Research Article

Construction and Standardization of the Syrian Version of the Word Recognition Test for Children with Hearing Impairment

Samer Mohammad Mohsen[1,2*], **Farah Jabri**[2], **Shaza Al Maidani**[3], **Abeer Ammar**[4], **Rahaf Alkhatib**[4]

[1.] *Department of Audiology, Faculty of Health Sciences, Damascus University, Damascus, Syria*
[2.] *Department of Audiology, Syrian Organization for Persons with Disabilities-Aamal, Damascus, Syria*
[3.] *Department of Psychological Testing and Evaluation, Syrian Organization for Persons with Disabilities-Aamal, Damascus, Syria*
[4.] *Department of Speech and Language Therapy, Syrian Organization for Persons with Disabilities-Aamal, Damascus, Syria*

Use your device to scan and read the article online

Citation: Mohsen SM, Jabri F, Al Maidani S, Ammar A, Alkhatib R. Construction and Standardization of the Syrian Version of the Word Recognition Test for Children with Hearing Impairment. Aud Vestib Res. 2022;31(4):264-74.

https://doi.org/10.18502/avr.v31i4.10730

Highlights

- Word Recognition test`s word list was established in the Syrian Arab language
- The test materials are valid, reliable, and stable for test-retest conditions
- The Syrian speech material is essential for national hearing rehabilitation programs

Article info:
Received: 10 Nov 2021
Revised: 26 Jan 2022
Accepted: 02 Feb 2022

[*] **Corresponding Author:**
Department of Audiology, Faculty of Health Sciences, Damascus University, Damascus, Syria.
samer_yn@yahoo.com

ABSTRACT

Background and Aim: This study aimed to construct and validate the test material of speech recognition test for children. Such a material helps in conducting the speech tests for native Arab children with Syrian accent. This work is a part of a complete project on preparing and validating the necessary materials for speech and central auditory processing tests for the Arabic Syrian community.

Methods: In this cross-sectional comparative study, two sets of phonetically balanced monosyllabic word lists were constructed. The two selected 25-monosyllabic word lists were presented to 50 normal hearing participants with the age range of 5 to 10 years. Tests of validity and reliability were conducted to assure the suitability of this material for the word recognition test. The test was repeated for the same children after three weeks.

Results: The two word lists have shown high face and content validity, Cronbach`s alpha coefficient was 0.79, 0.85 for the two lists respectively. Factor validity has shown that the test materials are one-dimensional, the first factor achieved 57.25 and 53.13 of the variance. Test-retest reliability ($p<0.001$) and Intraclass Correlation Coefficients reliability ($p<0.001$) were also significant.

Conclusion: The data analysis shows that the Syrian version of the word recognition test has high degree of validity and reliability. Therefore, developing the Syrian version of the scale has good psychometric properties, which makes it appropriate for the local use.

Keywords: Speech discrimination test; monosyllabic; arabic; syrian accent

Introduction

Normal hearing levels and natural language development are always linked throughout the early stages of child maturity [1]. It is well known, that the speech development process starting with babbling at six months of age, and the number of words that are specified for each age group, is delayed depending on the severity of hearing loss [2]. Hence, the proper identification, intervention and follow-up demands of a full test battery to be available, in order to guarantee that the child is keeping up with his peers pragmatically, academically, and socially [3]. According to literature, Jones and Knudsen were the firsts to develop the audiometer that allows puretone and speech audiometry [4]. Nowadays, speech audiometry is considered the most effective tool in measuring the extent to which hearing impairment has functionally affected the patient's communication, and the effectiveness of the re/habilitation programs [5].

Speech discrimination test is a crosschecking behavioral test used to assess hearing loss in addition to the puretone audiometry, which helps determine the degree and type of hearing loss. The speech discrimination test is a measurement tool that includes a set of procedures subject to specific conditions and rules. Its objective is to determine the degree to which an individual speech recognition is affected through the individuals' responses to a selected list of items (words) representing the ability to be measured [5, 6]. It provides information about the patient's ability to recognize speech, the uncomfortable levels of speech stimuli and helps in evaluating hearing aid performance in patients with moderate to severe sensorineural hearing loss [7, 8]. It is also possible, via this test, to identify the patient's ability to hear and understand speech sounds in noisy environments. During the speech discrimination test, the patient is presented with speech stimuli in different ways (live audio, headphones) and he is asked to repeat the words he hears. As a result, he gets a total score called the speech discrimination score or word recognition score (WRS), which is classified as normal, slight difficulty, poor and very poor speech discrimination levels [5, 8].

Hearing impairment has become a concern for specialists and members of society alike [9]. Therefore, it has become necessary to provide this group with the best rehabilitation services and find measures that will benefit in improving the hearing-related problems to enable the persons with hearing disability adapting and reaching the level of ordinary individuals. Children with hearing impairment have a clear deficiency in both receptive and expressive language processes, which is clearly manifested in the inconsistency of speech and lack of coherence in the use of vocabulary items appropriate to the child's age [10-12].

Based on the above, there has become an urgent need to prepare a set of tools and measurement scales through which speech discrimination among children with hearing disabilities can be assessed, including the word recognition scale that is the focal topic in the current research. The purpose of the present research was to further contribute to the development of Arabic speech audiometry materials by producing standardized Arabic words lists in the Syrian dialect for the use of measuring the WRS. This will provide speech test materials for audiologists in Syria for testing individuals whose native language is Syrian Arabic. The importance of the test is that it provides accurate information about the degree and consequences of hearing problems in children, and as a verbal test, it helps in assessing the mental abilities of the child [5].

Methods

This cross-sectional comparative research is a type of test-development study, conducted and validated at the Syrian Organization for persons with Disability-Aamal, in Collaboration with Damascus University.

Item construction

In constructing the test items, audiologists from the Aamal Organization who hold Master's and Ph.D. degrees were consulted to examine the foreign and Arab studies on word recognition test construction. Additionally, to develop alternative word lists comparable to the original test that are phonetically balanced (PB). The Audiology center team suggested a whole list of about 275 famous monosyllabic words from the Syrian Arabic dialect and elementary school`s books. The target was to find two phonetically balanced lists, each list consisting of 25 monosyllabic words and suitable for children aged between 5 and 10 years. Therefore, a preliminary review of the list and evaluation in terms of the words fulfilling the required conditions; some duplicate words were deleted and new words were added. After that, the Speech center team, which included 36 speech language pathologists, has conducted a mini-study on the prevalence of sounds in the Arabic language (standard and colloquial). The study included the selection of various random samples such as articles, news, books, various samples of spontaneous conversations, basic concepts in the rehabilitation plan, songs, extracts from interviews, clips posted on social networking sites, and samples from the rehabilitation sessions. Each sample consists of a minimum of 100 words, about 500 to 600 sounds, and a study on the frequency of each sound in the sample (according to the form) to determine whether it is rare, common, or very common based on the percentage was established [13]. Rare sounds are the ones that recorded a frequency rate of 0%-2% with all the results collected from specialists, which were (/θ/, /x/, /ð/, /z/, /sˤ/, /dˤ/,/tˤ/, /ðˤ/, /ɣ/,

/h/, /ʒ/). The common sounds are the ones that recorded a frequency rate of 2.1%-5% and they were (/ʔ/, /ħ/, /d/, /s/, /ʃ/, /ʕ/, /f/, /q/, /k/). As well, the very common sounds that recorded a frequency rate of (5.1% and above) were the sounds (/b/, /t/, /r/, /l/, /m/,/n/).

Then, the speech pathologists reevaluated the word list and added new words to achieve the obtained popularity. They determined the number of words in which the sound is targeted according to the frequency (vowels, consonants) proceeding to the next stage, which was the selection of the words for each of the two lists in proportion to what was previously specified. The words were divided into words of one syllable of the pattern consonant vowel consonant (CVC), and words of one syllable of the pattern consonant vowel consonant-consonant (CVCC), and words of one syllable of the pattern consonant-consonant vowel consonant (CCVC); noting that there are no words of this pattern. After that, they added the words in each list written in phonetic symbols, and balanced them in terms of the type of previous syllable. Using the percentage of each sound frequency in the Arabic language, the ratio of the number of a group of (rare, common, and very common) sounds to the total number of sounds was calculated. Then, by experimenting them with numerous words on children and adults, a number of final two word lists were prepared; each of which had a balanced percentage of rare sounds ranged between 16.7% and 19.2%; a balanced percentage of common sounds ranged between 33.9% and 35.9%; a balanced percentage of the very common sounds ranged between 44.7% and 47.8%. They also contained an equal number of words consisting of one syllable of the pattern CVC; an equal number of words consisting of one syllable of the pattern CVCC. However, no words of the pattern CCVC were included.

Regarding the vowels distribution, the two lists had an equal number of vowel /æ/, which was the most common used vowel; in addition to equal numbers of the vowel /u/ and /i/. Lastly, the word lists were formulated and refereed by a committee of experts in the field of Speech therapy and Audiology in Aamal Organization. The team judged the word items and its alternatives alongside the time needed to complete each item. The necessary amendments were made to the items as requested by the committee, then, they were presented again to the panel of referees to reach a 90% conformity before its final approval. The final test items were formed in 50 items; they were divided into two lists; 25 items for each list. The two sets that approved for use in children between 5 and 10 years old were selected for conducting this study; Appendix 1.

Subjects

The research community included all children (boys and girls) aged from 5 to 10 years, and the current research sample consisted of 50 Syrian children who are native speakers of Arabic language for the purposes of extracting the psychometric indications for the test. Those 50 children were enrolled through online advertising of participation on social media. All participants had passed the inclusion criteria as they had pure-tone air-conduction thresholds 5-15 dB HL at octave and mid-octave frequencies from 125 to 8000 Hz, and had static acoustic admittance between 0.3 and 1.4 mmhos with peak pressure between -100 and +50 daPa [14]. Additionally, each subject passed a screening exam, which included an otoscopic evaluation, presence of an ipsilateral acoustic reflex of 95 dB HL or better in the test ear at 1000 Hz, and a comprehensive language assessment by the speech pathologists. The exclusion criteria contained any case of not completing the test or retest sessions, or having any medical condition, which might affect the test results.

Talker and test presentation

Two females from the Audiologists of the team volunteered to present the test. The referee from the speech language center examined their voices and the top ranked talker was selected to say the monosyllabic Arabic words. The test was presented using monitored live voice, which is more suitable for the tested sample. A carrier sentence of "say the word" was used before each of the words. The presentation level was set on the most comfortable level of speech for each child, which ranged between 40 and 50 dB HL. The GSI AudioStar Pro (USA). Audiometer was used as the stimulus words were routed from the audiometer to the participant through a single TDH-50P headphone. Testing of each participant was conducted in a double-walled sound booth that met ANSI S3.1 standard [15]. Before testing each participant, the external inputs to the audiometer were calibrated to 0 VU by using the 1 kHz calibration tone on the wav file. The audiometer was calibrated before, weekly, and after data collection. This calibration was completed according to ANSI S3.6 specifications [16].

Validity and reliability testing

To ensure the validity and reliability of the Syrian version of speech discrimination word lists, we used the essential components of validity and reliability tests. For validity, the face validity, content validity, and factorial validity were established. As well, the split half, Cronbach's alpha, internal consistency, and the difficulty and discrimination coefficient factors have been established. Moreover, for reliability testing, the test-retest reliability was used. All used statistical

tests were performed using the statistical software package SPSS version 20 (SPSS Inc, Chicago, IL)), with p-value less than 5% was considered as significant.

Content validity and face validity

It is called the validity of the referees or experts, as the test is presented to a committee of expert university professors to judge the clarity, readability of the vocabulary items and its linguistic structure, its children age appropriateness and its effectiveness in measuring what it was designed to measure [17]. The test was presented to seven of expert specialists in Audiology, Speech therapy, and Linguistics in Aamal organization and Damascus University. They judged the validity of each test item to ensure that it measures what it was supposed to do. They corrected and modified the inappropriate or not essential items. There suggestions (correction, modification, addition, etc.) were taken into consideration and necessary amendments were made. To study the content validity, we used Lawashe method in which we added the experts` opinion on each one of the items then we calculated the content validity ratio (CVR) using the formula CVR=ne–N/2/N/2. As well, we calculated the content validity index (CVI), which is the proportion of content experts giving item a relevance rating of being "essential". The CVI is simply the mean of the CVR values of the retained items. It expresses the proportion of agreement on the relevancy of each item, which is between zero and one. After calculating the CVR we reported the total scale CVI which can be calculated by the formula CVI=(sum of CVR scores)/(number of item) [18, 19].

For measuring face validity, opinions of ten speech language specialists and children about the clearness of each item were collected; Then, the agreement between opinions were calculated using Cooper's equation. Pa=[Ag/(Ag+Dg)]*100; where Pa refers to percentage of agreement, Ag the number of those who agreed, and Dg=number of those who disagreed. The results showed that the percentage of the face validity of the groups ranged between 90% and 100%.

Difficulty coefficients analysis

The difficulty and discrimination coefficients show the suitability of the items to the level of the children who took the test. Item difficulty measures whether answering the item was very easy or very hard. However, the discrimination index measures how the item can discriminate between participants. Using SPSS, we calculated the difficulty and discrimination index for each item and the p value, which refers to the probability of correct answer, was calculated. Indices fallen between 0.3 and 0.7 are considered of moderate difficulty. Thus, in the word recognition test, the word lists should be of moderate difficulty therefore, items, which have index between 0.3-0.7, will be accepted. On the other hand, the items discrimination indices fallen between 0.2-0.4 are considered to be discriminating, and those of index more than 0.4 are strongly discriminating; we accepted all indices above 0.2.

Internal consistency validity

The internal consistency was calculated between each item in each one of the two lists and the total score of it.

Cronbach's alpha

A method depends on the internal consistency and shows the consistency of the items with each other and with the total score of each one of the two word lists. This method depends on the consistency in the performance of the examinees from one item to another and is based on the standard deviation of the scale and the standard deviations of individual items [20].

Split-half coefficient

After applying the test to the research sample, the half-part method is followed, where the test items are divided into two parts, each part (semi-set). Then, we continue the analysis by finding the correlation coefficient between the two halves of the test.

Factorial validity

The factor analysis sample included 50 children aged between 5 and 10 years. The exploratory factor analysis was conducted according to the Principle Component Method. The interpreted variance test was used to determine the number of factors. Depending on The Kaiser Eigen Value root`s rule only the factor whose root exceeded 1 was considered as the general factor [21]. To ensure the accuracy of the test, the criterion percent variance was extracted. This test depends on the percentage of cumulative variance that the factors explain. This method is considered important in practice because it indicates the ability of the selected factors to absorb or represent information (interpretation of variance) in the data, which contains the largest percentage of variance and the least number of factors.

Reliability

It is defined as the degree of consistency between two tests or two measures of the same trait among the subjects [22]. Test-retest reliability was calculated using Pearson's correlation coefficient between the two phases of administering the questionnaire. T-test also was used to compare the means of test scores for all children between the test and retest phases and between the two lists. Lastly, the intraclass correlation

coefficients (ICC) estimates and their 95% confident intervals were calculated using SPSS statistical package.

Results

Before initiating the statistical processes of the hypotheses, it was ensured that the data follows the normal distribution, in order to choose the appropriate statistical methods. The measures of central tendency alongside some measures of dispersion were calculated. The skewness coefficients of the sample ranged between (+1 and -1), while the kurtosis coefficients ranged between +3 and -3, which are within the normal range. Therefore, this indicates that the distribution was moderate to some extent, supported by the graph of the individuals' score distribution, Figure 1. Frequency statistics for the test`s scores according to gender with means comparison are shown in Table 1.

Validity analysis

Content validity

The CVR for each item in the first word list was acceptable; 5 items had the CVR of 0.71 that is one of the referee had considered the item as "not essential" while the other 20 items had CVR of 0.99 with the agreement of the seven referee to be essential. For this list, the CVI was 0.93. The same for the second word list, the CVR was 0.99 for 23 items, while only two items had the CVR of 0.71 and the CVI was 0.94. Therefore, the content validity of the two word lists was established.

Difficulty coefficients analysis

All items indices of difficulty and discrimination are within the accepted range, which is the moderate level of difficulty. For that, we accepted the items and could proceed with the other validity and reliability analysis (Table 2).

Internal consistency validity

All correlations, between each item in each one of the two lists and the total score of it, were high. Correlation coefficients were between 0.7 and 0.9, and all statistically significant at p-value of 0.01.

Cronbach's alpha coefficient analysis

The Cronbach's alpha coefficient was 0.79 and 0.85 for the both lists respectively.

Split-half coefficient analysis

The correlation coefficient between the two halves of the test was 0.97, and 0.84 for the two lists respectively, and because the value of this correlation coefficient represents the split-half coefficient of the half of the test, it was modified by calculating the reliability coefficient of the test using the Spearman-Brown equation. The correlation coefficient is corrected by the modified Spearman-Brown equation, and thus the reliability coefficient value is 0.89 and 0.75 respectively, a parameter that can be trusted. The following table shows the results. Hence, the Cronbach's alpha and split-half coefficient are very high; the high internal consistency of the test was assured (Table 3).

Factorial validity

The suitability of the sample size range was investigated using the Kaiser-Meyer-Olkin (KMO) procedure. KMO value was 0.69 which is more than 0.6 that allows to proceed with the factorial analysis without remedial actions [23]. After applying the items of each test to the main sample of 50 children, the validity coefficients were extracted by conducting a factorial analysis of the items for each test separately.

The two graphs in Figure 2 show the one-dimensional distribution of the test; that is the test is calculating only one factor, which is the speech discrimination for which it was designed. It appears in the Figure 2 that the first factor is prominent that it reached the number of 57.25 and 53.13 of the variance for both lists respectively, while the second factor is very smaller less than 10, which is thus achieving one-dimensionality.

Reliability

Test-retest reliability

The test was re-conducted on the same children 23 days later. In each session, the child was tested with the two 25-words. The correlation coefficient between the two tests was calculated in order to establish the test-retest reliability coefficient. Results are shown in (Table 4).

It is obvious from the above table that the results of the correlation coefficients between the first and second application of the test on the same children were high enough to allow the use of test items.

Mean score comparison between the two word-lists test in all conditions: Paired t-test was used to compare the test scores for the children sample to ensure that the two tests are similar and stable. Table 5 shows that there are no significant differences between the two applications of the test in all conditions.

Table 1. Descriptive statistics, and means comparisons for the scores of the children on the test total score according to gender

	Gender	Sample	Mean (SD)	Mean standard error	Independent t-test
Word recognition test	Boys	21	44.90 (4.45)	0.68	t=0.608, df=72 p>0.05
	Girls	29	44.53 (0.35)	0.77	
	Total	50	44.17 (4.38)	0.52	

df; degree of freedom

Intraclass correlation coefficients

Table 6 shows the output of ICC reliability analysis in SPSS. The obtained ICC value from the first word list is 0.89 (indicating good reliability), its 95% confidence interval ranges between 0.78 and 0.93, meaning that there is 95% chance that the true ICC value lands on any point between 0.78 and 0.93. On the other hand, the obtained ICC value from the second word list is 0.90 (indicating excellent reliability), its 95% confidence interval ranges between 0.90 and 0.91. Therefore, based on statistical inference, it is appropriate to conclude the level of reliability of the two word lists to be good to excellent [24].

Discussion

The main purpose of this study was to develop a phonetically balanced set of Syrian monosyllabic words that can be used to measure the WRS for children between 5 and 10 years, whose native language is Arabic with Syrian accent. Two lists of words were developed, each containing 25 monosyllabic words that are phonetically balanced and suitable for the children sample. The data analysis shows that the Syrian version of the word recognition test has high degree of validity and reliability. Therefore, developing the Syrian version of the scale has high and good psychometric properties, which makes it appropriate for local use. Degrees of skewness and kurtosis were within normal range. The results also

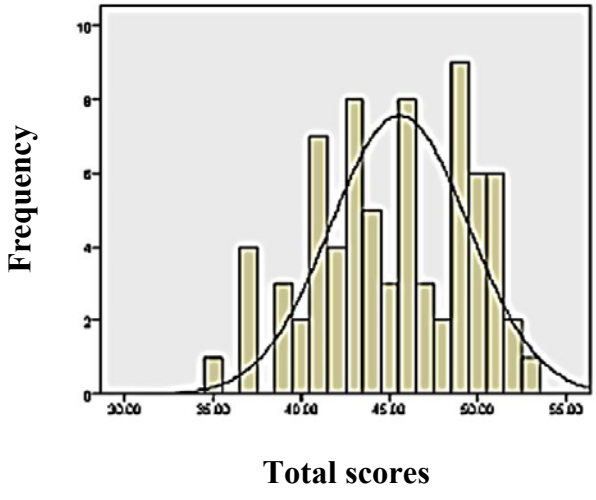

Figure 1. The normal distribution graph of the data of the research sample, total scores (50 items) were used

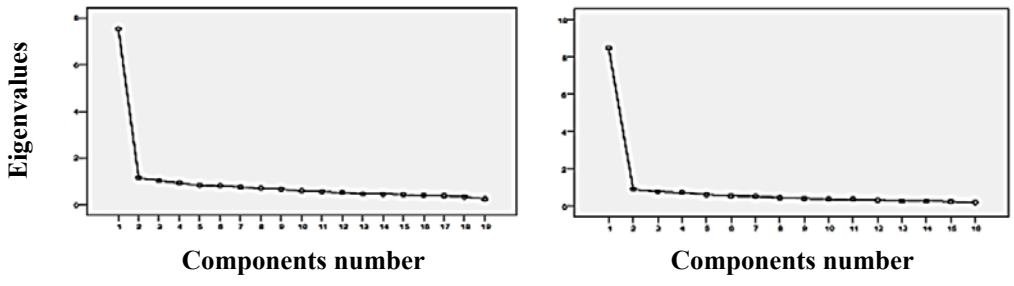

Figure 2. The factorial analysis graph for the first word list on the right and the second one on the left

Table 2. Discrimination and difficulty index with p value for each item in the two word lists

Item	First list			Second list		
	Difficulty index	p	Discrimination index	Difficulty index	p	Discrimination index
1	65.10	0.65	0.48	65.10	0.65	0.52
2	51.30	0.51	0.39	54.30	0.54	0.62
3	57.20	0.57	0.45	54.90	0.55	0.55
4	63.90	0.64	0.60	65.40	0.65	0.45
5	41.40	0.41	0.32	64.40	0.64	0.40
6	65.10	0.65	0.36	67.80	0.68	0.49
7	44.10	0.44	0.53	48.20	0.48	0.45
8	51.80	0.52	0.23	69.30	0.70	0.60
9	58.20	0.58	0.52	56.40	0.56	0.32
10	69.10	0.69	0.62	55.30	0.55	0.36
11	64.30	0.64	0.55	54.30	0.54	0.53
12	55.20	0.55	0.35	57.10	0.57	0.63
13	54.70	0.55	0.50	60.40	0.60	0.52
14	40.10	0.40	0.32	59.30	0.59	0.52
15	61.30	0.61	0.61	61.10	0.61	0.62
16	51.10	0.51	0.48	55.10	0.55	0.55
17	56.80	0.57	0.66	62.80	0.63	0.55
18	70.20	0.70	0.35	60.60	0.61	0.59
19	64.30	0.64	0.54	46.80	0.47	0.53
20	70.20	0.70	0.32	64.80	0.65	0.63
21	67.60	0.68	0.55	67.10	0.67	0.52
22	63.90	0.64	0.65	60.30	0.60	0.42
23	60.80	0.61	0.59	45.20	0.45	0.55
24	65.20	0.65	0.53	61.10	0.61	0.55
25	61.10	0.61	0.23	60.80	0.61	0.40

showed that there are no statistically significant differences between the performance averages of the children sample on the test based on gender.

It is very essential for any test material to examine the real number of factors for which the test is constructed. In this study, we constructed two words list in which we expect to examine the word recognition score and no other factors. For this purpose, we used the factor validity test; The test results have shown that there is one factor that reached the percent of 57.25 and 53.13% of the variance for the two lists respectively, while the next factors were very smaller less than 10, which is thus achieving one-dimensionality. Based on the evidence practice of factorial analysis, a factor can be considered as true if it achieved at least more than 40% of the variance, however, the score of 80% or above is much preferable to claim having one general factor [21]. The small test sample will affect the results where at least 100 participants is needed. The sample size of 50 children is one limitation for having this test in the current study even though we examined the suitability of the sample size using the KMO procedure. Adding to all, the interpretation of the factor analysis results can be built on both theoretical and statistical bases [21]. In our test we did not expect to have other factors which may affect the pureness of the test and the first factor could exceed the minimum acceptable level variance, for that, it can be claimed that this factor represents the goal of the test which is speech discrimination.

Table 3. Split-half coefficients of the two word lists

The test material (n=25 for each list)	Guttman coefficient value	Spearman-Brown coefficient value	Correlation coefficient between the two halves of each dimension and the scale as a whole	Cronbach's alpha coefficient	Section
First list	0.97	0.89	0.78	0.83	12
				0.87	13
Second list	0.84	0.75	0.87	0.85	12
				0.95	13

Table 4. Value of test-retest reliability coefficients of the two word lists

Test material	Number of items	Pearson`s coefficient	p
First list	25	0.78	<0.001
Second list	25	0.83	<0.001

Table 5. Compare means of test scores for the two lists in both conditions using paired t-test

Comparison	t-value	Degree of freedom	p
First list test-retest	0.86	24	0.501
Second list test-retest	0.82	24	0.498

Table 6. The intraclass correlation coefficients reliability and 95% confidence interval for the two word lists

Word list	Test	Intraclass correlation coefficients	95% confidence interval		F test with true value=0			
			Lower bound	Upper bound	Value	df1	df2	p
First	Single measurement	0.89	0.78	0.93	43.26	26	61	<0.001
Second		0.90	0.90	0.91	43.51	23	57	0.001

df; degree of freedom

Table 7. The values of the percentiles and the corresponding raw scores driven from the children's scores on the word recognition test

Raw grade	Percentiles	Raw grade	Percentiles	Raw grade	Percentiles
34.0	75	16.5	40	1.0	05
36.5	80	19.0	45	1.5	10
39.0	85	21.5	50	4.0	15
41.5	90	24.0	55	6.5	20
44.0	95	26.5	60	9.0	25
46.5<	100	29.0	65	11.5	30
		31.5	70	14.0	35

Decades, the development of proper speech audiometry materials has improved, especially for the English language [25]. The test materials are available in many languages, however, the development of these materials in many other languages is still lacking [26-28]. The selected Syrian monosyllabic words should be used in further research to study the WRS in the hearing-impaired populations, which is our target. For such a reason, we had to answer the question of what are the percentile standards corresponding to the raw scores on the Syrian version of the vocabulary knowledge test.

The process of deriving standards is the last and most important step in the process of constructing psychological tests preparation and construction; this allows the use of this scale on other groups and samples in society different from the ones on which the scale was initially based [29]. The standards are used to compare the score an examinee gets and determine its ratio to the whole community. The raw score on the scale has no meaning and significance unless the individual's position is determined in relation to the standardized community and building standards formulation. Since standard derivation is completed from the representative sample of the community for which the test is being prepared, therefore, the standards of the current test must be derived a representative sample of the original community.

Standard derivation is regarded as the main feature of the psychological scales and tests. without which a test cannot be valid and the level of an individual will not be determined in relation to his/her peers because the standards are specific levels of measurement that we refer to in order to understand the raw scores obtained by e subjects of the research sample [30]. To determine the level of children's performance on the word recognition test in its current form, measures of central tendency and measures of dispersion were used, as well as t-scores and percentiles corresponding to the raw scores obtained by the research sample members. Percentiles are specific points in a continuous distribution that falls at or below a certain percentage of the targeted score of a group or sample. The percentile rank of an individual is the place of an individual on a scale of 100 that qualifies it for the degree it obtains in this distribution. The percentiles corresponding to the raw scores of the basic research sample of 50 children were calculated depending on the raw scores and their frequencies.

After verifying the validity of the test and the reliability of the total scores resulting from its application. The above mentioned question was answered by finding the percentile performance standards of its application on the sample, which rank the examinee and indicate his position in the measured trait represented in the main nine percentile standards. The percentiles of 5, 10, 15, 20, 25, 50, 75, 90, 95, and 100 and the corresponding raw grades were demonstrated according to the Table 7 depending on the total score (50 items) the child achieved.

Conclusion

The monosyllabic word lists for word recognition score calculation in Syrian accent were established. They have a good validity and reliability indicator for the use as a test material. Two other 25-monosyllabic words sets and Bi-syllabic materials are prepared to complete the speech audiometry materials in Syrian accent. This work is very important for assessment of hearing and central auditory processing in children and adults with hearing disability.

Ethical Considerations

Compliance with ethical guidelines

All parents have signed an informed consent form to participate in this study, which is in accordance with the Helsinki Declaration of Ethics and was approved by the Scientific Board of Aamal Organization.

Funding

This research did not receive any grant from funding agencies in the public, commercial, or non-profit sectors

Authors' contributions

SMM: Study design, supervision, interpretation of results, critical revision of the manuscript; FJ: Study design, acquisition of data, drafting the manuscript; SA: Statistical analysis, drafting and revision the manuscript; AA: Preparing the PB word lists, data acquisition; RA: Preparing the PB word lists, data acquisition.

Conflict of interest

The authors declare that there is no conflict of interest to be reported.

Acknowledgements

The authors would like to thank Professor David L. McPherson, Brigham Young University, Provo, Utah, USA; for his help in supervision of many important parts of this work. Many thanks for the staff of Speech Language Department who helped in gathering the samples of Syrian dialect. Additionally, we thank the Huda Al-Obaid for her high contribution in language assessment of the children. Many thanks for the parents and Aamal staff who cooperated in doing this work even in COVID-19 circumstances.

Appendix 1. The two word lists, written in Arabic with their meanings and transcription

Second list			First list			Item
Meaning	Transcription	Word	Meaning	Transcription	Word	
Damas	/ʃæ:m/	شام	Dad	/ʔæb/	أب	1
Door	/bæ:b/	باب	Lettuce	/xæs/	خس	2
Hill	/tæl/	تل	guy	/ʃæ:b/	شاب	3
Light	/nu:r/	نور	Length	/t̪u:l/	طول	4
Wind	/ri:ħ/	ريح	FIG	/ti:n/	تين	5
Name	/mæjs/	ميس	Arch	/qæws/	قوس	6
dress	/Θæwb/	ثوب	Board	/læwħ/	لوح	7
fire	/næ:r/	نار	water	/mæ:ʔ/	ماء	8
Crown	/tæ:ʒ/	تاج	Gas	/ɣæ:z/	غاز	9
Duck	/bæt̪/	بط	Land	/bær/	بر	10
Nest	/ʕuʃ/	عش	Love	/ħub/	حب	11
Reem	/ri:m/	ريم	Socket	/fi:ʃ/	فيش	12
Dog	/kælb/	كلب	Dam	/sæd/	سد	13
Guest	/d̪æjf/	ضيف	Night	/læjl/	ليل	14
True	/s̪æħ/	صح	Luck	/ħæð̪/	حظ	15
Name	/ʒæ:d/	جاد	Mouth	/fæm/	فم	16
Trap	/fæx/	فخ	Neighbor	/ʒæ:r/	جار	17
Bean	/fu:l/	فول	Stick	/ʕu:d/	عود	18
rooster	/di:k/	ديک	Cat	/hir/	هر	19
time	/wæqt/	وقت	Date	/tæmr/	تمر	20
Fence	/su:r/	سور	Worms	/du:d/	دود	21
button	/zir/	زر	Mile	/mi:l/	ميل	22
Right	/ħæq/	حق	Palm	/kæf/	كف	23
Mother	/ʔum/	أم	Wool	/s̪u:f/	صوف	24
Nile	/ni:l/	نيل	Soft	/li:n/	لين	25

References

[1] Davis JM, Elfenbein J, Schum R, Bentler RA. Effects of mild and moderate hearing impairments on language, educational, and psychosocial behavior of children. J Speech Hear Disord. 1986;51(1):53-62. [DOI:10.1044/jshd.5101.53]

[2] Eilers RE, Wilson WR, Moore JM. Developmental changes in speech discrimination in infants. J Speech Hear Res. 1977;20(4):766-80. [DOI:10.1044/jshr.2004.766]

[3] ASHA. Hearing Screening and Testing. American Speech-Language-Hearing Association. ASHA; [cited 2020 Oct 22]. Available from: https://www.asha.org/public/hearing/hearing-testing/

[4] Jones IH, Knudsen VO. LXXXI. Facts of Audition. Ann Otol Rhinol Laryngol. 1925;34(4):1013-27. [DOI:10.1177/000348942503400401]

[5] Fee J. Speech Discrimination. B U L Rev. 2005;85:1103.

[6] Boothroyd A. Statistical theory of the speech discrimination score. J Acoust Soc Am. 1968;43(2):362-7. [DOI:10.1121/1.1910787]

[7] CDC. Types of Hearing Loss. Centers for Disease Control and Prevention. 2020 [cited 2021 Apr 2]. Available from: https://www.cdc.gov/ncbddd/hearingloss/types.html

[8] Guidelines for determining threshold level for speech. ASHA. 1988;30(3):85-9.

[9] World Health Organization. Addressing the rising prevalence of hearing loss. World Health Organization; 2018. Available from: https://apps.who.int/iris/handle/10665/260336. Licence: CC BY-NC-SA 3.0 IGO.

[10] Schmucker C, Kapp P, Motschall E, Loehler J, Meerpohl JJ. Prevalence of hearing loss and use of hearing aids among children and adolescents in Germany: a systematic review. BMC Public Health. 2019;19(1):1277. [DOI:10.1186/s12889-019-7602-7]

[11] Bess FH, Dodd-Murphy J, Parker RA. Children with minimal sensorineural hearing loss: prevalence, educational performance, and functional status. Ear Hear. 1998;19(5):339-54. [DOI:10.1097/00003446-199810000-00001]

[12] CDC. Data and Statistics About Hearing Loss in Children. Centers for Disease Control and Prevention. 2020 [cited 2020 Oct 23]. Available from: https://www.cdc.gov/ncbddd/hearingloss/data.html

[13] Storkel HL, Morrisette ML. The Lexicon and Phonology: Interactions in Language Acquisition. Lang Speech Hear Serv Sch. 2002;33(1):24-37. [DOI:10.1044/0161-1461(2002/003)]

[14] Roup CM, Wiley TL, Safady SH, Stoppenbach DT. Tympanometric Screening Norms for Adults. Am J Audiol. 1998;7(2):55-60. [DOI:10.1044/1059-0889(1998/014)]

[15] Frank T, Durrant JD, Lovrinic JM. Maximum Permissible Ambient Noise Levels for Audiometric Test Rooms. Am J Audiol. 1993;2(1):33-7. [DOI:10.1044/1059-0889.0201.33]

[16] ANSI. ANSI S3.6-2004 Specification for Audiometers. New York: American National Standard Institute; 2004.

[17] Ostergard CA. Factors Influencing the Validity and Reliability of Speech Audiometry. Semin Hear. 1983;4(3):221-39. [DOI:10.1055/s-0028-1091427]

[18] Lawshe CH. A quantitative approach to content validity. Pers Psychol. 1975;28(4):563-75. [DOI:10.1111/j.1744-6570.1975.tb01393.x]

[19] Yusoff MSB. ABC of content validation and content validity index calculation. Education in Medicine Journal. 2019;11(2):49-54. [DOI:10.21315/eimj2019.11.2.6]

[20] Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika. 1951;16(3):297-334. [DOI:10.1007/BF02310555]

[21] Watkins MW. Exploratory Factor Analysis: A Guide to Best Practice. J Black Psychol. 2018;44(3):219-46. [DOI:10.1177/0095798418771807]

[22] Crocker L, Algina J. Introduction to Classical and Modern Test Theory. 1st ed. New York: Holt, Rinehart and Winston; 1986.

[23] Hill BD. The sequential Kaiser-Meyer-Olkin procedure as an alternative for determining the number of factors in common-factor analysis: A Monte Carlo simulation. [Doctoral Dissertation]. Oklahoma State University; 2011.

[24] Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016;15(2):155-63. [DOI:10.1016/j.jcm.2016.02.012]

[25] Pampel F. Toward Evidence-Based Practice in the Speech Audiometric Assessment of Linguistic Minorities. [BA Thesis]. Swarthmore College, PA; 2009.

[26] Ashoor AA, Prochazka Jr T. Saudi Arabic speech audiometry for children. Br J Audiol. 1985;19(3):229-38. [DOI:10.3109/03005368509078977]

[27] Mahdavi ME, Rabiei A. Psychometric function characteristics of Persian consonant vowel-consonant words. Aud Vestib Res. 2021;30(1):50-5. [DOI:10.18502/avr.v30i1.5311]

[28] Hirsh IJ, Davis H, Silverman SR, Reynolds EG, Eldert E, Benson RW. Development of materials for speech audiometry. J Speech Hear Disord. 1952;17(3):321-37. [DOI:10.1044/jshd.1703.321]

[29] McBride GB, Ellis JC. Confidence of compliance: A Bayesian approach for percentile standards. Water Res. 2001;35(5):1117-24. [DOI:10.1016/s0043-1354(00)00536-4]

[30] Domino G, Domino ML. Psychological Testing: An Introduction. 2nd ed. New York: Cambridge University Press; 2006.