Aud Vestib Res (2020);29(1):39-47.
DOI: 10.18502/avr.v29i1.2368

**RESEARCH ARTICLE**

# Auditory recognition of Persian digits in presence of speech-spectrum noise and multi-talker babble: a validation study

**Amin Ebrahimi, Mohammad Ebrahim Mahdavi*, Hamid Jalilvand**

Department of Audiology, School of Rehabilitation, Shahid Beheshti University of Medical Sciences, Tehran, Iran

Received: 25 Sep 2019, Revised: 30 Oct 2019, Accepted: 9 Nov 2019, Published: 15 Jan 2020

## Abstract

**Background and Aim:** Digits are suitable speech materials for evaluating recognition of speech-in-noise in clients with the wide range of language abilities. Farsi Auditory Recognition of Digit-in-Noise (FARDIN) test has been developed and validated in learning-disabled children showing dichotic listening deficit. This study was conducted for further validation of FARDIN and to survey the effects of noise type on the recognition performance in individuals with sensory-neural hearing impairment.

**Methods:** Persian monosyllabic digits 1−10 were extracted from the audio file of FARDIN test. Ten lists were compiled using a random order of the triplets. The first five lists were mixed with multi-talker babble noise (MTBN) and the second five lists with speech-spectrum noise (SSN). Signal- to- noise ratio (SNR) varied from +5 to −15 in 5 dB steps. 20 normal hearing and 19 hearing-impaired individuals participated in the current study.

**Results:** Both types of noise could differentiate the hearing loss from normal hearing. Hearing-impaired group showed weaker performance for digit recognition in MTBN and SSN and needed 4−5.6 dB higher SNR (50%), compared to the normal hearing group. MTBN was more challenging for normal hearing than SSN.

**Conclusion:** Farsi Auditory Recognition of Digit-in-Noise is a validated test for estimating SNR (50%) in clients with hearing loss. It seems SSN is more appropriate for using as a background noise for testing the performance of auditory recognition of digit-in-noise.

**Keywords:** Auditory recognition; hearing loss; speech perception in noise; digit recognition in noise

* **Corresponding author:** Department of Audiology, School of Rehabilitation, Shahid Beheshti University of Medical Sciences, Damavand Ave., Tehran, 1616913111, Iran. Tel: 009821-77542057, E-mail: mahdavime@sbmu.ac.ir

**Citation:** Ebrahimi A, Mahdavi ME, Jalilvand H. Auditory recognition of Persian digits in presence of speech-spectrum noise and multi-talker babble: a validation study. Aud Vestib Res. 2020;29(1):39-47.

## Introduction

There are many studies that confirm noisy environments are challenging for patients with hearing loss. People with hearing loss usually complain that they hear the speech but cannot understand it properly and interfering noise or speech exacerbates their problem [1-4]. Speech recognition in quiet is routinely performed for patients and can provide valuable results. However, its results cannot predict the patient's performance in everyday communication situations unless the recognition score in quiet is poor [5-9].

Audiometry using speech stimuli, as its name implies, uses speech to evaluate the auditory system. Speech audiometry assesses two components: sensitivity to speech and clarity of the heard speech. Hearing loss is partly related to the ability to be heard (auditory sensitivity) and partly to the distortion component (speech clarity). The part related to auditory sensitivity is determined by assessing speech recognition in quiet [3-5] that are rated by speech recognition score (SRS) or word recognition score (WRS). In this test, the correct percentage of words recognized by a person at a sound intensity level (dB SL) is calculated relative to the speech recognition threshold (SRT). The speech recognition test in quiet aims to assess how one's speech is understood in a quiet environment while the speech intensity is high enough for the person to earn the highest score. The highest score a person gets in this test is called $PB_{MAX}$. Various materials such as nonsense sentences and syllables are used to evaluate the ability to recognize speech in quiet, but the most common materials are monosyllabic words [4].

Speech recognition testing in quiet is performed in many clinics every day. A review of studies in this area shows that 91% of audiologists do this test every day, and 92% of them use testing of suprathreshold monosyllabic word recognition in quiet. One of the limitations of speech recognition test in quiet (at a single level of presentation) is its failure to reflect one's performance in the real world. Moreover, most hearing-impaired patients do not complain of not understanding words in a quiet environment [5]. Carhart and Tillman (1970) for the first time suggested that in addition to the usual audiometric tests and speech recognition in quiet, the ability to recognize speech in competitive noise also needs to be assessed, and by describing a hearing loss as signal-to-noise ratio (SNR), an estimate of the patient communication disability is obtained [6].

The second part of hearing loss (distortion component) refers to impaired speech recognition ability, especially in the presence of noise, and is determined by the correct percentage of speech recognition in noise [3-5]. Recently, another form of SNR has been used, and it is called signal-to-noise ratio for 50% correct score (SNR (50%)). The score is calculated as the SNR required to obtain a 50% recognition score [4]. Distortion in the auditory signal may occur anywhere in the signal transduction pathway from the peripheral to the auditory cortex. For example, varying degrees of hearing loss in the peripheral auditory system that results in signal filtering can lead to distortion. Regardless of where the distortion occurs in the speech signal, the most distorting manifestation is the decrease in one's ability to perceive speech in the presence of background noise. However, a reverse relationship exists between the number and size of distortion in the auditory system and one's ability to recognize speech in noise [3].

Although sentences are more commonly used in everyday situations and considered as the most appropriate test material with higher face validity than words and syllables, however, using sentence has some limitations; the listener needs a high SNR to understand the whole sentence, which is not possible for some people with severe or profound hearing loss or people with poor language proficiency. Also, it is not simply possible to perform a sentence recognition in noise test in children and many users of cochlear implant [10]. Sentences cannot be used to test the elderly population because sentence-level stimuli require the listener to memorize and retrieve several words, which limits the use of this test often due to poor memory in the elderly [5]. The mentioned constraints do not apply for monosyllabic words, and the effect of memory on the results can be minimized [5,10].

When using digits 1−10 for speech audiometry, a few alternatives are available for each digit (eight alternatives for each monosyllabic digit between 1−10) [3]. However, if the digits are presented in a triplet (e.g. 2-3-5), the number of permutations with repetition would be 729, and without repetition would be 504. If we want to construct triplets with repetition, the number of combinations would be 165 and without repetition (unique triplets) would be 84, which would be far more than the eight alternatives. On the other hand, the items of a triple is much smaller than the working memory capacity or the Miller number [7], and has no considerable memory load. Digits are familiar and simple speech materials that can be used to evaluate a wide range of patients with language abilities and especially in cochlear-implanted children and bilinguals that language vocabulary is a concern. Using the conventional audiometric test, Ramikssoon et al. attempted to determine SRT in non-native English speakers using disyllabic digits instead of words. They found that disyllabic digits could determine SRT in non-native English speakers more accurately than disyllabic words [11].

Although broadband noise (BBN) has been used for speech-in-noise tests in the past [12-15]. The frequencies located lower and higher than the frequency band of the target speech are found to be ineffective on masking, and a tendency to use narrow-band noise was observed in the literature. Researchers have recently used multi-talker babble noise (MTBN), speech-spectrum noise (SSN) and speech-weighted noise (SWN) [10-18]. Therefore, studies should be conducted on the recognition of Persian digits [17] in the presence of different noises previously evaluated in children with learning disabilities. This study was conducted with two purposes. First, we wanted to know whether the recognition of digits in the presence of different noises can differentiate between people with hearing loss and those with normal hearing (validation) secondly, whether digit recognition is affected by noise type, and which noise is more appropriate for evaluating Persian digits recognition in noise.

## Methods

This research is a descriptive-analytical study based on observation. The study subjects included 20 people with normal hearing with mean age of 23.4 years (NH group) and 19 sensory-neural hearing loss (HL group) participants with mean age of 50.7 years referred to the audiology clinic of Amir A’alam Hospital affiliated to Tehran University of Medical Sciences (TUMS) in Iran. They were selected using convenience sampling method, which is a non-random sampling technique. All subjects were right-handed and the test was performed on their right ear. Also, all obtained WRS test scores ≥ 90% and had a healthy status of the tympanic membrane and middle ear based on tympanometry test results. Their maximum gap between air- and bone-conduction thresholds for each frequency was 10 dB, and they had no history of any neurological diseases. For people with normal hearing, the age ranged between 19 and 25 years, and their hearing threshold of 250 to 8000 Hz frequencies was ≤ 20 dB HL. For HL group, the age range was 20-65 years and the threshold of 500, 1000, and 2000 Hz frequencies were in the range of mild to moderate. Their pattern of hearing loss was symmetric such that the difference in hearing thresholds at frequencies of 500, 1000, 2000, and 4000 Hz in the left and right ears were ≤ 10 dB (Fig. 1).

Persian monosyllabic digits of 1 to 10 were extracted from the audio file of Farsi Auditory Recognition of Digit-in-Noise (FARDIN) test [17], and 10 lists of digit triplet were compiled using a randomization software. The audio file of the test material was saved as a “wav” file. To create the Persian SSN, we recorded the voices of 10 young men and 10 young women reading a newspaper. Those newspaper texts were selected that contained all Persian phonemes. After recording the audio for one minute, a duration of 40-second of each recorded voice was selected and cut. The root mean square (RMS) of this file was calculated at different frequencies of 100 to 10000 Hz, and then, the amount of energy at each frequency was determined. In the end, white noise was filtered based on this pattern. The MTBN of the six-talker used in the present study was the same as that used in the study of Heidari et al. [10]. Finally, the first 5 lists of triplet sets were combined with MTBN, and the other 5 lists were combined with SSN. Fig. 2 shows the electric spectrum of MTBN and SSN in dB relative to full scale (dB FS). The noise intensity level was fixed at 85 dB SPL for both groups and triplet sets were presented to the right ear of the subjects at SNRs of −15, −10, −5, 0, and 5 dB [10]. In each SNR, 6 triplet sets of digits were presented and the SNR change was in descending order.

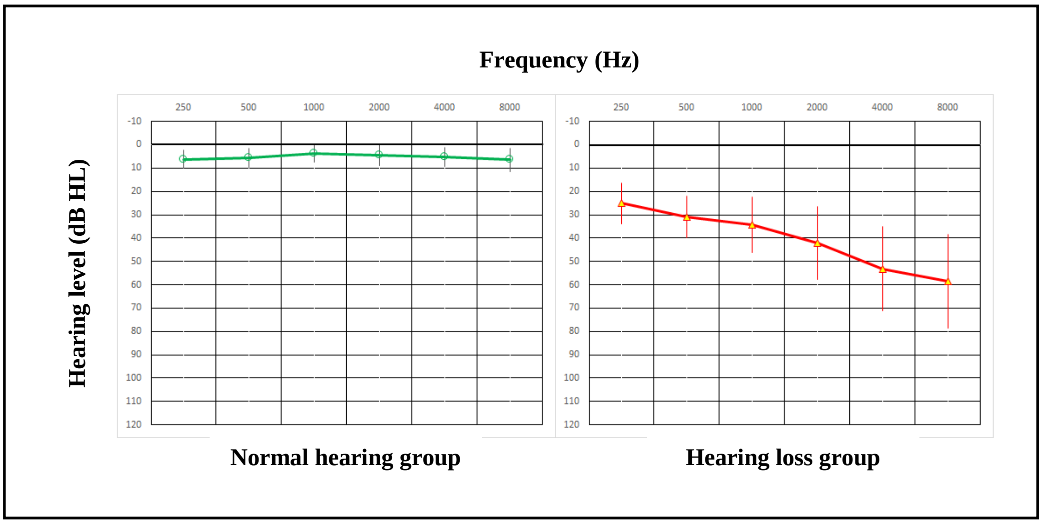


**Fig. 1. Mean and standard deviation of test ear (right ear) AC hearing thresholds in sensory-neural hearing loss (right panel) and normal hearing (left panel) groups.**

The data were entered into SPSS 23, and the mean and standard deviation of the correct recognition performance of each group were calculated for each digit and also for all digits. The non-parametric Mann-Whitney test was used to compare the mean recognition performance of each group, and the level of significance was set at p = 0.05. The SNR (50%) and the slope of 20%-80% were calculated on the third-degree polynomial psychometric function fitted to measured data.

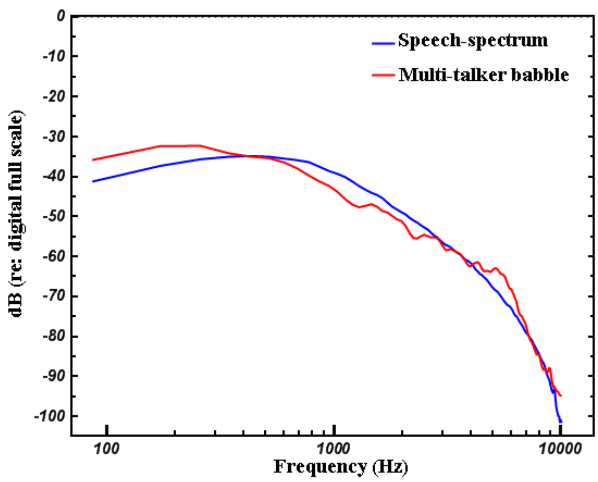


**Fig. 2. Electrical spectrum of the multi-talker babble noise (red) and speech-spectrum noise (blue) based on dB relative to full scale (dB FS).**

## Results

Table 1 shows the SNR (50%) and the slope of section 20%−80% on the psychometric function for different digits, the mean function of all digits, as well as the difference in these characteristics between NH and HL groups. As it can be seen, the HL group, compared to the NH group, required a higher SNR of 2−7 dB to obtain a 50% recognition score for each digit in the presence of SSN. This difference for MTBN varied from 1−8 dB. These values for the mean function of all digits for SSN and MTBN noise are 5.65 dB and 4.04 dB, respectively.

Fig. 3 shows the psychometric function of the mean recognition of all Persian monosyllabic digit 1−10 in MTBN noise and SSN, in the two groups. As it can be seen from the values in Table 1, the slope of the psychometric function of all digits in the section of 20%−80% in the NH group for MTBN and SSN is 7.30 and 7.11 %/dB and in HL group is 7.35 and 6.30 %/dB respectively. The slope becomes slightly slower in HL group for SSN than MTBN while, it is similar for the both noise in NH group with a liberal bias for SSN.

Fig. 4 illustrates the psychometric function of

**Table 1. SNR (50%) and slope of 20%-80% on the psychometric function of monosyllabic Persian digits 1-10 and mean function of all digits in presence of multi-talker babble noise and speech-spectrum noise in the both groups**

| | SSN | | | | | | MTBN | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | SNR (50%) (dB) | | | Slope (20%-80%) (%/dB) | | | SNR (50%) (dB) | | | Slope (20%-80%) (%/dB) | | |
| Digit | Normal hearing | Hearing loss | Difference | Normal hearing | Hearing loss | Difference | Normal hearing | Hearing loss | Difference | Normal hearing | Hearing loss | Difference |
| **1** | -13.32 | -8.54 | 4.78 | 11.09 | 8.53 | -2.55 | -9.31 | -7.27 | 2.03 | 7.51 | 8.43 | 0.92 |
| **2** | -6.51 | -2.34 | 4.17 | 6.81 | 6.95 | 0.14 | -4.94 | -3.01 | 1.93 | 11.11 | 8.63 | -2.48 |
| **3** | -13.36 | -7.41 | 5.95 | 10.68 | 8.15 | -2.53 | -12.42 | -5.40 | 7.02 | 8.80 | 7.76 | -1.04 |
| **5** | -14.61 | -12.54 | 2.06 | 12.58 | 9.90 | -2.68 | -12.00 | -6.81 | 5.19 | 10.19 | 8.15 | -2.03 |
| **6** | -10.88 | -8.00 | 2.88 | 9.84 | 2.95 | -6.88 | -11.85 | -6.29 | 5.56 | 8.65 | 5.78 | -2.87 |
| **7** | -12.47 | -9.15 | 3.32 | 10.44 | 8.57 | -1.87 | -12.00 | -10.21 | 1.80 | 12.07 | 6.66 | -5.41 |
| **8** | -14.90 | -7.52 | 7.37 | 13.31 | 6.51 | -6.79 | -12.80 | -4.09 | 8.71 | 10.92 | 8.06 | -2.86 |
| **9** | -9.12 | -5.45 | 3.67 | 5.97 | 8.97 | 3.00 | -6.24 | -4.19 | 2.05 | 9.52 | 10.51 | 0.99 |
| **10** | -8.32 | -5.72 | 2.60 | 9.08 | 6.36 | -2.72 | -6.43 | -5.06 | 1.37 | 8.92 | 7.36 | -1.57 |
| **All digits** | -12.97 | -7.32 | 5.65 | 7.35 | 6.30 | -1.06 | -9.74 | -5.70 | 4.04 | 7.30 | 7.11 | -0.19 |

SSN; speech-spectrum noise, MTBN; multi-talker babble noise, SNR; signal-to-noise ratio

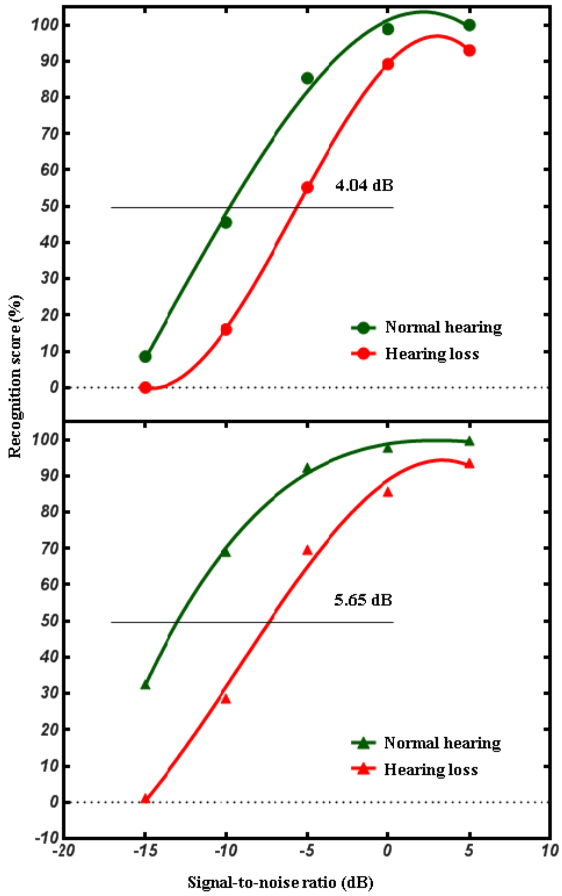


**Fig. 3. Psychometric function of mean recognition score for all Persian monosyllabic digits 1-10 divided in multi-talker babble noise (top) and speech-spectrum noise (down) in both groups.**

the mean recognition of all Persian monosyllabic digits from 1 to 10 in the presence of MTBN and SSN categorized by study groups. Based on this comparison, it is more challenging for NH group to recognize digits in the presence of MTBN than SSN. In other words, on average, NH group needed 3.2 dB higher SNR to recognize digits in the presence of MTBN compared to SSN. This finding is also true for people with hearing loss, who needed a higher SNR of 1.6 dB for digit recognition in the presence of MTBN. This finding suggests that MTBN creates a greater challenge for the NH group.

Fig. 5 shows the mean correct recognition with a 95% confidence interval for different digits in the NH and HL groups in the presence of MTBN and SSN. Mean correct recognition score in the presence of SSN in the HL group was significantly weaker than that of the NH group ($p < 0.05$ for digits 1, 2, 5, 7, 9, 10, and $p < 0.001$ for digits 3, 6, and 8). However, the mean correct recognition of digits 1, 2, 9, and 10 in the presence of MTBN did not show a significant difference between NH and HL groups. This difference was significant in the other digits ($p < 0.05$ for digits 5 and 7 and $p < 0.001$ for digits 3, 6, and 8). The mean ± SD of recognition score for all digits in the presence of SSN for the HL group (55.7% ± 38) was significantly weaker than that of the NH group (78.3% ± 27) ($p < 0.001$). In the presence of MTBN, the mean ± SD recognition score of the HL group (50.7% ± 40) was also significantly weaker than that of the NH group (67.7% ± 37) ($p < 0.05$).

## Discussion

The present study showed that HL group required an average SNR (50%) of 5.65 dB and 4.04 dB for all digits in the presence of SSN and MTBN, respectively, to have hearing performance equal to NH group. This difference in SNR (50%) between the two study groups confirms the complaint of hearing-impaired people of not understanding speech in the presence of noise. The current result is in line with many other studies [2,3,18]. The comparison of the psychometric function of two noises showed that a slope of 20%‒80% for SSN in the HL group for all digits (except 2 and 9), and their overall mean function was slower compared to the NH group. This finding was also true for MTBN except that the slope reduction was not observed for digits 1 and 9. Digit recognition in the presence of MTBN for both groups was more difficult than that in the presence of SSN. In other words, NH group needed 3.2 dB higher SNR, and HL group needed 1.6 dB higher SNR to recognize digits in the presence of MTBN compared to SSN.

The six-talker babble used in this study was taken from Heidari et al., and the level of noise intensity was similar to that of the study. As a

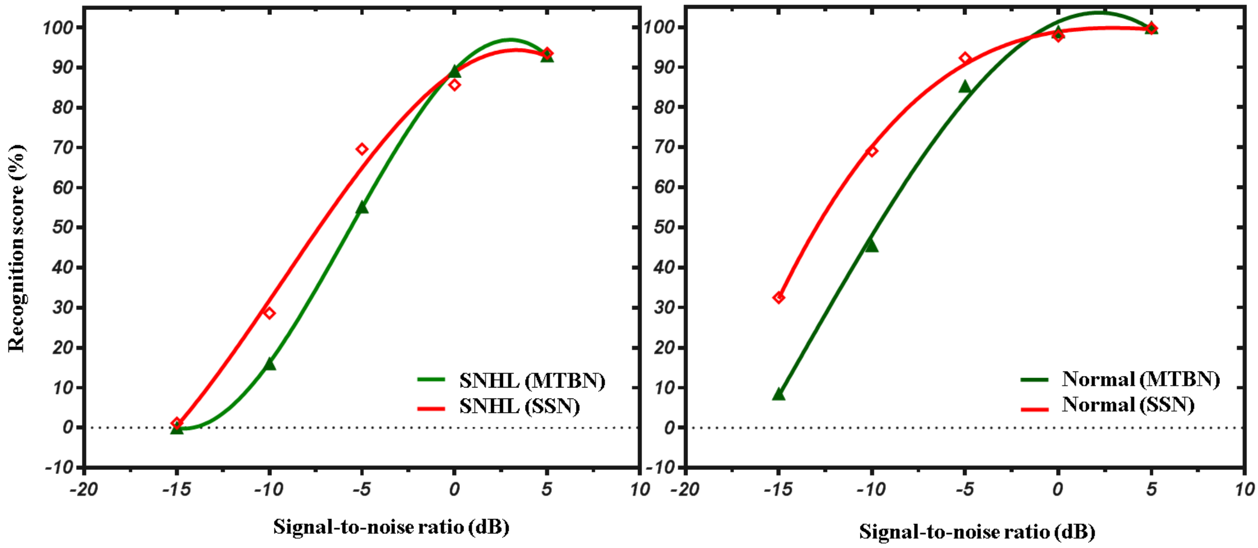


**Fig. 4. Psychometric function of mean recognition score for all Persian monosyllabic digits 1-10 in presence of multi-talker babble noise and speech-spectrum noise divided in normal (right) and hearing loss (left) groups.**

result, our results in the NH group for MTBN (−9.7 dB for mean function) were very similar to the Heidari et al. study (−9.5 and −9.6 dB using Spearman-Karber method and Probit regression respectively) [10].
McArdle et al. investigated the effect of various

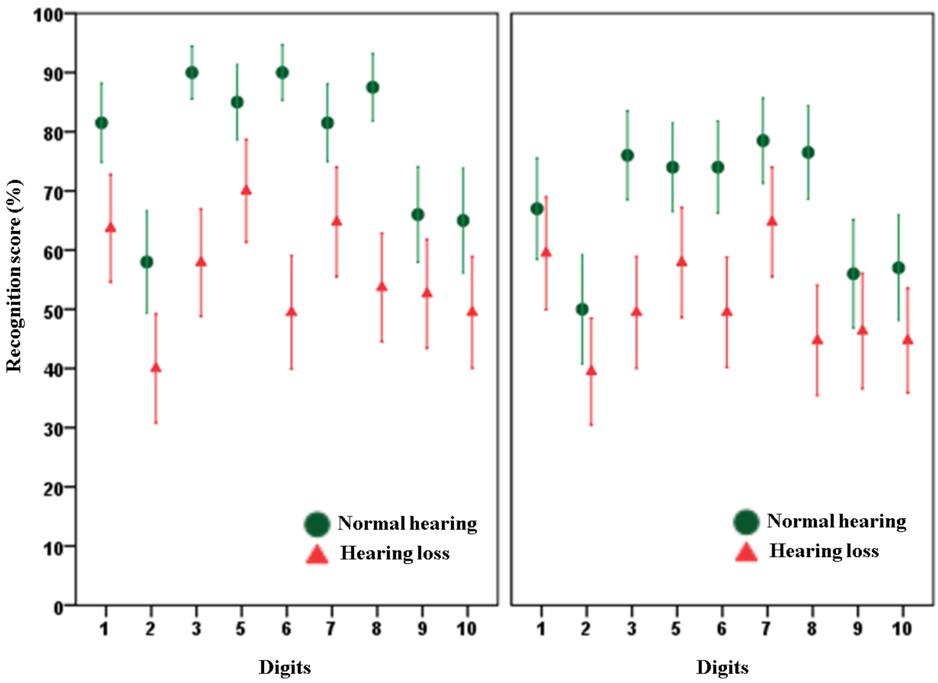


**Fig. 5. Mean (confidence: 95%) recognition score for all Persian monosyllabic digits 1−10 in presence of multi-talker babble noise (right) and speech-spectrum noise (left) in both groups.**

speech materials, including digits, monosyllabic words, and sentences presented in MTBN. They found that NH listeners and HL group needed less SNR to recognize digits than word and sentence stimuli. This finding was attributed to factors such as differences in linguistic complexity of different spoken materials, closed-set response for digits and semantic and calibration issues. They also found out that all the test materials could differentiate between the recognition results of individuals with NH and HL. McArdle et al. reported an 8 dB difference between their normal and hearing loss groups for digits recognition [5]. In our study, this difference was 4 dB.

Wilson et al. investigated the difference between MTBN and SSN as a background noise for the word-in-noise testing. They found that people with NH perform slightly better in the presence of MTBN than in SSN noise. This difference was related to the amplitude modulation of MTBN noise, which resulted in a slight improvement in their performance. Both types of noise could make a difference in the performance of HL and NH groups. Eventually, MTBN noise appeared to be better noise only because of its face validity in mimicking everyday listening conditions. However, the results of hearing-impaired subjects were the same for both types of noise [18].

In the present study, MTBN was more challenging for normal group than SSN, so the mean recognition for some digits (1, 2, 9 and 10) in MTBN did not show a significant difference between the two study groups, while the mean recognition of the hearing-impaired group in the presence of SSN was significantly weaker than the normal group for all digits (Fig. 5). Therefore, the present study suggests that SSN noise is more effective background noise to differentiate between HL and NH.

## Conclusion

FARDIN test has an acceptable validity for assessing speech-in-noise recognition ability of patients with sensory-neural hearing loss. Compared to multi-talker babble noise, it seems speech-spectrum noise to be more appropriate for Persian digit recognition testing in noise by yielding a greater difference between the performance of the normal hearing and hearing loss groups.

## Acknowledgments

This article has been extracted from A. Ebrahimi MSc. thesis on Audiology submitted to Shahid Beheshti University of Medical Sciences with Ethical Code No. IR.SBMU. RETECH.REC.1398.342.

## Conflict of interest

The authors declare that they have no conflict of interest.

**References**

1. Wilson RH, Abrams HB, Pillion AL. A word-recognition task in multitalker babble using a descending presentation mode from 24 dB to 0 dB signal to babble. J Rehabil Res Dev. 2003;40(4):321-7. doi: 10.1682/jrrd.2003.07.0321
2. Wilson RH, Burks CA, Weakley DG. A comparison of word-recognition abilities assessed with digit pairs and digit triplets in multitalker babble. J Rehabil Res Dev. 2005;42(4):499-510. doi: 10.1682/jrrd.2004.10.0134
3. Wilson RH, Weakley DG. The use of digit triplets to evaluate word-recognition abilities in multitalker babble. Semin Hear. 2004;25(1):93-111.
4. McArdle R, Hanth-Chisolm T. Speech audiometry. In: Katz J, Chasin M, English K, Hood LJ, Tillery KL, editors. Handbook of clinical Audiology. 7th ed. New York: Wolters Kluwer; 2015. p. 61-76.
5. McArdle RA, Wilson RH, Burks CA. Speech recognition in multitalker babble using digits, words, and sentences. J Am Acad Audiol. 2005;16(9):726-39. doi: 10.3766/jaaa.16.9.9
6. Carhart R, Tillman TW. Interaction of competing speech signals with hearing losses. Arch Otolaryngol. 1970;91 (3):273-9. doi: 10.1001/archotol.1970.00770040379010
7. Plomp R, Duquesnoy AJ. A model for the speech-reception threshold in noise without and with a hearing aid. Scand Audiol Suppl. 1982;15:95-111.
8. Martin FN, Champlin CA, Chambers JA. Seventh survey of audiometric practices in the United States. J Am Acad Audiol. 1998;9(2):95-104.
9. Mueller HG. Speech audiometry and hearing aid fittings: going steady or casual acquaintances? Hear J. 2001; 54(10):19-29. doi: 10.1097/01.HJ.0000294535.51460.0c
10. Heidari M, Mahdavi ME, Heidari F, Akbarzadeh Baghban A. Auditory recognition of Persian digits in multi-talker babble noise: a preliminary study. Aud Vestib Res. 2015;24(3):134-40.
11. Ramkissoon I, Proctor A, Lansing CR, Bilger RC. Digit speech recognition thresholds (SRT) for non-native speakers of English. Am J Audiol. 2002;11(1):23-8. doi: 10.1044/1059-0889(2002/005)
12. Spreen O, Borkowski JG, Benton AL. Auditory

word recognition as a function of meaningfulness, abstractness and phonetic structure. J Verbal Learning Verbal Behav. 1967;6(1):101-4. doi: 10.1016/S0022-5371(67)80057-4
13. Spreen O, Borkowski JG, Gordon AM. Effects of abstractness, meaningfulness, and phonetic structure on auditory recognition of nouns. J Speech Lang Hear Res. 1966;9(4):619-25. doi: 10.1044/jshr.0904.619
14. Jones D, Spreen O. The effects of meaningfulness and abstractness of word recognition in educable retarded children. Am J Ment Defic. 1967;71(6):987-9.
15. Mefferd RB Jr, Lester JW, Wieland BA, Falconer GA, Pokorny AD. Influence of distraction on the reproduction of spoken words by schizophrenics. J Nerv Ment Dis. 1969;149(6):504-9. doi: 10.1097/00005053-196912000-00006
16. Wilson RH, Cates WB. A comparison of two word-recognition tasks in multitalker babble: Speech Recognition in Noise Test (SPRINT) and Words-in-Noise Test (WIN). J Am Acad Audiol. 2008;19(7):548-56. doi: 10.3766/jaaa.19.7.4
17. Mahdavi ME, Pourbakht A, Parand A, Jalaie S, Rezaeian M, Moradiju E. Auditory recognition of words and digits in multitalker babble in learning-disabled children with dichotic listening deficit. Iran Red Crescent Med J. 2017;19(4):e42817. doi: 10.5812/ircmj.42817.
18. Wilson RH, Carnell CS, Cleghorn AL. The Words-in-Noise (WIN) test with multitalker babble and speech-spectrum noise maskers. J Am Acad Audiol. 2007; 18(6):522-9. doi: 10.3766/jaaa.18.6.7