

# Speech Perception Across The Lifespan by Means of Artificial Intelligence

Ulrich Hoppe (✉ [ulrich.hoppe@uk-erlangen.de](mailto:ulrich.hoppe@uk-erlangen.de))

University of Erlangen-Nuremberg

Thomas Hocke

Cochlear Deutschland GmbH & Co KG

Heinrich Iro

University of Erlangen-Nuremberg

---

## Research Article

**Keywords:** Hearing loss, clinical, audiometry, presbycusis

**Posted Date:** February 1st, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-148304/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

Hearing loss is one of the most common diseases worldwide. It affects communicative abilities in all age groups. However, it is well known that elderly people suffer more frequently from hearing loss. The aim of this study was to describe the relationship between hearing loss, age, and speech perception. The model that we employed used Random Forest Regression. It was applied to a large clinical data set of 19,801 ears, covering all degrees of hearing loss. It allows the estimation of age-related decline in speech perception in quiet, with the effect of pure-tone hearing loss completely separated. Our results show that speech scores depend on the specific type of hearing loss and age.

We found age effects for all degrees of hearing loss. A deterioration in speech perception of up to 25 percentage points across the whole life span was observed. The largest decrease was 10 percentage points per life decade. The decline can be attributed to a distortion component of presbycusis, which is not measured by pure-tone audiometry.

## Introduction

More than 5% of the world's population, approximately 360 million people, suffer from disabling hearing loss. Hearing disability is associated with reduced speech perception and, in consequence, reduced communication function. Hearing deteriorates with age<sup>1</sup>. In particular, auditory sensitivity and speech perception decline after age 60. Given our ageing society and the prevalence of age-related hearing loss, it is clear that hearing loss is a common public-health issue affecting almost all older adults. This age-related hearing loss (ARHL) has no adequate causal therapy, and technical solutions such as hearing aids, cochlear implants and assistive devices improve hearing thresholds but do not restore speech recognition sufficiently.

Speech perception deficits in persons with hearing impairment are mainly attributable to decreased audibility of the speech signal over part or all of the speech frequency range. A second contribution stems from impaired processing of the audible speech signal, which results in reduced clarity<sup>2</sup>. Similarly, Plomp<sup>3</sup> subdivided the components of speech recognition deficits in sensorineural hearing loss into Class A (attenuation) and Class D (distortion). The Class A component relates to Carhart's acuity deficit<sup>2</sup> and can be compensated for by frequency-dependent gain aiming at improving audibility. The Class D component relates to Carhart's clarity deficit<sup>2</sup> and represents the distortion that is produced by damage to the auditory periphery and the mechanisms of neural processing. Evidently, this distortion cannot be compensated for by acoustic amplification. Plomp<sup>3</sup> concluded that the Class D component may limit the benefit of hearing aids.

Both attenuation and distortion are part of ARHL. A large number of studies have focussed upon hearing in the elderly and have investigated pure-tone loss, speech perception, and other suprathreshold abilities such as spatial processing and recognition of distorted speech. In many studies<sup>1,4-17</sup> speech recognition in older adults has been investigated. The interpretation of these results remains challenging, as pure-

tone thresholds change substantially with increasing age and pure-tone loss is the primary predictor for speech recognition. Hence, it is necessary to correct for the effect of pure-tone loss when investigating the effect of age on speech perception. One of the first attempts to do this was described by Jerger<sup>4</sup> in a report on speech recognition in a large group of older subjects. He analysed scores from the clinical records of 2162 patients. With subjects grouped according to age and average hearing loss at 0.5, 1 and 2 kHz, results suggested that speech recognition, defined as the maximum score obtained using a monosyllabic word list ( $WRS_{max}$ ), declines above the age of sixty. In particular, he found that age had an effect on speech recognition for individuals with mild hearing loss of approximately 4% per life decade, but it had a greater effect upon those with higher degrees of hearing loss, e.g. 10% per decade. Unfortunately, he did not report on hearing loss at higher frequencies. There is a long-standing consensus that hearing thresholds at these higher frequencies are, in particular, worse for older subjects<sup>1, 5</sup>. Consequently, there is a systematic overestimation of the age effect<sup>12</sup>.

Several other studies have revealed evidence that deterioration in speech understanding occurs in addition to deterioration in hearing sensitivity and includes components beyond elevated hearing thresholds<sup>6-12</sup>. However, the authors<sup>11, 12</sup> highlighted the challenge of separating varying auditory thresholds from age, a factor affecting all sensory modalities<sup>13</sup>. In recent studies<sup>14-16</sup>, speech perception and its relation to age were investigated either by correcting for pure-tone loss<sup>15, 16</sup> or by using a longitudinal study design<sup>14</sup>. In a clinical population Hoppe *et al.*<sup>15</sup> investigated speech perception with hearing aids and  $WRS_{max}$  for different age groups in relation to average pure-tone loss at 0.5, 1, 2 and 4 kHz (4FPTA). They found a monotonic decrease for speech perception with increasing age and a significant drop of about 2–4% per decade. Müller *et al.*<sup>16</sup> investigated the  $WRS_{max}$  as a function of age. After correcting for pure-tone loss they found a significant, though smaller, drop for people aged above 70 years of about 2–3% per decade. Neither study included hearing threshold beyond 4 kHz, and therefore a small overestimation of the influence of age cannot be excluded. However, Dubno *et al.*<sup>14</sup> found a larger effect, around 7–8% per life decade. They performed a longitudinal study including 256 subjects with age-related hearing loss, aged 50–82 years, over a period of 3–15 years. The speech recognition scores were corrected for by changing hearing thresholds during the observation phase; this was done by using the individuals' articulation index as an importance-weighted metric for speech audibility. Unfortunately, longitudinal studies suffer from other disadvantages relating to population size, loss of follow-up etc., and their duration can approach the limits of the clinician's working life-span. The special characteristics of the study population and methods – neither the  $WRS_{max}$  nor hearing-aid scores were measured – differ from the studies mentioned above. This impedes a direct comparison with the above-mentioned studies and therefore does not imply a contradiction amongst them<sup>14-16</sup>. In summary, there is evidence that speech recognition in older subjects declines with increasing age more rapidly than their pure-tone thresholds would suggest. The effect seems to be small and depends on age and degree of hearing loss.

The goal of this study was the development of a tool to describe the relation between hearing loss, age, and speech perception. It uses a machine-learning algorithm, Random Forest Regression (RFR)<sup>18</sup>. This

algorithm was applied to a large data set from routine clinical audiometry in order to investigate the influence of age. The result is a representation of the relation between pure-tone thresholds and age, with clear distinction of both factors on the input side and speech perception on the target side. The tool allows description of the influence of age, with hearing thresholds kept fixed. RFR is an algorithm that uses an ensemble method of decision-tree-based regressions to determine a response from a set of predictor variables. It does not rely on any particular assumptions regarding data distribution. Hence, it is ideally suited for modelling a large clinical audiometric database.

## Results

The stacked bar plot (Fig. 1) shows the case distribution in our clinical population (N = 19,801). The mean ages of the different groups were 50, 61, 66, 65 and 59 years for WHO<sub>0</sub>, WHO<sub>1</sub>, WHO<sub>2</sub>, WHO<sub>3</sub> and WHO<sub>4</sub>. The vast majority (77%) of cases involved persons between 40 and 80 years of age. The subjects aged 40 to 80 years dominated all WHO grades except WHO<sub>0</sub>. The smallest data coverage with respect to age and hearing loss was observed for very young adults in the WHO<sub>4</sub> group and for subjects above 80 years of age in the WHO<sub>0</sub> group.

The speech audiometric results for the model's target scores, WRS<sub>65</sub> and WRS<sub>max</sub>, are shown in Fig. 2A and 2B, respectively. For both measures the median decreased monotonically with increasing degrees of hearing loss. The variability for WRS<sub>65</sub> was largest for WHO<sub>1</sub>, while for WRS<sub>max</sub> the largest variability was found for WHO<sub>3</sub>. In this rather rough classification the interpretation of some outliers may benefit from additional information about the actual cause of hearing loss. In particular, the WHO classification employs the hearing thresholds at only four frequencies, while other frequencies are not considered. The lowest quartile of the WHO<sub>0</sub> cases shows a WRS<sub>65</sub> lower than 95%. In this subgroup the mean threshold for high frequencies (> 4 kHz) was 48 dB, while the overall mean threshold for high frequencies was 30 dB in the WHO<sub>0</sub> group. Among the 1075 WHO<sub>3</sub> cases, 41 exhibited a WRS<sub>65</sub> larger than 0%. Those cases had a mean threshold for low frequencies (< 0.5 kHz) of 32 dB, while the overall mean threshold for high frequencies was 54 dB.

## Model setup

The RFR was trained with 80% of the data (training group). The model was then tested in the remaining 20% of the study population (evaluation group). Before assignment to groups was performed, the data sets were sorted according to 4FPTA and age. Subsequently, every fifth data set was assigned to the evaluation group. The pure-tone thresholds at all frequencies and the patients' age were input (predicting) variables, while the WRS<sub>65</sub>, WRS<sub>max</sub> and L<sub>max</sub> were targets. For each of the three output variables a separate RFR-model was built. The most important parameter with respect to overfitting is the number of decision trees within a random forest. As a parameter for estimating the RFR performance the median absolute error (MAE, resulting from measured minus predicted score) was used for both the training group and the evaluation group. The MAE in the evaluation group was larger for lower numbers (< 10) of

decision trees, while remaining at a certain level for a large range (50–1000) of numbers of trees. Therefore, the Matlab default of 100 was retained. Other parameters for setup of a single decision tree were found to be of less importance for the RFR performance and were therefore kept at default values. The number of nodes per binary decision tree varied for each model: around 2150 for  $WRS_{max}$ , around 2700 for  $WRS_{65}$ , and around 3650 for  $L_{max}$ .

Table 1 summarises the performance of the model as assessed by MAE for both the training and the evaluation group. Owing to the composition of our study population the  $WHO_0$  is by far the largest group. The MAE of this group would have dominated the overall summary. For this reason, Table 1 summarises the error estimation for each grade of hearing loss separately. Evidently, there is a great variation of the MAE among the WHO groups and between the target variables. The largest prediction errors were observed in  $WHO_2$  for the  $WRS_{65}$  group and in  $WHO_3$  and  $WHO_4$  for  $WRS_{max}$ . For those WHO groups the MAE of the training and evaluation groups differed by a factor of 1.5–1.7.

Table 1

Median absolute error of the machine-learning prediction model for the monosyllabic score at a presentation level of 65 dB<sub>SPL</sub> ( $WRS_{65}$ ), the maximum word recognition score ( $WRS_{max}$ ) and the presentation level for the maximum word recognition score,  $L_{max}$ .

Target	Quality measure	Subgroup	$WHO_0$	$WHO_1$	$WHO_2$	$WHO_3$	$WHO_4$
$WRS_{65}$	MAE	Training	0.9	5.6	8.2	0.1	0.0
	[percentage points]	Evaluation	1.5	8.8	11.9	0.1	0.0
$WRS_{max}$		Training	0.0	1.6	5.6	11.0	7.0
		Evaluation	0.0	2.4	9.4	16.8	12.1
$L_{max}$	MAE	Training	3.4	3.2	3.4	3.1	2.9
	[dB]	Evaluation	5.2	5.3	5.6	5.0	5.0

## Application of the prediction model

One possible application of the prediction model is shown in Fig. 3. For the Bisgaard audiogram types<sup>19</sup> the model input was a pure-tone audiogram according to the flat (N-type) or steep (S-type) audiograms. The pure-tone thresholds were kept constant and the subjects' age varied between 18 and 99 years. Owing to the relation between age and hearing thresholds hardly any subjects were in our population aged > 85 for groups  $N_1$  and  $S_1$ . Consequently, the model could not be trained in this range, so it was excluded from calculations. Linear regression models for all curves were fitted and then tested against a constant model including a Bonferroni–Holm correction. The determined slopes of the linear regression analysis are shown in Fig. 3D, E, and F, for  $WRS_{65}$ ,  $WRS_{max}$ , and  $L_{max}$ , respectively.

Figure 3A shows that the model results in  $WRS_{65}$ , which decreases, mostly monotonically, with age. However, the model suggests that age effects depend on specific age range, and also on the degree and

kind of hearing loss as indicated by the Bisgaard types<sup>19</sup>. For types  $N_3$ ,  $S_1$ , and  $S_3$  the model output indicates a decrease of  $WRS_{65}$  with age. The model results become more complex if the  $WRS_{max}$  and  $L_{max}$  are considered as shown in Fig. 3B and C, respectively. The presentation level shows, for all types except  $N_6$ , an increased presentation level for  $WRS_{max}$  with increasing age. A considerable decrease in score can be observed in  $N_6$ , accompanied by a slight but significant decrease of  $L_{max}$ . For the  $N_4$  and  $S_3$  types the model gives a significant decrease in  $WRS_{max}$  which is somehow weakened by an increased presentation  $L_{max}$  for this type. For all other types the  $WRS_{max}$  did not change significantly with age. However, for these types the model results in an increased presentation level.

## Discussion

The machine-learning approach allows description of the age-related decline of speech perception. In comparison with previous studies, more detailed information about the time course and amount of degradation was achieved. The RFR model allows the quantitative description of the two basic impacts of hearing loss and its relation to age: On the one hand the impact of the attenuation component, and on the other the impact of the age-related increase of the distortion components. The RFR model is able to take into account frequency-specific differences and to separate pure-tone threshold completely from age. It therefore offers the opportunity to overcome an immanent bias in previous investigations<sup>4, 8, 11, 15, 16</sup> by isolating age-related threshold shifts from age-related decline in speech perception as such.

This study should not be misunderstood as an attempt to predict speech perception scores on the basis of pure-tone loss. These scores have in any case to be measured individually. Already the large variability of individual scores necessitates speech audiometry. The purpose of the model in this study is the prediction for larger patient populations with respect to specific audiogram types and age. It can be seen in Fig. 3 that those age-related changes are present for the entire duration of adulthood. However, apart from the fact that higher age relates to lower speech perception scores, no common quantitative trend, for all age groups and pure-tone hearing losses, can be discerned. This may be regarded as the major outcome of the model calculations. The measurable age-related decline in speech perception depends on the age range considered, the specific audiogram, and the specific application of speech audiometry. Owing to saturation effects of the  $WRS_{65}$  measured at typical conversation level, we observe the largest age effect for moderate hearing losses ( $N_3$ -type audiograms). For the  $WRS_{max}$  measured at substantially higher levels, the largest effects were observed for audiogram types corresponding to severe hearing losses ( $N_4$ ,  $N_5$ ,  $N_6$ ). However, a potential decrease of the  $WRS_{max}$  can be substantially counterbalanced by an increased presentation level for some audiogram types. This finding is not consistent across all hearing losses. For the  $N_6$  group, with the largest age-related decline, we also see a decreased tolerance for higher presentation levels. This might reflect certain underlying pathomechanisms which are more likely to result in this audiogram type compared with others. Complementary to the attenuation and distortion component of hearing loss<sup>2, 3</sup> a causal and more differentiated breakdown with respect to presbycusis was proposed early on<sup>20, 21</sup>. Initially, five main types were proposed, namely sensory, neural,

metabolic, mechanical, and vascular presbycusis<sup>20</sup>. This was complemented by the term central presbycusis in order to reserve the term neural for degeneration of the cochlear nerve<sup>21</sup>. Sensory presbycusis is congruent with the attenuation component, and is, as pointed out above, completely separated from the relationships in Fig. 3. The effects of all the other types of presbycusis are not. Moreover, the specific and different root causes may potentially explain why, for some degrees of hearing loss, different changes in speech perception occur in different life decades. However, possible interactions between the main types of presbycusis are still not completely understood<sup>22</sup>.

It is not possible to confirm all these explanatory hypotheses by retrospective data analysis, a fact that clearly underlines the limits of our study design. We found differences in age effects in comparison with some of the studies referred to above. This is partly due to the neglect of hearing loss at higher frequencies for the elderly in those studies. On the other hand, for some hearing losses and audiogram types, this study may underestimate age effects, as ceiling effects of speech tests in quiet are included. Another unique feature of this study is the inclusion of a considerable number of subjects with mild hearing loss, as seen in group  $S_1$ . Even in that group, age effects play a part. Especially the  $WRS_{65}$  illustrates how everyday communicative ability in quiet might be already affected by mild to moderate hearing loss in a population in which the use of hearing aids does not reach the penetration level needed<sup>23</sup>.

Other possible applications of the model are related to acoustic amplification with hearing aids: As shown in Fig. 3, in all groups but  $N_6$  the level for best speech perception  $L_{max}$  increases with age at about 0.5 dB per decade. This may indicate that older people may benefit from higher levels for speech perception, i.e. greater amplification, when provided with a hearing aid. As far as we know, current amplification strategies do not take account of age.

The age dependence of the  $WRS_{max}$  found in our study may be used to improve studies evaluating the outcome of hearing-aid use: The  $WRS_{max}$  or an equivalent measure is often used as reference for the measurement of successful hearing-aid provision or other acoustic amplification<sup>15, 23-26</sup>, for investigation of age-related changes in cognition<sup>17</sup>, and for speech-perception-related studies in general<sup>27</sup>. A consideration of both age and specific audiogram type could potentially decrease the variability of results. Other studies<sup>28-31</sup> have focussed on the  $WRS_{max}$  as one criterion within the process of evaluation of candidates for cochlear implant or for predicting the outcome of such an implantation<sup>32, 33</sup>. A model may provide more insights into hearing-loss-related pathologies amongst hearing aid users, especially in the cases in which the scores with a hearing aid are far below the  $WRS_{max}$  or if the  $WRS_{max}$  is unexpectedly low with respect to clinical experience or model prediction. Furthermore, epidemiological studies<sup>34-38</sup> on hearing loss can be complemented, providing indirect functional assessment.

## Limitations of the study

An important limitation of this study is the restriction to a specific language and test. With respect to other languages and speech material the comparison of recent studies<sup>32, 39</sup> suggests that the test we used is comparable to the English Consonant-Vowel-Nucleus-Consonant test.

Secondly, the outdated but established calibration<sup>40</sup> procedure for the Freiburg monosyllable test at 65 dB<sub>SPL</sub> is roughly comparable to a level of 60 dB<sub>A</sub>. Consequently,  $L_{\max}$  should be corrected by about 5 dB<sup>40</sup> for a comparison e.g. with CNC results.

The disadvantage of binary decision trees is the high chance of overfitting. The use of a random-forest method does decrease this risk. However, a factor of up to 1.7 between the MAEs in the evaluation group in relation to the training group still indicates some degree of overfitting. Even the considerable size of the study population and the clustering of input variables do not entirely prevent this risk. Additionally, there are some intrinsic sources of unexplained variability. Even though we excluded unreasonable cases, the population may still have included mild cases of aggravation, simulation or dissimulation. There was also a small number of cases with retrocochlear lesions. This number can be estimated as less than 0.5% in our population by comparison with our patient files and the reported incidence<sup>41</sup>.

An RFR model inevitably reflects the characteristics of the clinical population that contributed to the training. The group characteristics might be different from their peers outside a clinic. Finally, the model reflects statistical characteristics of a population, and not causal relationships.

## Conclusion

A random-forest regression model allowed the estimation of age-related decline of speech perception in quiet, completely separated from the effect of pure-tone hearing loss. Small but considerable declines were found across the whole duration of adulthood and for all audiogram types. The decline can be attributed to a distortion component of presbycusis which is not measurable by pure-tone audiometry. The careful derivation of working hypotheses from our data has the potential to provide more insights into the relationships between pure-tone hearing loss, specific audiogram types and age.

## Methods

Audiometric data were retrieved from a clinical data base at the Audiological Department of Erlangen University Hospital. From the routine audiometric measurements pure-tone-thresholds for both bone and air conduction and speech perception for monosyllabic words (Freiburg Test) were extracted. All measurements were conducted in clinical routine in sound-shielded booths with clinical class A audiometers (AT900 / AT1000 AURITEC Medizindiagnostische Systeme GmbH, Hamburg, Germany). Approval for this study was received from the Institutional Review Board of the University of Erlangen.

All methods were carried out in accordance with relevant guidelines and regulations. In particular, informed consent was obtained from all patients.

Among 91,991 patients who underwent audiometry at our centre from April 2002 to June 2020 we identified 53,782 adults aged at least 18 years at the time of first investigation. Initially, the data were screened for repeated measurements. Only the first audiometric assessment was retained. Subsequently, the data from 107,564 ears (hereinafter “cases”) were checked for a complete set of air and bone conduction thresholds. After removal of incomplete data sets there remained 107,010 cases. In the next step, cases with missing or incomplete speech audiometry data were deleted, whereafter 26,324 cases remained. The data were then screened for cases of mixed hearing loss; the latter was defined as a difference between air and bone conduction thresholds greater than 10 dB for frequencies within the range 0.5–3 kHz. After removal of mixed-hearing-loss cases, the remaining 19,929 cases were checked for inconsistent results (< 1%), caused e.g. by simulation or lack of collaboration by the patient. If, within the discrimination function for monosyllabic words, a score larger than zero was observed while the presentation level was below the hearing threshold, the data set for that case was not used. For some cases it was observed that the measurement of the discrimination function had not been fully completed, so that a score of 100% was not reached, with the presentation level well (> 15 dB) below the discomfort level. The 19,801 cases finally remaining were used for model-building and for error analysis.

The following data were used for analysis:

1. Air-conduction hearing thresholds at 0.125, 0.25, 0.5, 0.75, 1, 1.5, 2, 3, 4, 6 and 8 kHz,
2. Word recognition score at 65 dB<sub>SPL</sub> ( $WRS_{65}$ ),
3. Maximum word recognition score ( $WRS_{max}$ ) and corresponding level ( $L_{max}$ ).

$WRS_{65}$  describes speech recognition at a typical conversational level. While  $WRS_{65}$  is primarily dependent on the attenuation and reflects the loss of speech perception ability in everyday life,  $WRS_{max}$  describes the maximum information that can be processed to the auditory system. The difference  $WRS_{max} - WRS_{65}$  can be used to estimate the acceptance of acoustic amplification<sup>23</sup>.

In order to summarise audiometric constellation of our study population and to describe the error of the model, we used the WHO classification. The average of hearing thresholds, measured at 0.5, 1, 2 and 4 kHz, refers to the 4FPTA. This average was used to classify according to the WHO categories: WHO<sub>1</sub> (26 dB < 4FPTA ≤ 40 dB), WHO<sub>2</sub> (40 dB < 4FPTA ≤ 60 dB), WHO<sub>3</sub> (60 dB < 4FPTA ≤ 80 dB) or WHO<sub>4</sub> (80 dB < 4FPTA).

As an example of application of the model we used a reference audiogram data set described by Bisgaard *et al.*<sup>19</sup>. Those authors identified ten standard audiograms with different slopes for a large clinical set. While seven of them have a moderate sloping audiogram ( $N_1 - N_7$ ) three have a steeper slope ( $S_1 - S_3$ ) with a threshold difference of at least 60 dB.

## Data analysis

For data analysis, model calculation, statistics and figures, the software Matlab R2019B (The Mathworks Inc. Natick, Massachusetts) was used. Data were rounded before the model calculation: hearing thresholds were rounded to 5 dB<sup>42</sup>, and for linear regression the patients' ages were classified into decades corresponding to the spacing of the RFR model input.

RFR was used to relate speech scores to hearing thresholds and age. To test the model results for significant age dependence, linear regression models were fitted to all curves. They were then tested against a constant model including a Bonferroni–Holm correction.

## Declarations

## Acknowledgements

We would like to thank Peter Woolley for his careful revision of language.

## Author contributions

Conceptualization: UH

Formal analysis: UH, TH

Investigation: UH, HI

Methodology: TH, UH

Project administration: UH, HI

Ressources: HI, UH

Software: TH, UH

Validation: UH, TH

Visualization: TH, UH

Writing original draft: UH, TH

Writing, review and editing: UH, TH, HI

## References

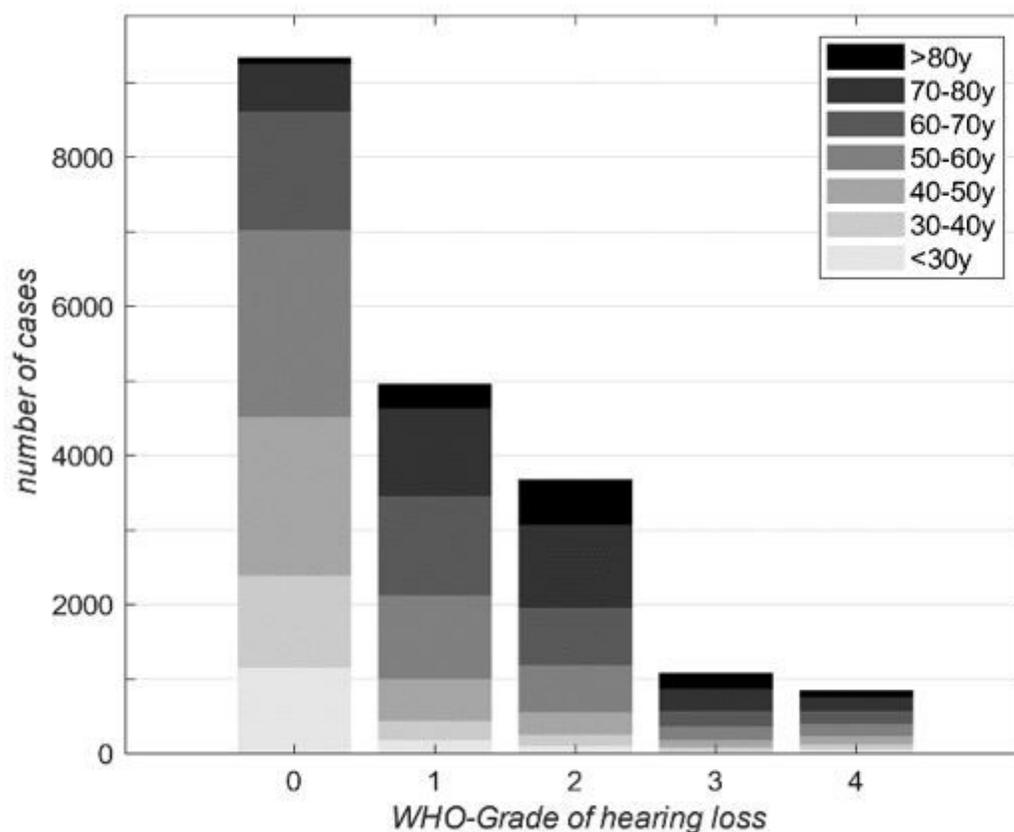
1. Zwaardemaker, H. Der Verlust an hohen Tönen mit zunehmendem Alter, ein neues Gesetz. *Arch Ohr. Nas.-KehlkHeilk.* **47** (1899).
2. Carhart, R. Basic Principles of Speech Audiometry. *Acta Otolaryngol.* **40**, 62–71 (1951).

3. Plomp, R. Auditory handicap of hearing impairment and the limited benefit of hearing aids. *J. Acoust. Soc. Am.* **63**, 533–549 (1978).
4. Jerger, J. Audiological findings in Aging. *Adv. Oto-Rhino-Laryng.* **20**, 115–124 (1973).
5. ISO DE. Akustik - Statistische Verteilung von Hörschwellen als Funktion des Alters. *ISO 7029:2000* (2001).
6. Jerger, J. & Hayes, D. Diagnostic Speech Audiometry. *Arch Otolaryngol.* **103**, 216–222 (1977).
7. Bergman, M., Blumenfeld Cascardo, D., Dash, B., Levitt, H. & Margulies, M. K. Age-Related Decrement in Hearing for Speech Sampling and Longitudinal Studies. *Journal of Gerontology.* **31**, 533–538 (1976).
8. Marshall, L. & Bacon, S. P. Prediction of speech discrimination scores from audiometric data. *Ear Hear.* **2**, 148–155 (1981).
9. Pedersen, K. E., Rosenhall, U. & Moioerb, M. B. Longitudinal Study of Changes in Speech Perception between 70 and 81 Years of Age. *Audiology.* **30**, 201–211 (1991).
10. Divenyi, P. L. & Haupt, K. M. Audiological correlates of Speech Understanding Deficits in Elderly Listeners with Mild-to-Moderate Hearing Loss. III Factor Representation. *Ear Hear.* **18**, 189–201 (1997).
11. Dubno, J. R., Lee, F. S., Matthews, L. J. & Mills, J. H. Age-Related and Gender-Related Changes in Monaural Speech Recognition. *JSHLR.* **40**, 444–452 (1997).
12. Humes, L. E. The Contributions of Audibility and Cognitive Factors to the Benefit Provided by Amplified Speech to Older Adults. *J Am Acad Audiol.* **18**, 590–603 (2007).
13. Humes, L. E. & Young, L. A. Sensory-Cognitive Interactions in Older Adults. *Ear Hear.* **37** (Suppl 1), 52–61 (2016).
14. Dubno, J. R. *et al.* Longitudinal changes in speech recognition in older persons. *J. Acoust. Soc. Am.* **123**, 462–475 (2008).
15. Hoppe, U., Hast, A. & Hocke, T. Sprachverstehen mit Hörgeräten in Abhängigkeit vom Tongehör. *HNO.* **62**, 443–448 (2014).
16. Müller, A., Hocke, T., Hoppe, U. & Mir-Salim, P. The age effect in evaluation of hearing aid benefits by speech audiometry. *HNO.* **64**, 143–148 (2016).
17. Kronlachner, M., Baumann, U., Stover, T. & Weissgerber, T. Investigation of the quality of hearing aid provision in seniors considering cognitive functions. *Laryngorhinootologie.* **97**, 852–859 (2018).
18. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
19. Bisgaard, N., Vlaming, M. S. & Dahlquist, M. Standard audiograms for the IEC 60118-15 measurement procedure. *Trends in amplification.* **14**, 113–120 (2010).
20. Schuknecht, H. F. Further observations on the pathology of presbycusis. *Arch. Otolaryngol.* **80**, 369–382 (1964).
21. Johnsson, L. G. & Hawkins, J. E. Jr. Sensory and neural degeneration with aging, as seen in microdissections of the human inner ear. *Ann. Otol. Rhinol. Laryngol.* **81**, 179–193 (1972).

22. Bao, J. & Ohlemiller, K. K. Age-related loss of spiral ganglion neurons. *Hear. Res.* **264**, 93–97 (2010).
23. Halpin, C. & Rauch, S. R. Improvement in Word Recognition Score with Level Is Associated with Hearing Aid Ownership among Patients with Hearing Loss. *Audiol. Neurotol.* **17**, 139–147 (2012).
24. Müller, A. *et al.* Influence of Floating-Mass Transducer Coupling Efficiency for Active Middle-Ear Implants on Speech Recognition. *Otol. Neurotol.* **38**, 809–814 (2017).
25. Maier, H., Lenarz, T., Dolleza, L. V. & Busch, S. Direct Acoustic Cochlear Implants Lead to an Improved Speech Perception Gap Compared to Conventional Hearing Aid. *Otol. Neurotol.* **39**, 1147–1152 (2018).
26. Maier, H. *et al.* Minimal Reporting Standards for Active Middle Ear Hearing Implants. *Audiol. Neurotol.* **23**, 105–115 (2018).
27. Müller, J., Plontke, S. K. & Rahne, T. Speech audiometric outcome parameters in clinical trials on hearing improvement. *HNO.* **65**, 211–218 (2017).
28. Hoppe, U., Hast, A. & Hocke, T. Audiometry-Based Screening Procedure for Cochlear Implant Candidacy. *Otol. Neurotol.* **36**, 1001–1005 (2015).
29. McRackan, T. R., Ahlstrom, J. B., Clinkscales, W. B., Meyer, T. A. & Dubno, J. R. Clinical Implications of Word Recognition Differences in Earphone and Aided Conditions. *Otol. Neurotol.* **37**, 1475–1481 (2016).
30. McRackan, T. R. *et al.* Earphone and Aided Word Recognition Differences in Cochlear Implant Candidates. *Otol. Neurotol.* **39**, e543–e549 (2018).
31. Franks, Z. G. & Jacob, A. The speech perception gap in cochlear implant patients. *Cochlear Implants Int.* **20**, 176–181 (2019).
32. Hoppe, U., Hocke, T., Hast, A. & Iro, H. Maximum preimplantation monosyllabic score as predictor of cochlear implant outcome. *HNO.* **67** (Suppl 2), 62–68 (2019).
33. Hoppe, U., Hocke, T., Hast, A. & Iro, H. Cochlear Implantation in Candidates With Moderate-to-Severe Hearing Loss and Poor Speech Perception. *Laryngoscope*. Online ahead of print(2020).
34. Sohn, W. & Jörgenshaus, W. Schwerhörigkeit in Deutschland. Repräsentative Hörscreening-Untersuchung bei 2000 Probanden in 11 Allgemeinpraxen. *Z. Allg. Med.* **77**, 143–147 (2001).
35. Chang, Y. S. *et al.* Standard Audiograms for Koreans Derived through Hierarchical Clustering Using Data from the Korean National Health and Nutrition Examination Survey 2009–2012. *Sci. Rep.* **9**, 3675 (2019).
36. Löhler, J. *et al.* The prevalence of hearing loss and use of hearing aids among adults in Germany: a systematic review. *Eur. Arch. Otorhinolaryngol.* **276**, 945–956 (2019).
37. von Gablenz, P., Hoffmann, E. & Holube, I. Prevalence of hearing loss in Northern and Southern Germany. *HNO.* **65**, 130–135 (2017).
38. von Gablenz, P., Hoffmann, E. & Holube, I. Gender-specific hearing loss in German adults aged 18 to 84 years compared to US-American and current European studies. *PLoS One.* **15**, e0231632 (2020).

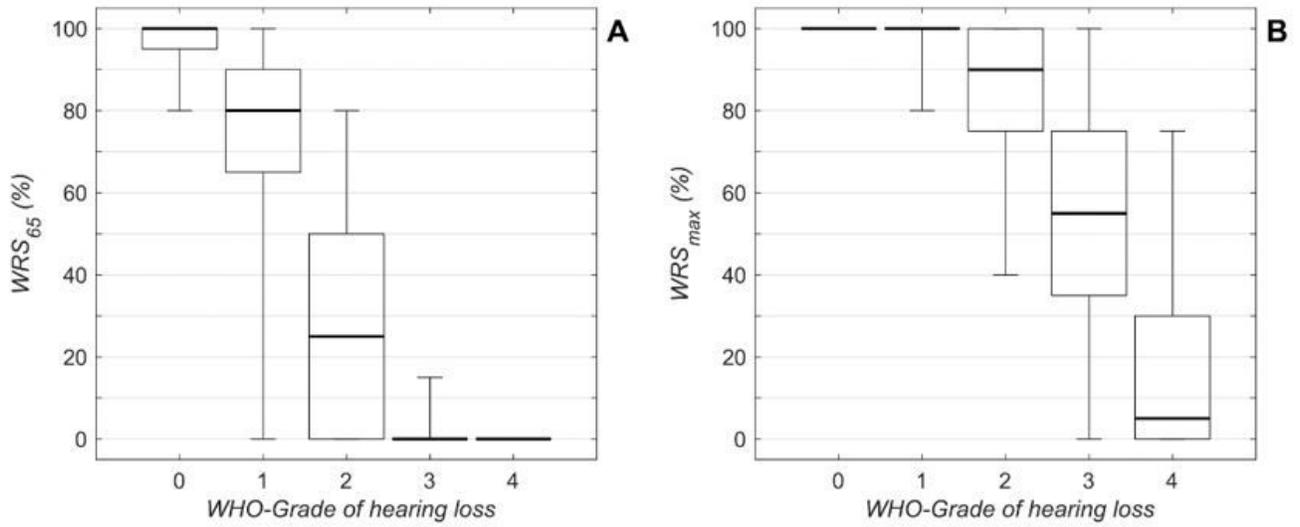
39. Holden, L. K. *et al.* Factors affecting open-set word recognition in adults with cochlear implants. *Ear Hear.* **34**, 342–360 (2013).
40. Holube, I., Steffens, T. & Winkler, A. Calibration of the Freiburg monosyllabic speech test. *HNO.* **67**, 304–305 (2019).
41. Lin, D., Hegarty, J. L., Fischbein, N. J. & Jackler, R. K. The prevalence of "incidental" acoustic neuroma. *Arch. Otolaryngol. Head Neck Surg.* **131**, 241–244 (2005).
42. von Wedel, H. Fehlermöglichkeiten in der Ton- und Sprachaudiometrie. *HNO.* **49**, 939–959 (2001).

## Figures



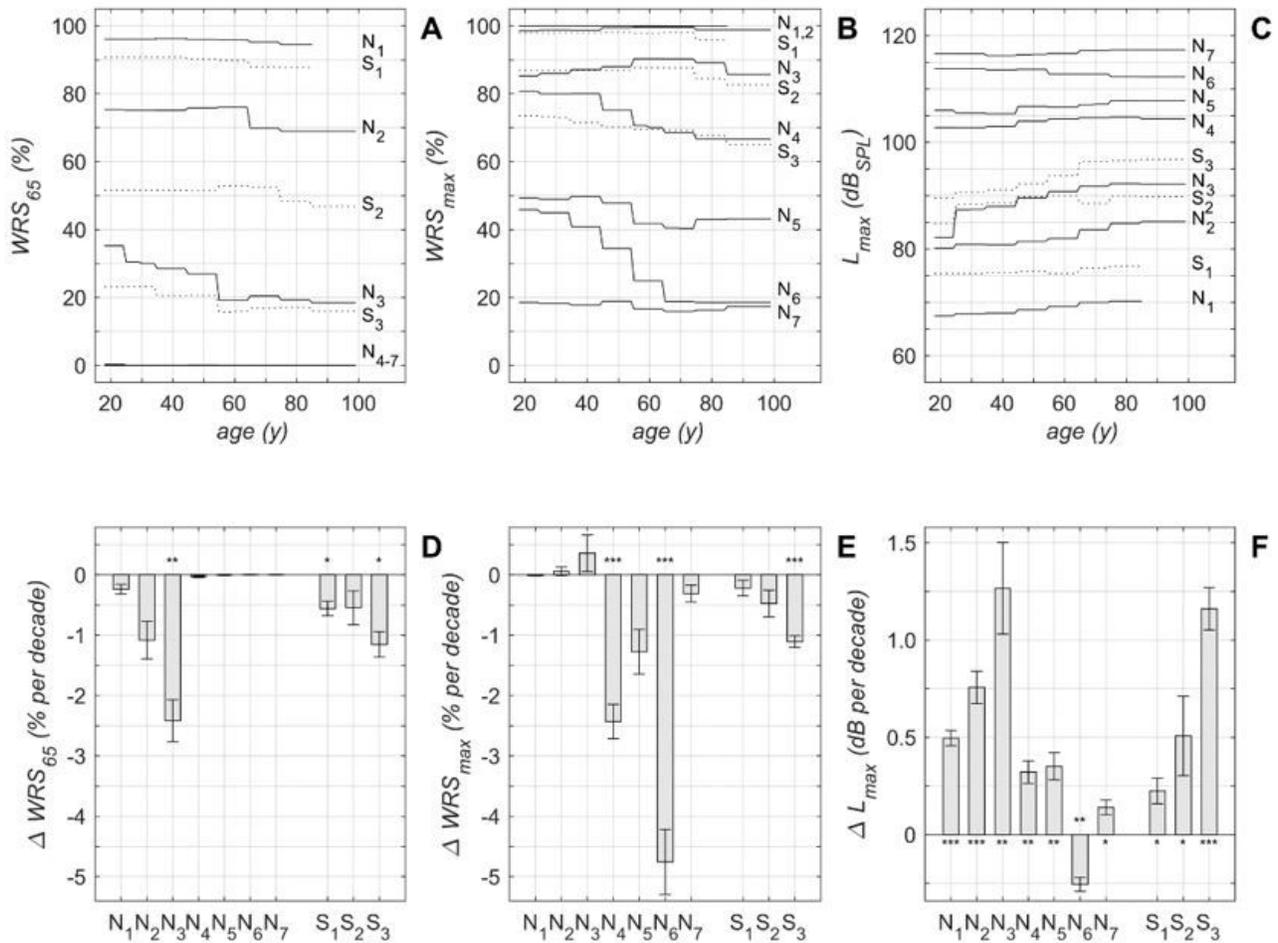
**Figure 1**

Distribution of 19,801 cases with respect to age for different WHO grades of hearing loss.



**Figure 2**

The monosyllabic score (A) at 65 dB SPL presentation level (WRS<sub>65</sub>) and (B) the maximum word recognition score (WRS<sub>max</sub>), for different WHO grades of hearing loss. The boxplots show medians with 1st and 3rd quartiles. The whiskers denote the 0.025 and 0.975 quantiles.



**Figure 3**

Results of the model prediction: (A) age dependence of monosyllabic score at 65 dB SPL presentation level ( $WRS_{65}$ ), (B) maximum word recognition score ( $WRS_{max}$ ) and (C) applied presentation level for  $WRS_{max}$ . The abbreviations N<sub>1</sub> to N<sub>7</sub> and S<sub>1</sub> to S<sub>3</sub> refer to the Bisgaard types of audiograms of the model subjects. The second row (D–F) summarises the age dependences of the above target variables by means of the slope resulting from a linear regression analysis. A negative value corresponds to a decline of the calculated score with age. A positive value corresponds to an increase of the calculated presentation level. Asterisks indicate a significant change of the output variable with age.