

## The Aphasia Rapid Test: an NIHSS-like aphasia test

C. Azuar · A. Leger · C. Arbizu · F. Henry-Amar ·  
S. Chomel-Guillaume · Y. Samson

Received: 1 February 2013/Revised: 21 April 2013/Accepted: 26 April 2013/Published online: 15 May 2013  
© The Author(s) 2013. This article is published with open access at Springerlink.com

**Abstract** The Aphasia Rapid Test (ART) is a 26-point scale developed as a bedside assessment to rate aphasia severity in acute stroke patients in <3 min. We tested its inter-rater reproducibility, its sensitivity to detect changes from Day 1 to Day 8, and the predictive value of D8 ART scores on the 3-month aphasia outcome assessed with the Aphasia Handicap Score (AHS), a 0–5 “Rankin-like” score for aphasic disability. The reproducibility was tested in 91 aphasic patients within one week of stroke onset. The inter-rater concordance coefficient was 0.99 and the weighted Kappa value ( $\kappa_w$ ) was 0.93. The sensitivity was tested in 70 aphasic patients by measuring changes in ART values between D1 and D8. Improvement occurred in 46 patients (66 %) and aggravation in three patients (4 %). In these patients, a logistic regression analysis

showed that D8 ART was the only significant predictor of good (AHS 0–2) or poor (AHS 4–5) outcome. The ROC curves analyzes showed areas under the curve above 0.9 for good and poor outcome and revealed D8 ART best cut-off values of <12 for good and >21 for poor outcome, with more than 90 % sensitivity and 80 % specificity. The ART is a simple, rapid and reproducible language task, useful in monitoring early aphasic changes in acute stroke patients and highly predictive of the 3-month verbal communication outcome. It should be easy to adapt to other languages.

**Keywords** Acute stroke · Outcomes · Functional recovery · Language outcome · Rating scale

**Electronic supplementary material** The online version of this article (doi:10.1007/s00415-013-6943-x) contains supplementary material, which is available to authorized users.

C. Azuar · A. Leger · Y. Samson  
APHP Service des Urgences Cérébro-Vasculaires, Groupe Hospitalier Pitié-Salpêtrière, COGIMAGE CR ICM, ICM UPMC Paris VI Inserm U 975, CNRS UMR 7225, 75013 Paris, France

C. Azuar · C. Arbizu · F. Henry-Amar  
APHP-Fédération de Neurologie, Institut de la Mémoire et de la Maladie d’Alzheimer, Groupe Hospitalier Pitié-Salpêtrière, 75013 Paris, France

C. Azuar (✉)  
Institut de la Mémoire et de la Maladie d’Alzheimer, Groupe Hospitalier Pitié-Salpêtrière, 47, Boulevard de l’Hôpital, 75013 Paris, France  
e-mail: carole.azuar@psl.aphp.fr

S. Chomel-Guillaume  
Service de Neurologie, Centre Hospitalier de Versailles, Le Chesnay, 78157 Paris, France

### Introduction

The severity of neurological deficit is now routinely quantified in acute stroke patients using the NIH stroke scale (NIHSS), a standard in stroke care and research. This bedside neurological examination is simple, rapid, and reproducible. It is sensitive enough to detect early changes in neurological status, and has been shown to be highly predictive of stroke outcome (<http://www.nihstrokescale.org>). We reasoned that an NIHSS-like scale of language would be useful in grading the severity of aphasia in acute stroke patients. Most aphasia rating scales are too long to be used in acute stroke patients [6]. They are designed for trained speech therapists and require specific material. Some bedside assessments of aphasia have recently been developed [2–4, 9], but their sensitivity in detecting rapid changes in the severity of aphasia and in predicting language recovery is limited or has not been evaluated [3, 4, 9].

We designed the Aphasia Rapid Test (ART) as an NIHSS-like aphasia scale, based on the scoring of items that

are commonly used in the neurological examination of language in acute stroke patients. It has been designed to be easy to translate into any language, and to be as little language-specific as possible. It can be administered by any health care professional after brief training, without requiring any specific test material. The ART should not be used as a diagnostic tool since it does not discriminate between aphasia, apraxia of speech and dysarthria. However, we reasoned that the ART may be useful to monitor changes in aphasia severity during the acute stage and to predict aphasia prognosis. Here we first describe the ART, its scoring system, and its reliability across two different examiners. Next, we detail the sensitivity of the ART in detecting change in language skills during the first week post-stroke. Finally, we present the value of the ART in predicting language ability at three months using a different measure, the Aphasia Handicap Scale (AHS).

## Methods

### Scoring systems

The ART was designed by two neurologists (YS and AL) with extensive experience with aphasic patients, in an attempt to quantify the severity of aphasia during the acute phase of a stroke. Earlier, more complex versions were discarded at preliminary stages, because they were found to be unsuitable for bedside examination in acute stroke patients or were found to have low reproducibility in preliminary investigations. The ART score ranges from 0 to 26, with higher values indicating more severe impairment. The patient is successively asked to follow two simple orders (maximum 2 points), one more complex order (3 points), repeat three single words (6 points), repeat one sentence (2 points) and name three common objects (6 points). This is followed by a 1-min verbal semantic fluency task (4 points). The examiner additionally scores dysarthria (3 points) using the same scoring system as in the NIHSS. Since the ART has been designed as a bedside clinical tool, there is no explicit time limit for patient responses. Table 1 shows an English version of the ART and explains the scoring system (the Online Resource 1 shows the French version).

The outcome of aphasia was assessed three months post-stroke using the Aphasia Handicap Scale (AHS), a modified Rankin-score-like five-point scoring system for handicap in verbal communication, designed by two of the authors (YS and SCG) [1]. The scores are as follows: 0 = normal language, 1 = minor difficulties of language without disability (no impact on normal life), 2 = mild language-related disability (without restrictions in the autonomy of verbal communication in daily life), 3 = moderate language-related disability (restricted

autonomy of verbal communication), 4 = severe language-related disability (lack of effective verbal communication), 5 = mutism or total loss of verbal expression and comprehension. The scoring system is patient-oriented. If oral communication is preserved, the examiner asks the patient or his/her proxy, if necessary, to self-rate his/her language abilities using a semi-structured interview, as described in the Online Resource 2. In case of ambiguity between moderate (score 3) and severe (score 4) disability, the rule of thumb is to assign a score of 3 if it is possible to score activities of daily living using a scale such as the Barthel index by oral communication with the patient alone, and to assign a score of 4 if the help of a proxy is necessary. Since the AHS has not been published in a peer-reviewed journal, we retrospectively compared these results with a conventional and well-established language testing battery. The search for patients of the follow-up study, who received a formal assessment of aphasia by a speech therapist three months post-stroke using the French version of the Boston Diagnostic Aphasia Examination (BDAE) [8], yielded 37 patients. It should be stressed however that functional scales (such as the AHS) and impairment scales (such as the BDAE) are of course correlated (see Online Resource 2) but are clearly different.

### Patients and studies

All patients were recruited through our stroke unit. All were right-handed, with French as their first language. The study was approved by the local ethics committee, and in agreement with French legislation, informed consent was waived since assessing the severity of aphasia is part of standard care in stroke patients.

### *Inter-rater reliability of the ART*

We included 91 patients with acute stroke confirmed by MRI, and considered as aphasic by the neurologists and speech therapists of the stroke unit. Patients with impaired consciousness were excluded. The patients were tested at a median post-stroke delay of eight days (inter-quartile range, IQR: 7–10) by two independent examiners, who administered the ART on the same day at a maximum interval of 12 h. The examiners were stroke neurologists or speech therapists who were not involved in the development of the ART, and were blind to each other's ratings.

The inter-rater reproducibility of the ART was assessed by computing the coefficient of concordance, the weighted Kappa values ( $\kappa_w$ ), and constructing a Bland–Altman plot of the total ART scores of both examiners. In addition, weighted Kappa values ( $\kappa_w$ ) were calculated for the scores of each item. All statistical analyses were carried out using MedCalc for Windows (version 11.6.1.0; <http://www.medcalc.be>).

**Table 1** Instructions and scoring system for the Aphasia Rapid Test

Instructions	Score
1a. Execution of simple orders: «Open and close your eyes» «Give me your left hand»	0 = performs both tasks correctly. 1 = performs one task correctly. 2 = performs neither task correctly.
1b. Execution of a complex order: «Put your left hand on your right ear»	0 = performs the task in less than 10 s. 1 = performs the task in more than 10 s or requires the order to be repeated. 2 = performs the task partially: moves the hand across the median line or performs the task on the wrong side. 3 = does not perform the task: does not move the hand across the median line or does not move at all.
2. Repetition of words:	Each word scores from 0 to 2 (total 0-6), as follows:
2a. « button»	0 = normal repetition.
2b. « macaroon»	1 = abnormal repetition but the word is correct and recognizable by the examiner *.
2c. « luggage»	2 = non-repetition or unrecognizable word *. <i>*Note: Phonemic, apraxic or pronunciation errors can be scored 1 if the word is recognizable by the examiner, or 2 if the word is unrecognizable.</i>
3. Repetition of a sentence: «The boy is singing in the woods.»	0 = normal repetition. 1 = abnormal repetition but the sentence is recognizable by the examiner *. 2 = non-repetition or unrecognizable sentence*. <i>*Note: Phonemic, apraxic or pronunciation errors can be scored 1 if the sentence is recognizable by the examiner, or 2 if the sentence is unrecognizable.</i>
4. Object naming:	0 = normal naming.
4a. « watch»	1 = abnormal naming but the word is correct and recognizable by the examiner.*
4b. « pen»	2 = wrong naming or unrecognizable word. *
4c. « coat »	<i>*Note: Phonemic, apraxic or pronunciation errors can be scored 1 if the word is recognizable by the examiner. An unrecognizable word or lexical error must be scored 2.</i>
5. Scoring of dysarthria :	0 = normal. 1 = minor dysarthria. 2 = moderate dysarthria: patient can be understood. 3 = severe dysarthria: unintelligible speech.
6. Verbal semantic fluency task: «Name as many animals as you can in one minute. »	0 = more than fifteen words. 1 = between eleven and fifteen words. 2 = between six and ten words. 3 = between three and five words. 4 = between zero and two words.
Total Score	/26

### Sensitivity of the ART during the first week post stroke

We included another population of 70 consecutive patients who met the following criteria: admission to the stroke unit within 12 h of the onset of a first-ever stroke; left MCA infarct confirmed by MRI; acute aphasia noted by the neurologist on duty; lack of consciousness disorders; ART performed within 24 h of stroke onset (D0) and at eight - days (D8); and an AHS score obtained at three months during patient follow-up by a stroke neurologist blind to the D0 and D8 ART scores. The ART data were not used to plan speech therapy strategies and all patients received speech therapy as usual in the stroke unit and during rehabilitation. We determined the changes in ART between

D0 and D8 in the whole group of patients and the proportion of patients who had a significant change between D0 and D8 ART values. We also investigated whether patients with good, intermediate or poor 3-month outcome differed in D0 and D8 ART values by running a two-way repeated measure ANOVA.

### Prediction of 3-month aphasia outcome

This was done in the same population of 70 patients. We conducted two stepwise logistic regression models, the first predicting good language outcome (AHS 0–2) and the second predicting poor language outcome (AHS score 4–5). In both models, the independent variables were

gender, age, recombinant tissue plasminogen activator (rt-PA) treatment, and D0 and D8 ART scores. The variables were retained in the final model at  $p < 0.01$ . In addition, Receiver Operating Characteristic (ROC) curves were generated to compare the predictive value of the ART score at D0 and D8 for good and poor outcomes. Since AHS has not been published in a peer review journal, we performed similar analyses in the subgroup of 37 patients, which had a formal assessment of aphasia by a speech therapist three months post-stroke using the French version of the Boston Diagnostic Aphasia Examination (BDAE) [8].

## Results

### Inter-rater reliability of the ART

The mean age ( $\pm$ SD) of the 91 patients (52 men and 39 women) was  $63.96 \pm 19.3$  years. They had ischemic ( $n = 80$ ) or hemorrhagic ( $n = 11$ ) stroke. The median ART value was 11 (IQR: 4.25–24) for Rater 1 and 12 (4.25–24) for Rater 2, and the mean ( $\pm$ SD) ART value was  $13.4 (\pm 9.51)$  for Rater 1 and  $13.49 (\pm 9.52)$  for Rater 2.

The inter-rater agreement was good, with a coefficient of concordance of 0.990 (95 % confidence interval, CI: 0.985–0.993;  $p < 0.0001$ , Fig. 1a) and a  $\kappa_w$  of 0.934 (95 % CI: 0.909–0.958). The  $\kappa_w$  of each item is shown in Table 2, and ranged from 0.967 for the denomination of the watch to 0.854 for the scoring of dysarthria. The Bland–Altman plot (Fig. 1b) showed that there was no test–retest effect and that ART reproducibility was stable across all degrees of aphasia severity. A difference of more than two points indicated a significant change in aphasia severity. The mean duration of ART administration was calculated for the first 58 patients and was found to be 177 s, including the 1-min fluency task. Thus, the ART is quick to administer, with no test–retest effects, and good rater reliability.

### Sensitivity of the ART during the first week post stroke

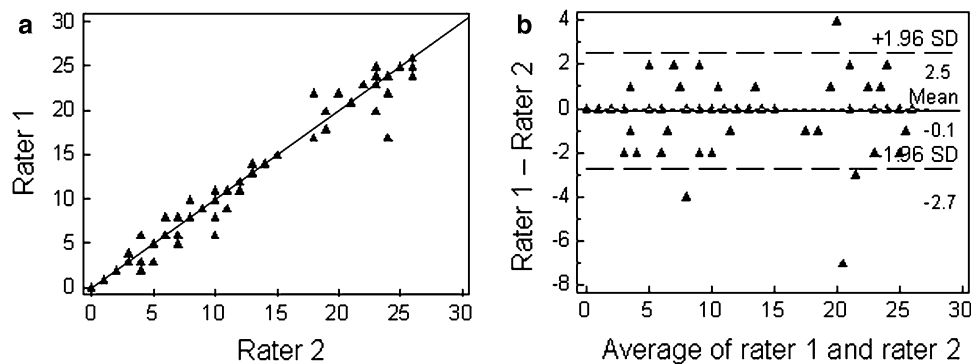
The mean age ( $\pm$ SD) of the 70 patients (41 men, 29 women) was  $61.2 \pm 15.6$  years, and they had a median initial NIHSS score of 16 (IQR: 8–22). Thirty-two patients (46 %) were treated with intravenous rt-PA. At D0, the mean ( $\pm$ SD) ART score was  $19.6 (\pm 7.8)$  and the median score 24 (IQR: 13–25, range 1–26). The ART score correlated with the initial NIHSS score at D0 ( $r: 0.635$ ,  $p < 0.0001$ ) and at D8 ( $r: 0.525$ ,  $p < 0.0001$ ) but not with age, gender or rt-PA treatment (stepwise multiple regression). At D8, the ART score had significantly decreased ( $p < 0.0001$ ) with a mean ( $\pm$ SD) value of  $12.5 (\pm 9.5)$  and a median of 10 (IQR: 4–23). The difference in the ART score between D0 and D8 (i.e., the score on D0—the score

on D8,  $\Delta$ ART) was above 2 points (i.e., revealing an improvement) in 46 patients (66 %) and below  $-2$  points (i.e., revealing an aggravation of aphasia) in three patients (4 %), who all suffered from an enlargement of their infarct during the first few days of stroke. Figure 2 shows the ART values at D0 and D8 in the subgroups of patients with good (AHS 0–2, 33 patients, 47 %), intermediate (AHS 3, 22 patients, 31 %), and poor (AHS 4–5, 15 patients, 21 %) language 3-month outcome groups. The two-way ANOVA for group and time showed significant group ( $F(2,134) = 39.5$ ,  $p < 0.0001$ ) and time ( $F(2,134) = 52.9$ ,  $p < 0.0001$ ) effects. The Fig. 2 also shows that D0 ART was lower in the good recovery group and that the changes between D0 and D8 in ART scores differed across the three groups, as confirmed by a significant group X time interaction [ $F(2,134) = 4.8$ ,  $p < 0.01$ ]. This is also shown by an ANOVA for  $\Delta$ ART ( $F: 12.5$ ,  $p < 0.0001$ ). Post hoc tests showed significant differences ( $p < 0.05$ ) between good ( $\Delta$ ART =  $11.2 \pm 7.9$ ), intermediate ( $\Delta$ ART =  $6.2 \pm 8.3$ ) and poor ( $\Delta$ ART =  $-0.7 \pm 6.1$ ) recovery groups. In summary, ART appears to be highly sensitive to change in the first week post-stroke.

### Prediction of 3-month aphasia outcome

#### Good (AHS 0–2) and poor (AHS 4–5) outcome

In the logistic regression analysis, the ART score at D8 remained the only significant predictor of good (odds ratio, OR: 0.75, 95 % CI: 0.66–0.85,  $p < 0.0001$ , accuracy: 88.6 %), or poor outcome (OR: 1.60, 95 % CI: 1.15–2.23,  $p = 0.005$ , accuracy: 88.6 %), and age, gender, rt-PA treatment, and D0 ART scores were not retained in the final logistic regression models. The ROC analysis of the D8 ART score showed that the area under the curve (AUC) was very high for good (0.926, 95 % CI: 0.838–0.975,  $p < 0.0001$ ) and poor (0.955, 95 % CI: 0.876–0.990,  $p < 0.0001$ ) outcomes. A comparison of the AUC generated with ART scores at D8 and D0 confirmed that ART was a better predictor at D8 than at D0 of both good ( $p = 0.02$ ) and poor ( $p = 0.007$ ) outcomes (Fig. 3). The best prediction of good recovery (AHS 0–2) was yielded by a D8 ART score of  $<12$ , which was associated with a sensitivity of 93.9 % (95 % CI: 79.7–99.1), a specificity of 83.8 % (95 % CI: 68.0–93.8), a positive predictive value of 83.8 % and a negative predictive value of 93.9 %, whereas the best prediction of poor recovery (AHS 4–5) was observed with a D8 ART score of  $>21$ , associated with a sensitivity of 93.3 % (95 % CI: 68.0–98.9), a specificity of 89.1 % (95 % CI: 77.7–95.9), a positive predictive value of 70.0 %, and a negative predictive value of 98.0 %. Figure 4 shows the distribution of the AHS at three months as a function of these D8 ART thresholds. Note that most of

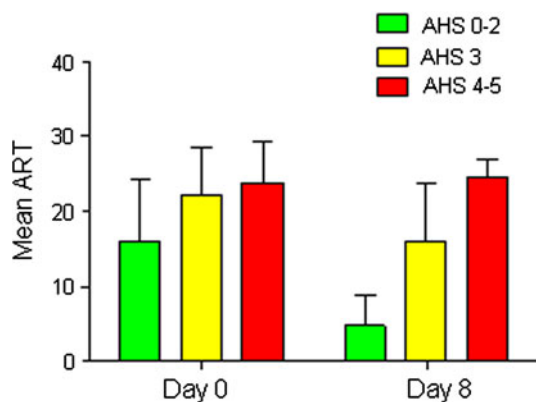


**Fig. 1** High inter-rater reproducibility of the ART tested in 91 aphasic patients. **a** ART scores rated on the same day by two independent examiners (Raters 1 and 2) showing a coefficient of concordance of 0.990 and a weighted kappa value of 0.934. **b** Bland–

Altman plot showing that ART reproducibility is stable across all degrees of aphasia severity, with no test–retest effect. Note that a difference of >2 points indicates a significant change in aphasia severity

**Table 2** Weighted kappa value for each item of the ART

ART items	1a	1b	2a	2b	2c	3	4a	4b	4c	5	6
Weighted kappa	0.874	0.863	0.874	0.921	0.833	0.842	0.967	0.889	0.885	0.854	0.925



**Fig. 2** D0 and D8 ART in patients with good, intermediate and poor 3-month language outcomes. In *green*, ART values (mean  $\pm$  SD) at Day 0 and Day 8 in patients with good outcome ( $n = 33$ , 47 % of patients). In *yellow*, ART values in patients with intermediate outcome ( $n = 22$ , 31 % of patients). In *red*, ART values in patients with poor outcome ( $n = 15$ , 21 % of patients). A two-way ANOVA shows that group and time effects and the group X time interaction were significant ( $p < 0.0001$ )

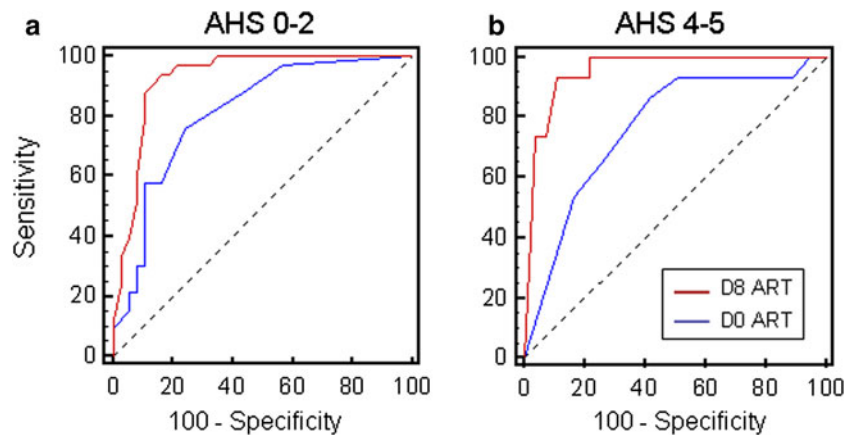
the patients (77 %) with intermediate D8 AHS score had moderate language related disability. In summary, D8 ART appears to be a good predictor of 3-month post-stroke language-related disability perhaps because it integrates the D0 ART value and also the changes occurring during the first week post stroke.

#### BDAE aphasia severity rating scale

Data were available for 37 patients at three months. The predictive value of the ART was tested by generating two

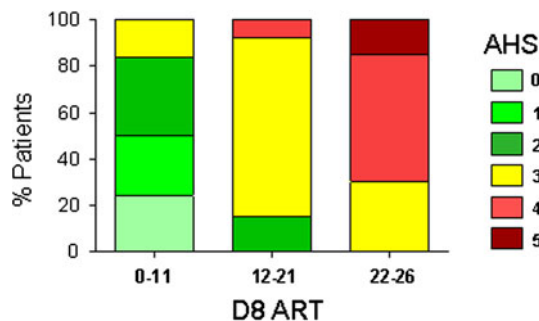
ROC curves, the first predicting good language outcome (BDAE 4–5) and the second predicting poor language outcome (BDAE 0–1). The ROC analysis of the D8 ART score showed that the AUC was very high for good (0.946, 95 % CI: 0.818–0.992,  $p < 0.0001$ ) and poor (0.93, 95 % CI: 0.795–0.986,  $p < 0.0001$ ) outcomes. The prediction of good recovery (BDAE 4–5) yielded by a D8 ART score of <12 (best prediction for AHS 0–2) was associated with a sensitivity of 94.4 % (95 % CI: 72.6–99.1), a specificity of 78.9 % (95 % CI: 54.4–93.8), a positive predictive value of 81 % and a negative predictive value of 93.8 %, whereas the prediction of poor recovery (BDAE 0–1) yielded by a D8 ART score of >21 (best prediction for AHS 4–5) was associated with a sensitivity of 70 % (95 % CI: 34.8–93.0), a specificity of 92.6 % (95 % CI: 75.7–98.9), a positive predictive value of 77.8 %, and a negative predictive value of 89.3 %.

We also compared in these patients the BDAE aphasia severity ratings with AHS scores with kappa statistics. The weighted Kappa obtained by comparing BDAE aphasia severity ratings of 0–5 with AHS scores of 5–0 was 0.89 (95 % CI: 0.84–0.94). The classification of the poor outcome group, defined by AHS 4–5 and BDAE 0–1, had a 91.9 % concordance level and a Kappa value of 0.79 (95 % CI: 0.56–1). The classification of the favorable outcome group, defined by AHS 0–2 and BDAE 3–5, had an 86.5 % concordance level and a Kappa value of 0.72 (95 % CI: 0.49–0.95). However, when the comparison was restricted to AHS 0–2 and BDAE 4–5 (instead of 3–5) the agreement was higher: concordance 91.9 %, Kappa value of 0.84 (95 % CI: 0.66–1).



**Fig. 3** ROC curves for good (AHS 0–2) and poor (AHS 4–5) 3-month language outcomes. **a** ROC analysis of D0 (blue curve) and D8 ART scores (red curve) for good outcome (AHS 0–2). The area under the curve (AUC) was significantly larger at D8 (0.926, 95 % CI: 0.838–0.975) than at D0 (0.811, 95 % CI: 0.700–0.895,  $p = 0.02$ ). A D8 ART value of <12 predicted good outcome with

93.9 % sensitivity and 83.8 % specificity. **b** ROC analysis of D0 (blue curve) and D8 ART scores (red curve) for poor outcome (AHS 4–5). The AUC was significantly larger at D8 (0.955, 95 % CI: 0.876–0.990) than at D0 (0.766, 95 % CI: 0.650–0.859,  $p = 0.007$ ). A D8 ART value of >21 predicted poor outcome with 93.3 % sensitivity and 89.1 % specificity



**Fig. 4** Distribution of 3-month AHS as a function of D8 ART thresholds predicting different language outcomes. The left bar corresponds to patients with a D8 ART score of <12 ( $n = 38$ ), the right bar to patients with a D8 ART score of >21 ( $n = 20$ ), and the middle bar to patients with intermediate scores (12–21)

**Discussion**

The ART was designed to quantify the severity of aphasia in acute stroke patients by assessing, in <3 min, comprehension, repetition, naming, and verbal fluency, the four major components affected in classic aphasic syndromes [6]. This short duration is partially explained by the simplicity of the task and partially by the rapidity of scoring in acute global aphasic patients, who fail most of the items. Like the NIHSS, it is a bedside test that does not require any specific material and can be taught to residents and nurses in a few minutes. The simplicity and rapidity of the scale are required to achieve the exhaustive monitoring of acute stroke patients, who tire easily, and who are hospitalized and treated round the clock in intensive care stroke units. However, as mentioned earlier, the ART should not be used as a diagnostic test, since the score of certain items can be affected by non-aphasic speech disorders (such as

speech apraxia or dysarthria). It was not used to screen patients who should be referred to a speech therapist or to decide speech therapy strategies, and should not, therefore, be used for these purposes without further studies.

The reliability sub-study showed high inter-rater reliability and the lack of a test–retest effect, with inter-rater variability being independent of the severity of aphasia. Furthermore, a change of  $\geq 3$  points (i.e., more than 2 points) corresponded to a significant change in the severity of aphasia, close to the change of  $\geq 4$  points (i.e., more than 3 points) required for a significant variation in the NIHSS score.

We found, as previously described [11], that the severity of the initial aphasia was correlated with the neurological deficit as measured by the NIHSS, but not with age or gender. The severity of aphasia markedly improved during the first week, with the median ART value decreasing from 24 to 10 between D0 and D8. The extent of this improvement may appear surprising. However, this is perhaps explained by the items of the ART, which score very basic language functions. Furthermore, this early recovery phase only occurred in two-thirds of the patients, and was more pronounced in the “good recovery” subgroup of patients, which had also the lower D0 ART scores. In other patients, the ART score remained stable or even increased. These findings are also consistent with the few published reports on early post-stroke changes in the severity of aphasia. In a series of 41 first-ever stroke patients with aphasia tested on naming, reading and repetition tasks at 24–48 h and seven days, some degree of improvement in overall performance was found in 61 % of the patients [2]. In three other reports based on the 3-point aphasia sub-score of the NIHSS, which may have more limited sensitivity, an early

improvement was found in 36–57 % of the patients [5, 7, 10].

The severity of initial aphasia is considered the best clinical predictor of the outcome of language function [11, 12, 14]. The prognosis value of the D8 ART score is consistent with other studies, where initial aphasia is often assessed around one week after stroke [6, 11–13]. However, several new findings emerge from our study. First, it is not only the ART score per se but also the recovery of basic language functions during the first week post-stroke (i.e., the difference in the score of some items tested by the ART between D0 and D8) that appears to be an important predictor of later language outcome. A comparison of ART values in good, intermediate, and poor outcome groups shows not only significant group and time effects but also a significant group-time interaction, indicating that early recovery differs across groups. This difference is confirmed by a comparison of  $\Delta$ ART across the three groups. This also explains why the D8 ART score, which integrates D0 ART and  $\Delta$ ART, is a better predictor of language outcome than the D0 ART score, as shown by logistic regression and ROC curve analyses. Second, the ROC curve-based predictions concerning good (AHS 0–2) and poor (AHS 4–5) outcomes were surprisingly accurate, with an AUC greater than 0.9 and accuracy greater than 85 % in each case. Indeed the cut-off values (D8 ART <12 for good and >21 for poor outcome) yielded a sensitivity of >90 % and a specificity of >80 %, better than results obtained using more complex language tests [6], and similar to those recently reported by a sophisticated model combining language tests and functional MRI results [13]. In addition, as shown in Fig. 4, most patients with intermediate D8 ART scores had an intermediate outcome (AHS 3). The good predictive value of D8 ART was also observed when good and poor outcome groups were defined using the BDAE aphasia severity rating scale. Age and gender were not predictors of outcome in this study, in agreement with a recent review that concluded that gender and age did not significantly impact recovery patterns in post-stroke aphasia [12]. It should be noted that the lack of an impact of rt-PA treatment on recovery cannot be interpreted as a lack of an effect of rt-PA on aphasia outcome, since all patients eligible for rt-PA were thus treated, instead of being randomized into treatment and non-treatment groups. In addition, dramatic recovery of aphasia may have occurred in some rt-PA treated patients before the D0 ART scoring, which was done after thrombolysis.

The study also has certain obvious limitations. First, as already stated, the ART should not be used as a diagnostic test for aphasia, since some of the scored items may also be affected by dysarthria, speech apraxia, buccofacial apraxia, ideomotor apraxia, executive dysfunction, or attentional fluctuations. Second, the 3-month language outcome was

based on the AHS, an unpublished verbal communication handicap score adapted from the modified Rankin score. However, more than half of the patients had a BDAE 3 months post-stroke, and the ART appeared to be in good agreement with the BDAE aphasia severity rating scale. Third, since the ART was developed in French, although it was designed to be easy to translate, it needs to be tested and validated in other languages by independent studies. Finally, the ART has not been directly compared with comprehensive aphasia rating scales, and does not allow us to classify patients in classic aphasic syndromes. Considering that our subjects are consecutive acute stroke patients, such an analysis may not be feasible, and in any case, the use of ART scoring in an acute stroke unit cannot and should not replace comprehensive language assessment in stabilized patients in speech therapy departments.

In summary, the ART appears to be a simple, rapid, and reproducible language-focused stroke scale to quantify the severity of initial aphasia and to monitor early changes in acute stroke patients. It is an accurate predictor of verbal communication outcome at three months. This may be of importance for patient stratification in future trials testing the effect of early therapeutic intervention after stroke on aphasia recovery. In addition, since the only language-specific items on the test are three words and one sentence to be repeated, it should be easy to adapt to other languages.

**Conflicts of interest** On behalf of all authors, the corresponding author states that there is no conflicts of interest.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## References

1. Chomel-Guillaume S, Laloup G, Bernard I (2010) Les aphasies evaluation et reeducation. Elsevier Masson, Issy les moulineaux
2. Cloutman L, Newhart M, Davis C, Heidler-Gary J, Hillis AE (2009) Acute recovery of oral word production following stroke: patterns of performance as predictors of recovery. *Behav Neurol* 21:145–153
3. Enderby P, Crow E (1996) Frenchay aphasia screening test: validity and comparability. *Disabil Rehabil* 18:238–240
4. Flamand-Roze C, Falissard B, Roze E, Maintigneux L, Beziz J, Chacon A, Join-Lambert C, Adams D, Denier C (2011) Validation of a new language screening tool for patients with acute stroke: the language screening test (LAST). *Stroke* 42:1224–1229
5. Inatomi Y, Yonehara T, Omiya S, Hashimoto Y, Hirano T, Uchino M (2008) Aphasia during the acute phase in ischemic stroke. *Cerebrovasc Dis* 25:316–323
6. Laska AC, Bartfai A, Hellblom A, Murray V, Kahan T (2007) Clinical and prognostic properties of standardized and functional aphasia assessments. *J Rehabil Med* 39:387–392

7. Maas MB, Lev MH, Ay H, Singhal AB, Greer DM, Smith WS, Harris GJ, Halpern EF, Koroshetz WJ, Furie KL (2012) The prognosis for aphasia in stroke. *J Stroke Cerebrovasc Dis* 21:350–357
8. Mazaux JM, Orgogozo JM (1982) Echelle d'évaluation de l'aphasie. Editions Scientifiques et Psychologiques, Issy les Moulineaux
9. Nakase-Thompson R, Manning E, Sherer M, Yablon SA, Gontkovsky SL, Vickery C (2005) Brief assessment of severe language impairments: initial validation of the Mississippi aphasia screening test. *Brain Inj* 19:685–691
10. Payabvash S, Kamalian S, Fung S, Wang Y, Passanese J, Kamalian S, Souza LC, Kemmling A, Harris GJ, Halpern EF, Gonzalez RG, Furie KL, Lev MH (2010) Predicting language improvement in acute stroke patients presenting with aphasia: a multivariate logistic model using location-weighted atlas-based analysis of admission CT perfusion scans. *AJNR Am J Neuroradiol* 31:1661–1668
11. Pedersen PM, Vinter K, Olsen TS (2004) Aphasia after stroke: type, severity and prognosis. The Copenhagen aphasia study. *Cerebrovasc Dis* 17:35–43
12. Plowman E, Hentz B, Ellis C (2012) Post-stroke aphasia prognosis: a review of patient-related and stroke-related factors. *J Eval Clin Pract* 18:689–694
13. Saur D, Ronneberger O, Kummerer D, Mader I, Weiller C, Kloppel S (2010) Early functional magnetic resonance imaging activations predict language outcome after stroke. *Brain* 133:1252–1264
14. Wade DT, Hewer RL, David RM, Enderby PM (1986) Aphasia after stroke: natural history and associated deficits. *J Neurol Neurosurg Psychiatry* 49:11–16