Development and measurement of psychometric properties of the Persian test of speech consistency in children with typical development

Article *in* Applied Neuropsychology · July 2020

DOI: 10.1080/21622965.2020.1786831

CITATIONS
0

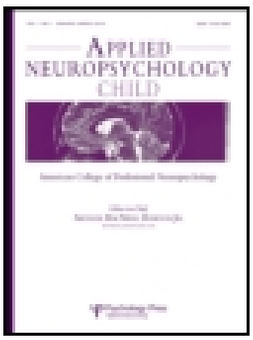
READS
67

3 authors:

Talieh Zarifian
University of Social Welfare and Rehabilitation Sciences
46 PUBLICATIONS **60** CITATIONS
SEE PROFILE

Akram Ahmadi
Babol University of Medical Sciences
23 PUBLICATIONS **42** CITATIONS
SEE PROFILE

Abbas Ebadi
Baqiyatallah University of Medical Sciences
596 PUBLICATIONS **3,455** CITATIONS
SEE PROFILE



Applied Neuropsychology: Child

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hapc20

Development and measurement of psychometric properties of the Persian test of speech consistency in children with typical development

Talieh Zarifian , Akram Ahmadi & Abbas Ebadi

To cite this article: Talieh Zarifian , Akram Ahmadi & Abbas Ebadi (2020): Development and measurement of psychometric properties of the Persian test of speech consistency in children with typical development, Applied Neuropsychology: Child, DOI: 10.1080/21622965.2020.1786831

To link to this article: https://doi.org/10.1080/21622965.2020.1786831

Published online: 10 Jul 2020.

APPLIED NEUROPSYCHOLOGY: CHILD
https://doi.org/10.1080/21622965.2020.1786831

Development and measurement of psychometric properties of the Persian test of speech consistency in children with typical development

Talieh Zarifian[a], Akram Ahmadi[b,c], and Abbas Ebadi[d,e]

[a]Department of Speech Therapy, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran; [b]Department of Speech Therapy, School of Rehabilitation, Babol University of Medical Sciences, Babol, Iran; [c]Mobility Impairment Research Center, Babol, Iran; [d]Behavioral Sciences Research Center, Life Style Institute, Faculty of Nursing, Baqiyatallah University of Medical Sciences, Tehran, Iran; [e]Faculty of nursing, Baqiyatallah University of Medical Sciences, Tehran, Iran

ABSTRACT

The current research pursues the study of developing the Persian test of speech consistency and investigating speech consistency in normally developing children in Iran. After developing the Persian test of speech consistency, the obtained speech samples of 317 participants (163 girls and 154 boys) were analyzed. This test of consistency includes 27 items. Experts' consensus over the characteristics of the Persian test of speech consistency was above 85%. Different age groups had statistically significant differences for the mean scores of the variable responses ($p < .001$), consistent correct ($p < .001$), and consistent incorrect responses ($p = .007$) in the Persian test of speech consistency. Also, there was a high value for the inter-rater reliability (ICC = .88, $p < .001$) as well as the moderate value for the test-retest reliability (ICC = .89, $p < .001$). The Persian test of speech consistency is regarded as a reliable and valid scale for measuring the speech consistency in Persian-speaking children. Future studies are proposed to investigate the speech consistency in Persian-speaking children with various kinds of speech sound disorders.

KEYWORDS

Inconsistency; Persian; test; validity; reliability

Introduction

Speech Sound Disorders (SSD), as a subgroup of speech disorders, are identified by the difficulty of producing one or more than one speech sound(s) (Ahmadi et al., 2019). The accuracy of speech sound production in children with SSD is less than that in children with normal speech and, subsequently, intelligibility is influenced by this condition (Lousada et al., 2014).

Several studies have reported different rates for the prevalence of SSD (Eadie et al., 2015; Keating et al., 2001; McKinnon et al., 2007) depending on diagnosis procedures and age range of the studied population. Considering the high heterogeneity of the speech and language skills, including types of speech sound errors, the severity of the disorder, and auditory perception in individuals with SSD, different dimensions of the speech, like sound production, intelligibility, stimulability, and consistency, should be measured (De Castro & Wertzner, 2011). Another point that should be considered in differential diagnosis of children with SSD is choosing the best intervention procedure in terms of diagnostic classification of SSD (Terband et al., 2019).

Assessment tasks in SSD

In general, the common tasks for the assessment of speech in children include two main groups: assessment of children in spontaneous speech and single word tests. It seems that evaluation in spontaneous speech shows the speech characteristics in a more realistic situation compared with the single word assessment. However, some drawbacks are considered for the speech assessment in the spontaneous speech level. Some children may avoid using words that contain speech sounds they cannot produce correctly or their speech sample may be inadequate. So, these children's spontaneous speech sample may not represent the complete image of their speech characteristics. On the other hand, there are some merits to using single word tests in clinical settings. Administration of these procedures is easy and using norm-referenced tests make it possible to compare the assessment outcome of the child with SSD and their peers in terms of age and sex. Identifying the words in the speech of children whose speech is highly unintelligible is so challenging. Therefore, administration of single word tests is more helpful in children with SSD. Given the differences in characteristics of speech sounds system across languages, the evaluation process needs to be specific for each language (Ahmadi et al., 2018). Despite similarities among them, all languages are universal in the development of phonology and language; these unique developmental characteristics are referred to as individual preferences (Tamer Abou-Elsaad et al., 2019).

Assessment of consistency

Speech inconsistency or intra-word variability is referred to as variability in producing one word from one token to

CONTACT Akram Ahmadi slp347@gmail.com Department of Speech Therapy, School of Rehabilitation, Babol University of Medical Sciences, Ganj Afrooz Ave, Babol, 47745-47176, I.R, Iran.

another (De Castro & Wertzner, 2011; Sosa, 2015). In literature, this term is occasionally referred to as variability or inconsistency (De Castro & Wertzner, 2011). This characteristic has been discussed as a potential clinical indicator of specific subtypes of SSD; inconsistent production of speech sounds (especially token to token inconsistency) has been associated with both childhood apraxia of speech (CAS) (American Speech and Hearing Association, 2007) and inconsistent phonological disorder. In both disorders, production inconsistency is typically defined as phonemically variable production of individual words from one token to the next (Crosbie et al., 2005).

Also, some studies reported the existence of high degrees of variability in the speech of normally developing children (Macrae, 2013; Macrae & Sosa, 2015; Sosa & Stoel-Gammon, 2006). The term intra-word variability is usually used for variability in the speech production of children with typical development (Sosa, 2015).

Speech inconsistency is generally analyzed in two levels: token to token and in the level of phonemes (Iuzzini-Seigel et al., 2017). Beside diverse assessment, different techniques have been utilized to measure the inconsistency of speech. In general, these techniques can be classified as acoustic, articulatory, and perceptual measures. Terband et al (Terband et al., 2019) described the variety of evaluation techniques in their structured tutorial to quantify the speech inconsistency in CAS. Literature witnesses different types of evaluation procedures suggested and used for each measure. They concluded that these techniques are complementary for obtaining accurate information about the nature of speech inconsistency in individuals with CAS. One subtype of perceptual measure that seems to be easy to use, reliable, and practical is token to token assessment. This is the most common procedure that allows SLPs to have high consensus in measuring speech consistency (Betz & Stoel-Gammon, 2005; De Castro & Wertzner, 2011). Dodd (1995) introduced and suggested this method for the first time and developed a scale to assess the inconsistency. The scale characteristics of this test are provided in the next section of the introduction. In the token to token assessment of inconsistency, the repeated productions of single words in the same context are compared. If a child with SSD always produces the word, in the same manner, he or she will be categorized in a consistent group and if he produces that word variably in multiple productions, he or she will be considered as inconsistent (Holm et al., 2007).

The obtained information of measurement of speech inconsistency helps Speech-Language pathologists (SLPs) to identify the skills that should be considered in the intervention programs of children with SSD.

Previous works on speech consistency in typically developing children

Up to now, speech consistency has been investigated in several studies (Burt et al., 1999; De Castro & Wertzner, 2011; Holm et al., 2007; Kenney & Prather, 1986; Macrae, 2013; Macrae & Sosa, 2015; Martikainen et al., 2019; Sosa & Stoel-Gammon, 2006) (the main results are summarized in Table 1). As provided in Table 1, the majority of these studies have been carried out for English-speaking children.

In the study by Kenney and Prather (Kenney & Prather, 1986), speech consistency of some consonants was examined in children with normal development. They used 72 two-word phrases to assess the consistency. These phrases included the following nine phonemes:/l, ʃ, θ, t, f, b, r, tʃ, and s/. The participants were asked to imitate these phrases with delay. The results of their study revealed that the position of the speech sound and age had significant effects on speech consistency.

In addition, four published papers investigated the speech consistency using Inconsistency Assessment (IA) in English speaking children (Burt et al., 1999; Holm et al., 2007; Macrae & Sosa, 2015; Sosa, 2015). This tool includes 25 words with two to four syllables and covers most of English consonants and vowels (Dodd, 1995).

Burt et al assessed the speech consistency of preschool English speaking children (Burt et al., 1999) and showed that the production consistency increased in older children. They reported that boys and girls had no significant difference in terms of speech consistency. Sosa (2015) examined the intra-word variability in children with typical development in the age range of 2 to 3 years. She reported 78% variability in the age range of 30–35 months and this rate dropped to 57% in 42–47 month children. The interaction of speech variability and sex was not reported in this research. Holm et al (Holm et al., 2007) studied the speech variability of the children between 3.0 to 6.11 years. They found that values of consistency were greater in older children and girls had more consistent speech compared with boys. Macrae and Sosa (Macrae & Sosa, 2015) investigated the predictors of the speech variability in the children with normal development of speech and language in the age range of 30 to 50 months. They studied the effects of some variables like receptive and expressive language, participants' abilities in the articulation of speech sounds, and age on speech consistency. Goldman Fristoe Test of Articulation 2, Expressive Vocabulary Test-2, Peabody Picture Vocabulary test-4, and IA were used in the assessment procedure. They demonstrated that the only variable that moderately correlated with consistency was expressive vocabulary.

Stackhouse and Williams (2000) studied the consistency of speech by the diadochokinetic tasks, naming, and repetition of non-words. They showed that the type of task did not have any significant effect on the consistency of the responses and speech consistency progressed in older children.

De Castro et al (De Castro & Wertzner, 2011) described the inconsistency of the speech in Brazilian Portuguese speaking children. They used Brazilian Portuguese version of speech inconsistency index (BP SIT), involves 25 target words for eliciting all of the speech sounds of the Brazilian Portuguese to assess speech inconsistency. The word length of this test ranges from 1 to 4 syllables, which varies in terms of syllable structure of Brazilian Portuguese. They designed this index according to the criteria reported in a

Table 1. The summary of the published studies on speech consistency in typically developing children.

Author/s	Task	Study Design	Age of the participants (month)	No. of the participants	Language	Country	findings
Kenney and Prather (1986)	Delayed imitation	Cross-sectional	30–60	360	English	USA	Speech consistency improved in older children. The mean of number of errors in consistency was 35.10 in thirty month old children and decreased to 14.60 in 60 month old children. The mean value of the errors was higher in boys than girls.
Burt et al. (1999)	Picture naming	Cross-sectional	46–58	57	English	UK	The mean value of inconsistency was 5.7 in 46–51 months and was 3.8 in 52–58 months. The girls and boys performed similarly in assessment of speech consistency.
Stackhouse and Williams (2000)	diadochokinetic tasks, naming, and repetition of non-words	Cross-sectional	36–60	30	English	UK	Speech consistency was 84.5% in 3 years old children and reached to 91.6% in 5 years old children.
Sosa and Stoel-Gammon (2006)	Picture naming	Longitudinal	12–24	4	English	USA	Variability is a common phenomenon in typically developing children at 12–24 month old children.
Holm et al. (2007)	Picture naming	Cross-sectional	36–83	409	English	UK	Typically developing children produced consistently correct speech at 3 years old at the value of more than 75%. Girls had more consistent production than boys.
(Sosa, 2015)	Picture naming	Cross-sectional	30–47	33	English	USA	The mean value of the variability in 30–35 month children was 77% that reduced to 57% in 42–47 month children.
Martikainen et al. (2019)	Picture naming	Cross-sectional	36–83	80	Finnish	Finland	The median of 75% for 36 month old children that increased to 95% in the older children.

previous study in English (Crosbie et al., 2005). They compared the speech inconsistency in children with typical phonology and with phonological disorders and found that speech inconsistency is an efficient marker to differentially diagnose phonological disorders. Martikainen et al (Martikainen et al., 2019) studied the word consistency of Finnish speaking children. The participants were asked to name 20 target words three times in a single assessment session. The assessed words had 2 to 5 syllables and their syllable structures were those that are frequently used by 38–64 months children. They reported a high value for the consistent correct responses even for 38-month children.

Extant tests for assessment of speech consistency

During our study, we found two published tests for the assessment of speech consistency in other languages. As mentioned in the previous section, IA is the routine scale for the consistency evaluation of English speakers, that was developed by Dodd (1995). The properties of this scale were explained previously. The other test was developed by De Castro and Wertzner for Brazilian Portuguese speakers, whose properties were described previously. Apart from these two, we could not find any other test for the assessment of speech consistency published in the literature.

Phonological features of the Persian language

Persian or Farsi, as one of the Indo-European languages, is the first spoken language in Iran, Tajikistan, Uzbekistan, and the Pamir mountain region. It is reported that 110'000'000 people speak this language all over the world. Persian constitutes one or more than one syllabic words. This language is considered a syllable-timed language and has three-syllable structures, consisting of CV, CVC, and CVCC. There are 23 consonants and 6 vowels in the speech sound system of Persian. Vowels cannot occur in the initial position of the word and there is always a glottal stop/ʔ/before the vowels in the initial position of the words. Another specific property of the Persian language is the type of consonant clusters that are only in the final position of

the words and have one syllable structure, i.e. CVCC (Samareh, 1992).

The present study

Speech consistency has not been addressed in Persian speaking children and there is no valid and reliable scale to measure this language property. Conducting studies on speech consistency across different languages to increase our knowledge about this phenomenon is a necessity (Martikainen et al., 2019). Therefore, in the current investigation, the validity and reliability of the Persian Test of Speech Consistency (PTOSC) and evaluation of speech consistency in Persian speaking children with typical development were targeted. To establish the construct validity, the children were compared in different age groups. The data related to the effects of sex on speech consistency is insufficient and contradictory. Additionally, we examined the construct validity (known group validity) of PTOSC by comparing boys and girls.

Method

Study design

A cross-sectional, descriptive, and analytical project was conducted to validate the PTOSC and to investigate the speech consistency in Persian speaking children.

Participants

The sampling procedure was launched after obtaining permission from the Welfare Organization of the city of Tehran, the capital of Iran, in 2011. Participants were recruited from nursery schools. They included 317 children with normal development, aged from 36 to 72 months classified into six groups with six months of interval. Group 1 included children who aged between 36–41 months, group 2 involved children who aged between 42–47 months, group 3 included those between 48–53 months, group 4 was assigned for children who aged between 54 – 59 months, group 5 was for the age range between 60 to 65 months, and group 6 was for children who aged between 66–71 months. First, a questionnaire that included the information required to verify inclusion and exclusion criteria was designed by the authors. Considering the inclusion and exclusion criteria, the participants were monolingual Persian speaking children with normal hearing and typical development, in the abovementioned age range, who had no disorders or diseases including language disorders, mental retardation, autism spectrum disorders, history of seizure or epilepsy, craniofacial abnormalities, like cleft lip and palate, hearing impairment, and none of them had a history of speech therapy sessions. The inclusion and exclusion criteria were confirmed by a form that the children's parents completed, hearing screening assessment, screening that was performed by the SLP, and participants' medical reports, which were available.

Development of the PTOSC

The PTOSC development was carried out carefully to select culturally and linguistically suitable and familiar words. The developers of PTOSC applied some criteria including selecting the appropriate items regarding the age of the participants, items' familiarity, and other properties that are critical to developing a valid scale for children. First, an item pool was generated including 120 items with one to five syllables. All of the items were generated through an extensive literature review (Derakhshandeh, 1997; Damarchi et al., 2009; Dodd et al., 2002; Jalilevand, 2011; Hasani, 2006; Mansouri, 2008; Mehdipour Shahrivar & Nematzadeh, 2013; Mohammadi, 2004; Nematzadeh et al., 2012; Shokri, 2004). Then, 120 words were obtained. In this stage, research team investigated and selected the most appropriate items based on the phonological characteristics of Persian, the concept of inconsistency, the frequency of phonemes in Persian, and including words that have all of the Persian consonants and vowels. Ultimately, the number of items was narrowed down to 27 items: 23 items were categorized in the noun class, 2 words as adjective, and the other 2 words were in the verb class. Next, a professional graphic artist, who was a specialist in graphic design for children, prepared colorful pictures in the size of 15×20 cm for the items. The pictures were reviewed by the members of the research team to review and detect their probable imperfections. In this stage, we asked the experts to rate the questions that were related to the characteristics of the PTOSC items in the range of 0 to 100. Then, descriptive statistics were calculated for their ratings for every question. After this phase, PTOSC was administrated on 57 children to calculate the difficulty coefficient. The transcribed forms of the target words of the PTOSC and their English equivalents are given in Appendix 1.

The number of syllables for the items ranged from 1 to 5. SLPs and linguists who were expert in research, evaluation, and intervention of children with SSD were invited to comment on the items, their pictures, and the manuals of the test using VAS, which was then transformed into percentages.

Construct validity

Known group validity, or the ability of the PTOSC to investigate the difference among several age groups, was considered as the evidence of construct validity in the current study. Thus, for this purpose, the mean scores of the variable percentages were compared among the six age groups. Furthermore, the consistently correct and consistently incorrect responses of the children were compared in the six age groups.

Reliability measurement

Two types of reliability measurements were calculated to investigate the reliability analysis in this research: test-retest and inter-rater reliability. Totally, 52 children were

randomly selected from the studied population for the test-retest analysis. Then, administration of the PTOSC was repeated with 14 days of interval. As for the inter-rater reliability, two SLPs were instructed about scoring, recording the data, and transcription of the speech samples. Afterwards, they recorded, transcribed, and scored the produced samples independently, and the correlation between their scorings was considered the inter-rater reliability index (Table 2).

Data analytic approach

The statistical analysis of the data was performed using Statistical Package for the Social Sciences, version 24.0 (SPSS, Inc., Chicago, IL, USA; IBM Corp., Armonk, NY, USA). The procedure used for the item generation was qualitative. Then, the descriptive statistics were obtained for the content validity of the PTOSC. Obtained values of.15–.85 were considered acceptable for measuring of difficulty coefficient (Anastasi, 1998). Also, the Kolmogorov-Smirnov test was used to check the normality of the data distribution. Given the normal distribution of the data, One-way ANOVA was applied to explore the construct validity (comparison of the variable response among age groups) and to compare the consistently correct as well as the consistently incorrect responses among different age groups. Test-retest and inter-rater reliability were measured using Intraclass correlation coefficient (ICC). In addition, independent sample t-test was used to compare the mean scores of the PTOSC in boys and girls. The significance level was set at ($p < .05$).

To analyze the speech responses, code 1 was considered for the state that the child produced the intended words correctly three times. Code 2 was intended for the state that the child produced that word consistently, but his or her production was not correct. The productions that were variables were coded as 3. Consistently correct, consistently incorrect, and variable percentage scores were obtained by dividing the number of words that were produced in each state by 27 (the total items of the PTOSC were multiplied by 100).

Table 2. Descriptive statistics of the participants for calculating the construct validity of PTOSC in terms of age and sex.

	Girls		Boys		*N* (total)	
Age group	*N*	Percent	*N*	Percent	*N*	Percent
1	20	6.3	15	4.7	35	11
2	31	9.8	21	7.6	54	17
3	25	7.9	21	6.6	46	14.5
4	29	9.1	32	10.1	61	19.2
5	30	9.5	26	8.2	56	17.6
6	24	7.6	25	7.9	49	15.4
Total	163	51.4	154	48.6	317	100

Procedure

Each participant was evaluated separately in a quiet room at the nursery school by the first author. First, she called their names, communicated with the examinee, and explained to him or her that we were going to look at some images. She prompted the child to name the images with question sentences including "What is it?", "What is he doing?", and "What color is it?" The examiner asked the child to repeat the words she did not hear. The child was requested to have three repetitions for each word within the consistency assessment in that session, producing 27 words in one session, in three nonconsecutive trials, separated by another activity. For the words that the child could easily name the pictures, the second and the third trials were elicited through name repeating. The required time for implementing the PTOSC was 15–20 min. The first author transcribed the speech samples of the children in a broad manner. She transcribed the speech samples of each child simultaneously. All of the samples were audio recorded using a COBY (Model: MPC-7405) voice recorder and a Sony VAIO (Model: PCG-5p6p) laptop. The audio recording was made to check the accuracy of the transcriptions.

Results

Content validity

The mean values of Table 3 show the agreements among the experts, which were higher than .85 for the questions related to the content validity of the PTOSC.

Difficulty coefficient

The results of calculating the difficulty coefficient of the items of PTOSC showed that 25 out of 27 items had difficulty index between .3–. 7. Two other items had difficulty coefficient of .28 and .25.

Construct validity

The frequency of the consistently correct responses was higher than those of the consistently incorrect and variable responses. The descriptive statistics of all responses are given in Table 4. As seen in this Table, the consistently correct responses were produced more than the other responses by the children.

Table 3. Descriptive statistics of the agreement percentage among experts for the content validity of the PTOSC.

Characteristics of PTOSC	Minimum	Maximum	Mean	SD
The number of items	60	100	88.42	16.26
The syllable structure of the items	97	100	99.28	1.25
Appropriateness of the target words for assessment of consistency	94	100	98.28	2.97
Picturability of the target words	87	100	92.71	5.18
Appropriateness of the test's manual	88	100	98.28	4.53

Table 4. Descriptive statistics of the different types of participants'answers in PTOSC.

Response	Minimum	Maximum	Mean (SD)
Consistently correct	0	100	72.16 (18.73)
Consistently incorrect	0	85.19	13.43 (13.34)
Variable	0	29.63	14.41 (8.35)

Table 5. Evidence of construct validity (discriminative validity in terms of age) of PTOSC by the One-way ANOVA test.

Age groups	Consistently correct Mean (SD)	Consistently incorrect Mean (SD)	Variable Mean (SD)
1	62.85 (13.24)	16.41 (10.72)	20.74 (7.78)
2	68.55 (16.43)	14.68 (12.80)	16.77 (8.14)
3	67.65 (19.41)	17.29 (16.53)	15.06 (8.15)
4	77.53 (13.98)	11.46 (13.68)	11.01 (7.44)
5	77.71 (13.57)	10.19 (10.45)	12.10 (7.52)
6	78.70 (15.94)	10.49 (13.71)	10.81 (6.44)
p value	<.001	.007	<.001

Note: *p value $< .05$.

Table 6. Results of the analysis of the PTOSC in terms of sex by the Independent sample t -test.

Sex	Consistently correct Mean (SD)	Consistently incorrect Mean (SD)	Variable Mean (SD)
Boys	71.66 (19.11)	13.63 (13.80)	14.71 (8.39)
Girls	72.20 (16.32)	13.58 (14.06)	14.22 (8.24)
p value	.52	.97	.87

Note: *p value $< .05$.

Supporting evidence for the construct validity (known group validity) of the PTOSC is given in Table 5. The six age groups were significantly different in terms of the mean value of the percentage of the variable responses ($p < .001$). Thus, PTOSC could discriminate among different age groups.

There were also statistically significant differences among the six groups regarding the consistently correct responses ($p < .001$). Likewise, they were statistically significant differences among them in terms of the consistently incorrect response in PTOSC ($p = .007$).

As demonstrated in Table 6, the participants did not demonstrate significant differences for the consistently correct response ($p = .52$), consistently incorrect responses ($p = .97$), and variable responses ($p = .87$) with regard to sex.

Results of the reliability measurement

There was a significant but moderate correlation between two administrations with 14-day intervals of the PTOSC as calculated by ICC (ICC = .89, $p < .001$) and a high value was obtained for the inter-rater reliability (ICC = .88, $p < .001$).

Discussion

The current research aimed to develop and measure the psychometric properties of PTOSC. Using this scale helps SLPs to measure the speech consistency in the Persian speaking children which could be the scientific innovation of our study.

To prepare this scale, the essential steps of developing an instrument were taken, including literature review, content validity, and measurement of the difficulty coefficient values for the items of PTOSC.

The experts revealed an acceptable level of agreement about the characteristics of the PTOSC. Similar disseminated studies have used agreement percentage for the content validity of the instruments, too (Abou-Elsaad et al., 2009; Ahmadi, 2018).

Among the different responses, the frequency of consistently correct responses was higher than those of the other responses. This finding of the present study is in accordance with the previous studies (Holm et al., 2007; Martikainen et al., 2019) but not with that reported in Sosa study (Sosa, 2015). This discrepancy is possibly related to the participants' age and characteristics of the target words used to assess the word variability. The possible reason for the different findings between our study and the study by Sosa could be the high frequency of the poly-syllabic words in the target words that were utilized to assess the speech variability in her study. This attribute and the younger age of the participants in her study, compared to the other similar studies, probably explains the difference in the finding (Sosa, 2015).

Older children had fewer values for the variable percentage response. This finding implies that the degree of speech variability decreased as age increased, which is consistent with those of the earlier studies (Holm et al., 2007; Martikainen et al., 2019; Sosa, 2015). The obtained values for the variable percentages in the present research are rather higher than those of the previous studies (Holm et al., 2007; Martikainen et al., 2019) and lower than the values reported by some others (Macrae & Sosa, 2015; Sosa, 2015). Given the differences across languages, we cannot directly compare the obtained findings of languages with each other (Martikainen et al., 2019); however, one possible reason for this difference may be the phonological and phonetic features of the target words (Martikainen et al., 2019; Sosa, 2015). About 30% of the target words contained the consonant cluster CVCC, which is somewhat challenging to produce in the Persian speaking children with typical development, according to the existing evidence (Ghasisin et al., 2011; Jalilevand et al., 2014). Moreover, a notable number of the children produced consonants that are acquired at a later time in Persian (Zarifian & Fotuhi, 2020). Another reason that could be considered for this finding is the age group recruited in the previous research (Holm et al., 2007; Martikainen et al., 2019). In the studies conducted by Holm et al (Holm et al., 2007) and Martikainen et al (2019), a broader age range was examined including older children in their studies. Hence, this difference may have an influence on the fewer variable percentages obtained in their studies.

Another result of the present study showed that the participants did not obtain significant values in terms of sex. To put it differently, boys and girls acquired similar percentages for different kinds of responses in this research. This finding is not consistent with those reported in some

previous works (De Castro & Wertzner, 2011; Holm et al., 2007) but in line with the results of some other works (Burt et al., 1999). The reason for the difference may be attributed to some factors including studied age group and the type of tasks used to assess the inconsistency (Holm et al., 2007). In other similar studies (Ahmadi et al., 2018; Fatemi Syadar et al., 2019; Zarifian, 2014), too, conducted to study the other dimensions of speech production skills in Persian speaking children, such findings were reported. A body of evidence emerged showing that effect of sex on phonetic and phonological capabilities was apparent in older children (Dodd, 1995). For example, Poole demonstrated that the interaction of sex and phonological skills was evident in children aged 66 months. In general, the evidence of sex differences in speech inconsistency is relatively little and should be targeted in future studies (Holm et al., 2007).

The current study had some limitations that should be taken into consideration before any generalization. First of all, the main aim of the study was to design and validate an assessment tool for Persian speaking children. Considering time limitations, the study of the speech consistency in different kinds of SSD, although necessary, was not feasible in the present research. So, future surveys can investigate this phenomenon in various types of SSD including inconsistent phonological disorders, and CAS in Persian speakers. Second, in the current study, we studied speech consistency perceptually. Utilizing some instrumental analysis, like kinematic and acoustic measures, and exploring the relationship among these analyses may strengthen our knowledge about discrepancy in the obtained results in speech inconsistency in the performed studies (Martikainen et al., 2019). Third, we recruited Persian-speaking children above three years old. The profile of younger children seems different from those of the older ones (Sosa & Stoel-Gammon, 2006). Further studies are suggested to investigate the speech consistency in younger Persian-speaking children with normal development.

Conclusions

Following the necessary steps to develop a valid and reliable measurement and obtaining acceptable values for the psychometric properties of the PTOSC proved that the Persian SLPs can utilize this scale to measure the speech variability in typically developing children. The findings of the present study showed that Persian-speaking children with typical development within the age range of 3–6 year olds are consistent in speech production to a relatively great extent.

Statement of ethics

The Ethical Committee of the University of Social Welfare and Rehabilitation Sciences approved this research by the reference code of (801.91.ت30082.1). The participants' parents were given a consent form to complete and sign for the recruitment of their children in the study. Moreover, explanations were provided for the parents about the aim of the current work.

Acknowledgements

The authors would like to express their gratitude to the children, their parents, and the teachers in nursery schools who assisted us in conducting the present research. We also extend our deep gratitude to Welfare Organization for its cooperation in sampling procedure.

Disclosure statement

The authors have no conflict of interests to declare.

ORCID

Talieh Zarifian http://orcid.org/0000-0002-6067-829X

References

Abou-Elsaad, T., Afsah, O., & Rabea, M. (2019). Identification of phonological processes in Arabic–speaking Egyptian children by single-word test. *Journal of Communication Disorders*, *77*, 80–93. https://doi.org/10.1016/j.jcomdis.2018.10.003

Abou-Elsaad, T., Baz, H., & El-Banna, M. (2009). Developing an articulation test for Arabic-speaking school-age children. *Folia Phoniatrica et Logopaedica : official Organ of the International Association of Logopedics and Phoniatrics (Ialp))*, *61*(5), 275–282. https://doi.org/10.1159/000235650

Ahmadi, A. (2018). *Development and validation a package for assessment of speech sound production in children with age range of 3 to 5 Year old in Tehran City* [thesis]. Department of Speech and Language Pathology, Iran University of Medical Sciences. A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy, 216. [In Persian].

Ahmadi, A., Ebadi, A., Kamali, M., Zarifian, T., Kazemi, M. D., & Mohamadi, R. (2018). Single word test for the assessment of speech sound production in Persian speaking children: Development, validity and reliability. *International Journal of Pediatric Otorhinolaryngology*, *114*, 61–66. https://doi.org/10.1016/j.ijporl.2018.08.018

Ahmadi, A., Mohamadi, R., Ebadi, A., Kamali, M., Zarifian, T., & Kazemi, M. D. (2019). Development and validation of a new scale for assessment of speech stimulability in Persian speaking children. *International Journal of Pediatric Otorhinolaryngology*, *122*, 180–184. https://doi.org/10.1016/j.ijporl.2019.04.018

American Speech and Hearing Association. (2007). *Childhood apraxia of speech*. https://www.asha.org/practice-portal/clinical-topics/childhood-apraxia-of-speech/

Anastasi, A. (1998). *Psychological testing* (1998 ed.). Macmillan.

Betz, S. K., & Stoel-Gammon, C. (2005). Measuring articulatory error consistency in children with developmental apraxia of speech. *Clinical Linguistics & Phonetics*, *19*(1), 53–66. https://doi.org/10.1080/02699200512331325791

Burt, L., Holm, A., & Dodd, B. (1999). Phonological awareness skills of 4-year-old British children: An assessment and developmental data. *International Journal of Language & Communication Disorders*, *34*(3), 311–335. https://doi.org/10.1080/136828299247432

Crosbie, S., Holm, A., & Dodd, B. (2005). Intervention for children with severe speech disorder: A comparison of two approaches. *International Journal of Language & Communication Disorders*, *40*(4), 467–491. https://doi.org/10.1080/13682820500126049

Damirchi, Z., Jalileh Vand, N., Mahmoudi Bakhtiyari, B., & Keyhani, M. Z. (2009). Development of phonetic inventory in 2–6 year-old Farsi speaking children. *Journal of Research in Rehabilitation*, *5*, 1–6.

De Castro, M. M., & Wertzner, H. F. (2011). Speech inconsistency index in Brazilian Portuguese-speaking children. *Folia Phoniatrica et Logopaedica: Official Organ of the International Association of Logopedics and Phoniatrics (Ialp))*, *63*(5), 237–241. https://doi.org/10.1159/000323183

Derakhshandeh, F. (1997). *The sound system of persian speaking children* [Unpublished thesis database, MA]. Iran University of Medical Sciences, Tehran.

Dodd, B. (1995). *Differential diagnosis and treatment of children with speech disorder*. Whurr.

Dodd, B., Zhu, H., Crosbie, S., Holm, A., & Ozanne, A. (2002). *Diagnostic evaluation of articulation and phonology (DEAP)*. Psychology Corporation.

Eadie, P., Morgan, A., Ukoumunne, O. C., Ttofari Eecen, K., Wake, M., & Reilly, S. (2015). Speech sound disorder at 4 years: Prevalence, comorbidities, and predictors in a community cohort of children. *Developmental Medicine & Child Neurology*, *57*(6), 578–584. https://doi.org/10.1111/dmcn.12635

Fatemi Syadar, S., Zarifian, T., Modarresi, Y., Zahedi, M. S., Ebrahimipour, M., & Biglarian, A. (2019). Kurdish Speech Test: A validation study for children aged 3–5 years. *International Journal of Pediatric Otorhinolaryngology*, *117*, 61–66. https://doi.org/10.1016/j.ijporl.2018.10.009

Ghasisin, L., Ghayoumi, Z., & Mahmoudi Bakhtiyari, B. (2011). *Study of developmental consonant cluster in 2–4 years old Farsi speaking children*. http://jrrs.mui.ac.ir/index.php/jrrs/article/view/374

Hasani, H. (2006). *The most frequent words of today Persian: Inclding over 8000 frequent words based on a one-million-word corpus*. Iranian Language Institute.

Holm, A., Crosbie, S., & Dodd, B. (2007). Differentiating normal variability from inconsistency in children's speech: Normative data. *International Journal of Language & Communication Disorders*, *42*(4), 467–486. https://doi.org/10.1080/13682820600988967

Iuzzini-Seigel, J., Hogan, T. P., & Green, J. R. (2017). Speech inconsistency in children with childhood apraxia of speech, language impairment, and speech delay: Depends on the stimuli. *Journal of Speech, Language, and Hearing Research: JSLHR*, *60*(5), 1194–1210. https://doi.org/10.1044/2016_JSLHR-S-15-0184

Jalilevand, N. (2011). Development of speech and language in Farsi speaking children.

Jalilevand, N., Kamali, M., Tavakoli, M., Mahmudi, Z., Amiri, M., Hadavi, S., & Javanbaxt, M. (2014). The ability of 3-to 6-years-old Persian-speaking children in production of consonant clusters in mono-syllable CVCC words. *Audiology*, *23*(3), 23.

Keating, D., Turrell, G., & Ozanne, A. (2001). Childhood speech disorders: Reported prevalence, comorbidity and socioeconomic profile. *Journal of Paediatrics and Child Health*, *37*(5), 431–436. https://doi.org/10.1046/j.1440-1754.2001.00697.x

Kenney, K. W., & Prather, E. M. (1986). Articulation development in preschool children: Consistency of productions. *Journal of Speech, Language, and Hearing Research*, *29*(1), 29–36. https://doi.org/10.1044/jshr.2901.29

Lousada, M., Jesus, L. M., Hall, A., & Joffe, V. (2014). Intelligibility as a clinical outcome measure following intervention with children with phonologically based speech–sound disorders. *International Journal of Language & Communication Disorders*, *49*(5), 584–601. https://doi.org/10.1111/1460-6984.12095

Macrae, T. (2013). Lexical and child-related factors in word variability and accuracy in infants. *Clinical Linguistics & Phonetics*, *27*(6–7), 497–507. https://doi.org/10.3109/02699206.2012.752867

Macrae, T., & Sosa, A. V. (2015). Predictors of token-to-token inconsistency in preschool children with typical speech-language development. *Clinical Linguistics & Phonetics*, *29*(12), 922–937. https://doi.org/10.3109/02699206.2015.1063085

Mansouri, M. (2008). *The Frequency system of phonemes in Persian*. Ilam University.

Martikainen, A.-L., Savinainen-Makkonen, T., & Kunnari, S. (2019). Intra-word consistency and accuracy in Finnish children aged 3–6 years. *Clinical Linguistics & Phonetics*, *33*(9), 815–816. https://doi.org/10.1080/02699206.2019.1576770

McKinnon, D. H., McLeod, S., & Reilly, S. (2007). The prevalence of stuttering, voice, and speech-sound disorders in primary school students in Australia. *Language, Speech, and Hearing Services in Schools*, *38*(1), 5–15. https://doi.org/10.1044/0161-1461(2007/002)

Mehdipour Shahrivar, N., Shirazi, T., & Nematzadeh, S. (2013). Most frequent expressing words of Farsi-speaking children ages between 18–24 months. *J Speech Language Pathology*, *1*(1), 71–79.

Mohammadi, M. (2004). *Picture-Riddle Dictionary*. Cheesta Publishing Co.

Nematzadeh, S., Dastjerdi Kazemi, D. M., M., & Mansoorizadeh, M. (2012). *Basic Persian words from the language of Iranian children*. Announcer of Education.

Samareh, Y. (1992). Avashenashi-e Zabane-e Farsi [Persian language phonetics]. University Publications Centre.

Shokri, G. (2004). *The Persian Dictionary for school children*. Institute for Humanities and Cultural Studies.

Sosa, A. V. (2015). Intraword variability in typical speech development. *American Journal of Speech-Language Pathology*, *24*(1), 24–35. https://doi.org/10.1044/2014_AJSLP-13-0148

Sosa, A. V., & Stoel-Gammon, C. (2006). Patterns of intra-word phonological variability during the second year of life. *Journal of Child Language*, *33*(1), 31–50. https://doi.org/10.1017/S0305000905007166

Stackhouse, P., & Williams, J. (2000). Rate, accuracy and consistency: diadochokinetic performance of young, normally developing children. *Clinical Linguistics & Phonetics*, *14*(4), 267–293.

Terband, H., Namasivayam, A., Maas, E., van Brenk, F., Mailend, M.-L., Diepeveen, S., van Lieshout, P., & Maassen, B. (2019). Assessment of childhood apraxia of speech: A review/tutorial of objective measurement techniques. *Journal of Speech, Language, and Hearing Research: Jslhr*, *62*(8S), 2999–3032. https://doi.org/10.1044/2019_JSLHR-S-CSMC7-19-0214

Zarifian, T. (2014). *Adaptation of diagnostic evaluation of articulation and phonology for 3–6 persian speaking children and determining it's psychometric properties* [Thesis]. Department of Speech-Language Pathology, Tehran, University of Social Welfare and Rehabilitation. A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy, Tehran, Modarresi, Yahya, Gholami Tehrani, Laya, [In Persian], (Tehran).

Zarifian, T., & Fotuhi, M. (2020). Phonological development in Persian-speaking children: A cross-sectional study. *International Journal of Speech-Language Pathology*, 1–12.

Appendix 1. Target words of PTOSC

Phonetic Transcription of Target Word	Target Word(Eng.)	Phonetic Transcription of Target Word	Target Word(Eng.)
l ɑ k p o ʃ t	turtle	ʤ ɑ r u b æ r G i	Vacuum cleaner
ʔ ɑ d æ m b æ r f i	snowman	h æ v i ʤ	carrot
d o k t o r	doctor	k o l ɑ h	hat
r ɑ n æ n d e g i	driving	g ɑ v	cow
z æ r d	yellow	d æ r s	study
ʃ of ɑ ʒ	radiator	z i p	zip
ʒ e l e	jelly	k æl ɑ G	crow
l i v ɑ n	glass	t æ s ɑ d o f	crash
χ æ r g u ʃ	rabbit	s i b z æ m i n i	potato
m o b l	sofa	t o f æ n g	gun
j æ χ ʧ ɑ l	refrigerator	s æ b z	green
n e j	straw	p i ʧ	screw
ʧ æ r χ	wheel	x æ m i r d æ n d an	toothpaste
z i r p u ʃ	underwear		