



# Research Paper

## Development of the Persian Northwestern Anagram Test: Content Validation



Fahimeh Poormohammadi<sup>1</sup> , Reyhane Mohamadi<sup>1</sup> , Mohammad Hassan Torabi<sup>2</sup>, Mona Ebrahimipour<sup>3</sup> , Arezoo Saffarian<sup>1\*</sup>

1. Department of Speech Therapy, Rehabilitation Research Centre, School of Rehabilitation Science, Iran University of Medical Sciences, Tehran, Iran.

2. Saadi Foundation, Tehran, Iran.

3. Elisabeth Bruyere Hospital, Bruyere Health Services, Ottawa, Canada.



Copyright: © 2025 The Author(s). This is an open access article distributed under the terms of the [CC BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/)

### ABSTRACT

**Background and Objectives:** Agrammatism, a disorder characterized by difficulties in producing complex syntax, reduced verb use, and the omission of grammatical morphemes, significantly impacts speech production and comprehension. Given the prevalence and impact of grammatical deficits, assessing agrammatism is essential for diagnosing and treating affected individuals. This study aimed to translate, culturally adapt, and develop the northwestern anagram test (NAT) in Persian to assess agrammatism in Persian-speaking populations.

**Methods:** The syntactic structures of Persian were analyzed to design test items, including 33 sentences and related images. In addition, content validity was evaluated by 11 experts, comprising seven speech-language pathologists and four linguists, using the content validity ratio (CVR) and content validity index (CVI). Sentences and images that scored below threshold values were revised or replaced, and the mean CVR and CVI scores were calculated for the remaining items to confirm validity.

**Results:** The Persian version of NAT (P-NAT) demonstrated satisfactory content validity, with a mean CVR of 0.85 and a CVI exceeding 0.79 for all items. Expert feedback was incorporated to refine test items and illustrations, ensuring cultural and linguistic appropriateness.

**Conclusion:** P-NAT is a promising tool for assessing agrammatism. Future research should evaluate its psychometric properties, including construct validity and reliability, to establish its clinical applicability.

**Keywords:** Agrammatism, Psychometrics, Grammar, Persian

#### Article info:

**Received:** 05 Jan 2025

**Accepted:** 25 Jan 2025

**Available Online:** 10 Nov 2025



Use your device to scan and read the article online

**Cite this article as** Poormohammadi F, Mohamadi R, Torabi MH, Ebrahimipour M, Saffarian A. Development of the Persian Northwestern Anagram Test: Content Validation. *Function and Disability Journal*. 2025; 8:E302. <http://dx.doi.org/10.32598/fdj.8.302>

<http://dx.doi.org/10.32598/fdj.8.302>

#### \* Corresponding Author:

**Arezoo Saffarian, Associate Professor.**

**Address:** Department of Speech Therapy, Rehabilitation Research Centre, School of Rehabilitation Science, Iran University of Medical Sciences, Tehran, Iran

**Tel:** +98 (912) 056784

**E-mail:** [saffarian.a@iums.ac.ir](mailto:saffarian.a@iums.ac.ir); [asaffarian\\_slp@yahoo.com](mailto:asaffarian_slp@yahoo.com)

↑ *What is “already known” in this topic:*

*Agrammatism is a prevalent feature of various aphasia types, characterized by impaired syntactic processing. The northwestern anagram test (NAT) is a brief, validated tool for evaluating sentence production and has been adapted into multiple languages. However, there is currently no standardized Persian version available for clinical or research use.*

→ *What this article adds:*

*This study developed and validated the Persian version of the NAT (P-NAT) to assess syntactic production in individuals with agrammatism. The P-NAT provides a reliable, culturally adapted tool for clinical evaluation, enabling cross-linguistic comparisons in agrammatism research.*

## Introduction

Language is a vital tool for communication, encompassing both understanding and expression. It consists of three main components: form (syntax, morphology, and phonology), content (vocabulary and semantics), and use (pragmatics) [1, 2].

Grammar plays a crucial role in language structure, governing how individuals comprehend and produce speech [3-5]. Agrammatism, a disorder characterized by difficulties in producing complex syntax, reduced verb usage, and omission of grammatical morphemes, significantly impacts speech production and comprehension [6, 7]. For instance, a person with agrammatism may struggle to understand sentences, such as “A girl pushes a boy,” and may find it difficult to distinguish between intended and reversed meanings, such as “a boy pushing a girl” [8].

In Persian, research on agrammatism has revealed patterns similar to those found in other languages, including reliance on nouns and the use of /ast/ (“is”) as a default verb. Additionally, the present tense is often utilized as a default, even though it is longer than the corresponding past tense forms. For example, the sentence /mard [darad] sombane ast/ (“[the] man is breakfast.”) demonstrates the use of auxiliary verbs and noun phrase constructions that omit more complex syntactic structures [9].

Agrammatism is commonly associated with nonfluent aphasias, such as Broca’s aphasia and transcortical motor aphasia, which result from vascular lesions in the left frontal lobe [10]. However, deficits in syntactic processing also occur in other aphasia types, including anomic

and Wernicke’s aphasia, indicating that agrammatism is not confined to Broca’s aphasia. Nearly 25% of stroke survivors with aphasia experience agrammatism [11-14]. Furthermore, agrammatism is observed in primary progressive aphasia (PPA), associated with neurodegenerative diseases (e.g. frontotemporal lobar degeneration and certain forms of Alzheimer’s disease). Frontotemporal lobar degeneration is relatively rare, with an estimated prevalence of 2.7–15 cases per 100,000 people [15].

Given the prevalence and impact of grammatical deficits, assessing agrammatism is essential for diagnosing and treating affected individuals. Current assessment methods for sentence comprehension and production include tasks, such as sentence-to-picture matching and picture description, which help evaluate syntactic understanding and expression. However, these methods have limitations, including challenges with spontaneous speech analysis and reliance on intact memory or speech production abilities [16, 17].

The northwestern anagram test (NAT), first developed by Weintraub et al. [16], provides an alternative assessment tool for evaluating syntactic expression in agrammatism, particularly in PPA patients. The NAT uses word cards to construct meaningful sentences and is available in both full and shortened versions. It has been adapted into Italian and German and is a valid, brief instrument for assessing sentence production [16]. This test is especially beneficial for patients with motor speech difficulties or cognitive impairments because it reduces the need for working memory and motor speech abilities.

To the best of our knowledge, there is currently no validated Persian tool for assessing syntactic production in individuals with agrammatism. Given the importance of

identifying grammatical deficits in clinical practice, this study aimed to create and validate a Persian version of the NAT (P-NAT). The development of the P-NAT will likely provide a reliable assessment tool for clinical settings and research, allowing for studies comparable to international research on agrammatism.

## Materials and Methods

This cross-sectional descriptive-analytical study followed a methodological approach consisting of five steps based on the “model of systematic test development” [18]. In this study, only the first two stages are reported: Extracting syntactic structures, creating test items, and determining content validity.

The administration protocol and scoring method for the P-NAT are identical to the original NAT [16]. The same principles have been followed for presenting images, arranging words, and recording responses. The instructions and steps for test administration, along with the scoring criteria, ensured consistency between the two tests. During administration, the participants are given a set time to respond (30 s), and the provided words must be used to form a sentence that matches the image. Practice trials were included to familiarize participants with the procedure, and scoring involved marking responses as correct or incorrect, calculating totals, and determining percentages for canonical and non-canonical sentences. The images employed in the P-NAT are designed to match the size of the original test, measuring 20×15 cm, to maintain uniformity.

### Step 1: Extracting syntactic structures and creating test items

In this stage, after correspondence with the original authors of the test, permission was obtained to use the NAT [16], and the Persian name of the test was also suggested by the original authors. The P-NAT is not merely a direct translation of the original test; it has been designed considering the Persian language’s high-frequency and relevant syntactic structures.

A combined approach was adopted in this study due to the lack of a definitive corpus of Persian syntax and the presence of differing and sometimes contradictory views among Persian linguists regarding Persian morphosyntactic structures. First, the syntactic structures of Persian were examined using prominent grammar books, articles, and language tests [19-28], along with consultations with various experts in Persian syntax. Moreover, efforts were made to achieve a consensus by synthesiz-

ing insights from these sources. In addition, interviews were conducted with individuals who were authorities in Persian linguistics, including members of the Persian Language and Literature Academy, to simultaneously consider cultural perspectives and clinical points.

High-frequency structures were selected based on this comprehensive analysis, and their necessity for testing was confirmed. The grammatical structures included in the P-NAT (active, passive, subjective cleft, objective cleft, subject-extracted wh-questions, and object-extracted wh-questions) are broadly aligned with the canonical and non-canonical structures in the original test. However, these structures have been specifically extracted and localized based on the unique patterns and features of the Persian language, ensuring that the items are both linguistically and clinically relevant. This process guaranteed that the syntactic structures were culturally and linguistically appropriate for Persian, forming the basis for extracting and designing P-NAT syntactic sentences.

Therefore, while the P-NAT conceptually aligns with the original test in terms of assessment goals, its detailed syntactic structures and sentences are tailored to Persian rather than identical to those of the original NAT. This process ensured that the syntactic structures were linguistically and culturally suitable for Persian, and the final set was selected based on their frequency and clinical relevance. In designing the test items, five to eight appropriate sentences were created for each syntactic structure and then included in the test. The criteria for constructing these sentences were as follows:

**Vocabulary limitation:** To focus exclusively on expression without involving working memory, the vocabulary was limited to ensure it remained manageable. This criterion was critical because assessing vocabulary expression and understanding was not among the test’s goals. Limiting vocabulary also enhances our ability to identify grammatical problems.

**Imageability:** The second criterion was that the sentences needed to be imageable, meaning they should be structured to allow for the creation of a mental image. This was essential because the sentences and associated images in the original test were ineffective when translated into Persian.

**Adaptation with Persian language characteristics:** Finally, the sentences had to be common and relevant to everyday Persian language use, thereby ensuring that the test items were linguistically and culturally appropriate.

Ultimately, clear black-and-white images were created for the sentences. All images were designed by a professional illustrator who had previously collaborated on similar tests and whose work was approved by the test development team. Further, efforts were made to ensure that the images were clear and simple, closely resembling the original test (as one of the original authors' conditions was to remain as faithful to the original test as possible) while also being tailored to the structure of the Persian language and Iranian culture. In the image design process, great care was taken to keep the background simple, free of any extraneous elements, and clearly representative of the corresponding sentences.

### Determining content validity

To assess the test's content validity, 13 experts were invited to participate in this study, including seven speech-language pathologists and six linguists. Of these, 11 experts accepted the invitation and provided their feedback. All participants had at least five years of clinical or research experience in Persian syntax with Persian-speaking adults.

To evaluate content validity, a list of syntactic structures, sentences, and test images was emailed to the experts, along with a detailed explanation of the test's purpose, scoring criteria for the syntactic structures and test sentences, and additional information to clarify any ambiguities or questions.

The experts were tasked with assessing the necessity of each syntactic structure in the test by categorizing it into one of three options: necessary, useful but unnecessary, or unnecessary. In addition, they were asked to provide specific comments on the sentences associated with the structures they deemed essential. Furthermore, a section was included next to each sentence for the experts to write their own comments and suggestions about improving the relevant syntactic structures or sentences.

Additionally, two practice sentences were included at the beginning of the test to prepare the participants. These sentences were not scored but were solely intended for this preparatory purpose. Content validity assessments were also conducted on these practice items. The minimum content validity ratio (CVR) was determined based on the Lawshe approach [29].

After calculating the CVR, the next step was to provide modified test images to 12 experts, including linguists and speech therapists who were not part of the research team. The experts were asked whether they believed the

images effectively conveyed the sentences for which they were designed and to suggest any changes that could enhance their clarity. At this stage, test images and a content validity table were sent to the experts via email. This message also included additional explanations of the test's purpose and scoring method, addressing any ambiguities or questions they might have had. Experts were instructed to evaluate each image on simplicity, relevance, and clarity. They assigned scores on a scale from one to five. Next to each image, there was a space for the experts to write their comments and suggestions for improvement. With these guidelines, they were encouraged to provide their feedback.

Eventually, the minimum content validity index (CVI) was determined based on the approach outlined by Waltz and Bausell [30].

### Results

Table 1 presents the CVR for Persian sentences designed for the P-NAT. The content validity of these sentences was assessed using the Lawshe model. According to this model, the CVR is calculated using Equation 1 as follows:

$$1. CVR = \frac{n_e - N/2}{N/2}$$

Where  $n_e$  represents the number of experts who deemed the item necessary, and  $N$  is the total number of experts involved in the evaluation process.

In general, if more than half of the experts consider an item essential, but not all, the CVR will range from 0 to 0.99. In this study, 11 experts participated in the evaluation, yielding a CVR of 0.63 using the Lawshe model [30]. Based on the results (Table 1), any sentences that scored below this value were to be either removed or reassessed based on content validity changes. Consequently, 12 sentences were removed from the list: Sentences 1, 2, and 3 from image 1; sentences 7, 10, and 11 from image 3; sentences 14, 15, and 16 from image 4; sentences 24 and 25 from image 5; and sentence 33 from the final image. The mean CVR score estimated for the remaining items was 0.85.

Alternative sentences were designed for the mentioned sentences. Then, three speech-language pathologists, each with at least five years of clinical experience in syntax, were interviewed to ensure the quality of the alternative sentences.

**Table 1.** CVR for items

Images		Sentences	CVR
English	Persian		
Hugging	1/baghal kardan/	1	0.45
		2	0.45
		3	0.45
Tickling	2/ghelghelak dadan/	4	0.81
Chasing	3/donbal kardan/	5	0.63
		6	1
		7	0.45
		8	0.81
		9	0.81
		10	0.45
		11	0.45
		12	1
Watching	4/didan/	13	1
		14	0.45
		15	0.09
		16	0.27
		17	1
		18	0.63
		19	1
Saving	5/nejat dadan/	20	0.81
		21	1
		22	1
		23	0.81
		24	0.27
		25	0.09
		26	0.63
Pulling	6/keshidan/	27	1
		28	0.81
		29	0.63
		30	0.63
		31	1

Images		Sentences	CVR
English	Persian		
Kissing	7/busidan/	32	0.81
		33	0.45
		34	1
		35	0.81
		36	1
		37	0.63

CVR: Content validity ratio.

Table 2 presents the CVI for the images designed for the P-NAT test. The CVI was calculated using the Waltz-Bausell model [30]. Equation 2 details the method for calculating the CVI. In this formula,  $n_e$  represents the number of experts who evaluated the images for relevance, clarity, and simplicity, giving them a score of 3 or 4 on a scale of 1-4. In addition,  $N$  denotes the total number of experts who participated in the content validity assessment. After evaluating each image for clarity, simplicity, and relevance to the provided sentences, the final CVI score was derived from the average of these three indices.

$$2. CVI = \frac{n_e}{N}$$

According to Equation 2 and the Waltz and Bausell model, an image is considered suitable if its mean score for clarity, simplicity, and relevance exceeds 0.79. The image requires modification if the average falls between 0.7 and 0.79. An image is deemed unacceptable and should be removed if its average score is below 0.7. In this study, 12 experts participated in the assessment and provided their evaluations. These experts reviewed all images, and their CVI scores were calculated. Furthermore, all images met the threshold scores determined by CVI (Table 2).

Table 2. CVI for Images

Images	CVI														Total Score	
	Relevance				Clarity				Simplicity							
	Not Relevant (Sentences)	Somewhat Relevant (to Sentences)	Relevant (to Sentences)	Completely Relevant	Score	Not Clear	Fairly Clear	Clear	Quite Clear	Score	Not Simple	Relatively Simple	Simple	Quite Simple		Score
1			1	11	1			3	9	1			3	9	1	1
2		1		11	0.92	1	1	2	8	0.83	1	1		10	0.83	0.86
3			1	11	1		2	1	9	0.83		2		10	0.83	0.87
4			2	10	1		2	2	8	0.83		3		9	0.75	0.86
5		1	1	10	0.92		2	3	8	0.92		1	2	9	0.92	0.92
6				12	1			2	10	1			2	9	0.92	0.97
7		1		11	0.92		1	2	9	0.92		1	1	9	0.83	0.89

CVI: Content validity index.



## Discussion

The current study provides a detailed account of the preliminary steps involved in constructing and assessing the content validity of the P-NAT items. The results indicate that the P-NAT sentences and pictures possess acceptable content validity.

Content validity is a critical aspect of test development, ensuring that test items accurately reflect the construct they are intended to measure. In this study, the mean CVR of 0.85 and a CVI exceeding 0.79 demonstrate the robustness of the P-NAT's design. More precisely, these metrics confirm that test items effectively capture the syntactic structures relevant to assessing agrammatism in Persian-speaking populations.

The strong content validity of the P-NAT is particularly significant because it provides a foundation for reliable and valid measurements in both clinical and research settings. By ensuring that the test items align with the target construct, the P-NAT minimizes the risk of measurement errors, thus enhancing its diagnostic utility.

Although the P-NAT has recently been developed and is not yet widely used, it was adapted from the NAT [31], which has been researched in several languages, including German and Italian [32, 33]. The translation and cultural adaptation of valid tests commonly utilized worldwide can benefit Iranian researchers by enabling them to use these tools in their studies. This alignment will yield results largely comparable to those obtained in studies conducted in other languages. Furthermore, the P-NAT specifically addresses the needs of Persian-speaking populations by incorporating linguistic and cultural considerations, thereby increasing its relevance and applicability.

While some studies have developed tasks and scales for assessing syntactic production in Persian, these tools typically require verbal expression [34]. In contrast, the P-NAT allows patients to construct sentences non-verbally, using vocabulary cards rather than verbal expression. This feature makes the P-NAT particularly useful for individuals who may have difficulties with verbal output, such as those with Broca's aphasia or PPA.

It is essential to discuss the potential benefits of this assessment tool for both research and clinical applications. The non-verbal approach of the P-NAT not only facilitates the assessment of syntactic abilities in individuals with severe speech impairments but also provides a standardized instrument for comparing results across dif-

ferent populations and languages. By addressing agrammatism in Persian-speaking individuals, the P-NAT bridges a gap in diagnostic tools, offering researchers and therapists a valuable resource for studying and treating this condition. Furthermore, the distinction between agrammatism in Broca's aphasia and PPA highlights the importance of targeted assessments (e.g. P-NAT), which enables clinicians to differentiate between these conditions more effectively.

Despite these strengths, the study had some limitations. One limitation was the lack of a widely accepted corpus for Persian syntax, which necessitated the reliance on expert opinions and available literature to design the test items. Additionally, the small sample size for the content validity assessment may have limited the generalizability of the findings. Thus, future studies should address these limitations by increasing sample size and evaluating P-NAT's performance across diverse clinical populations.

In the following stages of this research, we plan to administer the P-NAT in a population of individuals with agrammatism. Our goal is to evaluate its face validity, construct validity, and reliability. We also hope to contribute to improved diagnosis of agrammatism by establishing standard scores and publishing the test. Furthermore, comparative studies should explore similarities and differences in agrammatism between Broca's aphasia and PPA in Persian-speaking populations, thereby providing deeper insights into the nature of these disorders and the utility of the P-NAT.

## Conclusion

This study introduced the P-NAT and provided evidence of its content validity. The findings indicate that the P-NAT effectively captures the syntactic structures essential for assessing agrammatism in Persian-speaking individuals. By offering a culturally adapted, non-verbal measure of sentence construction, the P-NAT represents a valuable addition to the diagnostic tools available for Persian-speaking populations. Future research should further examine its psychometric properties, establish normative data, and explore its diagnostic utility across various clinical groups, particularly for differentiating Broca's aphasia from PPA.

## Ethical Considerations

### Compliance with ethical guidelines

This study was approved by the Ethics Committee of [Iran University of Medical Sciences](#), Tehran, Iran (Code: IR.IUMS.REC.1402.039).

### Funding

The present article was extracted from a master thesis of Fahimeh Poormohammadi, approved by [Iran University of Medical Sciences](#), Tehran, Iran (1402-4-6-27466).

### Authors' contributions

Conceptualization: Arezoo Saffarian and Fahimeh Poormohammadi; Formal analysis: Reyhane Mohamadi, Arezoo Saffarian and Fahimeh Poormohammadi; Funding acquisition: Reyhane Mohamadi, Arezoo Saffarian and Mona Ebrahimipour; Investigation: Arezoo Saffarian, Mohammad Hassan Torabi and Fahimeh Poormohammadi; Methodology: Reyhane Mohamadi, Arezoo Saffarian and Mohammad Hassan Torabi; Project administration and supervision: Arezoo Saffarian; Resources and validation: Reyhane Mohamadi, Mohammad Hassan Torabi, Mona Ebrahimipour and Arezoo Saffarian; Visualization: Reyhane Mohamadi, Mohammad Hassan Torabi, Arezoo Saffarian and Fahimeh Poormohammadi; Writing the original draft, review & editing: Arezoo Saffarian and Fahimeh Poormohammadi; Data collection: All authors.

### Conflict of interest

The authors declared no conflicts of interest.

### Acknowledgments

The authors thank Sandra Weintraub, developer of the NAT, for granting permission to adapt the test into Persian and for her valuable support and also thank all the experts and participants who contributed to the implementation of this study.

## References

- [1] Moreira L, Schlottfeldt CG, Paula JJ, Daniel MT, Paiva A, Cazita V, et al. Normative study of the token test (short version): preliminary data for a sample of Brazilian seniors. *Arch Clin Psychiatry*. 2011; 38(3):97-101. [DOI:10.1590/S0101-60832011000300003]
- [2] Paul R. Language disorders from infancy through adolescence. St. Louis: Mosby; 2001. [Link]
- [3] Chomsky N. Minimalist inquiries: The framework (MITOPL 15). In: Bender EM, editor. Step by step: Essays on minimalist syntax in honor of Howard Lasnik. Cambridge: MIT Press; 2000. [Link]
- [4] Chomsky N. Aspects of the Theory of Syntax. Cambridge: MIT Press; 2014. [Link]
- [5] Chomsky N, Miller GA. Finite state languages. *Inf Control*. 1958; 1(2):91-112. [DOI:10.1016/S0019-9958(58)90082-2]
- [6] Zarifian T. [Descriptive dictionary of speech and language pathology (English-Persian). Ameri H, editor. Tehran: Farhang Moaser Publications; 2013 (Persian)]. *Lang Linguistics*. 10(19):123-8. [Link]
- [7] Svete A, Cotterell R. Recurrent neural language models as probabilistic finite-state automata. *ArXiv*. 2023; [Unpublished]. [Link]
- [8] Thompson CK, Mack JE. Grammatical impairments in PPA. *Aphasiology*. 2014; 28(8-9):1018-37. [DOI:10.1080/02687038.2014.912744] [PMID]
- [9] Nilipour R. Agrammatic language: two cases from Persian. *Aphasiology*. 2000; 14(12):1205-42. [DOI:10.1080/02687030050205723]
- [10] Daroff RB, Aminoff MJ. Encyclopedia of the neurological sciences. Cambridge: Academic Press; 2014. [Link]
- [11] Bastiaanse R, Hurkmans J, Links P. The training of verb production in Broca's aphasia: A multiple-baseline across-behaviours study. *Aphasiology*. 2006; 20(2-4):298-311. [DOI:10.1080/02687030500474922]
- [12] Cho-Reyes S, Thompson CK. Verb and sentence production and comprehension in aphasia: Northwestern Assessment of verbs and sentences (NAVS). *Aphasiology*. 2012; 26(10):1250-77. [DOI:10.1080/02687038.2012.693584] [PMID]
- [13] Farooqi-Shah Y, Thompson CK. Effect of lexical cues on the production of active and passive sentences in Broca's and Wernicke's aphasia. *Brain Lang*. 2003; 85(3):409-26. [DOI:10.1016/S0093-934X(02)00586-2] [PMID]
- [14] Gilmore N, Dwyer M, Kiran S. Benchmarks of significant change after aphasia rehabilitation. *Arch Phys Med Rehabil*. 2019; 100(6):1131-9.e87. [DOI:10.1016/j.apmr.2018.08.177] [PMID]
- [15] Grossman M. The non-fluent/agrammatic variant of primary progressive aphasia. *Lancet Neurol*. 2012; 11(6):545-55. [DOI:10.1016/S1474-4422(12)70099-6] [PMID]
- [16] Weintraub S, Mesulam MM, Wieneke C, Rademaker A, Rogalski EJ, Thompson CK. The northwestern anagram test: measuring sentence production in primary progressive aphasia. *Am J Alzheimers Dis Other Dement*. 2009; 24(5):408-16. [DOI:10.1177/1533317509343104] [PMID]
- [17] Links P, Hurkmans J, Bastiaanse R. Training verb and sentence production in agrammatic Broca's aphasia. *Aphasiology*. 2010; 24(11):1303-25. [DOI:10.1080/02687030903437666]
- [18] Downing SM. Twelve steps for effective test development. In: Downing SM, Haladyna TM, editors. Handbook of test development. Mahwah: Lawrence Erlbaum Associates; 2006. [Link]
- [19] Bastiaanse R, van Zonneveld R. Sentence production with verbs of alternating transitivity in agrammatic Broca's Aphasia. *J Neurolinguistics*. 2005; 18(1):57-66. [DOI:10.1016/j.jneuroling.2004.11.006]



- [20] Caplan D, Waters G, Dede G, Michaud J, Reddy A. A study of syntactic processing in aphasia I: Behavioral (psycholinguistic) aspects. *Brain Lang.* 2007; 101(2):103-50. [DOI:10.1016/j.bandl.2006.06.225] [PMID]
- [21] Colman KS, Koerts J, van Beilen M, Leenders KL, Poët WJ, Bastiaanse R. The impact of executive functions on verb production in patients with Parkinson's disease. *Cortex.* 2009; 45(8):930-42. [DOI:10.1016/j.cortex.2008.12.010] [PMID]
- [22] Dick J, Fredrick J, Man G, Huber JE, Lee J. Sentence production in Parkinson's disease. *Clin Linguist Phon.* 2018; 32(9):804-22. [DOI:10.1080/02699206.2018.1444791] [PMID]
- [23] Garraffa M, Fyndanis V. Linguistic theory and aphasia: An overview. *Aphasiology.* 2020; 34(8):905-26. [DOI:10.1080/02687038.2020.1770196]
- [24] Garraffa M, Grillo N. Canonicity effects as grammatical phenomena. *J Neurolinguistics.* 2008; 21(2):177-97. [DOI:10.1016/j.jneuroling.2007.09.001]
- [25] Grossman M, Kalmanson J, Bernhardt N, Morris J, Stern MB, Hurtig HI. Cognitive resource limitations during sentence comprehension in Parkinson's disease. *Brain Lang.* 2000; 73(1):1-16. [DOI:10.1006/brln.2000.2290] [PMID]
- [26] Thompson CK. Unaccusative verb production in agrammatic aphasia: The argument structure complexity hypothesis. *J Neurolinguistics.* 2003; 16(2-3):151-67. [DOI:10.1016/S0911-6044(02)00014-3] [PMID]
- [27] Thompson CK, Choy JJ. Pronominal resolution and gap filling in agrammatic aphasia: Evidence from eye movements. *J Psycholinguist Res.* 2009; 38(3):255-83. [DOI:10.1007/s10936-009-9105-7] [PMID]
- [28] Mehri A, Jalaie S. A systematic review on methods of evaluate sentence production deficits in agrammatic aphasia patients: Validity and reliability issues. *J Res Med Sci.* 2014;19(9):885-98. [PMID]
- [29] Lawshe CH. A quantitative approach to content validity. *Pers Psychol.* 1975; 28(4):563-75. [DOI:10.1111/j.1744-6570.1975.tb01393.x]
- [30] Waltz CF, Bausell BR. *Nursing research: Design statistics and computer analysis.* Philadelphia: FA Davis Company; 1981. [Link]
- [31] United Nations. *Handbook of vital statistics systems and methods: Legal, organizational, and technical aspects.* New York: United Nations; 1984. [Link]
- [32] Canu E, Agosta F, Imperiale F, Ferraro PM, Fontana A, Magnani G, et al. Northwestern anagram test-Italian (Nat-I) for primary progressive aphasia. *Cortex.* 2019; 119:497-510. [DOI:10.1016/j.cortex.2019.08.007] [PMID]
- [33] Ditzes R, Barbieri E, Thompson CK, Weintraub S, Weiller C, Mesulam MM, et al. German language adaptation of the NAVS (NAVS-G) and of the NAT (NAT-G): testing grammar in aphasia. *Brain Sci.* 2021; 11(4):474. [DOI:10.3390/brainsci11040474] [PMID]
- [34] Mehri A, Ghorbani A, Darzi A, Jalaie S, Ashayeri H. Comparing the production of complex sentences in Persian patients with post-stroke aphasia and non-damaged people with normal speaking. *Iran J Neurol.* 2016; 15(1):28-33. [PMID]