

Guidelines on applying APROCSA in a research context

1. Identify one or more raters with clinical and/or research experience with individuals with aphasia. All of the raters in our study were experienced with aphasia and confident in their knowledge of aphasia (Table 4), so we would not advise using raters who are not similarly experienced and confident. We recommend using at least three raters to improve reliability.
2. Join the AphasiaBank consortium (MacWhinney et al., 2011), by following the instructions on the AphasiaBank home page (<http://aphasia.talkbank.org>). Special attention should be paid to the “ground rules” (<http://talkbank.org/share/rules.html>). You may wish to consult your local institutional review board regarding the use of the human subjects data available through AphasiaBank.
3. Ask each rater to review Tables 1 and 2 of our paper, and ‘The APROCSA system’ subsection of the Methods. You may wish to implement training sessions to ensure that all raters have similar understandings of how the features are defined. Then ask raters to rate some or all of the AphasiaBank samples that we rated. On our website (<http://www.aphasialab.org/aprocsa>), we have provided a MATLAB script that takes as input feature ratings for any of the 24 patients we rated, or our training patient, and outputs a comparison between the input ratings and our experienced researcher ratings, including an ICC (A-1) measure of reliability. A training scheme might involve rating as many samples as are needed to bring reliability to an acceptable value.

4. Have each rater rate the connected speech of the individuals with aphasia in your study. If you have multiple raters, then scores should be averaged across the raters, and inter-rater reliability can be calculated and reported.

5. Carry out univariate analyses of features of interest, or multivariate analyses across the set of features. If the patient cohort under investigation is comprised of individuals with chronic post-stroke aphasia, factor scores for each patient can be obtained based on our factor analysis using our MATLAB script. If the patient cohort under investigation is different, a new factor analysis should be performed on the data obtained, because the generalization of our factors to other patient cohorts cannot be assumed.