

The Aphasia Rapid Test: adaptation and standardisation for Russian

Olga Buivolova^a, Oxana Vinter^b, Roelien Bastiaanse^{a,c}, and Olga Dragoy^{a,d}

^a Center for Language and Brain, National Research University Higher School of Economics, Moscow, Russia ^b Department of Neurology, City Clinical Hospital No. 31 of the Moscow Health Department, Moscow, Russia ^c Center for Language and Cognition Groningen (CLCG), University of Groningen, Groningen, The Netherlands ^d Department of Medical Rehabilitation, Center for Cerebrovascular Pathology and Stroke, Moscow, Russia

ABSTRACT

Background: The Aphasia Rapid Test (ART) is a screening test developed for fast speech/language assessment of people in the acute stroke period. This test has been developed for French and English and was recently adapted for Portuguese and Italian. Nowadays, such a standardised screening test is in a great need at clinics with Russian-speaking patients. To fill this gap, the ART was adapted for Russian. **Aims:** The current study investigated whether the Russian ART meets all the psychometric standards, and whether it is suitable for detecting speech/language disorders and estimating their severity, as well as for the evaluation of improvement in the acute post-stroke period.

Methods & Procedure: First, we evaluated the validity, sensitivity, specificity, accuracy, test-retest reliability, inter-item consistency and inter-rater reliability of the test in a group of people with chronic speech/language disorders (N = 55) and in an age-matched control group of non-brain-damaged individuals (N = 50). Participants performed the Russian ART, and their linguistic status was confirmed by the Russian e-version of the Token Test. Second, to test the appropriateness of the Russian ART in the acute post-stroke period, a clinical group of such individuals (N = 43) performed the ART and the Token Test, as well as the Vasserman's scale which is widely used in Russian clinics. Finally, 16 people in the acute stroke period performed the Russian ART twice to prove that the test can detect early changes in an acute patient's linguistic status.

Outcomes & Results: The results showed that the Russian ART can be considered as a valid, sensitive, specific, and accurate screening tool with the high test-retest reliability, inter-item consistency, and inter-rater reliability. In the acute post-stroke group, the correlation between the ART and the Token Test was high and significant; a moderate correlation and no significant correlation were found between the Vasserman's scale and the Russian ART and the Token Test correspondingly. The Russian ART also allowed us to detect the improvement in speech/language status in the acute post-stroke period.

Conclusion: The study confirmed that the Russian ART meets all required standards to be suggested for usage in a Russian-speaking clinical population. This test was relevant for detecting the presence and severity of speech/language disorders and to measure the improvement in the acute post-stroke period.


ARTICLE HISTORY

Received 9 June 2019

Accepted 21 January 2020

KEYWORDS

Screening test, standardised, assessment, acute stroke, aphasia, rating scale

CONTACT Olga Buivolova  obuivolova@hse.ru Center for Language and Brain, National Research University Higher School of Economics, 21/4 Staraya Basmanaya Street, Office 510, Moscow 105066, Russian Federation 105066

© 2020 Informa UK Limited, trading as Taylor & Francis Group

1. Introduction

Worldwide, 15 million people suffer from stroke each year; of those, 5 million die and another 5 million are affected by disabilities (Benjamin et al., 2018). Around 21-38% of strokes cause aphasia (Berthier, 2005), a language impairment influencing one's abilities to comprehend and produce speech, to read and write. To provide adequate post-stroke patient management, it is important to identify speech and language disorders

in the first days, ideally hours, following a stroke (Godecke et al., 2013). So-called aphasia screening tests (National Stroke Foundation, 2010) are used for this purpose in the acute stage. Individuals in this stage may be cognitively affected immediately after stroke. Furthermore, regular hospital staff of the neurological ward (i.e., nurses, general practitioners, neurologists) should be able to run the tests, because speech/language therapists may be unavailable. Additionally, an effective screening test must be valid, reliable, and accurate (El Hachoui et al., 2017).

Available aphasia screening tests are designed to identify post-stroke acute aphasia and its severity in 3–15 minutes. They all assess speech production, comprehension, and repetition. They are suitable to measure the improvement in speech/language in acute post-stroke period (Rohde et al., 2018) and, importantly, to measure initial aphasia severity which helps to predict recovery potential (Benghanem et al., 2019). Such instruments have been developed for different languages: Aphasia Rapid Test (ART) for French and English (Azuar, Leger, Arbizu, & Samson, 2013); Aphasia Bedside Check (ABC; Paemeleire, 2014) and ScreeLing (Doesborgh et al., 2003) for Dutch; Frenchay Aphasia Screening Test (FAST; Enderby, Wood, Wade, Langton, & Wade, 1987) for English; Language Screening Test (LAST) for French (Flamand-Roze et al., 2011); Mississippi Aphasia Screening Test (MAST) for English (Nakase-Thompson et al., 2005), Czech (Kostalova et al., 2008) and Spanish (Romero et al., 2012); and Ullevaal Aphasia Screening (UAS; Thommessen, Thoresen, Bautz-holter, & Laake, 1999) for Norwegian. All of these tests have the same purpose – to identify speech/language problems in the acute post-stroke stage, but they are designed differently. For example, MAST and UAS also assess reading and writing skills, MAST and ART examine the ability to follow instructions, and ScreeLing assesses impairments at different linguistic levels (phonology, semantics, and syntax).

Russian is the most geographically widespread language in Eurasia and one of the 10 most spoken languages in the world (Lewis, Simons, & Fennig, 2018). However, until now, there have been no valid screening instruments for the Russian-speaking clinical population in the early post-stroke period. The only available speech and language assessment battery used in Russian acute stroke departments has been the scale for assessing the severity of speech disorders in patients with local brain injuries (Vasserman's scale; Vasserman, Dorofeeva, & Meerson, 1997). The scale examines a patient's abilities on 21 functions, assessing speech production and comprehension, gnosis and praxis, memory and cognition, including a detailed linguistic and communicative assessment (a set of tasks evaluating phonological, lexical, semantic, syntactic and discourse linguistic levels). However, the Vasserman's scale is long and effortful. It can take several hours to complete and cannot be used for people with severe communicative and cognitive deficits. It also requires a trained speech/language pathologist or a neuropsychologist for the test administration and especially for the interpretation of results. Considering that the language status of stroke patients in the acute period is unstable and they can sometimes recover spontaneously, such a time- and resource-consuming assessment may not be suitable (El Hachoui et al., 2017). That is why there is a great clinical need for a quick but valid bedside assessment of language and speech in Russian-speaking acute post-stroke individuals.

To fill this gap, we present a version of the Aphasia Rapid Test adapted for Russian, and an assessment of its validity, sensitivity, specificity, test-retest reliability, inter-item consistency, and inter-rater reliability. The ART is successfully used in clinical practice for identification of early post-stroke aphasia signs in different languages (French and English: Azuar et al., 2013; Portuguese: Tábuas-Pereira et al., 2018; and Italian: Panebianco et al., 2019). It is short (takes less than 5 minutes) and simple enough for patients in their first hours and days after a stroke. The ART was originally designed as a screening tool allowing medical staff to detect speech/language disorders, to evaluate their severity, and to track changes in a patient's linguistic status in the acute period. It does not require a speech/language pathologist or a neuropsychologist to be involved.

The original ART (Azuar et al., 2013) is a 26-point scale assessing abilities to produce and comprehend oral speech. The ART consists of six tasks: two simple and one complex instructions; repetition of words; repetition of a sentence; object naming; scoring of dysarthria; and a verbal semantic fluency task. The scoring system of the ART is based on the NIH Stroke Scale (National Institute of Neurological Disorders and Stroke, 2011), which is commonly used for neurological evaluation of speech and language in people with acute stroke. The ART does not discriminate between speech and language disorders, such as aphasia versus apraxia of speech and dysarthria. However, it indicates the presence of such problems and shows the necessity of further assessment by a speech/language therapist.

2. Russian ART

Following the principles of the original ART (Azuar et al., 2013), we developed a set of stimuli relevant for Russian (for protocol and scoring details, see Appendix 1). Some tasks (the first, fifth, and the sixth) were directly translated from English to Russian; others were changed according to specific features of the Russian language but followed the main principles of the original ART. The first task assesses a person's ability to comprehend speech; it includes two simple and one complex instructions: to close and open the eyes, to present the left hand, to put the left hand on the right ear. The second task is repetition of nouns with a different number of articulatory transitions: *kit* "whale" (1 transition), *groza* "thunderstorm" (1 transition and 1 consonant cluster), *vorotnik* "collar" (3 transitions and 1 consonant cluster). Failure on this task reflects an impairment at the articulatory, phonological and/or lexical levels. The third task examines repetition of a simple sentence with an unmarked structure in Russian and containing a subject, a verb in past tense and an object with two prenominal modifiers (*Mama kupila dva zelenykh yabloka* "Mother bought two green apples"). This task evaluates the degree of impairment at the morphosyntactic level and the level of lexical phonology. The fourth task is to name three object pictures: *myach* "ball", *zvezda* "star", *kompas* "compass". Nouns are varied in their phonological complexity, frequency and age of acquisition, with increasing difficulty on all of these parameters from the first to the third word. The task evaluates word production at the phonological and lexical-semantic levels. Finally, the semantic fluency task (Azuar et al., 2013) invites the person to name as many animals as possible in one minute.

The repetition tasks as well as object naming include measures of articulation, but as in the original ART they do not allow distinction between aphasia versus motor speech disorders. The latter problems are scored on a separate four-point scale, estimating articulation disorder severity from 0 to 3, where 0 is an absence of problems with articulation and 3 indicates a severe disorder.

To investigate whether the Russian ART is an effective clinical tool for screening speech/language disorders in the post-stroke population, we performed three studies. In the first study, we tested its validity, sensitivity, specificity, accuracy and reliability. The second study was focused on its suitability for detecting speech/language disorders in the acute period. The third study tested whether the Russian ART is suitable for improvement estimation in the first days after a stroke. The studies and informed consent forms were approved by the HSE Committee on Interuniversity Surveys and Ethical Assessment of Empirical Research (National Research University Higher School of Economics, Moscow, Russia) and complied with the Declaration of Helsinki. All participants volunteered for the studies and signed either written informed consent forms or gave their agreement to participate orally (in the clinical group), as was discussed in the ethical application.

3. Study 1. standardisation and validation of the russian ART

The Russian ART was tested for validity (the results of the test correlate with a "golden standard"), sensitivity (all people with speech/language disorders should be diagnosed as such), specificity (non-brain-damaged individuals should not be diagnosed with speech/language disorders), accuracy (false positive and false negative results are not frequent), test-retest reliability (the results of the test are consistent over time),

internal consistency (the degree to which each item contributes to the final score) and inter-rater reliability (the degree of agreement between raters).

3.1. Method

3.1.1. Participants

Study participants were a clinical group of individuals who were aphasic due to a stroke and were in the chronic period, and a control group of non-brain-damaged speakers. The clinical group ($N = 51$; 23 females; mean age = 58.9 years ($SD = 10.6$, range 38–81); mean years of education = 14 ($SD = 2.13$, range = 8–18)) was recruited at the Center of Speech Pathology and Neurorehabilitation (Moscow, Russia). When admitted to the center, the patients were thoroughly tested for language and articulation disorders with a standardised test battery by a professional speech/language therapist. Taking into consideration the fact that the ART does not discriminate between aphasia and motor speech disorders, we included people who had aphasia without speech motor problems ($N = 36$), aphasias with speech motor problems ($N = 6$), or pure speech motor problems ($N = 9$) according to the diagnosis of well-experienced speech and language therapists (see Appendix 2 for detailed information). The mean time post stroke was 22 months ($SD = 27.3$, range 2–119).

The control group of 50 people consisted of friends and relatives of the experimenters who agreed to volunteer for the study ($N = 50$, 31 females; mean age = 42.6 years ($SD = 2.5$, range 18–79); mean years of education = 13.62 ($SD = 2.5$, range = 10–18)). They did not differ from the clinical group with respect to age, gender, and educational level. All participants were (premorbidly) right-handed, native speakers of Russian, with normal or corrected to normal vision and hearing, no diagnosed psychiatric and neurodegenerative disorders and no visual agnosia.

The control group and 40 individuals from the clinical group (P1-P16; P28-P51) participated in the validation study to measure the sensitivity, specificity, accuracy and validity of the Russian ART; 14 out of these 16 individuals with speech/language disorders (P1-P14) were tested twice, with an average interval of 23 days ($SD = 7.22$, range 16–38), to assess test-retest reliability and inter-item consistency. To measure inter-rater reliability, eight individuals from the clinical group (P8, P11-P14, P28, P49, P50) were selected randomly by tossing a coin from the group that participated in the validation study and 11 others (P17-P27) were recruited additionally. These latter participants were not included in the group that participated in the validation analysis, because they only performed the Russian ART but not the Token Test. Detailed demographic information is presented in Appendix 2.

3.1.2. Materials and procedure

To estimate the concurrent validity of the adapted Russian ART, participants were tested with the ART and the Token Test (De Renzi & Faglioni, 1978), one of the most widely used tests for detecting aphasia and estimating its severity. The Token Test is a tool designed for differential diagnosis of aphasia and is very suitable for measuring the general severity of language impairment (De Renzi & Vignolo, 1962). The iPad-version of the Token Test (Token Test App: Bastiaanse, Raaijmakers, Satoer, & Visch-Brink, 2016; Russian version: Akinina et al., 2015) was used. To date, the Token Test has not been validated for Russian, neither for the original Perspex version nor for the App, but the results of the ongoing validation study show that scores on the Perspex and App versions are very similar (Akinina, Buivolova, Soloukhina, & Bastiaanse, 2019).

Participants were tested individually by a clinical linguist in a quiet place and performed the two tests in a randomised order. The responses for the ART were registered in a paper protocol and scored by the examiner (see Appendix 1). For measuring inter-rater reliability, participants' responses were scored independently by two clinical linguists, who were both present during testing. Token Test responses were

scored automatically in the App, according to the original scoring system. For performing the Token Test, participants should be able to use an iPad.

3.1.3. Data analysis

Sensitivity, specificity and accuracy, as well as positive and negative predictive values (Lalkhen & McCluskey, 2008) were calculated according to the formulae suggested by Greenhalgh (1997). The validity of the test was measured with the Spearman's correlation coefficient between the ART and Token Test scores. To estimate internal consistency, Cronbach's α (Cronbach, 1951) was employed, using the participants' first testing time-point. To measure inter-rater reliability, weighted kappa (κ_w) and the Spearman's correlation coefficient were calculated. All statistical analyses were carried out using Statistical Package for the Social Sciences (SPSS, Version 20.0) (IBM SPSS, Chicago, IL).

Table 1. Russian ART and Token Test (TT) scores obtained by participants of Study 1 (Table view)

Clinical group			Control group		
Participant ID	ART	TT	Participant ID	ART	TT
P01	1	19	C01	0	32
P02	5	20	C02	0	29.5
P03	9	20.5	C03	0	35
P04	6	12.5	C04	0	33.5
P05	9	27	C05	0	33.5
P06	1	29	C06	0	31
P07	2	30	C07	0	30
P08	6	26.5	C08	0	34
P09	15	14.5	C09	0	33.5
P10	9	11	C10	0	34
P11	13	6	C11	0	33.5
P12	9	12	C12	0	29.5
P13	2	27.5	C13	0	33.5
P14	2	19.5	C14	0	32.5
P15	2	25	C15	0	33.5
P16	5	24.5	C16	0	32.5
P27	1	21.5	C17	0	31.5
P28	19	7	C18	0	29.5
P29	0	36	C19	0	29
P30	1	22	C20	0	32.5
P31	2	18.5	C21	0	29
P32	0	32.5	C22	0	32
P33	2	31.5	C23	0	31.5
P34	0	32.5	C24	0	34
P35	3	21.5	C25	0	35
P36	1	28	C26	0	29
P37	5	29.5	C27	0	34
P38	4	11.5	C28	0	31.5
P39	4	16.5	C29	0	30.5
P40	0	32	C30	0	36
P41	17	3.5	C31	0	36
P42	1	30	C32	0	34
P43	12	7	C33	0	36

Clinical group			Control group		
Participant ID	ART	TT	Participant ID	ART	TT
P44	1	30.5	C34	0	33
P45	15	3	C35	0	36
P46	2	29.5	C36	0	34
P47	2	25.5	C37	0	36
P48	6	8	C38	0	36
P49	4	17.5	C39	0	36
P50	14	7.5	C40	0	35
P51	0	31.5	C41	0	33
			C42	0	34
			C43	0	35
			C44	0	34
			C45	0	32
			C46	0	36
			C47	0	35
			C48	0	35
			C49	0	36
			C50	0	29.5
Mean	6	19.9	Mean	0	33.15
SD	4.2	8.1	SD	0	2.2

Table 2. Sensitivity, specificity, positive and negative predictive values and accuracy of the Russian ART according to the formulae suggested by Greenhalgh (1997) (Table view)

TT (<29)		TT (≥ 29)
ART (>0)	A True positive (28)	B False positive (7)
ART (= 0)	C False negative (0)	D True negative (55)
	Sensitivity $A/(A + C)$	1
	Specificity $D/(B + D)$	0.89
	Positive predictive value $A/(A + B)$	0.8
	Negative predictive value $D/(C + D)$	1
	Accuracy $(A + D)/(A + B + C + D)$	0.92

Note: ART (>0): the score on the Russian ART is higher than 0, indicating speech/language problems. ART (= 0): the score on the Russian ART is 0, indicating the absence of speech/language problems. TT (<29): the Token Test score is lower than 29, indicating aphasia. TT (≥29): the Token Test score is equal to or higher than 29, indicating the absence of aphasia.

3.2. Results

In the clinical group, the mean overall ART score was 5.17 (SD = 5.28, range 0–19) and the mean overall Token Test score was 20.9 (SD = 9.3, range 3–36). The control group performed at ceiling on the ART: all participants scored 0 points. The mean Token Test score for the control participants was 33.15 (SD = 2.2, range 29–36). Detailed information is presented in Table 1.

3.2.1. Validity, sensitivity, specificity, accuracy and reliability

Results of the Spearman's correlation indicated that there was a significant negative association between performance on the ART and the Token Test ($r_s = -.807$, $p < .001$). The correlation was negative because more severe aphasia corresponds to higher scores on the ART but lower scores on the Token Test. Such an

association indicated that the Russian ART measured the severity of speech/language disorders as reliably as the Token Test does. This confirmed the concurrent validity of the Russian ART with the Token Test.

Detailed information on the sensitivity, specificity and accuracy of the Russian ART is presented in Table 2. The test has very high sensitivity (100% true positives) and thus allows the examiner to detect speech/language disorders if they are present. The specificity of the test is also high (89% true negatives). Thus, the ART has high positive (80%) and negative (100%) predictive values, meaning that when a participant has a positive result on the Russian ART, the score reflects the presence of the disorder detected with the Token Test. The overall accuracy of the test (true positives and negatives) is also high (92%).

3.2.2. Test-retest reliability, inter-item consistency and inter-rater reliability

Since the data are not normally distributed, statistical testing was done non-parametrically. The results of the Wilcoxon Signed-Rank Test showed that there was no significant difference between the performance of individuals in the clinical group on the Russian ART at the first and second testing time points ($Z = -1.2032$, $p = .205$). The results indicated high test-retest reliability of the Russian ART. In the inter-item consistency analysis, Cronbach's α was moderately high ($\alpha = .766$), meaning that items are well-correlated to each other (Tavakol & Dennick, 2011). Hence, the Russian version of the ART can be considered internally consistent. The inter-rater reliability measures ($r_s = .95$, $k_w = .93$) demonstrated that the two raters highly agreed in their scoring, showing that the scoring system is reliable.

3.3. Interim discussion

The results showed that the Russian version of the ART meets the standards for clinical tests (Ivanova & Hallowell, 2013) and can be considered a valid, sensitive, and reliable screening tool for identifying speech/language disorders. Scores of the clinical group on the Russian ART and the Token Test had a strong and significant correlation, suggesting high correspondence between the ART and the “golden standard”, confirming the validity of the Russian ART. In the clinical group, all participants with speech/language disorders were detected correctly with the ART, meaning that the test is highly sensitive. The specificity of the ART is also high. 17% of the cases (7 out of 41 participants) showed speech/language disorders on the ART but not on the Token Test. According to a speech/language pathologist's diagnosis, these were cases who suffered from motor speech disorders and not from aphasia. Thus, the ART showed to be a highly accurate screening test because false positive and false negative results hardly ever occurred.

Additionally, the ART showed high test-retest and inter-rater reliability in people with chronic aphasia, which means that the results of the test are consistent over time and examiners. The inter-item consistency was acceptable. Unfortunately, the paper on the original version of the ART (Azuar et al., 2013) did not provide information about the internal consistency of the French or English versions of the test, making it impossible to compare results.

4. Study 2. verification of the Russian ART in an acute clinical population

This study aimed to determine whether the new Russian ART and the test currently used in Russian clinical settings (Vasserman's scale; Vasserman et al., 1997) allows for the detection of speech/language problems as well as the “golden standard” Token Test, and which of the two former tests is more suitable for screening in the acute stroke period.

4.1. Method

4.1.1. Participants

A clinical group of people in the acute post-stroke period ($N = 43$, 18 females, mean age = 56 years ($SD = 11.3$, range 40–88), mean post onset time = 4.95 days ($SD = 2.5$, range 1–12), was recruited at the

City clinical hospital No. 31 of the Moscow Health Department (for detailed information, see Appendix 3). All participants were native Russian speakers, (premorbidity) right-handed and had no neurological or psychiatric history.

4.1.2. Materials and procedure

Participants were tested with the Russian ART and the Token Test in a randomised order. The results of the ART and the Token Test were scored in the same manner as in Study 1. On the same day, all participants were examined by a speech/language therapist using the Vasserman's scale. The severity of speech and/or language disorders was rated from 0 to 6, where the higher score reflected a more severe disorder.

4.1.3. Data analysis

All statistical analyses were carried out using the Statistical Package for the Social Sciences (SPSS, Version 20.0; IBM SPSS, Chicago, IL). Due to the small dataset, Spearman's correlation was used to evaluate the association between the ART and the Vasserman's scale, between the ART and the Token Test results, and between the Vasserman's scale and the Token Test results. Only 8 out of 43 individuals in the clinical group were able to perform the Token Test due to the general physical condition.

4.2. Results

The correlation between scores on the ART ($N = 43$, mean score = 6.2 (SD = 4.9, range 0–18)) and the Vasserman's scale ($N = 43$, mean score = 3.4 (SD = 1.84, range 0–6)) was moderate, but significant ($r_s = .547$, $p < .001$). The correlation between scores on the ART ($N = 8$, mean score = 6.62 (SD = 5.6, range 1–18)) and the Token Test ($N = 8$, mean score = 24.7 (SD = 10.8, range 0–34)) was significant and high ($r_s = -.970$, $p = .000$). Spearman's correlation between the Vasserman's scale ($N = 8$, mean score = 3.37 (SD = 1.99, range 1–5)) and the Token Test ($N = 8$, mean score = 24.7 (SD = 10.8, range 0–34)) was not significant ($r_s = -.536$, $p = .171$).

4.3. Interim discussion

Not all participants could complete the Token Test, because using the iPad version of the test was too complicated for people in the acute stroke period. However, for the 8 individuals who were able to do it, the results on the Token Test and the ART were highly correlated. That means that the ART can reliably detect language problems in the acute stroke period and is more suitable when only the linguistic status needs to be identified. The ART's strong correlation with the Token Test (even though the group was very small) shows that it does not only differentiate between stroke patients with and without speech/language problems, but the score gives an indication of the severity as well: the correlation with the Token Test suggests that the higher the score on the ART, the more severe the speech and language problems are. In combination with the Token Test, the ART detects not only the presence of aphasia but assesses articulation disorders such as dysarthria. People who have a combination of ART scores higher than 0 and Token Test scores equal to or higher than 29, do not have aphasia but need attention for their articulation problems. Thus, this can provide direction for a more detailed assessment by a speech/language pathologist.

Scores on the Vasserman's scale had no significant correlation with the "golden standard" Token Test. The Token Test is language specific, conceptually elementary, short and easy to administer, and its items are easy to memorise (De Renzi & Vignolo, 1962). The Vasserman's scale, however, consists of 107 items, estimating not only speech/language performance, but also other domains as gnosis, praxis, and attention. Performing this test can be time consuming and tiresome for patients, especially if they have cognitive deficits. However, the iPad version of the Token Test also has some limitations in the acute stroke period: people with severe motor disorders cannot perform the test on the tablet screen. The ART was designed to

estimate the presence and severity of a language disorder, just like the Token Test. Additionally, it has a specific scale which reflects the presence of articulation disorders. That is why the correlation between these two scales is moderate, but significant. All these facts demonstrate that the ART is more speech/language specific than the Vasserman's scale and easier for administration in acute settings than both the Token Test and the Vasserman's scale. Thus, the ART is more suitable for screening speech/language problems than the Vasserman's scale, which also requires a trained speech/language therapist for assessment while the ART does not.

5. Study 3. detecting changes in linguistic status in the acute period

As reported in Study 1, the Russian ART has high test-retest reliability: chronic post-stroke individuals showed the same results at two time points that were not too far apart. However, in the acute post-stroke period, spontaneous recovery is possible. This study tested whether the Russian ART is sensitive to these changes.

5.1. Method

5.1.1. Participants

The clinical group that took part in this study consisted of 16 individuals in the acute stroke period ($N = 16$, 9 females; mean age = 71.4 ($SD = 9.6$, range 49–87), mean days post onset 7.6 ($SD = 3.7$, range 3–14)), who were admitted to the City clinical hospital No. 31 of the Moscow Health Department (for detailed information, see Appendix 4). Inclusion criteria were the same as in Study 2.

5.1.2. Materials and procedure

The materials were the same as the materials of Studies 1 and 2. The participants were tested individually with the Russian ART while in bed in their hospital wards, by a clinical linguist. They performed the test twice during their stay in the hospital: the first time at the beginning of the hospitalisation in the neurology department (T1), and the second time before they left the hospital (T2): 4 days passed between testing points, on average ($SD = 1.69$, range = 2–6).

5.2. Results

The mean score on the ART at admission (T1) was 9.6 ($SD = .8$, range 2–25) and when leaving the hospital (T2) it was 7.0 ($SD = 5.9$, range 1–23). According to Wilcoxon Signed-Rank Test, this difference was significant ($Z = -3.18$, $p = .001$) at the group level. Additionally, we compared the results of participants' performance on each task and found significant improvement on three tasks: 1b – following complex instructions ($Z = -2.0226$, $p = .05$); 2a – repetition of a word without articulatory switches ($Z = -2.0226$, $p = .04$); and 6 – the semantic fluency task ($Z = -2.0226$, $p = .04$). The improvement was measured as ΔART (T2 score – T1 score). All participants improved or retained the same score (for details, see [Figure 1](#)).

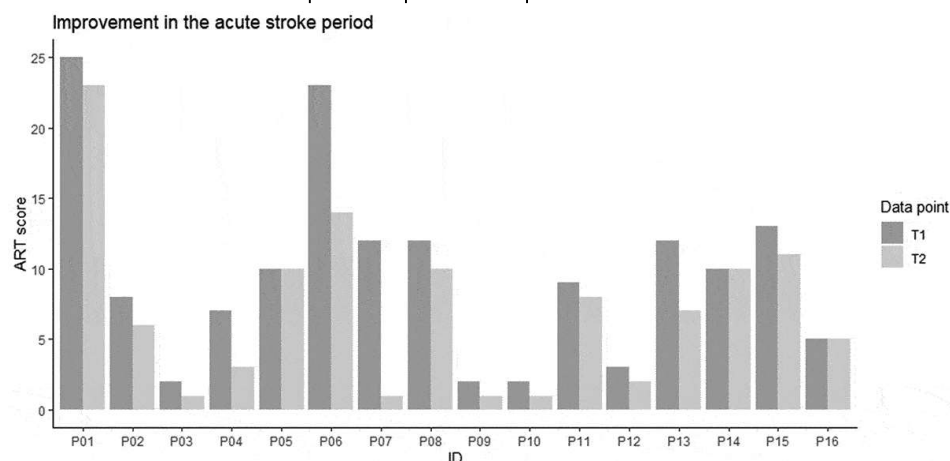


Figure 1. Changes in Russian ART scores in the acute stroke period

5.3. Interim discussion

According to the results, the Russian ART is sensitive to early changes in post-stroke individuals' linguistic profiles. In a follow up study going beyond the scope of the present paper, we will investigate whether the improvement detected with the Russian ART during the acute phase can be a reliable predictor of aphasia outcomes in the long run.

6. General discussion

Three studies have been run to test the Russian adaptation of the Aphasia Rapid Test as a relevant screening tool for speech/language problems in the acute post-stroke stage. As in the original ART (Azuar et al., 2013), its Russian adaptation is a 26-point bedside screening test allowing medical staff of neurological ward to detect the presence of speech/language disorders and to estimate their severity. This test assesses language comprehension, repetition, object naming, semantic fluency and evaluates the severity of articulation disorders as well. The test can be effectively used in acute clinical settings, because assessment takes approximately 3 minutes and does not require any special equipment or materials. However, the nature of a patient's deficit cannot be established with this tool, because there is no discrimination between aphasia and problems with articulation.

In Study 1, which was done in the group of people with chronic aphasia, we demonstrated the Russian ART to be valid, sensitive, specific, accurate and reliable with respect to test-retest effects and individual raters, whereas the internal consistency was moderately high. This suggests that the ART should be a standardised screening test, which can be further used in Russian speaking post-stroke individuals with speech/language impairments.

Study 2 showed that the tools which are now in use in Russian clinical settings are suboptimal for aphasia screening, compared to the Russian ART. The Vasserman's scale is a non-standardised, time consuming and very detailed test including measurements for a wide range of cognitive skills. Concurrently, we found that the tablet version of the Token Test cannot be used in the acute period either: it is difficult to perform for people with motor disabilities and paralyses and is tiring for most of them. Instead, the Russian ART can fill the niche and become an easy to use screening test for speech/language disorders. Although the ART does not discriminate between language and motor speech disorders, the combination of scores obtained by the ART and the Token Test can provide direction for the speech/language therapist. For example, if a patient scored above the cut-off on the Token Test, he or she probably suffers from isolated motor speech disorders, which is valuable information for a speech/language therapist.

In Study 3, we provided evidence that the Russian ART is sensitive enough to detect early changes in acute post-stroke individuals' speech/language status. Individuals in an acute hospital unit were tested at two

time points of their stay in the hospital – when they were admitted and were able to perform the test, and when they were discharged. The results confirmed that the ART can detect these early changes over time.

Considering the results of our research, the Russian ART can be recommended for usage in the acute post-stroke population. This test meets all the psychometric standards and fills the gap in standardised and normalised screening tests. Thus, all members of medical staff of neurological ward working with Russian speakers in acute hospitals can use this instrument for detecting speech/language disorders and assessing the improvement in speech/language.

7. Limitations

Even though sample sizes are relatively small in each of the three studies, 51 people with chronic aphasia participated in Study 1, 66 people with acute aphasia took part in Studies 2 and 3, and 50 healthy individuals were recruited as a control group. Overall, the size of our database resembles the sample size in the original study (Azuar et al., 2013), with 161 participants overall. However, further data collection, especially combined with Token Test data in the acute stage, will further strengthen our claims that the ART is a valuable addition as a screening tool for speech and language impairments.

We used the Token Test app because it is sensitive and reliable, even though the Russian iPad version has not been standardised yet. Preliminary results (Akinina et al., 2019) show that all psychometric properties of the test are high, and this instrument is as reliable as its original version. Not all participants could perform the Token Test due to their motor or cognitive disabilities. This is not only the case for the iPad version, with our clinical experience revealing that the Perspex version is also challenging for individuals in the acute period. We hoped that the iPad version would make it easier for this patient group, but that is not the case.

We had access to two facilities: one hospital with a stroke unit (City clinical hospital No. 31 of the Moscow Health Department) and the Center for Speech Pathology and Neurorehabilitation for people suffering from chronic speech and language disorders. Both hospitals are located in Moscow. We did not have access to hospitals in other Russian regions. We will continue our data collection in other centers outside of the Moscow Region, recruiting stroke victims with other educational and cultural backgrounds.

8. Future directions

The developers of the original version of the ART claimed that results obtained by this test can predict aphasia outcome in the chronic phase (Azuar et al., 2013). It will be helpful when the initial severity of speech/language disorders measured by the Russian ART can be used as a predictor for the outcome of aphasia in the long term. So far the results look promising, but further investigation is needed.

We standardized the Russian ART in a post-stroke population. It may be worthwhile to investigate whether this test is suitable for usage in post-neurosurgical populations as well. Also, there is a great need for a short and simple screening test that can be used in people with very severe aphasia, especially in settings when there is no speech/language therapist available in the hospital. The group of aphasic individuals in the current study included some severely aphasic individuals, but a study focusing on this specific population may show whether the ART is a useful tool in this population.

There is a great need for screening tests that detect language disorders in the acute stage. The ART only gives direction for future diagnostics, not for the kind of treatment that is needed, so it is necessary to develop more specific language screening instruments for Russian. We are now working on Russian versions of two tests (ABC: Paemeleire, 2014; ScreeLing: Doesborgh et al., 2003). We will compare the results of these three instruments to find out which test can be used in which population and in which period.

Acknowledgments

We are grateful to Roman A. Cheremin and Galina E. Ivanova for their help with organisation of the study and for providing access to people admitted to the Center of Speech Pathology and Neurorehabilitation and City clinical hospital

No. 31 of the Moscow Health Department. The authors thank Olga A. Soloukhina, Anastasia A. Shlyakhova and Victoria A. Pozdnyakova for their help with patient recruitment and data collection. We acknowledge all colleagues from the Center for Language and Brain of the National Research University Higher School of Economics for their comments on the manuscript. The authors are grateful to Anna Linden for illustrations for the test. We would also like to express our gratitude to all the participants of the present study.

Funding

The study was supported by the Center for Language and Brain NRU Higher School of Economics, RF Government grant, ag. № 14.641.31.0004.

Supplementary material

Supplemental data for this article can be accessed [here](#).

References

- Akinina, Y., Buivolova, O., Soloukhina, O., & Bastiaanse, R. (2019). Psychometric properties of the Token Test App. *Stem-, Spraak- En Taalpathologie*, 24, 74–76.
- Akinina, Y., Dragoy, O., Raaijmakers, S., Satoer, D., & Bastiaanse, R. (2015). *The e-Token Test: Russian version*. Groningen (NL): Groningen Expert Center for Language and Communication Disorders.
- Azuar, C., Leger, A., Arbizu, C., & Samson, Y. (2013). The aphasia rapid test: An NIHSS-like aphasia test. *Journal of Neurology*, 260, 2110–2117. [Crossref](#)
- Bastiaanse, R., Raaijmakers, S., Satoer, D., & Visch-Brink, E. (2016). The Multilingual Token Test. *Aphasiology*, 30, 508. [Crossref](#)
- Benghanem, S., Rosso, C., Arbizu, C., Moulton, E., Dormont, D., Leger, A., ... Samson, Y. (2019). Aphasia outcome: The interactions between initial severity, lesion size and location. *Journal of Neurology*, 266, 1303–1309. [Crossref](#)
- Benjamin, E. J., Virani, S. S., Callaway, C. W., Chamberlain, A. M., Chang, A. R., Cheng, S., ... Muntner, P. (2018). Heart disease and stroke statistics - 2018 update: A report from the American Heart Association. *Circulation*, 137, e67–e492. [Crossref](#)
- Berthier, M. L. (2005). Poststroke aphasia: Epidemiology, pathophysiology and treatment. *Drugs and Aging*, 22, 163–182. [Crossref](#)
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297–334. [Crossref](#)
- De Renzi, E., & Faglioni, P. (1978). Normative data and screening power of a shortened version of the Token Test. *Cortex*, 14, 41–49. [Crossref](#)
- De Renzi, E., & Vignolo, L. A. (1962). The Token Test: A sensitive test to detect receptive disturbances in aphasics. *Brain*, 85, 665–678. [Crossref](#)
- Doesborgh, S. J. C., van de Sandt-koenderman, W. M. E., Dippel, D. W. J., van Harskamp, F., Koudstaal, P. J., & Visch-Brink, E. G. (2003). Linguistic deficits in the acute phase of stroke. *Journal of Neurology*, 250, 977–982. [Crossref](#)
- El Hachoui, H., Visch-Brink, E. G., de Lau, L. M. L., van de Sandt-koenderman, M. W. M. E., Nouwens, F., Koudstaal, P. J., & Dippel, D. W. J. (2017). Screening tests for aphasia in patients with stroke : A systematic review. *Journal of Neurology*, 264, 211–220. [Crossref](#)
- Enderby, P. M., Wood, V. A., Wade, D. T., Langton, R., & Wade, T. (1987). The frenchay aphasia screening test: A short, simple test for aphasia appropriate for non-specialists. *International Rehabilitation Medicine*, 8, 166–170. [Crossref](#)
- Flamand-Roze, C., Falissard, B., Roze, E., Maintigneux, L., Beziz, J., Chacon, A., ... Denier, C. (2011). Validation of a new language screening tool for patients with acute stroke: The Language Screening Test (LAST). *Stroke*, 42, 1224–1229. [Crossref](#)
- Godecke, E., Rai, T., Ciccone, N., Armstrong, E., Granger, A., & Hankey, G. (2013). Amount of therapy matters in very early aphasia rehabilitation after stroke: A clinical prognostic model. *Seminars in Speech and Language*, 34, 129–141. [Crossref](#)
- Greenhalgh, T. (1997). Papers that report diagnostic or screening tests. *British Medical Journal*, 315, 540–543. [Crossref](#)
- Ivanova, M. V., & Hallowell, B. (2013). A tutorial on aphasia test development in any language: key substantive and psychometric considerations. *Aphasiology*, 27(8), 891–920.
- Kostalova, M., Bartkova, E., Sajgalikova, K., Dolenska, A., Dusek, L., & Bednarik, J. (2008). A standardisation study of the Czech version of the Mississippi Aphasia Screening Test (MASTcz) in stroke patients and control subjects. *Brain Injury*, 22, 793–801. [Crossref](#)
- Lalkhen, A. G., & McCluskey, A. (2008). Clinical tests: Sensitivity and specificity. *Continuing Education in Anaesthesia, Critical Care and Pain*, 8, 221–223. [Crossref](#)
- Lewis, M. P., Simons, G. F., & Fennig, C. D. (2018). Statistical Summaries. Summary by language size. Language size. In *Ethnologue: Languages of the World* 21st, Dallas: SIL International. Retrieved July 16th, 2018 from <https://www.ethnologue.com/statistics/summary-language-size-18>
- Nakase-Thompson, R., Manning, E., Sherer, M., Yablon, S. A., Gontkovsky, S. L. T., & Vickery, C. (2005). Brief assessment of severe language impairments: Initial validation of the Mississippi aphasia screening test. *Brain*

- Injury*, 19, 685–691. [Crossref](#)
- National Institute of Neurological Disorders and Stroke (U.S.). (2011). *NIH stroke scale*. Bethesda, Md.: National Institute of Neurological Disorders and Stroke, Dept. of Health and Human Services, USA.
- National Stroke Foundation. (2010). *Clinical Guidelines for Stroke Management 2010*. Australia: Melbourne.
- Paemeleire, F. (2014). De Afasie Bedside Check (ABC). Een nieuwe hulp voor het detecteren van afasie door artsen. *Wat Zeg Je?*, 6, 2–4.
- Panebianco, M., Zavanone, C., Dupont, S., Perrone, P., Pavone, A., Napolitano, S., ... Samson, Y. (2019). The inter-rater reliability of the Italian version of Aphasia Rapid Test (ART) for acute ischemic stroke. In *Neurological Sciences*, 40(10), 2171–2174.
- Rohde, A., Worrall, L., Godecke, E., O'Halloran, R., Farrell, A., & Massey, M. (2018). Diagnosis of aphasia in stroke populations: A systematic review of language tests. *PLoS ONE*, 13, 1–17. [Crossref](#)
- Romero, M., Sanchez, A., Marin, C., Navarro, M. D., Ferri, J., & Noe, E. (2012). Clinical usefulness of the Spanish version of the Mississippi Aphasia Screening Test (MASTsp): Validation in stroke patients/Utilidad clinica de la version en castellano del Mississippi Aphasia Screening Test (MASTsp): Validacion en pacientes con ictus. *Neurology*, 79, 216–224. [Crossref](#)
- Tábuas-Pereira, M., Freitas, S., Beato-Coelho, J., Ribeiro, J., Parra, J., Martins, C., ... Santana, I. (2018). Aphasia Rapid Test: Estudos de Tradução, Adaptação e Validação para a População Portuguesa. *Acta Médica Portuguesa*, 31, 265. [Crossref](#)
- Tavakol, M., & Dennick, R. (2011). Making Sense of Cronbach's Alpha. *International Journal of Medical Education*, 2, 53–55. Editorial [Crossref](#)
- Thommessen, B., Thoresen, G. E. V. A., Bautz-holter, E., & Laake, K. (1999). Screening by nurses for aphasia in stroke — Ullevaal Aphasia Screening (UAS) test. *Disability and Rehabilitation*, 21, 110–115. [Crossref](#)
- Vasserman, L. I., Dorofeeva, S. A., & Meerson, Y. A. (1997). *Metody neiropsikhologicheskoy diagnostiki: Prakticheskoe rukovodstvo. [Neuropsychological assessment methods: a Practical guide]*. St. Petersburg: Stroylespechat'.