Authors' accepted manuscript
Article published in *The Journal of the Acoustical Society of America* 152, 3155 (2022)
The final published article is available at https://doi.org/10.1121/10.0015237

Title: Morphologic clustering of earcanals using deep learning algorithm to design artificial ears dedicated to earplugs attenuation measurement

Running title: Deep learning-based earcanals morphologic clustering

Bastien Poissenot-Arrigoni,[1] Chun Hong Law,[2] Djamal Berbiche,[3] Franck Sgard,[2] and Olivier Doutres[1]

[1] *Department of Mechanical Engineering, École de Technologie Supérieure (ÉTS), 1100 Rue Notre-Dame O, Montréal, Québec, H3C 1K3, Canada*

[2] *Institut de recherche Robert-Sauvé en santé et sécurité du travail, 505 Boulevard de Maisonneuve O, Montréal, Québec, H3A 3C2, Canada*

[3] *Département des Sciences de la Santé Communautaire. Faculté de Médecine et des Sciences de la Santé. Université de Sherbrooke. Centre intégré de santé et de services sociaux de la Montérégie-Centre. Centre de recherche Charles-Le Moyne (CRCLM). Campus de Longueuil, 150 Place Charles-Le Moyne. bureau 200, C.P. 11, Longueuil, Canada J4K 0A8*

ABSTRACT

Designing earplugs adapted for the widest number of earcanals requires acoustical test fixtures (ATFs) geometrically representative of the population. Most of existing ATFs are equipped with unique sized straight cylindrical earcanals, considered representative of average human morphology, and are therefore unable to assess how earplugs can fit different earcanal morphologies. In this study, a methodology to cluster earcanals as a function of their morphologies with the objective of designing artificial ears dedicated to sound attenuation measurement is developed and applied to a sample of a Canadian workers' earcanals. The earcanals morphologic indicators that correlate with the attenuations of 6 models of commercial earplugs are first identified. Three clusters of earcanals are then produced using statistical analysis and artificial intelligence-based algorithm. On the sample of earcanals considered in this study, the identified clusters differ by the earcanals length, and surface and ovality of the first bend cross-section. The cluster that comprises earcanals with small girth and round first bend cross-section shows earplugs induced attenuation significantly higher than the cluster that includes earcanals with bigger and more oval first bend cross-section.

I. INTRODUCTION

Commercial disposable and reusable earplugs are widely used to prevent noise-induced hearing loss by attenuating the surrounding noise. To efficiently attenuate noise, the shape and material of the earplugs must match the earcanal morphology and provide a tight seal. However, due to the wide variability in humans' morphology it is difficult for designers to achieve a universally acceptable product (Ferguson et al., 2015). Thus, designing efficient and adapted earplugs that fit the widest range of earcanals morphologies remains extremely challenging. The-one-size-fit-the-most approach has been used by manufacturer for many years to design earplugs. However, earplugs available in one size may provide either physical discomforts to extra-small earcanals due to a too-tight fit (e.g., pain inside the earcanal) or even functional discomfort (e.g., earplug falling out) and low attenuation to extra-large earcanals (Berger and Voix, 2022; Doutres et al., 2020). Today, more inclusive design approaches tend to be favored to ensure safety and comfort for all (not only in the hearing protection field, but also in clothing and architecture for example). To ensure the best fit for the widest variety of users, a common solution consists in providing some earplugs models in two or more sizes. For example, some models of foam earplugs are available in regular and small size. These sizes correspond to different earplug diameters, but targeted user groups of each size are not clearly identified on the packaging, making the selection and use of these earplugs much less convenient. As for premolded earplugs (usually made of flanges affixed to a stem), that may also be available in a range of sizes, it has been shown that the greater the number of flanges, the fewer the sizes required to fit the population (Berger and Voix, 2022). However, this is more a general trend than a practical designing rule. Designing for the outliers and introducing diversity into the design process requires inclusive methods and tools.

Acoustical test fixture (ATFs) (that comply with the ANSI S12.42 standard) are good candidates for earplugs design tools because they allow for rapid and repeatable attenuation measurements. However, existing ATFs are equipped with unique sized straight cylindrical earcanals in which some earplugs (for example flangeless bullet-shaped earplug of small diameter) cannot be properly fitted (Smith et al., 1980; Berger et al., 1986). Furthermore, for a given earplug model to be tested, artificial straight cylindrical earcanals poorly capture the intra-individual variability in sound attenuation due to earplug fit (Benacchio, 2019) and cannot capture the inter-individual variability caused by large differences between human earcanal morphology (e.g. extra-small, regular and extra-large earcanals). An ATF intended to test how earplugs can fit different users should therefore allow for a variety of shapes of earcanals (Berger, 2005). There is thus a need for more realistic artificial ears available in a variety of sizes and shapes, characteristic of targeted populations and instrumented to measure sound attenuation. It would allow for the design of earplugs that are better suited to a wide range of earcanal sizes and shapes or better identify the population for which the earplug is best suited.

The objective of this study is to develop a methodology to cluster earcanals as a function of their morphologies with the objective of designing artificial ears dedicated to sound attenuation measurement and apply it to a sample of a Canadian workers' earcanals. In this context, having a comprehensive view of the earcanals morphology and its relation to earplugs sound attenuation is crucial, but the number of studies on this subject are limited. In 1988, Abel et al. found significant differences between women and men in attenuation of four commercially available earplugs: two foam earplugs and two premolded earplugs. The attenuations of earplugs available in a single size were lower when measured on women, whereas no gender effect was observed for earplugs available in a range of sizes. Gender differences in attenuation were therefore partly attributed to earcanal morphology differences between men and women. Abel et al. (1990) examined the correlation between the real attenuation at threshold (REAT) of three earplugs measured on 93 subjects and three morphologic parameters of earcanals estimated from the earmolds of these subjects. These parameters were: (i) the areas of two cross-sections of the earcanal estimated at the conchomeatal angle (first bend region) and at the cartilaginous-bony junction (second bend region), (ii) the conicity (called degree of funneling in the Abel study) calculated as the ratio between these two section areas and finally (iii) the tortuosity (which quantifies if the earcanal is more tortuous or straight), estimated visually. Results showed that a mismatch between the earcanal and the protector shapes could affect the attenuation. These earplug/earcanal mismatches were mainly attributed to the tortuosity and the conicity. Moreover, Abel et al. found that attenuation is linearly related to the cross-sectional area of the earcanal at the cartilaginous-bony junction. A gender effect was observed since the correlation between the cross-sectional area of the earcanal at the cartilaginous-bony junction and the attenuation was found positive for women and negative for men. The effects of the morphology on sound attenuation were found higher at medium frequencies (3150 Hz) than at low frequencies (500 Hz). Viallet et al. (2015) found similar tendencies on the effects of morphology on sound attenuation. Using a numerical approach, Viallet et al. were able to investigate the effects of earcanal morphology and acoustic leakage between the earcanal and earplugs. They showed that the important variability in the simulated sound attenuation of a foam and silicone earplugs was mainly due to acoustic leakage for frequencies below 1 kHz and by the inter-individual variability of the earcanal morphology between 1 and 5 kHz. More recently, Mououdi et al. (2018) measured 918 external ears dimensions of 153 operational workers and found that the design of molded type earplugs should be improved to better match earcanal entrance shape and diameter to avoid inducing acoustic leaks. The literature thus suggests that the inter-individual variability in earcanal morphology contributes significantly to the inter-individual variability in sound attenuation. However, none of these studies provides a comprehensive description of earcanals through morphologic indicators quantified objectively together with their relations with attenuations of earplugs from the three earplugs family: roll-down-foam, premolded and push-to-fit. Thus, there is a lack of data and methods to design artificial ears representative of the wide variability in earcanals morphologies of a given population and able to mimic the sound attenuation measured on these earcanals.

In this work, a methodology to cluster earcanals as a function of their morphologies with the objective of designing artificial ears dedicated to sound attenuation measurement is developed and applied to a sample of Canadian workers' earcanals. The paper is organized as follows. Section II presents the morphologic and attenuation data acquisition and details the proposed methodology. Section III discusses the results and presents the limitations of this study. Finally, some concluding remarks are given in Section 4.

II. METHODOLOGY

The general description of the methodology used to cluster earcanals is shown in FIG. 2. In short, it starts with a verification of the main hypothesis of this work (step 0), followed by the clustering process (steps 1 and 2) and ends by the evaluation of the proposed clusters (steps 3 and 4).

Sections II.A. to II.C describe the sample of participants and the acquisition of morphologic and attenuation data on which the clustering process is applied. Based on the literature, morphologic indicators supposedly correlated to attenuation are proposed and extracted from the sample of 242 earcanals. Attenuations of six different earplugs are objectively measured on these same earcanals. The clustering process is described in section II.D. In step 0 (section II.D.1), correlations between morphologic indicators and attenuation are evaluated to check that earcanals morphology is effectively related to inter-individual variability in sound attenuation. In step 1, a pre-processing of the morphologic dataset is performed: n combinations of morphologic indicators relevant for the clustering are selected following the rules detailed in section II.D.2. These combinations are then set as input to the clustering algorithm (see section II.D.3 about the k-means clustering algorithm) which is executed in step 2 to obtain $2.n$ clustering proposals based on earcanal morphologies: n proposals of k=2 clusters and n proposals of k=3 clusters. The next two steps, aim at choosing the clustering proposal which is the most relevant to be used as a basis to the design of realistic artificial ears representative of a sample of earcanals and dedicated to sound attenuation measurement. To do so, statistical analyses are performed to check that morphologic indicators are significantly different from one cluster to another (step 3, referred to as internal validation) and that PAR data are significantly different from one cluster to another (step 4, referred to as external validation).

A. Participants

A total of 121 persons (18 females, 103 males) working in three different Canadian companies participated in this study. Participants are aged between 21 and 64 years old (mean 46, standard deviation 10 years). They are exposed to noise at work and used to wear earplugs before being involved in the study. They did not have antecedents of ear or neurological pathologies and did not have an important amount of earwax in their earcanals. This study usesthe secondary data of morphologic and attenuation data collected during a field survey on earplugs comfort (Doutres et al., 2018) [Grant IRSST #2015-0014, Principal Investigators: Doutres and Sgard] approved by the ethical committee of the ÉTS[1] (ethic certificate H20171101).

B. Morphologic data acquisition

1. Earcanals morphology sampling and scanning

The left and right earcanal morphology of each participant was obtained by scanning earmolds of earcanals. Earmolds were casted by two different custom earplugs manufacturer: Laviolette auditory laboratory, QC, Canada (manufacturer #1) and Custom protect ear Inc, BC, Canada (manufacturer #2). The manufacturing process of custom earplugs involved remake of earmolds prior to the fabrication of custom earplugs. Among the 242 earmolds of this study (2 times 121 participants), 64 were cast and scanned by manufacturer #2 before being reworked. Manufacturer #2 casted and scanned 52 others after earmolds being remade. Remaking operations performed on these earmolds

[1] École de technologie supérieure

included cutting the lateral part of earmold to keep only the earcanal plus the concha and a portion of helix, chamfering the medial part of the mold and carrying out a hole to introduce acoustic filters. The remaining 126 earmolds casted by manufacturer #1 were slightly modified before being scanned in our laboratory using a 3D Scanner Einscan-SP (Hangzhou Shining 3D Tech Co., China). Scans were hole-filled and smoothed using the EinScan-S Series v2.6.0.8 software. Operations performed on these earmolds included cutting the lateral part of earmold to keep only the earcanal plus the concha and a portion of helix. These simple operations did not modify the shape of the earcanal part of the mold.

The assumption is made that obtained earcanal scans accurately describes the participants' earcanals morphology: the modifications of the real earcanal morphology due to the acquisition process (i.e., the earmold casting process, the 3D scanner model, and the earmold reworking process) are considered negligible and the difference between scans is only attributed to the difference between participants' earcanal morphology.

2. Extraction of morphologic indicators of shape and size of earcanals from scans

The earcanal is an "S-shaped" duct that extends between the concha on its lateral side and the tympanic membrane on its medial side. The cross-section shape and size vary along the duct curvilinear axis (axis that passes through the centroid earcanal cross-sections, as seen in FIG. 1). As an overall trend, cross-sections become smaller and more circular in the medial direction. Different characteristic sections are usually used to describe earcanal morphology (Lee et al., 2018; Fan et al., 2021; Abel et al., 1990). In this study, three characteristic cross-sections that cover all the earcanal portion accessible through the casting process are used: the entrance (E), the first bend (FB) and the second bend (SB). The entrance is usually defined at the base of the concha. The first bend is located a few millimeters after the entrance in the cartilaginous part of the earcanal. The second bend is positioned deeper in the earcanal and close to the cartilaginous-bony junction.

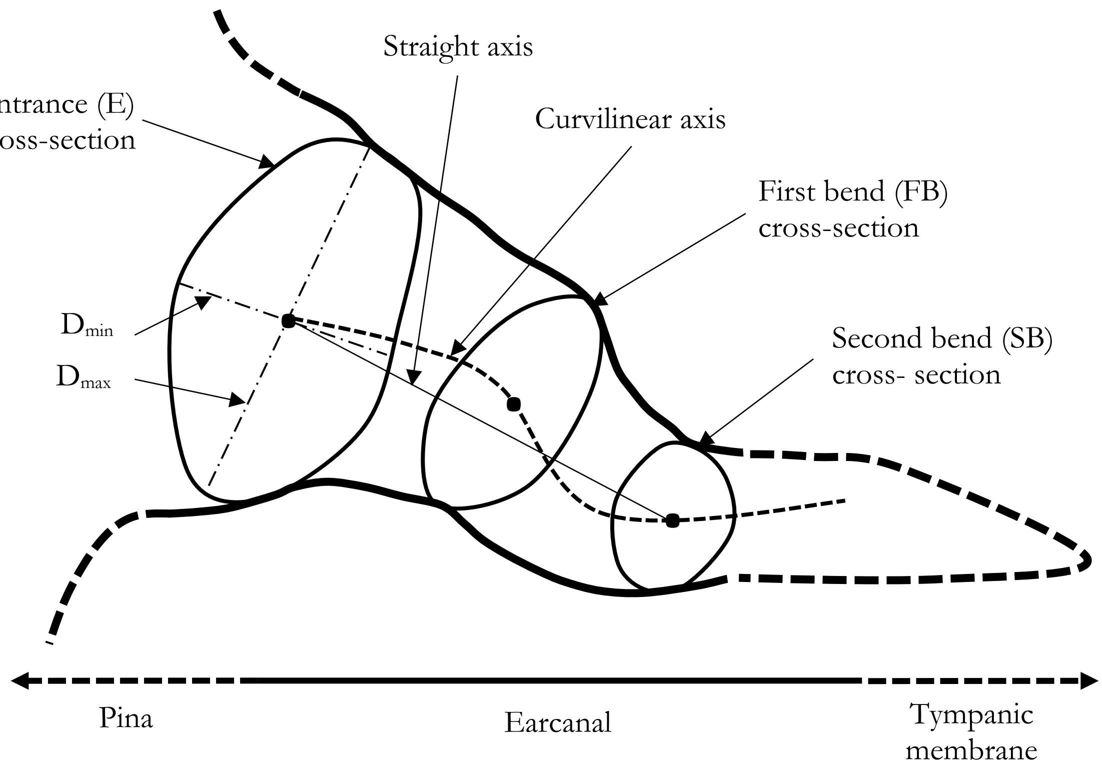

FIG. 1. Earcanal description. Dark thick solid lines represent earcanal walls in the region of interest for this study. Dark thick dotted lines represent earcanal regions that are ignored. Dark thin solid lines represent reference cross-sections of earcanal. Dark thin dotted line represents the curvilinear axis of the earcanal. Thin mixed lines represent the longest and shortest diameters of entrance cross-section (used to calculate shape indicators as described below in this section)

Two dimensions can be used to describe the morphology of the earcanal: size and shape. In this work, five features are chosen to characterize these two dimensions either because they have been shown to be relevant to the ergonomic design of an ear product (Lee et al., 2018, Fan et al., 2020) or correlated with earplugs attenuation (Abel et al., 1990). Each feature is quantified with one or several indicator(s). The calculation of all indicators belonging to the two aforementioned dimensions is based on the determination of the three cross-sections E, FB and SB. It is worth noting that these 3 characteristic cross-sections may or may not be involved in the fit of the earplugs (since the earplug fit associated to the measured PAR is unknown). For example, cross-section SB may not be involved in the fit of roll-down-foam earplugs for long earcanals, or if the earplug is not fitted deeply inside the earcanal. Similarly, cross-section E may not be involved in the fit of some push-to-fit-foam earplugs fitted deeply inside extra-large earcanals. The goal here is to describe the earcanal with morphologic indicators potentially related to earplugs attenuation (based on the limited literature on the subject). The relevance of these indicators will be discussed in section III.B.

The position of each cross-section (E, FB and SB) in the earcanal is located using an objective methodology to avoid inducing any experimenter's bias. This objective methodology is based on both the landmarks method and an objective method described below based on the positioning of cross-sections perpendicular to the curvilinear axis of the earcanal. First, the curvilinear axis is extracted using the Stinson and Lawton's (1989) method. For each earcanal, the curvilinear axis has two local maxima of curvature. The first local maxima of curvature (the closest to cross-section E) and the second (the closest to the tympanic membrane) correspond to the position on the curvilinear axis of the FB and SB respectively. Cross-sections FB and SB are identified as the intersection between the earcanal walls and the planes perpendicular to the curvilinear axis at these two positions. Some earmolds are not casted deep enough in the earcanal to reach the SB. For these earmolds, the most medial section of the earmold is chosen as the section of the SB. To identify cross-section E with a good repeatability, Lee et al. (2018) methodology (also used by Fan et al. (2021)) is adapted. This method is based on 4 different points (landmarks) to define the earcanal cross-section E. In the work presented here, cross-section E is defined as the intersection between the earcanal walls and a plane perpendicular to the curvilinear axis that passes through the most posterior point at cross-section E.

This specific point defined in Lee et al. (2018) is chosen because it is the most easily identifiable one. Indeed, this point is located right at the junction between the concha and the earcanal so in this zone, the earcanal surface has a high curvature. Curvy areas such as bumps and valleys can easily be located on a surface with a good repeatability.

The features used to describe the earcanal size are the length and girth. The earcanal length is characterized by the length of its curvilinear axis (in mm) between cross-sections E and SB (because the bony portion of the earcanal was not accessible through the molding process). The girth of the three earcanal cross-sections (i.e. E, FB and SB) are described by two indicators that are either their area (in mm^2) or circumference (in mm).

The features used to describe the earcanal shape are the tortuosity, the conicity and the shape of cross-sections. The tortuosity measures if the earcanal is straight or crooked (i.e., being more "S-shaped"). It is computed as the ratio between the curvilinear and the Euclidean length of the earcanal between the E and SB cross-section centroids (see FIG. 1). A tortuosity equal to 1 indicates that the duct is perfectly straight whereas a tortuosity greater than 1 indicates that the duct has an "S" shape. Conicity measures how much the earcanal shrinks in the medial direction. It is computed similarly to Abel et. al. (1990) as the ratio between the cross-sections E and SB areas (S_E/S_{SB}): A ratio close to 1 indicates that the earcanal is non-conical whereas a higher ratio indicates that the earcanal significantly shrinks in the medial direction. The indicator of conicity computed as a simple ratio between the cross-sections E and SB is an important simplification of the morphology of the earcanal. It simply describes the global diminution of earcanal cross-section surface between the cross-sections E and SB. A discussion about the relevance of this indicator can be found in section III.A. Finally, the shape of a cross-section gives an information about its circularity. Usually, cross-sections between E and FB are triangular or elliptical whereas those close to the SB are more circular. The isoperimetric ratio is used to evaluate the circularity of these sections. It is defined as the ratio between the area and the squared perimeter multiplied by four times π and varies between 0 and 1 (the closer to 1, the more circular the section). The aspect ratio of these cross-sections is also computed to quantify their ovality. It is defined as the ratio between the longest and the shortest diameters of the cross-section. Here, a diameter refers to a segment joining two opposite points on the cross-section circumference and passing through its centroid. An example is shown in FIG. 1. where the aspect ratio of the cross-section E is calculated as D_{min}/D_{max}.

All indicators are determined using Polyworks (InnovMetric Logiciels Inc, Canada) and Matlab R2017b (MathWorks, Inc., USA). After a data inspection, two earcanals were discarded from the database because the curvilinear axis could not be computed with the Stinson and Lawton's method. Because cross-section FB determined with the proposed method intersects the concha leading to very unusual shapes and very large perimeters which yielded outliers for the statistical analysis, three more earcanals were removed.

C. Attenuation data acquisition

As mentioned previously, this study uses the secondary data of attenuation measurements collected during a field survey on earplugs comfort. The original project included nine earplugs of different families and different manufacturers but only 6 of them, for which attenuation measurements were carried out, are considered in this secondary study. Of these 6 earplugs, three belong to the "roll-down-foam" earplugs family, one to the "premolded" family and two to the "push-to-fit foam" family. References names of these earplugs can be found in TABLE I. Participants of the original project tested 4 different earplugs models in their work environment for 7 weeks. At the beginning of each week, each worker had a one-on-one meeting with an audiologist to train him/her on the model of earplugs to be tested and to measure and verify the effective wearing of the earplugs. To this purpose, a field attenuation estimation system (FAES), the 3M™ E-A-Rfit™ Dual-Ear Validation System was used as a training tool and attenuation data measurement. This system uses surrogate earplugs (see pictures in FIG. 3 and FIG. 4) to instantly measure and display a personal attenuation rating (PAR) compliant with the ANSI/ASA S12.71 standard (2018). The PAR is the overall average A-weighted attenuation of an earplug for a given fitting in a large ensemble of representative industrial noise spectra (NIOSH 100) (Berger, 2010). This FAES system was chosen because it allows for quick measurements which was an essential selection criterion since training sessions occurred during the participants work shift and had to be limited in time.

TABLE I. Earplugs references.

Earplug family	Roll-down-foam			Premolded	Push-to-fit	
Earplug manufacturer's name	3M™ E-A-R ™ Classic uncorded	3M™ Foam Earplug 1100	3M™ E-A-R™ E-Z-Fit™	3M™ E-A-R™ UltraFit™	3M™ E-A-R™ Push-Ins	3M™ E-A-R™ Push-Ins earplugs, 318-1008, with grip rings
Simplified name in this study	Classic foam	1100 foam	E-Z-Fit foam	Premolded	Push-ins	Push-ins-grip-rings

Two different PARs provided by the FAES are used in this study: the $PAR_{50\%}$ and the $PAR_{84\%}$. The $PAR_{50\%}$ is a median PAR that represents the most statistically probable value of the PAR (Berger and Voix, 2022) and is used in the following to cluster the earcanals (see section II.D). The $PAR_{84\%}$ is computed from the $PAR_{50\%}$ from which uncertainties are subtracted (such as the fit variability that accounts for the fact that the next time the person fits the hearing protector, he or she may do it differently) in order to give a more conservative estimate of the protection that is likely to be achieved on the field (Berger and Voix, 2022). It was therefore used by the audiologists during the training sessions as described in more details in the next paragraph.

Details of the fit training procedure can be found in Martin et al. (2019) and are recalled here for completeness. The audiologist first reminded the worker how to put the earplugs in place, when to replace them and how to check if there was a proper fit. Then, the worker put the surrogate earplugs in place himself (or herself) for a first PAR trial. If both ears had an initial $PAR_{84\%}$ of minimally 50% of the manufacturer's NRR value (considered to as the first threshold value), the worker was considered adequately protected and the individual training was over. If not, the worker was asked to adjust the earplugs for a second PAR trial, still aiming for 50% of the NRR. Since, the $PAR_{84\%}$ data from the FAES takes into account uncertainties that act as a security factor (Berger, 2010), a second threshold value of $PAR_{84\%}$ = 10 dB was accepted. This threshold was chosen because most of workers participating in the study had an average daily sound exposure level for 8 hours less than 95 dBA. If the second trial reached at least this second threshold value of $PAR_{84\%}$ = 10 dB for each ear, the training was over. If this threshold value could not be obtained, a third placement was attempted by the audiologist. If this PAR trial was adequate, the worker was asked to replicate the proper placement to ensure that he or she was able to put it back in place (third trial, and more if needed). This is similar to the method described by Federman & Duhon (2016), where the participants learned successfully to reproduce the adequate placement (and similar PAR) after feeling the correct insertion by an expert. Finally, if both ears did not reach a $PAR_{84\%} \geq 10$ dB for all trials (fitted by the worker), the earplug model was considered unsuitable for this participant's ear(s). Most workers needed between one to three trials per session to properly fit their earplugs. For the roll-down-foam earplugs, 6 trials (for one ear) were sometimes needed. For a few participants, more than 10 trials were required to reach the safe-threshold attenuation values of the training.

For each ear of each worker and for each earplug, the test data leading to the best $PAR_{84\%}$ is kept, and the research team exported the associated $PAR_{50\%}$ value as attenuation data to test the main underpinning hypothesis (see step 0 in FIG. 2) and to evaluate the clusters (see step 4 in FIG. 2). For the ease of reading, in the remainder of the paper, the acronym "PAR", refers to the $PAR_{50\%}$. The distributions of $PAR_{50\%}$ for each earplug are plotted in FIG. 3. By considering both the fitting training process (similar for all participants) and the relatively high PAR values displayed on FIG. 3 (i.e., usually greatly superior to NRR/2 , see section III.A for more details), the research team hypothesized that participants inserted their earplugs correctly so that the inter-individual variability in measured PARs can be mostly primarily attributed to differences in earcanals' morphology and not to other sources of variability related to the psychosocial characteristics of the participant and of his/her work environment (Doutres et al., 2022) (ex., education, gender, support from family /colleagues, type of work, type and frequency of training…). As mentioned previously, this hypothesis is checked in step 0 of the methodology presented in this paper (see sections II.D.1 and III.B).

D. Earcanals clustering

1. Step 0: relations between earcanal morphology and sound attenuation

According to section II.C, the research team hypothesized that the inter-individual variability observed in the measured PARs is mainly induced by the differences in earcanals morphology. To check if this hypothesis is relevant (from the sample to which the methodology is applied in this paper), it is first checked if correlations between morphologic data and attenuation data obtained during the training session exist. To do so, Pearson's correlation coefficients are computed between the morphologic indicators and PARs data using IBM® SPSS® Statistics 27.

2. Step 1: Choice of combinations of morphologic indicators relevant for the clustering

All relevant combinations of input morphologic indicators of the clustering algorithm to be tested are identified based on correlation between morphologic indicators. Accounting for correlations between morphologic data is crucial to avoid choosing a combination of morphologic indicators that are strongly correlated as an input to the clustering algorithm (Negrini et al., 2020). Indeed, if two input morphologic indicators are strongly correlated, they would have a biggest weight in the clustering analysis than other morphologic indicators. To account for the correlations between morphologic indicators, Pearson's correlation coefficient is computed for each pair of morphologic indicators using IBM® SPSS® Statistics 27. Additionally, scatter plots of each pair of morphologic indicators are also drawn to visually check if non-linear correlations (not captured by the Pearson coefficient) between two morphologic indicators exist.

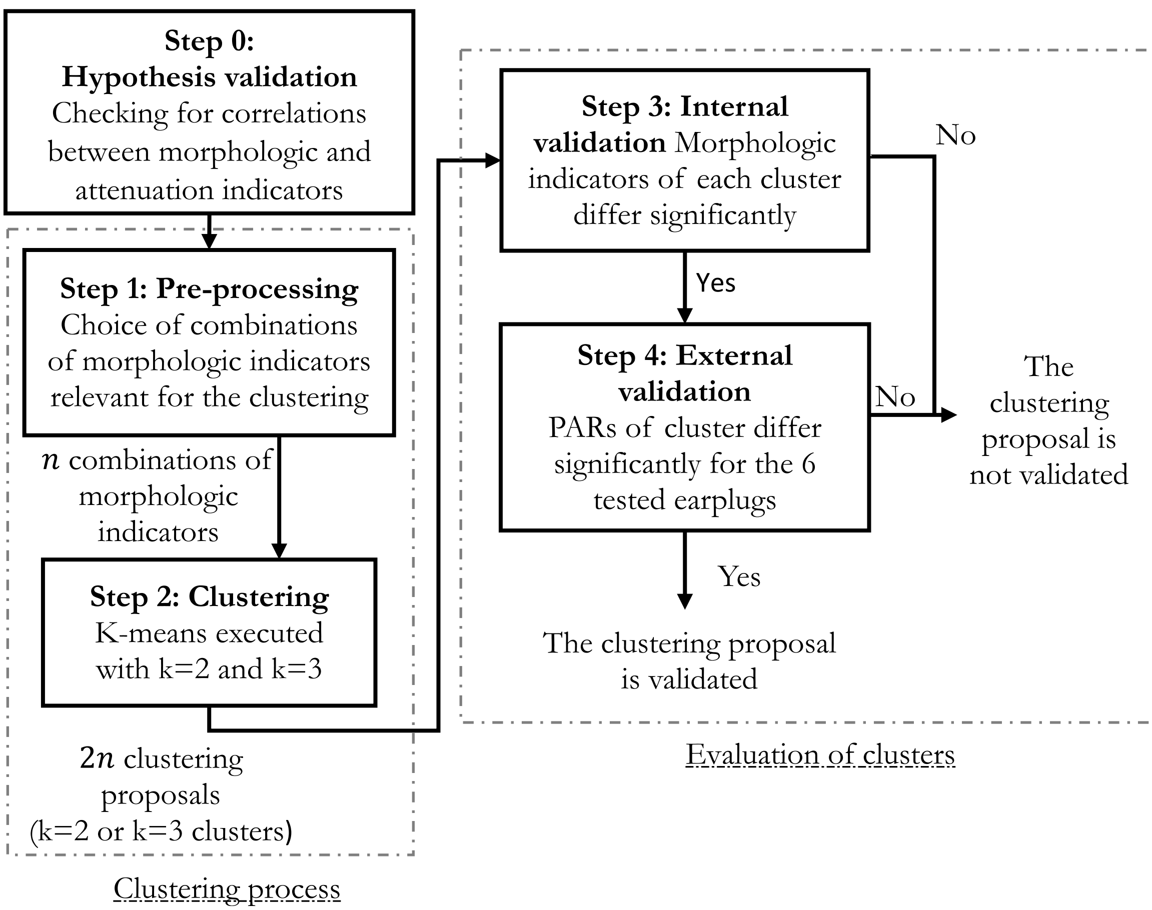

FIG. 2. Description of the clustering process

Following the correlation analysis, the combination of morphologic indicators to cluster earcanals is performed and based on three considerations. Firstly, the correlation between two morphologic indicators in the same combination should not be higher than 0.8. Secondly, as some features (girth and cross-section shapes) are described by several indicators, each combination must not have more than one indicator per feature (not to overweight a feature over the others). Thirdly, each combination must include a girth indicator. This choice is motivated by the objective of building two or three artificial ears to test as much as earplugs as possible. As several commercial earplugs are available in two sizes that differ in diameter, artificial ears should have appropriate earcanal girth to make it possible to test these earplugs.

3. Step 2: clustering algorithm

The k-means clustering algorithm is chosen to classify earcanals. K-means is a partitional algorithm that classifies a set of data points in two phases (Na et al., 2010). The first phase selects k centers randomly, where the value k is fixed in advance. In this work k is forced to be less than 3 for practical and economical reasons associated with the objective of building artificial ears. The next phase is to take each data point to the nearest center. In this study, the Euclidean distance is used to determine the distance between each data point and the cluster centers. When all the data points are included in some clusters, the first step is completed, and an early grouping is done. This iterative process continues repeatedly until a goal function is minimal. Here, the goal function is the sum of the squared distances between each data point and its cluster center. An advantage of k-means over other clustering algorithms, is that it minimizes the dispersion of data points around the cluster centroid and allows for determining the centroid of each cluster (Jain et al., 2000). Knowing the centroid of each cluster is essential to find earcanal morphologies representative of each cluster (for example, an existing earcanal with dimensions close to the centroid of the cluster).

The k-means algorithm is executed with all n selected morphologic indicators combinations (previously selected in step 1) as inputs with k=2 and k=3 clusters and provides $2n$ clustering proposals (n for k=2 plus n for k=3). All these proposals are then evaluated individually to choose the best clustering of earcanals.

4. Steps 3 and 4: clustering evaluation

The individual evaluation of each cluster is based on the following hypothesis: (i) it is possible to cluster 2 or 3 groups of workers' earcanals by combining relevant morphologic indicators; (ii) from these clusters, it is expected to observe significant differences in means showing that the level of PAR varies according to the morphologic indicators that characterize the groupings. The individual evaluation of each cluster proposal is therefore made using two consecutive validation procedures: (i) the internal validation (step 3) and (ii) the external validation (step 4). The internal validation is based on the following criterion: each morphologic indicator used to cluster earcanal must significantly differ from one cluster to another. This first criterion guarantees that artificial ears build based on these clusters will have significantly different morphologies. However, it does not guarantee that these artificial ears will enable to measure earplugs attenuations being different and representative of the inter-individual variability in sound attenuation. A second validation procedure, referred to as the external validation is therefore carried out. This validation is based on the following criterion: mean attenuations (PAR) of the 6 earplugs of this study must significantly differ from one cluster to another. This second criterion is relevant because PAR data are checked to be indeed correlated with earcanal morphology (in step 0), otherwise, significant differences in mean attenuation data of each cluster would not be expected.

Internal and external validations are performed using ANOVA and Bonferroni post-hoc test with a significance level set at 0.05.

III. RESULTS AND DISCUSSION

A. Data description

Descriptive statistics of morphologic data measured on the sample of a population of Canadian workers consisting of 237 earcanals are summarized in TABLE II.

TABLE II. Morphologic dimensions of earcanals and corresponding indicators names and descriptive values.

Dimension	Features	Indicator (s)	Earcanal region	Symbol	Mean	Median	Std	Min	Max
Size	Length	Curvilinear length (mm)	Between E and SB	L_{E-SB}	13.3	13.3	2.3	7.8	19.6
	Girth	area (mm^2)	Cross-section E	S_E	104.3	102.7	22.3	43.3	203.2
			Cross-section FB	S_{FB}	75.6	73.2	19.0	33.8	124.8
			Cross-section SB	S_{SB}	62.3	60.5	19.5	21.6	117.5
		Circumference (mm)	Cross-section E	C_E	39.6	39.9	4.50	23.9	52.1
			Cross-section FB	C_{FB}	32.2	32.4	4.2	21.4	42.4
			Cross-section SB	C_{SB}	28.6	28.6	4.6	17.1	39.5
Shape	Sections' shape	Isoperimetric ratio $4\pi \frac{Surface}{Circumference^2}$	Cross-section E	IR_E	0.83	0.84	0.07	0.62	0.96
			Cross-section FB	IR_{FB}	0.91	0.92	0.05	0.72	0.98
			Cross-section SB	IR_{SB}	0.93	0.94	0.04	0.79	0.99
		Aspect ratio D_{min}/D_{max}	Cross-section E	AR_E	0.64	0.62	0.12	0.32	0.96
			Cross-section FB	AR_{FB}	0.62	0.61	0.12	0.35	0.98
			Cross-section SB	AR_{SB}	0.72	0.71	0.11	0.45	0.99
	Tortuosity	Curvilinear length over Euclidian length	Between E and SB	T	1.06	1.06	0.03	1.01	1.19
	Conicity	area of E over area of SB $\frac{S_E}{S_{SB}}$	Between E and SB	$F_{E/SB}$	1.81	1.68	0.61	0.89	5.48

The earcanal size dimension is quantified through 2 features: the length and the earcanal girth. The length is comprised between 7.8 and 19.6 mm. The earcanal girth is quantified through 2 indicators that are the area and the circumference, both measured at the three cross-sections E, FB and SB. Their means (and standard deviations) circumferences are respectively C_E = 39.6 mm (s.d = 4.5 mm), C_{FB} = 32.2 mm (s.d = 4.2 mm) and C_{SB} = 28.6 mm (s.d = 4.6 mm) and their areas are S_E = 104.3 mm^2 (s.d = 22.3 mm^2), S_{FB} = 75.6 mm^2 (s.d = 19.0 mm^2) and S_{SB} = 62.3 mm^2 (s.d = 19.5 mm^2). As expected, the earcanal shrinks in the medial direction ($C_E > C_{FB} > C_{SB}$), confirmed by the conicity indicator $F_{E/SB}$ that is larger than 1. Other shape dimension indicators indicate that the earcanal becomes more circular in the medial direction ($IR_E < IR_{FB} < IR_{SB}$). The aspect ratio of cross-sections E and FB are similar, whereas that of cross-section SB is larger. Cross-sections E and FB differ in terms of their iso-perimetric ratio but have similar aspect ratios. This is because cross-section E is shaped like a triangle whereas cross-section FB (and SB) is shaped like an ellipse. Consequently, the aspect ratio and the iso-perimetric ratio are complementary to describe cross-section E. Overall, this dataset confirms the general description of an earcanal given in Alvord and Farmer (1997).

The medians of the distribution of the best PARs obtained during the fit training vary between 23 and 37 dB depending on the earplug. The histograms of PAR data are plotted in FIG. 3 and show that except for the push-ins grip-rings earplug, most workers were able to obtain a high PAR during the training session. Indeed, most of workers obtained PARs highly superior to 50% of the NRR values of the earplugs which is a typical derating factor applied to the earplugs NRR for estimating average protection levels for groups of users (see table 2 of the CSA Z94.2-14 standard). Considering that the workers received about 5 trainings in the insertion of disposable and reusable earplugs during the field study (see section II.C.), and that they obtained rather high PARs values after the training, it can be considered that the training sessions greatly reduced the inter-individual variability in sound attenuation related to psychosocial characteristics of the user and of his/her work environment (e.g., education, type and frequency of training…) (Doutres et al., 2022). It is therefore reasonable to hypothesize that the inter-individual variability observed in the PARs measured is mainly induced by the differences in the morphology of the earcanals (this hypothesis is checked during the step 0 presented in the next subsection).

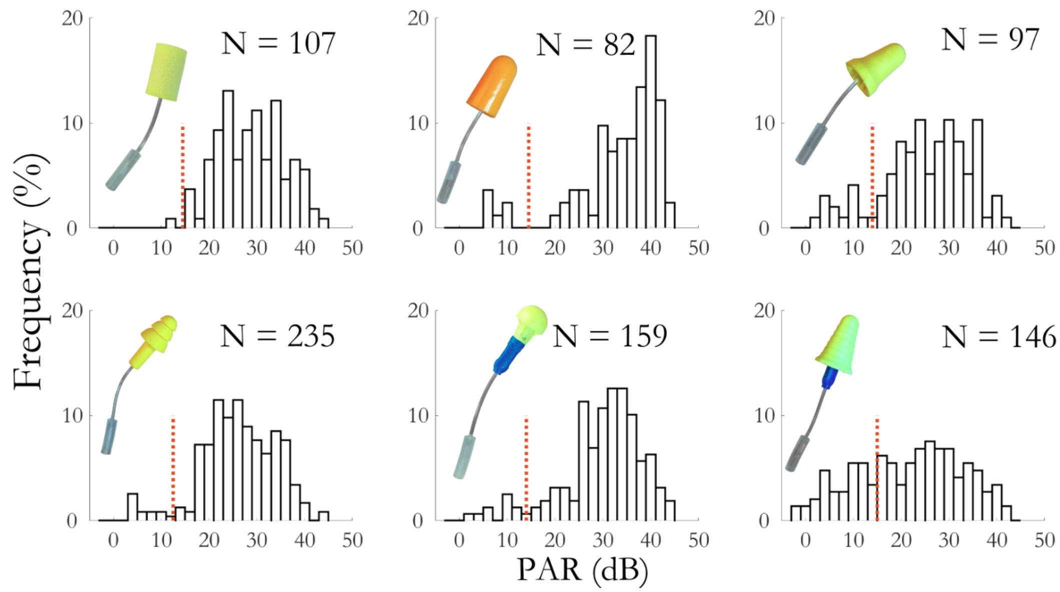

FIG. 3. Distribution of the best PARs obtained during a fit training for the 6 earplugs (clockwise from top left): Classic foam, 1100 foam, E-Z-Fit foam, premolded, Push-Ins and Push-Ins-Grip-Rings. Orange dotted lines show half of the NRR of each earplug.

Low and negative PARs values observed on the push-ins-grip-rings earplug histogram suggest that a certain number of workers cannot fit properly the push-ins-grip-rings earplug resulting in leaks and a poor attenuation. Large leaks may indeed act as a Helmholtz resonator and provide a gain effect in the low to middle frequencies range (Berger, 2014). The fact that some workers were not able to obtain a safe PAR, even with a fit training, is consistent with the statement of Franks et al. (1996): "Not every person can wear every hearing protector. Some people may be unable to wear certain types of earplugs because of the shape or size of their earcanals".

B. Step 0: Relations between earcanal morphology and sound attenuation

Correlations between earcanals morphology and PARs are evaluated (TABLE III) to confirm that the inter-individual variability in sound attenuation is related to the earcanal morphology and that the external validation described in section II.D is relevant on this dataset that characterizes a sample of Canadian worker' earcanals.

TABLE III. Pearson linear correlation between morphologic parameters of earcanals and maximum PAR obtained with trained participant fitting himself/herself its earplug. Dark gray boxes highlight a correlation higher than 0.4, gray boxes highlight a correlation between 0.3 and 0.4, white boxes highlight a correlation smaller than 0.3. Empty boxes indicate that the correlation between two variables is not significant at the level 0.05.

Morphologic parameters		Personal attenuation rating					
		Malleable			Premolded	Push-to-fit	
Position in the earcanal	morphologic indicator	Classic foam	1100 foam	E-Z-Fit foam	Premolded	Push-ins	Push-ins-grip-rings
Entrance cross-section	C_E	-.195*		-.298**	-.359**	-.234**	-.285**
	S_E				-.302**	-.246**	-.257**
	AR_E		.273*				
	IR_E	.269**		.330**	.198**		
First bend cross-section	C_{FB}		-.355**	-.330**	-.418**	-.311**	-.362**
	S_{FB}		-.281*	-.261**	-.413**	-.332**	-.340**
	AR_{FB}	.235*					
	IR_{FB}	.228*	.292**	.327**			
Second bend cross-section	C_{SB}		-.335**	-.308**	-.478**	-.347**	-.410**
	S_{SB}		-.226*	-.270**	-.470**	-.352**	-.381**
	AR_{SB}			.304**			.182*
	IR_{SB}	.228*	.649**	.306**			.211*
Along earcanal	L_{E-SB}	.223*				.177*	
	$F_{E/SB}$				.260**		.221**
	T						-.209*

*. The correlation is significant at the level 0.05 (bilateral).

**. The correlation is significant at the level 0.01 (bilateral).

TABLE III suggests that the girth of FB and SB cross-sections are moderately but significantly correlated to the sound attenuation of the push-to-fit and premolded earplugs (Pearson correlations coefficients inferior to 0.5). A significant correlation between these sections' girths and attenuation of two malleable earplugs is also found. These correlations are negative, which means that the larger the earcanal, the lower the attenuation. It can be hypothesized that a large earcanal leads to a lower compression of the earplug and surrounding tissues. As at low frequencies, the vibro-acoustic behaviour of the earplug coupled to the earcanal is governed by the equivalent rigidity of the system {earplug + earcanal skin} (Sgard et al. 2011), a lower earplug/skin compression induces a lower equivalent rigidity, and a lower sound attenuation. A lower mechanical pressure between earcanal skin and earplug may also introduce acoustic leakage.

Weak but significant correlations between the PAR and the cross-sections FB and SB shapes (IR_{FB}, IR_{SB} AR_{FB} and AR_{SB}) are also found especially with roll-down foam earplugs, except for IR_{SB} for which the correlation with the PAR of the 1100 foam earplug is fairly high (between 0.5 and 0.8). Correlations between sections shapes indicators and PAR are positive, meaning that the more circular the earcanal, the higher the PAR. It could be hypothesized that a circular earcanal allows for a better contact between earplug and earcanal walls, which avoids leaks between the earplug and the skin, leading to a higher attenuation. Lower but significant correlations between cross-section E size and shape and PAR of earplugs are observed.

The conicity is only correlated to the attenuation of the premolded and the push-ins-grip-rings earplugs (the more conical the earcanal, the higher the PAR). These two earplugs have the most conical shapes of the 6 earplugs, and it can be hypothesized that they better match the geometry of conical earcanals than straight cylindrical earcanals (because the contact surface between the earplug and the earcanal would be higher in the first scenario). As described in the methodology section, the conicity computed as the ratio between the surfaces of the cross-sections E and SB is an important simplification of the morphology: it does not describe how the cross-sections area changes in the medial direction (linearly or exponentially for example), and it is computed between two cross-sections that are not necessarily involved in the earplugs fit (but correlated with earplugs attenuation). In a preliminary study not shown in this paper, the conicity has also been computed as the ratio between the cross-sections E – FB and FB – SB. These two additional indicators were however shown to be less relevant for this study because they were not or very poorly correlated to the earplug's attenuation. Finally, the conicity indicator computed between cross-sections E and SB seems relevant to be included in the clustering process of this study because it is significantly correlated to the attenuation of two conical earplugs.

As for the parameters of length and tortuosity, they are poorly but statistically correlated to the attenuation of the Classic foam and push-ins earplug (length indicator) and the push-ins-grip-rings earplug (tortuosity indicator). Conversely, Abel et al., (1990) found a high correlation between tortuosity and attenuation of earplugs. This could be due to the fact that Abel et al., evaluated the tortuosity subjectively and selected only the 17^{th} most straight and the 18^{th} most twisted earcanals (over the 186 of his study) to compute Pearson's coefficient. Taking extrema values favour high linear correlation coefficients.

Finally, correlations between morphologic indicators and attenuations of the six earplugs given in Table III show that a given morphologic indicator is not equally relevant for the attenuation of different earplugs models. This underlines the interest of choosing indicators that characterize the open earcanal (step 3, internal validation), and then, to study the correlation with the attenuation (step 4, external validation) in order to build artificial ears dedicated to the measurement of the attenuation of a multitude of earplugs. Overall, correlations suggest that the morphologic variability of the earcanals induces a variability in the sound attenuation of earplugs correctly inserted. Therefore, it seems relevant to use attenuation data to validate clustering proposal (step 4). It is reasonable to expect that mean attenuations of clusters classified using morphologic data will differ significantly.

C. Step 1: Choice of combinations of morphologic indicators relevant for the clustering

To choose relevant combination of morphologic indicators as input for the k-means clustering algorithm, correlation coefficients are checked. Correlations between all morphologic indicators of this study are presented in TABLE IV.

TABLE IV shows that the two indicators of girth (i.e., circumference C and area S) of a given cross-section have Pearson coefficients higher than 0.8 (see blue border boxes), indicating that they are highly correlated. Consequently, with the objective of choosing morphologic indicators combinations as input for the clustering algorithm, a given combination should include either the circumference or the area indicators but not both. Otherwise, the girth feature would have more weight that other features in a given combination. Correlations between the girths of SB and E cross-sections are between 0.45 and 0.5 (orange border boxes) and the correlations between the girths of couples {FB, E} on the one hand and {FB, SB} on the other hand are close to 0.6 (green border boxes). Consequently, the girth of all the earcanal can be fairly well described by the FB cross-section only. A similar conclusion can be drawn from the shape features of cross-sections E, FB and SB.

TABLE IV. Pearson linear correlation between different morphologic indicators of earcanals. Dark gray boxes highlight a correlation higher than 0.5, gray boxes highlight a correlation between 0.3 and 0.5, white boxes highlight a correlation smaller than 0.3. Empty boxes indicate that the correlation between two variables is not significant at the level 0.05.

	S_E	AR_E	IR_E	C_{FB}	S_{FB}	AR_{FB}	IR_{FB}	C_{SB}	S_{SB}	AR_{SB}	IR_{SB}	L_{E-SB}	$F_{E/SB}$	T
C_E	0.92**		-0.35**	0.67**	0.65**			0.49**	0.48**			0.16*		0.25**
S_E		0.15*		0.66**	0.69**			0.47**	0.46**				0.19**	0.18**
AR_E			0.26**			0.28**	0.17**							
IR_E						0.55**	0.50**	-0.14*	-0.13*			-0.23**	0.18**	-0.23**
C_{FB}					0.97**		-0.19**	0.61**	0.58**		-0.18**		-0.14*	
S_{FB}								0.61**	0.59**		-0.13*			
AR_{FB}							0.43**							
IR_{FB}										0.32**	0.20**			
C_{SB}									0.98**			-0.15*	-0.72**	
S_{SB}													-0.71**	
AR_{SB}											0.43**			
IR_{SB}												0.25**	-0.14*	
L_{E-SB}													0.13*	0.21**
$F_{E/SB}$														

**. The correlation is significant at the level 0.01 (bilateral).

*. The correlation is significant at the level 0.05 (bilateral).

Considering that the earcanal girth is better represented by FB cross-section than E and SB ones, it is selected to calculate girth (S_{FB}, C_{FB}) and shape indicators (IR_{FB}, AR_{FB}). Either C_{FB} or S_{FB} are chosen as cross-section girth indicator and IR_{FB} or AR_{FB} as cross-section shape indicator. In the sample of earcanals used in this study, there are 48 combination of morphologic indicators that respect all criteria for the input combinations of the k-means algorithm. These 48 combinations of morphologic indicators are summarized in TABLE V. To check that there is no multicollinearity between morphologic indicators of a same combination, variable inflation factors (VIFs) are computed between all morphologic indicators. It is found that no VIFs are higher than 5 if the surface and the circumference of the cross-section FB are not together in the list of morphologic indicators. As no combination includes these two morphologic indicators together, the research team concludes that there is no multicollinearity between morphologic indicators of a combination used as input for the clustering algorithm.

D. Steps 3 and 4: Clusters evaluation

As described in section II.D.4, the evaluation of earcanals clustering is based on a two-step evaluation for each clustering proposal: the internal and external validations. This two-step evaluation is performed for both k=2 and k=3 earcanals clusters.

TABLE V summarizes the validation process for the 48 proposals of earcanals classifications in two different clusters (k=2). The second column "combination of morphologic indicators" contains all the 48 combinations of morphologic data selected as inputs for the k-means clustering algorithm. The third column "Internally validated? T-test" indicates if the morphologic indicators of a given combination are statistically different from one cluster to another. If the answer is "Yes" the external validation is performed. The next 6 columns display the p-values of the ANOVAs performed on PAR of the 6 tested earplugs. If the 6 p-values are below the significance threshold of 0.05, the clustering proposal is considered validated according to the external validation procedure.

TABLE V. Cluster evaluation for k=2 clusters. Gray boxes indicate that the p-value of the external validation ANOVA is significant at the level 0.05.

N°	Combination of morphologic indicators	Internally validated? T-test	External validation p-value of the ANOVA on earplugs PAR					
			Classic foam	1100 foam	E-Z-Fit foam	premolded	Push-ins	Push-ins-grip -rings
1	C_{FB}	Yes	0.409	0.021	0.003	0.000	0.004	0.000
2	S_{FB}	Yes	0.823	0.031	0.011	0.000	0.005	0.000
3	C_{FB} AR_{FB}	Yes	0.003	0.064	0.002	0.000	0.085	0.002
4	C_{FB} $F_{E/SB}$	Yes	0.698	0.001	0.010	0.000	0.005	0.000
5	C_{FB} T	No						
6	C_{FB} L_{E-SB}	Yes	0.023	0.020	0.166	0.000	0.000	0.007
7	C_{FB} IR_{FB}	Yes	0.001	0.017	0.000	0.000	0.061	0.042
8	S_{FB} AR_{FB}	Yes	0.891	0.007	0.002	0.000	0.001	0.001
9	S_{FB} $F_{E/SB}$	Yes	0.858	0.008	0.010	0.000	0.004	0.000
10	S_{FB} T	Yes	0.972	0.025	0.009	0.000	0.006	0.000
11	S_{FB} L_{E-SB}	No						
12	S_{FB} IR_{FB}	No						
13	C_{FB} AR_{FB} $F_{E/SB}$	Yes	0.217	0.084	0.009	0.000	0.020	0.001
14	C_{FB} AR_{FB} T	Yes	0.002	0.055	0.022	0.000	0.190	0.001
15	C_{FB} AR_{FB} L_{E-SB}	No						
16	C_{FB} $F_{E/SB}$ T	No						
17	C_{FB} $F_{E/SB}$ L_{E-SB}	Yes	0.199	0.018	0.108	0.000	0.002	0.002
18	C_{FB} T L_{E-SB}	Yes	0.285	0.800	0.789	0.659	0.000	0.727
19	C_{FB} IR_{FB} $F_{E/SB}$	Yes	0.197	0.004	0.012	0.000	0.006	0.007
20	C_{FB} IR_{FB} T	Yes	0.005	0.016	0.003	0.000	0.031	0.177
21	C_{FB} IR_{FB} L_{E-SB}	Yes	0.022	0.011	0.012	0.000	0.001	0.009
22	S_{FB} AR_{FB} $F_{E/SB}$	No						
23	S_{FB} AR_{FB} T	No						
24	S_{FB} AR_{FB} L_{E-SB}	Yes	0.156	0.262	0.475	0.258	0.018	0.743
25	S_{FB} $F_{E/SB}$ T	No						
26	S_{FB} $F_{E/SB}$ L_{E-SB}	Yes	0.307	0.038	0.139	0.000	0.004	0.000
27	S_{FB} T L_{E-SB}	Yes	0.274	0.802	0.429	0.809	0.001	0.508
28	S_{FB} IR_{FB} $F_{E/SB}$	No						
29	S_{FB} IR_{FB} T	Yes	0.847	0.347	0.286	0.000	0.032	0.000
30	S_{FB} IR_{FB} L_{E-SB}	No						
31	C_{FB} AR_{FB} $F_{E/SB}$ T	No						
32	C_{FB} AR_{FB} $F_{E/SB}$ L_{E-SB}	No						

33	C_{FB}	AR_{FB}	T	L_{E-SB}		No						
34	C_{FB}	$F_{E/SB}$	T	L		Yes	0.428	0.572	0.411	0.011	0.000	0.084
35	C_{FB}	IR_{FB}	$F_{E/SB}$	T		Yes	0.345	0.006	0.092	0.000	0.013	0.080
36	C_{FB}	IR_{FB}	$F_{E/SB}$	L_{E-SB}		Yes	0,011	0.019	0.010	0.000	0.000	0.010
37	C_{FB}	IR_{FB}	T	L_{E-SB}		Yes	0,005	0.041	0.093	0.002	0.000	0.143
38	S_{FB}	AR_{FB}	$F_{E/SB}$	T		Yes	0.580	0.003	0.056	0.000	0.002	0.000
39	S_{FB}	AR_{FB}	$F_{E/SB}$	L_{E-SB}		Yes	0,535	0.017	0.068	0.000	0.002	0.000
40	S_{FB}	AR_{FB}	T	L_{E-SB}		No						
41	S_{FB}	$F_{E/SB}$	T	L_{E-SB}		Yes	0,263	0.394	0.197	0.005	0.000	0.023
42	S_{FB}	IR_{FB}	$F_{E/SB}$	T		Yes	0.857	0.014	0.090	0.000	0.002	0.000
43	S_{FB}	IR_{FB}	$F_{E/SB}$	L_{E-SB}		No						
44	S_{FB}	IR_{FB}	T	L_{E-SB}		No						
45	C_{FB}	AR_{FB}	$F_{E/SB}$	T	L_{E-SB}	No						
46	C_{FB}	IR_{FB}	$F_{E/SB}$	T	L_{E-SB}	Yes	0.004	0.072	0.093	0.000	0.000	0.163
47	S_{FB}	AR_{FB}	$F_{E/SB}$	T	L_{E-SB}	No						
48	S_{FB}	IR_{FB}	$F_{E/SB}$	T	L_{E-SB}	No						

According to TABLE V, 29 combinations of morphologic indicators passed the internal validation step. Each of these 29 combinations are then tested with the external validation procedure with the objective to select a clustering proposal for which attenuations significantly differ from one cluster to another. This external validation is much more restrictive. Looking at grey boxes in TABLE V, it is worth noting that roll-down foam earplugs especially invalidate a lot of clustering proposals (this earplug is very restrictive for the external validation). The two push-to-fit and the premolded earplugs are much less restrictive. Indeed, the two push-to-fit earplugs invalidate only 9 clustering proposals over the 29 combinations internally validated whereas the Classic foam earplug invalidate 20 clustering proposals. Interestingly, earplugs for which PAR are poorly or not correlated to earcanals morphology invalidate more clustering proposals than earplugs for which PAR are moderately to highly correlated to earcanals morphology. Indeed, as the clustering is based upon morphologic classification, it is expected that earplugs PARs significantly correlated to earcanals morphology may have significantly different means between clusters. This supports the interest of an external validation based on attenuation data in the objective of building artificial ears for attenuation measurements.

Finally, only two clustering proposals lead to significantly different attenuations for all 6 earplugs. These two combinations are: $\{C_{FB} - IR_{FB} - L\}$ (line 21 of TABLE V) and $\{C_{FB} - IR_{FB} - L - F_{E/FB}\}$ (line 36 of TABLE V). These combinations are very close to each other, the only difference being the earcanal conicity which is present only in the second combination. As the objective is to design artificial ears representative and different between two clusters for a maximum of morphologic dimensions, it is the second proposal of clustering taking into account 4 morphologic dimensions that is retained (for k=2 clusters).

For this kept clustering proposal $\{C_{FB} - IR_{FB} - L_{E-SB} - F_{E/FB}\}$, TABLE VI shows that the cluster 0 comprises the largest earcanals (leading to the lower attenuation as shown in FIG. 4) and the one with the lower iso-perimetric ratios (also leading to the lower attenuation as presented in FIG. 4). Therefore, there is a double effect of morphology on attenuation for this clustering proposal: the most circular and smallest earcanals have the best attenuation whereas the more oval and larger earcanals have the poorer sound attenuation. This double effect explains why attenuations of the cluster 0 and cluster 1 differ significantly.

TABLE VI. Comparison of means of morphologic indicators between the two clusters of the best clustering proposal with k = 2.

Class	Earcanals number	C_{FB} (mm)	IR_{FB}	$F_{E/SB}$	L_{E-SB} (mm)	
0	83	35.2	.87	1.53	11.8	Mean
		3.2	.057	.42	2.0	std
1	154	30.5	.92	1.96	14.1	Mean
		3.6	.04	.64	2.03	std

It should be noted that for the sample of earcanals presented here, the cluster of largest earcanals also comprises the shortest earcanals. Finally, the most conical and straight cylindrical earcanals are grouped in cluster 1 and cluster 0 respectively.

For this clustering proposal, the PAR significantly differs from one cluster to another for the 6 earplugs as shown in FIG. 4.

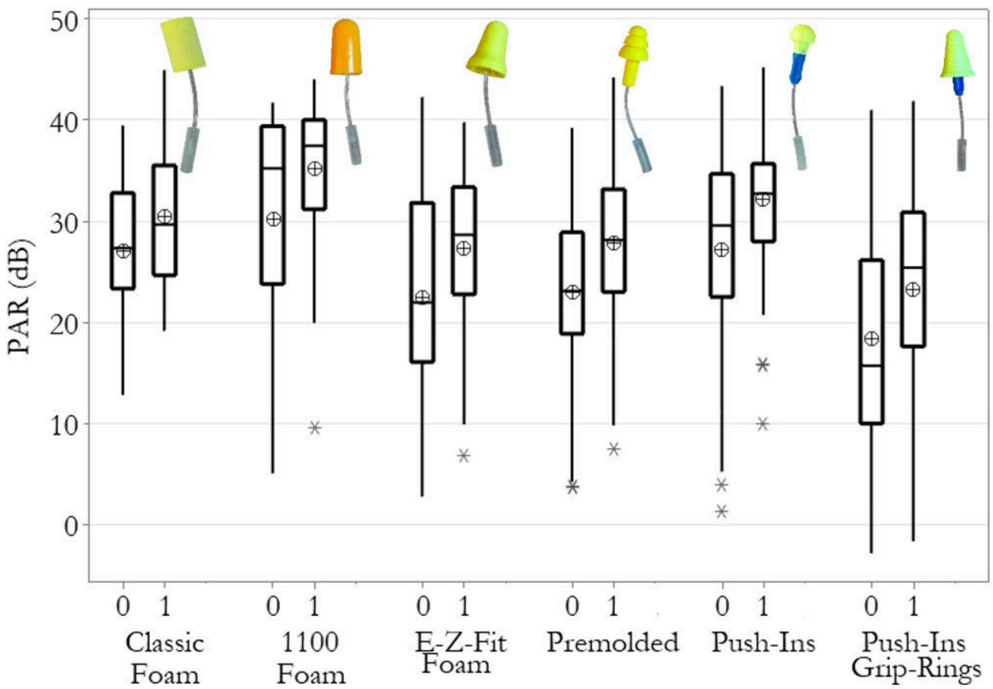

FIG. 4. Box plot of the PAR of 6 earplugs for the two clusters of the best clustering proposal for k=2

Regarding the evaluation of clustering proposal for k=3: the same two-steps validation process as for k=2 is followed. The only difference is that the internal validation is based on the Bonferroni post-hoc test. This is motivated by the fact that there are now three clusters of earcanals. This post-hoc test allows a pairwise comparison of clusters. The results of the external validation for the 12 combinations that passed the internal validation are listed in TABLE VII.

TABLE VII. Cluster evaluation for k=3 clusters. Only clusters that satisfied the internal validation (Bonferroni post-hoc test) are plotted in this table. Gray boxes indicate that the p-value of the external validation ANOVA is significant at the level 0.05.

N°	**Combination of Morphologic indicators**	**External validation p-value of the ANOVA on earplugs PAR**					
		Classic foam	**1100 foam**	**E-Z-Fit foam**	**premolded**	**Push-ins**	**Push-ins-grip -rings**
1	C_{FB}	0.015	0.135	0.227	0.000	0.001	0,029
2	S_{FB}	0.907	0.001	0.009	0.000	0.000	0,000
3	C_{FB} IR_{FB}	0.070	0.051	0.004	0.000	0.028	0,020
4	C_{FB} $F_{E/SB}$	0.317	0.006	0.007	0.000	0.004	0,001
5	C_{FB} T	0.586	0.061	0.001	0.000	0.015	0,000
6	C_{FB} L_{E-SB}	0.192	0.023	0.003	0.000	0.009	0,002
10	S_{FB} T	0.524	0.169	0.002	0.000	0.004	0,000
15	C_{FB} AR_{FB} L_{E-SB}	0.022	0.173	0.021	0.000	0.026	0,032
16	C_{FB} $F_{E/SB}$ T	0.609	0.034	0.012	0.000	0.057	0,000
21	C_{FB} IR_{FB} L_{E-SB}	0.045	0.005	0.017	0.000	0.004	0,002
32	C_{FB} AR_{FB} $F_{E/SB}$ L_{E-SB}	0,344	0.159	0.027	0.000	0.030	0.003
33	C_{FB} AR_{FB} T L_{E-SB}	0,006	0.754	0.019	0.000	0.003	0.064

As for the external validation, an ANOVA is performed on earplugs PARs. The same trends than for k=2 clusters are observed. The two push-to-fit earplugs PAR significantly differ for most of clustering proposals. All clustering proposal internally validated are also validated with the PAR for the premolded earplug. Foam earplugs, however invalidated several clustering proposals, especially the classic foam earplug which invalidate 8 out of the 12 internally validated clustering proposals.

Finally, only one combination of morphologic indicators provides clusters that meet both the external and the internal validation: $\{C_{FB} - IR_{FB} - L_{E-SB}\}$ (line 21 in TABLE VII). The unique combination of morphologic indicators ($\{C_{FB} - IR_{FB} - L_{E-SB}\}$) that satisfies both validation criteria for k=3 clusters also meets both criteria for k=2 clusters (see line 21 of TABLE V).

As seen in TABLE VIII, the number of earcanals per cluster is well balanced for this clustering proposal. Again, there is a double effect of morphology on attenuation. Cluster 0 includes the earcanals for which the girth is the smallest (higher attenuation as seen in FIG. 5) and earcanals with the highest iso-perimetric ratios (higher attenuation as seen in FIG. 5). Cluster 2 comprises the largest and most oval earcanals. Finally, cluster 1 is in the middle of these two clusters for these two indicators (C_{FB} and IR_{FB}). This double effect explains why attenuations of clusters 0, 1 and cluster 2 differ significantly.

TABLE VIII. Comparison of means of morphologic indicators between the three clusters of the best clustering proposal with k = 3

Class	Earcanals number	C_{FB} (mm)	IR_{FB}	L_{E-SB} (mm)	
0	92	28.5	0.93	12,9	Mean
		28.2	0.93	13.0	Median
		2.6	0.03	1.6	std
1	69	33.4	0.91	15.7	Mean
		33.5	0.91	15.7	Median
		2.9	0.04	1.3	std
2	76	35.4	0.88	11.5	Mean
		35.1	0.88	11.7	Median
		3.0	0.06	1.6	std

FIG. 5, shows boxplots of the best PARs obtained during the fit training for the three clusters and each earplug. The sound attenuations of earplugs in cluster 0 are overall, higher (significantly at the level 0.05) than those in cluster 2. However, sound attenuations of earplugs in cluster 1 do not necessarily differ from those in other clusters. It is important to recall that attenuations have not been used as an input to cluster earcanals. The difference of attenuation of different clusters is just a consequence of the correlation between morphology and attenuation. A Bonferroni post-hoc test (not shown in this paper) has been conducted for sound attenuations of all earplugs. For the premolded earplug, attenuations in each cluster significantly differ from one another (at the level 0.032 between clusters 0 and 1, level 0.006 between clusters 1 and 2 and level <0.001 between clusters 0 and 2). Consequently, with the objective to build artificial ears for the measurement of attenuation on a maximum of earplug types, it seems relevant to use 3 different clusters of earcanals. Finally, it is this final clustering proposal, obtained with the k-means algorithm with morphologic indicators $\{C_{FB} - IR_{FB} - L_{E-SB}\}$ and k=3 different clusters of earcanals that seems the most relevant to help the design of realistic artificial ears dedicated to earplug measurement attenuation.

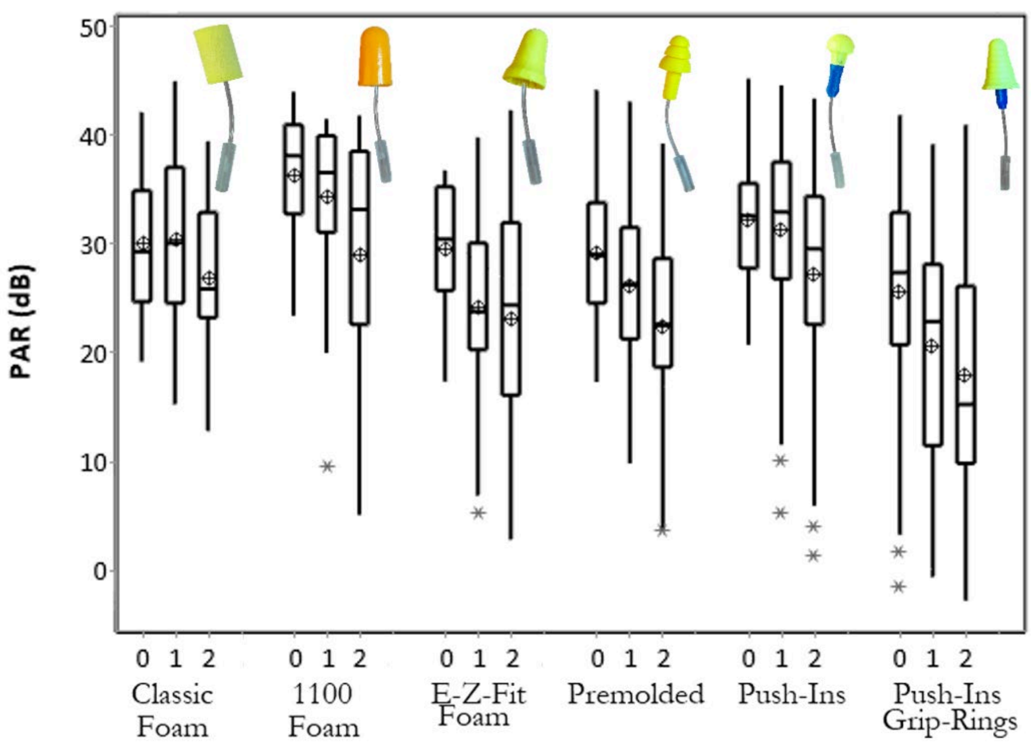

FIG. 5. Box plot of the PAR of 6 earplugs for the three clusters of the best clustering proposal for k=3

In order to check that the final k value of 3 achieves an optimal solution, complementary analysis (not associated with the objective of building 2 or 3 artificial ears) have been conducted. The same clustering methodology as for k=2 and k=3 has been applied with k=4. For 4 clusters, only two combinations of morphologic indicators successfully passed the internal validation ($\{S_{FB}\}$ and $\{S_{FB}$; $AR\}$). None of these two combinations successfully passed the external validation. This strongly suggests that the final clustering proposal obtained with k=3 clusters is an optimal solution.

E. Limits

Limitations of the clustering methodology and its application to a sample of earcanals are identified in this section.

The proposed methodology being applied on a limited number of earcanals, statistical limitations associated with generalizing results from a sample apply.

A single process of clustering and validation procedure is performed to cluster earcanals. Other clustering algorithms and/or statistical tests to validate clusters could have been used and may have lead to another clustering structure of earcanals. The method presented here makes it possible to select a clustering of earcanals relevant as a basis for the design of artificial ears dedicated to sound attenuation measurement.

The description of earcanals' morphology is here limited to 15 morphologic indicators (7 size indicators and 8 shapes indicators), that describe the earcanal portion where the earplugs are supposed to be fitted (between the entrance and the second bend). It is therefore assumed that these indicators are sufficient to comprehensively describe the earcanal morphology. Other anatomical properties that may be also responsible for inter-individual variability in sound attenuation such as mechanical properties of ear tissues, the position of the cartilaginous/bony junction or eardrum impedance are not considered here (note that some of them can be difficult to determine in the field, or even impracticable).

Comparison of earcanal morphologic differences between studies is complicated because methods to extract morphologic indicators differ and are not always objective. In this paper, the proposed method to extract the morphologic indicators has been designed to be as objective as possible (i.e. reducing the number of manually placed landmarks to locate characteristic cross-sections of the earcanal). However, this method is based on the use of cross-sections perpendicular to the curvilinear axis, which may not be equally relevant for all earplugs considered in this study. For example, the radial axis about which the flanges of a premolded earplug extend might not be centered on the curvilinear axis.

The earplugs insertion depth is unknown and a better knowledge of the position of each earplug in each ear could have been helpful to identify the most relevant cross-sections to be correlated with the measured sound attenuation.

In addition, the type of training used in the original field study has led to a PAR value that was considered high enough to assume that that measured inter-individual variability in PAR could mainly be attributed to the morphologic differences between earcanals. However, it can be hypothesized that the correlations between morphologic indicators and PAR could have been higher if the training session was designed specifically for this study (or if an experimenter fitting was performed for PAR measurements) and thus would have targeted the maximum PAR achievable for a given earplug.

IV. CONCLUSION

Most of existing acoustical test fixtures (ATFs) dedicated to earplugs sound attenuation measurement are equipped with unique sized straight cylindrical earcanals, considered as representative averaged morphology of humans, and thus are unable to assess how earplugs can fit different earcanal morphologies.

In this paper, a methodology to cluster earcanal as a function of their morphologies with the objective of designing artificial ears dedicated to sound attenuation measurement is developed and applied to a sample of Canadian workers' earcanals. Morphologic indicators are measured/computed on earmolds of earcanals and attenuation of 6 different earplugs were measured on these same earcanals. An artificial intelligence-based algorithm and statistical analysis were used to assess earcanal clusters the most relevant to help the design of realistic artificial ears dedicated to earplug attenuation measurement. The morphologic data of the population sample considered in this study proved to be consistent with the literature and significant correlations between some morphologic indicators and attenuation of earplugs were found. Considering this population sample, the best clustering proposal was obtained using the three following morphologic indicators as input for the k-means algorithm: (i) circumference of the first bend cross-section (ii) isoperimetric ratio of the first bend cross-section and (iii) length between the entrance and the second bend. This clustering proposal consists of three different clusters of earcanals. It was found that the cluster that comprises earcanals of smallest girth and the most circular is also the cluster where measured PAR are the highest, whereas the cluster that includes the largest and most oval earcanals has low measured PAR. This observation is coherent with the correlation morphology/attenuation observed both in the literature and confirmed by this study.

ACKNOWLEDGMENTS

The authors acknowledge the support of the Institut de recherche Robert-Sauvé en santé et en sécurité du travail (IRSST) (funding reference number 2015-0014) and the MITACS Accelerate program (funding reference number IT10643).

REFERENCES (BIBLIOGRAPHIC)

Abel, S. M., Alberti, P. W., & Rokas, D. (**1988**). "Gender differences in real-world hearing protector attenuation", The Journal of Otolaryngology, 17(2), 86-92.

Abel, S. M., Rockley, T., Goldfarb, D., & Hawke, M. (**1990**). "Outer ear canal shape and its relation to the effectiveness of sound attenuating earplugs", The Journal of otolaryngology, 19(2), 91-95.

Alvord, L. S., & Farmer, B. L. (**1997**). "Anatomy and orientation of the human external ear", Journal of the American Academy of Audiology, 8(6).

ANSI/ASA S12.42. (**2010**). "Methods for the Measurement of Insertion Loss of Hearing Protection Devices in Continuous or Impulsive Noise Using Microphone-in-Real-Ear or Acoustic Test Fixture Procedures".

ANSI/ASA S12.71. (**2018**). "Performance Criteria for Systems that Estimate the Attenuation of Passive Hearing Protectors for Individual Users".

Benacchio, S., Poissenot-Arrigoni, B., Martin, L., Saint-Gaudens, H., Sgard, F., and Doutres, O. (**2019**) "An artificial ear to assess objective indicators related to the acoustical comfort dimension of earplugs: comparison with attenuation and occlusion effect measured on subjects", In the proceedings of the 26th International Congress on Sound and Vibration, (ICSV26) (Montreal, QC, Canada, July 07-11, 2019) Canadian Acoustical Association.

Berger, E. H and Royster L. H. (**1996**). "In Search of Meaningful Measures of Hearing Protector Effectiveness", Spectrum Suppl. 1, 13, p.29.

Berger, E. H. (**1980**). "Hearing Protector Performance: How They Work - and - What Goes Wrong in the Real World", E.A.RLog, vol. 5.

Berger, E. H. (**1986**). "Methods of measuring the attenuation of hearing protection devices", The Journal of the Acoustical Society of America, 79(6), 1655-1687.

Berger, E. H. (**2005**). “Preferred methods for measuring hearing protector attenuation”, in Proceedings of Inter-Noise, vol. 5, p. 58).

Berger, E. H. (**2010**). “What is a personal attenuation rating (PAR)?”, EAR 07-21/HP Indianapolis (IN), 1-6.

Berger, E. H., Brown, J. D. and Smith, P. (**2014**). “3M™ E-A-Rfit™ Validation System Frequently Asked Questions (FAQs)”, E-A-R 13-17/HP.

Berger, E.H and Voix, J. (**2022**). “Chapter 11 : Hearing protection devices”, In The Noise Manual. 6th Edition. p. 255-308. American Industrial Hygiene Association.

CSA Group. (**2014**). Z94. 2-14. “Hearing Protection Devices-Performances, Selection, Care and Use”, Tech. rep.

Doutres O, Sgard F, Benacchio S, Terroir J, Perrin N, Trompette N, Negrini A, Gaudreau M.A, Jolly C, Berry A, Gauthier P.A, Padois T, Gauvin C. (**2018**). “Earplug comfort: From subjective assessment on the field to objective measurement and simulation using augmented artificial heads”, The Journal of the Acoustical Society of America, vol. 143, Minneapolis, MN, USA,: Acoustical Society of America; 2018, p. 1910–1910. https://doi.org/10.1121/1.5036223

Doutres O, Terroir J, Jolly C, Gauvin C, Martin L, Negrini A. (**2022**). “Towards a Holistic Model Explaining Hearing Protection Device Use among Workers”, IJERPH 2022;19:5578. https://doi.org/10.3390/ijerph19095578

Doutres, O., Sgard, F., Terroir, J., Perrin, N., Jolly, C., Gauvin, C., and Negrini, A. (**2020**). “A critical review of the literature on comfort of hearing protection devices: Analysis of the comfort measurement variability”, International Journal of Occupational Safety and Ergonomics, p. 1-36doi: 10.1080/10803548.2020.1772546

Fan, H., Yu, S., Wang, M., Li, M., Chu, J., Yan, Y., & Harris-Adamson, C. (**2021**). “Analysis of the external acoustic meatus for ergonomic design: part I–measurement of the external acoustic meatus using casting, scanning and rapid estimation approaches", Ergonomics, 64(5), 640-656.

Federman, J. and Duhon, C. (**2016**). "The viability of hearing protection device fit-testing at navy and marine corps accession points", Noise Health, Nov-Dec;18(85):303-311.

Ferguson, T., Greene, M., Repetti, F., Lewis, K., and Behdad, S. (**2015**). "Combining anthropometric data and consumer review content to inform design for human variability", In Proceedings of the ASME 2015 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference. Boston, Massachusetts, USA. p. V02BT03A022. doi: 10.1115/DETC2015-47640

Franks JR, Stephenson MR, Merry CJ. (**1996**), "Preventing occupational hearing loss - a practical guide", U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health.

Hayes ME, Hammond S, Montgomery AP, Stephenson L. (**2020**). "Improving Hearing Protection Device Noise Attenuation Through Fit-Testing in an Occupational Health Clinic", Workplace Health Saf 2022:21650799211067930. https://doi.org/10.1177/21650799211067927.

Jain, A.K., Murty, M. N. and P.J. Fynn. (**2000**). "Data clustering: A review", ACM Computing Surveys, 31(3), p. 264-323.

Lee, W., Yang, X., Jung, H., Bok, I., Kim, C., Kwon O., and You, H. (**2018**). "Anthropometric analysis of 3D ear scans of Koreans and Caucasians for ear product design", Ergonomics, 61:11, 1480-1495.

Liu, B. S. (**2008**). "Incorporating anthropometry into design of ear-related products", Applied Ergonomics, 39(1), 115-121.

Martin L., Negrini, A., Gaudreau, M-A., Sgard, F., Berbiche, D. and Doutres, O. (**2019**). "Earplug personal attenuation rating (PAR) in noise-exposed workers: evolution over a five weeks follow-up", In Proceedings of the 26th International Congress on Sound and Vibration (ICSV26) (Montreal, QC, Canada, July 07-11, 2019) Canadian Acoustical Association.

Mououdi, M. A., Akbari, J., & Mohammadi Khoshoei, M. (**2018**). "Measuring the external ear for hearing protection device design", Ergonomics in Design, 26(3), 4-8.

Na, S., Xumin, L. and Yong, G. (**2010**). "Research on k-means clustering algorithm: An improved k-means clustering algorithm", 3rd IEEE International Symposium on Intelligent Information Technology and Security Informatics (IITSI), pp. 63-67.

Negrini, A., Perron, J., Perron, B. (**2020**). "Analyse de classification par regroupement. Description et application à une problématique d'équilibre travail-famille et de détresse psychologique", (*Cluster analysis. Description and application to a work-family balance and psychological distress issue*). In, M. Corbière et N. Larivière. "Méthodes qualitatives, quantitatives et mixtes dans la recherche en sciences humaines, sociales et de la santé" (*Qualitative, quantitative and mixed methods in human, social and health sciences research*). p. 421-446. Québec, QC: Presses de l'Université du Québec.

Rabinowitz P, Galusha D, Cantley LF, Dixon-Ernst C, Neitzel R. (**2021**). "Feasibility of a daily noise monitoring intervention for prevention of noise-induced hearing loss", Occup Environ Med. https://doi.org/10.1136/oemed-2020-107351.

Sgard, F., Nélisse, H., Gaudreau, M-A., Boutin, J., Voix, J., and Laville, F. (**2011**). " Étude de la transmission sonore à travers les protecteurs auditifs et application d'une méthode pour évaluer leur efficacité en milieu de travail Partie 2 – Étude préliminaire d'une modélisation par éléments finis", *(Study of the sound transmission through hearing protectors and application of a method to assess their effectiveness in the workplace Part 2 – Preliminary study by finite element modeling),* IRSST - Rapport R-680. pp. 1–100.

Smith, C. R., Borton, T. E., Patterson, L. B., Mozo, B. T., and Camp, R. T., Jr. (**1980**). "Insert Hearing Protector Effects", Ear Hear. 1, 26-32.

Stinson, M. R., & Lawton, B. W. (**1989**). "Specification of the geometry of the human ear canal for the prediction of sound-pressure level distribution", The Journal of the Acoustical Society of America, 85(6), 2492-2503.

Takada MM, Rocha CH, Neves-Lobo IF, Moreira RR, Samelli AG. (**2020**). "Training in the proper use of earplugs: An objective evaluation", Work, 65(2), 401-407.

Viallet, G., Sgard, F., Laville, F., & Nelisse, H. (**2015**). "Investigation of the variability in earplugs sound attenuation measurements using a finite element model", Applied Acoustics, 89, 333-344.