Accepted Manuscript

Benchmarks of significant change after aphasia rehabilitation

Natalie Gilmore, MS, CCC-SLP, Michaela Dwyer, Swathi Kiran, PhD, CCC-SLP

PII: S0003-9993(18)31190-0

DOI: 10.1016/j.apmr.2018.08.177

Reference: YAPMR 57361

To appear in: *ARCHIVES OF PHYSICAL MEDICINE AND REHABILITATION*

Received Date: 30 March 2018

Revised Date: 30 May 2018

Accepted Date: 7 August 2018

Please cite this article as: Gilmore N, Dwyer M, Kiran S, Benchmarks of significant change after aphasia rehabilitation, *ARCHIVES OF PHYSICAL MEDICINE AND REHABILITATION* (2018), doi: https://doi.org/10.1016/j.apmr.2018.08.177.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Benchmarks of significant change after aphasia rehabilitation

Natalie Gilmore[1], MS, CCC-SLP*; Michaela Dwyer[1]; Swathi Kiran[1], PhD, CCC-SLP

[1]Boston University
Sargent College of Health and Rehabilitation Sciences
Speech, Language, and Hearing Sciences
Aphasia Research Laboratory, Room 326
635 Commonwealth Avenue, Boston, MA, 02215
*Corresponding author: Phone: 617-353-2706 Fax: 617-353-5074 Email: ngilmore@bu.edu

The manuscript has not been submitted elsewhere nor published elsewhere, except as an abstract: Gilmore, N., Dwyer, M., & Kiran, S. (2017). A systematic review of standardized outcome measures used in aphasia rehabilitation, including meta-analyses. *Archives of Physical Medicine and Rehabilitation*, *98*(10), e137. This abstract's content was presented in poster format on 10/26/2017 at American Congress of Rehabilitation Medicine (ACRM) 2017 Annual Conference in Atlanta, Georgia.

This work was supported by NIH/NIDCD grant T32DC0130170. None of the authors of this work has a financial conflict of interest with respect to this project.

Running Head: Change after post-stroke rehabilitation
Word Count for Manuscript: 2964
Word Count for Abstract: 265
Table:2
Figures: 4
Search terms: stroke; rehabilitation; outcome; speech therapy; aphasia

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Benchmarks of significant change after aphasia rehabilitation

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

ABBREVIATIONS

ASHA: American Speech-Language Hearing Association

BNT: Boston Naming Test

CI: Confidence interval

CETI: Communicative Effectiveness Index

COS: Core Outcome Set

ES: effect size

PRISMA: Preferred Reporting Items for Systematic Review and Meta-analyses

QOL: Quality of Life

SEM: Standard Error of Measurement

TPO: time post onset

WAB-AQ: Western Aphasia Battery-Aphasia Quotient

ABSTRACT

ACCEPTED MANUSCRIPT

Objective: To establish benchmarks of significant change for aphasia rehabilitation outcome measures (i.e., Western Aphasia Battery-Aphasia Quotient [WAB-AQ], Communicative Effectiveness Index [CETI], Boston Naming Test [BNT]) and assess if those benchmarks significantly differed across subgroups (i.e., time post onset, dose frequency, treatment type).

Data Sources: A comprehensive literature search of 12 databases, reference lists of previous reviews, and evidence-based practice materials was conducted.

Study Selection: Randomized-controlled trials, quasi-experimental studies, single-subject design, and case studies that used a standardized outcome measure to assess change were included. Titles and full-text articles were screened using a dual review process. 78 studies met criteria for inclusion.

Data Extraction: Data were extracted independently and 25% of extractions were checked for reliability. All included studies were assigned quality indicator ratings and an evidence level.

Data Synthesis: Random-effects meta-analyses were conducted separately for each study design group (i.e., within/between group comparisons). For within group designs, the summary effect size after aphasia rehabilitation was 5.03 points (95% confidence interval: 3.95-6.10, $p < .001$) on the WAB-AQ, 10.37 points (6.08-14.66, $p < .001$) on the CETI and 3.30 points (2.43-4.18, $p < .001$) on the BNT. For between group designs, the summary effect size was 5.05 points (1.64-8.46, $p = .004$) on the WAB-AQ, and .55 points (-1.33, 2.43, $p = .564$) on the BNT, the latter of which was not significant. Subgroup analyses for the within group designs showed no significant differences in the summary effect size as a function of dose frequency, or treatment type.

Conclusions: This study established benchmarks of significant change on three standardized outcome measures used in aphasia rehabilitation.

Key Words: stroke; rehabilitation; outcome; speech therapy; aphasia

Thirty to forty percent of stroke survivors experience aphasia.[1] While numerous systematic reviews and meta-analyses have demonstrated aphasia rehabilitation efficacy,[2,3] none have provided the average significant change, or summary effect size (ES) by outcome measure, a valuable metric for practitioners and researchers. Robey's hallmark meta-analyses[2,4,5] showed a positive aphasia treatment effect, but were segregated by study design and focused on identifying the effect size for different conditions (e.g., treated vs untreated recovery). Similarly, the most recent Cochrane review[3] demonstrating speech therapy efficacy, synthesized data from randomized controlled trials only, excluding a wealth of aphasia treatment data. Furthermore, effect sizes were represented as standardized mean differences for specific behaviors (e.g., verbal expression), not for specific outcome measures (e.g., Western Aphasia Battery-Aphasia Quotient[6] [WAB-AQ]).

Another option is to synthesize results by outcome measure to obtain a summary ES (i.e., raw unstandardized mean difference),[7] which can be used to interpret meaningful change on a specific assessment post-treatment. Clinicians and researchers frequently utilize standard error of measurement (SEM) to interpret a test score's meaningfulness after intervention. However, summary ES is a more appropriate metric. It reflects the treatment effect's size[7] and can be used to interpret group data, as opposed to SEM, which is more relevant for interpreting individual scores.[8]

Numerous aphasia assessment instruments exist[9] for assessing impairment (i.e., Body Structure/Function), functional communication (i.e., Activity/Participation), psychosocial functioning (i.e., Contextual Factors) and well-being (i.e., Quality of Life [QOL]). It is not surprising then that practicing speech-language pathologists[10–12] and researchers[13,14] use measures inconsistently making synthesis and comparison across trials challenging.

ACCEPTED MANUSCRIPT

Wallace and colleagues proposed a core outcome set (COS)[13,15–18] for aphasia, specifying a minimum set of outcomes that should be administered to persons with aphasia as standard practice (i.e., WAB, The Scenario Test, General Health Questionnaire-12, SAQOL-39g) to increase consistency. Yet, the summary ES for these measures remains unknown. Given the potential benefits to clinical and research practice, a systematic review of behavioral aphasia intervention studies with meta-analyses was conducted with two aims: 1) To calculate the summary ES reported on the most frequently-used and relevant outcome measures; and 2) To determine if the summary ES significantly differed across subgroups for each outcome measure (i.e., time post onset, dose frequency, treatment type).

METHODS

This study followed the Preferred Reporting Items for Systematic Review and Meta-analyses: the PRISMA Statement[19] guidelines and was registered at the International prospective register of systematic reviews, PROSPERO, under the identification number CRD42016039393.

Inclusionary Criteria

Randomized-controlled, quasi-experimental, single-subject design, and case studies with an $n \geq 3$ were included if they (1) assessed the effect of a behavioral aphasia intervention and (2) used a standardized outcome measure to evaluate change post-treatment as compared to pre-treatment (i.e., data from two time points).

Literature search

The following databases: PubMed, EMBASE, CINAHL, PsycINFO, SpeechBite, LLBA, PLoS, Worldcat, Web of Science, Ageline, Scopus, and Google Scholar were searched (see Supplementary Material 1 for sample search strategy) from 5/24/2016-08/26/2016. Reference lists of relevant systematic reviews, meta-analyses and professional organization materials were

reviewed. Search terms were modified to meet each database's requirements. Grey literature was removed during screening. All citations were managed using Zotero[20] and exported to Excel for screening and data extraction.

Study Selection and Data Extraction

Two reviewers (first two authors) independently screened 9,285 titles and abstracts against inclusionary criteria (96% inter-reviewer reliability). Full-text articles were obtained for records that met all criteria. Both reviewers screened 858 full-text articles against the inclusionary criteria (90% inter-reviewer reliability). Disagreements were resolved through discussion and searching the full-text. Study exclusion rationale was documented (Figure 1). When results from the same dataset were included in multiple publications, only the publication with the greatest sample size was included. Both reviewers extracted the following data from the full-text: the standardized outcome measure used to measure intervention-related change, presence/absence of data from two time points, study design, sample size, testing time points, and population treated (i.e., stroke survivors and/or caregivers).

The number of studies using each standardized outcome measure was calculated. Based on the measure's use frequency (Supplementary Material 2), field relevance (i.e., part of aphasia COS), and disability domain [21,22] measured (i.e., Body Structure/Function, Activity/Participation, Contextual Factors and/or QOL), the WAB-AQ, the Communicative Effectiveness Index[23] (CETI) and the Boston Naming Test[24] (BNT) were chosen for meta-analysis. To have a power of .80 to detect an effect size of $\geq$.50 using a random-effects model, outcome measures with cumulative sample sizes across within group studies < than 100 were excluded and/or if the measure was used in less than < 10 studies.[25] The contextual factor and QOL COS measures were excluded from meta-analysis because 1) the 12-item General Health Questionnaire was

only used in 1 study and 2) sensitivity to change had already been established[26,27] for the Stroke and Aphasia Quality of Life Scale-39. 78 studies met eligibility for meta-analysis. Both reviewers extracted the following data from these studies: age, sex, aphasia type and severity, time post onset, treatment type and description, session length, weekly session frequency, testing time points, treatment length, pre- and post-treatment test score correlation, and pre- and post-treatment mean (SD) on the WAB-AQ, CETI and/or BNT.

Studies were classified as including an acute (i.e., < 6 months post stroke onset) or chronic sample; providing a lower dose frequency (i.e., ≤ 4 hours/week) or a higher dose frequency; and utilizing an impairment-based (i.e., treated discrete deficits), activity/participation-based (i.e., targeted everyday communication) and/or integrated (i.e. combined impairment and activity/participation level approaches) treatment. According to Warren, Fey and Yoder, 2007,[28] dose frequency is the number of times an intervention was provided daily and weekly.

The same two reviewers responsible for screening divided the data extraction. Each reviewer extracted data for 25% of the others' studies (98% inter-reviewer reliability). Reviewers contacted original authors for additional data needed to calculate effect sizes as needed.

Quality Assessment

The same two reviewers independently appraised included studies' quality using indicators identified by the American Speech-Language Hearing Association (ASHA) level of evidence scheme.[29,30] See Supplementary Material 3 for quality indicator details. Quality indicator summative scores ≤ 1 for within group studies [Post-treatment Mean vs. Pre-treatment mean for the same group] and ≤ 2 for between group studies [Experimental group Post-treatment Change vs. Control group Post-treatment Change] were excluded for poor quality. Reviewers

ACCEPTED MANUSCRIPT

assigned each study's evidence level using ASHA[31] guidelines originally proposed by the Scottish Intercollegiate Guidelines Network[32] (i.e., IB: randomized controlled study; IIA: non-randomized controlled study; IIB: quasi-experimental study; III: non-experimental studies).

Data Analysis

Individual patient results from studies with sample sizes ≥ three were averaged to calculate a group mean and SD. Pre-post treatment correlation scores were calculated for studies providing individual subject data as follows: Pre-treatment SD + Post-treatment SD – Change SD/ 2 * Pre-treatment SD * Post-treatment SD.[33] When it could not be computed, the average of the observed pre-post treatment correlation coefficients was used.[34] For crossover designs, data were extracted after both treatment phases, as long as both involved the same treatment type (i.e., impairment, activity/participation and/or integrated). For the WAB-AQ within group analysis, a weighted mean and SD was calculated for the Cherney, 2010 study as the published results were split by severity and for the Mozeiko et al., 2016 study, data for the higher dose frequency and lower dose frequency groups were entered separately.

Meta-analyses were conducted independently for within and between group study designs to avoid methodological concerns involved in transforming to a common metric.[35] After group averages were calculated for both time points, single-subject design and case study data were included in the within group meta-analyses.

Meta-analyses for each outcome measure for both study designs were performed using Comprehensive Meta-Analysis software.[36] As heterogeneity between studies was anticipated, a random-effects model was used to combine individual study results into a summary ES (i.e., raw unstandardized mean difference). Raw unstandardized mean difference was calculated because clinicians and researchers interpret raw change on these outcome measures post-intervention,

making this effect size inherently meaningful to the field.[7] Q and I^2 statistics were examined to determine the extent of any remaining heterogeneity across studies. Even if the heterogeneity was low (i.e., non-significant and < 75%), subgroup analyses were conducted to assess summary ES differences depending on recovery stage, treatment type, and dose frequency. Sub-group analyses were corrected for multiple comparisons using the Bonferroni correction method.

Subgroup Analyses

Although no significant heterogeneity was present in the overall summary ESs, subgroup analyses were performed to investigate for summary ES differences due to these variables. As > 5 studies per subgroup are required to conduct a valid subgroup analysis,[7] the same subgroup analyses were not feasible for all outcome measures and study design groups. Subgroup analyses were conducted with the following variables, outcome measures, and study designs: 1) dose frequency for within group studies using the WAB-AQ, CETI, and BNT and 2) treatment type for within group studies using the WAB-AQ and BNT. No subgroup analyses were conducted to assess for differences in summary ES related to TPO as the nearly all of the within group studies included participants in the chronic phase. No subgroup analysis was conducted to assess for a difference in summary ES according to treatment type for within group studies using the CETI, or any of the between group study designs as there were < 5 studies in each subgroup.

Funnel plots for meta-analyses including > 10 studies were examined for asymmetry (i.e., within group meta-analyses only). Publication bias was objectively assessed using Begg and Mazumdar rank correlation, Egger's regression intercept and Duval and Tweedie's Trim and Fill.[7]

RESULTS

ACCEPTED MANUSCRIPT

Aim 1: What is the summary ES post-therapy on three commonly-used outcome measures in aphasia rehabilitation?

Study Identification/Description. 78 studies met criteria for inclusion in the meta-analyses (i.e., within group: 70; between group: 8). Descriptive information and references for these studies can be found in Supplementary Materials 4 through 9.

Within group study designs. Combining individual studies' findings resulted in a significant summary ES indicating a positive treatment effect across all three outcome measures. On the WAB-AQ (53 studies, n = 522), the summary ES on the raw unstandardized mean difference was 5.03 points, (95% confidence interval [CI]: 3.95-6.10, $p < .001$). No significant heterogeneity was found (Q = 50.79, df = 52, p = .52; I^2 = 0). The CETI summary ES (17 studies, n = 208), was 10.37 points (6.08-14.66, $p < .001$). No significant heterogeneity was found (Q = 16.47, df = 16, p = .42; I^2 = 2.86). The summary ES for the BNT (36 studies, n = 347), was 3.30 points (2.43-4.18, $p < .001$). No significant heterogeneity was found (Q = 42.17; df = 35; p =.19; I^2 = 17.01). See Figures 2 and 3 for forest plots depicting the variability across studies.

Publication bias for within group meta-analyses. No marked asymmetry was noted in funnel plots for any of these meta-analyses (Supplementary Materials 10). For the WAB-AQ, both the Egger's regression intercept (ß = 1.31, CI = (-.11, 2.72), t (51) =1.86, p = .04) and the Duval and Tweedie's Trim and Fill (Observed point estimate = 5.03(3.95, 6.10); Imputed point estimate = 5.88 (4.74, 7.02)) suggested the presence of publication bias for the WAB-AQ (i.e., missing positive studies). There was no significant presence of publication bias for the CETI meta-analysis (1-tailed p > .05). For the BNT, the Duval and Tweedie's Trim and Fill revealed the presence of publication bias (Observed point estimate = 3.30(2.43, 4.18); Imputed point

ACCEPTED MANUSCRIPT

estimate = 2.97(2.02, 3.92)) (i.e., missing negative studies). In both cases where publication bias, was indicated, the SES shifted only minimally (i.e., < 1 point, within the confidence interval), verifying that the within group SESs reported for all three outcome measures are valid and can be utilized with confidence.

Between group study designs. On the WAB-AQ (6 studies, Experimental n = 119; Control n = 99), the summary ES on the raw unstandardized mean difference between the experimental and control groups was 5.05 (1.64-8.46, $p < .01$). No significant heterogeneity was found (Q = 5.26, df = 5, p =.39; I^2 = 4.87). No between-group meta-analysis was conducted for the CETI as only one publication using it to measure post-intervention change was identified. On the BNT (5 studies, Experimental n = 66; Control n = 35), the raw unstandardized mean difference between the experimental and control groups at post-treatment was .55 (-1.33-2.43, p = .56). There was no significant heterogeneity between included studies (Q = .86, df = 4, p = .93; I^2 = 0). See Figure 4 for forest plots that illustrates the variability across studies.

Publication bias for between group meta-analyses. Due to the low sample size in the between group study design meta-analyses,[37] funnel plots could not be validly assessed for the presence of publication bias.

Aim 2: Does the summary ES vary according to time post onset, dose frequency and/or treatment type?

There were no statistically significant differences between summary ESs for any of the within group study design subgroup analyses completed (i.e., dose frequency for WAB-AQ, CETI, and BNT; treatment type for WAB-AQ and BNT). See Table 1 for results and Supplementary Materials 11 for forest plots.

Quality Appraisal

ACCEPTED MANUSCRIPT

For within group study designs, 73% of studies included in the meta-analyses were level III evidence,[29,31] 26% were IIB, and 1% were IIA. For between group study designs, 50% were classified as IB, 38% as IIA, and 13% as IIB level evidence. None of the 78 studies selected for meta-analysis were excluded from the analysis based on their quality, which is unsurprising as studies of poorer quality were likely excluded during the two initial screening phases. See Table 2 for summative quality indicator scores for both study designs. For within group studies, most studies had summative scores of 3, with higher scores indicating better quality. For between groups comparisons, the majority of studies using the WAB or BNT had summative scores of 7 or 5, respectively. Individual study ratings are included in Supplementary Materials 4-8. The percentage of studies meeting criterion for each specific quality indicator are available in Supplementary Material 12.

DISCUSSION

This study established benchmarks for significant change on three outcome measures used in aphasia rehabilitation to assess severity, functional communication, and naming ability. Practitioners can use these metrics to objectively demonstrate improvement in their clients following treatment, an essential element of clinical practice that directly influences reimbursement and clients' duration of services. Likewise, researchers can reference the reported summary ESs when quantifying change from experimental interventions, but also when conducting *a priori* power analyses for future studies. The latter analyses require estimating the effect size,[38] which is not consistently reported in published aphasia treatment studies,[39] further emphasizing the utility of this study's benchmarks.

The relationship between the summary ESs established in this study and each outcome measure's SEM must be discussed. WAB-AQ summary ESs (Within group: 5.03; Between

group: 5.05), were equivalent to its SEM of 5, which has been framed as a metric of clinically meaningful improvement.[40–42] On initial inspection, the adjacency of these two values suggests a diminished effect of aphasia rehabilitation as measured by the WAB-AQ. However, the seminal work of Hula, Donovan, Kendall & Gonzalez-Rothi, 2010,[42] demonstrating that the WAB-AQ's SEM was actually closer to 2 for AQs between 28-68, but much higher (i.e., up to 12) for scores outside that range (i.e., AQs of 0-27, 69-100) serves to clearly distinguish the summary ES established in this study from measurement error. Future research should examine how the WAB-AQ summary ES varies for persons with more mild or severe aphasia and examine which treatment approaches result in summary ESs well outside of the SEM for all severity groups. The CETI's summary ES of 10.37 was well above its SEM of 5.87,[23] suggesting that those improvements were not due to variations inherent to measurement alone. Lastly, the summary ES for the BNT of 3.30 was also higher than its SEM of 2.04,[43] supporting its validity as a metric of intervention-related improvement. Importantly, the summary ESs were consistent across treatment approaches and dose frequencies as none of the meta-analyses demonstrated significant heterogeneity, nor were any of the sub-group analyses significant.

This study provides a unique contribution to the literature on aphasia rehabilitation as it included studies according to the outcome measure used to assess change as opposed to by study design, as in previous systematic reviews and meta-analyses.[2,3] This methodological shift is valuable as rather than conducting only meta-analyses with between group comparisons, separate meta-analyses were also conducted using within group study comparisons, including single subject design studies. This approach allowed for the inclusion and synthesis of a larger body of the treatment literature in the field than previous reviews. In summary, this work adds to the body of literature that confirms a positive effect of aphasia treatment and further, provides

ACCEPTED MANUSCRIPT

benchmarks for significant change.

Nonetheless, some open questions remain. In order to maintain adequate power to conduct meta-analyses, a number of studies employing less-frequently used outcome measures were excluded (e.g., assessing contextual factors). Secondly, subgroup analyses could not be conducted between acute and chronic participant studies. Third, as the summary ES for the WAB-AQ was only notably higher than the SEM for a range of AQs (i.e., 28-68), it should be tested whether a higher benchmark for improvement should be used for individuals who are more mild or severe, or a different assessment measure altogether.

Study Limitations

All systematic reviews and meta-analyses are susceptible to publication bias. Although funnel plots for the within group designs were largely symmetric, publication bias was detected in the within-group WAB-AQ and BNT analyses. However, the point estimates varied minimally and thus, the observed summary ESs for those measures should be considered valid.

CONCLUSIONS

By combining evidence from existing treatment studies, the present systematic review and meta-analyses establishes valuable benchmarks of change for three frequently used outcome measures. Furthermore, it confirms that aphasia rehabilitation is indeed effective.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

REFERENCES

1. Dickey L, Kagan A, Lindsay MP, Fang J, Rowland A, Black S. Incidence and Profile of Inpatient Stroke-Induced Aphasia in Ontario, Canada. *Arch Phys Med Rehabil*. 2010;91(2):196-202. doi:10.1016/j.apmr.2009.09.020

2. Robey RR. A meta-analysis of clinical outcomes in the treatment of aphasia. *J Speech Lang Hear Res*. 1998;41(1):172-187.

3. Brady MC, Kelly H, Godwin J, Enderby P, Campbell P. Speech and language therapy for aphasia following stroke. In: The Cochrane Collaboration, ed. *Cochrane Database of Systematic Reviews*. Chichester, UK: John Wiley & Sons, Ltd; 2016. http://doi.wiley.com/10.1002/14651858.CD000425.pub4. Accessed July 22, 2016.

4. Robey R. The Efficacy of Treatment for Aphasia Persons: A Meta-analysis. *Brain Lang*. 1994;47:582-608.

5. Robey R, Schultz M, Crawford A, Sinner C. Review: Single-subject clinical-outcome research: designs, data, effect sizes, and analyses. *Aphasiology*. 1999;13(6):445-473.

6. Kertesz A. *Western Aphasia Battery (Revised)*. San Antonio, TX: PsychCorp; 2006.

7. Borenstein M, ed. *Introduction to Meta-Analysis*. Chichester, U.K: John Wiley & Sons; 2009.

8. Harvill LM. Standard error of measurement. *Educ Meas Issues Pract*. 1991;10(2):33-41.

ACCEPTED MANUSCRIPT

9. Wallace SJ, Worrall L, Rose T, Le Dorze G. Measuring outcomes in aphasia research: A review of current practice and an agenda for standardisation. *Aphasiology*. 2014;28(11):1364-1384. doi:10.1080/02687038.2014.930262

10. Hesketh A, Hopcutt B. Outcome measures for aphasia therapy: It's not what you do, it's the way that you measure it. *Eur J Disord Commun*. 1997;32(3, Spec Iss):189-202. doi:10.1080/13682829709177096

11. Worrall L, Egan J. A survey of outcome measures used by Australian speech pathologists. *Asia Pac J Speech Lang Hear*. 2001;6(3):149-162. doi:10.1179/136132801805576635

12. Simmons-Mackie N, Threats TT, Kagan A. Outcome assessment in aphasia: A survey. *J Commun Disord*. 2005;38(1):1-27. doi:10.1016/j.jcomdis.2004.03.007

13. Wallace SJ, Worrall L, Rose T, Le Dorze G. Measuring outcomes in aphasia research: A review of current practice and an agenda for standardisation. *Aphasiology*. 2014;28(11):1364-1384. doi:10.1080/02687038.2014.930262

14. Ali M, English C, Bernhardt J, Sunnerhagen KS, Brady M, VISTA-Rehab Collaboration. More outcomes than trials: a call for consistent data collection across stroke rehabilitation trials: Review. *Int J Stroke*. 2013;8(1):18-24. doi:10.1111/j.1747-4949.2012.00973.x

15. Wallace SJ, Worrall L, Rose T, et al. Which outcomes are most important to people with aphasia and their families? an international nominal group technique study framed within the ICF. *Disabil Rehabil*. June 2016:1-16. doi:10.1080/09638288.2016.1194899

ACCEPTED MANUSCRIPT

16. Wallace SJ, Worrall L, Rose T, Le Dorze G. Which treatment outcomes are most important to aphasia clinicians and managers? An international e-Delphi consensus study. *Aphasiology*. May 2016:1-31. doi:10.1080/02687038.2016.1186265

17. Wallace SJ, Worrall L, Rose T, Le Dorze G. Core Outcomes in Aphasia Treatment Research: An e-Delphi Consensus Study of International Aphasia Researchers. *Am J Speech Lang Pathol*. 2016;25(4S):S729. doi:10.1044/2016_AJSLP-15-0150

18. Wallace S, Worrall L, Rose T, Le Dorze G. Improving research outcome measurement in aphasia: Development of a core outcome set. Presented at the: International Aphasia Rehabilitation Conference (IARC); December 14, 2016; London.

19. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P). *Syst Rev*. 2015;4(1):1-9.

20. Center for History and New Media. Zotero Quick Start Guide. http://zotero.org/support/quick_start_guide.

21. Kostanjsek N. Use of The International Classification of Functioning, Disability and Health (ICF) as a conceptual framework and common language for disability statistics and health information systems. *BMC Public Health*. 2011;11(Suppl 4):S3. doi:10.1186/1471-2458-11-S4-S3

22. Kagan A, Simmons-Mackie N, Rowland A, et al. Counting what counts: A framework for capturing real-life outcomes of aphasia intervention. *Aphasiology*. 2008;22(3):258–280.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

23. Lomas J, Pickard L, Bester S, Elbard H, Finlayson A, Zoghaib C. The Communicative Effectiveness Index: Development and Psychometric Evaluation of a Functional Communication Measure for Adult Aphasia. *J Speech Hear Disord*. 1989;54(1):113-124. doi:10.1044/jshd.5401.113

24. Goodglass H, Kaplan E, Weintraub S. *Boston Naming Test*. Philadelphia: Lea & Febiger; 1983.

25. Liu J. Statistical Power in Meta-Analysis. 2015. https://scholarcommons.sc.edu/cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article=4230&context=etd.

26. Hilari K, Lamping DL, Smith SC, Northcott S, Lamb A, Marshall J. Psychometric properties of the Stroke and Aphasia Quality of Life Scale (SAQOL-39) in a generic stroke population. *Clin Rehabil*. 2009;23(6):544-557. doi:10.1177/0269215508101729

27. Guo YE, Togher LE, Power E, et al. Sensitivity to change and responsiveness of the Stroke and Aphasia Quality-of-Life Scale (SAQOL) in a Singapore stroke population. *Aphasiology*. 2017;31(4):427-446.

28. Warren SF, Fey ME, Yoder PJ. Differential treatment intensity research: A missing link to creating optimally effective communication interventions. *Ment Retard Dev Disabil Res Rev*. 2007;13(1):70-77. doi:10.1002/mrdd.20139

29. Faroqi-Shah Y, Frymark T, Mullen R, Wang B. Effect of treatment for bilingual individuals with aphasia: A systematic review of the evidence. *NEL J Neurolinguistics*. 2010;23(4):319-341.

30. Mullen R. The State of the Evidence: ASHA Develops Levels of Evidence for Communication Sciences and Disorders. *ASHA Lead*. 2007;12(March):8-25.

31. Steps in the Process of Evidence-Based Practice: Assessing the Evidence. American Speech-Language-Hearing Association. http://www.asha.org/Research/EBP/Assessing-the-Evidence/. Accessed October 4, 2017.

32. SIGN 50: A guideline developer's handbook - SIGN grading system 1999 – 2012. http://www.sign.ac.uk/guidelines/fulltext/50/annexoldb.html. Accessed October 4, 2017.

33. Higgins JP, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*. Vol 4. England: John Wiley & Sons Ltd.; 2011.

34. Laird KT, Tanner-Smith EE, Russell AC, Hollon SD, Walker LS. Short-term and Long-term Efficacy of Psychological Therapies for Irritable Bowel Syndrome: A Systematic Review and Meta-analysis. *Clin Gastroenterol Hepatol*. 2016;14(7):937-947.e4. doi:10.1016/j.cgh.2015.11.020

35. Morris SB, DeShon RP. Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychol Methods*. 2002;7(1):105-125. doi:10.1037//1082-989X.7.1.105

36. Borenstein M, Hedges L, Higgins J, Rothstein H. *Comprehensive Meta-Analysis*. Englewood, NJ: Biostat; 2014.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

412 37. Sterne JAC, Sutton AJ, Ioannidis JPA, et al. Recommendations for examining and
413 interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ*.
414 2011;343(jul22 1):d4002-d4002. doi:10.1136/bmj.d4002

415 38. Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice*. 3rd
416 Edition. Upper Saddle River, NJ: Prentice Hall Health; 2000. 2008.

417 39. Beeson PM, Robey RR. Evaluating Single-Subject Treatment Research: Lessons Learned
418 from the Aphasia Literature. *Neuropsychol Rev*. 2006;16(4):161-169. doi:10.1007/s11065-
419 006-9013-7

420 40. Elman RJ, Bernstein-Ellis E. The efficacy of group communication treatment in adults with
421 chronic aphasia. *J Speech Lang Hear Res*. 1999;42(2):411-419.

422 41. Katz RC, Wertz RT. The efficacy of computer-provided reading treatment for chronic
423 aphasic adults. *J Speech Hear Res*. 1997;40(3):493-507.

424 42. Hula W, Donovan NJ, Kendall DL, Gonzalez-Rothi LJ. Item response theory analysis of
425 the Western Aphasia Battery. *Aphasiology*. 2010;24(11):1326-1341.
426 doi:10.1080/02687030903422502

427 43. Flanagan JL, Jackson ST. Test-Retest Reliability of Three Aphasia Tests: Performance of
428 Non-Brain-Damaged Older Adults. *J Commun Disord*. 1997;30(1):33-43.

429

ACCEPTED MANUSCRIPT

FIGURE TITLES & LEGENDS

Figure 1. The PRISMA flow diagram[1] of study inclusion. *Note:* 1. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med*. 2009;6(7):6.

Figure 2. Summary effect sizes for within group studies reporting the Western Aphasia Battery-Aphasia Quotient (WAB-AQ). The difference in means column reflects the pre-treatment mean subtracted from the post-treatment mean. The lower and upper limits columns show the 95% confidence interval surrounding the difference in means. The p-value indicates the significance of the effect. The final row describes the summary effect size, 95% confidence interval, and p-value. The diamond represents the summary effect size. The squares reflect effect sizes of individual studies.

Figure 3. Summary effect sizes for within group studies reporting the Communicative Effectiveness Index (CETI) and Boston Naming Test (BNT). Figure details are the same as for Figure 2.

Figure 4. Summary effect sizes for between group studies reporting the Western Aphasia Battery-Aphasia Quotient (WAB-AQ) and Boston Naming Test (BNT). The diamond is the summary effect size. The squares reflect effect sizes of individual studies. The difference in means column reflects the post-treatment control group mean change subtracted from the post-treatment experimental group mean change. The lower and upper limits columns show the 95% confidence interval surrounding the difference in mean change. The p-value indicates the significance of the effect. The final row describes the summary effect size, 95% confidence interval, and p-value. The diamond represents the summary effect size. The squares reflect effect sizes of individual studies.

ACCEPTED MANUSCRIPT

Table 1. Results of subgroup analyses for within group study designs

Outcome Measure	LDF	HDF	IMP	A/P	INT
WAB-AQ	*n* = 35 4.50 3.64-5.36	*n* = 11 5.17 3.72-6.61	*n* =33 4.42 3.09-5.76	*n* = 6 5.10 1.73-8.47	*n* = 14 6.48 4.38-8.57
CETI	*n* = 10 10.05 3.83-16.28	*n* = 5 11.02 2.81-19.24	n/a	n/a	n/a
BNT	*n* = 25 3.55 2.33-4.76	*n* = 9 3.39 1.75-5.02	*n* = 24 3.18 2.09-4.27	*n* = 5 3.89 1.65-6.14	*n* = 7 3.34 1.18-5.49

Note: WAB-AQ=Western Aphasia Battery-Aphasia Quotient; CETI= Communicative Effectiveness Index; BNT= Boston Naming Test; LDF = lower dose frequency; HDF = higher dose frequency; IMP = impairment-based treatment; A/P = activity/participation-based treatment; INT= integrated treatment

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

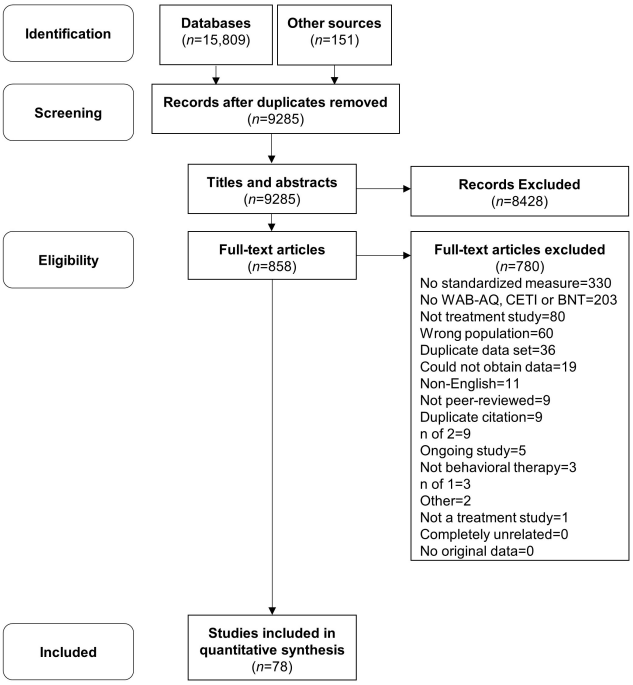
Table 2. Quality Indicator Summative Scores for Included Studies

Design	Test	N	7	6	5	4	3	2	1
Within Group	WAB	53	N/A	2	17	21	32	28	0
	CETI	17	N/A	12	24	35	67	18	0
	BNT	36	N/A	6	11	28	33	22	0
Between Group	WAB	6	50	33	17	0	0	0	0
	BNT	5	0	20	80	0	0	0	0

Note: Value in cell represents percentage of studies with that summative score. Within group studies could not obtain a rating of 7 because intention to treat is not a relevant parameter for that study design. Higher scores = higher methodological quality.

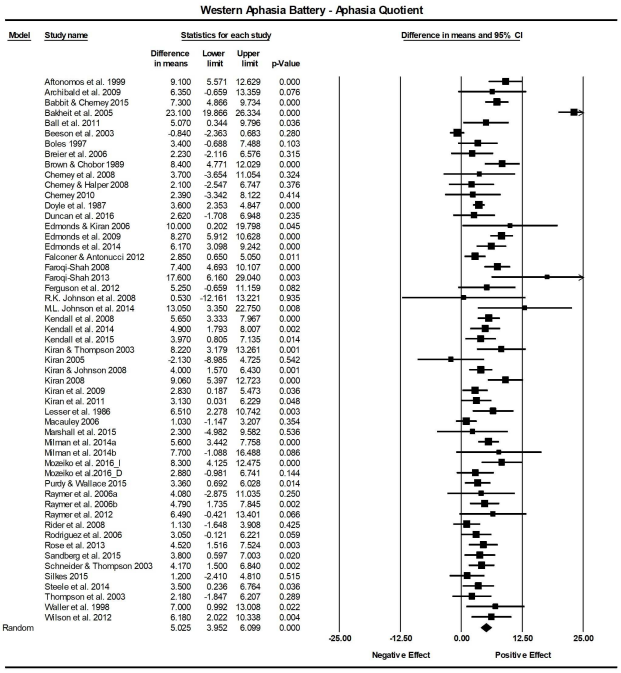
ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

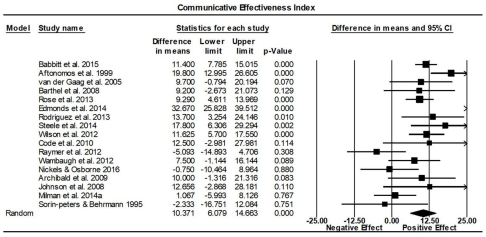


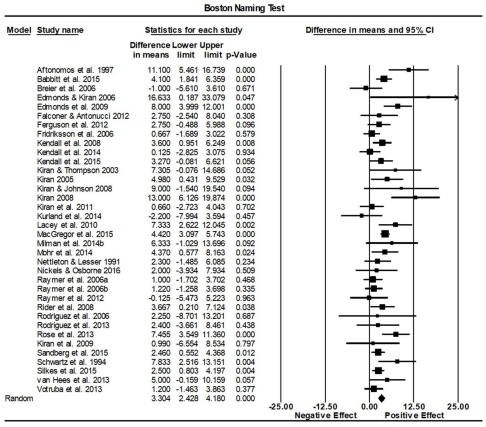
ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT



ACCEPTED MANUSCRIPT

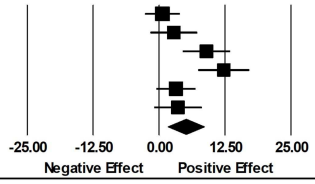




ACCEPTED MANUSCRIPT

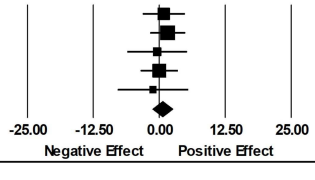
Western Aphasia Battery Aphasia Quotient

Model	Study name	Difference in means	Lower limit	Upper limit	p-Value
	Altmann et al. 2014	0.639	-2.692	3.969	0.707
	Des Roches et al. 2015	2.800	-1.678	7.278	0.220
	Godecke et al. 2012	8.980	4.449	13.511	0.000
	Godecke et al. 2014	12.270	7.394	17.146	0.000
	Katz & Wertz, 1997	3.200	-0.538	6.938	0.093
	Maher et al. 2006	3.565	-0.983	8.113	0.124
Random		5.047	1.638	8.456	0.004



Boston Naming Test

Model	Study name	Difference in means	Lower limit	Upper limit	p-Value
	Altmann et al. 2013	0.860	-3.156	4.876	0.675
	Des Roches et al. 2015	1.594	-1.750	4.938	0.350
	Maher et al. 2006	-0.400	-6.085	5.285	0.890
	Raglio et al. 2016	0.000	-3.568	3.568	1.000
	Wilssens et al. 2015	-1.200	-7.905	5.505	0.726
Random		0.554	-1.325	2.433	0.564



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Supplementary Material 1: Sample Search Strategy

In PubMed:

Line 1: aphasia Line 2: AND treatment OR therapy OR intervention OR rehabilitation OR outcome OR training Line 3: AND adult Line 4: NOT primary progressive aphasia OR dementia OR dysphagia OR transcranial magnetic stimulation OR transcranial direct current stimulation Line 5: NOT pharmaceutical preparations.

Article type was limited to Case Reports, Clinical Study, Clinical Trial, Clinical Trial, Phase I, Clinical Trial, Phase II, Clinical Trial, Phase III/CLASS IV, Clinical Trial, Phase IV, Comparative Study, Controlled Clinical Trial, Dataset, Meta-Analysis, Multicenter Study, Observational Study, Practice Guideline, Randomized Controlled Trial, Systematic Reviews, Validation Studies and Evaluation Studies. No other limits or filters were applied.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Supplementary Material 2: Frequency of outcome measure use

Outcome	Type	*n*
WAB-AQ	Study	80
	Subject	1276
BNT	Study	53
	Subject	673
CETI	Study	27
	Subject	458
CADL-2	Study	11
	Subject	89
Scenario Test	Study	1
	Subject	34
ACOM	Study	1
	Subject	73
SAQOL	Study	2
	Subject	34

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

SAQOL-39	Study	6
	Subject	87
SAQOL-39g	Study	1
	Subject	20
ALA	Study	2
	Subject	23
GHQ-12	Study	2
	Subject	14

Note: Indicates the outcome measure, the number of studies reporting the measure and the cumulative number of subjects reported for the measure. CADL-2= Communication Activities of Daily Living-Second Edition; ACOM= Aphasia Communication Outcome Measure; SAQOL= Stroke and Aphasia Quality of Life Scale; ALA= Assessment For Living With Aphasia; GHQ-12= 12-item General Health Questionnaire

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Supplementary Material 3. Quality indicators for assessing included studies

Indicator	Description
1. Study protocol	Adequate detail about the study protocol was given for the study to be replicated.
2. Blinding	Participants were blinded to condition. Assessors were blinded to condition/treatment.
3. Sampling/allocation	For example: random sample, convenience sample, not described, etc.
4. Treatment fidelity	Administrators established that the treatment protocol was delivered as planned.
5. Significance for primary outcome measure of interest (e.g., trained verbs)	Statistical analyses were conducted and p-values were reported.
6. Significance for standardized outcome measure of interest (i.e., WAB, CETI, BNT)	Statistical analyses were conducted and p-values were reported.
7. Precision	Effect size was reported or calculable.
8. Intention to treat	Data was analyzed based on the group to which the participants were originally assigned.

Note: Indicators: 1, 2, and 4-8 were scored as either (+) for present, or (-) for absent. Indicator 3 consisted of qualitative information. (+) values were tallied to create quality indicator summative scores. This table was adapted from Table S3. 1 in Faroqi-Shah, Y, Frymark, T, Mullen, R, & Wang, B. Effect of treatment for bilingual individuals with aphasia: A systematic review of the evidence. *Journal of Neurolinguistics*. 2010;23(4):319-341.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Supplementary Material 4: Demographic information for within group studies using the Western Aphasia Battery – Aphasia Quotient

Note: M=Male; F = Female; TCM = transcortical motor; TSM = transcortical sensory; sev. = severe; I = Impairment-based treatment; A/P = activity/participation-based treatment; INT = integrated treatment; st. = standardized; CS = convenience sample

Study Name	Study N; N for outcome measure; Sex	Mean Age(SD) range	Aphasia Type	Aphasia Severity (WAB-AQ)	Mean MPO range	Treatment(Tx)	Pre-Tx M(SD) Post-Tx M(SD) Change Score	Methodological Rigor
Aftonomos, Appelbaum, & Steele, 1999	60; 60; M = 35; F = 25	68.60(12.30) 24-86	Broca's = 21 Anomic = 13 Global = 11 Wernicke's = 8 Conduction = 3 TCM = 2 TSM = 1 Isolation = 1	Mod.-to-sev.	24.60 .24-144	**Type:** INT **Description:** Individual treatment using the Lingraphica (icon-based language system) to provide therapeutic exercises at the appropriate level for participants' severity. Also, focused on improving functional communication outside of the clinic as well as provided home exercises. **Intensity:** 2x/week, 60 min, 20.5 weeks	42.50(27.40) 51.60(28.70) 9.10	Level of Evidence: IIB/class III Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
Archibald, Orange, & Jamieson, 2009	8; 8; M = 6; F = 2	71.00(11.15) 55-87	Anomic = 4 Broca's = 2 Conduction = 1 Global =1	Mild = 3 Mild-to-Mod. = 2 Mod. = 1 Sev. = 2	48.38 7-150	**Type**: I **Description**: Computer-provided treatment via AphasiaMate across 8 modules (i.e., auditory comprehension, visual matching, reading comprehension, spelling, semantics, sentence processing). Patients used computer at home or in clinic with trained personnel. **Intensity**: 1x/week, 60 min, 15 weeks	60.29(33.37) 66.64(27.50) 6.35	Level of Evidence: IIB/class III Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Babbitt, Worrall, & Cherney, 2015	74; 74; M = 52; F = 22	54.10(16.30) 18-86	Nonfluent = 49 Fluent = 25	Mod.	15.5 3-87	**Type**: INT **Description**: Intensive Comprehensive Aphasia Program (ICAP): two individual therapy sessions and one session each of constraint-induced language therapy (CILT), reading/writing, computers and conversation group for six hours of daily programming. **Intensity**: 5x/week, 360 min, 4 weeks	51.30(21.80) 58.60(21.30) 7.30	Level of Evidence: IIB/ class III Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
Bakheit, Carrington, Griffiths, & Searle, 2005	67; 67; M = 31; F = 36	71.90(N/A) 38-92	Broca's = 21 Anomic = 18 Global = 15 Wernicke's = 9 Conduction = 3 TCM = 1	Mod.-to-sev.	12.72 n/a	**Type**: n/a **Description**: Individual "conventional" SLP sessions targeting comprehension and expression to improve functional communication. Tasks included selecting pictures/objects, naming objects, describing/recognizing associations between items, facilitating the expression of feelings and improving conversational ability. SLPs encouraged the use of gesture and other non-verbal communication including aids and equipment. **Intensity**: 2-5x/week, 40-60 min, 12 weeks	44.30(28.10) 67.40(25.50) 23.10	Level of Evidence: IIB/ class III Study protocol: + Blinding: + Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
Ball, de Riesthal, Breeding, & Mendoza, 2011	3; 3; M = 1; F = 2	70.67(3.21) 67-73	Global = 2 Conduction = 1	Mod. = 1 Sev. = 2	28.33 26-33	**Type**: I **Description**: Modified Anagram and Copy Treatment (ACT) and Copy and Recall Treatment (CART) (Beeson, Hirsch & Rewega, 2002 Beeson, Rising & Rolk, 2003) **Intensity**: 1x/week, 60 min, 12 weeks (daily home practice)	23.80(20.35) 28.87(17.83) 5.07	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Beeson, Rising, & Volk, 2003	8; 8; M = 5; F = 3	71.00(5.98) 64-79	Broca's = 7 Wernicke's = 1	Mod.-to-sev. = 1 Sev. = 7	39.75 24-84	**Type**: I **Description**: Copy and Recall Treatment (CART): 1) Show a picture 2) Have PWA write the word and support them in writing the word, if needed 3) Remove the word and show picture again and have them write three more times again. Stimuli (i.e., 20 words) was developed with family support to make it functionally relevant. **Intensity**: 1x/week, 60 min, 17 to 30 weeks	20.59(5.31) 19.75(4.81) -0.84	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: + Significance for st. outcome of interest: -

ACCEPTED MANUSCRIPT

								Precision: + Intention to treat: N/A
Boles, 1997	4; 4; M = 1; F = 3	56.00(15.38) 47-79	N/A	Mild = 3 Mod. = 1	28.75 7-84	**Type**: INT **Description**: Conversation partner therapy: Family member was coached by SLP to facilitate communication with PWA. **Intensity**: 2x/week, 60 min, 7 weeks	70.70(9.85) 74.10(8.62) 3.40	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Breier, Maher, Novak, & Papanicolaou, 2006	6; 6; M = 4; F = 2	61.33(8.80) 53-77	Broca's = 5 Conduction = 1	Mild-to-Mod. = 2 Mod. = 2 Mod.-to-sev. = 1 Sev. = 1	46.83 21-70	**Type**: A/P **Description**: Constraint Induced Language Therapy (CILT) = Only verbal expression was accepted and multi-modality communication was restricted, even self-cueing. Treatment was conducted in dyads and consisted of a dual card task with barrier present (i.e., PWA took turns requesting a card or responding another's request). Stimuli included four sets of cards of different semantic categories with two levels of difficulty (i.e., low- and high-frequency). Clinicians used shaping (i.e., increasing communicative demands of request/response from single words to lengthier sentences) and cueing for a successful production (i.e., semantic, phonemic, repetition). **Intensity**: 4x/week, 180 min, 3 weeks	52.22(21.99) 54.45(24.65) 2.23	Level of Evidence: IIB/ class III Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Brown & Chobor, 1989	10; 10; M = 7; F = 3	64.90(N/A) 55-76	Nonfluent = 9 Fluent = 1	Mild-to-Mod. = 1 Mod. = 2 Mod.-to-sev. = 3 Sev. = 5	77.99 36-120	**Type**: I **Description**: Writing treatment with right arm using a prosthesis which included four stages 1) geometric shapes 2) block letter alphabet 3) low- and high-frequency words and 4) two- and three-word short phrases. PWA went through three training phases: tracing, copying and writing to command within each of these stages. **Intensity**: 2x/week, 60 min, 12 weeks	36.40(19.75) 44.80(23.00) 8.40	Level of Evidence: IIB/ class III Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Cherney, Halper, Holland, & Cole, 2008	3; 3; M = 1; F = 2	69.33(7.51) 65-78	Broca's = 1 Wernicke's = 1 Anomic = 1	Mild-to-Mod. = 1 Mod. = 2	28.33 18-48	**Type**: A/P **Description**: AphasiaScripts software program was used for script training. 1) PWA listened to script while it is visible on the screen. 2) PWA reads the sentence twice chorally with avatar. PWA practice any words with which they had difficulty. 3)PWA reads each sentence aloud on their own. The computer records their response. 4) PWA can listen to the recorded sentence and then, practice and record again, if they want. PWA were trained sequentially on three scripts (i.e., three weeks each script). They practiced at home for 30 minutes daily. Clinician observed participants practicing once per week. **Intensity**: 5x/week, 30 min, 9 weeks	62.13(11.41) 65.83(7.96) 3.70	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Cherney & Halper, 2008	3; 3; M = 2; F = 1	64.00(12.77) 50-75	Nonfluent = 2 Fluent = 1	Mild-to-Mod. = 1 Mod. = 2	36 12-48	**Type**: I **Description**: AphasiaScripts software program was used for script training. 1) PWA listened to script while it is visible on the screen. 2) PWA reads the sentence twice chorally with avatar. PWA practice any words with which they had difficulty. 3)PWA reads each sentence aloud on their own. The computer records their response. 4) PWA can listen to the recorded sentence and then, practice and record again, if they want. PWA were trained sequentially on three scripts (i.e., three weeks each script). They practiced at home for 30 minutes daily. Clinician observed participants practicing once per week. **Intensity**: 5x/week, 30 min, 8 weeks	61.43(16.95) 63.53(13.51) 2.10	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Cherney, 2010	25; 25; M = 16; F = 9	55.38(9.49) 35-82	Nonfluent	Mod.	52.47 12-253	**Type**: I **Description**: Oral Reading for Language in Aphasia (ORLA): 1) PWA listened to the sentence twice while reading it on a card or on the computer and pointing to each word in the sentence 2) PWA read the sentence aloud with the SLP twice 3) PWA identified two or three words randomly and read them aloud 4) PWA and SLP read the whole sentence again together. Thirty different stimulus items of a certain length (i.e., 3-5 words, 8-12 words and 15-30 words) according to their severity level were practiced within the session. **Intensity**: 2-3x/week, 60 min, 8-12 weeks	54.59(29.68) 56.98(29.37) 2.38	Level of Evidence: IIB/class III Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Doyle, Goldstein, & Bourgeois, 1987	4; 4; M = 3; F = 1	55.75(9.32) 42-62	Broca's = 4	Mild-to-Mod. = 2 Mod. = 2	117.5 30-177	**Type**: I **Description**: Treatment incorporated Helm Elicited Language Program for Syntax Stimulation (HELPSS), which included sentence production training with Level A prompting (i.e., delayed repetition) and Level B prompting (i.e., verbal stimulus requiring response to a question). **Intensity**: 3x/week, 6 months max	65.68(4.93) 69.28(4.20) 3.60	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Duncan, Schmah, & Small, 2016	19; 19; M = 15; F = 4	53.50(11.70) 31-72	Broca's = 9 Anomic = 6 Conduction = 1 Wernicke's = 1 TSM = 1 TCM = 1	Mild-to-Mod.	chronic 5-130	**Type**: I **Description**: Imitation-based therapy wherein PWA listened to words and phrases produced by six different speakers and then, repeated them once or numerous times. Half of the PWA were also exposed to a video of the speaker. **Intensity**: 6x/week, 90 min, 6 weeks	67.72(20.00) 70.34(18.33) 2.62	Level of Evidence: III/ class IV Study protocol: + Blinding: + Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
Edmonds & Kiran, 2006	3;3; M = 1; F = 2	54.00(1.73) 53-56	Nonfluent = 3	Mod. =2 Sev. = 1	8.67 8-9	**Type**: I **Description**: Semantic feature analysis-based (SFA-based) treatment (Boyle & Coehlo, 1995 Kiran & Thompson, 2003) involving the following steps: 1) initial naming attempt 2) written feature verification 3) yes/no feature questions 4) second naming attempt. Treatment was administered in both languages. **Intensity**: 2x/week, 120 min, 7-34 weeks	48.33(24.66) 58.33(16.07) 10.00	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Edmonds, Mammino, & Ojeda, 2014	11;10; M = 6; F = 4	63.30(13.07) 35-81	Anomic = 5 Conduction = 2 TCM = 2 Wernicke's = 1	Mild = 1 Mild-to-Mod. = 8 Mod. = 1	57.5 14-144	**Type**: I **Description**: Verb Network Strengthening Treatment (VNeST): PWA were given a verb then, asked to retrieve related agents and patients. They are encouraged and supported to generate multiple pairs of agents and patients for each verb. **Intensity**: 2x/week, 120min, 10 weeks	75.91(10.36) 82.08(8.54) 6.17	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
Edmonds, Nadeau, & Kiran, 2009	4; 4; M = 1; F = 3	61.50(10.08) 52-75	TMA = 2 Conduction = 2	Mild-to-Mod. = 4	37.25 10-96	**Type**: I **Description**: Verb Network Strengthening Treatment (VNeST): 1) PWA were given a verb. 2) Asked to produce 3-4 thematic role pairs. 3) Picked a thematic role pair and answered wh-questions about it **Intensity**: 2x/week, 120 min, avg. 4.75 weeks (4-6 weeks)	74.83(3.41) 83.10(2.27) 8.28	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Falconer & Antonucci, 2012	4; 4; M = 3; F = 1	45.75(15.09) 31-62	Conduction = 2 Broca's = 1 TCM = 1	Mild-to-Mod. = 1 Mod. = 2 Mod.-to-sev. = 1	86.99 24-156	**Type**: INT **Description**: Modified Promoting Aphasics' Communication Effectiveness (PACE) approach: Within a small group, PWA took turns describing stimuli hidden from others with enough detail for others to guess the item) When word-retrieval difficulty occurred, the activity was briefly discontinued while PWA were led through the SFA chart (Boyle,2004) until they accessed the target. HW assignments included describing difficult-to-name pictured objects using SFA outside of treatment sessions. **Intensity**: 2x/week, 90-120 min, 7 weeks	54.15(15.39) 57.00(16.22) 2.85	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Faroqi-Shah, 2013	6; 6; M = 5; F = 1	47.67(8.71) 37-56	Broca's = 6	Mild-to-Mod. = 3 Mod. = 2 Sev. = 1	33.17 16-84	**Type**: I Morphosemantic treatment: 1) name action in 3 pictures 2) grammaticality judgment 3) match spoken sentence to picture 4) PWA were given a sentence and asked to write the verb inflection to match the picture 5) PWA arranged words in the correct order to form the sentence matching the picture. Trained past, present and future tenses of 20 verbs. **Intensity**: 4x/week, 60-120 min, 3 weeks	59.97(22.20) 77.57(12.86) 17.60	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
Faroqi-Shah, 2008	4; 4; M = 2; F = 2	64.50(3.87) 59-68	Broca's = 3 TCM = 1	Mild-to-Mod. = 1 Mod. = 3	56.99 12-108	**Type**: I Morphophonological treatment: 1) Naming the action from a picture 2) Auditory discrimination 3) Lexical decision 4) Morphology generation 5) Oral and written transformation 6) Repetition AND Morphosemantic treatment 1) Naming the action 2) Anomaly judgment (i.e., identifying mismatch between adverb & verb tense) 3) Auditory Comprehension (i.e., matching sentence to picture) 4) Sentence completion (i.e., fill in the blank with correct verb form) 5) Sentence construction (i.e., arranging words in the correct order) **Intensity**: 4-5x/week, 60-120 min, 3 weeks	64.65(2.83) 72.05(3.71) 7.40	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Ferguson, Evans, & Raymer, 2012	4; 4; M = 2; F = 2	57.75(14.20)	Broca's = 2 Conduction = 1 TCM = 1	Mild-to-Mod. = 2 Sev. = 2	34.75 22-41	**Type**: I Intention Gesture Treatment (IGT): 1) Participants generated L-hand gesture and pressed button to view target noun then, attempted to name. 2) If they were inaccurate, the SLP modeled the gesture and noun together and participant imitated 4-6 times. 3) SLP modeled again and PWA rehearsed gesture and verbal production 4-6 times. 4) PWA re-attempted to produce the target noun after producing gesture and pressing the red button. Pantomime Gesture treatment (PGT)) 1) PWA were trained to produce pantomime gestures. 2) SLP pushed button to change picture, then PWA attempted to name. 3) If they were inaccurate, SLP produced gesture and verbal model of target and PWA imitated 4-6x. 4) SLP modeled again and participant practiced the gesture and verbal target again. 5) They re-attempted production of the target after SLP pressed button. **Intensity**: 2-3x/week, 45-60 min, 3-5 weeks, 1 week break then 2-3x/week, 45-60 min, 3-5 weeks	50.45(30.39) 55.70(30.84) 5.25	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
R. K. Johnson, Hough, King, Vos, & Jeffs, 2008	3; 3; M = 1; F = 2	67.67(10.07) 57-77	Broca's =2 Mixed =1	Mod. = 1 Mod.-to-sev. = 1 Sev. = 1	52.68 27-93	**Type**: INT **Description**: Intensive therapy using computer-based augmentative alternative communication (AAC) (i.e., symbol identification, navigation, scenario role play, sentences). It involved training caregiver in	32.87(14.62) 33.40(6.48) .53	Level of Evidence: III/ class IV Study protocol: + Blinding: -

ACCEPTED MANUSCRIPT

						therapy and use of an AAC device to reduce the severity of the impairment and increase activities and participation. **Intensity**: 3-4x/week, 60 min, 12 weeks		Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
M. L. Johnson et al., 2014	4; 4; N/A	70.75(9.57) 60-83	Broca's = 4	Mild-to-Mod. = 2 Mod. = 2	46.79 16-96	**Type**: A/P Constraint-induced aphasia therapy (CIAT) (i.e., discouragement of gesture and nonverbal vocalizations). Daily tasks included 1) Completion of How Well scale of the Verbal Activity Log (VAL) 2) Speech Repetition Drills 3) Activities of Daily Living (ADL) phrase repetition drills 4) Language Card game 5) Picture description 6) Role play 7) Home skill assignment. Caregiver present for all therapy. **Intensity**: 7x/week, 195 min, 2 weeks	66.23(7.14) 79.28(11.29) 13.05	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: + Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Kendall et al., 2008	10; 10; M = 6; F = 4	52.40(11.40) 40-76	NS	Mild = 2 Mild-to-Mod. = 7 Mod. = 1	59.7 16-120	**Type**: I **Description**: Phonologically-based treatment: 1)Trains subjects on individual phonemes and 2) Trains phonological and orthographic sequence knowledge at the syllable level **Intensity**: 4x/week, 120 min, 12 weeks	77.12(14.47) 82.77(14.08) 5.65	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: - Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
Kendall, Raymer, Rose, Gilbert, & Gonzalez Rothi, 2014	8; 8; M = 4; F = 4	62.00(9.65) 46-72	N/A	Mild = 3 Mild-to-Mod. = 3 Mod. = 1 Mod.-to-sev. = 1	63.13 11-120	**Type**: I **Description**: Naming pictures with semantic, phonologic, repetition and orthographic cueing hierarchy including a delayed-recall step. **Intensity**: 3x/week, 60 min, 3.5 weeks	74.45(18.29) 79.35(20.03) 4.90	Level of Evidence: III/ class IV Study protocol: + Blinding: + Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: - Significance for st. outcome of interest: + Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Kendall, Oelke, Brookshire, & Nadeau, 2015	26; 26; M = 15; F = 11	56.04(14.53) 26-78	NS	Above cut-off = 5 Mild = 6 Mild-to-Mod. = 9 Mod. = 5 Mod.-to-sev. = 1	47.5 8-211	**Type**: I **Description**: Multimodal, phonologically-based therapy using phonemes in isolation and one-, two-, and three-syllable sequences in real words and nonword combinations. More specifically, Stage 1) targeted sounds in isolation and Stage 2) targeted sounds in syllables. Each stage involves an overview, introduction of sounds and sound sequences, perception tasks and production tasks. **Intensity**: 5x/week, 120 min, 6 weeks	78.68(16.53) 82.65(12.58) - 0.08	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
Kiran & Thompson, 2003	4; 4; M = 1; F = 3	68.50(5.92) 63-75	Fluent = 4	Mild-to-Mod. = 1 Mod. = 2 Mod.-to-sev. = 1	33.75 9-99	**Type**: I **Description**: Typicality-based SFA treatment involving 1) Naming 2) Category Sorting 3) Feature Verification 4) Answering yes/no questions **Intensity**: 2x/week, 120 min, 17-35 weeks	52.68(11.95) 60.90(12.81) 8.23	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Kiran, 2005	3; 3; M = 3	63.67(4.16) 59-67	TCM =1 Broca's = 1 Anomic = 1	Mild-to-Mod. = 2 Mod. = 1	156 24-288	**Type**: I **Description**: Phoneme-to-grapheme conversion: 1) writing to dictation of the word 2) copying the word 3) oral reading of the word 4) selecting and writing the sounds of the target 5) writing phonemes of the target word presented aloud 6) writing to dictation of the word **Intensity**: 2x/week, 120 min, 5-10 weeks	73.10(12.25) 70.97(14.33) - 2.13	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: - Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Kiran & Johnson, 2008	3; 3; M = 2; F = 1	62.33(11.15) 54-75	Anomic = 3	Mild = 1 Mild-to-Mod. = 2	18 7-36	**Type**: I **Description**: Typicality-based SFA treatment 1) Naming the picture 2) sorting pictures of target category 3) selecting written features for the target 4) answering written yes/no questions 5) naming the picture **Intensity**: 2x/week, 120 min, avg. 14 weeks (8-18 weeks)	84.70(2.42) 88.70(0.36) 4.00	Level of Evidence: III/ class IV Study protocol: - Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Kiran, 2008	5; 5; M = 1; F = 4	58.40(12.03) 47-77	Conduction= 3 Broca's = 2	Mild-to-Mod. = 1 Mod. = 3 Mod.-to-sev. = 1	8.2 7-10	**Type**: I **Description**: SFA-based treatment involved 1) naming the picture 2) sorting pictures by category 3) identify semantic features 4) answer yes/no feature questions **Intensity**: 2x/week, 60 min, 24 weeks	54.96(13.77) 64.02(12.91) 9.06	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Kiran, Sandberg, & Abbott, 2009	4; 4; M = 2; F = 2	56.75(15.63) 39-77	Anomic = 4	Mild = 4	25.5 8-43	**Type**: I **Description**: SFA-based treatment involving: 1) category sorting 2) feature selection 3) yes/no feature questions 4) word recall and 5) free generative naming. **Intensity**: 24 sessions	87.75(1.52) 90.58(1.63) 2.83	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Kiran, Sandberg, & Sebastian, 2011	6; 6; M = 3; F = 3	68.00(15.76) 39-84	Anomic = 4 Conduction=3	Mild-to-Mod. = 6	55.83 9-108	**Type**: I **Description**: SFA-based treatment involving 1) category generation 2) category sorting 3) feature generation and/or selection and 4) answering yes/no feature questions **Intensity**: 2x/week, 120 min, 10 weeks	78.85(6.06) 81.98(8.77) 3.13	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Lesser, Bryan, Anderson, & Hilton, 1986	13; 9; M = 3; F = 6	60.00(10.83) 40-76	Broca's = 5 Conduction = 4	Mild-to-Mod. = 3 Mod. = 4 Mod.-to-sev. = 2	15.56 2-33	**Type**: INT **Description**: Language Enrichment Therapy (LET): 160 basic words were pictured in line drawings in sets of eight within language tasks of increasing complexity from picture-matching to understanding a text (28 units of complexity). Each unit repeated the same exercise 20 times with different vocabulary. Exercises involve comprehension, repetition, naming, constructing sentences, reading and writing. Spouse/volunteer can use materials with PWA between therapy visits. LET was supplemented with conversation and counseling. **Intensity**: 1x/week, 60 min, 10-12 weeks	59.09(11.58) 65.60(14.79) 6.51	Level of Evidence: IIB/ class III Study protocol: + Blinding: + Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Macauley, 2006	3; 3; M = 3	63.00(4.00) 59-77	Nonfluent = 3	Mild-to-Mod. = 3 Mod. = 1	72 48-84	**Type**: INT **Description**: Traditional Therapy 1) PWA asked to name a picture 2) SLP asks for a phrase using the word Animal-assisted therapy: 1) Stimuli cards were arranged throughout the room with dog treats on them 2) PWA asks dog to "find treat" 3) SLP picks up card from dog who just ate the treat and asks client to name it. SLP asks PWA to tell dog phrase containing the target word. 4) Dog shakes hand or barks to say "well done" when the PWA says it accurately. All PWA had both treatments. **Intensity**: 1x/week, 30 min, 24 weeks	74.37(10.65) 75.40(8.73) 1.03	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Marshall, Laures-Gore, DuBay, Williams, & Bryant, 2015	3; 3; M = 1; F = 2	58.00(13.89) 49-74	Broca's = 2 Global = 1	Mod. = 1 Mod.-to-sev. = 2	18.33 12-22	**Type**: INT **Description**: PWA received conventional speech therapy while also practicing unilateral nostril breathing techniques (i.e., diaphragmatic breathing and close nostril on their affected side, inhale through the open nostril and exhale for twice as long than their inhalation). **Intensity**: 2x/week, 40 min, avg. 14 weeks (14-18 weeks)	38.33(12.52) 40.63(6.75) 2.30	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Milman, Clendenen, & Vega-Mendoza, 2014a	3; 3; M = 3	N/A(N/A) 56-68	Nonfluent = 3	Mod. =1 Mod.-to-sev. = 2	n/a 12-84	**Type**: INT **Description**: Functional use of adjectives to describe people in four different tasks: 1) single-word adjective production 2) single-word pronoun production 3) sentence training and 4) discourse production. Semantic, orthographic and phonemic cues were given to facilitate single word use. **Intensity**: 4x/week, 60 min, avg. 9 weeks (5-12 weeks)	43.70(8.75) 51.40(16.41) 7.70	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Milman, Vega-Mendoza, & Clendenen, 2014b	3; 3; M = 1; F = 2	62.33(6.35) 55-66	Non-fluent = 3	Mild-to-Mod. = 1 Mod.-to-sev. = 2	41 22-61	**Type**: INT **Description**: Each individual session targeted: 1) word retrieval 2) sentence production and 3) discourse-level communication to integrate training received in steps 1 and 2. Daily homework was assigned focusing on material from steps 1 and 2. Group session once weekly to transfer skills from individual therapy to conversational level. **Intensity**: 4x/week, 60 min, avg. 11 weeks (6-18)	47.07(27.67) 52.67(26.37) 5.60	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: - Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Mozeiko, Coelho, & Myers, 2016_Intensive	4; 4; M = 2; F = 2	54.50(20.63) 26-72	Broca's = 2 Not classifiable = 1 Global = 1	Mild-to-Mod. = 1 Mod.-to-sev. = 2 Sev. = 1	65.4 18-134	**Type**: A/P **Description**: Intensive Constraint Induced Language Therapy (CILT) (i.e., 5x/week): PWA participated in a Go Fish game wherein they had to ask one another for a card that matches one of their own. SLP increases the difficulty level by accepting different responses: Level 1) single word response with high frequency cards and Level 2) introduces a carrier phrase to the single word 3) adds an adjective to the carrier phrase with a single word response 4) adds two adjectives to the carrier phrase with a single word response. **Intensity**: 5x/week, 180 min, 2 weeks	38.05(20.00) 46.35(20.97) 8.30	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Mozeiko, Coelho, & Myers, 2016_Distributed	4; 4; M = 3; F = 2	59.50(13.50) 47-77	Broca's = 1 Anomic = 1 Conduction = 1 Not classifiable = 1	Mild-to-Mod. = 1 Mod. = 1 Mod.-to-Sev. = 2	36 13-96	**Type**: A/P **Description**: Distributed Constraint Induced Language Therapy (CILT) (i.e., 3x/week) PWA participated in a Go Fish game wherein they had to ask one another for a card that matches one of their own. SLP increases the difficulty level by accepting different response Level 1) single word response with high frequency cards and Level 2) introduces a carrier phrase to the single word 3) adds an adjective to the carrier phrase with a single word response 4) adds two adjectives to the carrier phrase with a single word response. **Intensity**: 3x/week, 60min, 10 weeks	59.20(24.72) 62.08(23.19) 2.88	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Purdy & Wallace, 2015	3; 3; M = 3	53.33(12.22) 40-64	Broca's = 3	Mod.-to-sev. = 1 Sev. = 2	26.67 10-48	**Type**: I **Description**: 1) Multimodality training of nouns and 2) training communicative use of the targets (i.e., Promoting Aphasic's Communication Effectiveness (PACE) (Davis & Wilcox, 1985) **Intensity**: 5x/week, 120-180 min, 2 weeks	22.27(4.97) 25.63(2.83) 3.37	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Raymer, Kohen, & Saffell, 2006a	5; 5; M = 2; F = 3	70.80(12.11) 51-82	Conduction = 2 Broca's = 2 Mixed = 1	Mild-to-Mod. = 2 Mod.-to-sev. = 3	18.4 4-42	**Type**: I **Description**: MossTalk Words (i.e., computer-assisted treatment program). PWA completed multi-modal matching exercises involving 1) spoken plus written word to picture matching 2) spoken word to picture matching 3) written word to picture matching **Intensity**: 1-2x/week, 60 min, 6-12 weeks, then, 3-4x/week, 60 min 3-4 weeks. 4 week break in between each 12-hour tx. phase.	53.32(19.14) 57.40(17.26) 4.08	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: + Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Raymer, Singletary, et al., 2006b	9; 9; M = 6; F = 3	60.67(9.08) 49-70	Broca's = 6 Wernicke's = 2 Conduction =1	Mild-to-Mod. = 1 Mod. = 4 Mod.-to-sev. = 3 Sev. = 1	29 5-62	**Type**: I **Description**: Gesture-Verbal Treatment (GVT): 1) SLP showed the picture and modeled the target word and a gesture. 2) PWA produced word and gesture three times 3) SLP showed gesture in isolation and participant imitated three times 4) SLP presented the target and PWA repeated it three times 4) After a 5-second delay, SLP prompted participant to show and tell them what happened in the picture. **Intensity**: 3-4x/week, 60 min, 10 weeks	47.64(16.89) 52.43(15.46) 4.79	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
Raymer et al., 2012	8; 8; M = 4; F = 4	58.13(14.30) 40-79	Broca's = 4 TMA = 2 TSA = 1 Wernicke's = 1	Mild-to-Mod. = 1 Mod. = 7	13.5 5-30	**Type**: I **Description**: Errorless Naming: 1) SLP modeled the picture name and PWA repeated 2)SLP showed the written word and PWA read it aloud three times 3)Written word was removed and PWA was given 5 seconds to hold onto it 4) SLP prompted PWA to name it again. Gestural Facilitation: 1) SLP modelled the name and a related gesture 2) SLP modelled the gesture alone for SLP to imitate three times 3) SLP modelled name and PWA repeated three times 4) Clinician modelled gesture while showing the picture 5) After 5 second delay SLP prompted PWA to provide name and gesture again. **Intensity**: 2-3x/week, 60 min, 10 weeks	56.91(5.43) 63.40(11.46) 2.50	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Rider, Wright, Marshall, & Page, 2008	3; 3; M = 2; F = 1	63.33(9.07) 55-73	Nonfluent =3	Mild-to-Mod. = 2 Mod. = 1	65.67 26-126	**Type**: I **Description**: Trained words related to 6-8 contexts (i.e., story retell and procedural narratives) using SFA (e.g., Boyle, 2004 Boyle & Coelho, 1995) **Intensity**: 2-3x/week, 60 min, 7-14 weeks	72.30(5.71) 73.43(8.00) 1.13	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
Rodriguez, Raymer, & Rothi, 2006	4; 4; M = 3; F = 1	65.00(9.76) 52-73	Conduction = 2 Wernicke's =1 Broca's = 1	Mild-to-Mod. = 1 Mod. = 1 Mod.-to-sev. = 2	34.25 8-96	**Type**: I **Description**: All PWA received both verb naming treatments. Gesture-Verbal Treatment (GVT): 1) SLP showed the picture and modeled the target word and a gesture. 2) PWA produced word and gesture three times 3) SLP showed gesture in isolation and PWA imitated three times 4) SLP presented the target and PWA repeated it three times 4) After a 5-second delay, SLP prompted participant to show and tell them what happened in the target picture. Semantic-Phonologic Treatment: 1) SLP showed PWA the picture and modeled the target word 2) PWA answered semantic and phonologic questions about the target 3) PWA produced the target three times 4) After a 5-second delay, PWA attempted to explain what was happening in the picture. **Intensity**: 2-3x/week, 60 min, 10-14 weeks	53.40(18.00) 56.45(20.18) 3.05	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Rose, Attard, Mok, Lanyon, & Foster, 2013	11; 11; M = 6; F = 5	58.09(10.63) 39-74	Broca's = 6 Anomic = 4, Conduction= 1	Mild = 2 Mild-to-Mod. = 2 Mod. = 6 Mod.-to-sev. = 1	44 17-88	**Type**: A/P **Description**: PWA targeted word retrieval in small groups through treatment activities including(i.e,. Go Fish, Memory, Request Role plays, Board games, rapid naming while playing snap, Who am I) In CIAT Plus: Verbal production was the goal but cueing was provided as needed (i.e,. phonemic cue, written cue). In multi-modal aphasia therapy (M-MAT): Verbal production was also the goal but, multi-modal cueing was provided (i.e., gesture, drawing, written model, verbal model). All PWA received both treatments. **Intensity**: 4x/week, 195 min, 2 weeks	66.26(18.29) 70.78(16.55) 4.52	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: - Significance for st. outcome of interest: + Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Sandberg, Bohland, & Kiran, 2015	10; 10; M = 3; F = 7	59.40(10.01) 47-75	Anomic = 6 Conduction = 2 Broca's = 1 TCM = 1	Above cut-off = 3 Mild = 1 Mild-to-Mod.= Mod.-to-sev. =	55.7 7-134	**Type**: I **Description**: PWA were trained on ten abstract words in a particular context category (e.g., courthouse) and ten untrained concrete words from the same context-category were monitored to measure generalization. Treatment steps included 1) Feature selection 2) Abstract/concrete lexical decision 3) Synonym generation **Intensity**: 2x/week, 120 min, 10 weeks	80.52(17.41) 84.32(15.00) 3.80	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Schneider & Thompson, 2003	7; 6; N/A	N/A(N/A) N/A	Broca's	Mild-to-Mod. = 5 Mod. = 1	N/A 39-132	**Type**: I **Description**: Semantic verb retrieval treatment or argument structure verb retrieval treatment was applied to a category of verbs. Semantic treatment focused on the meaning of verb and argument structure focused on number of argument structures pertaining to the verb and its thematic role assignment. Treatment involved 3 steps: 1) Presentation of the item 2) Presentation of the meaning or thematic role information for the verb being trained 3) PWA names the item. **Intensity**: 24 sessions	72.43(6.44) 76.60(4.39) 4.17	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
Silkes, 2015	4; 4; N/A	60.25(1.26) 49-60	Nonfluent = 1 Fluent = 3	Mild=1 Mild-to-Mod. = 3	58.5 24-96	**Type**: I **Description**: Masked repetition priming treatment: Each section PWA saw prime-picture pair 16 times and had four opportunities to name each picture. PWA were instructed to watch the screen and try to name the picture when they saw it for the 4th time. **Intensity**: 2x/day XX 12 days	75.13(9.56) 76.33(6.98) 1.20.00	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Steele, Baird, McCall, & Haynes, 2014	9; 9; M = 7; F = 2	61.44(10.42) N/A	Broca's = 5 TCM = 1 Wernicke's = 1 Isolation = 1 Conduction = 1	Mod.	66.72 16-230	**Type**: INT **Description**: Individual therapy (i.e., improving conversational skills using script training, sentence patterning and response elaboration), group therapy (i.e., word retrieval, improve speech intelligibility, train social exchanges, train longer and more complex sentences, increase conversational turns, improve well-being, increase life participation)and online language exercises (i.e., Talk Path: listening, speaking, reading and writing activities) **Intensity**: 2x/week 60 min 20.6 weeks	53.90(9.40) 57.40(10.40) 3.50	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: +

ACCEPTED MANUSCRIPT

								Precision: + Intention to treat: N/A
Thompson, Shapiro, Kiran, & Sobecks, 2003	4;4; N/A	N/A(N/A) N/A	Broca's = 4	Mild-to-Mod. = 2 Mod. = 2	N/A 12-132	**Type**: I **Description**: Trained to comprehend & produce different sentence **Type**s using Treatment of Underlying Forms (Thompson, 2001) **Intensity**: 2x/week 120 min 3-9 weeks	66.20(5.51) 68.38(7.04) 2.175.00	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Waller, Dennis, Brodie, & Cairns, 1998	4; 3; M = 2; F = 1	62.33(10.21) 55-74	Nonfluent = 3	Mod. = 2 Sev. = 1	31.56 15-78	**Type**: INT **Description**: PWA were trained to retrieve pre-programmed items in their TalkBac (i.e., word-based AAC device with personal sentences and stories) by SLP. Trained caregivers provided opportunities to elicit those phrases and SLP visited weekly to provide support. Three group sessions were organized to allow caregivers and subjects to meet and discuss pros and cons of the project. **Intensity**: 1x/week 90 min 52 weeks	38.77(28.34) 45.77(31.88) 7.00	Level of Evidence: III/ class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Wilson et al., 2012	9;9; M = 8; F = 1	52.22(9.73) 28-62	Anomic = 5 Nonfluent = 2 Global = 1 Mixed = 1	Above cut-off = 1 Mild = 3 Mild-to-Mod. = 2 Mod. = 1 Mod.-to-sev. = 1 Sev.= 1	22.68 6-66	**Type**: INT **Description**: PWA participated in Intensive Residential Aphasia Communication Theraprogram (InteRACT Carey et al. 2006). Five hours of daily treatment included focus on speech and language skills, functional communication strategy usage, community re-integration and communication partner training. **Intensity**: 5x/week 300 min 4 weeks	69.91(28.85) 76.09(25.25) 6.18	Level of Evidence: IIA/ class II Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Supplementary Material 5: Demographic information for within group studies using Communicative Effectiveness Index

Note: M=Male; F = Female; TCM = transcortical motor; TSM = transcortical sensory; sev. = severe; I = Impairment-based treatment; A/P = activity/participation-based treatment; INT = integrated treatment; st. = standardized; CS = convenience sample

Study Name	Study N; N for outcome measure; Sex	Mean Age(SD) range	Aphasia Type	Aphasia Severity (WAB-AQ)	Mean MPO range	Treatment(Tx)	Pre-Tx M(SD) Post-Tx M(SD) Change Score	Methodological Rigor
Aftonomos, Appelbaum, & Steele, 1999	60; 29; M=35; F=25	68.60(12.30) 24-86	Broca's= 21 Anomic = 13 Global = 11 Wernicke's = 8 Conduction = 3 TCM = 2 TSM = 1 Isolation=1	Mod.-to-sev.	24.6 .24-144	**Type**: INT **Description**: Individual treatment using the Lingraphica (icon-based language system) to provide therapeutic exercises at the appropriate level for participants' severity. Also, focused on improving functional communication outside of the clinic as well as provided home exercises. **Intensity**: 2x/week, 60 min, 20. 5 weeks	42.80(19.00) 62.60 (18.60) 19.80	Level of Evidence: IIB/class III Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
Archibald, Orange, & Jamieson, 2009	8; 3; M=2; F=1	77.67(9.02) 69-87	Broca's =1 Anomic =1 Global =1	Mod.=1 Mod.-to-sev. =2	44.66 10-105	**Type**: I **Description**: Computer-provided treatment via AphasiaMate across 8 modules (i.e., auditory comprehension, visual matching, reading comprehension, spelling, semantics, sentence processing). Patients used computer at home or in clinic with trained personnel. **Intensity**: 1x/week, 60 min, 15 weeks	41.67(3.79) 51.67(6.66) 10	Level of Evidence: IIB/class III Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Babbitt, Worrall, & Cherney, 2015	74;74; M=52; F=22	54.10(16.30) 18-86	Nonfluent = 49 Fluent = 25	Mod.	15.5 3-87	**Type**: INT **Description**: Intensive Comprehensive Aphasia Program (ICAP): two individual therapy sessions and one session each of constraint-induced language therapy (CILT), reading/writing, computers and conversation group for six hours of daily programming. **Intensity**: 5x/week, 240 min, 4 weeks	46.80(15.70) 58.20(16.20) 11.4	Level of Evidence: IIB/class III Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
Barthel, Meinzer, Djundja, & Rockstroh, 2008	12; 12; M=5; F=7	55.20(14.20) 35-76	Broca's=8 Anomic=1 Global=1 Non-standard=2	Mod.-to-sev.	64 13-156	**Type**: INT **Description**: Model-oriented aphasia therapy (MOAT) was provided on an individual basis. It combines model-oriented aphasia therapy (i.e., target semantic system), linguistic approach (i.e., target phonological errors), strategy approach (i.e., paraphrasing), communicative approach (i.e., role playing) and involvement of relatives. **Intensity**: 1x/day, 180 min, 10 days	42.60(21.30) 51.80(20.90) 9.20	Level of Evidence: IIB/class III Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
Code, Torney, Gildea-Howardine, & Willmes, 2010	8; 7; M=6; F=1	52.71(13.40) 36-73	Broca's= 3 Global= 2 Wernicke= 1 Amnesic= 1	Mod.	34.43 9-70	**Type**: INT **Description**: Individual and group therapy was administered based on participants' pre-treatment testing results. Weekly counseling was offered to caregivers & participants. **Intensity**: 5x/week, 4 weeks	45.00(10.10) 57.50(24.10) 12.50	Level of Evidence: IIB/class III Study protocol: - Blinding: + Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Edmonds, Mammino, & Ojeda, 2014	11; 9; M=6; F=4	63.30(13.07) 35-81	Anomic=5 Conduction=2 TCM=2 Wernicke's=1	Mod.	57.5 14-144	**Type**: I Desription: Verb Network Strengthening Treatment (VNeST): PWA were given a verb then, asked to retrieve related agents and patients. They are encouraged and supported to generate multiple pairs of agents and patients for each verb. **Intensity**: 2x/week, 120 min, 10 weeks	32.61(9.70) 65.28(11.20) 32.67	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
Johnson, Hough, King, Vos, & Jeffs, 2008	3; 3; M=1; F=2	67.67(10.07) 57-77	Broca's =2 Mixed =1	Mod.=2 Sev. = 1	52.68 27-93	**Type**: INT **Description**: Intensive therapy using computer-based augmentative alternative communication (AAC) (i.e., symbol identification, navigation, scenario role play, sentences). It involved training caregiver in therapy and use of an AAC device to reduce the severity of the impairment and increase activities and participation. **Intensity**: 3-4x/week, 60 min, 12 weeks	26.27(13.49) 38.93(3.09) 12.66	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Milman, Vega-Mendoza, & Clendenen, 2014	3;3; M=1; F=2	62.33(6.35) 55-66	Nonfluent=3	Mod.=2	41 40-66	**Type**: INT **Description**: Each individual session targeted: 1) word retrieval 2) sentence production and 3) discourse-level communication to integrate training received in steps 1 and 2. Daily homework was assigned focusing on material from steps 1 and 2. Group session once weekly to transfer skills from individual therapy to conversational level. **Intensity**: 4x/week, 60 min, avg. 9 weeks (5-12 weeks)	48.77(10.46) 49.83(14.91) 1.07	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: - Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Nickels & Osborne, 2016	4; 4; M=3; F=1	59.75(18.66) 34-74	Global= 2 Anomic= 1 TCS= 1	Mild-to-Mod.= 1 Mod.=1 Mod.-to-sev.= 2	25.8 15-42	**Type**: A/P **Description**: Constraint Induced Aphasia Therapy Plus (CIAT-plus): Therapy addressed verbal expression through Go Fish. 1) Volunteer played the game with two PWA. 2) PWA chose a card and asked the other players for a card. 3) Co-player then responded. PWA could use multi-modal communication, as needed. Shaping was included to increase the complexity of their verbal responses. **Intensity**: 2x/week, 90 min, 4 weeks	61.75(11.27) 61.00(11.60) -.75	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
Raymer et al., 2012	8; 6; M=2; F=4	60.33(14.49) 47-79	Broca's= 3 TMA =1 TSA =1 Wernicke's= 1	Mild= 1 Mild-to-Mod.= 2 Mod.=2 Mod.-to-sev.= 1	14 5-30	**Type**: I **Description**: Errorless Naming: 1) SLP modeled the picture name and PWA repeated 2)SLP showed the written word and PWA read it aloud three times 3)Written word was removed and PWA was given 5 seconds to hold onto it 4) SLP prompted PWA to name it again. Gestural Facilitation: 1) SLP modelled the name and a related gesture 2) SLP modelled the gesture alone for SLP to imitate three times 3) SLP modelled name and PWA repeated three times 4) Clinician modelled gesture while showing the picture 5) After 5 second delay SLP prompted PWA to provide name and gesture again. **Intensity**: 2-3x/week, 60 min, 10 weeks	64.37(19.91) 59.27(20.42) - 5.09	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Rodriguez et al., 2013	11;9; M=4; F=4	60.45(17.67) 18-79	N/A	Mod.-to-sev.	25.81 8-56	**Type**: INT **Description**: Individual treatment involved both impairment-based and functional therapy. PWA and family members were involved in group treatment (i.e., share information about available local services, facilitate discussions about "living with aphasia" , promoting social interaction and multi-modal communication) Computer-based therapy (i.e., Bungalow, REACT, Speech Sounds on Cue) Challenge Task: specific goal each PWA wanted to achieve by the end of the program. **Intensity**: 5x/week, 240 minutes, 2 weeks OR 5x/week, 300 minutes, 4 weeks	41.60(15.50) 55.30(16.60) 13.70	Level of Evidence: III/class IV Study protocol: + Blinding: + Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Rose, Attard, Mok, Lanyon, & Foster, 2013	11;10; M= 5; F = 5	59.30(10.37) 39-74	Broca's = 5 Anomic = 4 Conduction = 1	Mild - 1 Mild-to-Mod. - 2 Mod. – 3 Mod.-to-sev. - 4	46.21 7-88	**Type**: A/P **Description**: PWA targeted word retrieval in small groups through treatment activities including (i.e,. Go Fish, Memory, Request Role plays, Board games, rapid naming while playing snap, Who am I) In CIAT Plus: Verbal production was the goal but cueing was provided as needed (i.e,. phonemic cue, written cue). In multi-modal aphasia therapy (M-MAT): Verbal production was also the goal but, multi-modal cueing was provided (i.e., gesture, drawing, written model, verbal model). All PWA received both treatments. **Intensity**: 4x/week, 195 min, 4 weeks	53.00(22.13) 62.29(20.52) 9.29	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: - Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
Sorin-Peters & Behrmann, 1995	3;3; M=2; F=1	69.33(3.06) 66-72	Fluent =3	Mild=1 Mild-to-Mod.=1 Mod.=1	5 3-7	**Type**: INT **Description**: Individual Therapy goals: Targeted the impaired process and taught PWA to compensate using non-verbal techniques with the overall goal of increasing PWAs' participation in the community and their conversational skills. Group Therapy goals: Gave PWA the opportunity to use compensatory strategies in a more a natural setting. Individual goals were incorporated into group discussion. PWA also participated in groups at the Day Treatment Center, which were led by nursing, recreational therapy, occupational therapy and social work personnel that were trained to use compensatory strategies with PWA. **Intensity**: 2x/week, 60 min, avg. 22.14 weeks (4-6.5 mos)	72.00(13.45) 69.67(20.13) -2.33	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Steele, Baird, McCall, & Haynes, 2014	9; 8; M= 6; F = 2	60.75(10.91) 43-77	Broca's = 4 Conduction = 1 TCM = 1 Isolation = 1 Wernicke's = 1	Mod.	65.85 16-230	**Type**: INT **Description**: Individual therapy (i.e., improving conversational skills using script training, sentence patterning and response elaboration), group therapy (i.e., word retrieval, improve speech intelligibility, train social exchanges, train longer and more complex sentences, increase conversational turns, improve well-being, increase life participation)and online language exercises (i.e., Talk Path: listening, speaking, reading and writing activities) **Intensity**: 2x/week, 60 min, 20.6 weeks	49.90(18.60) 67.70(13.60) 17.80	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

van der Gaag et al., 2005	22; 14; M=6; F=16	58.00(13.40) 31-81	N/A	Mod.	Chronic N/A	**Type**: INT **Description**: Group therapy (i.e., conversation, communication skills, art, discussion, self-advocacy, monitoring communication skills of conversation partners) **Intensity**: 1x/week, 120 min, 7 weeks	48.40(20.43) 58.10(19.85) 9.70	Level of Evidence: IIB/class III Study protocol: + Blinding: + Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
Wambaugh, Wright, & Nessler, 2012	6; 6; M=4; F=2	57.83(8.26) 46-70	Nonfluent = 5 Fluent =1	Mild=1 Mild-to-Mod. = 3 Mod. = 2	44.5 19-96	**Type**: A/P **Description**: PWA received treatment in two picture **Description** contexts (i.e., "Tell me about the picture") and one personal recount context (i.e.,"Tell me something about anything you would like to talk about.") **Intensity**: 2-3x/week, 6 weeks	40.83(8.13) 48.33(10.46) 7.50	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Wilson et al., 2012	9;8; M= 7; F =1	51.38(10.04) 28-62	Anomic=5 TCM=1 Conduction=1 Broca's=1	Mild-to-Mod. = 2 Mod. = 4 Mod.-to-sev. =1 Sev. = 1	17.16 6-36	**Type**: INT **Description**: PWA participated in Intensive Residential Aphasia Communication Theraprogram (InteRACT Carey et al. 2006). Five hours of daily treatment included focus on speech and language skills, functional communication strategy usage, community re-integration and communication partner training. **Intensity**: 5x/week, 60 min, 4 weeks	53.63(16.41) 65.25(18.13) 11.63	Level of Evidence: IIA/class II Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Supplementary Material 6: Demographic information for within group studies using Boston Naming Test

Note: M=Male; F = Female; TCM = transcortical motor; TSM = transcortical sensory; sev. = severe; I = Impairment-based treatment; A/P = activity/participation-based treatment; INT = integrated treatment; st. = standardized; CS = convenience sample

Study Name	Study N; N for outcome measure; Sex	Mean Age(SD) range	Aphasia Type	Aphasia Severity (WAB-AQ)	Mean MPO range	Treatment(Tx)	Pre-Tx M(SD) Post-Tx M(SD) Change Score	Methodological Rigor
Aftonomos, Steele, & Wertz, 1997	23;10; M=7; F=3	62.20(10.62) 49-77	Anomic = 3 Wernicke's =2 Broca's=4 Conduction =1	Mild = 3 Mod. = 5 Sev. = 2	40.8 9-80	**Type**: INT: **Description**: Individual treatment to familiarize PWAs with their Lingraphica (LG) system and improve their performance in areas of weakness. Group treatment for 3 PWA involved PWAs using their LGs to respond to one another with a PACE treatment. PWA assigned exercises for home practice. **Intensity**: 1.96x/week, 60 min, avg. 14.1 weeks	23.00(12.03) 34.10(16.70) 11.10	Level of Evidence: IIB/class III Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
Babbitt, Worrall, & Cherney, 2015	74;74; M = 52; F = 22	54.10(16.30) 18-86	Nonfluent = 49 Fluent = 25	Mod.	15.5 3-87	**Type**: INT Intensive Comprehensive Aphasia Program (ICAP): two individual therapy sessions and one session each of constraint-induced language therapy (CILT), reading/writing, computers and conversation group for six hours of daily programming. **Intensity**: 5x/week, 360 min, 4 weeks	16.60(19.00) 20.70(20.50) 4.10	Level of Evidence: IIB/class III Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Breier, Maher, Novak, & Papanicolao,2006	6;6; M=4; F=2	61.33(8.80) 53-77	Broca's=5 Conduction =1	Above =2 Sev. =2 Mod. =1	47 21-70	**Type**: A/P **Description**: Constraint Induced Language Therapy (CILT) = Only verbal expression was accepted and multi-modality communication was restricted, even self-cueing. Treatment was conducted in dyads and consisted of a dual card task with barrier present (i.e., PWA took turns requesting a card or responding another's request). Stimuli included four sets of cards of different semantic categories with two levels of difficulty (i.e., low- and high-frequency). Clinicians used shaping (i.e., increasing communicative demands of request/response from single words to lengthier sentences) and cueing for a successful production (i.e., semantic, phonemic, repetition). **Intensity**: 4x/week, 180 min, 3 weeks	24.00(24.46) 23.00(24.54) - 1.00	Level of Evidence: IIB/class III Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Edmonds & Kiran, 2006	3 ;3; M=1; F=2	54.00(1.73) 53-56	n/a	Mild =1 Mod. =1 Sev. =1	8.66 8-9	**Type**: I **Description**: Semantic feature analysis-based (SFA-based) treatment (Boyle & Coelho, 1995 Kiran & Thompson, 2003) involving the following steps: 1) initial naming attempt 2) written feature verification 3) yes/no feature questions 4) second naming attempt. Treatment was administered in both languages. **Intensity:** 2x/week, 120 min 7-34 weeks	22.23(20.02) 38.87(8.21) 16.63	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
Edmonds, Nadeau, & Kiran, 2009	4; 4; M=1; F=3	61.50(10.08) 52-75	TMA=2 Conduction=2	Mild= 2 Mod.= 2	37.25 10-96	**Type**: I **Description**: Verb Network Strengthening Treatment (VNeST): 1) PWA were given a verb. 2) Asked to produce 3-4 thematic role pairs. 3) Picked a thematic role pair and answered wh-questions about it. **Intensity**: 2x/week, 120 min, avg. 4.75 weeks (4-6 weeks)	36.50(11.24) 44.50(9.40) 8.00	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Falconer & Antonucci, 2012	4; 4; M=3; F=1	45.75(15.09) 31-62	Conduction=2 Broca's=1 TCM=1	Mod.=1 Sev.=3	86.99 24-156	**Type**: INT **Description**: Modified Promoting Aphasics' Communication Effectiveness (PACE) approach: Within a small group, PWA took turns describing stimuli hidden from others with enough detail for others to guess the item) When word-retrieval difficulty occurred, the activity was briefly discontinued while PWA were led through the SFA chart (Boyle,2004) until they accessed the target. HW assignments included describing difficult-to-name pictured objects using SFA outside of treatment sessions. **Intensity**: 2x/week, 90-120 min, 7 weeks	9.00(8.29) 11.75(10.90) 2.75	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Ferguson, Evans, & Raymer, 2012	4;4; M=2; F=2	57.75(14.20) 40-74	Broca's=2 Conduction = 1 TCM=1	Above = 1 Mild = 1 Sev. = 2	34.75 22-41	**Type**: I **Description**: Intention Gesture Treatment (IGT): 1) Participants generated L-hand gesture and pressed button to view target noun then, attempted to name. 2) If they were inaccurate, the SLP modeled the gesture and noun together and participant imitated 4-6 times. 3) SLP modeled again and PWA rehearsed gesture and verbal production 4-6 times. 4) PWA re-attempted to produce the target noun after producing gesture and pressing the red button. Pantomime Gesture treatment (PGT)) 1) PWA were trained to produced pantomime gestures. 2) SLP pushed button to change picture, then PWA attempted to name. 3) If they were inaccurate, SLP produced gesture and verbal model of target and PWA imitated 4-6x. 4) SLP modeled again and participant practiced the gesture and verbal target again. 5) They re-attempted production of the target after SLP pressed button. **Intensity**: 2-3x/week, 45-60 min, 9-11 weeks	26.75(25.18) 29.50(27.09) 2.75	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Fridriksson, Morrow-Odom, Moser, Fridriksson, & Baylis, 2006	3 ;3; M=3	59.00(10.82) 47-68	Broca's = 2 Anomic = 1	Mild = 1 Sev. = 2	60 24-144	**Type**: I **Description**: A combination of spaced retrieval, errorless learning and massed practice techniques were used to treat naming. Target items were selected by participants. Treatment was administered in a group setting. PWA and clinicians played board games and took turns working on naming in between turns in the game. **Intensity**: 7x/week, 240 min, 2 weeks	13.67(22.81) 14.33(21.39) .67	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Kendall et al., 2008	10;10; M=6; F=4	52.40(11.40) 40-76		Mild=6 Mod.=1 Sev.=3	59.7 16-120	**Type**: I **Description**: Phonologically-based treatment: 1)Trains subjects on individual phonemes and 2) Trains phonological and orthographic sequence knowledge at the syllable level. **Intensity**: 4x/week, 120 min, 12 weeks	30.10(13.47) 33.70(12.61) 3.60	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: - Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
Kendall, Raymer, Rose, Gilbert, & Gonzalez Rothi, 2014	8;8; M=4; F=4	62.00(9.65) 46-72	Mild = 6 Mod. = 1 Severe = 1		63.13 11-120	**Type** I: **Description**: Naming pictures with semantic, phonologic, repetition and orthographic cueing hierarchy including a delayed-recall step. **Intensity**: 5x/week, 120 min, 6 weeks	30.63(12.72) 30.75(14.05) .13	Level of Evidence: III/class IV Study protocol: + Blinding: + Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: - Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
Kendall, Oelke, Brookshire, & Nadeau, 2015	26 ;26; M=15; F=11	56.04(14.53) 26-78	No aphasia=5 Mild=6 Mild-to-Mod.=9 Mod.=5 Mod.-to-severe=1		47.5 8-211	**Type**: I **Description**: Multimodal, phonologically-based therapy using phonemes in isolation and one-, two-, and three-syllable sequences in real words and nonword combinations. More specifically, Stage 1) targeted sounds in isolation and Stage 2) targeted sounds in syllables. Each stage involves an overview, introduction of sounds and sound sequences, perception tasks and production tasks. **Intensity**: 5x/week, 120 min, 6 weeks	34.34(18.11) 37.61(16.17) 3.27	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Kiran & Thompson, 2003	4;4 F=3; M=1	68.50(5.92) 63-75	Fluent =4	Sev. = 4	33.75 9-99	**Type**: I **Description**: Typicality-based SFA treatment involving 1) Naming 2) Category Sorting 3) Feature Verification 4) Answering yes/no questions **Intensity**: 2x/week, 120 min, 17-35 weeks	6.45(2.34) 13.76(6.99) 7.31	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Kiran, 2005	3;3; M=3	63.67(4.16) 59-67	TCM =1 Broca's = 1: Anomic = 1	Mild = 1 Mod. = 1 Sev. = 1	156 24-288	**Type**: I **Description**: Phoneme-to-grapheme conversion: 1) writing to dictation of the word 2) copying the word 3) oral reading of the word 4) selecting and writing the sounds of the target 5) writing phonemes of the target word presented aloud 6) writing to dictation of the word **Intensity**: 2x/week, 120 min, 5-10 weeks	23.02(16.52) 28.00(12.51) 4.98	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Kiran & Johnson, 2008	3 ;3; M=2; F=1	62.33(11.15) 54-75	Anomic=3	Mild =1 Mod. =1 Sev. =1	18 7-36	**Type**: I **Description**: Typicality-based SFA treatment 1) Naming the picture 2) sorting pictures of target category 3) selecting written features for the target 4) answering written yes/no questions 5) naming the picture **Intensity**: 2x/week, 120 min, 8-15 weeks	32.54(16.29) 41.54(7.05) 9.00	Level of Evidence: III/class IV Study protocol: - Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: - Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Kiran, 2008	5;5; M=1; F=4	58.40(12.03) 47-77	Conduction =3 Broca's=2	Mod.=2 Sev.=3	8.2 7-10	**Type**: I **Description**: SFA-based treatment involved 1) naming the picture 2) sorting pictures by category 3) identify semantic features 4) answer yes/no feature questions **Intensity**: 2x/week, 120 min, 24 weeks	14.00(9.06) 27.00(13.56) 13.00	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Kiran, Sandberg, & Abbott, 2009	4 ; 4; M=2; F=2	56.75(15.63) 39-77	Anomic =4	Mild=3 Sev. =1	25.5 8-43	**Type**: I **Description**: SFA-based treatment involving: 1) category sorting 2) feature selection 3) yes/no feature questions 4) word recall and 5) free generative naming. **Intensity**: 2x/week, 120 min, avg. 12 weeks (6-19)	36.51(15.48) 37.50(11.62) .99	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Kiran, Sandberg, & Sebastian, 2011	6 ;6; M=3; F=3	68.00(15.76) 39-84	Anomic=4 Conduction=2	Mild=2 Mod.=1 Sev.=3	43.16 6-108	**Type**: I **Description**: SFA-based treatment involving 1) category generation 2) category sorting 3) feature generation and/or selection and 4) answering yes/no feature questions **Intensity**: 2x/week, 120 min, 10 weeks	23.50(12.21) 24.16(13.02) .66	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Kurland, Wilkins, & Stokes, 2014	8;5; M=2; F=3	67.60(8.26) 58-80	Anomia = 3 Transcortical sensory =1 Wernicke =1	Mild = 4 Mod. = 1	44 17-84	**Type**: INT **Description**: Home practice implemented after two weeks of intensive language therapy (ILAT) or modified version of Promoting Aphasic Communicative Effectiveness (PACE). Each participant received two individualized iBook (i.e., objects and actions) to practice at home. Daily practice involved 20 words. Each word had a chapter with five interactive pages in the iBook targeting it. They also met with the SLP once weekly for informal conversation and trouble-shooting. **Intensity**: 5-6x/week, 20 min, 26 weeks	39.60(10.24) 37.40(11.72) -2.20	Level of Evidence: IIB/class III Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Lacey, Lott, Snider, Sperling, & Friedman, 2010	6; 3; F=3	56.00(15.39) 39-69	Anomic = 3	Mild = 1 Mod. = 2	54.33 13-114	**Type**: I **Description**: Multiple Oral Re-reading Treatment: 1) Read text passages three times aloud 2) Re-read words on which they made mistakes. Clinician read the word aloud for them if they could not. PWA re-read the whole sentence if they made so many errors that the sentence flow was disrupted. They were instructed to call the SLP daily and read the passage aloud over the phone. They also read it 5 times at home daily without assistance. **Intensity**: 1x/week, 60 min, 8 weeks	26.33(7.51) 33.67(3.79) 7.33	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
MacGregor, Difrancesco, Pulvermüller, Shtyrov, & Mohr, 2015	12;12; M=9; F = 3	57.00(15.64) 26-76	Non-fluent	Mod.	81.58 17-234	**Type**: A/P **Description**: Intensive Language Action Therapy (ILAT): Treatment involved language games (i.e., making requests) to improve their language and communication. Treatment stimuli included cards depicting scenes/objects. Verbal expression was encouraged and non-verbal communication was discouraged. **Intensity**: 1x/day, 180-240 min, 10 days	28.58(4.86) 33.00(4.22) 4.42	Level of Evidence: IIB/class III Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Milman, Clendenen, & Vega-Mendoza, 2014	3;3; M=3	N/A(N/A) 56-68	Broca's = 3	Sev.	12-84	**Type**: INT **Description**: Functional use of adjectives to describe people in four different tasks: 1) single-word adjective production 2) single-word pronoun production 3) sentence training and 4) discourse production. Semantic, orthographic and phonemic cues were given to facilitate single word use. **Intensity**: 4x/week, 60 min, avg. 11 weeks (6-18)	16.33(11.37) 22.67(8.50) 6.33	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Mohr, Difrancesco, Harrington, Evans, & Pulvermüller, 2014	8; 8; M=7; F = 1	62.38(12.75) 41-76	Nonfluent	Mild = 3 Mod. = 1 Sev. = 4	87 17-234	**Type**: A/P **Description**: Intensive Language Action Therapy (ILAT): Treatment involved language games (i.e., making requests) to improve their language and communication. Treatment stimuli included cards depicting scenes/objects. Verbal expression was encouraged and non-verbal communication was discouraged. **Intensity**: 1x/day, 180 min, 10 days	25.38(16.69) 29.75(13.31) 4.38	Level of Evidence: IIB/class III Study protocol: - Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
Nettleton & Lesser, 1991	6; 6; F=4; M=2	64.83(7.83) 55-74	Fluent = 4 Anomic = 1 Non-fluent = 1	Mod. = 1 Sev. = 5	47 6-96	**Type**: I **Description**: Semantic therapy involved word-picture matching, yes/no feature judgments and category sorting. Phonological therapy involved repetition of picture name, rhyme judgment and naming with progressive phonemic cues. **Intensity**: 2x/week, 60 min, 8 weeks	12.20(6.84) 14.50(6.26) 2.30	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Nickels & Osborne, 2016	4;4; M= 3; F= 1	59.75(18.66) 34-74	Global= 2 Anomic= 1 TCS= 1	Mod. = 3 Sev. = 1	25.8 15-42	**Type**: A/P **Description**: Constraint Induced Aphasia Therapy Plus (CIAT-plus): Therapy addressed verbal expression through Go Fish. 1) Volunteer played the game with two PWA. 2) PWA chose a card and asked the other players for a card. 3) Co-player then responded. PWA could use multi-modal communication, as needed. Shaping was included to increase the complexity of their verbal responses. **Intensity**: 2x/week, 90 min, 4 weeks	23.25(6.99) 25.25(8.18) 2.00	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
Raymer, Kohen, & Saffell, 2006a	5; 4; M=2; F=3	70.80(12.11) 51-82	Broca's=2 Conduction = 2 Mixed = 1	Mod. = 1 Sev. = 4	18.4 4-42	**Type**: I **Description**: MossTalk Words (i.e., computer-assisted treatment program). PWA completed multi-modal matching exercises involving 1) spoken plus written word to picture matching 2) spoken word to picture matching 3) written word to picture matching **Intensity**: 1-2x/week, 60 min, 6-12 weeks, then, 3-4x/week, 60 min 3-4 weeks. 4 week break in between each 12-hour tx. phase.	6.80(9.98) 7.80(7.92) 1.00	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: + Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Raymer, Singletary, et al., 2006b	9; 9; M=6; F=3	60.67(9.08) 49-70	Broca's = 6 Wernicke's = 2 Conduction =1	Mod. = 3 Sev. = 6	29 5-62	**Type**: I **Description**: Gesture-Verbal Treatment (GVT): 1) SLP showed the picture and modeled the target word and a gesture. 2) PWA produced word and gesture three times 3) SLP showed gesture in isolation and participant imitated three times 4) SLP presented the target and PWA repeated it three times 4) After a 5-second delay, SLP prompted participant to show and tell them what happened in the picture. **Intensity**: 3-4x/week, 60 min, 10 weeks	11.33(11.31) 12.55(14.11) 1.22	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Raymer et al., 2012	8;8; M= 4; F= 4	58.13(14.30) 40-79	Broca's = 4 TMA = 2 TSA = 1 Wernicke's = 1	Mod.= 4 Sev. = 4	13.5 5-30	**Type**: I **Description**: Errorless Naming: 1) SLP modelled the picture name and PWA repeated 2)SLP showed the written word and PWA read it aloud three times 3)Written word was removed and PWA was given 5 seconds to hold onto it 4) SLP prompted PWA to name it again. Gestural Facilitation: 1) SLP modelled the name and a related gesture 2) SLP modelled the gesture alone for SLP to imitate three times 3) SLP modelled name and PWA repeated three times 4) Clinician modelled gesture while showing the picture 5) After 5 second delay SLP prompted PWA to provide name and gesture again. **Intensity**: 2-3x/week, 60 min, 10 weeks	14.63(8.31) 14.50(6.78) -.13	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Rider, Wright, Marshall, & Page, 2008	3; 3; M=2; F =1	63.33(9.07) 55-73	Nonfluent =3	Mild = 1 Mod. = 2	65.67 26-126	**Type**: I **Description**: Trained words related to 6-8 contexts (i.e., story retell and procedural narratives) using SFA (e.g., Boyle, 2004 Boyle & Coelho, 1995) **Intensity**: 2-3x/week, 60 min, 7-14 weeks	30.33(11.06) 34.00(10.15) 3.67	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
Rodriguez, Raymer, & Rothi, 2006	4 ;4; M=3; F=1	65.00(9.76) 52-73	Conduction = 2 Wernicke's = 1 Broca's =1	Mild =1 Sev. =3	34.25 8-96	**Type**: I **Description**: All PWA received both verb naming treatments. Gesture-Verbal Treatment (GVT): 1) SLP showed the picture and modeled the target word and a gesture. 2) PWA produced word and gesture three times 3) SLP showed gesture in isolation and PWA imitated three times 4) SLP presented the target and PWA repeated it three times 4) After a 5-second delay, SLP prompted participant to show and tell them what happened in the target picture. Semantic-Phonologic Treatment: 1) SLP showed PWA the picture and modeled the target word 2) PWA answered semantic and phonologic questions about the target 3) PWA produced the target three times 4) After a 5-second delay, PWA attempted to explain what was happening in the picture. **Intensity**: 2-3/week, 60 min, 10-14 weeks	12.50(19.00) 14.75(23.08) 2.25	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Rodriguez et al., 2013	11;11; M=4; F=4	60.45(17.67) 18-79	N/A	Mod.	25.81 8-56	**Type**: INT **Description**: Individual treatment involved both impairment-based and functional therapy. PWA and family members were involved in group treatment (i.e., share information about available local services, facilitate discussions about "living with aphasia" promoting social interaction and multi-modal communication) Computer-based therapy (i.e., Bungalow, REACT, Speech Sounds on Cue) Challenge Task: specific goal each PWA wanted to achieve by the end of the program. **Intensity**: 5x/week, 240 min, 2 weeks OR 5x/week, 300 min, 4 weeks	20.10(19.10) 22.50(21.30) 2.40	Level of Evidence: IIB/class III Study protocol: + Blinding: + Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
Rose, Attard, Mok, Lanyon, & Foster, 2013	11;11; M= 6; F = 5	58.09(10.63) 39-74	Broca's = 6 Anomic = 4, Conduction = 1	Above =1 Mild = 1 Mod. = 4 Sev. = 5	44 17-88	**Type**: A/P **Description**: PWA targeted word retrieval in small groups through treatment activities including (i.e., Go Fish, Memory, Request Role plays, Board games, rapid naming while playing snap, Who am I) In CIAT Plus: Verbal production was the goal but cueing was provided as needed (i.e., phonemic cue, written cue). In multi-modal aphasia therapy (M-MAT): Verbal production was also the goal but, multi-modal cueing was provided (i.e., gesture, drawing, written model, verbal model). All PWA received both treatments. **Intensity**: 4x/week, 195 min, 4 weeks	20.18(16.28) 27.64(18.92) 7.45	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: - Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
Sandberg, Bohland, & Kiran, 2015	10; 10; M=7; F=3	59.40(10.01) 47-75	Anomic=6 2 Conduction=2 Broca's=1 TCM=1	Above =6 Mild= 3 Mod.=1	55.7 7-134	**Type**: I **Description**: PWA were trained on ten abstract words in a particular context category (e.g., courthouse) and ten untrained concrete words from the same context-category were monitored to measure generalization. Treatment steps included 1) Feature selection 2) Abstract/concrete lexical decision 3) Synonym generation **Intensity**: 2x/week, 120 min, 10 weeks	47.34(13.40) 49.80(13.57) 2.46	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Schwartz, Saffran, Fink, Myers, & Martin, 1994	8;6; M=4; F=2	60.00(8.37) 46-70	Nonfluent = 5	Above = 1 Mod. = 3 Sev. = 2	4.91 59-102	**Type**: I **Description**: Mapping Therapy: 1) PWA read a sentence aloud and then, was given assistance if needed by the SLP. 2) They were asked to identify the verb, the agent and the patient/theme. 3) PWA would underline the verb and head noun in the noun phrases. **Intensity**: 3x/week, 60-90 min, 17.38 weeks	24.33(14.25) 32.17(15.20) 7.83	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: + Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
Silkes, 2015	4; 4; N/A	60.25(1.26) 59-62	Fluent=3 Nonfluent=1	Mild = 1 Mod. = 1 Sev. = 2	58.5 24-96	**Type**: I **Description**: Masked repetition priming treatment: Each section PWA saw prime-picture pair 16 times and had four opportunities to name each picture. PWA were instructed to watch the screen and try to name the picture when they saw it for the 4th time. **Intensity**: 2x/day, 12 days	24.00(13.44) 26.50(14.55) 2.50	Level of Evidence: III/class IV Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: N/A
van Hees, Angwin, McMahon, & Copland, 2013	8; 8; F=5; M=3	56.38(9.15) 41-69	Anomic=6 Conduction=2	Mild	52.25 17-170	**Type**: I **Description**: Semantic Feature Analysis (SFA): 1) Name the item 2) produced semantic features (i.e., use, action, properties, location and association) 3) Name the item again Phonological Components Analysis (PCA) 1) Name the item 2) Produced phonological features (i.e., first sound, syllables, last sound, association and rhyme) **Intensity**: 3x/week, 60-90 min, 4 weeks	38.50(15.44) 43.50(14.04) 5.00	Level of Evidence: IIB/class III Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: + Significance for st. outcome of interest: - Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Votruba, Rapport, Whitman, Johnson, & Langenecker, 2013	50; 50; M=28; F=22	56.80(15.20) 20-85	N/A	Mod.	43.6 N/A	**Type**: INT **Description**: Outpatient speech language therapy (SLT): All PWA received mostly individual SLT and adjunct group SLT. Individual SLT addressed expressive language (100%), writing (92%), comprehension/reading/naming (83%), spelling (50%), repetition/prosody (8%). **Intensity**: Outpatient SLT university clinic schedule	25.70(18.60) 26.90(19.80) 1.20	Level of Evidence: IIB/class III Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: + Significance for st. outcome of interest: - Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Supplementary Material 7: Demographic information for between group studies using Western Aphasia Battery – Aphasia Quotient

Note: M=Male; F = Female; TCM = transcortical motor; TSM = transcortical sensory; sev. = severe; I = Impairment-based treatment; A/P = activity/participation-based treatment; INT = integrated treatment; st. = standardized; CS = convenience sample

Study Name	N; Sex	Mean Age (SD) Range	Mean MPO range	Treatment (Tx)	Pre-Tx M(SD) Post-Tx M(SD) Change Score	Mean Age(SD) Range	Mean MPO range	Treatment (Tx)	Pre-Tx M(SD) Post-Tx M(SD) Change Score	Methodological Rigor
Altmann et al., 2014	G=7; M=2; F=5 NG=7 M=6; F=1	72.14 (10.51) 62-92	6+ N/A	**Type**: I **Description**: Phase 1 and Phase 2: present pictures for naming Phase 3: PWA was presented with auditory/orthographic representations of category name and PWA generated a category member. Clinicians provided feedback and/or corrected them. Treatment started with L hand opening, reaching into a box to press a red button and making a non-meaningful circular gesture with Left hand during correction phase. **Intensity**: 5x/week, 120 min, 3 weeks	65.47 (8.34) 67.09 (9.09) 1.62	63.00 (9.22) 53-80	6 MPO+ N/A	**Type**: I **Description**: Phase 1 and Phase 2: present pictures for naming Phase 3: PWA was presented with auditory/orthographic representations of category name and PWA generated a category member. Clinicians provided feedback and/or corrected them. Treatment started with therapist pressing a button. There was no hand movement during the correction procedure **Intensity**: 5x/week, 120 min, 3 weeks	71.91 (11.80) 72.89 (14.50) 0.98	Level of Evidence: IIA/class II Study protocol: + Blinding: - Sampling/allocation:+ Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: +
Des Roches, Balachandran, Ascenso, Tripodis, & Kiran, 2015	E=40; N/A C= 9; M=7; F=2	62.98 (10.68) 38-83	53.3 1-178	**Type**: I **Description**: Participants completed a variety (between 2-11) of 37 different cognitive or language tasks in Constant Therapy (e.g., naming, rhyming, memory, symbol matching, etc.) **Intensity**: 5x/week, ~45 min, 10 weeks	68.90 (25.70) 72.90 (23.20) 3.91	67.11 (9.98) 53-87	98 13-359	**Type**: I **Description**: Participants completed a variety (between 2-14) of 37 different cognitive or language tasks in Constant Therapy (e.g., naming, rhyming, memory, symbol matching, etc.) **Intensity**: 1x/week, 40 min, 10 weeks	67.70 (31.70) 68.90 (33.60) 1.12	Level of Evidence: IIA/ class II Study protocol: + Blinding: - Sampling/allocation:+ Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest:+ Precision: + Intention to treat: +
Godecke, Hird, Lalor, Rai, & Phillips, 2012	DT=32; M=14; F=18 UC=27; M=15; F=12	70.30 (12.8)	6.1 DPO N/A	**Type**: I **Description**: Participants were administered impairment-based therapies (i.e., lexical-semantic (BOX), mapping, semantic feature analysis (SFA)) **Intensity**: 5x/week, 30-80 min, 4 weeks	33.78 (26.37) 56.42 (30.76) 22.64	67.70 (15.40)	3.4 DPO N/A	**Type**: I **Description**: Participants were administered impairment-based therapies (i.e., lexical-semantic (BOX), mapping, semantic feature analysis (SFA)). **Intensity**: 1x/week, 80 min, 4 weeks	20.46 (26.11) 34.12 (33.22) 13.66	Level of Evidence: IIB/ class III Study protocol: + Blinding: + Sampling/allocation:+ Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: +

ACCEPTED MANUSCRIPT

										Precision: + Intention to treat: +
Godecke et al., 2014	VER=20; M=12; F=8 UC=27; M=15; F=12	70.70 (14.30)	3.4 DPO N/A	**Type**: INT: **Description**: PWA received either individual or group therapy. Individual therapy consisted of Semantic Feature Analysis (SFA), Cued Naming Therapy, Lexical-Semantic (BOX), Mapping therapy and/or Phonological Feature Therapy. Group therapy consisted of Constraint Induced Aphasia Therapy (CIAT). 5x/week, 180-240 min, 4 weeks	43.53 (27.02) 67.55 (30.16) 24.02	67.7 (15.4)	3.2 DPO N/A	**Type**: I **Description**: 85% of participants did not receive direct speech and language therapy. When participants received therapy, it consisted of BOX therapy, Mapping therapy and Semantic Feature Analysis (SFA). **Intensity**: 11min, ~ 3 weeks	19.62 (26.26) 32.83 (45.62) 11.75	Level of Evidence: IIB/ class III Study protocol: + Blinding: + Sampling/allocation:CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: +
Katz & Wertz, 1997	CRT=21; N/A CS=29; N/A	61.60 (10.00) 48-83	74.4 12- 228	**Type**: I **Description**: Visual matching and reading comprehension software consisted of 10 matching activities (e.g., letters and words) and 22 reading comprehension activities (e.g., letters, words, phrases). The tasks varied in complexity (i.e., 8 difficulty levels). **Intensity**: 3x/week, 60 min, 26 weeks	68.90 (24.30) 73.60 (22.60) 4.70	66.40 (6.00) 53-76	64.80 21.6-228	**Type**: N/A **Description**: Computer stimulation software included games (e.g., Mini Putt) and cognitive rehabilitation tasks (e.g., Captain's Log). **Intensity**: 3x/week, 60 min, 26 weeks	61.90 (29.50) 63.40 (28.50) 1.50	Level of Evidence: IB/class I Study protocol: + Blinding: + Sampling/allocation:CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: +
Maher et al., 2006	CILT=4; M=3; F=1 PACE=5; M=3; F=2	48.25 (6.99) 40-55	38.75 24-48	**Type**: A/P **Description**: Constraint Induced Language Therapy (CILT): Treatment was given in groups of two or three with two clinicians also participating in the group. PWA were constrained to verbal expression only. Multi-modal communication was restricted and a barrier was placed between PWA to further reduce it. They played a card game in which they had to ask another participant for a card with the attempt of matching a card in their own hand. Speakers took turns and responses were shaped to increase complexity over time. **Intensity**: 4x/week, 180 min, 2 weeks	58.55 (12.99) 65.08 (10.22) 6.52	59.00 (12.81) 41-73	35.4 14-72	**Type**: A/P **Description**: Treatment used a modified Promoting Aphasics' Communication Effectiveness (PACE) approach. PWA could use multi-modal communication to perform the task. Intervention was provided to improve any incorrect responses regardless of the modality used. **Intensity**: 4x/week, 180 min, 2 weeks	53.94 (13.21) 56.90 (13.93) 2.96	Level of Evidence: IIA/ class II Study protocol: + Blinding: - Sampling/allocation:CS Treatment fidelity: + Significance of primary outcome measure: - Significance for st. outcome of interest: - Precision: + Intention to treat: +

ACCEPTED MANUSCRIPT

Supplementary Material 8: Demographic information for between group studies using Boston Naming Test

Note: M=Male; F = Female; TCM = transcortical motor; TSM = transcortical sensory; sev. = severe; I = Impairment-based treatment; A/P = activity/participation-based treatment; INT = integrated treatment; st. = standardized; CS = convenience sample

Study Name	N; Sex	Mean Age (SD) Range	Mean MPO range	Treatment (Tx)	Pre-Tx M(SD) Post-Tx M(SD) Change Score	Mean Age(SD) Range	Mean MPO range	Treatment (Tx)	Pre-Tx M(SD) Post-Tx M(SD) Change Score	Methodological Rigor
Altmann et al., 2014	G=7; F=5; M=2 NG=7; M=6; F=1	72.10 (10.50) 62-92	N/A 6+	**Type**: I **Description**: Phase 1 and Phase 2: present pictures for naming Phase 3: PWA was presented with auditory/orthographic representations of category name and PWA generated a category member. Clinicians provided feedback and/or corrected them. Treatment started with L hand opening, reaching into a box to press a red button and making a non-meaningful circular gesture with Left hand during correction phase. **Intensity**: 5x/week, 120 min, 3 weeks	24.71 (13.44) 28.57 (16.07) 3.86	63.00 (9.20) 53-80	N/A 6+	**Type**: I **Description**: Phase 1 and Phase 2: present pictures for naming Phase 3: PWA was presented with auditory/orthographic representations of category name and PWA generated a category member. Clinicians provided feedback and/or corrected them. Treatment started with therapist pressing a button. There was no hand movement during the correction procedure **Intensity**: 5x/week, 120 min, 3 weeks	30.86 (6.26) 33.86 (9.56) 3.00	Level of Evidence: IIA/class II Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A
Des Roches, Balachandran, Ascenso, Tripodis, & Kiran, 2015	E=40; N/A C= 9; M=7; F=2	62.98 (10.68) 38-83	53.3 1-178	**Type**: I **Description**: Participants completed a variety (between 2-11) of 37 different cognitive or language tasks in Constant Therapy (e.g., naming, rhyming, memory, symbol matching, etc.) **Intensity**: 5x/week, ~45 min, 10 weeks	28.29 (22.33) 29.45 (21.25) 1.16	67.11 (9.98) 53-87	98 13-359	**Type**: I **Description**: Participants completed a variety (between 2-14) of 37 different cognitive or language tasks in Constant Therapy (e.g., naming, rhyming, memory, symbol matching, etc.) **Intensity**: 1x/week, 40 min, 10 weeks	26.66 (24.40) 26.23 (23.86) -.43	Level of Evidence: IIA/ class II Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: N/A

ACCEPTED MANUSCRIPT

Maher et al., 2006	CILT=4; M=3; F=1 PACE=5; M=3; F=2	48.25 (6.99) 40-55	38.752 4-48	**Type**: A/P **Description**: Constraint Induced Language Therapy (CILT): Treatment was given in groups of two or three with two clinicians also participating in the group. PWA were constrained to verbal expression only. Multi-modal communication was restricted and a barrier was placed between PWA to further reduce it. They played a card game in which they had to ask another participant for a card with the attempt of matching a card in their own hand. Speakers took turns and responses were shaped to increase complexity over time. **Intensity**: 4x/week, 180 min, 2 weeks	18.00 (16.47) 21.00 (17.32) 3.00	59.00 (12.81) 41-73	35.4 14-72	**Type**: A/P **Description**: Treatment used a modified Promoting Aphasics' Communication Effectiveness (PACE) approach. PWA could use multi-modal communication to perform the task. Intervention was provided to improve any incorrect responses regardless of the modality used. **Intensity**: 4x/week, 180 min, 2 weeks	15.20 (19.64) 18.60 (21.98) 3.40	Level of Evidence: IIA/ class II Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: -
Raglio et al., 2016	E=10; M=7; F=3 C=10 M=7; F=3	61.30 (12.76) 42-89	chronic	**Type**: INT **Description**: Music therapy + Speech Therapy: Treatment involved a Promoting Aphasics' Communicative Effectiveness (PACE) approach with the addition of music therapy (i.e., play instruments sing with therapist) **Intensity**: 2x/week, 75 min, 15 weeks	23.00 (21.00) 26.00 (21.00) 3.00	70.90 (8.99) 61-89	chronic	**Type**: A/P **Description**: Treatment consisted of a Promoting Aphasics' Communicative Effectiveness (PACE) approach only. **Intensity**: 2x/week, 45 mins, 15 weeks	16 (19) 19 (19) 3.00	Level of Evidence: IB/class I Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: - Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: +
Wilssens et al., 2015	CIAT=5; N/A BOX=4; N/A	63.00 (8.00)	61 N/A	**Type**: INT **Description**: Constraint Induced Aphasia Therapy (CIAT): PWA participated in a communication-based group with card games (e.g., "Go Fish"). Verbal expression was encouraged in the game context. They were allowed to produce gestures when communicate, but these were hidden from view of other PWA with a screen. **Intensity**: daily, 120-180 min, 9-10 consecutive working days	30.20 (14.00) 39.8 (13.8) 9.6	71.00 (9.00) 60-81	52	**Type**: I **Description**: PWA were administered BOX, a treatment focused on semantic processing using written words, sentences and longer texts. **Intensity**: daily, 120-180 min, 9-10 consecutive working days	29.00 (20.10) 39.8 (13.9) 10.8	Level of Evidence: IIA/ class II Study protocol: + Blinding: - Sampling/allocation: CS Treatment fidelity: + Significance of primary outcome measure: + Significance for st. outcome of interest: + Precision: + Intention to treat: No attrition: N/A

ACCEPTED MANUSCRIPT

Supplementary Material 9: References for all included studies

Included Within Group Studies

Aftonomos, L. B., Appelbaum, J. S., & Steele, R. D. (1999). Improving outcomes for persons with aphasia in advanced community-based treatment programs. *Stroke; a Journal of Cerebral Circulation*, *30*(7), 1370–1379.

Aftonomos, L. B., Steele, R. D., & Wertz, R. T. (1997). Promoting recovery in chronic aphasia with an interactive technology. *Archives of Physical Medicine and Rehabilitation*, *78*(8), 841–846.

Archibald, L. M. D., Orange, J. B., & Jamieson, D. J. (2009). Implementation of computer-based language therapy in aphasia. *Therapeutic Advances in Neurological Disorders*, *2*(5), 299–311. https://doi.org/10.1177/1756285609336548

Babbitt, E., Worrall, L., & Cherney, L. R. (2015). Structure, Processes, and Retrospective Outcomes From an Intensive Comprehensive Aphasia Program. *American Journal of Speech-Language Pathology / American Speech-Language-Hearing Association*, *24*(4), 854–63.

Bakheit, A. M. O., Carrington, S., Griffiths, S., & Searle, K. (2005). High scores on the Western Aphasia Battery correlate with good functional communication skills (as measured with the Communicative Effectiveness Index) in aphasic stroke patients. *Disability and Rehabilitation: An International, Multidisciplinary Journal*, *27*(6), 287–291. https://doi.org/10.1080/09638280400009006

Ball, A. L., de Riesthal, M., Breeding, V. E., & Mendoza, D. E. (2011). Modified ACT and CART in severe aphasia. *Aphasiology*, *25*(6–7), 836–848. https://doi.org/10.1080/02687038.2010.544320

Barthel, G., Meinzer, M., Djundja, D., & Rockstroh, B. (2008). Intensive language therapy in chronic aphasia: Which aspects contribute most? *Aphasiology*, *22*(4), 408–421. https://doi.org/10.1080/02687030701415880

Beeson, P. M., Rising, K., & Volk, J. (2003). Writing treatment for severe aphasia: Who benefits? *Journal of Speech, Language, and Hearing Research*, *46*(5), 1038–1060. https://doi.org/10.1044/1092-4388(2003/083)

Boles, L. (1997). Conversation analysis as a dependent measure in communication therapy with four individuals with aphasia. *Asia Pacific Journal of Speech, Language & Hearing*, *2*(1), 43–61 19p.

Breier, J. I., Maher, L. M., Novak, B., & Papanicolaou, A. C. (2006). Functional Imaging Before and After Constraint-Induced Language Therapy for Aphasia Using Magnetoencephalography. *Neurocase*, *12*(6), 322–331. https://doi.org/10.1080/13554790601126054

Brown, J. W., & Chobor, K. L. (1989). Therapy with a prosthesis for writing in aphasia. *Aphasiology*, *3*(8), 709–715.

Cherney, L. R. (2010). Oral reading for language in aphasia: Impact of aphasia severity on cross-modal outcomes in chronic nonfluent aphasia. *Seminars in Speech and Language*, *31*(1), 42–51. https://doi.org/10.1055/s-0029-1244952

Cherney, L. R., & Halper, A. S. (2008). Novel technology for treating individuals with aphasia and concomitant cognitive deficits. *Topics in Stroke Rehabilitation*, *15*(6), 542–554. https://doi.org/10.1310/tsr1506-542

ACCEPTED MANUSCRIPT

Cherney, L. R., Halper, A. S., Holland, A. L., & Cole, R. (2008). Computerized script training for aphasia: Preliminary results. *American Journal of Speech-Language Pathology*, *17*(1), 19–34. https://doi.org/10.1044/1058-0360(2008/003)

Code, C., Torney, A., Gildea-Howardine, E., & Willmes, K. (2010). Outcome of a one-month therapy intensive for chronic aphasia: Variable individual responses. *Seminars in Speech and Language*, *31*(1), 21–23. https://doi.org/10.1055/s-0029-1244950

Doyle, P. J., Goldstein, H., & Bourgeois, M. S. (1987). Experimental analysis of syntax training in Broca's aphasia: A generalization and social validation study. *Journal of Speech & Hearing Disorders*, *52*(2), 143–155.

Duncan, E. S., Schmah, T., & Small, S. L. (2016). Performance Variability as a Predictor of Response to Aphasia Treatment. *Neurorehabilitation and Neural Repair Neurorehabilitation and Neural Repair*.

Edmonds, L. A., & Kiran, S. (2006). Effect of semantic naming treatment on crosslinguistic generalization in bilingual aphasia. *Journal of Speech, Language, and Hearing Research*, *49*(4), 729–748. https://doi.org/10.1044/1092-4388(2006/053)

Edmonds, L. A., Nadeau, S. E., & Kiran, S. (2009). Effect of verb network strengthening treatment (VNeST) on lexical retrieval of content words in sentences in persons with aphasia. *Aphasiology*, *23*(3), 402–424. https://doi.org/10.1080/02687030802291339

Edmonds, Lisa A., Mammino, K., & Ojeda, J. (2014). Effect of Verb Network Strengthening Treatment (VNeST) in persons with aphasia: Extension and replication of previous findings. *American Journal of Speech-Language Pathology*, *23*(2), S312–S329. https://doi.org/10.1044/2014_AJSLP-13-0098

Falconer, C., & Antonucci, S. M. (2012). Use of semantic feature analysis in group discourse treatment for aphasia: Extension and expansion. *Aphasiology*, *26*(1), 64–82. https://doi.org/10.1080/02687038.2011.602390

Faroqi-Shah, Y. (2008). A comparison of two theoretically driven treatments for verb inflection deficits in aphasia. *Neuropsychologia*, *46*(13), 3088–3100. https://doi.org/10.1016/j.neuropsychologia.2008.06.018

Faroqi-Shah, Y. (2013). Selective treatment of regular versus irregular verbs in agrammatic aphasia: Efficacy data. *Aphasiology*, *27*(6), 678–705. https://doi.org/10.1080/02687038.2012.751577

Ferguson, N. F., Evans, K., & Raymer, A. M. (2012). A comparison of intention and pantomime gesture treatment for noun retrieval in people with aphasia. *American Journal of Speech-Language Pathology*, *21*(2), s126–s139. https://doi.org/10.1044/1058-0360(2012/11-0076)

Fridriksson, J., Morrow-Odom, L., Moser, D., Fridriksson, A., & Baylis, G. (2006). Neural recruitment associated with anomia treatment in aphasia. *NeuroImage*, *32*(3), 1403–1412. https://doi.org/10.1016/j.neuroimage.2006.04.194

Johnson, R. K., Hough, M. S., King, K. A., Vos, P., & Jeffs, T. (2008). Functional communication in individuals with chronic severe aphasia using augmentative communication. *AAC: Augmentative and Alternative Communication*, *24*(4), 269–280. https://doi.org/10.1080/07434610802463957

ACCEPTED MANUSCRIPT

Kendall, D. L., Oelke, M., Brookshire, C. E., & Nadeau, S. E. (2015). The Influence of Phonomotor Treatment on Word Retrieval Abilities in 26 Individuals With Chronic Aphasia: An Open Trial. *Journal of Speech, Language & Hearing Research, 58*(3), 798–812 15p. https://doi.org/10.1044/2015_JSLHR-L-14-0131

Kendall, D. L., Raymer, A., Rose, M., Gilbert, J., & Gonzalez Rothi, L. J. (2014). Anomia treatment platform as behavioral engine for use in research on physiological adjuvants to neurorehabilitation. *Journal of Rehabilitation Research and Development*, *51*(3), 391–400. https://doi.org/10.1682/JRRD.2013.08.0172

Kendall, D. L., Rosenbek, J. C., Heilman, K. M., Conway, T., Klenberg, K., Gonzalez Rothi, L. J., & Nadeau, S. E. (2008). Phoneme-based rehabilitation of anomia in aphasia. *Brain and Language*, *105*(1), 1–17. https://doi.org/10.1016/j.bandl.2007.11.007

Kiran, S. (2005). Training phoneme to grapheme conversion for patients with written and oral production deficits: A model-based approach. *Aphasiology*, *19*(1), 53–76. https://doi.org/10.1080/02687030444000633

Kiran, S. (2008). Typicality of inanimate category exemplars in aphasia treatment: Further evidence for semantic complexity. *Journal of Speech, Language, and Hearing Research*, *51*(6), 1550–1568. https://doi.org/10.1044/1092-4388(2008/07-0038)

Kiran, S., & Johnson, L. (2008). Semantic complexity in treatment of naming deficits in aphasia: Evidence from well-defined categories. *American Journal of Speech-Language Pathology*, *17*(4), 389–400.

Kiran, S., Sandberg, C., & Abbott, K. (2009). Treatment for lexical retrieval using abstract and concrete words in persons with aphasia: Effect of complexity. *Aphasiology*, *23*(7–8), 835–853. https://doi.org/10.1080/02687030802588866

Kiran, S., Sandberg, C., & Sebastian, R. (2011). Treatment of Category Generation and Retrieval in Aphasia: Effect of Typicality of Category Items. *Journal of Speech Language and Hearing Research*, *54*(4), 1101. https://doi.org/10.1044/1092-4388(2010/10-0117)

Kiran, S., & Thompson, C. K. (2003). The Role of Semantic Complexity in Treatment of Naming Deficits Training Semantic Categories in Fluent Aphasia by Controlling Exemplar Typicality. *Journal of Speech, Language, and Hearing Research*, *46*(4), 773–787.

Kurland, J., Wilkins, A. R., & Stokes, P. (2014). iPractice: Piloting the effectiveness of a tablet-based home practice program in aphasia treatment. *Seminars in Speech and Language*, *35*(1), 51–63.

Lacey, E. H., Lott, S. N., Snider, S. F., Sperling, A., & Friedman, R. B. (2010). Multiple Oral Re-reading treatment for alexia: The parts may be greater than the whole. *Neuropsychological Rehabilitation*, *20*(4), 601–623.

Lesser, R., Bryan, K., Anderson, J., & Hilton, R. (1986). Involving relatives in aphasia therapy: An application of language enrichment therapy. *International Journal of Rehabilitation Research*, *9*(3), 259–267. https://doi.org/10.1097/00004356-198609000-00005

Macauley, B. L. (2006). Animal-assisted therapy for persons with aphasia: A pilot study. *Journal of Rehabilitation Research and Development*, *43*(3), 357–366.

ACCEPTED MANUSCRIPT

MacGregor, L. J., Difrancesco, S., Pulvermüller, F., Shtyrov, Y., & Mohr, B. (2015). Ultra-rapid access to words in chronic aphasia: The effects of intensive language action therapy (ILAT). *Brain Topography*, *28*(2), 279–291. https://doi.org/10.1007/s10548-014-0398-y

Marshall, R. S., Laures-Gore, J., DuBay, M., Williams, T., & Bryant, D. (2015). Unilateral forced nostril breathing and aphasia—Exploring unilateral forced nostril breathing as an adjunct to aphasia treatment: A case series. *The Journal of Alternative and Complementary Medicine*, *21*(2), 91–99. https://doi.org/10.1089/acm.2013.0285

Milman, L., Clendenen, D., & Vega-Mendoza, M. (2014). Production and integrated training of adjectives in three individuals with nonfluent aphasia. *Aphasiology*, *28*(10), 1198–1222. https://doi.org/10.1080/02687038.2014.910590

Milman, L., Vega-Mendoza, M., & Clendenen, D. (2014). Integrated training for aphasia: An application of part-whole learning to treat lexical retrieval, sentence production, and discourse-level communications in three cases of nonfluent aphasia. *American Journal of Speech-Language Pathology*, *23*(2), 105–119. https://doi.org/10.1044/2014_AJSLP-12-0054

Mohr, B., Difrancesco, S., Harrington, K., Evans, S., & Pulvermüller, F. (2014). Changes of right-hemispheric activation after constraint-induced, intensive language action therapy in chronic aphasia: fMRI evidence from auditory semantic processing. *Frontiers in Human Neuroscience*, *8*(November). https://doi.org/10.3389/fnhum.2014.00919

Mozeiko, J., Coelho, C. A., & Myers, E. B. (2016). The role of intensity in constraint-induced language therapy for people with chronic aphasia. *Aphasiology*, *30*(4), 339–363. https://doi.org/10.1080/02687038.2015.1070949

Nettleton, J., & Lesser, R. (1991). Therapy for naming difficulties in aphasia: Application of a cognitive neuropsychological model. *NEL Journal of Neurolinguistics*, *6*(2), 139–157.

Nickels, L., & Osborne, A. (2016). Constraint Induced Aphasia Therapy: Volunteer-led, unconstrained and less intense delivery can be effective. *NeuroRehabilitation*, *39*(1), 97–109. https://doi.org/10.3233/NRE-161341

Purdy, M., & Wallace, S. E. (2015). Intensive multimodal communication treatment for people with chronic aphasia. *Aphasiology Aphasiology*, 1–23.

Raymer, A. M., Kohen, F. P., & Saffell, D. (2006). Computerised training for impairments of word comprehension and retrieval in aphasia. *Aphasiology*, *20*(2–4), 257–268. https://doi.org/10.1080/02687030500473312

Raymer, A. M., McHose, B., Smith, K. G., Iman, L., Ambrose, A., & Casselton, C. (2012). Contrasting effects of errorless naming treatment and gestural facilitation for word retrieval in aphasia. *Neuropsychological Rehabilitation*, *22*(2), 235–266. https://doi.org/10.1080/09602011.2011.618306

Raymer, A. M., Singletary, F., Rodriguez, A., Ciampitti, M., Heilman, K. M., & Rothi, L. J. G. (2006). Effects of gesture+verbal treatment for noun and verb retrieval in aphasia. *Journal of the International Neuropsychological Society*, *12*(6), 867–882. https://doi.org/10.1017/S1355617706061042

ACCEPTED MANUSCRIPT

Rider, J. D., Wright, H. H., Marshall, R. C., & Page, J. L. (2008). Using semantic feature analysis to improve contextual discourse in adults with aphasia. *American Journal of Speech-Language Pathology*, *17*(2), 161–172. https://doi.org/10.1044/1058-0360(2008/016)

Rodriguez, A. D., Raymer, A. M., & Rothi, L. J. G. (2006). Effects of gesture+verbal and semantic-phonologic treatments for verb retrieval in aphasia. *Aphasiology*, *20*(2–4), 286–297. https://doi.org/10.1080/02687030500474898

Rodriguez, A. D., Worrall, L., Brown, K., Grohn, B., McKinnon, E., Pearson, C., … Copland, D. A. (2013). Aphasia LIFT: Exploratory investigation of an intensive comprehensive aphasia programme. *Aphasiology*, *27*(11), 1339–1361. https://doi.org/10.1080/02687038.2013.825759

Rose, M. L., Attard, M. C., Mok, Z., Lanyon, L. E., & Foster, A. M. (2013). Multi-modality aphasia therapy is as efficacious as a constraint-induced aphasia therapy for chronic aphasia: A phase 1 study. *Aphasiology*, *27*(8), 938–971. https://doi.org/10.1080/02687038.2013.810329

Sandberg, C. W., Bohland, J. W., & Kiran, S. (2015). Changes in functional connectivity related to direct training and generalization effects of a word finding treatment in chronic aphasia. *Brain and Language*, *150*, 103–116. https://doi.org/10.1016/j.bandl.2015.09.002

Schneider, S., & Thompson, C. (2003). Verb production in agrammatic aphasia: The influence of semantic class and argument structure properties on generalisation. *Aphasiology*, *17*(3), 213–241.

Schwartz, M. F., Saffran, E. M., Fink, R. B., Myers, J. L., & Martin, N. (1994). Mapping therapy: A treatment programme for agrammatism. *Aphasiology*, *8*(1), 19–54. https://doi.org/10.1080/02687039408248639

Silkes, J. P. (2015). Masked Repetition Priming in Treatment of Anomia: A Phase 2 Study. *American Journal of Speech-Language Pathology*, *24*(4), S895–S912 18p. https://doi.org/10.1044/2015_AJSLP-14-0138

Sorin-Peters, R., & Behrmann, M. (1995). Change in perception of communication abilities of aphasic patients and their families. *Aphasiology*, *9*(6), 565–575. https://doi.org/10.1080/02687039508248715

Steele, R. D., Baird, A., McCall, D., & Haynes, L. (2014). Combining Teletherapy and On-line Language Exercises in the Treatment of Chronic Aphasia: An Outcome Study. *International Journal of Telerehabilitation*, *6*(2), 3–20. https://doi.org/10.5195/ijt.2014.6157

Thompson, C. K., Shapiro, L., Kiran, S., & Sobecks, J. (2003). The role of syntactic complexity in treatment of sentence deficits in agrammatic aphasia: The complexity account of treatment efficacy (CATE). *Journal of Speech, Language, and Hearing Research*, *46*(3), 591–607. https://doi.org/10.1044/1092-4388(2003/047)

van der Gaag, A., Smith, L., Davis, S., Moss, B., Cornelius, V., Laing, S., & Mowles, C. (2005). Therapy and support services for people with long-term stroke and aphasia and their relatives: A six-month follow-up study. *Clinical Rehabilitation*, *19*(4), 372–380. https://doi.org/10.1191/0269215505cr785oa

van Hees, S., Angwin, A., McMahon, K., & Copland, D. (2013). A comparison of semantic feature analysis and phonological components analysis for the treatment of naming impairments in aphasia. *Neuropsychological Rehabilitation*, *23*(1), 102–132. https://doi.org/10.1080/09602011.2012.726201

ACCEPTED MANUSCRIPT

Votruba, K. L., Rapport, L. J., Whitman, R. D., Johnson, A., & Langenecker, S. (2013). Personality differences among patients with chronic aphasia predict improvement in speech-language therapy. *Topics in Stroke Rehabilitation*, *20*(5), 421–431. https://doi.org/10.1310/tsr2005-421

Waller, A., Dennis, F., Brodie, J., & Cairns, A. Y. (1998). Evaluating the use of TalksBac, a predictive communication device for nonfluent adults with aphasia. *International Journal of Language & Communication Disorders*, *33*(1), 45–70. https://doi.org/10.1080/136828298247929

Wambaugh, J. L., Wright, S., & Nessler, C. (2012). Modified response elaboration training: A systematic extension with replications. *Aphasiology*, *26*(12), 1407–1439. https://doi.org/10.1080/02687038.2012.702887

Wilson, K. R., O'Rourke, H., Wozniak, L. A., Kostopoulos, E., Marchand, Y., & Newman, A. J. (2012). Changes in N400 topography following intensive speech language therapy for individuals with aphasia. *Brain and Language*, *123*(2), 94–103. https://doi.org/10.1016/j.bandl.2012.06.005

Included Between Group Studies

Altmann, L. J. P., Hazamy, A. A., Carvajal, P. J., Benjamin, M., Rosenbek, J. C., & Crosson, B. (2014). Delayed stimulus-specific improvements in discourse following anomia treatment using an intentional gesture. *Journal of Speech, Language, and Hearing Research*, *57*(2), 439–454. https://doi.org/10.1044/1092-4388(2013/12-0224)

Des Roches, C. A., Balachandran, I., Ascenso, E. M., Tripodis, Y., & Kiran, S. (2015a). Effectiveness of an impairment-based individualized rehabilitation program using an iPad-based software platform. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.01015

Des Roches, C. A., Balachandran, I., Ascenso, E. M., Tripodis, Y., & Kiran, S. (2015b). Effectiveness of an impairment-based individualized rehabilitation program using an iPad-based software platform. *Frontiers in Human Neuroscience*, *8*. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2015-26028-001&site=ehost-live&scope=site

Godecke, E., Ciccone, N. A., Granger, A. S., Rai, T., West, D., Cream, A., … 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 13p. https://doi.org/10.1111/1460-6984.12074

Godecke, E., Hird, K., Lalor, E. E., Rai, T., & Phillips, M. R. (2012). Very early poststroke aphasia therapy: a pilot randomized controlled efficacy trial. *International Journal of Stroke: Official Journal of the International Stroke Society*, *7*(8), 635–644. https://doi.org/10.1111/j.1747-4949.2011.00631.x

Katz, R. C., & Wertz, R. T. (1997). The efficacy of computer-provided reading treatment for chronic aphasic adults. *Journal of Speech & Hearing Research*, *40*(3), 493–507.

Maher, L. M., Kendall, D., Swearengin, J. A., Rodriguez, A., Leon, S. A., Pingel, K., … Rothi, L. J. G. (2006). A pilot study of use-dependent learning in the context of Constraint Induced Language Therapy. *Journal of the International Neuropsychological Society*, *12*(6), 843–852. https://doi.org/10.1017/S1355617706061029

ACCEPTED MANUSCRIPT

Raglio, A., Oasi, O., Gianotti, M., Rossi, A., Goulene, K., & Stramba-Badiale, M. (2016). Improvement of spontaneous language in stroke patients with chronic aphasia treated with music therapy: a randomized controlled trial. *International Journal of Neuroscience*, *126*(3), 235–242. https://doi.org/10.3109/00207454.2015.1010647

Wilssens, I., Vandenborre, D., van Dun, K., Verhoeven, J., Visch-Brink, E., & Mariëna, P. (2015). Constraint-induced aphasia therapy versus intensive semantic treatment in fluent aphasia. *American Journal of Speech-Language Pathology*, *24*(2), 281–294. https://doi.org/10.1044/2015_AJSLP-14-0018

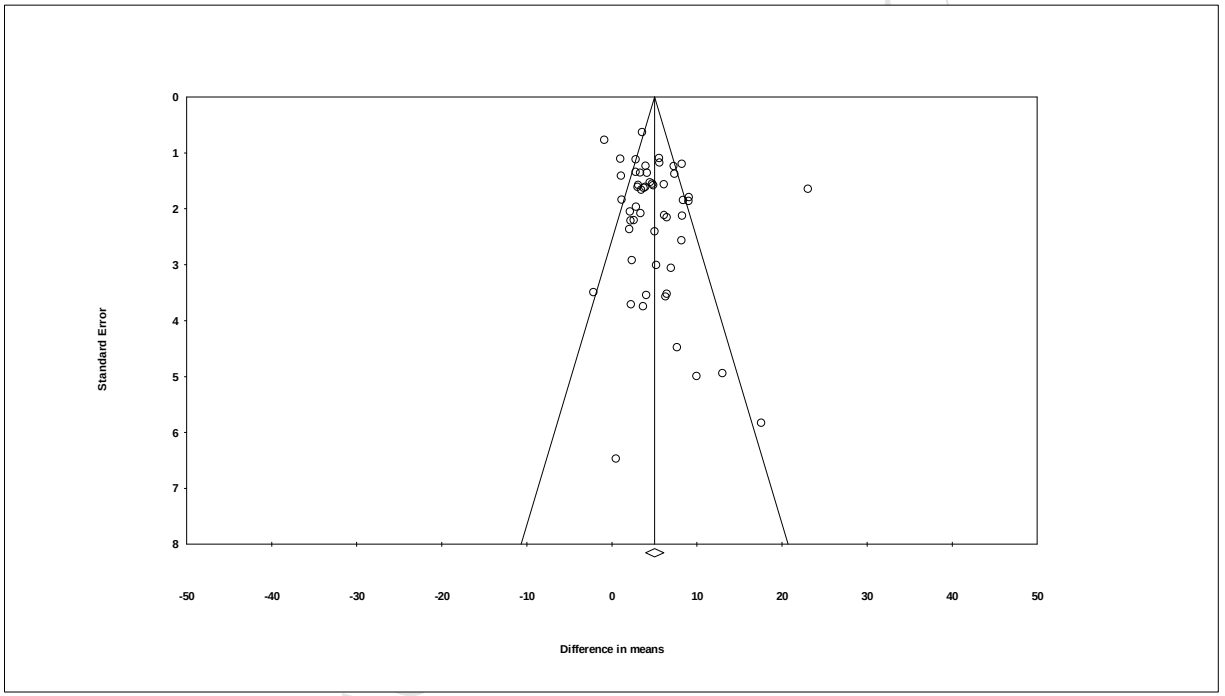
ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Supplementary Material 10: Funnel plots for meta-analyses using within group designs

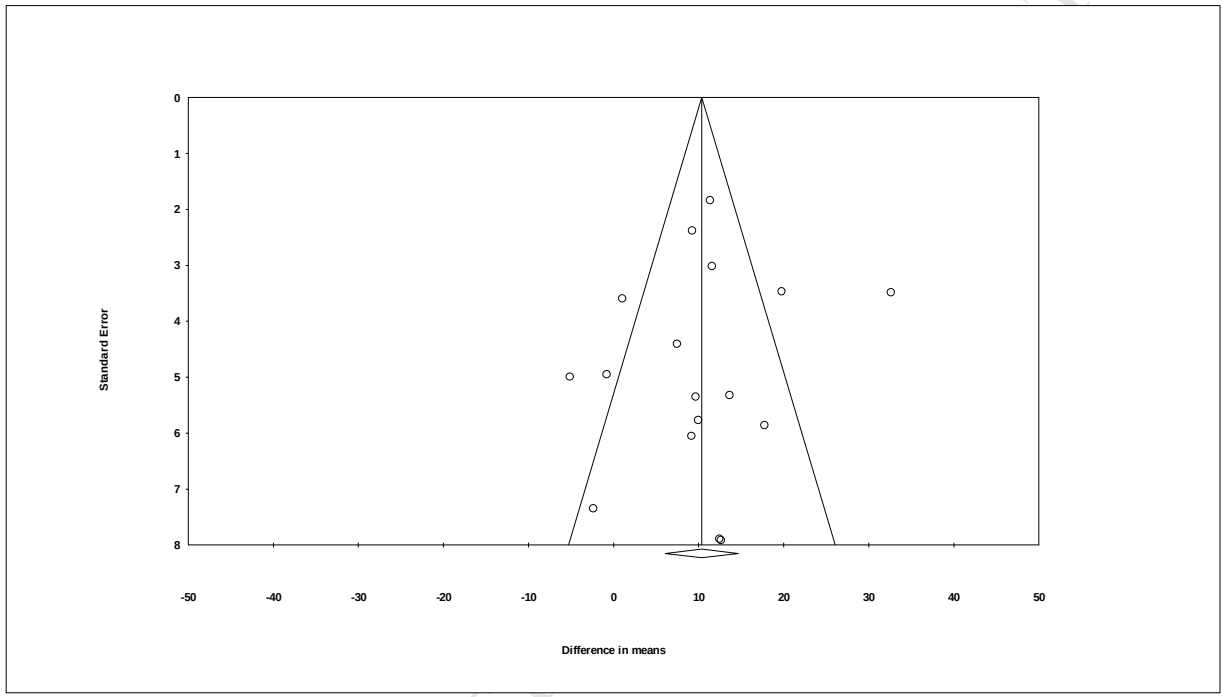
These plots reflect relatively symmetric distribution of studies on both sides of the mean, suggesting limited impact of publication bias on the overall summary effect size (SES) results.

Funnel plot for Western Aphasia Battery-Aphasia within group meta-analysis



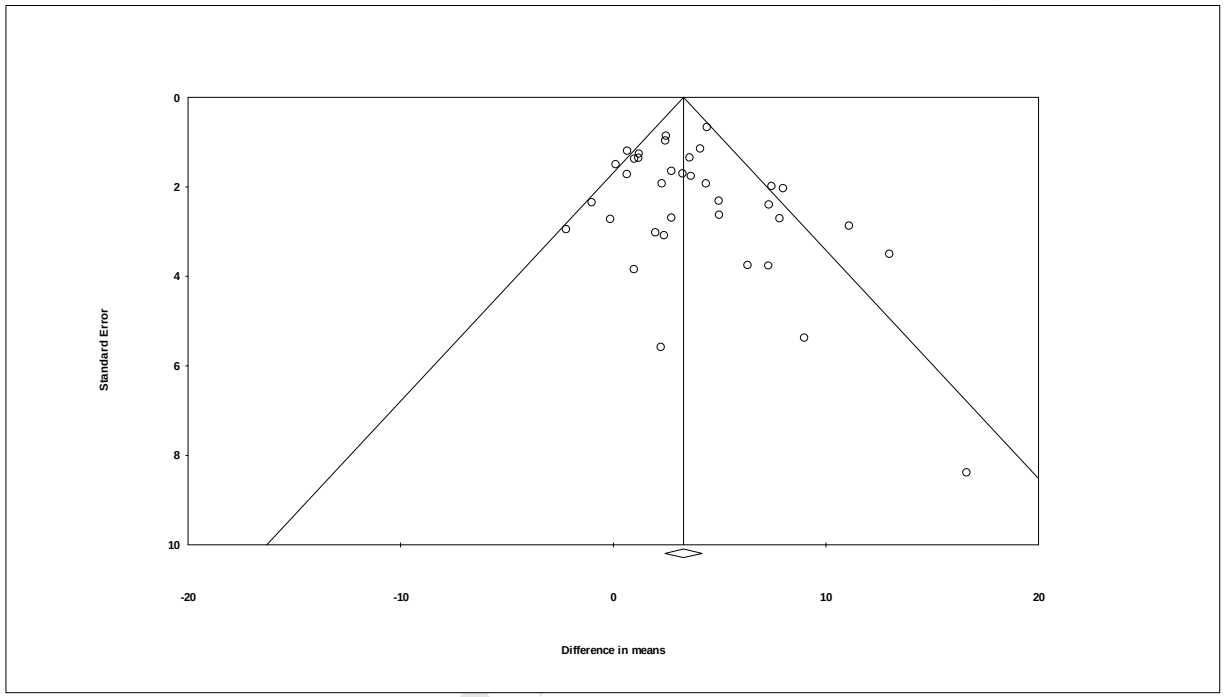
ACCEPTED MANUSCRIPT

Funnel plot for Communicative Effectiveness Index within group meta-analysis



ACCEPTED MANUSCRIPT

Funnel plot for Boston Naming Test within group meta-analysis



ACCEPTED

ACCEPTED MANUSCRIPT

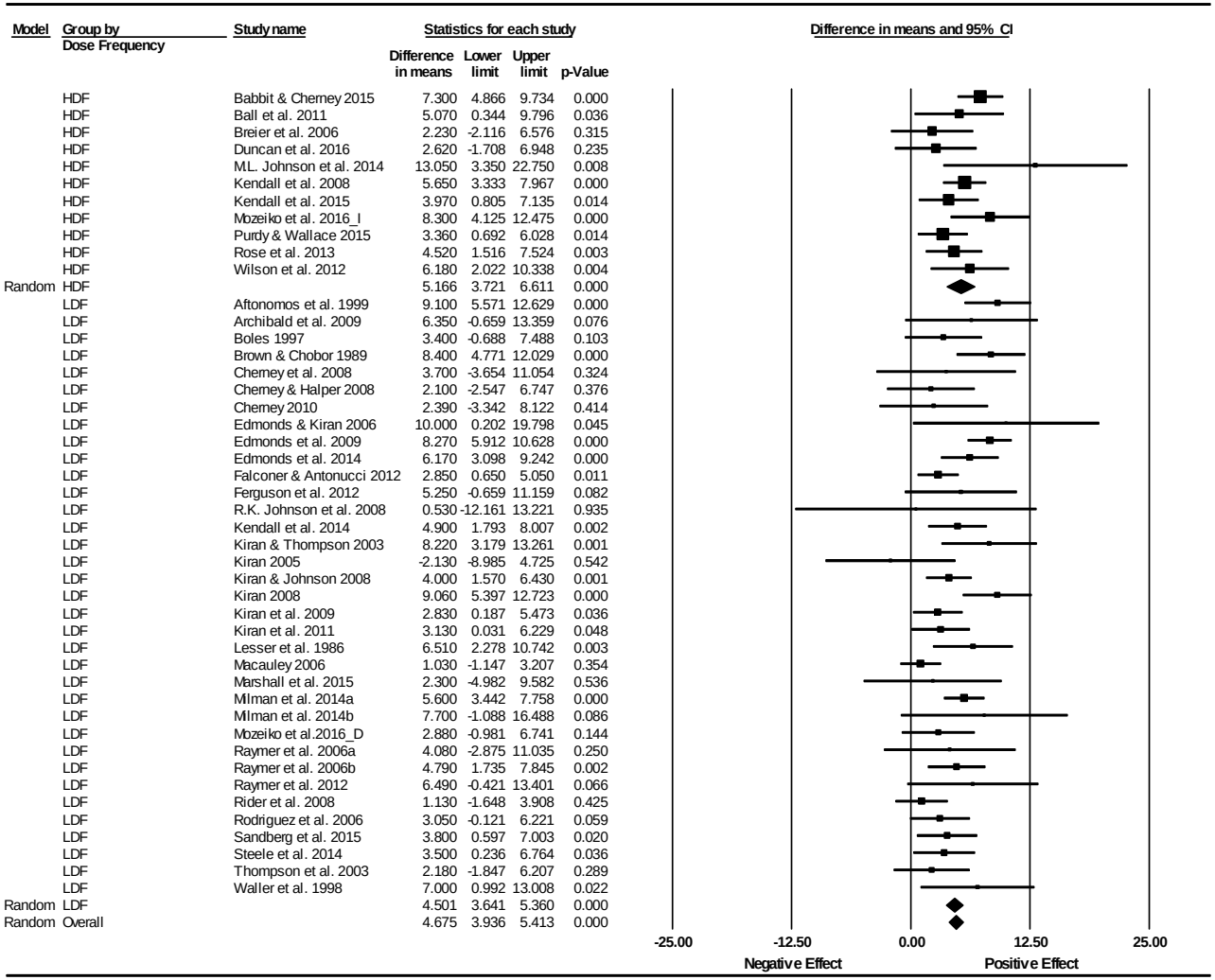
Supplementary Material 11: Forest plots for subgroup analyses using within group study designs

Summary effect sizes for each subgroup and for all of the studies are provided. The difference in means column reflects the pre-treatment mean subtracted from the post-treatment mean. The lower and upper limits columns show the 95% confidence interval surrounding the difference in means. The final row describes the summary effect size, 95% CI, and p-value. The diamond represents the summary effect size. The squares reflect effect sizes of individual studies.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Forest plot of Western Aphasia Battery –Aphasia Quotient subgroup analysis for dose frequency



Model	Group by Dose Frequency	Study name	Difference in means	Lower limit	Upper limit	p-Value
	HDF	Babbit & Cherney 2015	7.300	4.866	9.734	0.000
	HDF	Ball et al. 2011	5.070	0.344	9.796	0.036
	HDF	Breier et al. 2006	2.230	-2.116	6.576	0.315
	HDF	Duncan et al. 2016	2.620	-1.708	6.948	0.235
	HDF	M.L. Johnson et al. 2014	13.050	3.350	22.750	0.008
	HDF	Kendall et al. 2008	5.650	3.333	7.967	0.000
	HDF	Kendall et al. 2015	3.970	0.805	7.135	0.014
	HDF	Mozeiko et al. 2016_I	8.300	4.125	12.475	0.000
	HDF	Purdy & Wallace 2015	3.360	0.692	6.028	0.014
	HDF	Rose et al. 2013	4.520	1.516	7.524	0.003
	HDF	Wilson et al. 2012	6.180	2.022	10.338	0.004
Random	HDF		5.166	3.721	6.611	0.000
	LDF	Aftonomos et al. 1999	9.100	5.571	12.629	0.000
	LDF	Archibald et al. 2009	6.350	-0.659	13.359	0.076
	LDF	Boles 1997	3.400	-0.688	7.488	0.103
	LDF	Brown & Chobor 1989	8.400	4.771	12.029	0.000
	LDF	Cherney et al. 2008	3.700	-3.654	11.054	0.324
	LDF	Cherney & Halper 2008	2.100	-2.547	6.747	0.376
	LDF	Cherney 2010	2.390	-3.342	8.122	0.414
	LDF	Edmonds & Kiran 2006	10.000	0.202	19.798	0.045
	LDF	Edmonds et al. 2009	8.270	5.912	10.628	0.000
	LDF	Edmonds et al. 2014	6.170	3.098	9.242	0.000
	LDF	Falconer & Antonucci 2012	2.850	0.650	5.050	0.011
	LDF	Ferguson et al. 2012	5.250	-0.659	11.159	0.082
	LDF	R.K. Johnson et al. 2008	0.530	-12.161	13.221	0.935
	LDF	Kendall et al. 2014	4.900	1.793	8.007	0.002
	LDF	Kiran & Thompson 2003	8.220	3.179	13.261	0.001
	LDF	Kiran 2005	-2.130	-8.985	4.725	0.542
	LDF	Kiran & Johnson 2008	4.000	1.570	6.430	0.001
	LDF	Kiran 2008	9.060	5.397	12.723	0.000
	LDF	Kiran et al. 2009	2.830	0.187	5.473	0.036
	LDF	Kiran et al. 2011	3.130	0.031	6.229	0.048
	LDF	Lesser et al. 1986	6.510	2.278	10.742	0.003
	LDF	Macauley 2006	1.030	-1.147	3.207	0.354
	LDF	Marshall et al. 2015	2.300	-4.982	9.582	0.536
	LDF	Mlman et al. 2014a	5.600	3.442	7.758	0.000
	LDF	Mlman et al. 2014b	7.700	-1.088	16.488	0.086
	LDF	Mozeiko et al.2016_D	2.880	-0.981	6.741	0.144
	LDF	Raymer et al. 2006a	4.080	-2.875	11.035	0.250
	LDF	Raymer et al. 2006b	4.790	1.735	7.845	0.002
	LDF	Raymer et al. 2012	6.490	-0.421	13.401	0.066
	LDF	Rider et al. 2008	1.130	-1.648	3.908	0.425
	LDF	Rodriguez et al. 2006	3.050	-0.121	6.221	0.059
	LDF	Sandberg et al. 2015	3.800	0.597	7.003	0.020
	LDF	Steele et al. 2014	3.500	0.236	6.764	0.036
	LDF	Thompson et al. 2003	2.180	-1.847	6.207	0.289
	LDF	Waller et al. 1998	7.000	0.992	13.008	0.022
Random	LDF		4.501	3.641	5.360	0.000
Random	Overall		4.675	3.936	5.413	0.000

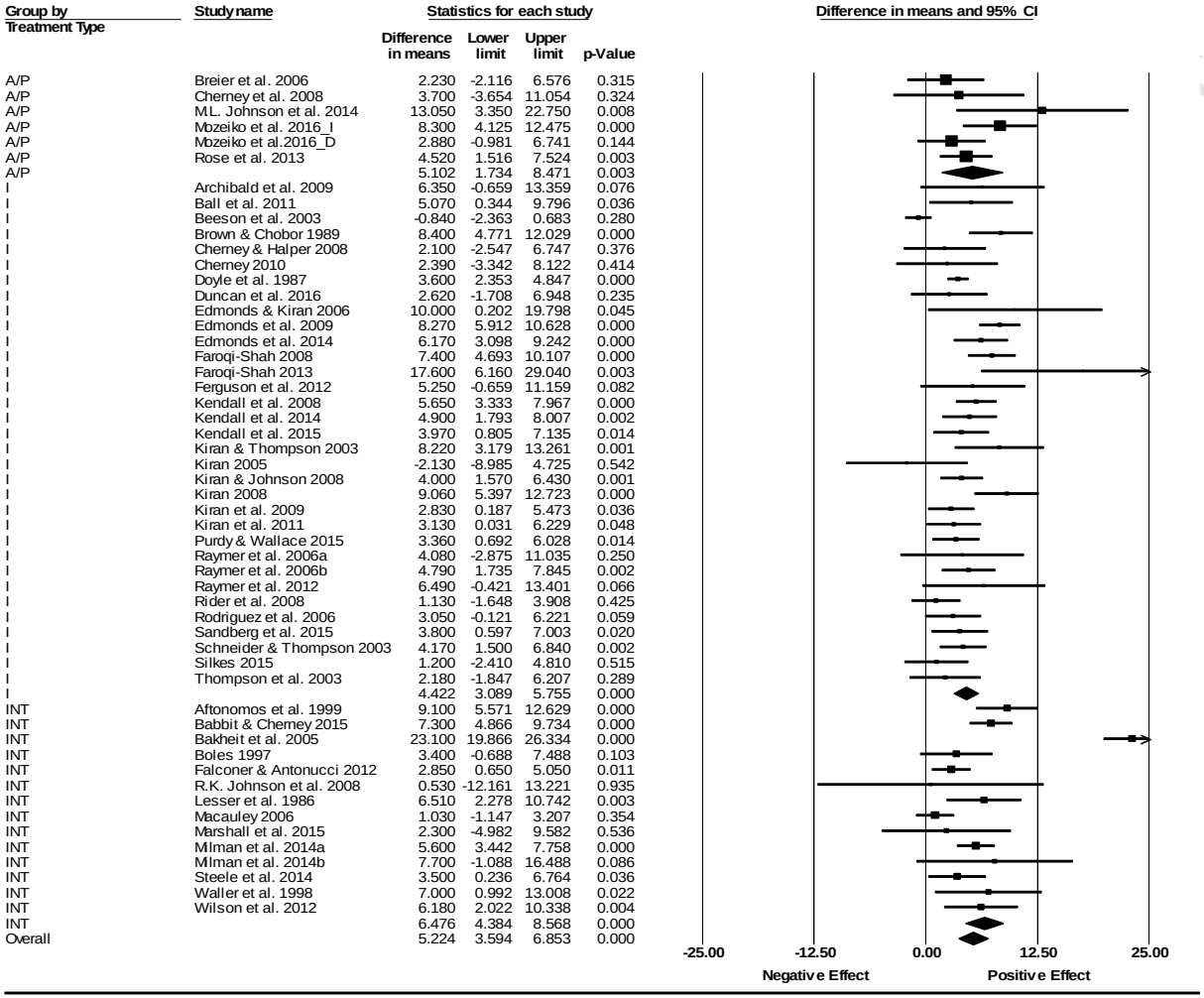
ACCEPTED MANUSCRIPT

Note: LDF = lower dose frequency treatment schedule, HDF = higher dose frequency treatment schedule. There were no statistically significant differences between summary effect sizes for the lower dose frequency and higher dose frequency subgroups (Q = .601, df = 1, $p > .05$).

Forest plot of Western Aphasia Battery –Aphasia Quotient subgroup analysis for treatment type

ACCEPTED MANUSCRIPT

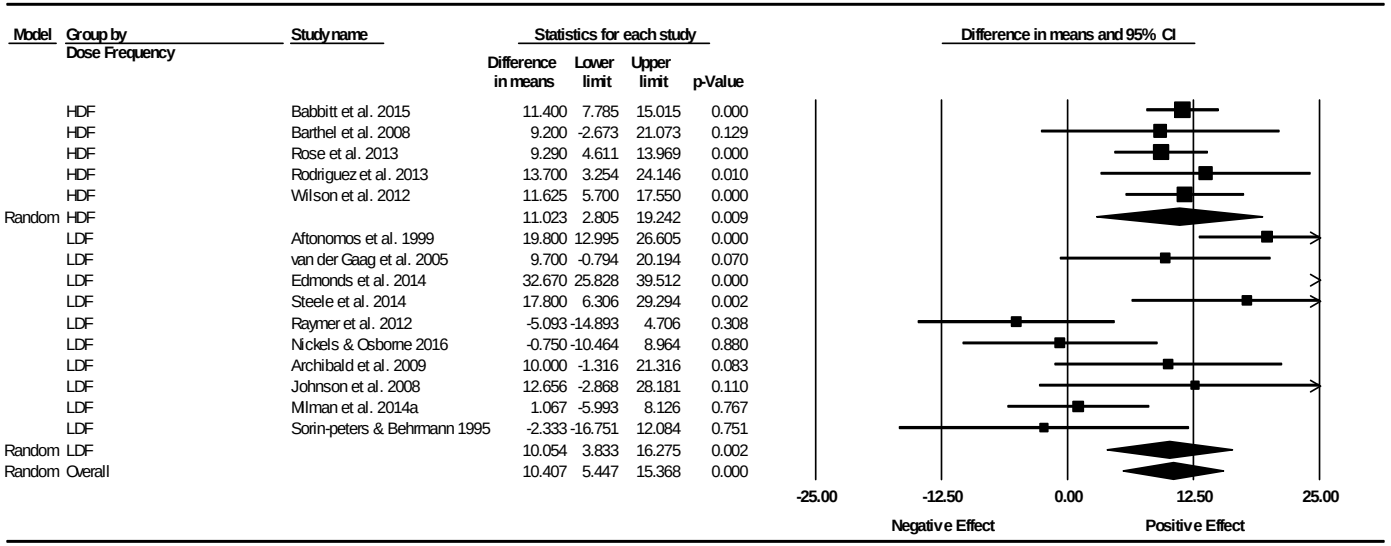
ACCEPTED MANUSCRIPT



Note: I = impairment-based treatment, A/P = activity/participation treatment, INT= integrated treatment. There were no statistically significant differences between summary effect sizes for the different treatment types (Q= 2.64, df= 2, $p > .05$).

ACCEPTED MANUSCRIPT

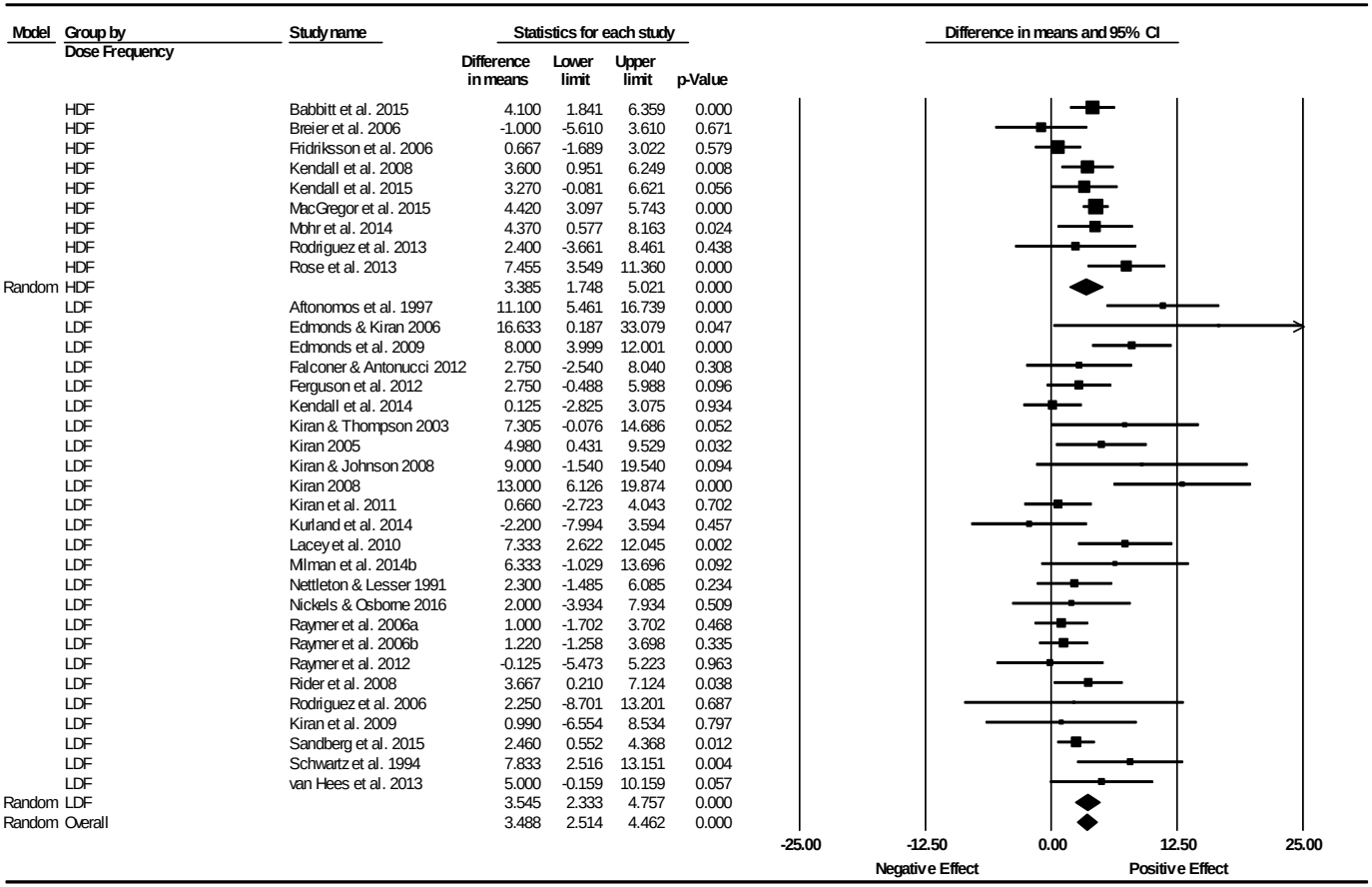
Forest plot of Communicative Effectiveness Index subgroup analysis for dose frequency



Note: LDF = lower dose frequency treatment schedule, HDF = higher dose frequency treatment schedule. There were no statistically significant differences between summary effect sizes for the lower dose frequency and higher dose frequency subgroups (Q = .034, df = 1, $p > .05$).

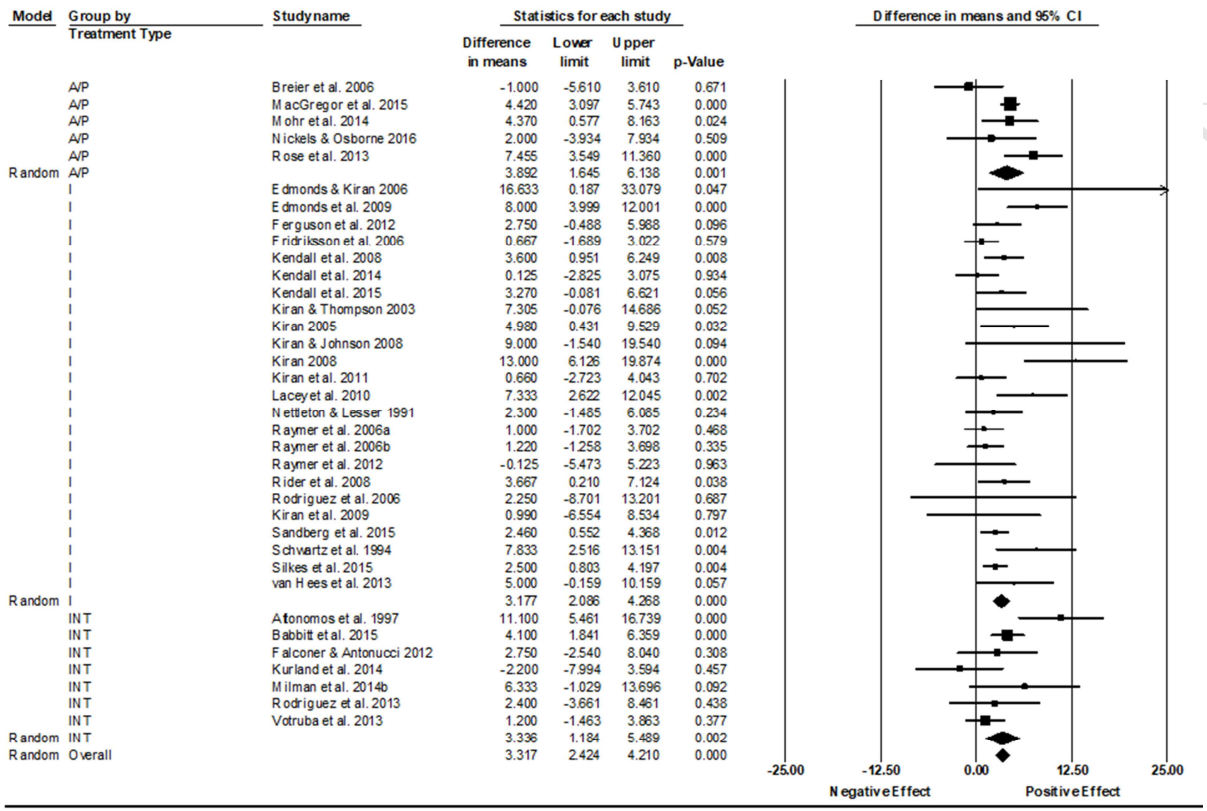
ACCEPTED

Forest plot of Boston Naming Test subgroup analysis for dose frequency



Note: LDF = lower dose frequency treatment schedule, HDF = higher dose frequency treatment schedule. There were no statistically significant differences between summary effect sizes for the lower dose frequency and higher dose frequency subgroups (Q =.024, df = 1, $p > .05$).

Forest plot of Boston Naming Test analysis for treatment type



Note: I = impairment-based treatment, A/P = activity/participation treatment, INT = integrated treatment. There were no statistically significant differences between the summary effect sizes across different treatment types (Q=.32, df= 2, *p*> .05).

ACCEPTED MANUSCRIPT

Supplementary Material 12: Summary of study quality ratings

	Study Protocol	Blinding	Sampling	Treatment Fidelity	Statistics-primary	Statistics-standardized	Precision	I-T-T
Within Group	96	10	99	47	50	40	100	N/A
Between Group	100	38	50	88	100	100	100	25

Note: Value reflects percentage of studies meeting criteria. Statistics-primary indicates that researchers tested for significance on the primary outcome measure; Statistics-standardized reflects that researchers tested for significance on the standardized outcome measure of interest (i.e., WAB, BNT, or CETI); Precision = provided adequate information to calculate an effect size; I-T-T = intention-to-treat analysis

ACCEPTED MANUSCRIPT