

Journal of Rehabilitation Sciences and Research

Journal Home Page: jrsr.sums.ac.ir

Original Article

Developmentand Psychometric Evaluation of Persian version of the Quick Speech in Noise Test in Persian Speaking 18-25 Years Old Normal Adults

Abdollah Moossavi¹, Mohanna Javanbakht², Hamideh Arbab Sarjoo^{2*}, Enayatollah Bakhshi³, Behrooz Mahmoodi Bakhtiari⁴, Yoones Lotfi²

¹Depament of Ear, Nose and throat, School of Medicine, Iran University of Medical Sciences, Tehran, Iran ²Department of Audiology, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran ³Departmentof Epidemiology, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran ⁴College of Fine Arts, University of Tehran, Tehran, Iran

ARTICLE INFO

Article History: Received: 28/812016 Revised: 15/10/2016 Accepted: 17/10/2016

Keywords: Quick speech-in-noise test SNR-50 Reliability List equivalency

ABSTRACT

Background: One of the common problems of those who refer to audiology clinics, is difficulty of hearing in noisy conditions. Different tests have been developed for diagnosis and quantification of reduced ability of speech perception in noise and the Quick Speech-in-Noise test is one of the most appropriate of them. The goal of the present study was to develop and validate the Persian version of the Quick Speech-in-Noise Test in normal, Persian speaking, 18-25 year-old participants.

Methods: This study had two main stages: first, developing the test materials and determining the content validity and the second, determining the equivalency of the test material lists, and assessing the reliability of them.

In the first stage, sentences having content validity were developed and recorded. In the second stage, in order to determine list equivalency, the lists were presented to 60 Persian-speaking (gender balanced), 18-25 year-old students and the average SNR-50 for each list was calculated using the Spearman-Kärber formula. In order to determine the reliability of the test, the test was again administered to15of the primary participants, two weeks later.

Results: among80 sentences developed, 60 received enough credits from audiologists and speech therapists. Average SNR-50 in the Persian language was calculated as -0.25 dB. Thirty six content valid sentences, according to special criteria, were divided into six lists each containing 6sentences. The SNR-50 for each list revealed that lists number 2,3, 4and 5 were equivalent. Examination of test-retest reliability indicated that lists number 1,2,3,4 and 6 were reliable (P<0.05). **Conclusion**: Present study was developing three equivalent and reliable lists (number 2,3,4) for the Persian version of the Q-SIN test which are useful for diagnosing and selecting the best rehabilitation method for people with reduced speech perception in noise.

2016© The Authors. Published by JRSR. All rights reserved.

Introduction

During many daily activities, speech is heard in noisy

E-mail: arbab.ha8888@gmail.com

conditions. In such situations, speech perception requires separating and differentiating the competing sounds being heard. Normal people are able to separate the signal i.e. speech from competitive background noise. Some people have difficulty in challenging hearing situations, like a noisy or echoic environment and face some problem in speech perception. Reduced ability to perceive speech-

^{*}Corresponding author: Hamideh Arbab Sarjoo, The Side of Ghods Mosque, 81 Mostafa Khomeini Avenue, Zahedan, Sistan va Baluchestan, Iran. **Tel:** +98 915 6342461

in-noise can influence a person's social communications resulting decrease quality of life, preventing them from leading an independent life [1,2].

many of hearing aid users, the elderlies, Those with central auditory processing disorders (CAPD) and children with attention deficit hyperactivity disorder (ADHD) may experience reduced ability to perceive speech-in-noise [3,4]. Identifying those who experience a reduction in their speech-in-noise perception, and quantifying the level of this reduction, will make a significant contribution to early diagnosis, provision of hearing rehabilitation services, consultations on available technologies, hearing aid prescription and fitting; so it will be helpful for both the patients and therapists (5,6). Speech perception in noise tests, as part of behavioral tests, can be used for hearing impairment diagnosis [7], especially in diagnosis of CAPD and prescription the hearing aid, it can be useful in providing proper rehabilitation programs [6] and help patients to have realistic expectations [8].

The commonly used tests for evaluation of speech perception in noise include the Consonant-Vowel in Noise test, Word-In-Noise (WIN) test, the Hearing-In-Noise Test (HINT), the Speech-In-Noise (SIN) test, and the Quick Speech-In-Noise (Q-SIN) test. The last one, i.e. the Q-SIN test, is the best and most commonly used tests for this goal [8,10]. Comparing the patient's performance with normal people, this test uses the signal-to-noise ratio loss (SNR loss) to assess speech perception in noise. SNR loss is the increased signal-to-noise ratio required by an individual with hearing impairment to correctly repeat 50% of the words in noise, as compared to normal individuals. It is calculated by subtracting the patient's SNR-50 from the average SNR-50 for normal people in any specific language.SNR50 which refers to a SNR that a person can correctly repeats 50% of key words, was used. The evident superiority of the Q-SIN test to other speech perception in noise tests is using test materials in the level of sentence to quantify the level of reduction in speech perception in noise similar to the natural environment condition [8], and being more sensitive to speech recognition in noise [9]. Another advantage of the Q-SIN test is the possibility to estimate the level of reduction in speech perception in noise for both ears in about one or two minutes [3].

The test materials of the Q-SIN are sentences presented in four-talker babble noise [9]. In its original English version, every list of the test has six sentences, and every sentence includes five key words. The key words have been selected in such a way that can hardly be guessed from the context of the sentence [13]. Each sentence is presented at one of SNRs: 0, +5, +10, +15, +20, and +25 dB. The overall SNR-50 is calculated for every list using the Spearman-Karber Formula [9]. According to this formula, SNR-50 for each list is calculated by subtracting 27.5(based on Spearman-Karber Formula for QSIN test) from the total number of correctly expressed words of every list. The SNR-50 was calculated separately for sentence and word using a particular method, which will be presented later on.

Khalili et al, examined the equivalency of four lists of the Persian version of the Q-SIN test; finding two of their lists being equivalent and reliable. They found that the test accuracy increased by increased the number of the lists [12]. However, no paper has been published on the details of the development of their lists, and the material properties of test was not available itself. Shayanmehr et al, conducted the "Development, validity and reliability of Persian quick speech in noise test with steady noise". They calculated the average SNR-50 in Persian-speaking, normal individuals as [-4] dB [11], which is significantly lower than the value calculated for the English language (2+) [13]. The key words used in this previous study were only two-syllabic words with the high cohort point [11]. Such method limits the selection of key words, and does not represent the real speech situation. In the present study, we tried to increase the number of the lists in the Persian version based on the framework of the original version in English. The psychometric properties of the Q-SIN test were assessed in normal Persian-speaking, 18-25 year-old participants, addressing some of the limitations in the previous studies.

Methods

This study has two main stages. 1-Test development and determine the content validity 2. Equivalency and reliability of lists.

The first stage included: A) selecting the key words, creating the sentences, and examining the content validity; B) recording and organizing the test materials; and C) organizing the lists.

The second stage included: A) evaluating the scores of the lists in normal, Persian-speaking, 18-25 year-old participants and determining the equivalent lists; and then B) determining the reliability of the test.

Stage 1: Test Development and Determine the Content Validity

A: Selecting the Key Words, Creating the Sentences, and Examining the Content Validity.

Based on previous studies for development of the Q-SIN test and the sample sentences provided in one study (including 360 sentences in English), the properties of the key words and sentences used in the test were determined [14]. The sentences advised by the Institute of Electrical and Electronic Engineers' (IEEE) [19] to be used in the lists of the Q-SIN test, while syntactically correct, must have difficult semantic structure in such a way that the key words cannot be guessed from the context of the sentence.

About 800 familiar words were selected from the Moein Persian dictionary to create sentences to construct the Persian lists of sentences. In order to make the sentences phonetically difficult, key words were selected in such a way that each sentence would contain pairs of voiced and voiceless consonants (such as /b/ and /p/ or /k/ and /g/), or a repetition of a consonant or a vowel according to linguistics adviser in this Persian version. For example, in the sample sentence "-/behtarin pargâretân râ be kelâse sevvom biyâvarid/ (Bring your best compass to class 3)." there are pairs of consonants (/b/ and /p/, and /k/ and /g/), and also the consonants /s/ and /b/ are repeated. A

total of 80 sentences with proper syntactic, semantic, and phonological properties were created, and approved by the linguistics adviser who analyzed them in terms of the above criteria, and also for phonological coherence. Each list contained six sentences, and five to nine words in every sentence to be used for 18-25 years old Persian-speaking individuals. Five words were selected as key words. Key words were a variety of monosyllabic to polysyllabic words, with different roles in the sentence, including subject, object, predicate, preposition, or verb. Then, the sentences were presented to 10 audiologists and speech therapists in the form a questionnaire to be analyzed for syntactic, semantic, and phonological properties. This questionnaire provided three options for every sentence, including 1) appropriate; 2) relatively appropriate; and 3) not appropriate with attention to test's criteria. Experts' suggestions were asked about any possible change. Then the content validity ratio (CVR) of the sentences were calculated according to Lawshe method.

B: Recording and Organization of the Test Materials

The developed sentences were recorded by a male talker with a familiar and expressive voice continuously, in IRIB (the Islamic Republic of Iran Broadcasting) center. In the Persian version of test, the babble noise was provided by four talkers (two men and two women) and intensity level of speech signal (sentences) was kept 70 dB HL (for normal hearers according to the original version of test). Using WavePad Sound Editor 5.03 NCH software, consecutive sentences were separated and their intensity level was normalized. To select the most proper sentences, according to the original version of test, SNR-50 criterion, which refers to a SNR that a person can correctly repeats 50% of key words, was used. Using the MATLAB software the approved sentences (60 sentences), were designed at SNRs: -6,-4, -2, 0, +2, and +4dB. Sound Forge 10 software was used to implement four-second interval between consecutive sentences to record what was repeated by participant. Each sentence was presented from the most difficult SNR (-6 dB) to the most comfortable one (+4 dB). The SNR was increased in 2 dB steps until reaching a SNR in which the person could repeat three of the key words in the sentence. This intensity level was considered as the threshold of SNR-50 for the sentence. To measure the threshold of SNR-50 of the sentences, they were presented with definition background noise using SHL-3100 PHILIPS calibrated headphone to 14 Persian speaking, 18-25 year-old normal hearing participants. They did not participate in other stages of the study, due to their familiarity with the test materials. To prevent the effect of ear advantage in the process of determining the SNR-50 threshold in 14 participants, the first 30 sentences were presented to the right ear, and the second 30 sentences were presented to the left ear for half of the participants and vice versa for the remainder.

Another criterion considered by the developers of the original version of the test, was the threshold for each key word in noise. To calculate this, we used the results of threshold determining of the sentences in noise; five participants from the 14 participants were randomly

C: Organizing the Lists

For the equivalency of final lists, the following three inclusion criteria were used to select each sentence:

1. Since in this study found the calculated threshold of speech perception in noise to be -0.25 dB in Persian language, those sentences within 2 dB above or below of this measure were used to develop the lists. Considering this criterion (an important criterion in homogenizing the sentences to make the final lists equivalent [13]), only 37 sentences out of 60 were kept.

2. Since the similarity of the threshold of key words would have resulted in a very simple or very difficult sentence, those sentences in which the SNR-50 for each of the five key words varied by less than or equal to 2 dB were discarded [13]. According to this, we had 36 sentences with appropriate quality.

3. In order to make the lists as equivalent as possible, the sentences were divided into three groups, based on theirSNR-50 value of each sentence. Sentences with SNR-50 value between -2 to -1dB ($-2 \le X \le -1$)considered as easy, those with SNR-50 value between -1 to +1dB ($-1 \le X \le +1$) considered as moderately difficult, and those between +1 to +2dB (+ $1 \le X \le +2$) were considered as difficult ones. The third criterion was based on the opinion of the authors of present study to make the Persian lists more equivalent. In each list the occurrence of three types (easy, moderately difficult, and difficult) kept balanced.

Thirty six proper sentences were arranged in 6 lists, each containing randomly chosen 6 sentences in the Persian version of the test. Then, each chosen sentence was adjusted to be presented at one SNRs: 0, +5, +10, +15,+20, and +25 dB according to their difficulty by using the MATLAB software and the recorded lists were presented to the target group.

Stage 2: Equivalency and Reliability of Lists

60 individuals from the students of the Zahedan University of Medical Sciences, Zahedan, Iran (30 men and 30 women), aged 18-25 years old (men's average age: 22.16 ± 1.11 ; women's average age: 21.56 ± 1.50), participated in this stage.

The inclusion criteria were:

Informed consent, normal and symmetrical hearing in both ears (pure-tone thresholds better than 20 dB HL at frequencies 500, 1000, 2000, and 4000 Hz for both ears, and less than 10 dB difference between the two ears), Persian-speaking, healthy outer ear canal and tympanic membrane, normal performance of the middle ear in both ears (determined by Otoscopy and Tympanometry), righthandedness in manual works(assessed by the Edinburgh Handedness Inventory), no psychological or neurological disorders(such as epilepsy or brain injury), no history of brain surgery; and no speech problems. Unwillingness to continue the study at any stage and/or losing any of the inclusion criteria (between two week for test-retest reliability) were considered as exclusion criteria.

A: Determining the Equivalent Lists

To evaluate equivalency, in the first step the recorded lists were presented to participants, after a training with two practice lists.

First, participants signed the consent forms. Then, a personal information questionnaire and the Edinburgh Handedness Inventory (to examine right-handedness) were completed by the examiner. In the next step, they who met the inclusion criteria underwent otoscopy (using a Riester otoscope) and immittance audiometry (using a Madsen Zodiac 901 audiometer), in order to analyze the performance of their middle and outer ear. Then, puretone audiometry was conducted through air conduction and at octave frequencies from 500 to 4000 Hz (using a Madsen Midimate 622 audiometer). The Q-SIN test was administered by using an Asus laptop model X44H and PHILIPS calibrated headphones. Test was carried out in an acoustic room-At first two practice lists were

administered to make them acquainted with the test procedure. The following explanation was given to each participant: "sentences expressed by the male talker along with background babble will be sent to your one of ears. Repeat each sentence you hear with the male talker voice." Then, the main lists, were monaurally presented to the participant's ears using the Sound Forge 10 software. The Mann–Whitney U test was used to compare average SNR-50 between the right and left ears and Wilcoxon signedrank test for comparing the results of men and women. Equivalency of List was determined by comparing the average SNR-50 between the lists and examination of standard deviation according to the original version of test.

B: Examining the Reliability of the Q-SIN Test

In order to examine test-retest reliability, two weeks after the first administration, the Q-SIN test was again administered to 15 persons out of 60 participants, and the interclass correlation (ICC) was calculated using the SPSS software v.19.

Figure 1 diagram, represents the different stages of



Figure 1: Diagram of different stages of developing and determining content validity, equivalency and reliability of the Q-SIN test.

developing and determining equivalency and reliability of the Q-SIN test in normal, Persian-speaking, 18-25yearold participants.

Results

The Results of the Firststage

The result of Stage 1: Part A: selecting the key words, creating the sentences, and examining the content validity

Those sentences reaching the score above 0.62 (CVR>0.62) according to the number of experts, were approved, being 60 out of primary 80 sentences.

The result of Stage 1: Part B: recording and organization of the test materials

After presenting each sentence at different SNR, the average SNR-50 of each sentence was calculated for both ears, and no significant differences was found between them (P>0.05). So each sentence received one point regardless of the side of test ear in all 14 participants. Other finding was that the sentences had different SNR-50 thresholds from -4 to +4 dB. The threshold for Key words in noise also were different ranging from -6 to +6. For example, for the sentence -/bahman ?ânruzkafšgerânašrâšost. /(On that day, Bahman washed his expensive shoes),"the sentenceSNR-50 threshold was 1.29 dB, and the key words in noise thresholds were 0.8, 3.6, 0.8, 1.6, and 0 dB respectively. This indicates that the threshold of key words were different between them and also from the sentence. These results were then used to create the main lists. In order to obtain the mean SNR value required by Persian speaking individuals to understand 50% of speech in noise found to be -0.25 dB.

The result of Stagel: Part c: organizing the lists

Based on the aforementioned criteria of this part, 36 sentences out of 60 were selected and randomly placed into six lists. The average SNR-50 for the lists were compared between the right and left ears, and no significant difference was found (P>0.05). The average SNR-50 was also compared between men and women, and again no significant difference was observed (P>0.05).

The Results of the Second Stage

The result of stage 2: Part A: determining the equivalent lists

In order to examine list equivalency, the average SNR-50 among six lists was calculated and found to be -0.95 dB and according to the original version of the test [13] this value was compared to the average SNR-50 of every list. Table 1 shows the difference between the average SNR-50 for each list and the average SNR-50 among six lists. Those lists with an average SNR-50 1 dB different from the average SNR-50 among six lists and with a standard deviation less than 1, were considered as equivalent. Therefore, the lists number 2, 3, 4, and 5were found to be equivalent.

The result of stage 2: Part B: examining the reliability of the Q-SIN test

The interclass correlation (ICC) was calculated to examine test-retest reliability; the results are presented in Table 2. The lists number 1, 2, 3, 4, and 6 were highly reliable (P<0.05).

Discussion

The average SNR-50 for speech perception in noise in the English language is about +2 dB [13, 14]. The slight difference between the SNR-50 in our study0.25-() and the original test can be due to higher enrichment and redundancy in the Persian language [11]. In contrast to original version by Killion et al, our study had fewer primary sentences, less final equivalency and reliability. The reason is that Killion et al. used phonetically balanced sentences (developed by Institute of Electrical and Electronics Engineers) which were previously selected in the development of SIN test. In the present research, however, we had to develop our own Persian key words and sentences in a limited time.

However, in the present study, we tried to follow the criteria of the original version of the test in the Persian version. Therefore, familiar and commonly used words (from monosyllabic to polysyllabic) in the Persian language were selected, and sentences were constructed in such a way that the key words would not be recognizable from the context of the sentence and

List number	Average SNR-50 for each	Average SNR-50 among six lists	Mean difference	Standard deviation
List 1	0.21	-0.95	1.16	1.2
List 2	-1.30	-0.95	0.35	0.57
List 3	-1.66	-0.95	0.71	0.69
List 4	-1.38	-0.95	0.43	0.9
List 5	-1.11	-0.95	0.5	0.76
List 6	-0.65	-0.95	0.3	1.10

fable 1: The difference between average SNR-50 fc	r each list and average SNR-50	among six lists in decibels
---	--------------------------------	-----------------------------

Table 2: The results of test-retest reliability examination					
List number	ICC	P value			
List 1	0.6	0.04			
List 2	0.83	0.001			
List 3	0.92	0.000			
List 4	0.70	0.012			
List 5	0.28	0.26			
List 6	0.75	0.005			

sentences to have limited SNR-50 threshold and also the threshold of key words perception in noise to be within the defined area. In the previous study by Shayanmehr et al on development of a Persian version of O-SIN test, some of these criteria (such as using from monosyllabic to polysyllabic words as keywords) were not considered [11]. Several differences between two studies in terms of test materials are noticeable, including talker's gender, criteria for selecting words (just two syllabic key words were selected that is in contrast with the primary goal of creating the Q-SIN test which requires different types of key words to represent the real world situation), sentences and presence or absence of noise during intervals, resulting in difference of SNR-50sin their study to be 11])4-(] and -0.25 dB in our study, the latter being closer to the mean value reported in other languages and the original version of the test in English13](2+(]. These differences in test construction, and also comparing the SNR-50 between the two studies, lead us to the conclusion that the test materials used in the present study were closer to the original version of the test in English regarding the phonological elements or possibility of guessing key words from the context of the sentence.

No significant difference was found between average SNR-50 in men and women, this finding is consistent with the results of the previous study conducted by Shayanmehr et al, and. Calaise et al. examined the effect of gender on the results of a speech-in-noise test, in 49 elderly people. They used filtered white noise at high and low frequency ranges, in order to simulate speech noise. SNR at +5 dB was used in this study; the results showed no significant difference between men and women which is in consistent with the findings of the present study [11,15].

Conclusion

The final result of the present study was developing three equivalent and reliable lists for the Persian version of the Q-SIN test which are useful for diagnosing people with reduced speech perception in noise, and also for quantifying the level of reduction in speech perception in noise for selecting the best rehabilitation method for an individual patient. At the present the Q-SIN test is a fast and the only test in Persian that evaluates patients in a condition similar to daily life. In addition, this test is useful in hearing aid consultation and prescription. Using this test, it is possible to assess speech perception in noise ability in acoustic fields, with or without hearing aid, and to examine the applicability of directional microphones. This test can be used to control and examine the effectiveness of rehabilitation programs, and, it can be used along with other behavioral tests in diagnosis of CAPD.

Acknowledgment

This article was extracted from a master's thesis, we thank the Department of Audiology of the University of Social Welfare and Rehabilitation Sciences. Part of the present study was done in the School of Rehabilitation, Zahedan University of Medical Sciences. Therefore, we sincerely thank all the authorities and students of this school for their kind cooperation in this project.

Conflict of interest: None declared.

References

- 1. Gussekloo, J., et al., Auditory rehabilitation of older people from the general population--the Leiden 85-plus study. British Journal of General Practice, 2003. 53(492): p. 536-540.
- Gopinath, B., et al., Hearing-impaired adults are at increased riskof experiencing emotional distress and social engagement restrictions five years later. Age and Ageing, 2012. 41(5): p. 618-623.
- Kochkin, S., MarkeTrak V." Why my hearing aids are in the drawer": The consumers' perspective. The Hearing Journal, 2000. 53(2): p. 34-36.
- Geffner, D., central auditory processing disorder definition ,description, and behaviors, in auditory processing disorder assessment management ,and treatmentplural publishing united states of America by McNaughton and Gunn.2007. p. 25-30
- Cord, M.T., et al., Performance of directional microphone hearing aids in everyday life. Journal of the American Academy of Audiology, 2002. 13(6): p. 295-307.
- Geffner, d., management strategyies, in auditory processing disorder assessment management ,and treatment 2007, plural publishing united states of America by McNaughton and Gunn. p. 207-215.
- Marni, L.J., J.B. Teri, and B. Cassie, Audiology assessment in auditory processing disorder assessment management ,and treatment 2007, plural publishing united states of America by McNaughton and Gunn. p. 76-84.
- Taylor, B., Speech-in-noise tests: How and why to include them in your basic test battery. The Hearing Journal, 2003. 56(1): p. 40-42.
- Wilson, R.H., R.A. McArdle, and S.L. Smith, An evaluation of the BKB-SIN, HINT, QuickSIN, and WIN materials on listeners with normal hearing and listeners with hearing loss. Journal of Speech, Language, and Hearing Research, 2007. 50(4): p. 844-856.
- Walden, T.C. and B.E. Walden, Predicting success with hearing aids in everyday living. Journal of the American Academy of Audiology, 2004. 15(5): p. 342-352.
- 11. Shayanmehr, S., et al., Development, validity and reliability of Persian quick speech in noise test with steady noise. Auditory and Vestibular Research, 2015. 24(4): p. 234-244.
- KhaliliM, FattahiJ,HajiabolhassanF,TahaaieAK,JalaieSH.,Testretest reliability and equivalency of the persian Quick Speech in Noise Test.in persian. Modera Rehabilitation, 2010. (3): p.16-21.
- Killion, M.C., et al., Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. The Journal of the Acoustical Society of America, 2004. 116(4): p. 2395-2405
- Bentler, R.A., List equivalency and test-retest reliability of the speech in noise test. American Journalof Audiology, 2000. 9(2): p. 84-100.
- Calais, Lucila Leal, Russo, Ieda Chaves Pacheco, Borges, Alda Christina Lopes de Carvalho., Performance of elderly in a speech in noise test. Journal of Pró-Fono Revista de Atualização Científica. 2008. 20(3): p. 147-152
- McArdle, R.A. and R.H. Wilson, Homogeneity of the 18 QuickSIN[™] lists. Journal of the American Academy of Audiology, 2006. 17(3): p. 157-167.
- Wilson, R. and R. Mcardle, speech in noise measures as necessary components of routine audiologic evaluations and auditory processing disorder evaluations, in Controversies in Central Auditory Processing Disorder (CAPD). 2009, Plural Publishing, Incorporated. p. 151-168.
- Ringleb, S., et al. Development and PreliminaryTesting of a Rehabilitation Game to Improve Listening in Background Noise. in ASME 2013 Summer Bioengineering Conference. 2013. American Society of Mechanical Engineers.
- Fikret-Pasa, Selda. The effect of compression ratio on speech intelligibility and quality. The Journal of the Acoustical Society of America.1994.2992