Research Paper: The Design and Validation of an Azeri Phonological Test for Children Aged 4-6 Years

Zeynab Khoshhal1, Ali Jahan2, Majid Mirzaee3, Fatemeh Haresabadi4*

1. Department of Speech and Language Pathology, Tabriz University of Medical Sciences, Tabriz, Iran.
2. Neuroscience, Tabriz University of Medical Science, Tabriz, Iran.
3. Department of Speech and Language Pathology, Faculty of Rehabilitation, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran.
4. Department of Speech Therapy, School of Paramedical Sciences, Mashhad University of Medical Sciences, Mashhad, Iran.

* Corresponding Author:
Fatemeh Haresabadi, PhD.
Address: Department of Speech Therapy, School of Paramedical Sciences, Mashhad University of Medical Sciences, Mashhad, Iran.
Tel: +98 (917) 7948964
E-mail: haresabadi.f@mums.ac.ir

Objectives: Since there is no reliable and valid tool to assess all features of children’s phonological skills whose first language is Azeri, speech-language pathologists usually use informal or Persian instruments. So, it is essential to develop a tool in Azeri language for assessing articulation and phonological skills. For this reason, this study aimed to develop an Azeri Phonological Test (APT) and then determine the validity and reliability of this test in 4-6 years old children whose first language is Azeri.

Methods: Participants were 77 (41 boys and 36 girls; 4-6 years old) typically-developing children. They were selected by cluster sampling method. Firstly, Tabriz City in Iran was geographically divided into three sections, namely south, north, and center. Then, the kindergartens of each area were randomly selected for sampling. The participants were assigned into 2 groups: 48-60 months (n=38) and 61-72 months (n=39) old. To assess the phonological skills, nine stories were designed. An expert panel with 15 speech-language pathologists investigated the content validity of the stories. Then the psychometric characteristics of the APT, including reliability (test-retest reliability, inter-rater, and internal consistency) and validity (construct –known group validity), were assessed. The statistical tests used were intraclass correlation coefficient (intra-rater reliability and test-retest), Kuder-Richardson 21 (internal consistency), and the independent samples t-test (known-groups validity).

Results: The content validity of the whole test was 0.91. Construct validity evaluation showed that age affects the test scores but not gender. There was a significant correlation between the results of test-retest and inter-rater reliability, and also, the test had a high internal consistency.

Discussion: Based on the findings, APT is a reliable and valid test for assessing the abilities of 4-6 years old children in terms of the production of sounds in the context of retelling stories.
1. Introduction

Children with speech sound disorders (SSD) comprise a large part of the clients who refer to speech therapy centers [1]. SSD refers to significant impairments in sounds or sound patterns production beyond the age of acquisition [2]. Four subtypes of SSD have been recognized based on Dodd’s classification system of the children’s speech disorders: inconsistent phonological disorder, phonological delay, consistent phonological disorder, and articulation disorder [3]. Inconsistent phonological disorder is the label applied when a child produces the same word differently; in other words, a range of different atypical errors every time the child produces it [4]. Phonological delay is used when children have speech production patterns seen in younger, typically developing children. Consistent phonological disorder is identified when a child produces unusual and idiosyncratic errors (specific to the child) and are not seen in typically developing children. An impaired ability to produce a perceptually acceptable type of particular phonemes, either in isolation or in any phonetic context, is known as an articulation disorder.

There is often no clear etiology for SSD [5]. These children form a heterogeneous group, differ in types of speech error patterns, severity, etiology, and other associated language difficulties [6]. Various prevalence rates of SSD have been reported in the literature [7]. SSD is a common communication disorder that affects 10%-15% of preschool children and 6% of school-aged children [8]. Also, it is estimated that the SSD prevalence is 13.1% in primary school students in Yazd City, Iran [9].

Speech sound disorders can have different side effects. It has been determined that the affected children show several problems, including reduced intelligibility, the risk for broader communication disorders, social stigma, and difficulties with acquiring phonological awareness [10, 11]. Phonological awareness has a positive correlation with literacy outcomes. Researchers have found that phonological deficits play an essential role in theories of both dyslexia and developmental language disorder [12]. These problems are a challenge for speech-language pathologists to deliver diagnostic and therapeutic services better. Furthermore, since there is a relation between aspects of speech and language development and the features of educational and social development, speech disorders of children should be accurately diagnosed and assessed to prevent future psychosocial, communication, and educational problems.

The purpose of phonological assessment procedures is to provide a sample of speech that reflects the child’s abilities in different contexts and situations to optimize assessment and treatment [13]. To study a child’s speech sounds, different contexts can be used, including connected speech, conversational speech, and single words [14]. Connected speech is commonly elicited through...
unbidden discussion, the retelling of a story, and oral reading of a text [5]. Compared to single word naming, spontaneous conversation tests are further enhanced by the accessibility of phonetic contexts, which are assumed to be important in phonological evaluation. In general, articulatory/phonological errors occur to a greater extent in conversation than naming words [13]. The spontaneous conversation is the best method for assessing connected speech. It can measure the child’s performance in the most natural conditions. In cases we could not gather this sample, the evaluation of the connected speech can be done by other procedures [15].

Although the story-retelling task can be different from the everyday conversational speech, it may use similar skills. In a story-retelling task, instant recall of a new story presented on-site by a clinician is required [16].

There are a variety of tools assessing the production of a speech sound at the level of a single word and connected discourse in various languages, including Arabic [17], English [18-20], European-Portuguese [21], Italian [22], Persian [14, 23, 24], and Turkish [25]. To choose the most suitable tool, different factors should be taken into accounts, such as the type of the instrument and its psychometric properties and the demographic features of the participants [26].

Also, cross-phonological studies of language acquisition show that there are only broad developmental patterns that can be described as universal. So the opinion that children are sensitive to the language of the environment is not a new matter. Although phonological inventions of children are not similar within and across languages, the similarities are seen in the main features [27]. Due to differences in the phonetic inventories or the frequency of particular phonetic and phonological properties, the results of other studies about a language cannot be generalized to other languages like Azerbaijani Turkish.

Azerbaijani Turkish or Azerbaijani language is a member of the Turkish language family spoken with two dialects in the Republic of Azerbaijan and the Azerbaijani regions of Iran. Azerbaijani Turkish as a spoken language is also spoken in eastern Turkey, southern Georgia, and Dagestan. Persian and Arabic have influenced the Azerbaijani Turkish language in the vocabulary, phonetic system, and grammar. In Iran, Tabriz, Urmia, and Ardabil are the largest cities where Azerbaijani Turkish is spoken, respectively. Azerbaijani has 9 vowels and 24 consonants. This language has a relatively complex syllable structure [28].

Unfortunately, no study has been done on phonological development in Azeri-speaking children. However, such a study is available for Turkish (known as Istanbul Turkish), one of the Turkish dialects. The results of Turkish language showed that its consonant development seems more rapid and error patterns are more predictable [27].

As it was mentioned, there is only one tool for the phonology assessment in Turkish language [25]. Since the sound system of the two languages is different so that Turkish version cannot be used for Azari.

Because there are no reliable and valid instruments to assess all features of children’s phonological skills whose first language is Azeri, speech-language pathologists usually use informal or Persian instruments. According to what was mentioned above, it is essential to develop a tool in Azeri for assessing articulation and phonological skills.

So, the present research aimed to design and validate an instrument for assessing speech sound production in Azeri-speaking children aged 4–6 years through story retelling and incorporated all Azeri consonants. The reason for selecting this age group was that speech disorders are most prevalent among children within this age range [29].

2. Materials and Methods

Study participants

The participants were 77 (41 boys and 36 girls; 4-6 years old), typically-developing children. Sampling was done by cluster sampling method. At first, Tabriz was geographically divided into three groups, namely south, center, and north. Then, the kindergartens of each area were randomly selected for sampling. Next, children were randomly chosen from the kindergartens. The children were assigned to two groups: group one (48-60 months old) and group two (61-72 months old). Table 1 summarizes the demographic features of the subjects (n=77) in each group with regard to gender and age. The inclusion and exclusion criteria were evaluated by interviewing participants and their parents and filling out an informal checklist of demographic and health characteristics. The inclusion criteria consisted of being monolingual, speaking Azeri as the first language, and being 48-72 months old. The participants had no problems like visual, hearing, emotional, physical, neurological, and speech. The diagnosis was according to the evaluation and examinations of an experienced language and speech pathologist, reports given by parents and teachers in kindergartens, and the children’s medical records.
Test design and psychometric properties

The test was developed, and its psychometric properties were evaluated based on the following criteria.

Item generation

In this step, nine stories were designed and written by 4 speech and language pathologists. The names of stories were as follows: socks /jurab/, book /kitab/, puppet /ᴂrusᴂk/, painting /nᴂGaʃi/, flower /gol/, car /maʃin/, candle /ʃᴂm/, balloon /badkonᴂk/, and umbrella /ʧᴂtr/. These stories incorporated all Azeri consonants, and for each consonant, more than one chance of expression was considered. The children had already been accustomed to the topics of stories and vocabularies that were provided for them. On average, the sentences in stories were 2-5 words long. The subjects of the stories were known to and liked by the children. It was tried to include all the sounds in the different positions of the words (three positions: initial, medial, and final). The sentences of the stories were not long so that a child could easily repeat them after the examiner. After developing the stories, a professional graphic artist was asked to design colorful and cartoon-style illustrations for each story in the A4 size. The aim of providing illustrations for stories was to help children in retelling them. Every story had four pictures.

For readers wishing to see the pictures used in this study, they are referred to the weblog, http://www.testtolideAzeri.blogfa.com.

Content and face validity

After the test was developed, the content validity of stories was determined using the viewpoints of experts. This panel consisted of 15 speech and language pathologists. The experts were chosen based on their experiences in the field of children’s speech sound disorders. To determine the quantitative content validity index (CVI) of each story, panel members were asked to judge whether each story is proper in terms of four indexes: “picture-content relevancy”, “concreteness”, “complexity”, and “phonological representativeness.” These indices were scored from 0 (the least) to 2 (the most). The CVI values of ≥0.8 were regarded as having sufficient content validity [30].

To ensure that participants can easily retell stories and the pictures are proper for the age group, 10 typically-developing children aged 48-72 months were chosen to check the face validity of the test.

Construct validity

Known-groups validity as one type of construct validity was investigated in this research. This type of validity investigates the ability of the test to distinguish the groups according to the studied construct. To investigate known-groups validity, the APT was administered to 77 participants, and then the participants’ responses were analyzed with regard to their age (48-60 months and 61-72 months) and gender.

Reliability

Test reliability was evaluated using three methods: test-retest, inter-rater reliability, and internal consistency. Test-retest reliability was examined by applying APT twice with a 2-week interval on 15 children, randomly chosen from the 77 subjects. The inter-rater reliability was established via transcription and scoring of the recorded audio tracks of 15 children by two speech and language pathologists (a PhD and an MS in Speech Therapy). The internal consistency was obtained through calculation of the relationships among the items by Kuder–Richardson 21.

Administration and scoring of the test

A language and speech pathologist conducted the APT in a friendly, individually, and calm environment in the nursery school which lasted approximately 20-30 minutes. The examiner showed the stories to the child and explained what needed to be done. After telling each sentence of each story, the participant was asked to retell it. The child’s voice was recorded for further evaluation. Speech samples were recorded for transcription using a voice recorder, and a laptop scoring responses were made as follows: if the phoneme is correctly spelled, the score was zero, and if the subject did not pronounce the phoneme or mispronounce it, the score was 1. So, the children who could accurately produce all the sounds received a final score of 0.

Statistical analysis

The findings were analyzed in the SPSS software, v. 22. The normality of the data distribution was examined using the Kolmogorov-Smirnov test. In the present research, the descriptive measures (Mean±SD) were investigated as well. Statistical tests used included intraclass correlation coefficient (test-retest and intra-rater reliability), Kuder-Richardson 21 (internal consistency), and the independent sample t-test (known-groups validity). A P value of less than 0.001 was considered statistically significant.
3. Results

Content and face validity

In total, 15 experts approved the content validity of the APT. As shown in Table 2, the CVI of each story was between 0.8 and 1, and that of the whole test was 0.91. All were within an acceptable level.

Also, children could retell all sentences after being told, so it can be said that the illustrations were helpful for them. Hence, the face validity of this test was considered good.

Construct Validity

As presented in Table 3, the mean values of female and male children regarding speech sound production of story retelling had no significant difference. Nevertheless, based on Table 3, a noticeable difference was observed between the mean values of speech sound production of the two age groups: there was a reduction in the mean values of children in the older group.

3.3. Reliability

The intraclass correlation coefficient was used to obtain test-retest reliability, and the scores of participants at two separate evaluations were found to be significantly correlated (r=0.901, P<0.001). Also, based on the findings of the inter-rater reliability, the scores of the two various raters had a significant correlation (r=0.98, P<0.001). Moreover, the result of Kuder–Richardson test 21 about the internal consistency indicated a strong correlation among the test items (r=0.89, P<0.001).

4. Discussion

The present research aimed to design and validate an Azeri phonological test for children of 4 to 6 years old whose first language is Azeri.

The speech sound production is evaluated most frequently by tools that include the production of the sounds at the word level. It is important to remember that speech sound production evaluation at a single word level cannot provide adequate information regarding the speech sound skills of children who have Speech Sound Disorder.

Table 1. Descriptive statistics of subjects in terms of age and gender (n=77)

<table>
<thead>
<tr>
<th>Age Range</th>
<th>Mean±SD</th>
<th>Female (n)</th>
<th>Male (n)</th>
<th>Total (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>48-60 mon</td>
<td>53±1.14</td>
<td>18</td>
<td>20</td>
<td>38</td>
</tr>
<tr>
<td>61-72 mon</td>
<td>64±0.15</td>
<td>18</td>
<td>21</td>
<td>39</td>
</tr>
</tbody>
</table>

Table 2. Content validity of azeri phonological test

<table>
<thead>
<tr>
<th>Story</th>
<th>Picture-Content Relevancy</th>
<th>Concreteness</th>
<th>Complexity</th>
<th>Phonological Representativeness</th>
<th>CVI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>2</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>3</td>
<td>1.00</td>
<td>0.85</td>
<td>1.00</td>
<td>0.85</td>
<td>0.92</td>
</tr>
<tr>
<td>4</td>
<td>1.00</td>
<td>0.85</td>
<td>1.00</td>
<td>1.00</td>
<td>0.96</td>
</tr>
<tr>
<td>5</td>
<td>0.85</td>
<td>0.71</td>
<td>0.85</td>
<td>0.85</td>
<td>0.81</td>
</tr>
<tr>
<td>6</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>7</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.85</td>
<td>0.96</td>
</tr>
<tr>
<td>8</td>
<td>0.85</td>
<td>0.85</td>
<td>0.71</td>
<td>0.85</td>
<td>0.81</td>
</tr>
<tr>
<td>9</td>
<td>0.71</td>
<td>0.85</td>
<td>0.85</td>
<td>1.00</td>
<td>0.81</td>
</tr>
<tr>
<td>Total</td>
<td>0.93</td>
<td>0.85</td>
<td>0.93</td>
<td>0.93</td>
<td>0.91</td>
</tr>
</tbody>
</table>

CVI: content validity index.
der (SSD). Also, at this level, the impact of other sounds is not considered. Indeed, to evaluate comprehensively, speech sound production should be assessed by the analysis of connected speech [5]. Pinkerton reported the beneficial effects of the connected-speech sampling compared to the single word production. She found that obtaining a speech sample provides an average for the natural production with the influence of co-articulation [31].

In the test that was developed in this study, the speech sample is provided by using pictorial stories. Therefore, the production of speech sounds can be investigated in the context of connected speech. Connected speech sampling allows the therapist to assess children’s ability to produce sounds in everyday conversation. The speech sample can provide information about voice quality, fluency, intelligibility/understandability, prosody, and mean length of response [32, 33].

An essential characteristic of the APT is that the pictures were designed to extract all phonemes. The authors tried to consider several positions for each phoneme (Three positions: initial, medial, and final).

The study of content validity of the APT showed an agreement percentage of higher than 0.8 among experts for all stories, which is considered high. This finding indicates that the stories could effectively represent and capture the speech sound production qualities of 4-6 years old children whose first language was Azeri. In addition to this, the face validity of the test was reported as good. The results were in line with those of the previous related research performed on other languages [24, 34]. It was attempted to write stories that contained all Azeri consonants and incorporated words suitable for this age range. These criteria were consistent with those used by Goldman and Fristoe, who wrote a story that described a familiar event to children [19].

Examination of the effect of gender on the scores of subjects showed a significant difference between the mean scores of male and female children. Also, based on the results, the mean scores of children in the two age groups were significantly different, so that the older children scored lower than the younger children. In other words, the older children showed fewer errors and accurate production. Age is among the most important factors that influence speech sound production. As the oral motor function of children matures with age, the children gain the ability to produce speech sounds more correctly [35]. Comparing the scores of the two gender groups also revealed that female children performed better in this instrument in terms of speech sound production; however, the mean scores of male and female children were not significantly different. This result matched with several studies worldwide [17, 24, 36]. It was indicated that gender constituted a relatively small variance (10%-15%) in the speech sound production of children [36].

The test-retest reliability findings showed the relatively excellent reliability of APT. Based on these results, the reproducibility of APT increases over time. Furthermore, it was found that the scoring of two examiners had a high level of inter-rater reliability. Based on the assessment of the internal consistency of the APT, the total instrument had an excellent internal consistency. The results revealed that this test is an appropriate test for assessing phonological ability in children.

In the present study, some limitations must be considered before the generalization of the results. First, this research was done only on 4-6 years old children, and other age ranges were not assessed. Second, the abilities of children with SSD and normal children were not compared by the APT. Third, standard scores were not obtained in this research. So, further research needs to be conducted to examine the administration of this test on children with SSD and collect normative data on a large sample of children in various age groups.

5. Conclusion

Based on the present research findings, the validity and reliability of Azeri phonological test (APT) were con-
firmed. Hence, it can be used to assess the ability of 4-6 years old children in terms of the production of sounds in the context of retelling stories. Because there is no tool for this purpose in Azeri language, it is a valuable tool for clinicians and researchers to get a distinct profile of speech sound disorders.

Ethical Considerations

Compliance with ethical guidelines

The research protocol was confirmed by the Ethics Committee of Tabriz University of Medical Sciences, Tabriz (Code: IR.TABMED.REC.1396.242) and was conducted in 2018-2019 in Tabriz City, Iran. Also an informed consent was obtained from the parents of subjects before the initiation of the data collection, and the names and other identifiers were not used in the questionnaire to maintain the confidentiality of information collected from each participant.

Funding

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

Authors’ contributions

All authors equally contributed to preparing this article.

Conflict of interest

The authors declared no conflict of interest.

References


[17] Abou-Elsaad T, Baz H, El-Banna M. Developing an articulation test for Arabic-speaking school-age chil-
I ranian R ehabilitation Journal
December 2021, V olume 19, Number 4


[28] Sarikaya M. [Since the Old Turkish period, singular in Turkish has been used for second person imperative conjugation (Turkish)]. Türkülük Bilimi Araştırmaları. 2009; (25):145-57. https://dergipark.org.tr/en/pub/tubar/isssue/16966/177193