Accepted Manuscript

Accepted Manuscript (Uncorrected Proof)

Title: Translation and Cross-Cultural Adaptation of the Aphasia Check List (ACL) for Farsi-Speaking Persons with Aphasia

Running title: Translation of Aphasia Check List (ACL) for persons with Aphasia

Authors: Amin Modarres Zadeh[1], Ahmad Reza Khatoonabadi[1*], Azar Mehri[1], Shohreh Jalaei[2], Elke Kalbe[3]

1.Department of Speech and Language Pathology, School of Rehabilitation Sciences, Tehran University of Medical Sciences, Tehran, Iran.

2.Department of Physical Sciences, School of Rehabilitation Sciences, Tehran University of Medical Sciences, Tehran, Iran.

3.Medizinische Psychologie | Neuropsychologie und Gender Studies, Centrum für Neuropsychologische Diagnostik und Intervention (CeNDI), Prodekanin für Akademische Entwicklung und Gender, Köln, Germany.

*Corresponding author: Ahmadreza Khatoonabadi

Corresponding author email: a.r.Khatoonabadi@gmail.com

To appear in: Basic and Clinical Neuroscience

Received date: 2019/06/17

Revised date: 2020/05/5

Accepted date: 2020/07/6

Please cite this article as:

Modarres Zadeh, A., Khatoonabadi, A. R., Mehri, A., Jalaei, Sh., & Kalbe, E. (In Press). Translation and Cross-Cultural Adaptation of the Aphasia Check List (ACL) for Farsi-Speaking Persons with Aphasia. *Basic and Clinical Neuroscience*. Just Accepted publication Sep. 1, 2020. Doi: http://dx.doi.org/10.32598/bcn.12.4.1906.1

DOI: http://dx.doi.org/10.32598/bcn.12.4.1906.1

Highlights:

- The ACL-P possesses medium level of item difficulty for participants with aphasia, Based on floor and ceiling analysis. That is, items are not too difficult that no patient can accomplish the tasks nor too easy that all patients can complete them.
- The ACL-P can differentiate between normal controls and individuals with aphasia participated in the study based on discriminant validity results.
- The language part of the ACL-P has an adequate internal consistency; however, the cognition part of the test has weak internal consistency because it consists of different cognitive domains.
- The test has adequate reliability in both language and cognition parts when administered different times on the same subjects.

Plain language summary:

In this study, an aphasia test called Aphasia check list has been translated from German into Farsi and then administered on a sample of normal and persons with aphasia to determine its psychometric properties. This test covers both language and cognition skills using different types of verbal and non-verbal stimuli at different levels of difficulty in a fast and convenient way. The comprehensiveness of this test and its convenient scoring system led the researchers to conduct this study. The main result of the study was that the translated version of the test, like its original version, has adequate discriminant validity, test-retest reliability and internal consistency.

Abstract

The Aphasia Check List (ACL) test is a comprehensive, time-saving tool for language evaluation in aphasia, including a cognitive assessment part. This cross-sectional study aimed to translate this test into Farsi and analyze the psychometric features of the translated version. The original version of the ACL was translated and adapted from German; its psychometric features were then determined. Twenty participants with aphasia (PWA) and 50 age- and education-matched, cognitively healthy controls participated. Possible floor and ceiling effects, discriminant validity, test-retest reliability, and internal consistency were analyzed in addition to the evaluation of internal correlations between the test parts (Language and Cognition).

Regarding the performance of PWAs in the language section and the cognitive subtests assessing attention, memory, and reasoning, there were no floor and ceiling effects. Adequate discriminant validities for the language section of the test (i.e., total score: [Mann-Whitney U= 6.000, p<0.001]; diagnostic subtests scores: [U=3.000, p<0.001]; and each subtest individually) and for the attention subtest of the cognition section [U=16.500, p<0.001] were observed. There was no difference between the control group and the patient group in the subtests of memory [U=497.500, p=0.973] and reasoning [U=3.000, p=308]. The test-retest reliability was acceptable in all subtests (ICCagreement =0.573-0.984). The ACL-P test showed appropriate internal consistency (Cronbach's alpha=0.761 for test and retest scores). There were also significant correlations between language and cognition in the control and patient groups. The ACL-P test showed sufficient reliability and validity for the evaluation of Farsi-speaking PWAs and used in studies on this population.

Keywords: Aphasia, Aphasia Check List, assessment, validation, reliability

Introduction:

Aphasia is an acquired disorder of verbal communication. It always occurs due to some damage to certain parts of the brain responsible for speech and language processing. It is manifested mainly by impairments in receptive and expressive modalities of spoken and written language to different extents (Hallowell & Chapey, 2008). It has been shown that deficits of memory (Vallila-Rohter & Kiran, 2013; Mayer & Murray, 2012), attention (Murray, 2012; Villard & Kiran), reasoning (Murray, 2012) and executive functions (Murray, 2017) can be comorbid with aphasia. Cerebrovascular accident is the most prevalent cause of aphasia. Ischemic infarction causes 80% of aphasic cases (Berthier, 2005). According to the studies conducted in Europe and the US, one-third of all patients suffering from strokes are prone to aphasia in the acute stages (Damico, Müller & Ball, 2010; Engelter et al., 2006; Laska, Hellborn, Murray, Kahan, & Von Arbin, 2001). In Iran, two studies have been released on the prevalence of aphasia. A cross-sectional study indicated that 5 out of 22 stroke patients (equal to 22.7% of the statistical population) had aphasia (Soltani, Khatoonabadi, Jenabi, & Piran, 2013). A retrospective study showed that 727 out of 1817 such patients (equal to 39.9% of the statistical population) had a particular type of aphasia (Zamani & Madjdinasab, 2013). Therefore, given the relatively high prevalence of aphasia among patients with a first-ever stroke and its adverse effects on the affected individuals' quality of life in the form of depression and social isolation (Kauhanen et al., 2000; Davidson, Howe, Worrall, Hickson, & Togher, 2008), it seems necessary to implement early intervention programs (Godecke et al., 2014), the first step of which is evaluation (Murray & Coppens, 2012). Evaluation can be conducted for such various purposes as determining the presence of aphasia and its classification, the severity of the disorder, the nature of the observed language disorders, setting therapeutic objectives and adopting adequate approaches, treatment prognosis and changes in the treatment process (Bruce & Edmundson, 2010). One aphasia test alone cannot meet all of the above goals. Therefore, clinicians have been advised to think about the exact purpose of their examination before choosing assessment tools (Spreen & Risser, 2003). When a therapist intends to collect samples from a wide range of verbal behaviors at different levels of assignment difficulty in the expressive and receptive modalities, comprehensive aphasia tests should be used (Patterson & Chapey, 2008).

The Persian Aphasia Test is the only comprehensive test which has been released in Iran. It evaluates the six primary language skills, i.e., oral expression, oral comprehension, repetition, available vocabulary, reading, and writing. Although according to the developers, this test has been used by speech and language pathologists for over 20 years within the country, its use may cause problems. This test consists of 25 subtests and 217 items. Thus, it might be time-consuming and tiring when conducted on an aphasic population because aphasia occurs mostly after the age of 65 years (Benjamin et al., 2017). Thus, patients may not be cooperative. As discussed earlier, aphasia is a multifaceted disorder which can affect other language-dependent aspects of cognitive functioning such as non-word processing (Luzzatti, Toraldo, Zonca, Cattani, & Saletta, 2006), numerical processing (Delazer & Bartha, 2001) and neuropsychological abilities including memory, attention, and reasoning (Murray, 2012; El Hachioui et al., 2014). Therefore, a comprehensive aphasia test should also cover these areas. Accordingly, the authors of this paper decided to translate the ACL, which is a German aphasia test, and evaluate its psychometric features. The ACL has certain advantages, the most important of which is the fact that it takes 30-40 minutes to administer the entire test. This feature enables the therapist to finish the evaluation process in one session; therefore, participants will experience less fatigue. Another feature of this test is the fact that it uses a variety of verbal stimuli (sentences, words [nouns and verbs], non-words, and numbers) for language evaluation. These stimuli become complicated systematically during every task, something which can significantly help in pinpointing the examinee's level of performance. Moreover, in order to minimize the interference of a faulty language system during the cognitive assessment, nonverbal stimuli have been used in the cognition part of the test (Kalbe, Reinhold, Brand, Markowitsch, & Kessler, 2005). This test has been provided with cut-off points for each subtest in the language and cognition sections to evaluate the presence of disorder in the respective area of functioning and also a total cut-off point for determining the presence of aphasia in the language section. It is easy to conduct, and has a convenient scoring method. Therefore, it appeared that translating and publishing this test could greatly help Iranian therapists increase the quality of intervention offered to patients with aphasia.

Materials and Methods

Translation and Cultural Adaptation of the ACL Test

Initially, permission to translate the test to Persian was obtained from the test developer. Then the translation process began based on the International Quality of Life Assessment (IQOLA) protocol for standard translation (Ware & Gandek, 1998).

Phase 1: The German-to-Persian translation process was started by two Farsi-speaking translators (Translator 1 and Translator 2). These two translators mastered both languages and linguistics terms and concepts. However, they were not familiar with the ACL test, and each one translated the test independently.

Phase 2: After finishing the forward translation process, each translator was asked to score all sections translated by the other translator (words, expressions, and sentences) concerning difficulty using a 100-point Likert scale, in which 0 would mean completely understandable, and 100 would mean unintelligible. If an item were scored above 30, it would be considered hard and then returned to its translator. In this section, none of the items was scored above 30.

Phase 3: In this phase, two other bilingual translators were asked to score the quality of each version of the translations. By quality, we mean the simplicity and clarity of translation. These translators had no connection to Translators 1 and 2 and lacked clinical experience. They gave scores independently from each other. Again, a 100-point visual scale was employed. In this scale, 0 meant the lowest quality of translation, and 100 meant the highest. With a score lower than 90, the translation was considered low-quality with the need for revision by its translator. In this phase, none of the items received a score below 90.

Phase 4: In this phase, an expert panel was held for Translators 1 and 2, and the researchers to evaluate and compare the difficulty and quality of the translations. At last, one translation was agreed upon.

Phase 5: The final translated version of the ACL in Farsi was back-translated by another translator (Translator 5) into German. This translator had not been present in the previous phases and knew nothing about the purpose and intended use of the ACL test.

Phase 6: After finishing the back-translation, the original version of the test was compared with the back-translated version by an expert panel for any necessary corrections or revisions. Finally, a satisfactory Persian version of the test was reached.

Phase 7: Ten speech and language pathologists working in the field of aphasia management were asked to score the pre-final Farsi version in terms of its following the required criteria using a 4-point Likert scale in terms of 1) comprehensibility of the commands and stimuli; 2) sociocultural suitability of the test stimuli; and 4) clarity and fluency of the used words, expressions, and sentences. As to this scale, 0 meant the absence of the criteria as mentioned earlier, and 4 meant full compliance with them. If 80% of the respondents gave less than 3 to any item, it had to be sent to the corresponding translator for revision. All of the verbal items received acceptable scores. The expert panel confirmed the adequacy of the results of this field testing, and the Persian version of the ACL (called ACL-P from now on) was prepared to be evaluated among a population of patients with aphasia.

Participants

In this study, the PWAs were recruited based on the following criteria from the caseload of clinics affiliated with Tehran University of Medical Sciences through the convenience sampling method. The inclusion criteria were

1. a first left-hemisphere stroke caused aphasia, and there had been no other previous strokes;

2. at least 3months had passed since the stroke before the initiation of the study;
3. maximum age of 65 years old;
4. the native language had to be Farsi, and the participant had to be literate before the occurrence of the lesion;
5. participants should have had no comorbid neurodegenerative or psychiatric diseases (e.g., dementia of the Alzheimer's type or clinical depression) based on their medical history, their prescribed medications and/or caregivers' reports of their problems;
6. participants had not received speech/language interventions at least one week before the testing sessions (they had already completed a program or were awaiting to be enrolled in a new program, or for whatever reason did not intend to take part in the speech therapy facilities available to them.

However, if they were receiving any program, they were excluded from this study. In addition to the usual evaluations of speech therapy, an experienced neurologist was available to confirm the diagnosis of aphasia. Neurologically healthy controls were selected from the clients referred to the same clinics. They were caretakers of participant patients or had other complaints than speech and language disorders (such as optometric and or lower-back problems). The inclusion criteria for healthy participants were

1. no history of verbal or cognitive disorders (diagnosed through the Farsi version of MMSE[cut-off point = 23])(Seyedian et al., 2008);
2. lack of any mental or psychological disorders disrupting test implementation ?;
   having Farsi as their native language (or using Farsi as their preferred language in case of being bi/multilingual);
3. being adequately literate to meet daily needs.

The research goals were explained to every participant and their caretakers. Then they were asked to sign an informed written consent, which was securely stored by one of the researchers.

Testing

The ACL-P test was given individually to every participant. The test was conducted in an acoustic well-lit room at the Rehabilitation School of Tehran University of Medical Sciences, away from any visual distractions. If a participant could not attend this location for any reason, the testing sessions were pursued at his/her home in full compliance with standard conditions. The language learning section was implemented by asking oral questions or showing pictures/printed words and recording responses. The cognitive section was paper-based. Since participants had to cross out the target items in the attention subtest and since PWAs used their non-dominant left hands to complete the task, before implementing the test in this group, they practiced crossing out geometrical items such as squares and triangles with pencils so that they would concentrate on the test itself instead of figuring out "how" to complete the task. The oral responses were immediately transcribed. However, these responses were also recorded by a device (Sony SO-ICD-PX240 4GB Voice Recorder). The retest session was held ten days after the first testing session for each participant by the same examiner in the same conditions.

The ACL-P test consists of two general sections: language and cognition. The language section consists of the seven following subtests, utilized to evaluate the primary modalities of verbal comprehension and expression, reading, writing, and repetition: 1) serial speech (2 tasks); 2) following commands (2 tasks); 3) the color-shape test; 4) word generation (2 tasks); 5) specific verbal abilities: a) confrontation naming; b) reading aloud; c) reading comprehension; d) listening comprehension; e) writing to dictation; f) repetition (45 tasks); 6) evaluation of general verbal communication ability; and 7) numerical processing (3 tasks). The cognitive section consists of three subtests i.e., memory, attention, and reasoning. Regarding the memory subtest, the participants looked at six geometrical shapes for 10 seconds. Then they were asked to recognize them from among an assortment of shapes immediately (short-term memory) and after 10 minutes (mid-term memory). The attention section has been designed to evaluate the speed and quality of selective attention through a task of cancelling out two geometrical shapes in a block of similar shapes. Regarding reasoning, there were 11 sequences of 9 objects, 8 of which were arranged according to a rule, and the examinee must determine the one that disrupted the arrangement (Kalbe et al., 2005).

Statistical Tests

The resultant data were statistically analyzed in SPSS (ver.23). The Shapiro-Wilk test was conducted to check the normality of data. The Mann-Whitney *U* test was carried out to evaluate the discriminant validity. The ICC coefficients (an absolute agreement, two-way mixed) were employed to determine the test-retest reliability. The lowest acceptable coefficient was considered 0.5, and values ranging between 0.5and 0.75, between 0.75 and 0.9, and over 0.9 were respectively considered mediocre, good, and excellent. Cronbach's alpha was utilized to analyze internal consistency. The coefficients ranging between 0.7 and 0.9 were regarded as excellent. The Spearman rank correlation coefficient test was carried out to determine the correlation between the language and cognition parts of the test. The minimum and maximum scores were taken into account to determine the floor and ceiling effects, respectively. The floor and ceiling effects were significant if they were observed to be over 15%.

Results

In this study, the statistical population included 20 aphasic patients (13 males and seven females with age mean of 56.90±7.51 years and the age range of 45-65 years) and 50 neurologically healthy individuals (29 males and 21 females with age mean of 59.80±5.20 years and the age range of 52-70 years). The mean years of education (standard deviation) in the patient and control groups were as follows: 11.85 (3.94) (with a range of 5-18 years) and11.34 (3.79) (with a range of 4-18 years). According to the Shapiro-Wilk test, the variable of age followed no normal distribution both in the patient group (W=0.864 and p=0.009) and in the control group (W=0.939 and p=0.013). The educational attainment had a normal distribution in the patient group (W=0.905 and p=0.052); however, it showed a significant difference from the normal distribution in the control group (W=0.906 and p=0.001). There were no significant differences between the two groups in terms of age (U=408 and p=0.230) and education (U=455 and p=0.543). The mean of the elapsed time from the onset of the disorder to the initiation of the study was 29.7±17.89 months, with a range of 8-60 months in the patient group. Based on hospital discharge records, MRI scans, and language evaluations via the P-WAB-1 test (Nilipour, Pourshahbaz, & Ghoryshi, 2014), the type of aphasia in all the subjects was categorized as non-fluent. It was due to lesions to the perisylvian areas of the left hemisphere. In the control group, the MMSE test was conducted to evaluate mental health. Their mean scores were 27.76±2.20, with a range of 24-30.

Floor and Ceiling Effects in PWA group

The total scores of all language assignments of the ACL-P were 213. For the language section of the test, the means and ranges of the total scores in the test (mean= 112.7±7.58 and range=63-171) and retest (mean=120.35±7.20 and range=69-174) sessions showed no floor and ceiling effects. The distribution of scores was normal in the patient group in both the test (W=0.935 and p=0.191) and the retest (W=0.952 and p=395). According to the test's manual, eight subtests (phonological fluency, semantic fluency, listening comprehension, reading comprehension, naming, reading aloud, dictation of words and sentences, repetition of words and sentences, and color-shape test) were considered the essential subtests (these subtests are henceforth called "diagnostic subtests"). The maximum total score for these subtests was 148. The scores of the patient group showed no such effects in the test (mean=78.10±24.01; Range=44-115) and the retest (mean=82.5±22.75; Range=42-120) for the diagnostic subtests. The distribution of scores was normal in the test (W=907 and p=0.56) and the retest (W=0.952 and p=0.395). Given the fact that the constructs are dependent on each other in the cognition part of the ACL-P, the scores of this section cannot be added together.

Therefore, each subtest's score was reported separately for this part of the test. No floor and ceiling effects were observed for immediate memory (mean=4.45; Range= -1- 6) in the test and in the retest sessions (mean=4.95; Range=3- 6) and also the case for delayed memory in the test (mean=3.95; Range=2-6) and the retest (mean=4.3; Range=2- 6) sessions (the minimum score was -6 on this subtest). In the attention section, three scores were determined: the total number of processed items (144 at most), the differentiation of the processed items and errors (no maximum), and error percentage (66.6% at most). There were no floor and ceiling effects either for the total number of processed items in the test (mean=61; Range=25-117) or the retest (mean=56.5; range=24-110) nor were there any for the error percentage in the test (mean=5.14; range=0-22.2) and the retest (mean=5.07; range=0-15.6) sessions. Regarding the subtest of reasoning (the maximum score was 11), the same was true in the test (mean=5.8, Range=4-10) and the retest (mean=6; Range=4-10) sessions. The next section deals with the floor and ceiling effects on each subtest.

Discriminant Validity

There were significant differences between the total score of the language section (*Mann–Whitney U*=6.000, *n1*=50, *n*2=20, *p*<0.001) and the scores of diagnostic subtests (*Mann–Whitney U*=3.000, *n1*=50, *n2*=20, *p*<0.001) for the control and PWA groups. Table 1 shows the control and patient scores on each subtest of the ACL-P in the first execution of the test and the significance of the difference between the mean or median scores of these two groups. Since the results of most of these subtests did not follow a normal distribution in both patient and control groups, the medians and score ranges have been reported. If the scores follow a normal distribution, the mean and standard deviation are reported and highlighted. Accordingly, the two groups were significantly different from each other on all subtests of the language section. In the second section of the test, a significant difference was only observed in the subtest of attention.

[Table 1 near here]

Internal Consistency

Cronbach's alpha was employed to test the internal consistency. The coefficient obtained was 0.761 for the entire test (both language and cognition sections were considered) in both test and retest phases. The values of Cronbach's alpha ranged from 0.777 to 0.804 in the test phase and from 0.727 to 0.895 in the retest phase in case of item deletion. Cronbach's alpha was 0.884 (ranged, if item deleted, from 0.867 to 0.884) for the language part and was 0.071 (ranged, if item deleted, from 0.060 to 0.563) for the Cognition part in the test phase. These values were 0.899 (range = 0.878 – 904) for the Language part and 0.120 (range=0.052-0.123) for the Cognition part in the retest. Regarding the diagnostic subtests of the language section, Cronbach's alpha was 0.852 in the test phase. However, it was 0.878 in the retest phase. If items were deleted, the coefficient ranged from 0.809 to 0.859 in the test phase and from 0.838 to 0.898 in the retest phase.

Test-Retest Reliability

The value of $ICC_{agreement}$ was high for the total scores of the language section ($ICC_{agreement}$=0.982, 95% CI =0.819-0.995, p<0.001). It was also high in the diagnostic subtests ($ICC_{agreement}$=0.981, 95%CI=0.863-0.995, p<0.001). Table 2 shows the values of $ICC_{agreement}$ determined for every ACL subtest in the language and cognition sections. Accordingly, ICC was acceptable with a 95% confidence interval in all ACL-P subtests. More precisely, it was between 0.573 and 0.984 at a significance level of $p<0.05$.

[Table 2 near here]

Analysis of the Cognition Part of ACL-P

As discussed earlier, the ACL-P test consists of two sections: language and cognition.

According to the results of the cognition part, all of the surveyed patients had dysfunctions in at least one of the cognitive areas (short memory, midterm memory, attention [the total processed items minus errors], and reasoning). A Performance lower than the cut-off points were observed in 4 participants on one subtest, in 9 participants on two subtests, in 3 participants on three subtests, and 4 participants in four subtests. The frequency of the presence of disorder in each domain is as follows: attention (19 participants), reasoning (15 participants), short-term memory (7 participants), and midterm memory (6 participants). Moreover, Table 3 indicates the amount of correlation between the essential subtests of the language section and the subtests of the cognition section in the healthy and aphasic group. Accordingly, the number of significant correlations in the control group was more extensive than that of the PWA group. In the control group, the only subtest without correlations with the cognition subtests was Repetition. In the patient group, the color-shape and repetition subtests had no correlations with any of the cognition subtests. The subtest of attention had the most significant number of correlations with the language subtests.

[Table 3 near here]

Discussion

This study aimed at the cultural adaptation and determination of the psychometric features of the ACL-P test so that it could be used for evaluation purposes in Farsi-speaking clinical environments where persons with aphasia attend. The coverage of a wide variety of language skills and the relatively short administration time make this test suitable for a time-saving but comprehensive assessment. The subtests include different numbers of tasks proportional to their relevance to the main characteristics of aphasia. For instance, the subtest for comprehension of words/sentences (which is essential for diagnosis and classification of aphasia syndromes) consisted of 6 tasks. In contrast, the subtest for reading numbers included three assignments. As a result, the therapist can screen lesser-important language skills and conduct a more comprehensive evaluation of these skills, only if it becomes necessary. The presence of nonverbal assignments for the evaluation of short-term memory, midterm memory, attention, and reasoning are among the other hallmarks of this test. According to recent studies, cognitive functions can affect the evaluation process and the response to treatment. Therefore, it is essential to know about a patient's cognitive status (Villard & Kiran, 2015).

Translation and Cultural Adaptation

The translation process was performed smoothly because a standard protocol was employed for translation. Also, the stimuli and guidelines in the source language (German) were void of any linguistic complexity (e.g., usage of words with multiple meanings or long sentences) so that the face and content validities of the test were appropriate. However, there were a few changes made in the verbal stimuli of the test and pictures. As for the reading, repeating, and writing of non-words, the test items matching the Persian phonotactic constraints were used. The pictures used for naming and reading comprehension were changed to match the Islamic culture governing Iran (Figure 1). Other than that, all of the other items matched the Persian language and Iranian culture.

Floor and Ceiling Effects

As a group, participants with aphasia did not obtain the minimum and maximum scores in any subtest of ACL-P during the test and retest sessions. Upon the analysis of every subtest of the language section, it was found that several participants in the aphasia group did obtain maximum scores in the subtests of automatic speech, following commands, and numerical processing in the test and retest phases. One reason could be the type of aphasia. Auditory comprehension is usually damaged less than other language functions in nonfluent aphasia (1). Since this was the type presented by all aphasic patients in the current study, the maximum score was observed in the direction-following subtest. Satisfactory performance of some aphasic participants in the automatic speech subtest (including reciting the days of the week and counting to 15) was consistent with the fact that people with aphasia show better performances in non-propositional (automatic) language than in propositional language assignments (Lum & Ellis, 1999). Regarding the maximum score of the numerical processing subtests, the results were consistent with the evidence that there can be distinctions between verbal functions and numerical processing functions of aphasic patients (Rossor, Warrington, & Cipolotti, 1995). In the original paper introducing this test, the score range of reading, writing, and repeating numbers were between 0 and 9, indicating that some patients with the disorder may reach the maximum score in these subtests (Kalbe et al., 2005).

Internal Consistency and Test-Retest Reliability

According to the Cronbach alphas of all subtests and diagnostic subtests, the ACL-P had an acceptable internal consistency. In the original paper of the test, the values of Cronbach's alpha were reported to range between 0.4 and 0.88 (Kalbe et al., 2005). Lower internal consistency of the cognition part of the ACL-P, compared to that of the language part, was expected because it consisted of distinct cognitive domains of memory, attention, and reasoning (Sachdev et al., 2014). Moreover, this test showed acceptable test-retest reliability. In the original normative study of the ACL, the test-retest reliability ranged between 0.5 and 0.91 at the significance level of $p<0.05$, a finding which was consistent with the findings of this study.

Correlation between Language and General Cognition

This test showed that the studied participants with aphasia performed more poorly in selective attention than the subjects of the control group. The attention subtest was significantly correlated with five language subtests (phonological fluency, reading, auditory comprehension, and writing to dictation) in the aphasia group. There were no significant differences between the two groups (patients and controls) - regarding the subtests of memory and reasoning. In the PWA group, each of these two subtests was significantly correlated with only one subtest of the language section. However, cognitive functions were correlated mainly with the color-shape test, phonological and semantic fluency, naming, reading aloud, reading comprehension, and auditory comprehension in the control group. These findings were in a completely reverse correlation with those of the original study on the test, in which there was a low correlation between the subtests of cognition and language in the control group, but a high correlation between these two sections in the aphasia group. One reason could be the number of patients, which was too small to show a higher correlation in the group. Therefore, more studies should be conducted on larger samples, including more diverse aphasia types.

Conclusion

Finally, it is concluded that the ACL-P test is reliable and valid for use in early evaluations, detecting progress throughout treatment courses, and prioritizing therapeutic goals in clinical environments. It can also be utilized in experimental studies on aphasia in the Farsi language.

Acknowledgement

Researchers hereby thank every aphasic patient participating in the study. The study was supported by Tehran University of Medical Sciences.

Conflict of Interests

The authors have no conflict of interests in this paper.

References:

Benjamin, E. J., Blaha, M. J., Chiuve, S. E., Cushman, M., Das, S. R., Deo, R., Floyd, J., Fornage, M., Isasi, C.R., & Jiménez, M. C. (2017). Heart disease and stroke statistics-2017 update: a report from the American Heart Association. *Circulation*, *135*(10), e146-e603.

Berthier, M. L. (2005). Poststroke aphasia. *Drugs & aging*, *22*(2), 163-182.

Bruce, C., & Edmundson, A. (2010). Letting the CAT out of the bag: A review of the comprehensive aphasia test. Commentary on Howard, Swinburn, and Porter,"putting the CAT out: What the comprehensive aphasia test has to offer". *Aphasiology*, *24*(1), 79-93.

Damico, J. S., Müller, N., & Ball, M. J. (Eds.). (2010). *The handbook of language and speech disorders*. West Sussex: Wiley-Blackwell.

Davidson, B., Howe, T., Worrall, L., Hickson, L., & Togher, L. (2008). Social participation for older people with aphasia: The impact of communication disability on friendships. *Topics in stroke rehabilitation*, *15*(4), 325-340.

Delazer, M., & Bartha, L. (2001). Transcoding and calculation in aphasia. *Aphasiology*, *15*(7), 649-679.

El Hachioui, H., Visch-Brink, E. G., Lingsma, H. F., van de Sandt-Koenderman, M. W., Dippel, D. W., Koudstaal, P. J., & Middelkoop, H. A. (2014). Nonlinguistic cognitive impairment in poststroke aphasia: a prospective study. *Neurorehabilitation and Neural Repair*, *28*(3), 273-281.

Engelter, S. T., Gostynski, M., Papa, S., Frei, M., Born, C., Ajdacic-Gross, V., Gutzwiller, F., & Lyrer, P. A. (2006). Epidemiology of aphasia attributable to first ischemic stroke: incidence, severity, fluency, etiology, and thrombolysis. *Stroke*, *37*(6), 1379-1384.

Godecke, E., Ciccone, N. A., Granger, A. S., Rai, T., West, D., Cream, A., Cartwright, J., & Hankey, G. J. (2014). A comparison of aphasia therapy outcomes before and after a Very Early Rehabilitation programme following stroke. *International journal of language & communication disorders*, *49*(2), 149-161.

Hallowell, B., & Chapey, R. (2008). Introduction to language intervention strategies in adult aphasia. In *Language intervention strategies in aphasia and related neurogenic communication disorders*(pp. 3-19). Philadelphia, PA: Lippincott, Williams, & Wilkins.

Kalbe, E., Reinhold, N., Brand, M., Markowitsch, H. J., & Kessler, J. (2005). A new test battery to assess aphasic disturbances and associated cognitive dysfunctions—German normative data on the aphasia check list. *Journal of clinical and experimental neuropsychology*, *27*(7), 779-794.

Kauhanen, M. L., Korpelainen, J. T., Hiltunen, P., Määttä, R., Mononen, H., Brusin, E., Sotaniemi K.A., & Myllylä, V. V. (2000). Aphasia, depression, and non-verbal cognitive impairment in ischaemic stroke. *Cerebrovascular Diseases*, *10*(6), 455-461.

Laska, A. C., Hellblom, A., Murray, V., Kahan, T., & Von Arbin, M. (2001). Aphasia in acute stroke and relation to outcome. *Journal of internal medicine*, *249*(5), 413-422.

Lum C., & Ellis, AW. (1999). Why do some aphasics show an advantage on some tests of nonpropositional (automatic) speech?. *Brain and Language*, 70(1), 95-118.

Luzzatti, C., Toraldo, A., Zonca, G., Cattani, B., & Saletta, P. (2006). Types of dyslexia in aphasia: A multiple single-case study in a shallow orthography language. *Brain and language*, *99*(1-2), 28-29.

Mayer, J. F., & Murray, L. L. (2012). Measuring working memory deficits in aphasia. *Journal of Communication Disorders*, *45*(5), 325-339.

Murray, L. L. (2012). Attention and other cognitive deficits in aphasia: Presence and relation to language and communication measures. *American Journal of Speech-Language Pathology*.

Murray, L., & Coppens, P. (2012). Formal and informal assessment of aphasia. , 66-91. In *Aphasia and related neurogenic communication disorders*(pp. 67-91). Burlington, MA: Jones & Bartlett Learning.

Murray, L. L. (2017). Design fluency subsequent to onset of aphasia: A distinct pattern of executive function difficulties?. *Aphasiology*, *31*(7), 793-818.

Nilipour, R., Pourshahbaz, A., &Ghoreyshi, ZS. (2014). Reliability and validity of bedside version of Persian WAB (P-WAB-1). *Basic and clinical neuroscience, 5(4)*, 253-258.

Patterson, J.P, & Chapey, R. (2008). Assessment of language disorders in adults. In *Language intervention strategies in aphasia and related neurogenic communication disorders* (5th ed., pp. 64-152). Philadelphia, PA: Lippincott Williams & Wilkins.

Rossor, MN., Warrington, EK., & Cipolotti. L. (1995). The isolation of calculation skills. *Journal of Neurology*, *242(2)*, 78-81.

Sachdev, P. S., Blacker, D., Blazer, D. G., Ganguli, M., Jeste, D. V., Paulsen, J. S., & Petersen, R. C. (2014). Classifying neurocognitive disorders: the DSM-5 approach. *Nature Reviews Neurology*, *10*(11), 634.

Seyedian, M., Falah, M., Nourouzian, M., Nejat, S., Delavar, A., & Ghasemzadeh, H. A. (2008). Validity of the Farsi version of mini-mental state examination. *Journal of medical council of I.R.I ,25(4),408-414*

Soltani, S., Khatoonabadi, A. R., Jenabi, M. S., & Piran, A. (2013). Frequency of aphasia resulting from stroke at hospitals affiliated to Tehran University of Medical Sciences. *Journal of Modern Rehabilitation*, *6*(4), 44-48.

Spreen, O., & Risser, A.H. (2003). Purposes of Assessment. In *Assessment of aphasia,* 22-32. Oxford University Press.

Vallila-Rohter, S., & Kiran, S. (2013). Non-linguistic learning and aphasia: Evidence from a paired associate and feedback-based task. *Neuropsychologia*, *51*(1), 79-90.

Villard, S., & Kiran, S. (2015). Between-session intra-individual variability in sustained, selective, and integrational non-linguistic attention in aphasia. *Neuropsychologia*, *66*, 204-212.

Ware Jr, J. E., & Gandek, B. (1998). Overview of the SF-36 health survey and the international quality of life assessment (IQOLA) project. *Journal of clinical epidemiology*, *51*(11), 903-912.

Zamani, P., & Madjdinasab, N. (2013). Time Variation in the Occurrence of Stroke-Induced Aphasia: A Report from Ahvaz. *Medical Journal of Tabriz University of Medical Sciences and Health Services*, *35*(2), 44-49.

References:

Table 1-Comparison of patients and controls in ACL subtests

Subtest	Max. score	PWA Group Median (Range)	Control Group Median (Range)	Significance
Part A) Language:				
1-Automatic speech	4	4 (2-4)	4 (-)[1]	p<0.001
2-following instructions	4	4 (1-4)	4 (-)[1]	p<0.001
3-color-figure test	20	10.35 (4.65)[2]	19 (10-20)	p<0.001
4- word generation:				
phonemic generation Raw scores	-	2 (0-7)	7.5 (4-20)	P<0.001
phonemic generation- Transformed Scores	10	0 (0-4)	4 (2-10)	p<0.001
Semantic generation- Raw scores	-	5.5 (0-17)	11.5 (5-26)	p<0.001
Semantic word generation- Transformed Scores	10	2 (0-8)	4 (2-8)	p<0.001
5-Picture naming	18	10.85 (4.43)[2]	18 (15-18)	p<0.001
6-Reading aloud:				
Reading words/ sentences	18	10.85 (4.43)[2]	18 (15-18)	p<0.001
Reading non- words	9	0 (0-6)	9 (6-9)	p<0.001
7-Reading comprehension	18	9.85 (4.59)[2]	18 (15-18)	p<0.001
8-Auditory comprehension	18	15 (4-18)	17 (15-18)	p<0.001
9- Writing to dictation:				
words/ sentences	18	6.6 (4.45)[2]	18 (9-18)	p<0.001
non-words	9	0 (0-7)	9 (0-9)	p<0.001
10- Repeating:				
words/ sentences	18	15 (3-18)	18 (17-18)	p<0.001
non-words	9	7.5 (0-9)	9 (7-9)	p<0.001
11-Communication Competence rating	3	3 (-)[1]	1 (0-2)	p<0.001

12-Numeracy skills				
Reading	9	4 (0-9)	9 (8-9)	p <0.001
writing to dictation	9	6 (0-9)	9 (7-9)	p <0.001
Repeating	9	9 (0-9)	9 (-)[1]	p <0.001
Part B) Cognition				
1- Memory:				
Immediate (Corrects minus errors)	6	5 (-1 – 6)	5 (1 – 6)	p=0.973
intermediate-term memory (Corrects minus errors	6	4 (2-6)	4 (2-6)	p=0.866
2- Attention				
total items processed	144	61 (25.06)[2]	128 (95-142)	p<0.001
total items processed minus errors	-	58 (24.02)[2]	126 (87-142)	p<0.001
Error percentage	66.6	3.65 (0-15.30)	3.15 (0-17.70)	p=0.990
3- Reasoning	11	5 (4-10)	6 (2-10)	p=0.308

1. Scores were constant. In other words, everyone obtained the maximum score.
2. The mean and standard deviation were reported.

Table 2: ICC values for inter-rater reliability of ACL-P

Subtest	ICC	95% *CI*	p
Part A) Language:			
1-Automatic speech	0.573	0.005-0.825	=0.02
2-following instructions	0.725	0.328-0.889	=0.003
3-color-figure test	0.830	0.577-0.932	<0.001
4- word generation:			
phonemic generation -Raw scores	0.907	0.763-0.963	<0.001
phonemic generation Transformed Scores	0.836	0.593-0.935	<0.001
Semantic generation-Raw scores	0.888	0.716-0.956	<0.001
semantic word generation-Transformed Scores	0.711	0.269-0.886	=0.005
5-picture naming	0.949	0.861-0.981	<0.001
6-reading aloud:			
Reading words/ sentences	0.984	0.938-0.994	<0.001
Reading non- words	0.931	0.758-0.976	<0.001
7-Reading comprehension	0.920	0.803-0.968	<0.001
8-Auditory comprehension	0.769	0.406-0.909	<0.001
9- Writing to dictation:			
words/ sentences	0.932	0.826-0.973	<0.001
non-words	0.747	0.349-0.901	=0.003
10- Repeating:			
words/ sentences	0.715	0.307-0.886	=0.003
non-words	0.872	0.401-0.960	<0.001
11-Communication Competence rating	0.592	0.012-0.836	=0.026
12-Numeracy skills			
Reading	0.913	0.784-0.965	<0.001
writing to dictation	0.949	0.861-0.981	<0.001
Repeating	0.967	0.916-0.987	<0.001
Part B) Cognition			
1- Memory:			
Immediate (Corrects minus errors)	0.567	-0.046-0.826	=0.034

intermediate-term memory (Corrects minus errors	0.592	0.012-0.836	=0.026
2- Attention			
total items processed	0.916	0.823-0.972	<0.001
total items processed minus errors	0.929	0.823-0.972	<0.001
Error percentage	0.916	0.790-0.967	<0.001
3- Reasoning	0.954	0.885-0.982	<0.001

Table 3: Correlation between the two parts of ACL-P

	Controls				PWAs			
ACL subtests	STM[1]	ITM[2]	Attention	Reasoning	STM	ITM	Attention	Reasoning
color-shape test	.130	0.182	0.395**	0.427**	-0.405	0.177	0.394	0.054
B Generation	0.099	0.297*	0.329*	0.553**	0.161	0.180	0.561*	0.017
Supermarket Generation	0.114	0.047	0.108	0.344**	0.479*	0.239	0.406	0.250
Naming	0.326*	0.315*	0.333*	0.296*	-0.160	0.172	0.373	0.447*
Reading	0.089	0.077	0.310*	0.344*	-0.197	0.198	0.483*	
Reading Comprehension	0.353*	0.224	0.404**	0.431**	-0.246	0.080	0.467*	0.275
Auditory Comprehension	0.491**	0.359*	0.279*	0.465**	-0.030	0.427	0.505*	0.222
Writing to Dictation	0.244	0.296*	0.350*	0.237	-0.195	0.012	0.483*	0.138
Repetition	0.228	0.220	0.223	0.240	-0.030	0.219	0.164	0.311

* $p<0.05$, ** $p<0.01$

[1]Short-term memory; [2]Intermediate-term memory

The original pictures	The modified Pictures
	
Figure 1. A modification to respect cultural constraints	

Authors' contribution

Conceptualization, Ahmad reza Khatoon abadi; Investigation, Amin Modarres Zadeh; Writing – Original Draft, Amin Modarres Zadeh; Writing – Review & Editing, Ahmadreza Khatoonabadi, Azar Mehri, and Elke kalbe

This research was supported by Tehran University of Medical sciences.