

# Development of an Arabic speech intelligibility test for children

Samia E. Bassiouny<sup>a</sup>, Mona A. Hegazi<sup>a</sup>, Jilan F. Nassar<sup>a</sup>, Mohsen A. Ibrahim<sup>b</sup>, Aml S. Saber<sup>a</sup> and Ahmed Abdel Hamid<sup>a</sup>

<sup>a</sup>Unit of Phoniatrics, Department of Otolaryngology, Faculty of Medicine, Ain Shams University, Cairo and  
<sup>b</sup>Department of Psychometry, Faculty of Arts, El Minya University, El Minya, Egypt

Correspondence to Jilan F. Nassar, MD, Unit of phoniatrics, Department of Otolaryngology, Faculty of Medicine, Ain Shams University, Cairo, Egypt  
Tel: + 01001433625;  
e-mail: jilan\_nassar@yahoo.com

**Received** 9 March 2013

**Accepted** 7 March 2013

**The Egyptian Journal of Otolaryngology**  
2013, 29:202–206

## Objectives

Speech intelligibility is usually expressed as a percentage of spoken words, sentences, or phonemes correctly identified by a listener or a group of listeners when spoken by a talker or a number of talkers. Communicative disorders could have a deleterious effect on speech intelligibility. To date, there are no Arabic tests measuring speech intelligibility.

## Aim

It is necessary to develop and standardize an objective Arabic speech intelligibility test that can be used to estimate the degree of deterioration of speech intelligibility in communicative disorders as an attempt to develop a valid and a reliable tool to assess the efficacy of different therapy programs for different communicative disorders.

## Patients and methods

This study was carried out on 200 children with an age range of 4–12 years whose language skills were either fully developed or equivalent for age. All children were selected randomly from the outpatient clinic of phoniatrics and had any of five selected speech disorders affecting speech intelligibility. Each child included in the study was subjected to two evaluations: a subjective rating of the child's speech intelligibility and the developed Arabic speech intelligibility test, which is meant to be an objective measure.

## Results

The results showed a highly significant correlation between the scores of the Arabic speech intelligibility test and the average scores given by raters, the average scores for words, and the average scores for sentences (0.92, 0.98, and 0.84, respectively).

## Conclusion

The developed test is proven to be valid and reliable for measuring speech intelligibility in children and could be categorically classified according to ranges of severity.

## Keywords:

Arabic test, communicative disorders, objective, rating, reliability, speech intelligibility, validity

Egypt J Otolaryngol 29:202–206  
© 2013 The Egyptian Oto - Rhino - Laryngological Society  
1012-5574

## Introduction

Speech intelligibility is defined as the degree to which the acoustical signal is understood by a listener. Speech intelligibility is usually expressed as a percentage of words, sentences, or phonemes correctly identified by a listener or a group of listeners when spoken by a talker or a number of talkers [1].

Speech intelligibility is affected by: the experience of the listener, the context in which the sentence was spoken, the visibility of the speaker, the phonological and syntactic difficulty of the material spoken, and the familiarity of the speaker with the material spoken. Monsen [2] discussed the magnitude of these influencing factors on speech intelligibility. He considered the listener's experience as the best known factor that can influence the intelligibility of speech.

Yorkston *et al.* [3] considered intelligibility to be the cumulative effect of the impairment and any compensatory strategies. Many communicative disorders have a deleterious effect on intelligibility.

Paralytic dysphonia [4], spasmodic dysphonia [5], and alaryngeal communication [6] are vocal disorders that may affect speech intelligibility. Fluency disorders such as stuttering [7] and cluttering [8] and impaired resonance [9] and motor speech disorders such as dysarthria (DYS) [10] and apraxia [11] are among the speech disorders that might impair speech intelligibility. In delayed language development [7] and dysphasia [12], speech intelligibility may also be affected.

An accurate estimation of speech intelligibility using an objective tool is important because it is one of the outcomes of therapy against communicative disorders. Improvement of speech intelligibility may sometimes be a unique goal in the treatment of some of these disorders such as DYS or alaryngeal speech.

## Patients and methods

### Patients

This study was carried out on 200 children within an age range of 4–12 years. All children were randomly selected

from the outpatient clinic of phoniatrics and had any of the following disorders: developmental phonological disorders, hearing impairment, velopharyngeal incompetence, disfluency disorders, and DYS.

These disorders were included because they are known to have a negative impact on the degree of speech intelligibility. Any child within the age range of 4–12 years whose language skills were either fully developed or equivalent for age (evidenced by a language test) and who had average intelligence quotient was included in the study. Any child whose language was considered delayed for age or who had a visual perception problem was excluded from the study.

## Methods

Each child included in the study was subjected to two evaluations:

### Subjective rating of the child's speech intelligibility

This was done by three individuals, including the clinician (phoniatrician) (D) and two outsiders who were not familiar with the child (r1, r2). The outsiders were neither phoniatricians nor logopedists, and thus the element of past experience with the child's speech and the nature of the disorder was excluded. After establishing rapport with the clinician, the child was asked to talk about a certain topic or activity and discuss it. Thereafter, the two external raters were admitted to the room and allowed to listen to the child's speech without prior knowledge about the topic the child was discussing about. All three raters were asked to rate for the child's intelligibility along a five-point scale [13].

Thereafter, the clinician administered the test to the child, which is a mixed word and sentence intelligibility test. The test is considered as a modified version of The CID Picture SPINE test [14].

### Test application

This Arabic speech intelligibility test is meant to be an objective measure. The test is composed of 100 cards carrying 50 pictures (each picture is repeated twice). The pictures are structurally organized into three sets as follows:

- (1) Set A includes 20 pictures of monosyllabic words that start with: bilabial, nasal, epicodental, and supra-alveolar consonants.
- (2) Set B includes 20 pictures of monosyllabic words that start with: dorsopalatal, uvular, velar, and pharyngeal consonants.
- (3) Set C includes 10 pictures indicating simple action-verb sentences.

The pictures of each set are shuffled carefully and presented one by one to the child who is asked to name what is in the picture. Thereafter, the clinician will write down what he heard from the child in the clinician response form. The Arabic speech intelligibility test is designed to provide an estimation of the overall speech

intelligibility of children by providing a total score in terms of percentage.

The child's task is to say the pictured word correctly enough that the listener can decide which word, among the 20 or 10 possible choices, he/she was trying to say. Although the examiner knows the set of 20 (or 10 in sentences) possibilities, the examiner does not know which word the child will try to say on any given test item. Therefore, the examiner does not have to evaluate how a word was said but only whether the quality of articulation was functionally good enough to make the word understandable.

A sample of 30 children within the same age group range of 4–12 years was tested using the Arabic speech intelligibility test and then retested with the same test after an interval of 3 weeks. A correlation between the results of both tests was calculated using the  $\kappa$ -test for agreement, and the results indicated a highly significant correlation and reliability of the test.

## Results

The distribution of the children as per the five disorders affecting speech intelligibility was as follows: 96 children with developmental phonological disorders, 33 children with hearing impairment, 34 children with velopharyngeal incompetence, 29 children with stuttering, and eight children with DYS.

The scores of all three raters showed a highly significant correlation. The average value represents the subjective score (Table 1).

Table 2 shows the correlation between the mean score of the Arabic speech intelligibility test and that of the subjective test for all patients according to the disorder.

**Table 1 Correlation among the raters**

Raters	r2	r1
r1	$r=0.76^{**}$ (HS, $P<0.01$ )	–
D	$r=0.87^{**}$ (HS, $P<0.01$ )	$r=0.82^{**}$ (HS, $P<0.01$ )

\*\*Correlation is highly significant at  $P<0.01$ .  
HS, highly significant.

**Table 2 Correlation between the subjective scores and the Arabic test scores according to disorders**

Disorder	Mean score of raters (the subjective test)	Mean score of Arabic speech intelligibility test	$r$
All patients	3.88	79.85	0.92**
DPD	4.05	82.54	0.93**
HI	2.96	62.94	0.83**
VPI	3.66	78.09	0.87**
Stuttering	4.77	97.24	0.91**
DYS	3.28	65.75	0.98**

DPD, developmental phonological disorder; DYS, dysarthria; HI, hearing impairment; VPI, velopharyngeal incompetence.

\*\*Correlation is highly significant at  $P<0.01$ .

**Table 3 Correlation between the percentage scores of the Arabic intelligibility test and the mean values given by the raters according to age**

Age group	Mean score of raters	Scores of Arabic intelligibility test	<i>r</i>
All patients ( <i>n</i> =200)	3.88	79.85	0.92**
4–6 years ( <i>n</i> =137)	3.71	76.77	0.91**
7–9 years ( <i>n</i> =48)	4.28	89.80	0.94**
10–12 years ( <i>n</i> =15)	4.07	78.33	0.96**

\*\*Correlation is highly significant at  $P < 0.01$ .

The correlation proved to be highly significant ( $P < 0.01$ ), with the highest correlation obtained in the age group of 10–12 years, followed by the age group of 7–9 years and then the age group of 4–6 years (Table 3). Table 3 shows correlation between the two test scores in the three age groups.

Test reliability was measured by the test–retest method, which used the ‘Kappa test for agreement’, and the scores obtained from 30 children after a 3-week interval are shown in Table 4. The results of the  $\kappa$ -test for agreement for test and retest of 30 patients indicated a highly significant correlation between the results of the test and retest ( $P < 0.01$ ) for all items assessed.

Table 5 shows the correlation between the total score of the Arabic speech intelligibility test and the average score given by raters, the average scores for words, and the average scores for sentences.

The results showed that:

- (1) The correlation between the total score of the Arabic intelligibility test and the subjective score (average score given by raters) is highly significant.
- (2) The correlation between the total score of the Arabic intelligibility test and that for words is highly significant.
- (3) The correlation between the total score of the Arabic intelligibility test and that for sentences is highly significant.

The highly significant correlation between the scores of the Arabic speech intelligibility test and the subjective test scores (which indicate different severity levels of speech intelligibility) allows the final intelligibility scores of the test to be revealed (if desired) in a categorical classification as shown in Table 6 and Fig. 1.

## Discussion

This study presents a design of an objective intelligibility test that can be used for Arabic speaking children in the age range of 4–12 years. This Arabic speech intelligibility test is meant to provide an estimation of the overall speech intelligibility of children. The following are the characteristics of the Arabic speech intelligibility test,

which make it ideally suited for clinical use in children with impaired speech:

- (1) It is easy and practical to administer.
- (2) It is reliable from test to retest.
- (3) It provides a quantitative index of how intelligible a child’s speech is in common communication disorders. It also provides a categorical value of speech intelligibility.

For determining the reliability of this test, the test–retest method was used because the most obvious method for finding the reliability of test scores is repeating the test on a second occasion [15]. The reliability coefficient in this case is simply the correlation between the scores obtained by the same individual on the two administrations of the test. The reliability coefficient expresses the degree of consistency in measurement of the test scores. The reliability coefficient ( $r$ ) ranges from 1.00 (indicating perfect reliability) to 0.00 (indicating absence of reliability) [16].

Lillvik *et al.* [17] in their study on normal speakers and speakers with different types and degrees of DYS assessed speech intelligibility by using qualitative (subjective) and quantitative (objective) measures; they found that the results of that assessment gave a reliable quantitative measure of word and sentence intelligibility in patients with dysarthric speech.

The validity is usually tested by comparing new test scores with previous test or tests scores measuring the same phenomenon. However, there are no Arabic tests measuring speech intelligibility. That is why the scores of the Arabic speech intelligibility test were correlated with those of the five-point scaling subjective test. The  $r$  value was 0.92, indicating that the Arabic speech intelligibility test is highly predictive of speech intelligibility. In their study, Beerends *et al.* [18] reported that there is a correlation between the objective and subjective tests of speech intelligibility of about 0.90 and recommended using objective tests as they are less time consuming compared with subjective tests.

In contrast, Monsen [19] after obtaining a correlation of 0.86 between the patients’ scores on the SPINE test and those obtained on an open-set sentence intelligibility test reported that the SPINE test is highly predictive of speech intelligibility.

According to Schiavetti [20], word identification tasks are considered valid and can be reliable. Validity of the Arabic speech intelligibility test can be assessed by several ways – by studying the total scores of the Arabic speech intelligibility test for correlation with the subjective score test (score given by raters) and by studying the total scores of the Arabic speech intelligibility test for correlation with its subtests (words and sentences). The results showed that:

- (1) The correlation between the Arabic intelligibility test and the subjective score (score given by raters) is highly significant.
- (2) The correlation between the Arabic intelligibility test score and those for words is highly significant.

**Table 4 Results of test standardization**

Test items	r	Test items	r	Test items	r
1. بحر	1.00	18. شاي	0.84	35. قطر	1.00
2. بنت	0.87	19. شمس	0.80	36. فرد	0.92
3. بيت	0.89	20. ريش	0.93	37. قفل	0.91
4. بوز	0.75	21. كاس	0.73	38. حوض	0.92
5. بيض	0.93	22. كلب	0.85	39. حبل	0.78
6. موز	0.79	23. كتف	0.93	40. ايد	0.78
7. مشط	1.00	24. كم	0.87	41. ولد بيركب عجلة	0.84
8. فول	0.76	25. كوع	0.93	42. بنت بتلعب كورة	0.85
9. فيل	0.90	26. جون	0.80	43. بنت بتسمع راديو	0.93
10. فار	0.75	27. جيب	0.93	44. ولد بيتفرج علي التلفزيون	0.93
11. ورد	0.84	28. خذ	0.79	45. ولد ببليس هدومه	0.70
12. ورد	0.91	29. خس	0.79	46. بنت بتاكل برتقان	0.83
13. ديل	0.79	30. خيط	0.82	47. ولد ببشوط الكورة	0.81
14. درج	0.87	31. عيش	0.90	48. ولد ببشرب مية	0.76
15. نار	0.91	32. عش	0.85	49. بنت بتقطع تورثة	0.78
16. نور	0.78	33. عقد	0.91	50. بنت بتمسك قطة	0.71
17. شعر	0.90	34. عين	0.84		

All r values are highly significant.

**Table 5 Correlation between the total score of the Arabic speech intelligibility test and the average score given by the raters, the average scores for words, and the average scores for sentences**

	Average of three raters' scores	Words subtest	Sentences subtest
Mean value	3.88	63.83	16.02
r value	0.92**	0.98**	0.84**
Significance	$P < 0.01$	$P < 0.01$	$P < 0.01$

Mean value of the total score of Arabic intelligibility test = 79.85.

\*\* $P < 0.01$  is highly significant.

(3) The correlation between the Arabic intelligibility test scores and those for sentences is highly significant.

The fact that the correlation produced highly significant results (Table 5) proves that the test is a valid measure of speech intelligibility.

### Final test scoring

The Arabic speech intelligibility test provides the examiner with a final score in terms of percentage, which is meant to be an objective measure of a child's speech intelligibility. This can be used for comparing speech intelligibility before and after intervention, and thus provides numerical values that can be used for statistical comparisons. On the basis of our results, the score may be changed into a categorical value of unintelligible speech (0–29%), poorly intelligible speech (30–50%), fairly intelligible speech (51–66%), good intelligibility (67–84%), and excellent intelligibility (85–100%).

### Arabic speech intelligibility test

- (1) It is proven to be valid and reliable.
- (2) It can be readily used for children and adults who cannot read.
- (3) It provides a percentage score that can be changed into a categorical score (% scores are important for research studies, whereas categorical scores are more clear to parents and families of speech disorder patients).

**Table 6 Categorical classification of the scores**

Score of the subjective test	% range of the Arabic speech intelligibility test	Severity
1	0–29	Unintelligible speech
2	30–50	Poor intelligibility
3	51–66	Fair intelligibility
4	67–84	Good intelligibility
5	85–100	Excellent intelligibility

**Figure 1**

The categorical values of the Arabic speech intelligibility test.

- (4) It is easy to apply and is not time consuming.
- (5) It does not need patient audio recordings.
- (6) It does not need panels of listeners or listeners with very high linguistic abilities.

### Acknowledgements

#### Conflicts of interest

There are no conflicts of interest.

### References

- 1 Kent RD. Speech intelligibility. In: Yoder DE, Kent RD, editors. *Decision making in speech – language pathology*. Philadelphia: Decker; 1988. pp. 140–143.
- 2 Monsen RB. The oral speech intelligibility of hearing-impaired talkers. *J Speech Hear Disord* 1983; 48:286–296.
- 3 Yorkston KM, Beukelman DR, Tice R. *Phoneme intelligibility testing for windows*. Lincoln, communication disorders software. Lincoln, NE: Tice Technology Services; 1999.
- 4 Stevens KN, Nickerson RS, Rollins AM. Suprasegmental and postural aspects of speech production and their effect on articulatory skills and intelligibility. In: Hochberg I, Levitt H, Osberger MJ, editors. *Speech of the hearing impaired: research, training and personnel preparation*. Baltimore: University Park Press; 1983. pp. 35–51.
- 5 Bender BK, Cannito MP, Murry T, Woodson GE. Speech intelligibility in severe adductor spasmodic dysphonia. *J Speech Lang Res* 2004; 47:21–32.

- 6 Liu H, Wan M, Ng ML, Wang S, Lu C. Tonal perceptions in normal laryngeal, esophageal, and electrolaryngeal speech of Mandarin. *Folia Phoniatri Logop* 2006; 58:340–352.
- 7 Justice L. *Communication science and disorders: an introduction*. 1st ed. Upper Saddle River, NJ: Pearson Education Inc.; 2006.
- 8 St. Louis KO, Myers FL, Faragasso K, Townsend PS, Gallaher AJ. Perceptual aspects of cluttered speech. *J Fluency Disord* 2004; 29: 213–235.
- 9 Van Lierde KM, De Bodt M, Van Borsel J, Wuyts FL, Van Cauwenberge P. Effect of cleft type on overall speech intelligibility and resonance. *Folia Phoniatri Logop* 2002; 54:158–168.
- 10 Kent RD, Weismer G, Kent JF, Rosenbek JC. Toward phonetic intelligibility testing in dysarthria. *J Speech Hear Disord* 1989; 54:482–499.
- 11 Shuster LI, Wambaugh JL. Consistency of speech sound errors in apraxia of speech accompanied by aphasia. Poster Presented at the Clinical Aphasiology Conference, Orca's Island, WA; 2003.
- 12 Hardcastle WJ, Edwards S. EPG-based description of apraxic speech errors. In: Kent RD, editor. *Intelligibility in speech disorders*. Philadelphia: John Benjamin; 1992. pp. 287–328.
- 13 Flipsen P Jr. Measuring the intelligibility of conversational speech in children. *Clin Linguist Phonet* 2006; 20:303–312.
- 14 Monsen R, Moog J, Jears AE. *CID Picture SPINE*. St Louis, MO: Central Institute for the Deaf; 1988.
- 15 Anastasi A. *Psychological testing*. 7th ed. Upper Saddle River, NJ: Prentice Hall; 1997.
- 16 Wodrich DL. *Children's psychological testing*. 3rd ed. Baltimore, MD: Paul H. Brookes Publishing Co. Inc.; 1997.
- 17 Lillvik M, Allemark E, Karlström P, Hartelius L. Intelligibility of dysarthric speech in words and sentences: development of a computerised assessment procedure in Swedish. *Logoped Phoniatri Vocol* 1999; 24:107–119.
- 18 Beerends JG, Van Buuren R, Van Vugt J, Verhave J. Objective speech intelligibility measurement on the basis of natural speech in combination with perceptual modeling. *J Audio Eng Soc* 2009; 57:299–308.
- 19 Monsen RB. A usable test for the speech intelligibility of deaf talkers. *Am Ann Deaf* 1981; 126:845–852.
- 20 Schiavetti N. Scaling procedures for the measurement of speech intelligibility. In: Kent RD, editor. *Intelligibility in speech disorders*. Philadelphia: John Benjamin; 1992. pp. 11–34.