## Appendix A:

Main steps:

1. TEOAE waveforms collection
2. RQA on TEAOAE waveforms to found 3 main measures: Rec, Det and Ent (see A.1)
3. RQA measures (REC, DET, ENT) are projected into PC1/PC2 references plane* by the factor score coefficients (see A.2)
4. RAD2D is reckoned as the square root of the sum of $P C 1{ }^{2}$ and $P C 2^{2}$ (see A.3)
5. WWR vs RAD2D plot is built with two thresholds lines: WWR $=70 \%$ (horizontal line) and RAD2D $=1.78$ (vertical line)
6. Right top square evidenciates true negative ears
${ }^{*}$ ) the reference plane and factor scores are defined by a representative data set from 118 signals, different from those studied in the present manuscript; consequently each ear-point is described by PC1 and PC2

## WWR vs RAD2D plot

| A <br> high WWR and low RAD2D <br> (true negative) | B <br> high WWR and high RAD2D <br> (false negative) |
| :---: | :---: |
| C <br> low WWR and low RAD2D <br> (false positive) | D <br> low WWR and high RAD2D <br> (true positive) |

## Appendix A1: Recurrence Quantification Analysis

To apply Recurrence Quantification Analysis (RQA) to a signal, its time behavior is represented by a set of $n$ points equally spaced in time (e.g. $\left\{a_{1} ; a_{2} ; \ldots . a_{n}\right\}$, where $a_{i}$ is the signal value at time $i$ ). Then, the series is represented in column format and copied into successive columns (the number of columns is defined as the "embedding dimension", N), each one shifted by a given number of points (lag). This procedure creates the so-called embedding matrix.

Finally, if the distance between the jth and the ith rows of the embedding matrix is less than a fixed value (called radius), a dot is drawn in the position that represents the distances between the corresponding rows (epochs of the series) thus producing a graph pattern (recurrence plot). ${ }^{25-26}$ In the recurrence plot, the horizontal and vertical axes represent the relative positions of the $n$ points of the time series.

The RQA descriptors are calculated on the basis of the number and location of dots (named recurrence points) on the graph. In particular, percent of recurrence (REC) is the percentage of recurrence points in a recurrent plot; percent of determinism (DET) is the percentage of recurrence points which form diagonal lines and it indicates the degree of deterministic structure of the signal; entropy (ENT) is the Shannon entropy of the probability distribution of the diagonal line lengths and it is linked to the richness of deterministic structure; laminarity (LAM) is the percentage of recurrent points forming vertical line segments in the recurrence plot. The presence of horizontal and vertical lines in the recurrence plot shows that part of the considered signal matches closely with a sequence farther along the time. To build the recurrence plots, RQA variables were set according to the experience gained in previous publications: ${ }^{22-26}$ the delay in the embedding procedure (lag) was set to 1 ; the number of the embedding matrix columns (embedding dimension, N ) was 10 ; the cut-off distance (radius) was set to 15 , line $=8 .{ }^{22-23}$

## Appendix A2: Principal Component Analysis

The descriptors obtained by the RQA were re-dimensioned by applying the Principal Component Analysis (PCA) technique. PCA is a common statistical technique ${ }^{37}$ which provides the possibility to (i) reduce the dimension of a data set without consistent loss of information; and (ii) to separate the different and independent features of the data. The PCA procedure describes the original data set with a lower number of parameters called main components (PC1, PC2), which explain more than $90 \%$ of the total variability in the data set. Details on this procedure can be found in previous publications. ${ }^{22,23}$

Having PC1 and PC2, by construction, zero mean and standard deviation equal to 1 , if a set of TEOAE signals from normal ears are studied, $96 \%$ of the them will fall within a circle centered in the origin of the PC1/PC2 plane with radius equal to 2 (reference circle). In the present study, the PC1/PC2 plane was defined starting from a representative data set made by 118 signals measured from normal hearing subjects, in analogy with previous papers. ${ }^{22,23}$ The explained variability of the representative data set used was $90.02 \%, 6.29 \%$, and $3.68 \%$ for PC1, PC2, and PC3, respectively. The correlation between RQA parameters and PCs is described by the "factor score coefficients". The signals measured from the subjects in the NH and ARHL groups considered in this study have then been projected in the same PC1/PC2 plane. ${ }^{22}$

## Appendix A3: RAD2D's mathematic formula

RAD2D is defined in the PC1/PC2 plane as the Euclidean distance of one point representing a TEOAE signal from the origin of the plane, i.e.:

$$
R A D 2 D=\sqrt{P C 1^{2}+P C 2^{2}}
$$

the factor scores carried on by factor loadings reckoned from 118 normal references signals (representative data set) permit to projected new TEOAE signals in the PC1/PC2 reference plane. ${ }^{22}$


