ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

A Randomized Controlled Trial of the Effects of Group Conversation Treatment on Monologic Discourse in Aphasia

Elizabeth Hoover,[a] Gayle DeDe,[b] and Edwin Maas[b]

[a] Department of Speech, Language & Hearing Sciences, Boston University, MA [b] Department of Communication Sciences and Disorders, Temple University, Philadelphia, PA

ARTICLE INFO

Article History:
Received January 13, 2021
Revision received June 15, 2021
Accepted July 28, 2021

Editor-in-Chief: Stephen M. Camarata
Editor: Julius Fridriksson

https://doi.org/10.1044/2021_JSLHR-21-00023

ABSTRACT

Purpose: Evidence has shown that group conversation treatment may improve communication and reduce social isolation for people with aphasia. However, little is known about the impact of conversation group treatment on measures of discourse. This project explored the impact of conversation treatment on measures of monologic discourse.

Method: In this randomized controlled trial, 48 participants with chronic aphasia were randomly assigned to dyadic, large group, or control conditions. Conversation group treatment was provided for 1 hr, twice per week, for 10 weeks. Discourse samples were collected and coded at pretreatment, posttreatment, and 6-week maintenance. There were three narrative tasks: (a) Comprehensive Aphasia Test (CAT) picture description, (b) Cat Rescue Picture, and (c) Cinderella retell. All narratives were coded using the percent correct information units (percent CIUs), the CAT standardized narrative analysis method, and the complete utterance (CU) method.

Results: No significant changes were observed on percent CIU, which was the primary outcome measure. The treated groups demonstrated improvement on aspects of the CU method following treatment, whereas the control group did not. Significant changes were observed for other CIU measures and the CAT standardized narrative analysis in both the treated and control groups.

Conclusions: The results suggest that the CU measures were more sensitive to the effects of conversation treatment in monologic discourse compared to CIU and CAT measures. Changes were more common in absolute rather than relative values, suggesting that conversation treatment impacts the overall amount of language produced rather than efficiency of production.

Aphasia, the loss or reduction of language following brain damage, is one of the most debilitating consequences of stroke (Simmons-Mackie et al., 2002). Conservative estimates reveal that over 2.5 million people are living with aphasia in the United States (Simmons-Mackie, 2018). People with aphasia (PWA) often recover some language abilities following the stroke. However, their communication difficulties are often chronic, limit reengagement in vocational and recreational activities, and negatively impact social relationships (Parr, 2007). Social isolation has been correlated with additional adverse health consequences such as infection (Cohen et al., 1997), depression (Heikkinen & Kauppinen, 2004), cognitive decline (Butler et al., 2004; Wilson et al., 2007), and morbidity and mortality (Brummett et al., 2001; Seeman, 2000). Given the significant and expansive consequences of aphasia, it is important to establish treatment approaches that both maximize communication outcomes and minimize the social isolation associated with aphasia.

Both support groups and conversation group treatment offer psychosocial benefits for PWA. For example,

Correspondence to Elizabeth Hoover: ehoover@bu.edu. ***Disclosure:*** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

the collaboration and psychosocial reinforcement inherent to groups positively affects health and well-being in both types of groups (Elman, 2004, 2007; Hilari & Northcott, 2006; Ryff et al., 2004; Vickers, 2010). Conversation group treatment, however, is a theoretically motivated approach in which a speech-language pathologist facilitates discourse about topics of interest to PWA. Communication goals are targeted in the context of natural discourse, specific supports are provided to increase communicative attempts and conversational success, and there is an opportunity to observe communication strategies employed by other group participants (DeDe et al., 2019; Elman & Hoover, 2013). Thus, conversation group treatment has the potential to address both psychosocial and communicative effects of aphasia (DeDe et al., 2019; Elman & Bernstein-Ellis, 1999a, 1999b; Hoover et al., 2015; McCall et al., 2014; Savage et al., 2014).

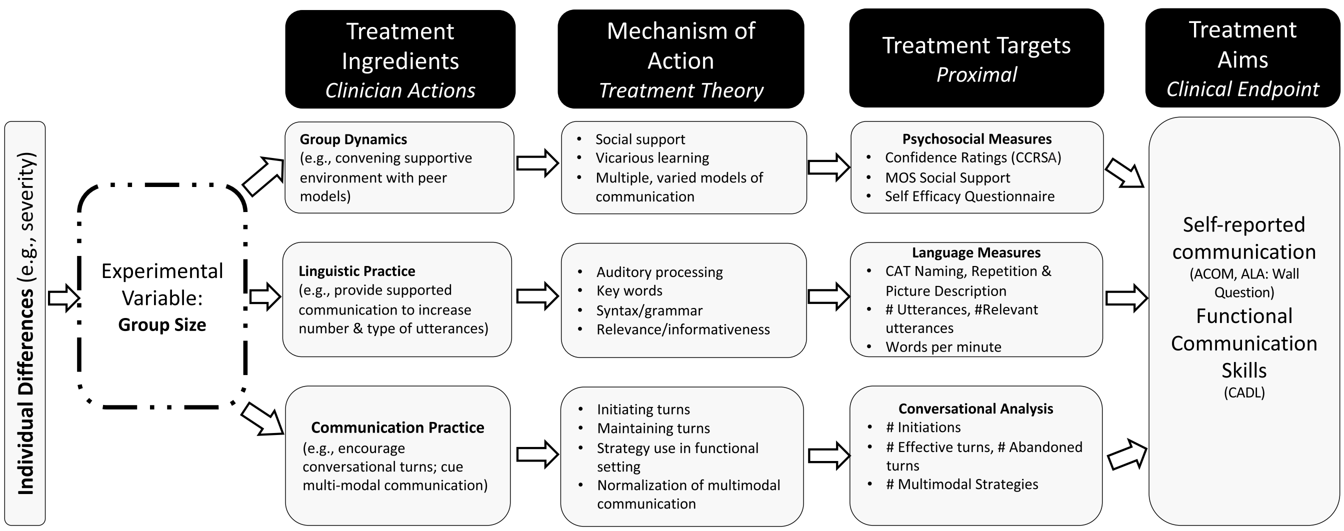
There are several theoretical constructs that motivate conversation group treatment. In general, conversation group treatment is informed by aspects of learning theory, which suggests that learning is enhanced when targets are selected by and relevant to the individual, and easily transferred to other environments (Siegler, 1991). These principles are supported in conversation group treatment because the treatment occurs in the context of naturalistic discourse, where the PWA directs the conversation, and the linguistic target (conversation) is easily generalized to many communication settings (Elman, 2007). In addition, PWA receive diverse language models from their conversation partners, allowing opportunities for vicarious (observational) learning. Conversation group treatment provides opportunities for language improvisation, as participants need to adjust to varied and changing topics (Elman, 2004). Because conversation is dynamic and evolving, it offers the opportunity for personally relevant and meaningful dialogue. The personalized nature of the conversation may capitalize on autonomous motivation, thereby facilitating superior goal progress, more effective coping, better learning and improved well-being (Biel, 2017; Deci & Ryan, 2000; Koestner et al., 2008). Finally, conversation group treatment follows the idea that targeting language at a higher level of linguistic complexity (i.e., discourse) may generalize to related but less complex language functions (e.g., lexical retrieval), and thus, induce neural plasticity (e.g., Complexity Account of Treatment Efficacy hypothesis; Kiran & Thompson, 2019; Thompson et al., 2003; cf. Elman, 2004).

Conversation group treatment, like other functional communication treatments, is interactive and multifaceted. The implementation of conversation group treatment also varies widely (cf. Simmons-Mackie et al., 2014). For example, conversation groups differ in size, with some studies reporting individual, rather than group, conversation treatment (e.g., Savage et al., 2014). Little is known about the relative importance of different conversation treatment ingredients, such as group size, in achieving desired outcomes. Further study of the relative impact of these ingredients on outcomes is a critical next step to better understand the mechanisms of action underlying this intervention.

Recently, the Rehabilitation Treatment Specification System (RTSS; e.g., Hart et al., 2019) was proposed as a method for specifying the pathway by which active ingredients (via presumed mechanisms of action) lead to changes in specific behaviors (*treatment targets*), and how changes in those specific behaviors may relate to ultimate goals (*treatment aims*) of treatment such as self-reported, functional ability (e.g., Hart et al., 2019; Van Stan et al., 2019 and cf. Czajkowski et al., 2015, for a similar system of describing treatment approaches). Figure 1 presents a hypothesized pathway that includes these treatment ingredients and the underlying, presumed mechanisms of action for conversation treatment. We hypothesize that some of the key treatment ingredients in conversation group treatment include group dynamics, as well as effects of practice on linguistic targets, turn taking, and use of multimodal communication. Presumed mechanisms of the *linguistic targets* component within conversation group treatment include opportunities to process auditory input, produce key words, and so on. Presumed mechanisms of group dynamics encompass vicarious (i.e., observational) learning, social support, and exposure to multiple, varied models of communication. Columns labeled *treatment targets* and *treatment aims* in Figure 1 reflect communication behaviors that may change after treatment. *Treatment targets* are quantifiable behaviors that may influence the *treatment aims* but are more proximal to the treatment itself, whereas the *treatment aims* refer to the ultimate, more distal goals of treatment (also sometimes referred to as clinical endpoints). *Treatment targets* may include, for example, psychosocial outcomes (e.g., Ratings on the Communication Confidence Rating Scale for Aphasia; Babbitt et al., 2011), and language measures at word, sentence, and discourse levels. This study is relevant to two aspects of this model: the types of language measures that change as a result of conversation treatment and the critical ingredients of treatment (i.e., effect of group size on group dynamics and language practice).

Our hypothesized pathway portrays a simplified view of which treatment targets are likely to be associated with different treatment ingredients. This pathway represents an initial step toward formally describing conversation treatment within the RTSS. The purpose is to systematically describe our conversation treatment approach with respect to its underlying theoretical underpinnings, in order to facilitate hypothesis testing about the effects of treatment as well as which treatment ingredients are critical and what treatment parameters are optimal.

Figure 1. Hypothesized pathway by which conversation treatment affects communication ability and life participation (based on the Rehabilitation Treatment Specification System, Van Stan et al., 2019). Group size is the experimental variable in this study. By manipulating group size, we hypothesize differential effects in other aspects of the model. CCRSA = Communication Confidence Rating Scale for Aphasia (Babbitt et al., 2011); MOS = MOS Social Support Survey (Sherbourne & Stewart, 1991); CAT = Comprehensive Aphasia Test (Swinburn et al., 2004); ACOM = Aphasia Communication Outcome Measure (Hula et al., 2015); ALA = Assessment of Living with Aphasia (Kagan et al., 2010, 2018); CADL = Communicative Activities of Daily Living -Third Edition (Holland, 1980; Holland et al., 2018).



We recently completed a randomized controlled clinical trial to investigate one ingredient of conversation treatment, namely, the effects of group size (DeDe et al., 2019). Participants with aphasia were randomly assigned to one of three conditions: dyad (two PWA with one graduate student clinician), large group (six to eight PWA with two graduate student clinicians), or the control group. We hypothesized that individuals who received conversation treatment in dyads would have more opportunities to take conversational turns, and thus more linguistic and communication practice. In contrast, we hypothesized that larger groups would offer more of the benefits of group dynamics. Thus, our manipulation of group size was intended to examine the relative contribution of three ingredients of conversation treatment (see Figure 1).

Both treatment groups (dyad and large group) showed significant improvements on several standardized tests of language and on patient-reported outcome measures, whereas the control group did not (DeDe et al., 2019). The data revealed a complex relationship between group size and outcome measures. Dyads were associated with improvement on discrete linguistic tasks of repetition and two different word retrieval tasks, whereas larger groups were associated with improvement on one naming measure, the picture description task from the Comprehensive Aphasia Test (CAT; Swinburn et al., 2004) and a patient-reported outcome measure of functional communication, the Aphasia Communication Outcome Measure (ACOM; Hula et al., 2015). This work suggests that the degree of change in treatment targets related to language ability (e.g., naming) and treatment aims (i.e., ACOM) varies based on group size.

This study is a companion article to DeDe et al. (2019), and focuses on the effect of conversation treatment on monologic discourse. In our hypothesized pathway (see Figure 1), discourse is identified as a subset of language measures that are proximal treatment targets. Dipper and Pritchard (2017) describe the following common types of monologic discourse: narrative (campfire story), procedural (how to make a cup of tea), personal (stroke story), descriptive discourse (picture description task), and expository (what is your favorite game or sport? Why is it your favorite game?). There are many reasons to examine the effects of conversation treatment on monologic discourse. Discourse is critical for human interactions, including the expression of one's feelings, thoughts, or ideas (Linnick et al., 2016). Telling stories is essential not just for communication but also for developing and maintaining social connections and establishing personal identity (e.g., Linnick et al., 2016; Shadden & Agan, 2004; Shadden & Hagstrom, 2007). Conversation treatment is well situated to address personal story telling and discourse comprehension and production, given its focus on authentic sharing of thoughts and ideas.

There are several measures available to evaluate discourse, many of which focus on one element of the spoken language, such as the proportion of informative words, or the syntactic complexity of the discourse (Bryant et al., 2016). PWA demonstrate unique and varied patterns of strengths and difficulties across linguistic elements, which complicates measurement selection (Boyle, 2020). In addition, measures of only one specific element within a discourse may not be sufficiently sensitive to capture change for a group of individuals with varied profiles of aphasia. As a result, little consensus exists in the field about which method of discourse measurement might work best for a heterogeneous cohort of PWA (Pritchard et al., 2017).

Discourse measures are often excluded from patient outcomes because clinicians report a lack of confidence regarding selection of appropriate measures and/or the coding methods are time consuming to complete (Cruice et al., 2020). Factors such as variability in task elicitation and client performance further complicate the use of discourse coding methods in practice. Conversation group treatment is also inherently unstructured, and clinicians may focus on different linguistic or multimodal targets for different participants (DeDe et al., 2019). Furthermore, discourse measures are often limited to spoken productions, which makes selection of a single discourse measure particularly challenging for a cohort with varied profiles of aphasia. Despite these challenges, improved discourse remains a critical outcome. PWA's self-identified goals include telling personal stories, participating in conversations, and creating other forms of spoken and written narratives. Improved discourse is essential for developing and maintaining social connections and personal identity (Linnick et al., 2016). To date, there is limited research investigating the effects of conversation treatment on measures of discourse. Mason et al. (2020) completed a pilot study on the impact of group conversation treatment for aphasia on connected speech. Three PWA participated in a 6-week group treatment program that focused on weekly conversational topics. Semistructured interviews were collected pre- and posttreatment and analyzed using measures of word counts, profile of word errors, and retrieval in speech, a measure of propositional idea density, and perceptual discourse ratings. Only one participant showed a statistical improvement on one discourse rating: ability to find adequate words during a discourse sample. No other significant changes were observed.

This article presents data on effects of conversation treatment on monologic discourse outcome measures from our randomized controlled trial (DeDe et al., 2019). Changes in monologic narratives were evaluated using three discourse measures. Percent correct information units (percent CIUs; Nicholas & Brookshire, 1993) was the predetermined primary outcome measure. In addition, we examined other measures derived from CIU analysis (e.g., number of CIUs, number of words), the narrative coding method from the CAT (Swinburn et al., 2004), and complete utterances (CUs; Edmonds et al., 2009).

This study addressed two research questions. The first question was whether participation in conversation treatment resulted in improved performance on monologic discourse. We predicted that the treated groups, but not the control group, would show significant improvements on discourse measures (CIU, CAT narrative coding, CU). This question is relevant to the types of language outcomes that are effectively targeted by conversation treatment (see Figure 1).

The second question focused on one ingredient of conversation treatment. Like DeDe et al. (2019), we asked whether the size of the group (dyads vs. large group conditions) influenced changes in discourse after conversation treatment. DeDe et al. (2019) showed that individuals in dyads have more conversational turns than individuals in a large group. However, large groups may offer more of the benefits of group dynamics than dyads. By manipulating group size, we were able to examine the relative contribution of language practice and group dynamics to the treatment outcomes. Based on the superior performance of the large group on the single picture description task (CAT standardized narrative analysis method [CAT-NA]; DeDe et al., 2019), we predicted that the large group would outperform the dyads on additional measures of monologic discourse. This study is, to our knowledge, the first group study to examine effects of treatment on multiple discourse measures, coded on three different narrative tasks (i.e., Cinderella retell and two picture descriptions). This is important because discourse coding is labor intensive, and it is critical to identify which measures are both stable and sensitive to effects of treatment.

Method

For a detailed description of treatment participants, development, fidelity, and methodology, see DeDe et al. (2019). This project was approved by the institutional review board (IRB) at Boston University, which served as a single IRB for all sites.

Participants

A total of 48 PWA were recruited across two sites for a multicenter, parallel-group unblinded randomized controlled trial with a control group and balanced randomization. Table 1 shows basic demographic data for participants.

Inclusion criteria were (a) 18 years of age or older, (b) at least 5 months postonset from a stroke, (c) English as first language (based on self-report), and (d) presence of aphasia (based on clinical judgment and scores on the CAT). Using a simple randomization sequence determined by a statistician, participants were randomly assigned to one of three conditions: large group (six to eight PWA), dyad (two PWA) or control group, after it was determined that they met inclusion criteria. One-way analyses of variance showed that the groups were equivalent with respect to age and years of education (all Fs < 1, $p > .40$). Following random assignment, dyads were matched on the

Table 1. Demographic data by condition and site.

	Boston (*n* = 24)			Temple (*n* = 22)			All (*n* = 46)		
Variable	**Dyad**	**Large group**	**Delay**	**Dyad**	**Large group**	**Delay**	**Dyad**	**Large group**	**Delay**
Age in years (*SD*)	70.5 (9.7)	61.4 (20.9)	71.4 (9.0)	60.8 (6.6)	61.4 (5.9)	58.4 (11.7)	65.6 (9.4)	61.4 (14.3)	65.8 (11.9)
Education in years (*SD*)	16.4 (3.5)	15.1 (3.1)	16.9 (2.1)	13.7 (1.5)	12.6 (2.2)	13.8 (3.7)	15.0 (2.9)	13.8 (2.9)	15.6 (3.2)
Gender (#M/#F)	7/1	5/3	8/0	5/3	6/1	1/6	12/4	11/4	9/6
Aphasia Severity[a] *M* (*SD*)	29.9 (27)	50.6 (21.8)	45.5 (23.6)	52.1 (24.5)	49.8 (21.5)	59.5 (16.7)	41 (27.4)	50.2 (20.9)	51.5 (21.5)
Aphasia Syndromes									
Anomic	1	2	1	4	4	2	5	6	3
Broca	1	3	2	2	4	2	3	7	4
Conduction	2	2	3			2	2	2	5
Sv MNF			2	1			1		2
TCM	1								
TCS	1						1		
Wernicke	2	1		1					

Note. Reprinted with permission (DeDe et al., 2019). M = male; F = female; Sv MNF = severe mixed nonfluent; TCM = transcortical motor; TCS = transcortical sensory.

[a]Comprehensive Aphasia Test (CAT) naming total score mean (*SD*).

basis of clinician judgment of compatibility (e.g., personality, shared interests). The decision to assign dyads in this way was made to maximize the clinical relevance of the study. Figure 2 shows the total number of participants at each site at each stage of the study. After attrition, there were 11 participants in the large group condition, 16 in the dyadic condition, and 14 in the delayed treatment control condition.

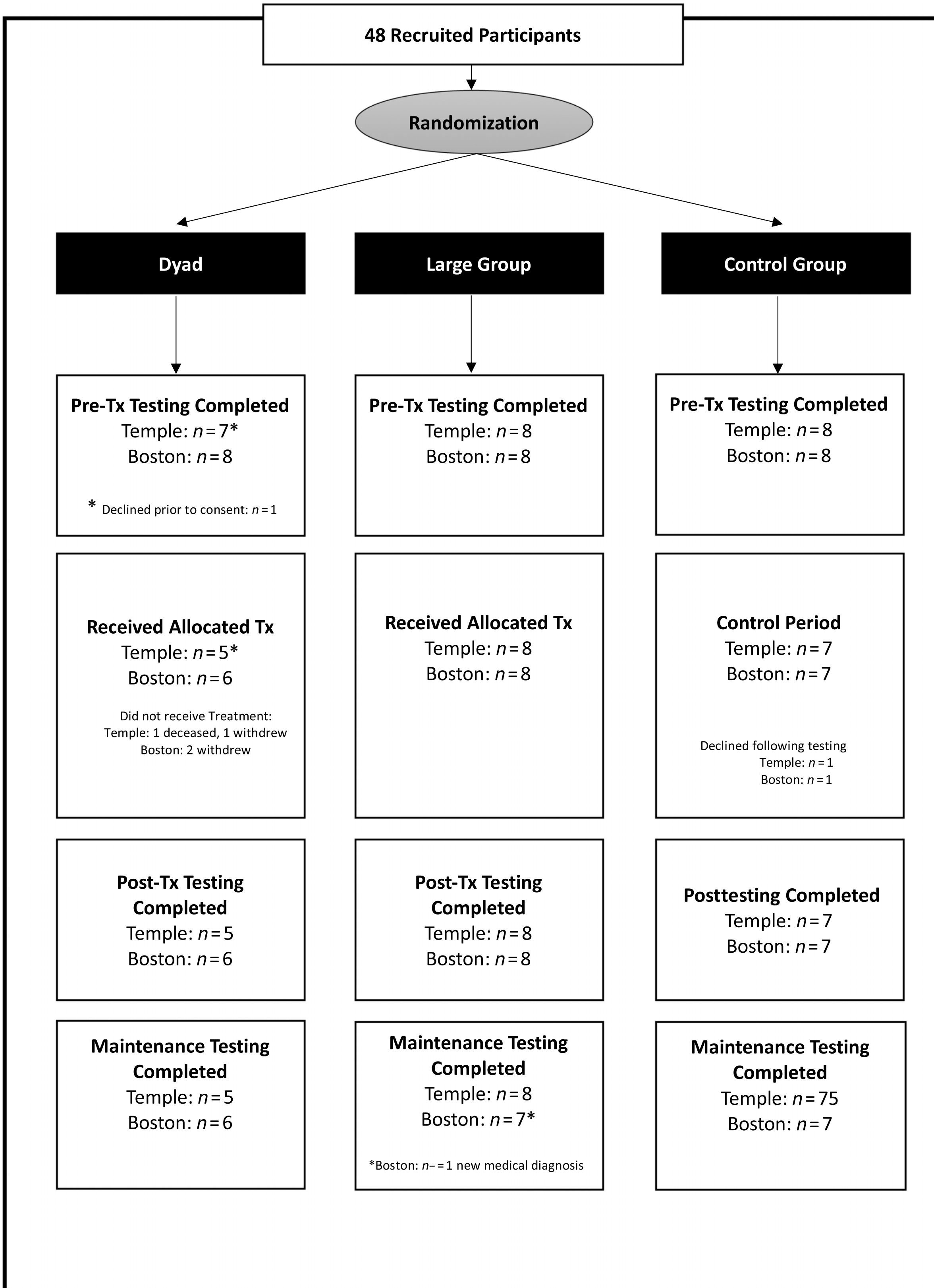
Treatment

Treatment sessions were 60 min, 2 times per week for 10 weeks. Experimental groups received treatment first. The control group received treatment approximately 6 weeks after cessation of the experimental treatment phase, meaning that they did not receive treatment during the time period of this study. The treatment protocol was administered by trained graduate student clinicians from Boston University and Temple University, under the supervision of certified and licensed speech-language pathologists (first and second authors) with previous experience administering conversation treatment. Hereafter, we refer to the graduate student clinicians as group facilitators.

The conversation group treatment followed the life-participation approach to aphasia (Chapey et al., 2008) where participants were encouraged to participate in meaningful, authentic conversations surrounding topics related to personal interests. Supported conversation techniques were incorporated in each session to facilitate engagement and communicative success (Kagan et al., 2001; Elman & Hoover, 2013). Strategies were individualized to the participants based on goals established at pretreatment testing.

To maximize the reproducibility of the study and maintain the consistency of treatment across sessions and sites, we provided detailed training to treatment providers, developed and used consistent supports across sites, and completed fidelity checks to ensure compliance during and following the treatment period. We developed five conversation topics to be discussed over the treatment period: personal history, dining, news/events, entertainment, and travel. Subthemes within the topics were generated, allowing each broad topic to be discussed 4 times. These topics were chosen given their potential applicability to personal interests and daily activities and thus for their potential to promote generalization of any treatment gains to other environments with other conversational partners. Conversation supports and prompts were developed and shared across each treatment site. Each session began with social questions such as, "How is everyone today?" Then conversation surrounding the daily theme was facilitated. Once a topic was introduced, the conversation flowed in a dynamic and authentic manner. The facilitators attempted to elicit an equal number of conversational turns within the group. This included using questions (e.g., *Liz, you've been quiet. What do you think?*) and monitoring for nonverbal cues that a participant wanted to initiate a turn (*Gayle, you looked like you wanted to say something*). PWA were cued to use supports as appropriate. Facilitators supplemented their speech with gestures and written key words at least 3 times per session to normalize multimodal communication. In addition, strategies such as writing or gesturing key words and repeating/rephrasing key information were used to facilitate understanding among participants.

Figure 2. Study recruitment and attrition (adapted from DeDe et al., 2019). Tx = treatment.



Each participant had two individual communication goals, which were based on client preference and results of testing. The facilitators created opportunities to address these individual goals during treatment sessions within the natural flow of the conversation. For example, consider a participant whose goal is to use multimodal communication and who wants to express that they watched a football game. The facilitator might prompt the client to use a gesture to communicate their idea. Alternatively, the facilitator may use question and answer sequences to facilitate communication. When offering suggestions, the facilitator might concurrently speak and gesture options ("Did you watch football?" and gesture throwing football, soccer/kicking ball to demonstrate how multimodal communication could be used).

Treatment Fidelity

The graduate student facilitators had all successfully completed at least one three-credit graduate-level course on characteristics, diagnosis, and treatment of aphasia. Furthermore, the facilitators at both sites received additional training, which included a treatment manual with information about conversation treatment, cueing hierarchies, and weekly topics along with a 2-hr large group orientation meeting to review procedures and goals. Facilitator performance was reviewed verbally or in writing each week to confirm correct implementation of the protocol, based on the supervisor's concurrent observation of the sessions. All facilitators provided treatment in both experimental conditions (large group and dyad). Finally, trained undergraduate observers used a checklist to record variables related to treatment fidelity each session. Specifically, they counted the number of facilitator models for multimodal communication strategies, participant attendance, and the total number of conversation turns for each PWA and facilitator. Any attempt at communication, regardless of modality or accuracy, was counted as a conversational turn. A second independent rater, also a trained undergraduate student, viewed 20% of the treatment sessions to determine reliability of the online coding. Fidelity data are presented in DeDe et al. (2019).

Discourse Measures

Discourse testing was administered as part of a larger testing battery at pretreatment, posttreatment, and 6 weeks posttreatment. There were three narrative tasks, all of which use picture prompts to elicit a monologic discourse: (a) CAT (Swinburn et al., 2004) picture description, (b) Cat Rescue Picture (Nicholas & Brookshire, 1993), and (c) Cinderella retell. Standardized prompts were given from the Aphasia Bank protocol for the Cinderella retell and Cat Rescue Picture and from the CAT for its picture description task (MacWhinney et al., 2011). Narratives were recorded using video and transcribed orthographically and segmented into utterances. All narratives were coded using percent CIUs (Nicholas & Brookshire, 1993), the CAT-NA (Swinburn et al., 2004), number of utterances, and number of CUs (Edmonds et al., 2009).

Percent CIU was chosen as the primary outcome measure because this is a widely used method to analyze linguistic discourse in PWA (Bryant et al., 2016) and because of its strong interrater reliability with well-trained coders (Nicholas & Brookshire, 1993; Leaman & Edmonds, 2019a, 2019b; Oelschlaeger & Thorne, 1999). Furthermore, percent CIU has been used frequently to evaluate improvements in connected speech in the aphasia treatment literature (e.g., Boyle & Coelho, 1995; Edmonds et al., 2009; Hoover et al., 2015). CIUs are words that add new information to the topic, are intelligible, accurate and relevant in context (Nicholas & Brookshire, 1993). CIUs are extracted from the overall word count. CIUs offer relevant, new, or syntactically required information about the pictured scene. Nicholas and Brookshire (1993) provided detailed examples and rules for which words to include or exclude. In addition to percent CIU, we calculated the number of words, number of CIUs, duration of sample, words per minute (WPM), and CIUs per minute.

The CAT-NA was used as part of the administration of the CAT in the larger testing battery. In DeDe et al. (2019), the CAT-NA coding showed sensitivity to change in one of the experimental treatment groups (large group). For this reason, we applied the CAT-NA method to the narratives from the Aphasia Bank protocol to increase the number of samples and provide a comparison to the percent CIU and CU methods. The CAT-NA score is derived from the total number of *appropriate* information carrying words (ICW) minus the total number of *inappropriate* information carrying words. Additional points from a Likert scale are added for syntactic variety, grammatical well-formedness, and speed.

The CU coding method (Edmonds et al., 2009, 2014) was chosen due to its clinical applicability and sensitivity to treatment change, ease of use, and inclusion of both a measure of informativeness and syntax. Utterances are coded for two components (Leaman & Edmonds, 2019a). First, the presence of an appropriate agent and verb and its obligatory arguments (± SV[O]). Second, the relevance of the items within the SV(O) sentence frame in relationship (REL) to the discourse context (± REL). Utterances containing both +SV(O) and + REL are referred to as *CUs.*

Interrater reliability. All narrative samples were coded by research assistants (undergraduate student or someone with a bachelor's degree) who were not involved in the treatment arm of the study. Research assistants received training on each coding method, which included group practice, references, and resource templates to aid with decision making regarding common vocabulary. For all measures, two research assistants independently scored each narrative sample. Codes were compared, and discrepancies were resolved by consensus. If the research assistants were unsure, they consulted one of the authors (DeDe or Hoover) to resolve the code. We do not report additional reliability data because all codes were reached via consensus of two independent raters.

Analysis Plan

Following Boyle's (2014) recommendations, the outcome measures for each of the three narratives were averaged together with equal weight prior to analysis to improve test–retest reliability. Statistical analyses were completed using R Version 3.6.3. Kruskal–Wallis tests were used to determine whether groups differed on outcome measures before treatment. Then pre- versus posttreatment and pretreatment versus 6-week maintenance difference scores were compared using Wilcoxon signed-ranks tests separately for each group, due to differences in group size across conditions and the small sample size. Note that R reports V-statistics for Wilcoxon signed-ranks tests. The alpha level was set at .05 for all analyses, with Holm corrections applied to control type I error due to multiple comparisons for each group. Only effects that were significant after the Holm correction were interpreted as significant. However, all effects that were significant at .05 are reported for full transparency.

Results

CIU. Means and standard deviations for all CIU measures for all three conditions are in Table 2. Kruskal–Wallis tests showed no significant differences between groups for any measure pretreatment (all chi-square ≤ 1.1). As shown in Table 2, there were no significant changes in the primary outcome measure, *percent CIU,* for any group.

The control group showed significant effects in the number of CIUs at Testing Time 2 ($V = 8$, $p = .01$) and at Testing Time 3 ($V = 3$, $p = .002$) and in the duration of the sample at Testing Time 2 ($V = 10$, $p = .02$). Additionally, the control group showed a numeric increase in number of words at Testing Time 2 ($V = 14$, $p = .05$) and Testing Time 3 ($V = 11$, $p = .03$); however, neither effect was significant after the Holm correction.

In the dyad condition, duration of the sample significantly increased from pre- to posttreatment ($V = 20$, $p = .02$). The following measures showed numerical changes which did not survive the Holm correction: number of CIUs from pre- to posttreatment ($V = 20$, $p = .04$) and pretreatment to maintenance ($V = 18$, $p = .05$); WPM decreased from pre- to posttreatment, $V = 98$, $p = .03$. No other analyses approached significance in the dyadic condition (all $ps \geq .07$).

The large group condition showed statistically significant changes from pre- to posttreatment in number of words ($V = 7$, $p = .009$) and duration of the speech sample ($V = 8$, $p = .01$). There was also a numeric increase in number of CIUs from pre- to posttreatment, but it was not significant after the Holm correction ($V = 12$, $p = .03$). No other analyses approached significance in the large group condition (all $ps \geq .10$).

CAT narrative coding. Table 3 presents mean and standard deviations for all groups. Kruskal–Wallis tests showed no significant differences between groups for this measure before treatment (chi-square < 1). The control group showed a statistically significant increase from Testing Time 1 to Testing Time 2 ($V = 10$, $p = .02$), and from Testing Time 1 to Testing Time 3 ($V = 3$, $p = .002$). Dyads' scores showed a numeric increase from pre- to posttreatment testing ($V = 25$, $p = .02$) and the large group showed an increase from pretreatment to maintenance testing ($V = 9$, $p = .03$). Neither change observed for the dyad nor large group condition was significant after the Holm correction.

CUs. Table 3 presents means and standard deviations for variables derived from coding for CUs. Kruskal–Wallis tests showed that prior to treatment, there were no significant differences between groups for any measure (all chi-square ≤ 3, all p values ≤ .23). There were no significant changes in the control group, for any measure (all $ps \geq .13$). For the dyad condition, number of CUs showed a significant increase ($V = 11$, $p = .01$) from pretreatment to maintenance testing. Number of relevant utterances showed a positive deflection, but was not significant after the Holm correction ($V = 16$, $p = .04$). For the large group condition, number of utterances ($V = 8$, $p = .02$), and number of relevant utterances ($V = 7$, $p = .02$) significantly increased from pre- to posttreatment. No other analyses approached significance (large group condition: all $ps \geq .06$).

Discussion

We undertook this project with the intention of answering two primary questions: Does participation in conversation treatment result in improved performance on monologic discourse, and does group size influence changes in discourse? Our primary outcome measure, percent CIU, showed no significant changes across time for any group. This null result raises the question of whether the lack of changes reflects limitations of the treatment or of the measure. To explore this question further, we reviewed and compared the data from all three coding methods.

Table 2. Mean (*SD*) correct information units (CIU) by condition.

	Control			Dyad			Large group		
Measure	**T1**	**T2**	**T3**	**T1**	**T2**	**T3**	**T1**	**T2**	**T3**
Number words	102.6 (42.9)	**131.7* (67.1)**	123.4† (54.7)	136.4 (84.4)	149.6 (75.9)	158.5 (82.4)	96.9 (52.6)	**123.9* (56.4)**	103.7 (52.4)
Number CIUs	47.1 (40.0)	**57.1* (54.3)**	**57.1* (48.8)**	54.8 (67.4)	62.4† (65.8)	71.3 (79.9)	40.5 (31.6)	56.7† (39.6)	48.7 (37.5)
Duration (s)	115.6 (42.0)	**154.6 (68.1)**	124.8 (52.4)	118.8 (55.5)	**152.0* (49.2)**	171.6† (72.5)	108.4 (33.2)	**146.1* (61.0)**	111.7 (47.3)
Words per min	62.4 (39.9)	59.1 (38.8)	66.1 (34.1)	67.3 (32.5)	60.1† (29.8)	61.6 (32.0)	60.5 (34.5)	57.5 (20.8)	58.8 (24.1)
CIU per min	30.8 (31.8)	27.2 (29.9)	31.1 (30.0)	25.0 (24.7)	23.2 (23.4)	23.2 (24.0)	28.1 (28.5)	28.5 (19.0)	28.7 (20.3)
Percent CIU	41.1 (21.0)	39.0 (23.0)	42.3 (23.0)	32.6 (28.6)	32.4 (27.3)	35.3 (31.2)	40.3 (23.3)	45.1 (19.4)	46.0 (20.7)

Note. T1 = Time 1 (pretreatment for treated groups); T2 = Time 2 (immediately posttreatment); T3 = Time 3 (6 weeks posttreatment).
*Significant effects that survived the Holm correction (bold). †*p* < .05 but did not survive Holm correction.

Table 3. Mean (*SD*) complete utterances and CAT narrative analysis by condition.

Measure	Control			Dyad			Large group		
	T1	T2	T3	T1	T2	T3	T1	T2	T3
Complete utterances									
Number utterances	12.1 (5.1)	14.3 (7.6)	13.2 (4.9)	14.3 (7.1)	16.2 (6.2)	16.9 (9.0)	11.7 (5.1)	**14.8* (7.0)**	12.7 (5.6)
Relevant	7.2 (4.9)	8.8 (6.1)	7.3 (4.4)	6.1 (4.9)	7.4 (5.8)	$7.8^{\dagger}$ (6.3)	5.9 (3.7)	**7.7* (4.7)**	7.1 (4.2
Subject–verb (SV)	9.5 (5.3)	10.6 (6.7)	10.1 (4.9)	11.8 (6.8)	12.2 (6.4)	13.3 (7.8)	8.9 (4.4)	10.5 (5.3)	9.3 (4.6)
Complete utterances (CU)	5.9 (5.2)	7.0 (6.3)	6.3 (4.6)	5.3 (5.0)	6.2 (5.6)	**6.8* (5.6)**	5.0 (3.8)	6.3 (4.6)	5.6 (4.1)
Percent relevant	56.7 (19.9)	58.1 (27.1)	53.0 (24.0)	41.9 (26.5)	41.7 (31.8)	44.5 (30.5)	50.7 (25.1)	54.6 (28.5)	58.4 (28.9)
Percent SV	78.5 (20.3)	74.8 (23.4)	76.7 (21.6)	75.5 (25.6)	70.8 (25.3)	73.6 (25.4)	75.6 (23.5)	72.8 (20.0)	74.5 (23.5)
Percent CU	46.1 (26.7)	47.7 (32.6)	50.0 (34.6)	34.6 (28.7)	33.4 (31.6)	39.7 (26.8)	41.6 (26.8)	45.7 (30.6)	46.7 (32.3)
CAT narrative analyses	24.4 (18.0)	**31.6* (25.2)**	**30.6* (20.1)**	23.5 (23.5)	$27.6^{\dagger}$ (24.7)	26.0 (29.6)	21.9 (14.0)	28.0 (18.1)	$28.0^{\dagger}$ (19.0)

Note. T1 = Time 1 (pretreatment for treated groups), T2 = Time 2 (immediately posttreatment), T3 = Time 3 (6 weeks posttreatment). CAT = Comprehensive Aphasia Test.
*Significant effects that survived the Holm correction (bold). $^{\dagger}p < .05$ but did not survive Holm correction.

Even though percent CIU showed no significant changes, both treated and control groups showed significantly increased scores in several related measures (e.g., number of CIUs, number of words) between testing times. The lack of treatment effect shown in the CIU measures could partly reflect limitations of the measure or our data set. First, CIUs may not have been a stable enough outcome measure for the data in this study. Boyle (2014) demonstrated that CIU analyses are most stable when at least five narrative samples are considered. Even then, percent CIU was not considered to show adequate test–retest reliability to interpret *individual* data. Brookshire and Nicholas (1994b) showed that percent CIU is stable when 10 narrative samples are collected. Collecting and analyzing five to 10 narrative samples was not feasible in this study. At each testing point, data from all participants were collected within a 2-week period (due to the randomized group design). This time constraint limited the number of measures we could collect. Furthermore, with 48 participants, five elicitation stimuli would translate into 240 samples at each time point.

Another consideration is the number of words in the three narrative samples. Brookshire and Nicholas (1994a) suggested that samples should contain 300–400 words to achieve adequate test–retest reliability. The average length of the three narrative samples (summed, rather than averaged) in this study ranged from 294 words to 505 words across the three groups and time points.[1] The fact that the narratives fell within the range suggested by Brookshire and Nicholas suggests that instability of the measure may not underlie our findings. However, we are hesitant to make strong claims because the greatest number of words came from the Cinderella retell stories (range: 159–300 words), and this narrative prompt differs from the others with respect to both memory demands and lexical frequency. Thus, we cannot rule out the possibility that the measures were not stable enough to detect meaningful change.

Studies that have reported posttreatment improvements in CIU typically use single-subject designs with no control group and a relatively small number of participants, which allows for a greater number of narrative samples to be collected and coded (e.g., Antonucci, 2009; Ballard & Thompson, 1999; Boyle & Coelho, 1995; Kendall et al., 2008; Rider et al., 2008). One exception is a recent analysis of narrative language samples collected as part of Elman and Bernstein-Ellis's (1999a, 1999b) project. Boyle et al. (2021) reported that CIUs per minute and percent CIU improved following group conversation treatment in 23 PWA. Participants in the delayed treatment condition did not show a significant change over the same time period, but this was based on a smaller sample (11 PWA). They used Bayesian analyses, which may mitigate the effects of a small sample size. Their treatment also differed considerably in duration, weekly intensity ("dose frequency"; Warren et al., 2007), and overall amount of treatment ("cumulative intervention intensity"; Warren et al., 2007) compared to this study. Participants in Boyle et al. (2021) received treatment over a period of 4 months (vs. 10 weeks in our study) and received 5 hr per week (vs. our 2 hr per week), for a total of about 80 hr of treatment (vs. our 20 hr). Thus, the lack of effects on the CIU measures in this study could reflect poor validity or reliability of these measures, the relatively small number of individuals and narrative samples within each study condition, or insufficient treatment amount or intensity. A related limitation of this study relates to the availability of reliability statistics. Two independent research assistants coded the discourse measures of all samples. These raters then met to resolve discrepancies. The downside of our approach was that we did not retain the original files, which prevented us from calculating formal reliability statistics for our discourse measures.

The second method of analyzing discourse samples was an adaptation of the narrative scoring method in the CAT. Our previous work, which focused on results of standardized testing, revealed evidence of treatment effects in this measure on the picture stimulus included in the CAT (DeDe et al., 2019). Thus, we decided to examine whether this discourse coding method was sensitive to treatment changes when applied to other stimuli. Unfortunately, when we averaged the results from two additional stimuli (Cat Rescue scene and Cinderella retell) for this study, we found that the control group showed more evidence of change than either treated group. This result is difficult to interpret. On the one hand, this result dampens our confidence in the previous finding because it was based on only one stimulus. On the other hand, the CAT was standardized using their picture stimulus, and the method may not be generalizable to other stimuli. Notably, the method involves combining a count of "appropriate information content words" with Likert scale measures of grammaticality, syntax, and speed. The relative contributions of these Likert scale measures may be disproportionately reduced in a Cinderella retell story, because there are a potentially larger number of appropriate ICWs.

Our third and final discourse method was CUs (Edmonds et al., 2009). Here, we found significant changes in the large group for number of utterances and number of relevant utterances (at posttesting), and for the dyads in number of CUs (at maintenance). The control group showed no significant changes on any measure between testing intervals. Interestingly, significant changes were observed only in measures that reflect raw number of productions rather than percentages. This pattern suggests that the PWA spoke more following treatment, but that the percent of CUs (contained both +SV and +REL) did not change.

[1]Values in Tables 2 and 3 are averaged across the three speech samples, which is why these numbers do not directly map onto the ones reported there.

Given the lack of consistent findings across these three methods and the relatively small sample size, it is not possible to draw firm conclusions regarding effects of conversation group treatment on monologic discourse. However, the results suggest that measures related to CUs are more sensitive to the effects of conversation treatment in monologic discourse compared to the CIU analyses and CAT narrative analysis method. One interesting trend was that changes were more common in absolute rather than relative values (e.g., number of CIU vs. percent CIU, number of relevant utterances vs. percent relevant utterances). It may be that conversation treatment affects the overall amount of language produced rather than efficiency or informativeness of discourse. It might also be that more hours of treatment are needed to cause changes in efficiency of verbal communication, as was observed by Boyle et al. (2021). One possibility is that conversation treatment first increases the amount of language produced, with increased efficiency emerging over time. Here we define efficiency as the proportion of relevant and informative words relative to the overall number of words produced (Savage & Donovan, 2017). PWA sometimes abandon utterances when they experience anomia. If conversation treatment increases their willingness to persist in the discourse, then they may increase output, without necessarily increasing efficiency, especially if the PWA uses circumlocutions as a strategy to facilitate word retrieval. Over time, this additional word retrieval practice could facilitate access to accurate content words, leading to increased efficiency. As such, it may be that a longer treatment period or an increased frequency of conversation group treatment is needed to achieve greater effects in measures of efficiency. Further research is needed to investigate these ideas.

Another consideration is whether the lack of significant changes reflects the diversity of the participants included in our study. Participants ranged from relatively mild to severe aphasia, meaning that very different goals were addressed for these participants within conversation treatment. As a result, it might be difficult to identify a single outcome measure that would be sensitive to group-level changes. Importantly, conversation goals for more moderately and severely impaired participants often included integration of multimodal communication, including gestures and drawing. None of the discourse measures reported here allow coding of nonverbal means of communication. Including nonverbal communication in the discourse measures might increase sensitivity to changes associated with conversation treatment.

A final notable consideration is the difference between conversational discourse and monologic discourse. Authentic conversation in the context of group treatment involves many different types of utterances, such as substantive, maintenance, and minimal turns. These utterances are tightly coupled; each turn is shaped by the previous utterance and will influence the subsequent turn (Clark, 1996; Horton, 2017). Conversations may incorporate narratives, if one individual shares information (e.g., a personal narrative) with their conversation partner. However, it might be that monologic discourse measures fail to capture other potential effects of conversation treatment such as number and/or type of multimodal communicative turns.

The issues discussed above are also relevant to the hypothesized pathway presented in Figure 1. DeDe et al. (2019) reported that treated groups showed significant changes on standardized outcome measures and a patient reported outcome. In contrast, the control group showed no significant evidence of change on the standardized outcome measures. Furthermore, this work suggested that group size influences the degree of change in different treatment targets (DeDe et al., 2019). Individuals in dyads tended to show greater change on more discrete language measures, such as noun and verb naming and repetition skills. In contrast, those in the large group condition showed greater change in broader measures such as the CAT picture description and ACOM. Data from the larger discourse analyses reveal mixed results. However, the CU method revealed changes in the dyad group at follow-up testing in number of CUs, and the large group showed significant changes on the number of utterances and the relevance of utterances at posttreatment testing. The lack of consistent results across coding different methods could reflect the measure, the treatment, or both. Further research is needed to investigate these issues.

In sum, conversation group treatment is a complex, multifaceted intervention. Evidence suggests that participation in conversation-based group treatment improves aspects of communication skills in a cost-effective manner, but we do not yet understand the critical ingredients within this intervention for varied profiles of aphasia. Additional research is needed to refine measures of monologic discourse. In addition, larger controlled studies are needed to systematically explore the relative contributions of variables highlighted in Figure 1 as well as effects of individual differences such as aphasia severity. Such work would have a significant impact on practice and assessment guidelines for clinicians, and thus, have the potential to improve quality of life for PWA.

Author Contributions

Elizabeth Hoover: Conceptualization (Equal), Data curation (Equal), Formal analysis (Equal), Funding acquisition (Equal), Investigation (Equal), Methodology (Equal), Validation (Equal), Visualization (Equal), Writing – original draft (Lead), Writing – review & editing (Lead). **Gayle DeDe:** Conceptualization (Equal), Data curation (Equal), Formal analysis (Equal), Funding acquisition (Equal),

Investigation (Equal), Methodology (Equal), Project administration (Equal), Resources (Equal), Writing – original draft (Supporting), Writing – review & editing (Equal). **Edwin Maas:** Conceptualization (Supporting), Writing – review & editing (Supporting).

References

Antonucci, S. M. (2009). Use of semantic feature analysis in group aphasia treatment. *Aphasiology, 23*(7–8), 854–866. https://doi.org/10.1080/02687030802634405

Babbitt, E. M., Heinemann, A. W., Semik, P., & Cherney, L. R. (2011). Psychometric properties of the Communication Confidence Rating Scale for Aphasia (CCRSA): Phase 2. *Aphasiology, 25*(6–7), 727–735. https://doi.org/10.1080/02687038.2010.537347

Ballard, K. J., & Thompson, C. K. (1999). Treatment and generalization of complex sentence Production in agrammatism. *Journal of Speech, Language, and Hearing Research, 42*(3), 690–707. https://doi.org/10.1044/jslhr.4203.690

Biel, M. (2017). *The role of personal motivation and self-determination.* Talk presented at Aphasia Access Summit, Baltimore, MD.

Boyle, M. (2014). Test–retest stability of word retrieval in aphasic discourse. *Journal of Speech, Language, and Hearing Research, 57*(3), 966–978. https://doi.org/10.1044/2014_JSLHR-L-13-0171

Boyle, M. (2020). Choosing discourse outcome measures to assess clinical change. *Seminars in Speech and Language, 41*(01), 001–009. https://doi.org/10.1055/s-0039-3401029

Boyle, M., Akers, C. M., & Elman, R. J. (2021). *Changes in informativeness following conversation-based group treatment for chronic aphasia* [Paper presentation]. 50th Annual Clinical Aphasiology Conference (virtual).

Boyle, M., & Coelho, C. A. (1995). Application of semantic feature analysis as a treatment for aphasic dysnomia. *American Journal of Speech-Language Pathology, 4*(4), 94–98. https://doi.org/10.1044/1058-0360.0404.94

Brookshire, R. H., & Nicholas, L. E. (1994a). Speech sample size and test-retest stability of connected speech measures for adults with aphasia. *Journal of Speech and Hearing Research, 37*(2), 399–407. https://doi.org/10.1044/jshr.3702.399

Brookshire, R. H., & Nicholas, L. E. (1994b). Test-retest stability of measures of connected speech in aphasia. *Clinical Aphasiology, 22,* 119–133.

Brummett, B. H., Barefoot, J. C., Siegler, I. C., Clapp-Channing, N. E., Lytle, B. L., Bosworth, H. B., Williams, R. B., Jr., & Mark, D. B. (2001). Characteristics of socially isolated patients with coronary artery disease who are at elevated risk for mortality. *Psychosomatic Medicine, 63*(2), 267–272. https://doi.org/10.1097/00006842-200103000-00010

Bryant, L., Ferguson, A., & Spencer, E. (2016). Linguistic analysis of discourse in aphasia: A review of the literature. *Clinical Linguistics & Phonetics, 30*(7), 489–518. https://doi.org/10.3109/02699206.2016.1145740

Butler, R. N., Forette, F., & Greengross, B. S. (2004). Maintaining cognitive health in an ageing society. *The Journal of the Royal Society for the Promotion of Health, 124*(3), 119–121. https://doi.org/10.1177/146642400412400312

Chapey, R., Duchan, J., Elman, R., Garcia, L., Kagan, A., Lyon, J., & Simmons-Mackie, N. (2008). Life participation approach to aphasia: A statement of values for the future. In R. Chapey (Ed.), *Language intervention strategies in aphasia and related neurogenic communication disorders* (5th ed., pp. 279–289). Lippincott, Williams & Wilkins.

Clark, H. (1996). *Using language*. Cambridge University Press.

Cohen, S., Doyle, W. J., Skoner, D. P., Rabin, B. S., & Gwaltney, J. M., Jr. (1997). Social ties and susceptibility to the common cold. *JAMA, 277*(24), 1940–1944. https://doi.org/10.1001/jama.1997.03540480040036

Cruice, M., Botting, N., Marshall, J., Boyle, M., Hersh, D., Pritchard, M., & Dipper, L. (2020). UK speech and language therapists' views and reported practices of discourse analysis in aphasia rehabilitation. *International Journal of Language & Communication Disorders, 55*(3), 417–442. https://doi.org/10.1111/1460-6984.12528

Czajkowski, S. M., Powell, L. H., Adler, N., Naar-King, S., Reynolds, K. D., Hunter, C. M., Laraia, B., Olster, D. H., Perna, F. M., Peterson, J. C., Epel, E., Boyington, J. E., & Charlson, M. E. (2015). From ideas to efficacy: The ORBIT model for developing behavioral treatments for chronic diseases. *Health Psychology, 34*(10), 971–982. https://doi.org/10.1037/hea0000161

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

DeDe, G., Hoover, E., & Maas, E. (2019). Two to tango or the more the merrier? A randomized controlled trial of the effects of group size in aphasia conversation treatment on standardized tests. *Journal of Speech, Language, and Hearing Research, 62*(5), 1437–1451. https://doi.org/10.1044/2019_JSLHR-L-18-0404

Dipper, L. T., & Pritchard, M. (2017). Discourse: Assessment and therapy. In F. Dreux & M. Fernandes (Eds.), *Advances in speech-language pathology*. IntechOpen. https://doi.org/10.5772/intechopen.69894

Edmonds, L. 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

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

Elman, R. J. (2004). Group treatment and jazz: Some lessons learned. In J. F. Duchan & S. Byng (Eds.), *Challenging Aphasia Therapies* (pp. 144–147). Psychology Press.

Elman, R. J. (2007). The importance of aphasia group treatment for rebuilding community and health. *Topics in Language Disorders, 27*(4), 300–308. https://doi.org/10.1097/01.TLD.0000299884.31864.99

Elman, R. J., & Bernstein-Ellis, E. (1999a). The efficacy of group communication treatment in adults with chronic aphasia. *Journal of Speech, Language, and Hearing Research, 42*(2), 411–419. https://doi.org/10.1044/jslhr.4202.411

Elman, R. J., & Bernstein-Ellis, E. (1999b). Psychosocial aspects of group communication treatment - Preliminary findings. *Seminars in Speech and Language, 20*(1), 65–72. https://doi.org/10.1055/s-2008-1064009

Elman, R. J., & Hoover, E. L. (2013). Integrating communication support into aphasia group treatment. In N. Simmons-Mackie, J. King, & D. Beukelman (Eds.), *Supporting communication for adults with acute and chronic aphasia* (pp. 189–220). Brookes.

Hart, T., Dijkers, M. P., Whyte, J., Turkstra, L. S., Zanca, J. M., Packel, A., van Stan, J. H., Ferraro, M., & Chen, C. (2019). A theory-driven system for the specification of rehabilitation treatments. *Archives of Physical Medicine and Rehabilitation, 100*(1), 172–180. https://doi.org/10.1016/j.apmr.2018.09.109

Heikkinen, R. L., & Kauppinen, M. (2004). Depressive symptoms in late life: A 10-year follow-up. *Archives of Gerontology and Geriatrics, 38*(3), 239–250. https://doi.org/10.1016/j.archger.2003.10.004

Hilari, K., & Northcott, S. (2006). Social support in people with chronic aphasia. *Aphasiology, 20*(1), 17–36. https://doi.org/10.1080/02687030500279982

Holland, A. (1980). *Communication activities of daily living*. Pro-Ed.

Holland, A. L., Fromm, D., & Wozniak, L. (2018). *Communication Activities of Daily Living–Third Edition (CADL-3)*. Pro-Ed.

Horton, W. S. (2017). Theories and approaches to the study of conversation and interactive discourse. In M. F. Schober, D. N. Rapp, & M. A. Britt (Eds.), *The Routledge handbook of discourse processes* (2nd ed., pp. 22–68). Routledge Press. https://doi.org/10.4324/9781315687384

Hoover, E. L., Caplan, D., Waters, G., & Budson, A. (2015). Effects of impairment-based individual and socially oriented group therapies on verb production in aphasia. *Aphasiology, 29*(7), 781–798. https://doi.org/10.1080/02687038.2014.989953

Hula, W. D., Doyle, P. J., Stone, C. A., Hula, S. N. A., Kellough, S., Wambaugh, J. L., Ross, K. B., Schumacher, J. G., & Jacque, A. S. (2015). The Aphasia Communication Outcome Measure (ACOM): Dimensionality, item bank calibration, and initial validation. *Journal of Speech, Language, and Hearing Research, 58*(3), 906–919. https://doi.org/10.1044/2015_JSLHR-L-14-0235

Kagan, A., Black, S. E., Duchan, J. F., Simmons-Mackie, N., & Square, P. (2001). Training volunteers as conversation partners using "supported conversation for adults with aphasia" (SCA): A controlled trial. *Journal of Speech, Language, and Hearing Research, 44*(3), 624–638. https://doi.org/10.1044/1092-4388(2001/051)

Kagan, A., Simmons-Mackie, N., Rowland, A., Huijbregts, M., Shumway, E., McEwen, S., & Dickey, L. (2010). *Assessment for Living with Aphasia (ALA)*. Aphasia Institute.

Kagan, A., Simmons-Mackie, N., & Victor, J. C. (2018). *Assessment for Living with Aphasia* [ALA] (2nd ed.). Aphasia Institute.

Kendall, D. L., Rosenbek, J. C., Heilman, K. M., Conway, T., Klenberg, K., Rothi, L. J. G., & 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., & Thompson, C. K. (2019). Neuroplasticity of language networks in aphasia: Advances, updates and future challenges. *Frontiers in Neurology, 10,* 295. https://doi.org/10.3389/fneur.2019.00295

Koestner, R., Otis, N., Powers, T. A., Pelletier, L., & Gagnon, H. (2008). Autonomous motivation, controlled motivation, and goal progress. *Journal of Personality, 76*(5), 1201–1230. https://doi.org/10.1111/j.1467-6494.2008.00519.x

Leaman, M. C., & Edmonds, L. A. (2019a). Conversation in aphasia across communication partners: Exploring stability of microlinguistic measures and communicative success. *American Journal of Speech-Language Pathology, 28*(1S), 359–372. https://doi.org/10.1044/2018_AJSLP-17-0148

Leaman, M. C., & Edmonds, L. A. (2019b). Revisiting the correct information unit: Measuring informativeness in unstructured conversations in people with aphasia. *American Journal of Speech-Language Pathology, 28*(3), 1099–1114. https://doi.org/10.1044/2019_AJSLP-18-0268

Linnick, A., Bastiaanse, R., & Hohle, B. (2016). Discourse production in aphasia: A current review of theoretical and methodological challenges. *Aphasiology, 30*(7), 765–800. https://doi.org/10.1080/02687038.2015.1113489

MacWhinney, B., Fromm, D., Forbes, M., & Holland, A. (2011). AphasiaBank: Methods for studying discourse. *Aphasiology, 25*(11), 1286–1307. https://doi.org/10.1080/02687038.2011.589893

Mason, C., Nickels, L., & McDonald, B. (2020). An exploration of the impact of group treatment for aphasia on connected speech. *Journal of the International Neuropsychological Society, 25,* 72–85.

McCall, D. M., Szabo, G., & Holland, A. (2014). *Assessing long-term impact of aphasia center participation*. Clinical Aphasiology Conference Presentation, St. Simons Island, GA, United States.

Nicholas, L. E., & Brookshire, R. H. (1993). A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. *Journal of Speech and Hearing Research, 36*(2), 338–350. https://doi.org/10.1044/jshr.3602.338

Oelschlaeger, M. L., & Thorne, J. C. (1999). Application of the correct information unit analysis to the naturally occurring conversation of a person with aphasia. *Journal of Speech, Language, and Hearing Research, 42*(3), 636–648. https://doi.org/10.1044/jslhr.4203.636

Parr, S. (2007). Living with severe aphasia: Tracking social exclusion. *Aphasiology, 21*(1), 98–123. https://doi.org/10.1080/02687030600798337

Pritchard, M., Hilari, K., Cocks, N., & Dipper, L. (2017). Reviewing the quality of discourse information measures in aphasia. *International Journal of Language & Communication Disorders, 52*(6), 689–732. https://doi.org/10.1111/1460-6984.12318

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)

Ryff, C. D., Singer, B. H., & Dienberg Love, G. (2004). Positive health: Connecting well-being with biology. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences, 359*(1449), 1383–1394. https://doi.org/10.1098/rstb.2004.1521

Savage, M. C., & Donovan, N. J. (2017). Comparing linguistic complexity and efficiency in conversations from stimulation and conversation therapy in aphasia. *International Journal of Language & Communication Disorders, 52*(1), 21–29. https://doi.org/10.1111/1460-6984.12252

Savage, M. C., Donovan, N. J., & Hoffman, P. R. (2014). Preliminary results from conversation therapy in two cases of aphasia. *Aphasiology, 28*(5), 616–636. https://doi.org/10.1080/02687038.2013.843153

Seeman, T. E. (2000). Health promoting effects of friends and family on health outcomes in older adults. *American Journal of Health Promotion, 14*(6), 362–370. https://doi.org/10.4278/0890-1171-14.6.362

Shadden, B. B., & Agan, J. P. (2004). Renegotiation of identity. *Topics in Language Disorders, 24*(3), 174–186. https://doi.org/10.1097/00011363-200407000-00005

Shadden, B. B., & Hagstrom, F. (2007). The role of narrative in the life participation approach to aphasia. *Topics in Language Disorders, 27*(4), 324–338. https://doi.org/10.1097/01.TLD.0000299887.24241.39

Sherbourne, C. D., & Stewart, A. L. (1991). The MOS social support survey. *Social Science & Medicine, 32*(6), 705–714. https://doi.org/10.1016/0277-9536(91)90150-B

Siegler, R. S. (1991). Strategy choice and strategy discovery. *Learning and Instruction, 1*(1), 89–102. https://doi.org/10.1016/0959-4752(91)90020-9

Simmons-Mackie, N. (2018). *The state of aphasia in North America: A white paper*. Aphasia Access. https://www.aphasiaaccess.org/white-papers/

Simmons-Mackie, N., Code, C., Armstrong, E., Stiegler, L., & Elman, R. J. (2002). What is aphasia? Results of an international survey. *Aphasiology, 16*(8), 837–848. https://doi.org/10.1080/02687030244000185

Simmons-Mackie, N., Savage, M. C., & Worrall, L. (2014). Conversation therapy for aphasia: A qualitative review of the literature. *International Journal of Language & Communication Disorders, 49*(5), 511–526. https://doi.org/10.1111/1460-6984.12097

Swinburn, K., Howard, D., & Porter, G. (2004). *The Comprehensive Aphasia Test (CAT)*. Psychology Press.

Thompson, C. K., Shapiro, L. P., 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 Stan, J. H., Dijkers, M. P., Whyte, J., Hart, T., Turkstra, L. S., Zanca, J. M., & Chen, C. (2019). The rehabilitation treatment specification system: Implications for improvements in research design, reporting, replication, and synthesis. *Archives of Physical Medicine and Rehabilitation, 100*(1), 146–155. https://doi.org/10.1016/j.apmr.2018.09.112

Vickers, C. P. (2010). Social networks after the onset of aphasia: The impact of aphasia group attendance. *Aphasiology, 24*(6–8), 902–913. https://doi.org/10.1080/02687030903438532

Warren, S. F., Fey, M. E., & Yoder, P. J. (2007). Differential treatment intensity research: A missing link to creating optimally effective communication interventions. *Mental Retardation and Developmental Disabilities Research Reviews, 13*(1), 70–77. https://doi.org/10.1002/mrdd.20139

Wilson, R. S., Krueger, K. R., Arnold, S. E., Schneider, J. A., Kelly, J. F., Barnes, L. L., Tang, Y., & Bennett, D. A. (2007). Loneliness and risk of Alzheimer disease. *Archives of General Psychiatry, 64*(2), 234–240. https://doi.org/10.1001/archpsyc.64.2.234