Title: **ASSIST: Child Apraxia Speech Treatment**

National Clinical Trial (NCT) Identified Number: **NCT03903120**

Date: **12/04/2018**

Statistical Design and Power

Power

As a Phase I trial of a novel treatment, this study is not powered to detect a definitive effect. Instead, realistic sample sizes for a summer camp at a single site will be used to estimate effect sizes for subsequent Phase 2 and Phase 3 trials and determine the potential need for a multisite trial. Nevertheless, for Aim 1, with sample sizes of 20 children per study, $\alpha = 0.05$ (two-tailed), we expect to have at least 80% power to detect an ASSIST treatment effect size (mean difference between groups over the study period, divided by the pooled standard deviation) of 0.909 using a two-sample t-test. The only prior randomized controlled trial for CAS,[28] with 26 children enrolled, reported effect sizes of 1.312 and 2.162 with only 12 hours of treatment (two different treatments), suggesting that the present study will have sufficient power to detect effects. For Aim 2, with sample sizes of 10 children per study, $\alpha = 0.05$ (two-tailed), we expect to have at least 80% power to detect condition/intensity treatment effect size of 1.325 using a two-sample t-test. No power estimates were computed for secondary outcome measures.

Statistical Design and Analysis Plan

Descriptive statistics (e.g., mean, median, standard deviation, skewness, and frequencies) will be computed for all variables to ensure data quality and evaluate assumptions of statistical tests. We will compare the pre-treatment performance between the immediate treatment groups and the control groups of each study to assess the success of randomization in producing two comparable groups. Although we do not expect any imbalance between the immediate treatment and control groups, we will adjust for any differences in our subsequent analyses to be sure any treatment effect is not due to potential confounds. For all analyses described in this study, we will use Statistical Analysis Software (SAS, v9.4, Cary, N.C.), with two-sided tests under presumed significant level of 0.05.

To address *Specific Aim 1*, independent two-sample t-tests will be used to compare gain scores (T1 to T2) from each proximal outcome measure, i.e., *speech accuracy of treated items* and *speech accuracy of generalization items*, between immediate treatment groups and the control groups with combined Studies 1 and 2 (***Figure 3***). A separate linear model will be fitted by adjusting for potential effect of study (Study 1/Study 2) to re-assess the initial efficacy of ASSIST.

To address *Specific Aims 2a and 2b,* independent two-sample t-tests will be used to compare gain scores from each proximal outcome measure, i.e., *speech accuracy of treated items* and *speech accuracy of generalization items*, between conditions (i.e., simple/complex, word/nonword). Gain scores will be based on T1 to T2 for Immediate groups and T2 to T3 for Delayed groups. Separate linear models will be fitted by adjusting for potential effect of ASSIST delivery time (Immediate/Delayed) to re-assess the parameters of *complexity* and *lexicality*.

To address *Specific Aim 2c,* independent two-sample t-tests will be used to compare gain scores (T1 to T3) from each proximal outcome measure, i.e., *speech accuracy of treated items* and *speech accuracy of generalization items*, between massed ASSIST group and the distributed ASSIST group. Because the Massed ASSIST group combines both Immediate Massed and Delayed Massed groups, no additional models will be fitted by adjusting potential effect of ASSIST delivery time (Immediate/Delayed) to re-assess the parameter of *intensity*.

To address *Specific Aims 3a and 3b,* the same analyses will be conducted as above for Studies 1-3, with the distal secondary outcome measures of parent ratings (communicative function, intelligibility) and objective intelligibility as dependent variables.