Telehealth Ecosystems in Practice
M. Giacomini et al. (Eds.)

doi:10.3233/SHTI230768

A Screening Platform for Hearing Loss and Cognitive Decline: WHISPER (Widespread Hearing Impairment Screening and PrEvention of Risk)

Alessia PAGLIALONGA[a,1], Edoardo Maria POLO[b,c], Marta LENATTI[a], Maximiliano MOLLURA[c] and Riccardo BARBIERI[c]
[a] *Cnr-Istituto di Elettronica e di Ingegneria dell'Informazione e delle Telecomunicazioni (CNR-IEIIT), 20133 Milan, Italy*
[b] *Sapienza University of Rome, DIAG, 00185 Rome, Italy*
[c] *Politecnico di Milano, DEIB, 20133 Milan, Italy*

Abstract. The WHISPER (Widespread Hearing Impairment Screening and PrEvention of Risk) platform was recently developed for screening for hearing loss (HL) and cognitive decline in adults. It includes a battery of tests (a risk factors (RF) questionnaire, a language-independent speech-in-noise test, and cognitive tests) and provides a pass/fail outcome based on the analysis of several features. Earlier studies demonstrated high accuracy of the speech-in-noise test for predicting HL in 350 participants. In this study, preliminary results from the RF questionnaire (137 participants) and from the visual digit span test (DST) (78 participants) are presented. Despite the relatively small sample size, these findings indicate that the RF and DST may provide additional features that could be useful to characterize the overall individual profile, providing additional knowledge related to short-term memory performance and overall risk of HL and cognitive decline. Future research is needed to expand number of subjects tested, number of features analyzed, and the range of algorithms (including supervised and unsupervised machine learning) used to identify novel measures able to predict the individual hearing and cognitive abilities, also including components related to the individual risk.

Keywords. Cognitive testing, speech-in-noise testing, cognitive decline, hearing loss, machine learning, multivariate classifiers, hearing screening

1. Introduction

Hearing loss (HL) is among the top leading disabilities, according to the World Health organization (WHO) [1]. It typically develops slowly and progressively in older adults and is frequently related to cognitive decline, also triggering a cascade of effects in terms of social isolation, anxiety, and depression [2]. Screening and early identification of HL and cognitive decline can help slow down or prevent the progression of these conditions and their consequences. Recently, there has been growing attention towards adult hearing

[1] Corresponding Author: Alessia Paglialonga, CNR-IEIIT, Piazza Leonardo da Vinci 32, 20133 Milan, Italy, E-mail: alessia.paglialonga@cnr.it

screening, and particularly towards self-administered speech-in-noise tests that, differently than conventional pure-tone audiometry, can be executed remotely and in non-clinical environments (e.g., [3]-[5]). However, most of the currently available speech-in-noise tests use language-specific stimuli and, thus, need to be adapted to the native language of the target population. Also, they typically use univariate rules to predict HL (for example, based on the measured speech recognition threshold, SRT). Moreover, despite the growing interest in speech-in-noise tests for screening for HL, comparatively limited attention has been given to concurrent identification of cognitive decline and to the individual risk as, for example, current multivariable approaches tend to focus on the auditory and perceptual aspects (e.g., [6]). Recently, we have developed a novel, modular system for screening for HL and cognitive decline, the Widespread Hearing Impairment Screening and PrEvention of Risk (WHISPER) platform. The aim of this contribution is to outline the main characteristics of the system and present the main results obtained through the currently available battery of tests.

2. Materials and Methods

The WHISPER platform was conceived as an integrated system to support screening and prevention of HL and cognitive decline in adults. The system is modular and is based on a battery of tests that can be administered locally and at a distance, and it provides an overall outcome based on the results of one or more tests, as outlined in Figure 1.

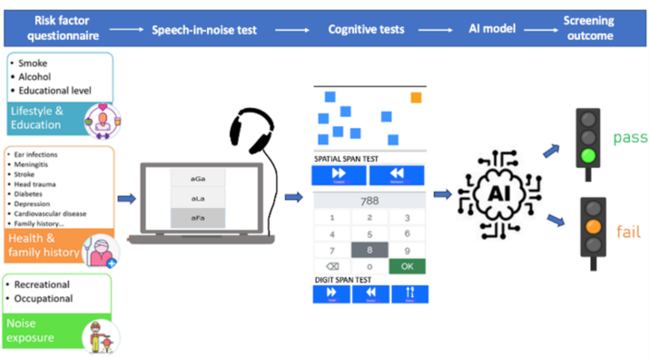

Figure 1. The WHISPER platform includes (i) a questionnaire about RF (lifestyle, noise exposure, and health history); (ii) a language-independent speech in noise test; (iii) a battery of cognitive tests (e.g., spatial span test, DST); (iv) AI models to classify individual performance; and (v) a traffic-light visualization of the test outcome.

Specifically, the platform includes:

- a short survey, including 18 multiple-choice questions about modifiable and non-modifiable risk factors (RF) (e.g., lifestyle, noise exposure, and health);
- a novel, validated user-operated speech-in-noise test [7]-[10]. The test is minimally dependent on the listeners' native language as it is based on an optimized, efficient adaptive procedure using meaningless vowel-consonant-vowel stimuli (e.g., aba, ada) presented in stationary speech-shaped noise. The level of noise is adaptively determined at each trial based on an efficient one-up/three-down staircase. In addition to the commonly measured SRT, the test

extracts a list of additional variables that can be used to identify HL, for example reaction times, percentage of correct responses, and self-adjusted test volume;

- a battery of cognitive tests, including the visual digit span test (DST) (forward, backward, and ordering versions) and the spatial span test (forward and backward versions). In the DST, the ability to recall sequences of digits of increasing length is tested. In the forward, backward, and ordering version, the subject is asked to dial/type the sequence in the same, opposite, or ordered version, respectively. In the spatial span test, the ability to recall spatially distributed sequences of boxes is tested, either in the forward or backward recall task. The digit span score (DSS) and the spatial span score are defined as the length of the longest sequence recalled by the subject. The lower the scores, the lower the performance, and the higher the chance that cognitive decline may occur. In a similar way as in the abovementioned speech-in-noise test, the platform extracts a variety of variables in addition to the scores, e.g., the average digit typing time and the pattern of responses.

The dataset collected using the WHISPER platform includes 442 records from 350 participants tested in laboratory environment and at hearing screening and awareness events, as of June 2023. The sample includes 117 men and 233 women (age: mean 50 years, range: 19-89 years) across 12 native languages. The study protocol was approved by the Politecnico di Milano Research Ethical Committee (opinion no. 2/2019, Feb 19, 2019; opinion no. 13/2022, Apr 13, 2022). The dataset includes 44 features, specifically:

- 18 from the RF questionnaire, e.g., presence/absence of a certain RF, and educational level (N=137 participants);
- 7 from the speech-in-noise test, e.g., SRT, average reaction time, percentage of correct responses (N=350);
- 7 from the forward DST, e.g., DSS, average digit time (N=78);
- 5 from the general individual and session characteristics, e.g., age, gender at birth, native language, transducers (N=350);
- 7 from pure-tone screening testing, e.g., pure-tone thresholds in the range 0.5-4kHz (N=350), and in the range 6-8kHz (N=235), ear asymmetry (N=92), as measured using pure-tone audiometry with a clinical audiometer.

As shown in Figure 1, a screening outcome is visualized using a traffic-light visualization and a pass/fail indication, computed using AI-based models that are currently under investigation. Recent research has shown that the features extracted from the speech-in-noise test can accurately predict HL [9]. Specifically, using different machine learning algorithms on different subsets of features on a population of 350 participants, the accuracy of binary classification of HL of mild and moderate degree was equal to 0.85 and 0.84, respectively. The accuracy for multiclass identification of normal hearing (NH), mild, and moderate HL was equal to 0.72, i.e., similar to other validated (but language-dependent) speech-in-noise tests (e.g., [3]). In the followings, preliminary results obtained from the analysis of features extracted from the short-term memory testing module (specifically, the DST) and from the RF survey are presented.

3. Results and Discussion

Figure 2 shows preliminary results obtained from the DST (N=78 participants; no. of NH/HL: 47/31). Panel A shows that the DSS was, on average, higher (i.e., better) in individuals with lower SRT and in individuals with NH. Nevertheless, four individuals with HL and with high SRT (i.e., SRT>0 dB SNR) obtained DSS scores higher than the cut-off value of 5 (up to 7), suggesting that good short-term memory can be observed in individuals with impaired speech recognition (as measured by the SRT) and impaired hearing sensitivity (as measured by pure-tone thresholds). Panel B show that the median age in individuals with DSS>5 was significantly lower than that of individuals with DSS≤5, with 75% of individuals with DSS>5 being younger than 55 years old. A similar finding was observed for the single digit typing time shown in Panel C, confirming a complex relationship between age, average reaction times, and short-term memory.

Regarding the RF questionnaire, Figure 3 shows a higher proportion of depression, high cholesterol, alcohol consumption, and lower educational level in individuals with HL, and a higher percentage of self-reported recreational noise exposure in individuals with NH. These findings are likely to be, at least in part, influenced by the different age range of NH and HL groups. Further analysis is necessary, on a larger, balanced sample, to investigate the differences in RF profiles between individuals with NH and HL.

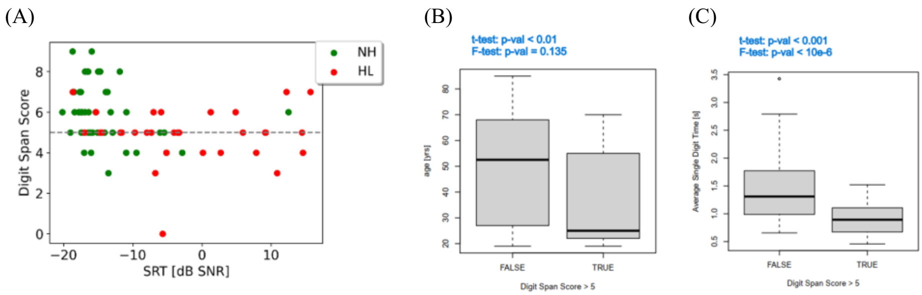

Figure 2. DSS as a function of SRT in individuals with NH and HL (Panel A); distribution of age (Panel B) and average single digit typing time (Panel B) as a function of the DSS class (N=78 participants).

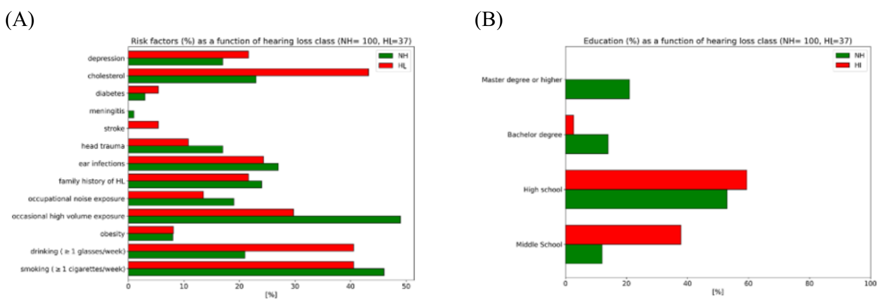

Figure 3. Distribution of RF related to health, noise exposure, lifestyle (Panel A) and education (Panel B) in individuals with NH and HL (N=137 participants).

4. Conclusions

Overall, the preliminary results described here suggest that the battery of tests included in the WHISPER platform may be able to capture different aspects of the individual performance, not limited to hearing sensitivity (as measured by pure-tone testing) and speech recognition performance (as measured by speech-in-noise testing). Future research is needed to investigate more deeply the relationships between the different measures extracted from the test battery and their role in defining an individual profile. Moreover, the analysis of the most relevant confounding factors to correct for, e.g., subject's age, is of primary interest. Further analysis of the set of features extracted from these two tests using unsupervised and supervised machine learning approaches, as well as causal models, could help identify homogeneous groups of individuals and unravel the causal associations between the identified RF and the different components of individual performance. This could, in turn, lead to a novel score that combines the speech-in-noise test and the DST, describing the individual overall performance. Further research is also needed to understand if, and to what extent, such score could be integrated with the RF questionnaire to define an global risk score that considers both the current overall performance and the risk of developing HL and cognitive decline.

Acknowledgments

This work was supported in part by the Capita Foundation through Project WHISPER, Widespread Hearing Impairment Screening and PrEvention of Risk (2022 and 2020 Capita Foundation Auditory Research Grants). Marta Lenatti is a PhD student enrolled in the National PhD in Artificial Intelligence, XXXVIII cycle, course on Health and life sciences, organized by Università Campus Bio-Medico di Roma.

References

[1] World Health Organization. (2021). Deafness and hearing loss. https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss.

[2] Fulton SE, Lister JJ, Bush AL, Edwards JD, Andel R. Mechanisms of the hearing-cognition relationship. Semin Hear. 2015 Aug;36(3):140-9. doi: 10.1055/s-0035-1555117.

[3] Watson CS, et al. Telephone screening tests for functionally impaired hearing: current use in seven countries and development of a US version. J Am Acad Audiol. 2012 Nov-Dec;23(10):757-67.

[4] Leensen MC, de Laat JA, Dreschler WA. Speech-in-noise screening tests by internet, part 1: test evaluation for noise-induced hearing loss identification. Int J Audiol. 2011 Nov;50(11):823-34. doi: 10.3109/14992027.2011.595016.

[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. Comput Biol Med. 2014 Sep;52:66-72. doi: 10.1016/j.compbiomed.2014.06.012

[6] Saak S, Huelsmeier D, Kollmeier B, Buhl M. A flexible data-driven audiological patient stratification method for deriving auditory profiles. Front Neurol. 2022 Sep 15;13:959582. doi: 10.3389/fneur.2022.959582.

[7] Paglialonga A, Polo EM, Zanet M, Rocco G, van Waterschoot T, Barbieri R. An Automated Speech-in-Noise Test for Remote Testing: Development and Preliminary Evaluation. Am J Audiol. 2020 Sep 18;29(3S):564-576. doi: 10.1044/2020_AJA-19-00071.

[8] Rocco G, Bernardi G. Ali R, van Waterschoot T, Polo EM, Barbieri R, Paglialonga A. Characterization of the Intelligibility of Vowel-Consonant-Vowel (VCV) Recordings in Five Languages for Application in Speech-in-Noise Screening in Multilingual Settings. Appl Sci. 2023; 13:5344.

[9] Polo EM, Mollura M, Barbieri R, Paglialonga A. Multivariate Classification of Mild and Moderate Hearing Loss Using a Speech-in-Noise Test for Hearing Screening at a Distance. In: Spinsante, S., Iadarola, G., Paglialonga, A., Tramarin, F. (eds) IoT Technologies for HealthCare. HealthyIoT 2022. LNICS 456:81-92. Springer, Cham, 2023. doi: 10.1007/978-3-031-28663-6_7

[10] Lenatti M, Moreno-Sánchez PA, Polo EM, Mollura M, Barbieri R, Paglialonga A. Evaluation of Machine Learning Algorithms and Explainability Techniques to Detect Hearing Loss from a Speech-in-Noise Screening Test. Am J Audiol. 2022 Sep 21;31(3S):961-979. doi: 10.1044/2022_AJA-21-00194.