking (Ragozini et al., 2017).

The proposed method showed a potential for reanalyzing other existing data sets. The new exploratory approach can help test specific hypotheses by dividing the listeners into meaningful groups before analyzing the data. However, some requirements about the data are needed in order to reach consistent conclusions about a general characterization of hearing deficits. The data set should contain a representative sample of different degrees of hearing loss and a NH reference, as well as a substantial variability in performance in other tests, such as speech-in-noise intelligibility, which should be performed unaided. In this way, both audibility- and non-audibility-related factors would influence the performance of the listeners. As the method is sensitive to the input variables, a representative number of supra-threshold outcome measures should also be considered, including measures of loudness perception, binaural processing abilities, as well as outcome measures of spectral and temporal resolution, as it has been suggested in this and previous studies. In addition, cognition as well as physiological indicators of hearing loss, such as auditory brainstem responses or middle-ear response, may be included to further characterize a listener's auditory profile. If these requirements are not fulfilled, the method would still categorize listeners into four subgroups, but the results may be misleading and difficult to interpret.

Overall, the present method provided results in line with the initial hypotheses. The two types of distortions were found to be related to spectral and temporal auditory processing deficits, which supports the idea of considering two independent dimensions instead of previous models based on audibility and additional factors. The analysis of further and more extensive existing data sets with the data-driven method proposed here, provided that they contain a representative population of listeners and outcome measures, may help refine the definition of the two distortion types and improve future characterization of individual hearing loss. Such a characterization may be useful in future clinical practice toward better classification of patients in terms of hearing-aid rehabilitation.

Conclusion

The data-driven statistical analysis provided consistent evidence of the existence of two independent sources of distortion in hearing loss and, consequently, different "auditory profiles" in the data. While Distortion Type I was more related to audibility loss at high frequencies, the origin of Distortion Type II was connected to reduced binaural and temporal fine-structure processing abilities. The most informative predictors for profile identification beyond the audiogram were related to

temporal processing, binaural processing, compressive peripheral nonlinearity, and speech-in-noise perception. The current approach can be used to analyze other existing data sets and may help define an optimal test battery to achieve efficient auditory profiling toward more effective hearing-loss compensation.

Acknowledgments

The authors thank R. Ordoñez, T. Behrens, M. L. Jepsen, M. Gaihede, and the rest of BEAR partners for their input during the realization of this study. The authors also thank E. López-Poveda and P. Johannesen for proving access to their data (Study 2), as well as G. Lócssei, O. Strelcyk, A. Papakonstantinou, and T. Neher, who provided data for preliminary analyses not presented here. The authors would also like to thank the two anonymous reviewers who suggested some improvements in an earlier version of the manuscript.

Declaration of Conflicting Interests

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


Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by Innovation Fund Denmark Grand Solutions 5164-00011B (Better hEARing Rehabilitation project).

Notes

1. As the data from Study 1 were already in that form, the data from Study 2 were processed accordingly.
2. As the data from Study 2 were collected only in the better ear, the data from Study 1 were processed accordingly for a better comparison.

ORCID iD

Raul Sanchez Lopez  <http://orcid.org/0000-0002-5239-2339>

References

- Brand, T., & Hohmann, V. (2002). An adaptive procedure for categorical loudness scaling. *The Journal of the Acoustical Society of America*, *112*(4), 1597–1604. doi:10.1121/1.1502902
- Cutler, A., & Breiman, L. (1994). Archetypal analysis. *Technometrics*, *36*(4), 338–347. doi: 10.1080/00401706.1994.10485840.
- 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.
- Elberling, C., Ludvigsen, C., & Lyregaard, P. E. (1989). DANTALE: A new Danish speech material. *Scandinavian Audiology*, *18*(3), 169–175.
- Glasberg, B. R., & Moore, B. C. (1989). Psychoacoustic abilities of subjects with unilateral and bilateral cochlear hearing impairments and their relationship to the ability to understand speech. *Scandinavian Audiology*, *32*, 1–25.
- Houtgast, T., & Festen, J. M. (2008). On the auditory and cognitive functions that may explain an individual's elevation of the speech reception threshold in noise. *International Journal of Audiology*, *47*, 287–295.
- Humes, L. E. (2007). The contributions of audibility and cognitive factors to the benefit provided by amplified speech to older adults. *Journal of the American Academy of Audiology*, *18*(7), 590–603.
- Humes, L. E., & Lee, L. W. (1994). Response to “Comments on ‘Evaluating a speech-reception threshold model for hearing impaired listeners’” [J. Acoust. Soc. Am. 96, 586–587 (1994)]. *The Journal of the Acoustical Society of America*, *96*(1), 588–589. doi:10.1121/1.410446
- Jürgens, T., Kollmeier, B., Brand, T., & Ewert, S. D. (2011). Assessment of auditory nonlinearity for listeners with different hearing losses using temporal masking and categorical loudness scaling. *Hearing Research*, *280*(1–2), 177–191. doi:10.1016/j.heares.2011.05.016
- 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*, 1–14. Available at: <http://tia.sagepub.com/cgi/doi/10.1177/23312165166641055>.
- Killion, M. C., & Niquette, P. A. (2000). What can the pure-tone audiogram tell us about a patient's SNR loss? *The Hearing Journal*, *53*(3), 46–53.
- Kollmeier, B., & Kiessling, J. (2016). Functionality of hearing aids: State-of-the-art and future model-based solutions. *International Journal of Audiology*, *57*(sup3), S3–S28. doi:10.1080/14992027.2016.1256504
- 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. Retrieved from <http://journal.frontiersin.org/journal/10.3389/fnins.2014.00348/full>
- Lopez-Poveda, E. A., & Johannesen, P. T. (2012). Behavioral estimates of the contribution of inner and outer hair cell dysfunction to individualized audiometric loss. *Journal of the Association for Research in Otolaryngology*, *13*, 485–504.
- Lopez-Poveda, E. A., Johannesen, P. T., Pérez-González, P., Blanco, J. L., Kalluri, S., Edwards, B. (2017). Predictors of hearing-aid outcomes. *Trends in Hearing*, *21*, 1–28. doi:10.1177/2331216517730526
- Moore, B. C. J., Vickers, D. A., Plack, C. J., & Oxenham, A. J. (1999). Inter-relationship between different psychoacoustic measures assumed to be related to the cochlear active mechanism. *The Journal of the Acoustical Society of America*, *106*(5), 2761–2778.
- Mørup, M., & Hansen, L. K. (2012). Archetypal analysis for machine learning and data mining. *Neurocomputing*, *80*, 54–63. Retrieved from <http://linkinghub.elsevier.com/retrieve/pii/S0925231211006060>
- Nelson, D. A., Schroder, A. C., & Wojtczak, M. (2001). A new procedure for measuring peripheral compression in normal-

- hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 110(October), 2045–2064.
- 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. Retrieved from <http://link.aip.org/link/?JAS/63/533/1>
http://asadl.org/jasa/resource/1/jasman/v63/i2/p533_s1
- Plomp, R. (1994). Comments on “Evaluating a speech-reception threshold model for hearing-impaired listeners” [J. Acoust. Soc. Am. 93, 2879–2885 (1993)]. *The Journal of the Acoustical Society of America*, 96(1), 586–589. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8064024>
- 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. doi:10.1002/sam.11325
- Rönnerberg, J., Lunner, T., Ng, E. H., Lidestam, B., Zekveld, A. A., Sörqvist, P., . . . Stenfelt, S. (2016). Hearing impairment, cognition and speech understanding: exploratory factor analyses of a comprehensive test battery for a group of hearing aid users, the n200 study. *International Journal of Audiology*, 2027(September), 1–20. doi:10.1080/14992027.2016.1219775
- 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.
- Strelcyk, O., & Dau, T. (2009). Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing. *The Journal of the Acoustical Society of America*, 125, 3328–3345.
- Summers, V., Makashay, M. J., Theodoroff, S. M., & Leek, M. R. (2013). Suprathreshold auditory processing and speech perception in noise: Hearing-impaired and normal-hearing listeners. *Journal of the American Academy of Audiology*, 24(4), 274–292.
- Thorup, N., Santurette, S., Jørgensen, S., Kjærboel, E., Dau, T., & Friis, M. (2016). Auditory profiling and hearing-aid satisfaction in hearing-aid candidates. *Danish Medical Journal*, 63(10), 1–5. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/27697129>
- van Esch, T. E. M., & Dreschler, W. A. (2015). Relations between the intelligibility of speech in noise and psychophysical measures of hearing measured in four languages using the auditory profile test battery. *Trends in Hearing*, 19, 233121651561890doi:10.1177/2331216515618902
- van Esch, T. E. M., Kollmeier, B., Vormann, M., Lyzenga, J., Houtgast, T., Hällgren, M., . . . 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.
- Vlaming, M. S. M. G., Kollmeier, B., Dreschler, W. A., Martin, R., Wouters, J., Grover, B., . . . Houtgast, T. (2011). Hearcom: Hearing in the communication society. *Acta Acustica United With Acustica*, 97(2), 175–192.