



REYHANE MOHAMADI¹, MAJID RAFIEE ALAVIJE¹, ASGHAR MINAYI²,
YAHYA MODARESI³, MEHDI KAZEMI DASTJERDI⁴, MAHDI GHADERI⁵

¹ University of Social Welfare and Rehabilitation Science, Tehran

² Allameh Tabataba'i University, Tehran

³ Institute for Humanities and Cultural Studies, Tehran

⁴ Research Institute of Education, Tehran

⁵ Qazvin University, Qazvin

GENERATION AND CONTENT VALIDATION OF A PERSIAN SYNTAX COMPREHENSION TEST

Lack of formal tests for assessing Persian language proficiency is one of the main problems for speech and language pathologists in Iran. The purpose of this study was to generate an item pool for a syntax comprehension test based on the characteristics of the Persian language and Iranian culture, as well as the spoken variant for 4-6 years old native Persian speaking children. We first extracted 41 syntactic structures of Persian for the syntax comprehension test, of which 8 structures were excluded in the first phase. Then, 198 items were developed for the remaining 33 syntactic structures. To determine content validity, 14 experts assessed the structures and 12 experts expressed their views on the items. Content Validity Ratio (CVR) was calculated. 24 structures with CVR > 0.50 were selected. 107 items were selected based on their CVR value (CVR > 0.47). Eleven new items were regenerated to replace items with a CVR below 0.47. According to the results, the test turned out to have a good content validity.

Key words: item pool, content validity, syntactic structure, test development

Introduction

Generally, linguistic knowledge is evaluated and analyzed in five areas of grammar, syntax, phonology, semantics and pragmatics, as well as in two levels

of comprehension and production. The most common way to assess expressive syntax is using speech samples (Kemp & Klee, 1997; Paul, 2007), but assessing linguistic knowledge at the comprehension level has some advantages. The first advantage is that it makes linguistic assessment feasible in preverbal children. Second, even children who can speak do not necessarily use all comprehensible linguistic structures in their speech. Comprehension assessment, the evaluation of grammatical structures is made possible before it emerges in speech. The third advantage is the ability to apply more control over comprehension assessments (Hirsh-Pasek & Golinkoff, 1996; McDaniel, McKee, & Cairns, 1998). Since comprehension assessment usually presents a decontextualized framework, it contributes to an accurate examination of linguistic knowledge (Hirsh-Pasek & Golinkoff, 1996; McDaniel et al., 1998; Paul, 2007).

The benefit of formal and norm-referenced test administration is obvious to researchers and clinicians. The lack of specialized tests in Persian has made Iranian speech-language pathologists utilize self-made tasks for conducting most of their research or clinical activities. Under such circumstances, although developing many of these tasks takes much time, they are not usable in other studies or even clinical activities due to various reasons, such as the limited age range and small size of the study samples, failure in covering the desired specialized field, and failure to investigate complete psychometric characteristics. Although the role of informal assessments in speech and language disorders cannot be ignored, complete and excessive dependence on these assessments, which generally rely on therapists and require a high degree of clinical skill and experience (Shipley & McAfee, 2009), can lead to improper clinical judgments and inaccurate research findings. Thus, conducting accurate and rigorous research requires formal and norm-referenced tests.

Adaptation of tests from one language and culture to another language and culture is a common scientific practice, which has some benefits such as facilitating comparative studies and reducing time and cost (Hambleton, 1993; Van de Vijver & Leung, 1997). However, some cross-cultural researchers have admitted that a high percentage of the studies in this area are invalid due to weakness in test adaptation and lack of construct equivalence. In test adaptation, sources of error are divided into three general areas: 1 – Cultural and linguistic differences, 2 – Design and methods, and 3 – Interpretation of the results (Hambleton, Merenda, & Spielberger, 2004). It cannot be easily assumed that instruments developed for measuring concepts and constructs of a language and culture can be readily used in other languages and cultures (Hui & Triandis, 1985).

The comparability of scores and measures of adapted measurement tools and tests depends on their validity and equivalence level (Hui & Triandis, 1985; Van de Vijver & Leung, 1997). One of the necessary conditions for the application of language tests belonging to another language and culture is equalizing its concepts and constructs. Thus, there is always the danger that a valid test version would not be reproduced in another language. This danger seems to be

greater if language itself is the subject of study and testing. Linguistic issues in the field of speech therapy are not only strongly culture-dependent, but also completely language-dependent.

The considerations above make the norms obtained in other countries not usable in Iran, where the official language is Persian. Modern Persian belongs to the Iranian language family that is a branch of the Indo-Iranian languages. Persian is a null subject language and is generally verb-final. The unmarked Persian word order is SOV, but the order of constituents, especially in the spoken language, is free (Karimi, 2005). Unlike English, in Persian the object is placed before the verb. English language lacks any object markers, but in Persian “*rä*” acts as the object marker. Persian, compared to English language, has a rich inflectional system; morphological inflection of verbs in Persian is especially rich and complex. Persian is a pro-drop language, where pronominal subjects are often omitted (Mahootian, 1997).

Clearly, the developmental assessment tools of language and syntactic comprehension that are available in English contain elements that are not compatible with the Iranian culture and its linguistic structures. Thus, the direct translation or back translation and adaption of language tests in English into Persian does not bridge the existing gap in Iran, and it may lead to the creation of a tool that is by no means suitable for assessing native Persian children in the context of Iranian culture. Theoretical studies have shown that culture has an impact on children’s development, particularly on their social development (Gladstone et al., 2008), however useful taking advantage of the general framework of the existing tests could be. Unique features of Persian require a completely native test in accordance with its unique features. Therefore, this study was aimed at generating an item pool for the development of a comprehension test of Persian syntax. Although the steps provided in this article are the first and of course the most important ones in developing a test, the authors are aware that in order to name a set of items a test, in addition to the content validity, an assessment of the items, their difficulty and discriminative index are needed. Items with good, etc. content validity may be removed in this process due to inappropriate difficulty or discriminative indices. In addition, a test should be of an acceptable validity and reliability. Indeed, in this article we explain the necessary stages for the item pool generation and examine their content validity to be applied as a comprehension test of Persian syntax. Thus, we examine the content validity of the items as well. Therefore, in this paper the term “test” means an initial version of items developed for testing Persian language syntax comprehension and evaluating their content validity.

Test development

The steps of Persian syntax comprehension test development were conducted based on a model of systematic test development (Downing, 2006).

Although these steps are listed linearly and successively from the beginning to the end, in practice some of these steps may be conducted simultaneously. This model is composed of 12 steps. The five steps followed in this research are presented here in order.

Step 1: Overall plan

Due to the lack of syntactic comprehension tests and the importance of evaluating the comprehension of syntactic structures in pre-school children, a decision was made to develop a receptive language test to specifically evaluate the comprehension of Persian syntax. In the first step, the construct of Persian syntax was defined as a latent variable, and content domains of the test were determined (DeVellis, 2003; Lawshe, 1975). It is suggested that content domains should be specified and then necessary steps should be taken for the content and construct validities. In this test, syntax refers to word order in sentences and processes that show grammatical information of person, number and tense of a word. The general format of the test was defined. The syntax comprehension test is a paper and pencil test that uses visual stimuli (pictures) to elicit a response to measure one of the subsystems of the Persian language, namely syntax.

Step 2: Content definition

Since the syntax of any language consists of different structures, it must first be determined which syntactic structures of Persian should be included in the syntax comprehension test. Three main criteria for the selection of the structures were considered, as follows:

1. Frequency: Structures with different frequencies were selected, but more frequent structures were prioritized. Since there was no study on the frequency of structures in Persian, highly frequent structures were identified by their early emergence in 2-6 years old children's speech and linguists' opinion on the frequency of these structures.
2. Structural simplicity/complexity: In addition to simple structures, we considered complex structures presumably comprehended later (Rahmany, Marefat, & Kidd, 2011, 2013).
3. Clinical significance: Clinical experiments and research findings confirm the importance of some structures. If a structure was of low-frequency but important, it was included in the early version of the test.

Step 3: Test specification

At this stage the test characteristics were determined. Syntax comprehension is a receptive language test. For each item there were four simple, clear and colored pictures. We included four pictures for each item because it reduced the chance level of correct responses. Items and images were designed to have minimal

redundancy to reduce nonverbal cues. A research assistant was asked to read the items based on the spoken variant of Persian. The choice of speech variant was made following consultations with experts in the field of speech and language. In Persian there is a kind of spoken language that is a mixture of both oral and written. Here, the structure of grammar and vocabulary relates to the written language and pronunciation, and phonetic features including pronunciation of sounds, intonation and stress relate to the spoken language. We used such a speech variant because it was a more familiar variant for children aged 4-6 years old, who did not yet begin formal instructions in reading and writing, compared to the formal written variant (Sammie & Razavi, personal communication). In fact, one of the most important characteristics of this test that differentiates it from other tests is that the research assistant read the items to the examinees using the spoken variant. After hearing each item the examinee pointed to one of the four pictures or said the picture number. Unlike the Persian version of the bilingual aphasia test and the Test of Language Development-Primary, 3rd Edition (TOLD-P: 3), in our test we did not use individual items to assess comprehension, but we first identified some syntactic structures and put the items as sub-sets to the syntactic structures.

Step 4: Item development

For the fourth step, three consecutive phases were conducted: 1 – Determining syntactic structures, 2 – Screening syntactic structures, 3- Item pool generation, and 4 – Determining content validity.

Phase 1: Syntactic Structure determination

In order to select syntactic structures according to the three above-mentioned criteria, 7 steps were followed. 1 – Books and articles about Persian grammar were reviewed (Anvary & Givy, 2012; Arzhang, 2008; Mahootian, 1997; Meshkato Dini, 2010). 2 – Articles and books related to syntactic development in Persian were reviewed and some fundamental structures produced in the speech of children under 5 years old were extracted. It is worth noting here that most developmental studies in Persian have been conducted on production, and little research is available on comprehension of syntax by children (Rahmany et al., 2011, 2013). 3 – Empirical evidence for the target groups of this test was collected and their tasks were reviewed (Chapman & Kohn, 1978; Rahmany et al., 2011, 2013; van der Lely & Harris, 1990). 4 – The clinical experience of the research team was considered. 5 – A database obtained from the study of core vocabulary in Persian was used to analyze the oral statements of 70 children in the first grade of primary school to find some extra syntactic structures (Nematzadeh, Dadras, Dastjerdi Kazemi, & Mansoorizadeh, 2011). 6 – A semi-structured face to face interview was conducted with six experts in Persian linguistics and speech therapy to find out their views concerning important syntactic structures in Persian. Experts were selected based on their publications and research and/or clinical expertise. Both

linguists and speech experts were interviewed in order to consider theoretical views and clinical issues. 7 – Tasks in Persian were based on previous studies (Rahmany et al., 2011, 2013). The Persian version of the bilingual aphasia test (Paradis, Paribakht, & Nilipour, 1987), the English version of the Test of Language Development-Primary: 3rd Edition (TOLD-P: 3) (Newcomer & Hammill, 1997), and the Test for Reception of Grammar-Version 2 (TROG-2) (Bishop, 2003) were also reviewed to extract the structures. In cases where the English tasks or tests were used, the items were not translated into Persian. However, taking into consideration the availability of syntactic constructions in Persian, the items were adapted. In many cases, the syntactic structure of the item was used and items for that syntactic structure were created based on Persian core words.

These procedures aimed to identify all possible structures that could be used in this test. We extracted 23 structures based on a database of core Persian vocabulary, a review of books on Persian grammar, the literature on development of syntax, the clinical experience of the research team, and empirical evidence from Persian concerning the comprehension of test items by those with hearing impairment and speech and language disorders. Five new structures were extracted from expert interviews. 13 new structures were extracted through studying the available tests and tasks, which were then adapted to suit our purpose. Finally, 41 syntactic structures that were considered suitable to be included in the test were extracted.

Phase 2: Structure reduction

In this phase, some structures were excluded for various reasons : one structure was removed because it was marked in Persian (scrambling), as we preferred not to use marked structures in the test. Also, five structures (one-person intransitive verbs, indirect quotation, cause clauses, and “from”... “to” preposition, either... or...) were eliminated due to the impossibility of converting them to pictures. Two structures were removed (bound pronouns in the function of genitive and complement) due to their overlap with another structure. The final number of included structures was 33.

Phase 3: Item pool generation

The goal of this phase was to develop items that represent well-formed sentences in Persian. On average, 6 items were generated for each structure. The number of expected items for the test was about 90 to 100. So, at this stage 198 items were generated, which is nearly twice the final items. We determined the criteria for generating the items as follows:

1. *Limited variety of vocabulary*: the first criterion was that the items should be generated with a limited number of content words, because our comprehension test was intended to evaluate syntactic relations and not the

- lexical ones. The use of different words in syntactic tests, especially for children before school age, increases the risk of giving incorrect answers due to their inability to comprehend the words used in the tests.
2. *Use of core words*: the second criterion taken into account to generate the items of the syntax comprehension test was for the words used in the tests. All of the words were selected on the basis of Persian developmental studies and core words. This criterion, along with the first criterion, were used to minimize the effects of word comprehension on comprehending syntax. Research has shown the impact of lexical frequency on sentence comprehension by children with specific language impairment (Leclercq, Majerus, Jacob, & Maillart, 2014).
 3. *Ease of conversion to representative pictures*: Items were generated so that they could be illustrated clearly with no ambiguities.
 4. *Reversibility*: items were generated so that they were reversible (unless the goal was to examine irreversible sentences). In assessing syntax comprehension, sentences in which agent and patient can replace each other and in which this replacement can cause semantic differentiations are called reversible (O'Grady, 2005). Under such circumstances generation of items is difficult and time consuming, but precise and specific assessment of syntax is possible.
 5. *Compliance with features of Persian*: The last criterion taken into account in generating the items was for the items to be generated based on the Persian syntactic sentences and to show no impact of translation from another language, especially from English.

The preliminary item pool, which included 198 items, was made up of 1032 words (108 different content words). To choose nouns, adjectives, pronouns, adverbs and verbs we used Persian core words of Iranian children's speech. The nouns included concrete objects and animals familiar to Iranian children and also humans (girl, boy, man, woman). 68% of verbs were action verbs and 32% of verbs were attributive. 30% of action verbs were compounded. The items usually described the relationship between a man and an animal, two men, a man and objects, animals and objects, and objects.

Phase 4: Determining content validity

In the current study, experts were asked to present their views concerning syntactic structures and items. To examine the content validity of the test, a list of syntactic structures and items from the test was presented to 14 experts (5 speech therapists and 9 linguists), in two separate forms. Selection of the experts was based on their specialized field and history of research and clinical expertise. Given that a number of experts lived in other cities and countries, we were not able to hold a panel of experts. Thus, the forms were sent to their

workplaces via mail or in face to face meetings. Furthermore, all explanations as to the purpose of the test and how to submit written comments concerning the structures and items were presented in writing. In delivering forms to the experts, all of the explanations were also presented orally or via phone and any questions or ambiguity in the test were answered. Concerning the structures and items, experts should have expressed their views on the appropriateness of each structure or item pertinent to their area of expertise, and also should have chosen one of the three options including essential, essential but not useful, and not necessary in the case of any structure and item. The forms concerning the syntactic structures were completed by all 14 experts, the item forms were filled out by 12 experts. Experts were given the opportunity to provide any modification of structures or items and to suggest additional views at the end of each form. Systematic analysis (item-by-item analysis) of the experts' ratings was conducted by the research team and decisions on excluding, revising, amending or replacing the new items were made.

Phase 4-1: Content validity of structures

In this study, to determine the content validity of 33 structures and 198 items, the study of Ayre and Scally (2014) was used. In their study, the content validity ratio (CVR) originally proposed by Lawshe (1979) was modified (Ayre & Scally, 2014). The minimal values of CVR were 0.47 and 0.50 for structures and items, based on 14 and 12 experts' reviews (Ayre & Scally, 2014). As it is clear from Table 1, structures 25-33 were excluded from the final list of structures.

Table 1. CVR values of the judgments and results of acceptance or rejection of syntactic structures

Structure number	Structure	CVR
1	reversible sov	0.86
2	intransitive basic sentence	1.00
3	simple negative sentences	1.00
4	subject relative clauses	1.00
5	transitive active simple sentences	1.00
6	prepositional phrases	1.00
8	direct object relative-clauses	1.00
9	comparative adjectives	1.00

10	passive sentences	0.71
11	omitted object in compound sentences	0.86
12	free pronouns	1.00
13	omitted subject in compound sentences	1.00
14	negative conjunction in compound sentences	0.67
15	A phrase not B phrase	0.57
16	locative adverbs	0.67
17	positive conjunction in compound sentences	1.00
18	tense-aspect-mood of verbs	0.71
19	bound pronoun of verbs	0.71
20	adjective genitive sequences	0.71
21	object deletion	1.00
22	two object verbs	0.71
23	superlative adjectives	0.71
24	subject-verb agreement	1.00
25	reflexive pronouns	0.43
26	causative verb	-0.71
27	subject cleft sentences	0.00
28	object cleft sentences	0.00
29	object deletion with different subject	0.29
30	left dislocation	0.14
31	indefinite pronouns	0.14
32	conditional sentences	0.14
33	not only... but also	-0.14
CVI: 0.87		

Phase 4-2: Content validity of items

Among 198 designed items, 107 items were selected on the basis of their CVR values (CVR > 0.50). 11 new items were generated instead of the items whose CVR values were below 0.50 (4 items for passive sentences, 1 item for omitted subject in compound sentences, 2 items for reversible SOV, 2 items for tense-aspect-mood of verbs, and 1 item for object deletion). 2 items (two object verbs) had CVR values

between 0.33 and 0.50, and thus had to be replaced with new items. Due to the fact that making new reversible items was impossible, pictures were designed for these items so as to be kept or excluded after the result of the pilot study.

Table 2. Shows the mean CVR for accepted items for each structure

Structure number	Structure number	CVR
1	reversible sov	0.70
2	intransitive basic sentence	0.75
3	simple negative sentences	0.78
4	subject relative clauses	0.79
5	transitive active simple sentences	0.76
6	sentences with noun coordinated phrase	0.66
7	prepositional phrases	0.66
8	direct object relative-clauses	0.73
9	comparative adjectives	0.83
10	passive sentences	1.00
11	omitted object in compound sentences	0.83
12	free pronouns	0.69
13	omitted subject in compound sentences	0.63
14	negative conjunction in compound sentences	0.80
15	A phrase not B phrase	0.67
16	locative adverbs	0.67
17	positive conjunction in compound sentences	0.95
18	18- tense-aspect-mood of verbs	0.75
19	bound pronoun of verbs	0.71
20	adjective genitive sequences	0.62
21	object deletion	0.61
22	two object verbs	0.71
23	superlative adjectives	0.67
24	subject-verb agreement	0.86

Step 5: Design of test pictures

We used a picture selection test with four pictures for each item. 120 four-picture items were designed based on the Iranian culture; two pictures on the top half of the page and two in the bottom half. The pictures were not cartoons, but all were colored. An attempt was made to make colors and faces of characters as attractive to 4-6 year old children as possible. The target and distractor pictures were of particular importance. Distractor pictures were semantic or syntactic or both, depending on the item. The distractors for structures 1, 2, 3, 5, 9, 10, 12, 19, and 23 (Table 1) were semantic.

After drawing four pictures for each item, corrective feedback was provided from the authors of the study. Modifications were made to the pictures accordingly. In the next step, the pictures were evaluated by the research team. The resolution, color, position and alignment of the target picture and three distractor pictures on the page were examined and corrective feedback was provided. All comments were collected and necessary changes were made to the pictures. Checking the pictures continued until they were considered as clear and expressive, and necessary modifications were made at each stage. After the three-stage modification of the pictures, they were given to five speech and language pathologists and a professional photographer. They expressed their considerations about resolution by deciding on whether the pictures were clear, relatively clear or unclear. They were also asked what to do to clarify the pictures. Their comments were examined by the research team and were conveyed to the illustrator. The agreement between ratings concerning the clarity of pictures was calculated using SPSS 16 (contingency coefficient 0.77, $p < 0.001$).

Pilot study

A pilot study was then performed to make decisions on eliminating some items. The test was administered by the first author to thirty 4-6 year old children in a quiet room. The goal of the pilot study was to analyze the items and pictures qualitatively; therefore not only children's responses to test items, but also all verbal statements through the test administration were recorded. The results of the pilot study were as follows: the pictures for 16 items were modified. Necessary revisions were made, including deleting some distracting frills, changing colors and their intensity, changing the position of the characters' hands and legs and the characters' postures illustrating the verbs, fitting or equalizing the size of the four images for an item. Comprehension of two words (a verb and a noun) was difficult for some children and influenced the sentence comprehension. Thus they were eliminated and replaced with more appropriate ones. Six items related to structures that included more than 5 items were deleted.

Discussion

The purpose of the current study was to generate an item pool for a Persian syntax comprehension test. The syntactic structures for the Persian comprehension test were extracted through a systematic process, an interview with an expert team, and using books on Persian syntax as well as available tests and tasks. Given that it is the first syntax comprehension test in Persian, it was necessary to design various structures that had different levels of complexity and ultimately choosing the structures according to the experts' views. After receiving the experts' reviews 8 structures were deleted, leaving us with 24 structures that were suitable for evaluating preschool children's comprehension. The experts believed that 6 out of 8 structures were too difficult for preschool children. Approximately half of the structures (11 structures) were selected to be included in the test by all 14 experts. The other half of the structures (except for structures 14 and 16) with a content validity ratio of greater than 0.70 were included in the test. The reason why the experts deleted structures 25 and 31 was that they mostly showed lexical and semantic relations, and did not have some of the syntactic features required in this test. The mean content validity ratios of test items were greater than 0.60, which indicates relatively good content validity.

Conclusions

The current study described in detail the successive steps in item generation and content validation for developing a new test in Persian. The syntax comprehension test is a test that specifically deals with syntax comprehension in Persian. The syntactic structures and items are fully based on characteristics of Persian, and an attempt to draw pictures based on the Iranian culture, as far as possible, was also made. According to the results, the test has appropriate content validity.

Future development of syntax test

In the next step, as a second pilot study, the items will be administered to 100 children aged 4 to 6 years old whose native language is Persian, and difficulty and discrimination indices will be calculated. All inappropriate items will be eliminated according to their discriminative index. The structures and the items of each structure will be arranged in an ascending order of difficulty, the final format of the test will be prepared for administration on a large sample of 4-6 Persian children, and the psychometric characteristics of the test will be examined.

Acknowledgements

This study was supported by University of Social Welfare And Rehabilitation Science with ethical code USWR.REC.1393.10. The authors appreciate all profes-

sors and experts that helped us with their scientific opinions on the structures and content validity of this study, especially Dr. Hossein Samei, Dr. Mohammad Reza Razavi and Dr. Shahin Nematzadeh.

References

- Anvary, H. & Ahmadiye Givy, H. (2012). *Persian Grammar 1,2 (4th ed.)*. Tehran: Fatemi Publishing.
- Arzhang, G.H. (2008). *Modern Persian Grammar*. Tehran: Ghatreh Publishing.
- Ayre, C. & Scally, A.J. (2014). Critical Values for Lawshe's Content Validity Ratio: Revisiting the Original Methods of Calculation. *Measurement and Evaluation in Counseling and Development*, 47 (1), 79–86.
- Bishop, D.V.M. (2003). *Test for Reception of Grammar: Version 2 (TROG-2)*. London: Pearson Assessment.
- Chapman, R.S. & Kohn, L.L. (1978). Comprehension strategies in two and three year olds: animate agents or probable events? *Journal of Speech And Hearing Research*, 21 (4), 746–761.
- DeVellis, R.F. (2003). *Scale Development: Theory and Applications (2nd Ed.)*. Thousand Oaks, CA: Sage.
- Downing, S.M. (2006). Twelve Steps for Effective Test Development. In S.M. Downing & T.M. Haladyna (Eds.), *Handbook of Test Development* (pp. 3–25). Mahwah, NJ: Lawrence Erlbaum Associates.
- Gladstone, M.J., Lancaster, G.A., Jones, A.P., Maleta, K., Mtitimila, E., Ashorn, P., & Smyth, R.L. (2008). Can Western developmental screening tools be modified for use in a rural Malawian setting? *Archives of Disease in Childhood*, 93 (1), 23–29.
- Hambleton, R.K. (1993). Translating achievement tests for use in cross-national studies. *European Journal of Psychological Assessment*, 9 (1), 57–68.
- Hambleton, R.K., Merenda, P.F., & Spielberger, C.D. (2004). *Adapting Educational and Psychological Tests for Cross-Cultural Assessment*. Mahwah: Lawrence Erlbaum Associates.
- Hirsh-Pasek, K. & Golinkoff, R.M. (1996). *The Origins of Grammar: Evidence from Early Language Comprehension*. Cambridge, MA: MIT Press.
- Hui, C.H. & Triandis, H.C. (1985). Measurement in Cross-Cultural Psychology: A Review and Comparison of Strategies. *Journal of Cross-Cultural Psychology*, 16 (2), 131–152.
- Karimi, S. (2005). *A Minimalist Approach to Scrambling: Evidence from Persian*. Berlin: Mouton de Gruyter.
- Kemp, K. & Klee, T. (1997). Clinical language sampling practices: results of a survey of speech-language pathologists in the United States. *Child Language Teaching and Therapy*, 13 (2), 161–176.
- Lawshe, C.H. (1975). A Quantitative Approach To Content Validity. *Personnel Psychology*, 28 (4), 563–575.

- Leclercq, A., Majerus, S., Jacob, L., & Maillart, C. (2014). The impact of lexical frequency on sentence comprehension in children with specific language impairment. *Research in Developmental Disabilities, 35* (2), 472–481.
- Mahootian, S.H. (1997). *Persian Descriptive Grammars*. London: Routledge.
- McDaniel, D., McKee, C., & Cairns, H.S. (1998). *Methods for Assessing Children's Syntax*. Cambridge, MA: MIT Press.
- Meshkato Dini, M. (2010). *Introduction to persian transformational syntax (2nd ed.)*. Mashhad: Ferdowsi University of Mashhad Press.
- Nematzadeh, S.H., Dadras, M., Dastjerdi Kazemi, M., & Mansoorizadeh, M. (2011). *Persian Core Vocabulary Based on Iranian Children*. Tehran: Madreseh.
- Newcomer, P.L. & Hammill, D.D. (1997). *The Test of Language Development-Primary: Third Edition (TOLD-P:3)*. Austin, TX: Pro-Ed.
- O'Grady, W. (2005). *How Children Learn Language (Cambridge Approaches to Linguistics)*. Cambridge: Cambridge University Press.
- Paradis, M., Paribakht, T., & Nilipour, R. (1987). *The Bilingual Aphasia Test (Farsi Version)*. Hillsdale, NJ: Lawrence Erlbaum.
- Paul, R. (2007). *Language Disorders from Infancy Through Adolescence: Assessment & Intervention (3rd ed.)*. St. Louis, MO: Mosby.
- Rahmany, R., Marefat, H., & Kidd, E. (2011). Persian speaking children's acquisition of relative clauses. *European Journal of Developmental Psychology, 8* (3), 367–388.
- Rahmany, R., Marefat, H., & Kidd, E. (2013). Resumptive elements aid comprehension of object relative clauses: evidence from Persian. *Journal of Child Language, 41* (4), 937–948.
- Shipley, K. & McAfee, J. (2009). *Assessment in Speech-Language Pathology: A Resource Manual (4th ed.)*. Clifton park, NY: Cengage Delmar Learning.
- Van de Vijver, F.J.R. & Leung, K. (1997). Methods and data analysis of comparative research. Vol. 1: Theory and method. In J.W. Berry, Y.H. Poortinga & J. Pandey (Eds.), *Handbook of Cross-Cultural Psychology (2nd ed., pp. 257–300)*. Des Moines, IA: Allyn & Bacon.
- Van der Lely, H.K. & Harris, M. (1990). Comprehension of reversible sentences in specifically language-impaired children. *Journal of speech and hearing disorders, 55* (1), 101–117.

Figure 1. Phases of item pool generation for development of syntax comprehension test

