Technical University of Denmark

DTU

Data-driven approach for auditory profiling

Sanchez Lopez, Raul; Bianchi, Federica; Fereczkowski, Michal; Santurette, Sébastien; Dau, Torsten

Link back to DTU Orbit

DTU Library
Technical Information Center of Denmark

Data-driven approach for auditory profiling

RAUL SANCHEZ[1,*], FEDERICA BIANCHI[1], MICHAL FERECZKOWSKI[1], SÉBASTIEN SANTURETTE[1,2], AND TORSTEN DAU[1]

[1] *Hearing Systems, Department of Electrical Engineering, Technical University of Denmark, Kgs. Lyngby, Denmark*

[2] *Department of Otorhinolaryngology, Head and Neck Surgery & Audiology, Rigshospitalet, Copenhagen, Denmark*

Nowadays, the pure-tone audiogram is the main tool used to characterize hearing loss and to fit hearing aids. However, the perceptual consequences of hearing loss are typically not only associated with a loss of sensitivity, but also with a clarity loss that is not captured by the audiogram. A detailed characterization of hearing loss has to be simplified to efficiently explore the specific compensation needs of the individual listener. We hypothesized that any listener's hearing 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 our hypothesis, the individual performance data from two previous studies were re-analyzed using an archetypal analysis. Unsupervised learning was used to identify extreme patterns in the data which form the basis for different auditory profiles. Next, a decision tree was determined to classify the listeners into one of the profiles. The new 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 non-linearity and speech-in-noise perception. The current approach is promising for analyzing other existing data sets in order to select the most relevant tests for auditory profiling.

INTRODUCTION

Supra-threshold distortions can be defined as the abnormal perception of stimuli when these are audible, i.e., above the audiometric threshold. While amplification can effectively compensate for the loss of sensitivity, supra-threshold distortions may require more advanced signal processing to improve hearing, particularly, when the hearing-aid user holds a conversation in noisy environments (Kollmeier and Kiessling, 2016; Plomp, 1978). Several studies have attempted to shed light on the underlying mechanisms responsible for these distortions by means of correlations of psychoacoustic tests with speech in noise performance (Glasberg and Moore, 1989; Strelcyk and Dau, 2009; Johannesen *et al.*, 2016). While these tests could provide a better characterization of the hearing deficits, they are infeasible in a clinical set up.

---

*Corresponding author: rsalo@elektro.dtu.dk

Raul Sanchez, Federica Bianchi, Michal Fereczkowski, Sébastien Santurette, and Torsten Dau

A recent study (Thorup *et al.*, 2016) proposed a test battery of new outcome measures used in a clinical set up for hearing aid candidates. Furthermore, Johannesen *et al.* (2016) investigated the influence of cochlear mechanical dysfunction, as well as the effects of temporal processing deficits and age on speech intelligibility in hearing-impaired (HI) listeners while wearing hearing aids (HA). The goal of the present study was to investigate how the characterization of hearing loss can be simplified by means of auditory profiling. For this purpose, a new analysis of these two datasets using archetypal analysis was performed. This analysis represents a useful tool for identifying patterns in the data and it has been proposed for prototyping and benchmarking (Ragozini *et al.*, 2017). The advantage of the proposed method is that the analysis involves the performance of the patient in different tests rather than correlations or regression analysis (Glasberg and Moore, 1989; Johannesen *et al.*, 2016).

Hearing deficits and auditory profiles

The characterization of hearing deficits can be complex and it needs to be simplified to efficiently explore the specific compensation strategies for the individual. Several authors have suggested classifications of the listeners to reduce this complexity. Three of such approaches served as inspiration for the hypothesis of the present study:

1. I: Plomp (1978) suggested that the hearing deficits in relation to speech intelligibility can be divided into two components: an attenuation (A) and distortion (D) component.

2. II: Lopez-Poveda (2014) reviewed the mechanisms that produce hearing loss and their perceptual consequences for speech. In a bi-dimensional space, the hearing loss can be understood as the sum of an outer-hair-cell (OHC) loss and an inner-hair-cell (IHC) loss.

3. III: Dubno *et al.* (2013) suggested four auditory phenotypes for explaining age-related hearing loss based on animal studies. Although this can be informative revealing the origin of the hearing loss, pure-tone audiometry has remained the main contributor in this classification and no information about speech or other tasks were considered.

Here, we hypothesize that any listener's hearing can be characterized along two dimensions: distortion type I and distortion type II (Fig. 1). While distortion type I can cause a loss of audibility, distortion type II is considered to reflect a non-audibility-related distortion, referred to as a *clarity loss*. In this space, four profiles may be identified: a sensitivity loss (Profile A), a sensitivity loss with associated distortions (Profile B), a sensitivity loss with a severe clarity loss (Profile C) and a mild-moderate clarity loss (Profile D). Figure 1 shows the four profiles in the two-dimensional space, where normal-hearing (NH) listeners are placed at the bottom-left corner since they would not exhibit any type of distortion. It is hypothesized that while the distortions of type I (in the vertical dimension) are due to a loss of frequency selectivity and a
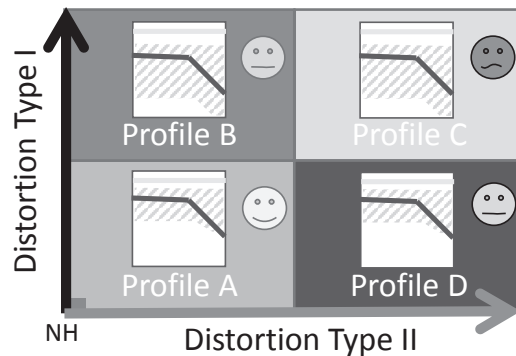
Fig. 1: Sketch of the hypothesis. Hearing deficits can happen due to two types of distortions that are independent. Distortion type I: distortions with consequent loss of sensitivity. Distortion type II: non-audibility related distortions. Profile A: sensitivity loss. Profile B: sensitivity loss and distortion. Profile D: moderate clarity loss and Profile C: severe clarity loss.

loss of cochlear compression, the distortions of type II (in the horizontal dimension) are related to inaccuracies in terms of temporal coding, in line with the conclusions from other studies (Glasberg and Moore, 1989; Strelcyk and Dau, 2009).

The aim of the present study was to investigate whether listeners can be grouped in the four different profiles by identifying trends in the results from the behavioral tests using a data-driven approach. The analysis is expected to help identify the underlying mechanisms and perceptual consequences of the hearing deficits that characterize an individual auditory profile. Moreover, it is of interest to reduce the test battery using only the most relevant tests for classifying the subjects in the proposed auditory profiles.

METHOD

The method used in the analysis is depicted in Fig. 2. The data-driven approach is based on two stages. First, *unsupervised learning* was used to identify the trends in the data that can be used to categorize the subjects in different profiles. The second stage consisted of *supervised learning*. Once the subjects were segmented by profiles, the data were analyzed again to find the best classification structure that can 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 dataset are explored. In contrast to linear regression, unsupervised learning does not aim to predict a specific output, for example, speech intelligibility. In the present approach, identified auditory profiles were eventually inferred using different unsupervised learning techniques.
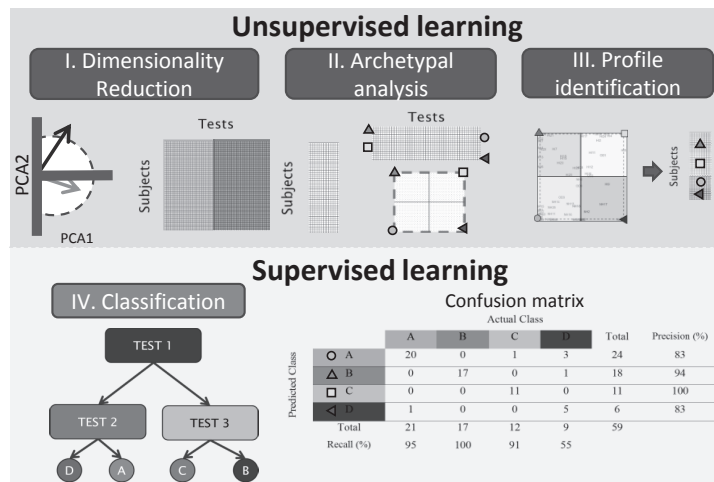
Raul Sanchez, Federica Bianchi, Michal Fereczkowski, Sébastien Santurette, and Torsten Dau

Fig. 2: Sketch of the method. Upper panel shows the supervised learning techniques applied to the whole dataset.The bottom panel shows the supervised learning, which uses the original data as the input and the identified profiles from the archetypal analysis as the output.

I. *Dimensionality reduction:* According to our hypothesis, two dimensions, corresponding to the two types of distortions, should be sufficient for auditory profiling. Therefore, the subset of variables that were strongly correlated to the first principal component (PCA1) and the second principal component (PCA2) and can explain most of the variance were chosen for the next step. The optimal number of variables in each of the two principal components was chosen using a leave-one-out cross-validation in an iterative principal component analysis (PCA).

II. *Archetypal analysis:* This technique combines characteristics of matrix factorization and cluster analysis. The aim of the analysis was to identify extreme patterns in the data (archetypes). This has the advantage that the subjects are no longer defined by the quantified performance in each of the tests, but by their similarity to the extreme exemplars contained in the data, i.e., the *archetypes*.

III. *Profile identification:* Based on the archetypal analysis, the subjects were placed in a simplex plot (square visualization). Here, the archetypes are located at each corner and the subjects 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 that cluster. Consequently, each subject was labelled with the letter of an auditory profile according to the nearest archetype.

Supervised learning

Once the profiles have been identified, supervised learning can be performed. Now, the joint probability density of the dataset and the output (the identified profiles) can be used to select the tests that are most relevant for the classification of the subjects in the auditory profiles.

IV. *Classification:* Decision trees are able to predict the class that corresponds to a given observation. Here, each relevant test is used in the nodes forming a logical expression and dividing the observations accordingly (Fig. 2 IV). A classification tree needs to be trained with a subset of the data and a known output. In the present study, the identified auditory profiles (III) were used as the response variable and a 5-fold cross-validation was used for training the classifier.

RESULTS

Dimension	Thorup (2016)		Johannesen (2016)	
	Variable	Test	Variable	Test
Distortion I	HL_{LF}	Hearing loss at low frequencies	HL_{HF}	Hearing loss at high frequencies
	HL_{HF}	Hearing loss at high frequencies	$BMComp_{HF}$	Basilar membrane compession at high frequencies
	SRT_Q	Speech reception threshold (SRT) in quiet	$OHCloss_{HF}$	Outer hair cell loss estimated at high frequencies
	SRT_N	SRT in noise using hearing in noise test (HINT)	$IHCloss_{HF}$	Inner hair cell loss estimated at high frequencies
	SRT_{ISTS}	SRT in noise using international speech test signal		
Distortion II	Bpdio	Binaural Pitch (BP) diotic control condition	HL_{LF}	Hearing loss at low frequencies
	Bpdicho	BP dichotic condition	FMDT	Frequency modulation discrimination threshold
	Bptot	BP diotic + dichotic	$OHCloss_{LF}$	Outer hair cell loss estimated at low frequencies
	MR	Masking release (SRTN-SRTISTS)	$IHCloss_{LF}$	Inner hair cell loss estimated at low frequencies
	$ACALOS_{Slope3k}$	Slope of growth of loudness at 3 kHz		

Table 1: Result from the dimensionality reduction of the two datasets. Variables strongly correlated to PCA1 (distortion type I) and PCA2 (distortion type II).

The whole dataset was reduced to the variables that were strongly correlated to Dimension I (PCA1) or Dimension II (PCA2), as summarized in Table 1. In Thorup *et al.*'s study, the dimensionality reduction revealed that the binaural tests were orthogonal to most of the tests related to audibility. Principal component analysis could explain 83.82% of the variance with two components. In Johannesen *et al.*'s study, Dimension II seemed to be dominated by low-frequency processing and Dimension I by high-frequency processing. PCA could explain 67% of the variance with the chosen variables.

The archetypal analysis was used to identify four archetypes using the variables from Table 1. As shown in Fig. 3, in both studies, Profile A (archetype A) exhibits the best performance in both dimensions and Profile C the worst. Profile B shows poor performance only in Dimension I while Profile D shows poor performance only in Dimension II. Based on these archetypes, each listener was considered to belong to the auditory profile of the closest archetype. Figure 3 B1 (Thorup *et al.*'s dataset) illustrates how the listeners are divided in clusters in the two-dimensional study. However, Figure 3 B2 (Johannesen *et al.*'s dataset) shows how the listeners are spread out among the profiles.
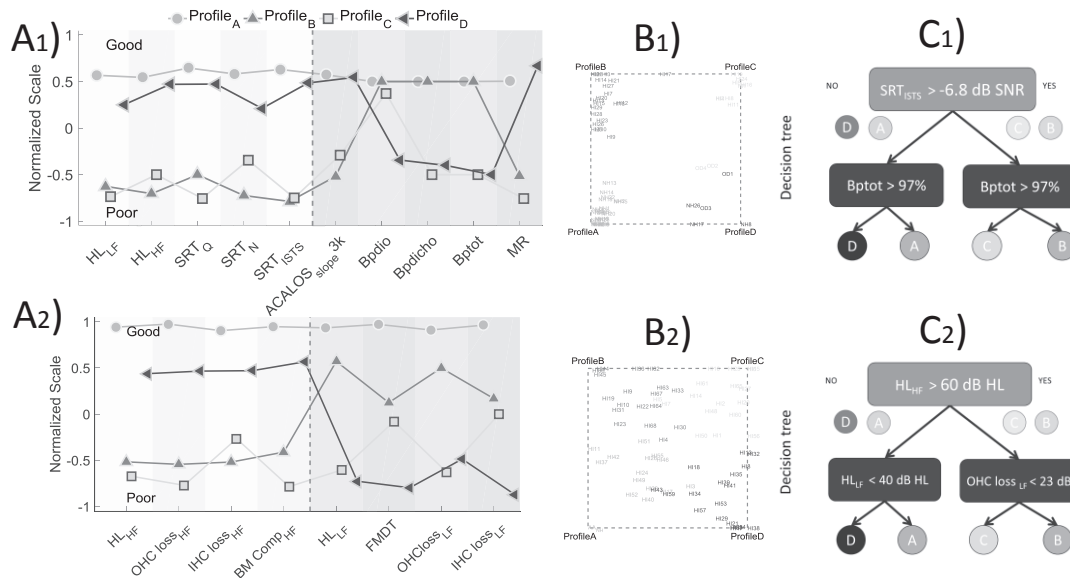

Fig. 3: Archetypes: Extreme exemplars of the different patterns found in the data. A1) Normalized performance of each of the 4 archetypes from Thorup *et al.*'s study. B1) Simplex representation of the listeners of Thorup *et al.*'s study. C1) Decision tree result of the supervised learning. A2, B2, and C2) are the same as A1, B1, and C1) but for Johannesen *et al.*'s study.

Decision trees were obtained by using the raw data as an input and the auditory profiles as the output. In Thorup *et al.*'s study, the classification tree based on SRT_{ISTS} and binaural pitch showed a very high precision (58 out of 59 true positives). In Johannesen *et al.*'s study, the classification was based not only on the audibility loss at high and low frequencies but also the estimate of OHC loss at low frequencies. The precision of this classifier was lower (66%).

DISCUSSION

The hypothesis in terms of the proposed four auditory profiles was evaluated in a data-driven approach. It was assumed that one dimension (distortion type I) may be related to a reduced frequency selectivity, while the second dimension would be related to

temporal processing reflecting to a non-audibility related distortions (distortion type II). First, the analysis of a dataset with a population of near-normal-hearing and hearing-impaired listeners revealed that binaural processing tests were highly sensitive for the classification of the listeners and a main contributor in the distortion type II dimension. Second, the analysis of a dataset with only hearing-impaired listeners showed that the distortion type I was related to high-frequency processing and the distortion type II was related to low-frequency processing. The two analyses cannot directly be compared because the tests and the listeners differed across studies. It should be noted that Johannesen *et al.*'s study did not consider any test concerning binaural processing. Therefore, the difference in the explained variance across the two studies is mostly caused by the lack of such information in their study.

Pure-tone audiometric thresholds are used to quantify the hearing loss but they can, in fact, be the consequence of different factors. As shown in Fig. 3 A2, dimension I does not only contain the high-frequency hearing loss but also estimated cochlear compression. Dimension II contains the low-frequency hearing loss and the outcome of the frequency modulation detection task which has been suggested to reflect temporal processing abilities. Therefore, it is important to bear in mind that there are interactions between the audibility and the two types of distortions proposed here. One approach to disentangle this interaction may be made based on the effects related to the OHC vs IHC processing.

If a substantial population of IHC or neural fibers is affected, the thresholds can be elevated (Lobarinas *et al.*, 2013), leading to temporal distortions as well as degraded binaural processing (Profiles D and C). However, the temporal acuity can also be compromised while audiometric thresholds are normal or close-to-normal (Zeng *et al.*, 1999) (Profile D). OHC loss is typically associated with basilar membrane (BM) compression loss (reduced frequency selectivity) as well as elevated audiometric thresholds (Ahroon *et al.*, 1993). Although reduced compression leads to a threshold elevation (Profile B), listeners with elevated thresholds can still have a nearly-normal BM compression (Profile A).

It is likely that the two types of deficits (degraded frequency selectivity vs degraded temporal processing) affect speech perception in different ways. The signal processing strategies that can be applied to compensate for each type of impairment can be assumed to be different. Therefore, both loss of audibility and outcome measures reflecting spectral and temporal distortions should be part of a clinical test battery for characterizing hearing deficits.

Conclusion

The new analysis provides consistent evidence of the existence of two sources of distortion and 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 processing abilities and low-frequency hearing loss. The most informative predictors for the 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 and might help define a test battery to achieve an efficient auditory profiling.

REFERENCES

Ahroon, W.A., Davis, R.I., and Hamemik, R.P. (**1993**). "The role of tuning curve variables and threshold measures in the estimation of sensory cell loss," Audiology, **32**, 244-259. doi: 10.3109/00206099309072940

Dubno, J.R., Eckert, M.A., Lee, F.S., Matthews, L.J., and Schmiedt, R.A. (**2013**). "Classifying human audiometric phenotypes of age-related hearing loss from animal models," J. Assoc. Res. Otolaryngol., **14**, 687-701. doi: 10.1007/s10162-013-0396-x

Glasberg, B.R., and Moore, B.C. (**1989**). "Psychoacoustic abilities of subjects with unilateral and bilateral cochlear hearing impairments and their relationship to the ability to understand speech," Scand. Audiol. Suppl., **32**, 1-25.

Johannesen, P.T., Pérez-González, P., Kalluri, S., Blanco, J.L., and 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 Hear., **20**. doi: 10.1177/2331216516641055

Kollmeier, B., and Kiessling, J. (**2016**). "Functionality of hearing aids: State-of-the-art and future model-based solutions," Int. J. Audiol., 1-26. doi: 10.1080/14992027.2016.1256504

Lobarinas, E., Salvi, R., and Ding, D. (**2013**). "Insensitivity of the audiogram to carboplatin induced inner hair cell loss in chinchillas," Hear. Res., **302**, 113-120. doi: 10.1016/j.heares.2013.03.012

Lopez-Poveda, E.A. (**2014**). "Why do I hear but not understand? Stochastic undersampling as a model of degraded neural encoding of speech," Front. Neurosci., **8**. doi: 10.3389/fnins.2014.00348

Plomp, R. (**1978**). "Auditory handicap of hearing impairment and the limited benefit of hearing aids," J. Acoust. Soc. Am., **63**, 533-549. doi: 10.1121/1.381753

Ragozini, G., Palumbo, F., and D'Esposito, M.R. (**2017**). "Archetypal analysis for data-driven prototype identification. Stat. Anal. Data Min., **10**, 6-20. doi: 10.1002/sam.11325

Strelcyk, O., and Dau, T. (**2009**). "Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing," J. Acoust. Soc. Am., **125**, 3328-3345. doi: 10.1121/1.3097469

Thorup, N., Santurette, S., Jørgensen, S., Kjærbøl, E., Dau, T., and Friis, M. (**2016**). "Auditory profiling and hearing-aid satisfaction in hearing-aid candidates," Dan. Med. J., **63**, A5275.

Zeng, F.G., Oba, S., Garde, S., Sininger, Y., and Starr, A. (**1999**). "Temporal and speech processing deficits in auditory neuropathy," Neuroreport, **10**, 3429-3435. doi: 10.1097/00001756-199911080-00031