Aud Vestib Res (2019);28(3):173-181.

RESEARCH ARTICLE

Development of the Persian version of high-frequency emphasis quick speech in noise

Jamileh Fatahi[1*], Maryam Amiri Jahromi[1], Fahimeh Hajiabolhassan[1], Amirsalar Jafarpisheh[2], Nariman Rahbar[3], Elham Faghihzadeh[4]

[1]- Department of Audiology, School of Rehabilitation, Tehran University of Medical Sciences, Tehran, Iran

[2]- Department of Ergonomics, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran

[3]- Department of Audiology, School of Rehabilitation Sciences, Iran University of Medical Sciences, Tehran, Iran

[4]- Department of Biostatistics, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

Received: 11 Nov 2018, Revised: 19 Jan 2019, Accepted: 21 Jan 2019, Published: 15 Jul 2019

Abstract

Background and Aim: The quick speech in noise **(**Q-SIN) test shows the difficulty of speech perception in noise by specifying signal to noise ratio (SNR) loss. Although the Persian version of Q-SIN has been already constructed, the high-frequency emphasis version of this test is not available. The present study aimed to construct six lists with high-frequency emphasis and implement it.

Methods: We are going to prepare a high-frequency emphasis version of Q-SIN and then test it on a small sample. First, researchers designed the relevant sentences; then experts examined their content and face validity. According to the criteria for developing the Q-SIN test, six lists with high-frequency emphasis were prepared. The test was examined on 26 (13 male and 13 female), 18−35 years old individuals with normal hearing. To determine the test reliability, it was re-administered three weeks later with the same conditions.

Results: Of 76 sentences prepared, 36 sentences received enough credit after determination of their content and face validity. These 36 sentences were used to make 6 lists. The mean value of SNR50 in the Persian language was obtained -4 dB. The mean values of SNR loss in 6 lists were -1.65, -1.8, -2.23, -1.61, -2.38 and -2.07. The results showed equivalency of lists 1, 2, 3, 4, and 6. Examination of test-retest reliability indicated that all lists except the list 2were reliable.

Conclusion: The lists of 1, 3, 4, and 6 are reliable and equivalent and can be used in clinical application.

Keywords: Equivalency; normal hearing; quick speech-in-noise test; reliability; signal to noise ratio loss

Citation: Fatahi J, Amir Jahromi M, Hajiabolhassan F, Jafarpisheh A, Rahbar N, Faghihzadeh E. Development of the Persian version of high-frequency emphasis quick speech in noise. Aud Vestib Res. 2019;28(3):173-181.

Introduction

Communication in the presence of background noise is a crucial skill, investigated by many studies. Communication in human life is vital and mostly performed by speech. Because people

* **Corresponding author:** Department of Audiology, School of Rehabilitation, Tehran University of Medical Sciences, Piche-Shemiran, Enghelab Ave., Tehran, 1148965141, Iran. Tel: 009821-77533939, E-mail: jfatahi@tums.ac.ir

sometimes talk in the presence of the competing background noise, listeners must reinforce their cognitive, neural and sensory sources to control interfering noise and achieve speech perception and successful communication [1]. Communication in a noisy environment is a challenging task for both normal-hearing and hearing-impaired subjects [2].

One of the main difficulties in hearing-impaired subjects is speech perception in noise [3,4]. Outer hair cell destruction reduces hearing sensitivity to soft sounds, but inner hair cells deficit lowers hearing resolution and consequently results in difficulty in the listening in noise [5,6]. Therefore, hearing-impaired subjects have low satisfaction with their hearing aids in these environments [7].

The conventional audiometry, which is based on pure tone threshold seeking, cannot estimate auditory system performance in the presence of wideband signals and speech perception in noise [8]. There are many tests for the evaluation of speech perception in noise, including speech in noise (SIN), the word in noise (WIN), hearing in noise test (HINT), and quick speech in noise test (Q-SIN) [9].

Among these tests, Q-SIN is more acceptable and effective and enables audiologists to identify speech perception difficulties in noise in the listeners [10].

This test was developed in the Etymotic research Inc. and commercially available since 2001 [11]. In this test, the signal to noise ratio (SNR) loss is used as a diagnostic index [8]. SNR loss is defined as the amount of SNR increment necessary to achieve correct word recognition in 50% of times compared to normal subjects. Conventional audiometry cannot measure or predict patients' speech perception in noise, and without knowing SNR loss in hearing-impaired subjects, prediction of their improvement following hearing aid prescription is impossible [8].

Q-SIN is an open-set sentence recognition test, in which multi-talker babble noise and target speech are presented simultaneously in a competing condition. Each list includes 6 sentences, and each sentence consists of 5 keywords, and every key word that has been repeated correctly is being scored. Test SNRs are 0, 10, 15, 20 and 25 dB and SNR reduces from sentence 1 to 6 via increasing noise level. The test presentation level is 70 dB HL for normal hearing subjects and subjects with hearing loss ≤ 45 dB HL. For subjects with hearing loss > 45 dB HL, the test is conducted at the most comfortable level (MCL) [8]. English Q-SIN consisted of 18 six-sentence lists, 12 lists of 30 dB with emphasis on high frequency and 12 low-pass filtered lists [10].

Khalili et al. developed the Persian version of Q-SIN in Iran and studied its reliability and equivalency in normal hearing subjects and the elderly [12]. Shayanmehr et al. developed five new lists for the Q-SIN test [13]. Haniloo et al. studied these new lists on normal-hearing as well as sensory-hearing-loss subjects [14]. Moosavi et al. studied psychometrics of the Persian Q-SIN test and introduced four equivalent lists out of six available lists [15].

The Q-SIN test aims to evaluate speech perception in the presence of the multi-talker babble noise. Although the original Q-SIN test has 12 high-frequency lists, there has been no high-frequency list for Q-SIN in the Persian version. The present study was a test development attempt to prepare six Persian high-frequency lists. Then we were going to test it on a small sample and determine its validity and reliability on normal subjects (age range of 18−35 years old).

Methods

We aimed to develop the new Q-SIN test in three main steps: developing the test, testing on the samples, and retesting it.

The first step comprises three minor steps: designing sentences and determining the test power, recording sentences and preparation of babble noise in studio, and editing every list and combining them with babble noise and filtering lists.

Designing sentence and determining the test power

In this test, we should use frequently-spoken words, i.e. the words used in everyday speaking. For this purpose, we used high-frequency

Persian words prepared by Assi et al. Key words used in the sentence must be at most three syllables long [10] and the sentence should not be more than 7−8 words long. Because long sentences are problematic in terms of auditory memory [13]. Using rare or foreign words is not permissible, and the sentences should be unpredictable. As utilizing the monosyllable words reduces the predictability of the sentence, we used the monosyllable words as much as possible in this study. According to the mentioned features, 76 sentences were prepared.

After designing the sentences, their content validity was checked. 15 audiologist, speech therapists, and linguistics experts checked the content validity of test sentences. They were all experts in the evaluation of the central auditory system. They scored the sentences based on a 5-point Likert-type scale and calculated with a quantitative Lawshe's method concerning grammatical correctness, unpredictability, and acceptability of the sentences. The content validity ratio (CVR) index was used for each sentence in this method. The face validity is often examined by the test participants. Since this test is designed for young people with a moderate level of education, the face validity of the sentences was evaluated by several young people with a high-school diploma or bachelor’s degree. We used the confirmed sentences for developing the final lists [13].

Recording sentences and preparation of noise in the studio

In the next step, the approved sentences were recorded by female speakers in the virtual college studio. The speakers had the required conditions such as correct pronunciation and similarity to the regular conversational speech [13]. Afterward, the recorded sentences were offered to three skilled audiologists in devising a central auditory test for approval. The noise level of the room was monitored (2230 Bruel and Kjaer of SLM) to record the speech signal and the background noise below 25 dB during the test. The recording was performed in the studio by the Universal Audio Solo 0/610 amplifier of Mac System (Universal Audio, Inc., USA) voice card of Mbox Pro, and Neumann TLM49 microphone. The voice was played at 44.1 kHz and 16-bit depth with 5 seconds time interval between sentences so that the listener had enough time to repeat each sentence. The voice was recorded by SBW-06D2X-U’s CD ROM of Asus from Taiwan at 8x speed on a compact disk. In the next step, the four-talker babble noise was prepared by one man and three women speakers. Expert speakers were not necessary to read and record the noise. To record the noise, a microphone was placed in the middle, and four speakers on the four sides; then they all started reading four different texts at the same time. After the recording, sequential sentences were separated by Sound Wave Pad software, and their intensity was normalized.

Editing all lists and combining them with babble noise

For editing test material, we need indicators to compare sentences. This indicator is the SNR50 for each sentence. That is the ratio of signal to a noise level that a person can recall and say at least 50% of the key words of a sentence [10]. The confirmed sentences in the previous step were defined +4, +2, 0, -2, -4, -6 dB in signal to noise ratios (SNR) for determining this index. Sound Forge 10 software was used to implement 5 seconds interval between sentences, because the listener should have enough time to repeat the sentences. The presentation started in -6 dB SNR and continued with increasing SNR in 2-dB step to determine the level of SNR that a person can express at least three words out of the key words of the test sentence. This intensity level was considered as the SNR50 threshold of each sentence. The defined sentences in SNRs were presented to 10 people at their most comfortable loudness (50 dB) by Philips calibrate headphone to determine the SNR50 threshold of each sentence and select the appropriate sentences and final examination listing (Because at this stage the participant got familiar with the test subject matter, they were not included in the next stages of the test). You can add or subtract to apply the SNR50 amount in any language, or you can change the intensity of the noise in the same amount on the test compact disk [16].

The next step was the even distribution of sentences in the lists so that the lists were matched by the length and difficulty of their sentences. After that, sentences were applied with 30 dB amplification and up to 2.5 kHz frequency into the MP3 Audio Editor, and the lists were made with 30 dB high-frequency emphasis [10]. Every list was combined by four talker's babble individually in a 5-dB step by MATLAB Software. Each list consists of six sentences and each sentence has 5 key words. If each key word is repeated correctly, it scores one point. Finally, the number of correct words was recorded by the examiner and SNR loss for each list was obtained by the following formula:

SNR loss = 27.5 – (the total number of correct words – SNR50).

The second main step was performing the test. The convenience sampling method was used to recruit a total of 26 normal-hearing volunteers aged 18 to 35 years. The test was conducted at the Audiology Clinic of School of Rehabilitation of Tehran University of Medical Sciences.

After obtaining the participants' consent, we explained the test procedure to them, and their personal history was taken by the audiologist. The inclusion criteria were being right-handed; lacking any diseases in the conduction ear system, otologic, neurological problems; or hearing problems in the noise-induced hearing loss (NIHL); having normal hearing (the pure tone average threshold less than 25 dB) [17], and not having a history of head injury. The exclusion criteria were no attention or poor cooperation during the test and unwillingness to continue the test.

The test was performed via circumaural headphone (SHL3100MGY headphone by Philips Company and PCG-61211, Sony laptop, China). The computer sound level was set at 70 dB HL (B&K 2235 sound level meter). First, we calibrated the system at a frequency of 1000 Hz by Cool Edit Pro 2.1, which was used as a measuring reference for volume control of the computer. Before starting the test, the calibration was performed by using a 1 kHz calibration tone to set up the laptop, and the zero should be seen on the VU-meter [10].

The Q-SIN is the recognition test of a sentence that runs as an open set, and the test can perform via headphone and sound field [10]. Because the reflection of the sound from the test space levels can affect the speech comprehension threshold, the design and standardization of Q-SIN test must be basically done by the headphone.

Before starting the test, we explain the procedure to the samples, i.e. the samples hear the speaker who reads some sentences with a background of a babble (humming) noise. Then they should repeat any sentence they hear. First, as a practice, a list was presented to the samples. The test was performed at a level of 70 dB HL [10]. The samples should respond verbally, and the examiner wrote their responses in a sheet. Encoded lists were randomly presented for everyone to control the order effect [18]. Each list had six sentences, and each sentence had five key words, and if each keyword were repeated correctly, the sample would award one point. Finally, the number of correct words was recorded for each sample, and the examiner calculated the SNR loss for each list by using the formula mentioned above.

The third main step was performing the re-test. For the final reliability check, all the lists of Q-SIN test were evaluated in the same way as the first run, three weeks later and the SNR loss was calculated.

For analyzing the obtained data, descriptive statistics were used to determine mean values. Also, the Kolmogorov-Smirnov test was used to determine the normal distribution of the collected data. Since the distribution of data was not normal to study the effect of gender, we used the Mann-Whitney test and to evaluate the reliability, two methods were done including the Wilcoxon and the Spearman correlation coefficients. To check the equivalency of SNR loss score in normal people, we used the Friedman non-parametric test and Wilcoxon test corrected by Bonferroni (because of abnormality of distribution in the six lists scored). The obtained data were analyzed in SPSS 19, and the significance level was set at 0.05.

Table 1. Mean (standard deviation) SNR loss of six lists in total population and comparison of the scores in men and women with Mann Whitney test

List	Mean (SD)			p	Statistical power
	Total (n = 26)	Female	Male		
1	-1.69 (0.633)	-1.73 (0.599)	-1.65 (0.688)	0.794	0.56
2	-1.76 (0.666)	-1.96 (0.66)	-1.57 (0.64)	0.126	0.54
3	-2.23 (0.533)	-2.26 (0.438)	-2.19 (0.63)	0.917	0.57
4	-1.57 (0.796)	-1.73 (0.832)	-1.42 (0.759)	0.315	0.63
5	-2.38 (0.325)	-2.34 (0.375)	-2.42 (0.277)	0.547	0.63
6	-2.07 (0.577)	-2.11 (0.506)	-2.03 (0.66)	0.88	0.65

Results

The results of this study are presented in three separate sections: developing the test, performing the test, and re-testing.

The CVR value, according to quantitative Lawshe's method, was 0.49 determined by 15 expert people. Accordingly, 48 out of 76 sentences were accepted, but after performing the face validity, only 43 sentences were selected. The content validity index (CVI) which is a quantitative index of the CVR mean value over the items for a specific test, was also determined. CVI was 0.736 for 48 sentences, an acceptable value.

The confirmed sentences in the previous step were defined +4, +2, 0, -2, -4, -6 dB in SNRs by MATLAB software to determine SNR50 threshold. The presentation started from -6 dB SNR and continued with increasing SNR in 2-dB steps to be determined the level of SNR that a person can recall 3 key words of the sentence. This intensity level was considered as the SNR50 threshold of each sentence. The defined sentences in SNRs were presented to 10 people at the most comfortable loudness (50 dB HL) to determine the SNR50 threshold of each sentence and select the appropriate sentences for the final test study. The amount of SNR50 in the Persian language is -4 dB that should be used it in SNR Loss formula. In this study, we added 4 dB to noise intensity for confirming Q-SIN test, and it was used in the formula to calculate the SNR Loss without applying SNR50. After calculating and checking SNR50, the easier sentences were deleted, and finally, 36 sentences were selected.

Q-SIN test was performed with emphasis on the high frequency on 26 right-handed (13 female and 13 male) Persian speaking volunteers with normal hearing. Table 1 presents the mean SNR loss of the samples separated by gender. As shown in this table, there is no significant difference between men and women in the mean obtained scores ($p > 0.05$).

The comparison of test lists was performed by checking the significant differences between mean SNR loss of the test lists in the samples. To check the equivalency of SNR loss lists, we used the Friedman nonparametric test. The results showed a significant difference between the SNR loss mean in all six lists ($p < 0.001$). Then the Wilcoxon test by Bonferroni correction was used. The results of all six lists are presented in Table 2.

According to simultaneous testing of the lists to reduce concurrency effects and increase the power of the test, we should correct the maximum error rate to the number of lists. That is why we used the Bonferroni test. Also, we used a more stringent alpha level for each comparison to maintaining alpha value in all tests at a reasonable level. To achieve this, the alpha level

Table 2. The equivalency of six Persian quick speech in noise test lists with in high frequency emphasis normal individuals with Wilcoxon c test corrected by Bonferroni (n = 26)

List	List	p	Statistical power
	2	0.593*	0.61
1	3	0.006*	0.55
	4	0.653*	0.70
	5	< 0.001	0.60
	6	0.012*	0.33
	3	0.026*	0.56
2	4	0.248*	0.45
	5	0.001	0.41
	6	0.101*	0.48
	4	0.006*	0.93
3	5	0.157*	0.35
	6	0.248*	0.47
	5	< 0.001	0.61
4	6	0.012*	0.61
5	6	0.011*	0.41

*Equivalent lists

was divided into the number of performed comparisons, and this new value was considered as the required alpha level. Since in the equivalency section, we made 15 comparisons, the new alpha level was obtained by dividing 0.05 by 15, which equals 0.003. Therefore, p values greater than 0.003 are acceptable. According to the findings of Table 2, only p values of 4, 8, 13 lines are less than 0.003. Therefore, there is no equivalency in these lists but p values of other lists in Table 2 are larger than 0.003, and there is equivalency between lists that marked with a star.

Two methods were used to check the reliability of Q-SIN test. We used the Wilcoxon test to compare the SNR loss means of six lists for test-retest in the first step. The results of the Wilcoxon test showed that the difference in mean of test-retest results was not statistically significant for lists number 1, 3, 4, 5, and 6 ($p > 0.05$) and the test-retest results in these lists approved the reliability of the lists Table 3. In the second method, we used the Spearman correlation coefficients for the reliability analysis and the results showed 0.99 for the list of 5 which indicates a significant correlation between test and retest Table 4. This coefficient was 0.82 for list 6 indicating a strong relationship. The Spearman correlation coefficient was 0.58, 0.61, 0.62, and 0.45 respectively for list 1, 2, 3 and 4 indicating a significant difference. However, there was a significant difference between test and retest in list 2, showing no reliability in this list.

Discussion

The most important factor that must be controlled in selecting Persian sentences for the test is the predictability of the sentences. The first step in preparing sentences is to select appropriate keywords. The words must be familiar, frequently-used, and mainly mono-syllabic as they have lower predictability. Also, the sentences should be developed in a way that the last part of the sentences was not predictable because of the beginning and the semantic relations of the words were as limited as possible. Based on the search in the published journals and websites, there was no study conducted on high-frequency Q-SIN. Therefore, the results of the present study will be compared to the results of the studies on ordinary Q-SIN lists. The comparisons were made regarding reliability, validity, SNR50, SNR loss, equivalency, and gender.

Validity

After preparing sentences, the content validity of the test was investigated. Evaluation of the content validity is the first step in evaluating an essential test. About the general content, validity is the indicator of whether the chosen sentences satisfy our stated targets or not [19]. In the present study, the expertise comments were used for determining content validity.

Face validity

Table 3. The comparison of test-retest SNR loss in six lists with high frequency emphasis with Wilcoxon test

		Mean (SD)			
List	**N**	**Test**	**Re-test**	**p**	**Statistical power**
1	26	-1.69	-1.73	0.705	0.71
2	26	-1.76	-2.07	0.011	0.46
3	26	-2.23	-2.38	0.102	0.59
4	26	-1.57	-1.76	0.244	0.58
5	26	-2.38	-2.34	0.317	0.37
6	26	-2.07	-2.03	0.655	0.69

Face validity refers to the relevance of the test instrument, its acceptance, and reasonability for what is supposed to be studied [13]. Face validity of sentences was investigated by the expertise and several undergraduate students as the examinees. In this stage, 5 sentences were put aside, and 43 sentences remained for the final step. Noise and the filtration method were according to Q-SIN test standard and three experts checked the recorded test, and they all confirmed its quality.

SNR50

Calculation of SNR50 is an important step for developing this test in every language. In the present study, the SNR50 was obtained as -4 dB in Persian. Therefore, 4 dB was added to the noise intensity, and for SNR loss calculation, the original formula was used, without applying SNR50. SNR in the English language is 2 dB, and this value must be applied in the formula [10,18]. In a study on Serbian Q-SIN, the SNR50 was -4 dB [20].

In a study conducted by Shayanmehr et al., SNR50 in Persian was -4 dB [13] and in the study of Moosavi et al., SNR50 in Persian was -0.25 [15]. SNR50 value in the present study was the same as Shayanmehr et al. study. The reason for the different results from Moosavi et al. study might be attributable to the different test materials. In addition, they did not use common and frequent sentences, and many sentences had unusual meaning therefore the SNR50 was lower than the present study. In other studies, SNR50 had not been reported.

SNR loss

For the first time, 6 high-frequency lists for quick speech in noise (Q-SIN) test were developed in the present study. Mean score of SNR loss for the list 1 to 6 were -1.65, -1.8, -2.23, -1.61, -2.38, and -2.07 dB, respectively. In addition, the total mean score of SNR loss in all lists was -1.96 dB.

Based on the reports of the Etymotic Research Center that developed Q-SIN test for the first time, the mean SNR loss in normal hearing subjects was 1.9 dB [18]. The reason for the difference between the results of the present study (-1.96 dB) and the original study (1.9 dB) might be due to the language difference. The Persian language has a higher level of redundancy than English does. In addition, it is difficult to make sentences with correct syntactic and grammatical rules when we want to make sentences unpredictable. The authors of the original Q-SIN test stated that the SNR loss must be from -2.5 to 2.5 dB and the present result is in agreement with that amount for 6 lists [18].

There are some studies on Persian Q-SIN test using ordinary lists that are reviewed here. In the first study conducted by Khalili et al. on 36 normal-hearing young adults, the mean values of SNR loss for 4 lists were -0.69, -1.63, -1.52 and -2.19 dB, respectively. Total SNR loss was -1.5 dB [12] which is in agreement with the present study finding. Shayanmehr et al. studied 5 new lists on 35 normal hearing subjects and the mean SNR loss for the list 1 to 5 were 0.32, 0.35, 0.47, 0.41 and 0.24 dB, respectively and total SNR loss was 0.35 [13]. Haniloo et al. tested these 5 lists (developed by Shayanmehr et al.) on 36 normal and 36 hearing-impaired subjects and showed that the mean values of SNR loss in normal-hearing subjects were 1, -0.13, 0.77, -1.41 and 0.58 dB, respectively. The total SNR loss was 0.16 [14]. The present study results are not in agreement with the

Table 4. The comparison between test-retest SNR loss means in six lists with high frequency emphasis with Spearman correlation coefficients (n = 26)

List	Correlation coefficient	p
1	0.58	0.002
2	0.61	0.001
3	0.62	0.001
4	0.45	0.021
5	0.99	0.000
6	0.82	0.000

Shayanmehr et al. and Haniloo et al. results. This disagreement might be due to the difference in test development method and different samples. Gheisi et al. tested 4 lists out of 5 lists of Shayanmehr et al. in Persian monolinguals and Persian-Azari bilinguals, and the mean values of SNR loss were -0.13, -1.07, -1.48 and -1.79, respectively and total SNR loss was -1.19 [21]. The present study is in agreement with their findings. In Moosavi et al. study, of 6 new Q-SIN lists developed and studied on 35 young subjects with normal hearing, the mean SNR values were 0.21, -1.3, -1.66, -1.38, -1.11 and -0.65, respectively and total SNR loss was -0.95 [15]. The present study finding is in line with this study.

Equalization

In the present study, the equalization of the 6 lists was investigated by comparing their SNR loss. Due to the lack of normal distribution of data, the Friedman test was used for mean comparison, which showed significant differences among lists ($p < 0.001$). To determine which two lists were equal, the mean SNR loss of all lists were compared. The equality was found among the following lists: list 1 with lists 2, 3, 4 and 6; list 2 with lists 3, 4, and 6; list 3 with lists 4, 5 and 6; list 4 with list 6; and list 5 with list 6. Therefore, it was concluded that list 1, 2, 3, 4 and 6 were equal.

Study on the equality of 18 English lists has shown that 9 lists were equal, including list 1, 2, 6, 8, 10, 11, 12, 15, and 17 [16].

The study was conducted by Haniloo et al. showed that only 2 lists from Shayanmehr et al.' study were equal [14]. In the study performed by Moosavi et al., it was shown that 4 lists out of 6 were comparable [15].

Gender effects

In the present study, Q-SIN results in normal subjects showed no significant sex-related differences, but according to the test power, there might be some effects by increasing the sample size. Shayanmehr et al. [13] and Moosavi et al. [15] showed no sex-related effects. There is not any study on the sex effects on English Q-SIN but Calais et al. investigated gender effects on speech perception in noise (SPIN) test and suggested no sex effects [22]. The present study results are in agreement with Calais et al. results.

Reliability

Based on the present study, all lists were reliable except list 2. The correlation coefficient was fair for list 1, 2, 3, and 4, very strong for list 5 and list 6. Given the test power, it might be possible to increase the correlation coefficient for the first 4 lists by increasing the sample size. As it is a speech test, if there were a 1-month interval instead of a 3-week interval, there would be a better correlation.

In the first Persian study, Khalili et al. studied Persian Q-SIN and showed that lists 2, 3 and 4 were reliable [12]. In the study conducted by Shayanmehr et al., the correlation for test-retest reliability was 0.8. All 5 lists had high reliability [13]. Haniloo et al. studied 5 lists of Shayanmehr et al. study and found no significant test-retest difference among lists [14]. Findings of Moosavi et al. showed that 5 lists (list 1, 2, 3, 4, and 6) out of 6 were reliable [15].

Conclusion

In the present study, 6 new Q-SIN lists with emphasis on high frequency were developed. Lists 1, 3, 4, 5 and 6 were reliable and lists 1, 2, 3, 4, and 6 were equal. Therefore, list 1, 3, 4 and 6 were reliable and equalized and they can be used in a clinical setting. Gender has no effects on the test results and test has essential standards for clinical usage in subjects with speech in noise complaint. Q-SIN is one of the few Persian tests designed for real-life evaluation of speech perception in noise. In addition, it plays an important role in monitoring and measuring the effectiveness of auditory rehabilitation as well as consulting patients and initiating rehabilitation interventions. According to test power, conducting research in a larger sample size is highly recommended.

Acknowledgments
This paper is extracted from M. Amiri Jahromi's MSc thesis supported by grant No. 97-03-32-39575 from TUMS, with Code of Ethics Committee of IR.TUMS.FNM.REC. 1397.003.
The authors would like to thank all the participants who helped us with this research.

Conflict of interest
The authors declared no conflicts of interest.

References
1. Wong PC, Uppunda AK, Parrish TB, Dhar S. Cortical mechanisms of speech perception in noise. J Speech Lang Hear Res. 2008;51(4):1026-41. doi: 10.1044/1092-4388(2008/075)
2. Lucks Mendel L, Widner H. Speech perception in noise for bilingual listeners with normal hearing. Int J Audiol. 2016;55(2):126-34. doi: 10.3109/14992027.2015.1061710
3. Theunissen M, Swanepoel de W, Hanekom J. Sentence recognition in noise: Variables in compilation and interpretation of tests. Int J Audiol. 2009;48(11):743-57. doi: 10.3109/14992020903082088
4. Taylor B. Speech-in-noise tests: How and why to include them in your basic test battery. Hear J. 2003;56(1): 40,42-46. doi: 10.1097/01.HJ.0000293000.76300
5. Killion MC. The SIN report: Circuits haven't solved the hearing-in-noise problem. Hear J. 1997;50(10): 28-30,32,34.
6. Killion MC, Niquette PA. What can the pure-tone audiogram tell us about a patient's SNR loss. Hear J. 2000;53(3):46-53.
7. Kochkin S. MarkeTrak V: "Why my hearing aids are in the drawer": The consumers' perspective. The Hearing Journal. 2000;53(2):34-41.
8. Nilsson M, Soli SD, Sullivan JA. Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. J Acoust Soc Am. 1994;95(2):1085-99.
9. Killion MC, Christensen LA. Principles of high-fidelity hearing aid amplification. In: Sandlin RE, editor. Textbook of hearing aid amplification. 2nd ed. San Diego: Singular Publishing Group Thomson Learning; 2000. p. 171-208.
10. Niquette P, Gudmundsen G, Killion M. QuickSIN Speech-in-Noise Test Version 1.3. Elk Grove Village, IL: Etymotic Research. 2001.
11. Duncan KR, Aarts NL. A comparison of the HINT and Quick SIN tests. J Speech Lang Pathol Audiol. 2006;30(2):86-94.
12. Khalili M, Fatahi J, Hajiabolhassan F, Tahaei AA, Jalaei S. Test-retest reliability and list equivalency of the Persian quick speech in noise test. MRJ. 2010;3(3-4):16-21. Persian.
13. Shayanmehr S, Tahaei AA, Fatahi J, Jalaie S, Modarresi Y. Development, validity and reliability of Persian quick speech in noise test with steady noise. Aud Vestib Res. 2015;24(4):234-44.
14. Hanilou J, Fatahi J, Tahaei AA, Jalaie S. List equivalency of the Persian quick speech in noise test on hearing impaired subjects. Aud Vestib Res. 2016;25(1):7-13.
15. Moossavi A, Javanbakht M, Arbab Sarjoo H, Bakhsh E, Mahmoodi Bakhtiari B, Lotfi Y. Development and psychometric evaluation of Persian version of the quick speech in noise test in Persian speaking 18-25 years old normal adults. Journal of Rehabilitation Sciences and Research. 2016;3(3):51-6.
16. McArdle RA, Wilson RH. Homogeneity of the 18 QuickSIN™ lists. Journal of the American Academy of Audiology. 2006;17(3):157-67. doi; 10.3766/jaaa.17.3.2
17. Carhart R, Jerger JF. Preferred method for clinical determination of pure-tone thresholds. J Speech Lang Hear Res. 1959;24(4):330-45. doi: 10.1044/jshd.2404.330
18. Killion MC, Niquette PA, Gudmundsen GI, Revit LJ, Banerjee S. Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. J Acoust Soc Am. 2004;116(4 Pt 1):2395-405.
19. Aiken LR. Content validity and reliability of single items or questionnaires. Educational and psychological measurement. 1980;40(4):955-9. doi: 10.1177/001316448004000419
20. Vojnoviü M, Subotiü M. Specifiþnosti QiuckSIN testa za srpski jezik. Telekomunikacioni forum TELFOR. 2010;23-5:1033-6.
21. Gheissi E, Fatahi J, Farahani S, Jalaie S, Tahaei AA. Test-retest reliability and list equivalency of Persian quick speech in noise test in Azari-Persian bilinguals. Aud Vestib Res. 2017;26(3):157-62.
22. Calais LL, Russo IC, Borges AC. Performance of elderly in a speech in noise test. Pro Fono. 2008;20(3):147-53.