Preliminary evaluation of a novel language independent speech-in-noise test for adult hearing screening

Edoardo Maria Polo[1-2], Marco Zanet[3], Alessia Paglialonga[3], Riccardo Barbieri[2]

[1] DIAG, Sapienza University of Rome, Italy
[2] Politecnico di Milano, Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB), Milan, Italy
[3] National Research Council of Italy (CNR), Institute of Electronics, Computer and Telecommunication Engineering (IEIIT), Milan, Italy
polo@diag.uniroma1.it (corresponding author)
marco.zanet@ieiit.cnr.it
alessia.paglialonga@ieiit.cnr.it
riccardo.barbieri@polimi.it

Abstract. This article presents a preliminary evaluation of a novel language independent Speech-in-Noise test for adult screening in terms of Speech Reception Threshold (SRT) estimates and prediction of hearing sensitivity.
The test is based on multiple-choice recognition of meaningless Vowel-Consonant-Vowel words and was administered to 26 normal hearing young adults and 58 unscreened adults who also underwent pure-tone audiometry. Receiver operating characteristics were built using the World Health Organization criteria for "slight/mild" and "moderate" hearing loss as gold standards and SRTs as test outcome. Both curves showed very good test performance in predicting success/failure in pure-tone audiometry (area under the curve: 0.79 for "slight/mild" and 0.83 for "moderate" hearing loss). A complete generalized linear model including SRT, age, and their interaction showed that the SRT and the interaction between SRT and age were significant predictors of pure-tone audiometry outcomes, whereas age alone was not a significant predictor of the degree of hearing loss. Moreover, preliminary results from test-retest data showed that the test was reliable in repeated measures (Spearman's rank-order correlation coefficient = 0.72; Cohen's kappa = 0.83 for "slight/mild" and 0.64 for "moderate" hearing loss). Further research is needed to fully assess test performance in a larger sample of participants, also including subjects with higher degrees of hearing loss (e.g. "severe" and "profound").

Keywords: Speech-in-Noise test, hearing screening, Vowel-Consonant-Vowel, pure-tone audiometry, receiver operating characteristics.

1 Introduction

According to the World Health Organization (WHO), hearing impairment is one of the top ranked leading causes of number of years lived in disability worldwide [1]. Nowadays, about half billion people experience disabling hearing loss and this number is expected to double by 2050 [2].

The substantial implications of hearing loss in terms of increased costs and decreased quality of life are well known [3]. Nevertheless, it is frequently under-diagnosed [4]. Hearing tests, indeed, are not usually included in routine health care examinations in adults. Moreover, a decrease in hearing is often believed as an inevitable consequence of aging and individuals typically do not seek help, or seek help very late. In this context, Speech-in-noise (SIN) tests are a very good option to screen people who may have initial difficulties but may not be ready to seek help [5]-[7]. SIN tests can be implemented in an automatic way and do not need experienced operators or a clinical environment so that they can be performed in remote settings (e.g. at home). Despite the increasing importance of SIN tests for hearing screening, pure-tone audiometry is still the standard test for hearing assessment and, based on pure-tone audiometry outcomes, it is possible to define different grades of hearing loss (as in [8]).

Recently, we have developed a new automated, language independent SIN test for adult hearing screening, based on multiple-choice recognition of meaningless Vowel-Consonant-Vowel (VCV) words spoken by a professional native English speaker [9]-[10]. The test consists in a novel adaptive staircase procedure which takes into account the real shape and SNR range of the different VCV psychometric curves which are obtained by means of the Short Time Objective Intelligibility measure (STOI). Differently than conventional procedures, which use the SNR as the difficulty variable, the novel procedure uses the stimuli's intelligibility, as estimated on the psychometric curves [11]. Preliminary results suggest that the test is fast, reliable in normal hearing and hearing impaired individuals, and accurate in native and non-native listeners.

The aim of this study was to evaluate the ability of the SIN test to predict hearing sensitivity in individuals with varying hearing sensitivity and age. Specifically, we investigated the relationship between the degree of hearing loss defined using the Pure Tone threshold Average (PTA), as measured by pure-tone audiometry, and the Speech Reception Threshold (SRT), as measured by the proposed SIN test.

2 Material and Methods

2.1 Participants

In order to span a wide range of hearing sensitivity and age, two different samples of subjects were recruited and tested. The first sample was composed by a group of normal hearing Italian young adults (YA) (N=26) who were recruited at Politecnico di Milano. The second sample included a group of unscreened adults and older adults (UA) of varying native language (N=58; Italian: 41 subjects; English: 10 subjects; French, African: 2 subjects; German, Spanish and Filipino: 1 subject), who were recruited during local health screening initiatives (i.e. at universities of senior citizens, health prevention and awareness events for the general public).

2.2 Procedures

Participants in the YA and UA groups underwent pure-tone audiometry at 0.5, 1, 2, and 4 kHz on both ears, the Hearing Handicap for the Elderly - Screening version (HHIE-S) and the SIN test. Due to the opportunistic nature of the local screening initiatives, UA were given the option to choose in which ear(s) to perform the test. Seven UA performed the test sequentially in both ears whereas 51 UA and 26 YA performed the test only in one ear, for a total of 91 ears tested.

The test was run on an Apple® Macbook Air® 13'' (OS X Yosemite version 10.10.5 and macOS High Sierra version 10.13.6) connected to Sony MDRZX110APW headphones. Through a volume slider on the graphical user interface, all subjects were given the option to adjust the output levels to a comfortable level before running the SIN test.

2.3 Data Analysis

To evaluate if SRTs measured by the SIN test could predict the outcomes of pure-tone audiometry, receiver operating characteristic (ROC) curves were computed for the SIN test against the WHO criteria [8] for hearing impairment as gold standards.

The WHO criteria are shown in Table 1. The higher the grade of hearing impairment, the higher the PTA, defined as the mean of hearing thresholds measured at the four frequencies tested. Accordingly, each ear was labeled in a binary way: "1" if the WHO criterion was fulfilled (pure-tone audiometry failure) and "0" otherwise. The number of ears that fulfilled each of the criteria is reported in the bottom row in Table 1. Only one ear had "severe" hearing loss and no ears had "profound" hearing loss. Since the purpose of the study was to assess if a SIN test for screening could be able to predict hearing sensitivity of individuals who seek for an initial indication about their hearing loss, criteria (3) and (4) were considered out of scope, since they are related to significant (disabling) hearing impairment, whereas criteria (1) and (2) were addressed.

The $SRT_{cut\text{-}off}$ of the SIN test was varied from 10 to -20 dB SNR to build the ROC curves. Ears that obtained a $SRT < SRT_{cut\text{-}off}$ were classified as "pass" (i.e., good speech recognition performance) and those who obtained a $SRT \geq SRT_{cut\text{-}off}$ as "fail" (i.e., bad speech recognition performance).

The area under the ROC curve (AUC), its 95% confidence interval and the test accuracy were computed for criterion (1) and criterion (2). The best candidate $SRT_{cut\text{-}off}$ for the test was identified as the point on the ROC curve closest to the ideal (0,1) point, and the resulting sensitivity, specificity, and accuracy of the SIN test were addressed.

Table 1. WHO criteria for hearing impairment. First row: WHO criteria as related to PTA. Second row: grades of hearing impairment. Third row: number of ears fulfilling the correspondent criterion out of the total number of ears tested.

WHO criteria	(1) PTA>25 [dB HL]	(2) PTA>40 [dB HL]	(3) PTA>60 [dB HL]	(4) PTA>80 [dB HL]
Hearing loss	Slight/Mild	Moderate	Severe	Profound
Number of ears/ Total number of ears	31/91	7/91	1/91	0/91

To further investigate the relationship between SRT and PTA, a generalized linear model (GLM) was computed. SRT and age were considered as predictors and the binary result coming from the WHO criteria was considered as outcome variable. Age was also addressed as possible predictor to evaluate its possible influence on the probability of fulfilling the WHO criteria. Due to the binary nature of the outcome variable, the logit function was chosen as link function for the GLM. The significance level α was set at 0.05 throughout the study.

A preliminary evaluation to assess the repeatability of the SIN test in terms of SRT estimates and in terms of ability to rank each ear into the same class ("pass" or "fail") in repeated measures was carried out on 20 ears, one per subject, 10 from each group of participants. The Spearman's rank-order correlation coefficient was computed between test and retest SRTs and the Cohen's Kappa statistic (k) was computed between test and retest SIN binary outcomes (i.e. "pass"/"fail") as a function of the $SRT_{cut\text{-}off}$.

3 Results

Table 2 shows the age, the HHIE-S score, the PTA and the SRT obtained in the YA sample and in the UA subgroups defined by the WHO criteria: criterion (1) ($UA_{PTA\leq25}$ and $UA_{PTA>25}$) and criterion (2) ($UA_{PTA\leq40}$ and $UA_{PTA>40}$).

Table 2. Mean, standard deviation and range of: age, HHIE-S score, PTA and SRT in YA and in UA subgroups, classified based on the WHO criteria (1) and (2).

		YA	$UA_{PTA\leq25}$	$UA_{PTA>25}$	$UA_{PTA\leq40}$	$UA_{PTA>40}$
Subjects	number	26	32	26	52	6
	(m/f)	(11/15)	(14/18)	(9/17)	(20/32)	(3/3)
Age	mean	24.2	54.81	66.23	58.92	70.33
[years]	(s.d)	(0.60)	(14.81)	(7.93)	(13.46)	(6.41)
	range	23÷26	24 ÷ 79	48÷80	24÷79	63÷80
HHIE-S	mean	0.0	6.37	11.78	7.52	19.20
[score]	(s.d)	(0.0)	(7.08)	(10.86)	(8.26)	(10.45)
	range	0 ÷ 0	0 ÷ 30	0÷36	0÷34	8÷36
PTA	mean	-0.5	21.29	37.90	26.91	48.21
[dB HL]	(s.d)	(6.46)	(2.79)	(7.90)	(7.58)	(9.15)
	range	-10.0÷20.0	16.25÷25	26.25÷67.50	16.25÷40	41.25÷67.50
SRT	mean	-15.3	-11.24	-6.77	-9.77	-3.61
[dB	(s.d)	(1.87)	(4.01)	(6.87)	(5.55)	(6.81)
SNR]	range	-18.7÷-11.8	-17.75÷-1.25	-18.25÷9.75	-18.25÷9.75	-15÷6.50

Figure 1 shows the ROC curves and the candidate $SRT_{cut\text{-}off}$ associated to criterion (1) (panel (a)) and criterion (2) (panel (b)). Table 3 reports the relevant parameters associated with each ROC curve.

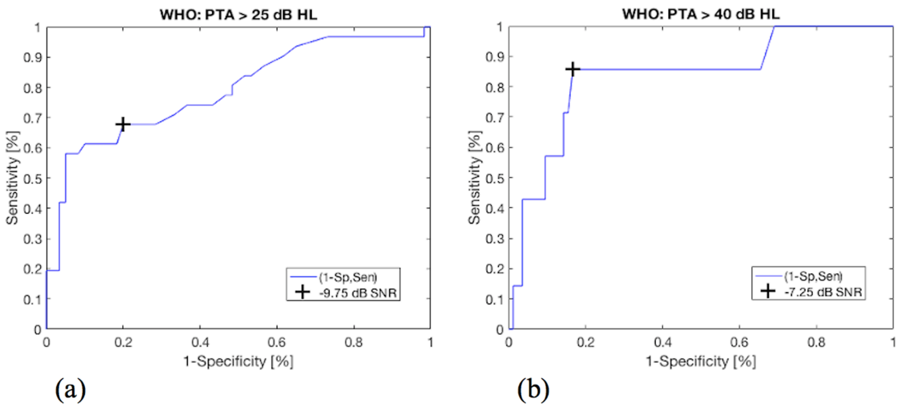

Figure 1. ROC curves related to criterion (1) (panel (a)) and criterion (2) (panel (b)). The black cross mark represents the point closest to the ideal (0,1) point and is associated to the candidate $SRT_{cut\text{-}off}$.

Table 3. Values of SRT_{cutoff}, sensitivity, specificity, accuracy, AUC and AUC confidence interval (CI) for criterion (1) and (2).

WHO criteria	Criterion (1)	Criterion (2)
SRT_{cutoff} [dB SNR]	-9.5	-7
Sensitivity [%]	68	86
Specificity [%]	80	83
Accuracy [%]	76	82
AUC	0.79	0.83
AUC CI	0.70-0.88	0.74-0.93

The GLM highlighted a significant contribution of the SRT in predicting the grade of hearing impairment for criterion (1) ($p << 0.01$) and a significant contribution of the interaction (i.e., the quadratic term in the GLM) between age and PTA ($p << 0.01$), whereas age was not a significant predictor ($p = 0.20$). For criterion (2), the GLM showed that neither SRT nor age nor the interaction between SRT and age were significant predictors of PTA ($p > 0.05$).

The preliminary analysis of test reliability from the subset of subjects in the YA group ($YA_{t/r}$) and in the UA group ($UA_{t/r}$) who underwent the SIN test twice is shown in Table 4. The Spearman's rank-order correlation coefficient was found equal to 0.72 and Cohen's *k* was equal to 0.83 for criterion (1) and 0.64 for criterion (2).

Table 4. Mean, standard deviation and range of: Age, HHIE-S score, PTA, SRT and SRT absolute variation between test and retest trials of YA and UA tested twice.

		$YA_{t/r}$	$UA_{t/r}$
Subjects	number	10	10
	(m/f)	(4/6)	(3/7)
Age [years]	mean	24.2	60.80
	(s.d)	(0.79)	(10.53)
	range	23÷26	42÷80
HHIE-S [score]	mean	0.0	8.4
	(s.d)	(0.0)	(10.49)
	range	0÷0	0÷28
PTA [dB HL]	mean	1.87	29.62
	(s.d)	(3.13)	(7.24)
	range	-1.25÷8.75	20÷41.25
SRT test [dB SNR]	mean	-15.77	-11.02
	(s.d)	(1.98)	(5.88)
	range	-18.75÷-12	-18.25÷0.25
SRT retest [dB SNR]	mean	-15.25	-12.02
	(s.d)	(1.20)	(4.38)
	range	-17.75÷-13.25	-17.25÷-3.5
$SRT_{t/r}$ absolute variation [dB SNR]	mean	1.47	2.25
	(s.d)	(1.04)	(1.63)
	range	0.5÷3.5	0.25÷4.75

4 Discussion

In this study, we performed a preliminary evaluation of a novel language independent SIN screening test in terms of predicting the degree of hearing impairment based on the estimated speech recognition performance and in terms of test reliability.

As a first step, we investigated the relationship between SRT and PTA to assess if the SIN test could predict the possible outcomes of pure-tone audiometry in terms of degrees of hearing impairment, as defined by the WHO criteria for "slight/mild" (criterion (1)) and "moderate" (criterion (2)) hearing loss. As shown by Figure 1 and Table 3, the AUC and confidence intervals were high for both criteria, suggesting a satisfactory performance of the SIN test in predicting the pure-tone audiometry outcomes. As expected (see Table 3), the candidate $SRT_{cut\text{-}off}$ for criterion (1) was worse (i.e. 2.75 dB SNR higher) than the $SRT_{cut\text{-}off}$ for criterion (2), mainly due to the fact that criterion (2) is aimed at classifying higher degrees of hearing loss.

As a second step, a GLM was estimated by considering SRT and age as predictors and the PTA outcome as the binary output variable of the model. For criterion (1), the SRT and the interaction between SRT and age were significant predictors, whereas age alone was found to be not a significant predictor. This suggests that the SRT is significantly related to PTA, including in combination with age, whereas age alone could not predict the success/failure in the pure-tone audiometry criterion. For criterion (2), the GLM showed that none of the variables was a significant predictor. This may be related to the fact that hearing loss increases from criterion (1) ("slight/mild" hearing loss) to criterion (2) ("moderate" hearing loss), and consequently the relationship between SRT and PTA becomes weaker due to the increasing involvement of higher level centers and functions (e.g. auditory processing, short-term memory, and so on) [12]-[13]. Moreover, this may also be related to the fact that the distributions of PTA and SRTs become wider as hearing loss increases, due to higher inter-subject variability [14]-[15].

Lastly, to analyze the test-retest repeatability of the SIN test, i.e. the ability of the test to rank ears in the same "pass" or "fail" class in repeated measures, a subset of subjects in the YA and UA groups performed the SIN test twice. As shown in Table 4, the mean SRTs measured in test and retest trials were similar and the mean absolute SRT variation between test and retest trials was low (i.e. 1.47 dB SNR in YA and 2.25 in UA). Indeed, this variation is lower than the just-noticeable difference in SNR for speech (i.e. 3 dB SNR) [16]. This result suggests a good performance of the test in repeated measures. The Spearman's rank-order correlation coefficient between test and retest SRTs was high, reflecting a strong relationship between the test outcomes in repeated trials, i.e. a high intra-individual repeatability. Similarly, the Cohen's Kappa showed a near perfect agreement for criterion (1) and substantial agreement for criterion (2), suggesting that the pass/fail outcomes of the test are reliable in repeated measures.

5 Conclusion

This work presents the performance of a recently developed automated, language independent SIN test in predicting hearing sensitivity. The SIN test showed a good performance in terms of reliability of SRT estimated in repeated trials and in terms of ability to predict pure-tone audiometry outcomes, particularly for subjects with "slight/mild" hearing loss. Further research is needed to fully validate the SIN test by investigating in a larger population PTA-SRT relationship and test performance in ranking subjects even more impaired than those already tested.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgements

The Authors are grateful to the Lions Clubs International for their support in the organization and management of experiments in the unscreened population of adults.

References

1. World Health Organization (WHO): World report on disability. WHO press, Geneve, Switzerland (2011). Retrieved from: http://whqlibdoc.who.int/publications/2011/97892406852115_eng.pdf
2. World Health Organization (WHO): Deafness and hearing loss, Fact Sheet n. 300. Available from: http://www.who.int/mediacentre/factsheets/fs300/en/
3. Graydon, K., Waterworth, C., Miller, H., & Gunasekera H.: Global burden of hearing impairment and ear disease. Journal of Laryngology and Otology, 133(Special Issue 1), 18–25 (2018).
4. Davis., A., Smith, P., Ferguson, M., Stephens, D., & Gianopoulos, I.: Acceptability, benefit and costs of early screening for hearing disability: a study of potential screening tests and models. Health Technology Assessment, vol. 11, no. 42 (2007).
5. Paglialonga, A., Tognola, G., & Grandori, F.: A user-operated test of suprathreshold acuity in noise for adult hearing screening: The SUN (Speech Understanding in Noise) test. Computers in Biology and Medicine, 52, 66-72 (2014).
6. Leensen M.C, de Laat J.A , Snik A.F, & Dreschler W.A.: Speech-in-noise screening tests by internet, part 2: improving test sensitivity for noise-induced hearing loss. International Journal of Audiology, 50(11), 835-848 (2011).
7. Smits, C., Goverts, T.S., & Festen, J.M.: The digits-in-noise test: assessing auditory speech recognition abilities in noise. The Journal of the Acoustical Society of America, 133(3), 1693-1706 (2013).
8. World Health Organization (WHO), Grades of hearing impairment, https://www.who.int/pbd/deafness/hearing_impairment_grades/en/#
9. Polo, E.M., & Zanet, M.: Development and evaluation of a novel adaptive staircase procedure for automated speech-in-noise testing. M.Sc. Thesis Dissertation, Department of Electronics Information and Bioengineering, Politecnico di Milano, Milan, Italy (2019). Retrieved from https://www.politesi.polimi.it/handle/10589/146105
10. Zanet, M., Polo, E.M., Rocco, G., Paglialonga, A., & Barbieri, R.: Development and preliminary evaluation of a novel adaptive staircase procedure for automated speech-in-noise testing. In: 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 2019, pp. 6991-6994 (2019).
11. Paglialonga, A., Polo, E.M., Zanet, M., Rocco, G., van Waterschoot, T., Barbieri, R.: An automated speech-in-noise test for remote testing: development and preliminary evaluation. The American Journal of Audiology, 2020. In press. https://doi.org/10.1044/2020_AJA-19-00071
12. Bosman, A.J., & Smoorenburg, G.F.: Intelligibility of Dutch CVC syllables and sentences for listeners with normal hearing and with three types of hearing impairment. Audiology, 34(5), 260-284 (1995).
13. Smoorenburg, G.F.: Speech reception in quiet and in noisy conditions by individuals with noise-induced hearing loss in relation to their tone audiogram. The Journal of the Acoustical Society of America, 91(1), 421-437 (1992).
14. Leensen M.C, de Laat J.A , & Dreschler W.A.: Speech-in-noise screening tests by internet, part 1: Test evaluation for noise-induced hearing loss identification. International Journal of Audiology, 50(11), 823-834 (2011).
15. Summers, V., Makashay, M.J., Theodoroff, S.M., & Leek, M.R.: Suprathreshold auditory processing and speech perception in noise: hearing-impaired and normal-hearing listeners. Journal of the American Academy of Audiology, 24, 274-92 (2010).

16. Akeroyd, M.A., McShefferty D., Whitmer, W.M.: The just-noticeable difference in speech-to-noise ratio. Trends in hearing, 19, 2331-2165 (2015).