Original Article
http://mjiri.iums.ac.ir

Medical Journal of the Islamic Republic of Iran (MJIRI)
Iran University of Medical Sciences

The Persian developmental sentence scoring as a clinical measure of morphosyntax in children

Nahid Jalilevand[1], Mohammad Kamali*[2], Yahya Modarresi[3], Yalda Kazemi[4]

Received: 24 May 2016 Accepted: 23 July 2016 Published: 31 October 2016

Abstract

Background: Developmental Sentence Scoring (DSS) was developed as a numerical measurement and a clinical method based on the morphosyntactic acquisition in the English language. The aim of this study was to develop a new numerical tool similar to DSS to assess the morphosyntactic abilities in Persian-speaking children.

Methods: In this cross-sectional and comparative study, the language samples of 115 typically developing Persian-speaking children aged 30 - 65 months were audio recorded during the free play and picture description sessions. The Persian Developmental Sentence Score (PDSS) and the Mean Length of Utterance (MLU) were calculated. Pearson correlation and one – way Analysis of variance (ANOVA) were used for data analysis.

Results: The correlation between PDSS and MLU in morphemes (convergent validity) was significant with a correlation coefficient of 0.97 ($p< 0.001$). The value Cronbach's Alpha ($\alpha= 0.79$) in the grammatical categories and the split-half coefficient (0.86) indicated acceptable internal consistency reliability.

Conclusion: The PDSS could be used as a reliable numerical measurement to estimate the syntactic development in Persian-speaking children.

Keywords: Language Measurement, Validity, Reliability, Typically Developing Children, Persian Language.

Cite this article as: Jalilevand N, Kamali M, Modarresi Y, Kazemi Y. The Persian developmental sentence scoring as a clinical measure of morphosyntax in children. *Med J Islam Repub Iran* 2016 (31 October). Vol. 30:435.

Introduction

In a natural setting, typically developing children usually tend to speak about their needs, ideas and feelings, or ask their parents about their environment. Children's language samples provide information about their language abilities, which can be systematically elicited in a natural conversation by a speech-language pathologist. Three numerical measures that have often been used for analyzing language samples are Mean Length of Utterance (MLU) (1), Developmental Sentence Score (DSS) (2,3) and Index of Productive Syntax (IPSyn) (4).

MLU is a common and valid tool used by researchers and clinicians to measure syntax (5). Rice et al. concluded that the MLU, as an index of general language development, is valid and reliable (6). On the other hand, some limitations of MLU were reported by some researchers. Miller and Chapman reported that different normal children have different MLUs at the same age (7). Klee and Fitzgerald believed that although there is a high correlation between MLU and chronological age in typically developing children, MLU is a valid index of development until approximately 3.0 morphemes in Brown's stage II (8). They suggested that MLU is a gross indicator of grammatical development (8).

Lee presented a standard method known as the DSS (3). She used this quantitative method to evaluate standard grammatical features, including eight grammatical cate-

[1]. PhD, Assistant Professor, Department of Speech and Language Pathology, School of Rehabilitation Sciences, Iran University of Medical Sciences, Tehran, Iran. jalilevand.n@iums.ac.ir
[2]. (**Corresponding author**) PhD, Associate Professor, Department of Rehabilitation Management, School of Rehabilitation Sciences, Iran University of Medical Sciences, Tehran, Iran. kamali.mo@iums.ac.ir
[3]. PhD, Professor, Department of Linguistics, Institute for Humanities and Cultural Studies, Tehran, Iran. ymodarresi@yahoo.com
[4]. PhD, Assistant Professor, Department of Speech Therapy, Faculty of Rehabilitation, Isfahan University of Medical Sciences, Isfahan, Iran. y.kazemi@hotmail.com

gories: Indefinite pronouns, personal pronouns, main verbs, secondary verbs, negatives, conjunctions, interrogative reversals, and WH-questions in the spontaneous speech of 200 typically developing children aged 2- to 6 years 11 months. DSS was calculated, using a sample of 50 complete sentences. To calculate the DSS, the sum of the scores from at least 50 complete sentences of speech sample was divided by 50 (3). The validity and reliability of the DSS were determined by Koenigsknecht (9). He suggested that a significant difference between the DSS of different age groups proved its validity, and concluded that the validity and reliability of DSS strongly support its usefulness as a measure of syntax development in children (9).

Various studies illustrated that DSS provides valuable information for the clinical setting. Hux et al. found that DSS is the most common standard and analytic method used by American speech-language Pathologists (10).

Toronto developed the Developmental Assessment of Spanish Grammar (DASG), which is similar to the DSS. The DASG evaluates grammatical performance of Spanish-speaking children in spontaneous speech. The six grammatical categories employed were as follows: Indefinite pronouns and noun modifiers, personal pronouns, primary verbs, secondary verbs, conjunctions and interrogative words. The DASG average score was reported on 128 Spanish-speaking children between the ages of 3 and 6 years 11 months. Toronto reported that the DASG can discriminate between one-year age levels (11).

The Developmental Sentence Score for Japanese (DSSJ) is the newest method to evaluate the syntax based on DSS in a non-English language (12). Miyata et al. developed DSSJ and reported the overall DSSJ in 84 typically developing children in age intervals of 2 years 8 months to 5 years 2 months. They collected 100 sentences during child-adult conversation and free play. A statistical analysis revealed that DSSJ and MLU were highly correlated. They introduced DSSJ as a valuable tool to study language acquisition (12).

Scarborough developed the Index of Productive Syntax (IPSyn) as a grammatical measure (4). IPSyn is based upon the developmental stage of 56 subcategories of noun phrase, verb phrase, questions, negations and sentence structure. She studied the speech samples of 15 children at the ages of 24, 30, 36, 42 and 48 months longitudinally, and gathered 75 speech samples. The comparison between the mean of the Index of Productive Syntax and MLU at each age demonstrated the reliability of the Index of Productive Syntax (4).

Gathering spontaneous language samples allows multiple analyses of speech and language aspects. To analyze language samples, Persian speech-language pathologists (SLPs) usually use the MLU as a tool to measure grammatical development. Jalilevand et al. conducted a longitudinal study on two Persian-speaking children (a girl and a boy) aged 12–60 months. The spontaneous speech of these two children was recorded and analyzed. The MLU increased with the age of the children. The variation slope in the MLU was steeper around 24–42 months (13). Kazemi et al. examined the MLU in 171 Persian-speaking children, aged 2.5–5.5, who lived in Esfahan. The results of this study indicated that the rate of variation was not high after 3–3.5 years, possibly due to slow syntactic growth after this age (14).

The Persian SLPs cannot use the DSS to analyze sentences in language samples, because the grammatical structure and morphosyntactic items in the Persian language are not similar to the English language. For example, Persian is a pro-drop language with canonical SOV word order (15), but English is not. The aim of this study was to develop a numerical measurement to analyze Persian sentences based on the Developmental Sentence Scoring model in English (3), and to evaluate its psychometric properties.

Methods

This descriptive-analytic cross-sectional study was conducted in two steps. The first step was constructing the Persian developmental sentence scoring chart. We needed to collect the morphosyntactic items and their acquisition time in typically developing Persian speaking children. Some studies have examined the morphosyntactic acquisition in three Persian-speaking children, who had been investigated longitudinally (13,16-18). The acquisition of some grammatical items were reported in these literatures (13,16-18), which were selected and grouped in eight Persian grammatical categories: Pronouns, question words, prepositions and conjunctions, verb morphology, modal and compound verbs, and grammatical morphemes, sentence types and sentence structures based on the Persian grammar sources (15,19-21). According to the time range acquisition, the grammatical subcategories were grouped into six levels based on the child's age (or weighted from 1 to 6). Then the PDSS chart was created with nine columns and six rows. The eight columns demonstrate eight grammatical categories, and the six rows indicate subcategories, and one column represents the sentence point. Each category was divided to some grammatical items (subcategories). For example, the first row in the question word category column has four items: /ku/ and /koĴa/ (where), / či/ (what), /ki/ (who), each item was weighted 1.

Data Collection and Statistical Analysis

One hundred fifteen monolingual Persian-speaking children (55 girls and 60 boys), whose parents signed the parental consent form and replied to the Age & Stage Questionnaire (ASQ), were selected from 18 kindergartens managed under the supervision of the Welfare Organization in Tehran (22). The children's ASQ scores on communication, fine motor, gross motor, personal, social and problem solving skills were normal. Each domain was scored separately, and the scores were compared to the screening cut-off score of each domain (22). The children had no history of neurological problems, seizures, brain damage, or any other disorder and no symptoms of movement delay. The children were divided into six age groups of six month intervals.

This study was approved by the Ethics Committee of Iran University of Medical Sciences. The participants were free to withdraw at any stage of the study.

Sampling and Instruments

Language sampling was done via natural conversations in free play and picture description context (2,12,23-25). One examiner collected the language samples in all the children in an appropriate (with minimum noise and enough light) room at each kindergarten. A free play session was provided for each child, and their conversations were recorded to be used as the child's language sample. One of the toys was a dollhouse with furniture and included a bedroom set, a dining room set, a bathroom and a restroom. Four small dolls (10 centimeters tall) were used as well. Two sets of dolls' clothes, an orderly and complete set and a messy and incomplete set provided more stimulant settings for the child's language elicitation. Other toys included kitchen utensils, balls, animals, cars and toy soldiers and guns. The description task included 30 colourful pictures (20×25 cm), representing daily family activities with the mother, father and children at home, at the park, at the doctor's office, at the birthday party and at the beach. During the conversation stage, the children were encouraged to describe the pictures. Language samples of 20–30-minute conversations between children and the examiner were recorded, using a digital voice recorder (Kingston-DVR-902) that was kept in a container to avoid distracting the child. The sampling conversation started with 15 minutes of free play, followed by a 15-minute of picture description.

The language samples of children, who were reluctant to communicate, (30 to 35-

month-old children with less than 50 intelligible utterances and of 36-month and older children with less than 100 intelligible utterances) were excluded from the sample set.

We randomly selected 15 children from the 115 participants to examine test–retest reliability, with an interval of 10 days to two weeks (25). Eliciting speech samples was done again through free play conversation and picture description.

Transcription, Coding Reliability and Scoring

The language samples were orthographically transcribed based on the transcription conventions (Persian Transcription Convention Protocol: PTCP), considering the utterance and morpheme segmentation criteria (26). Lexical morphemes, functional morphemes, inflectional morphemes and clitics (19,21) were segmentated based on the Persian-adapted instruction of the Systematic Analysis of Language Transcripts (SALT) software (26,27). The criteria to consider an accepted utterance for analysis are as follows:

The full intelligible utterances were accepted. The repeated utterances, one-word utterances and utterances without a verb were excluded. Each part of speech in the language sample was divided into more than an utterance with greater than two seconds pause, or terminal intonation rising. A maximum of 50 consecutive sentences from each setting of free play and picture description sessions were transcribed to meet the criterion of 100 sentences. Each sentence in the language sample was checked for the grammatical items listed in the PDSS table and coded based on the coding rules of the SALT. Every sentence that was produced by all the sentence rules of an adult standard sentence was scored with a sentence point.

The Traditional method of calculating MLU is to divide the total number of morphemes or words by the total number of utterances (1,5,24). The following Persian utterance is segmented to four morphemes and two words:

Utterance
bæstæni = o xord-i?
Ice cream= OM eat.PAST-2SG.SU
Did you eat ice cream?

Segmentation
bæstæni= 1 morpheme, 1 word;
= o = 1 morpheme;
xord-i= 2 morphemes, 1 word.

In this study, the MLUs of the language samples were calculated by the SALT. The PDSS was calculated from the total sentence scores of the language sample and dividing it by the total number of the sentences. To calculate the total sentence scores, we used the PDSS codes that were sorted by SALT.

To examine the inter-rater point-to-point agreement reliability, the percent agreement was calculated. The first 10 minutes of 20% of the language samples were re-transcribed by a trained transcriber (28). The agreements and disagreements were calculated for the utterances, morphemes, and coding.

Total agreement / total agreement + total disagreement × 100, was used as the formula for percentage agreement calculation (29). This calculation indicated 94%, 96% and 93% inter-rater agreement for utterances, morphemes, and coding, respectively. The values demonstrate an acceptable level of agreement (11).

Statistical Analysis

Statistical analysis was carried out, using the SPSS Software (Version 17). The mean and standard deviations for age, the MLU in morpheme, the MLU in word and the PDSS were calculated for every participant. A Kolmogorov-Smirnov test at $p>0.05$ was conducted, which indicated that the data were normally distributed; and hence, parametric testing was conducted. A one-way Analysis of Variance (ANOVA) was performed to evaluate the effect of age groups as an independent variable, and the PDSS

as dependent variables. Post hoc comparisons were done by the Tukey HSD test. To study the correlation between the total score of the PDSS with the total score of the MLUs (Convergent Validity) and the total score of the PDSS and age (age discriminative validity), Pearson's coefficients were calculated. The internal consistency of the scale was calculated through the Cronbach's alpha coefficient, which was calculated for the PDSS total score. Cronbach's alpha reliability coefficient normally ranges between 0 and 1. The test-retest reliability (Temporal reliability, inter-rater reliability) of the PDSS was evaluated by the repeatability coefficient of relative or Intraclass Correlation Coefficient (ICC) at $p<0.05$. The Standard Error of Measurement (SEM) was also calculated ($SEM= SD.\sqrt{1-r_{test-retest}}$).

Results

The mean and standard deviations of the children's ages are shown in Table 1. The mean and standard deviations of the PDSS and MLUs of the six age groups are shown in Table 2. Figure 1 displays the mean PDSS per age group. A significant positive correlation coefficient between age and MLUs and PDSS is shown in Table 2. Figure 2 shows the relationship between MLU in morphemes and PDSS. The correlation coefficient r between age and PDSS was 0.69 ($p<0.001$). The correlation coefficient r between age and MLU in morphemes and MLU in words was 0.68 ($p<0.001$) and 0.67 ($p<0.001$), respectively. The correlation between the PDSS and the MLU in morphemes was significant where the correlation coefficient r was 0.97 ($p<0.001$); and the correlation between the PDSS and the MLU in words was significant where the correlation coefficient r was 0.95 ($p<0.001$) as well. A one-way ANOVA was conducted to compare the effect of age groups as an independent variable and PDSS as a dependent variable. There was a statistically significant difference among age group means ($F(5,109)= 24.691$, $p<0.001$). Post hoc comparisons, using the Tukey HSD test indicated that the PDSS mean for the 30-35 months age group (mean= 9.2, SD= 1.2) were significantly different from the other age groups (Appendix A shows the significant results in Tukey HSD test).

Cronbach's Alpha was 0.79 for the grammatical categories, and Spearman-Brown Coefficient split-half (internal consistency) was 0.86. The correlation between the grammatical categories (items) and the PDSS (item-total) was calculated to test the scale homogeneity. There was a significant correlation between the grammatical categories and the PDSS ($p<0.05$), except for the category of question words. To examine

Table 1. Age groups, Gender and the Age Means and Standard Deviations of the Participants (n=115)

Age Groups In month	Gender		Total Participants	Age
	Girls	Boys		Mean (SD)
30-35	10	10	20	32 (1.6)
36-41	10	10	20	39.4(1.6)
42-47	8	10	18	44.6(1.6)
48-53	10	10	20	51.4 (1.1)
54-59	9	10	19	56.5(1.7)
60-65	8	10	18	62.5 (1.7)
Total	55	60	115	47.4(10.4)

Table 2. The Mean and Standard Deviation of MLUm, MLUw and PDSS in Six Age Groups in a Month (n=115)

Age Groups		30-35	36-41	42-47	48-52	53-59	60-65
MLUm	Mean	5.9	7.06	7.64	7.89	8.94	8.89
	SD	1.1	1.03	0.79	0.9	1.41	1.12
MLUw	Mean	3.7	4.27	4.74	4.93	5.58	5.67
	SD	0.6	0.59	0.65	0.5	0.94	0.85
PDSS	Mean	9.2	10.83	11.61	11.95	13.21	13.07
	SD	1.2	1.3	1.03	1.1	1.76	1.27

The Mean Length of Utterance in Morphemes (MLUm), Mean Length of Utterance in Words (MLUw), and Persian Developmental Sentence Scoring (PDSS)

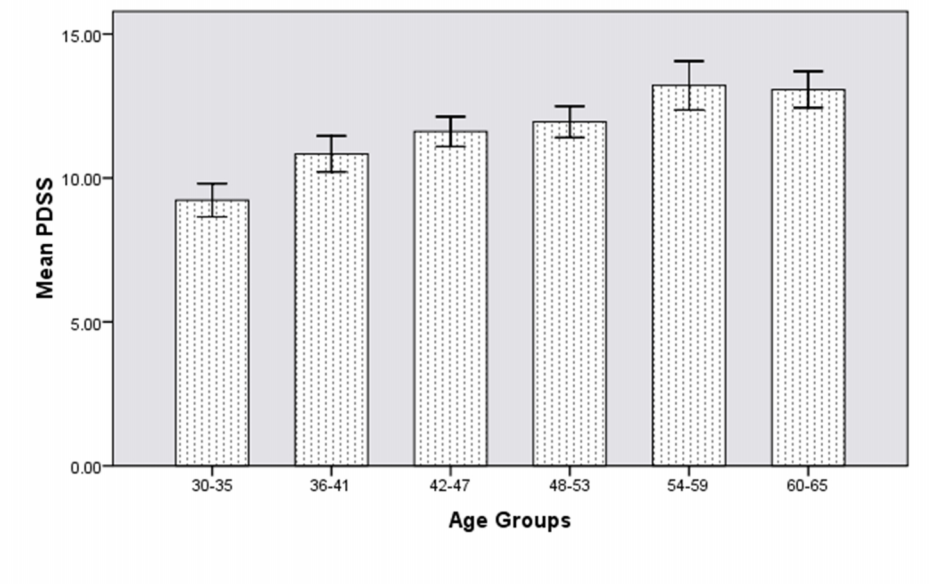

Fig. 1. Mean PDSS per Age Group (in 6 - Month intervals)

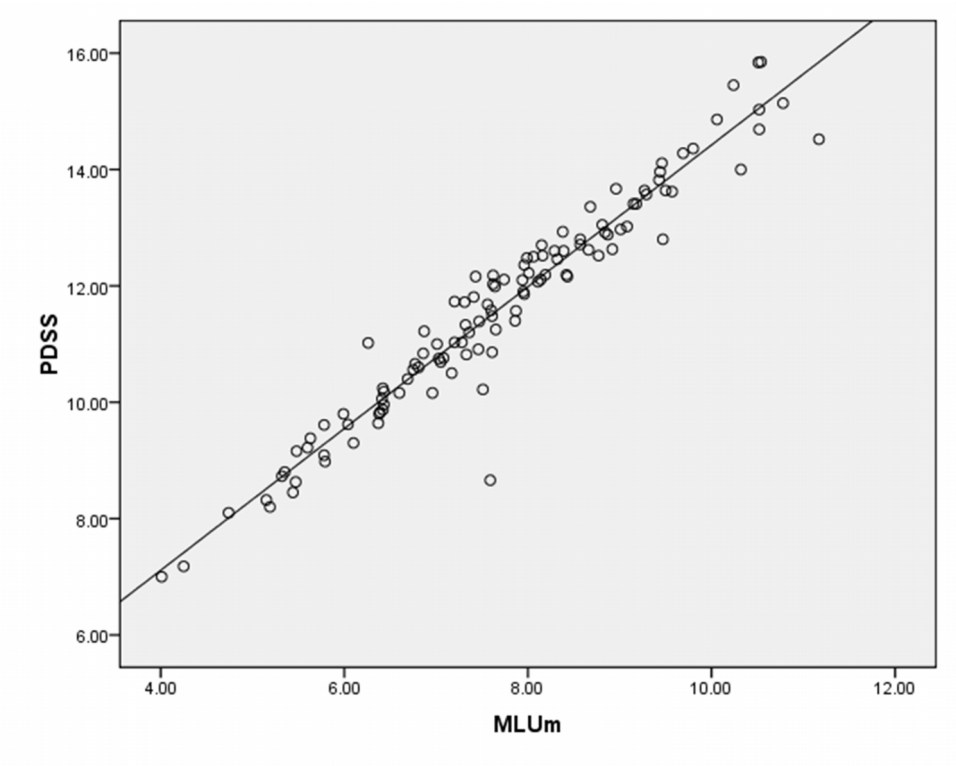

Fig. 2. The Relationship between PDSS and MLU in Morphemes

test–retest reliability, language sample with an interval of 10 days to 2 weeks was elicited from 15. The summary of item statistics are as follows: The mean (variance) and range were 12.06 (0.94) and 0.43, respectively. It was determined that the intraclass correlation coefficient (ICC) was 0.77 with 95% confidence interval (0.46-0.91), and the SEM was 0.96.

Discussion

The aim of this study was to develop PDSS as a numerical measurement to estimate the syntactic development in Persian-speaking children and evaluate its psychometric features. The results revealed acceptable psychometric properties. We compared the statistical features of this study with those reported for DSS, DASG, and DSSJ. The strong correlation between

MLU in morphemes and words with PDSS score suggested good convergent validity. Koenigsknecht reported a strong correlation (r=0.87) between MLU and DSS of complete sentences (6). Similarly, Miyata et al. found a strong correlation between DSSJ and the mean level of utterance in morphemes (r=0.94).

Cronbach's alpha (α) is the most common measure for assessing reliability and determining internal consistency (35). The significant and high correlation between items and the significant correlation between the two sets of items (split-half) indicate acceptable reliability. Koenigsknecht examined the overall internal consistency of DSS, using the coefficient α, (0.71), and the split-half coefficient (0.73), and demonstrated a positive correlation between individual grammatical categories and overall DSS scores, which supported its validity (6).

The significant correlation between grammatical subcategory scores and PDSS (item-total correlation) confirmed the internal consistency reliability and construct validity of the PDSS. The relationship between age and PDSS supports and confirms the age discriminative validity. A relationship also exists between age and DSS (6), DASG (14), and DSSJ (15).

Koenigsknecht concluded that significant differences that exist between the age groups is an indicator of DSS validity (6). Torento believed that increasing DASG scores with age can indicate the instrument's validity (14). Miyata et al. found a strong correlation between average DSSJ scores and age (15), but reported that the oldest age group (5; 2) scored slightly lower than the 4;8 age group (15). In this study, the coefficient of the correlation indicated a strong correlation with PDSS and age. In addition, a statistically significant difference was detected in the means of different age groups with average PDSS, but post hoc comparisons indicated no statistically significant differences between the groups aged 60–65 and 54–59 months. The average PDSS scores of the group aged 54–59 months was slightly more than that of the oldest group. Therefore, we can conclude that PDSS could not reflect morphosyntax development of Persian-speaking children in 6-month intervals, as Lee reported DSS in 1-year interval age groups (4).

A significant positive correlation coefficient confirmed the relationship between age and grammatical subcategories, but all coefficients were not equal. The verb morphology, grammatical morpheme, sentence structure, and preposition and conjunction scores had the highest correlation with age. The sentence structure scores were related to the complexity of sentences; therefore, increasing the sentence structure scores also increased the complexity of sentences. A significant positive correlation coefficient also confirmed the relationship between PDSS and the verb morphology, grammatical morpheme, sentence structure, and preposition and conjunction scores. Similar to Koenigsknecht's report, our study found significant overall differences in four of the eight DSS grammatical category scores: Indefinite pronouns and noun modifiers, personal pronouns, main verbs, and conjunctions (6). In addition, Miyata et al. reported that verb final inflection, verb middle inflection, conjunctions, copula, case, and other particles contributed to changes in the DSSJ score (15).

We observed poor correlation coefficient scores between age and question words. However, there were no significant correlation coefficient scores between question words and PDSS. This result was similar to that reported by Koenigsknecht (6) and Toronto (14).

In this study, the ICC value was 0.77 and indicated excellent inter-rater reliability (36). Koenigsknecht used the repeated applications of the DSS sampling procedure to examine temporal stability. He compared four measures during two weeks period and found that the mean DSS score for 10 participants increased progressively from first sampling to forth sampling (9). Toronto reported no differences in scoring between children tested by different examiners (11).

The result showed that by increasing MLUs, PDSS score also increases. Thus, the number of morphemes and words could increase PDSS score, which was similar to the MLUm and MLUw. The number of morphemes or words can add to the length of the utterance and increase the MLUm and MLUw, respectively. However, sentences with an equal number of morphemes may have different developmental sentence scores. According to Klee et al., MLU may not be a sensitive measure of any linguistic construct other than utterance length itself (8). Thus, the differences between sentences with an equal number of morphemes should be measured by PDSS because the weight or scores of morphemes or grammatical subcategories are not equal in the method of sentence scoring. Therefore, PDSS could determine the developmental value of sentences with different grammatical subcategories. The subcategories with high scores in a sentence could increase its total PDSS score. Thus, we could compare Persian-speaking children not only in terms of the total PDSS score, but also in grammatical subcategory scores. Consequently, the investigators and SLPs could analyze the grammatical subcategories by employing Persian-speaking children in a study, and by using PDSS as a clinical measure of morphosyntax.

An important limitation of this study was that we could not have parent–elicited conversational samples for all participants in the kindergartens. Thus, language sample was elicited by an examiner.

Conclusion

Here, PDSS was developed as a tool to assess the morphosyntactic abilities of Persian-speaking children based on the developmental hierarchy of the Persian morphosyntactic rules, and its validity and reliability were examined. The findings confirmed that PDSS may be used as a reliable numerical measurement to analyze the following Persian grammatical categories: Verb morphology, modal and compound verbs, grammatical morphemes, pronouns, question words, prepositions and conjunctions, sentence structure, and sentence type. The investigators and Persian SLPs might be able to use this method as a clinical measure to estimate the syntactic development in Persian-speaking children with language delay and compare them with typically developing Persian-speaking peers before and after speech therapy.

Acknowledgments

This investigation was supported by the education grant number of 320-3246-93-9-9 granted by the Vice-Chancellor for Research at Iran University of Medical Sciences. The authors would like to thank Amir Hossain Mashreghi, Navid Samavatian, Sahar Mafian, Tahmine Maleki, Fateme Babajani,Samane Babajani and Marziye Faraji for their contributions.

References

1. Brown R. A first language: The early stages. Cambridge, MA: Harvard University Press; 1973.
2. Lee LL, Canter SM. Developmental Sentence Scoring: A Clinical procedure for estimating syntactic development in children's spontaneous speech. J Speech Lang Hear Dis 1971; 36:315-40.
3. Lee LL. Developmental sentence analysis: A grammatical assessment procedure for speech and language clinicians. Evanston, IL: Northwestern University Press; 1974.
4. Scarborough HS. Index of Productive Syntax. Applied Psycholinguistics 1990;11:1-22.
5. Paul R. Language Disorders from infancy through Adolescence. Mosby Inc. Philadelphia; 2007.
6. Rice ML, Redmond SM, Hoffman L. Mean length of utterance in children with specific language impairment and in younger control children shows concurrent validity and parallel growth trajectories. J Speech Lang Hear Res 2006; 49(4):793 - 808.
7. Miller JF, Chapman RS. The Relation between Age and Mean Length of Utterance in Morphemes. J Speech Lang Hear Res 1981;24:154-161.
8. Klee T, Fitzgerald MD. The relation between grammatical development and mean length of utterance in morphemes. Child Lang 1985;12:251-269.
9. Koenigsknecht RA. Statistical information on development sentence analysis. In LL Lee. Developmental sentence Analysis: A grammatical assessment procedure for speech and language clinicians.. Evanston, IL: Northwestern University Press; 1974. PP. 222-268.

10. Hux K, Morris-Friehe M, Sanger DD. Language sampling practices: A Survey of nine states. Lang Speech Hear Serv Sch 1993;24:84 -91.
11. Toronto AS. Developmental assessment of Spanish grammar. J Speech Lang Hear Dis 1976; 41(2):150-171.
12. Miyata S, MacWhinney B, Otomo K, Sirai H, Oshima-Takane Y, Hirakawa M, et al. Developmental sentence scoring for Japanese. First Lang 2013;33(2):200-16.
13. Jalilevand N, Ebrahimipour M, Purqarib J. Mean length of utterance and grammatical morphemes in the speech of two Farsi-speaking children. Audiol 2012;21(2):96-108.
14. Kazemi Y, Taheri A, Kianfar F, Shafiei M, Eslamifard R, Pirmoraian M, et al. Mean length of utterance (MLU) in typically–developing 2:6-5:6 year–old Farsi–speaking children in Iran. JRRS 2012;8(5):928-937.
15. Mahootian S. Persian. London: Routledge; 1997.
16. Jalilevand N, Ebrahimipour M. Pronoun acquisition in Farsi-speaking children from 12 to 36 months. JCLAD 2013;1(1):1–9.
17. Jalilevand N, Ebrahimipour M. Use of question words in two Persian speaking children's speech from 12 to 36 months. Audiol 2010;19(2):93-101.
18. Meshkatodini M. Investigating the development of inflectional affixes and connective structures in developmental articulation of children. Journal of Linguistics 2004;19:10–26.
19. Kalbasi I. The derivational structure of word in Modern Persian. Tehran: Institute for Humanities and Cultural Studies; 2008.
20. Mace J. Persian Grammar: For reference and revision, RoutledgeCurzon, NewYork; 2003.
21. Meshkatodini M. Persian Grammar, the lexical categories and merge (3rd Ed.). Tehran: SAMT; 2008.
22. Vameghi R, Sajedi F, Mojembari AK, Habibollahi A, Lornezhad HR, Delavar B. Cross-cultural adaptation, validation and standardization of Ages and Stages Questionnaire (ASQ) in Iranian Children. Iranian Journal of Public Health 2013;42 (5):522–528.
23. Eisenberg SL, Guo LY. Sample Size for Measuring Grammaticality in Preschool Children from Picture-Elicited Language Samples. Lang Speech Hear Serv Sch 2015;46(2):81–93.
24. Rice ML, Smolik F, Perpich D, Thompson T, Rytting N, Blossom M. Mean Length of Utterance Levels in 6-month Intervals for Children 3 to 9 Years with and without Language Impairments. J Speech Lang Hear Res 2010 Apr;53(2):333–349.
25. Cole KN, Mills P, Dale PS. Examination of Test-Retest and Split-Half Reliability for Measures Derived from Language Samples of Young Handicapped Children. Lang Speech Hear Serv Sch 1989;20(3):259.
26. Kazemi Y, Klee T, Stringer H. Diagnostic accuracy of language sample measures with Persian-speaking preschool children. Clinical Linguistics & Phonetics 2015:1-15.
27. Miller JF, Iglesias A. Systematic Analysis of Language Transcripts (SALT), Research version 2012. Middleton, WI: SALT Software LLC; 2012.
28. Perakyla A. Reliability and validity in research based on tapes and transcripts. In: C S, editor. Social research methods: A reader. London: Routledge; 2004.
29. Wright HH, Capilouto GJ, Koutsoftas A. Evaluating measures of global coherence ability in stories in adults. Int J Lang Commun Disord 2013; 48(3):249-256.
30. Finestacka LH, Payesteha B, Dishera JR, Julien HM. Reporting Child Language Sampling Procedures. J Speech Lang Hear Res 2014;57(6):2274-2279.
31. Santos JR. Cronbach's Alpha: A Tool for Assessing the Reliability of Scales, Journal of Extension 1999;37(2):1-5.
32. Evans JD. Straightforward statistics for the behavioral sciences. Pacific Grove: Brooks/Cole Pub. Co; 1996.
33. Cicchetti D, Bronen R, Spencer S, Haut S, Berg A, Oliver P, et al. Rating scales, scales of measurement, issues of reliability: resolving some critical issues for clinicians and researchers. J Nerv Ment Dis 2006;194:557-64.

Appendix A. The Significant Results in Tukey HSD Test

Age Groups	Age Groups	Mean Difference	Std. Error	p	Confidence Interval 95%	
					Lower Bound	Upper Bound
30-35	36-41	-1.61	0.41	0.003	-0.82	-0.39
	42-47	-2.38	0.42	0.001	-3.63	-1.14
	48-53	-2.72	0.41	0.001	-3.93	-1.51
	54-59	-3.98	0.42	0.001	-5.21	-2.76
	60-65	-3.84	0.42	0.001	-5.08	-2.6
36-41	30-35	-1.61	0.41	0.003	-0.82	-0.39
	54-59	-2.37	0.42	0.001	-3.60	-1.15
	60-65	-2.23	0.42	0.001	-3.47	0.98
42-47	30-35	2.38	0.42	0.001	1.14	3.63
	54-59	-2.59	0.43	0.005	-2.85	- 0.33
	60-65	-1.45	0.44	0.016	-2.73	-0.17
48-53	30-35	2.38	0.42	0.001	1.14	3.63
	54-59	-2.59	0.43	0.005	-2.85	-0.33
54-59	30-35	3.98	0.42	0.001	2.76	5.21
	36-41	2.37	0.42	0.001	1.15	3.60
	42-47	1.59	0.43	0.005	0.33	2.85
	48-53	1.26	0.42	0.004	0. 35	2.48
60-65	30-35	3.84	0.42	0.001	2.60	5.08
	36-41	2.23	0.42	0.001	0.98	3.47
	42-47	1.45	0.44	0.016	0.17	2.73