# Conversation Therapy with People with Aphasia and Conversation Partners using Video Feedback: A Group and Case Series Investigation of Changes in Interaction

Wendy Best[1]*, Jane Maxim[1], Claudia Heilemann[1], Firle Beckley[1], Fiona Johnson[2], Susan I. Edwards[1,3], David Howard[4] and Suzanne Beeke[1]

[1] Department of Language and Cognition, Division of Psychology and Language Sciences, University College London, London, UK, [2] Homerton University Hospital NHS Foundation Trust, London, UK, [3] Psychology and Clinical Language Sciences, University of Reading, Reading, UK, [4] Speech and Language Sciences, School of Education, Communication and Language Sciences, University of Newcastle upon Tyne, Newcastle upon Tyne, UK

Conversation therapies employing video for feedback and to facilitate outcome measurement are increasingly used with people with post-stroke aphasia and their conversation partners; however the evidence base for change in everyday interaction remains limited. We investigated the effect of Better Conversations with Aphasia (BCA), an intervention that is freely available online at https://extend.ucl.ac.uk/. Eight people with chronic agrammatic aphasia, and their regular conversation partners participated in the tailored 8 week program involving significant video feedback. We explored changes in: (i) conversation facilitators (such as multi-modal turns by people with aphasia); and (ii) conversation barriers (such as use of test questions by conversation partners). The outcome of intervention was evaluated directly by measuring change in video-recorded everyday conversations. The study employed a pre-post design with multiple 5 minute samples of conversation before and after intervention, scored by trained raters blind to the point of data collection. Group level analysis showed no significant increase in conversation facilitators. There was, however, a significant reduction in the number of conversation barriers. The case series data revealed variability in conversation behaviors across occasions for the same dyad and between different dyads. Specifically, post-intervention there was a significant increase in facilitator behaviors for two dyads, a decrease for one and no significant change for five dyads. There was a significant decrease in barrier behaviors for five dyads and no significant change for three dyads. The reduction in barrier behaviors was considerable; on average change from over eight to fewer than three barrier behaviors in 5 minutes of conversation. The pre-post design has the limitation of no comparison group. However, change occurs in targeted conversational behaviors and in people with chronic aphasia and their partners. The findings suggest change can occur after eight therapy sessions and have implications for clinical practice. A reduction in

barrier behaviors may be easier to obtain, although the controlled case series results demonstrate a significant increase in conversation facilitators is also possible. The rehabilitation tool is available online and video technology was central to delivering intervention and evaluating change.

# INTRODUCTION

Interventions with people with aphasia (PWA) harness technology in increasingly diverse ways to deliver individualized impairment based programs (see for example Nouwens et al., 2013; Des Roches et al., 2015; Palmer et al., 2015), rehabilitate functional communication (Bilda, 2011; Marshall et al., 2013), promote social participation (Wilson et al., 2015), provide education (Rose et al., 2010) and enable remote delivery and continued rehabilitation in the context of limited resource (van de Sandt-Koenderman, 2011; Woolf et al., 2016).

In recent years, there has been increasing interest in conversation-focused aphasia interventions underpinned by a number of different theoretical perspectives (see Simmons-Mackie et al., 2014) and with a primary focus on training the communication partner (CP). A systematic review by Simmons-Mackie et al. (2016) concludes that CP training is effective, and is likely to improve participation in conversations for people with chronic aphasia. Although perhaps not emphasizing the use of technology, from the outset conversation-focused aphasia interventions have deployed video feedback for raising CP awareness of maladaptive conversation behaviors, and for outcome measurement (see reviews by Turner and Whitworth, 2006; Simmons-Mackie et al., 2014, 2016). Approaches based on conversation analysis (CA) view video feedback as key to the success of this therapy, and make use of it to enhance the learning of PWA as well as CPs (for a review see Wilkinson, 2014). However, whilst there is evidence that video feedback results in increased self-awareness when compared to verbal feedback alone in traumatic brain injury rehabilitation targeting occupational participation in daily activities (Schmidt et al., 2011, 2012), there is limited evidence for the effectiveness of video feedback in aphasia rehabilitation. In addition, a recent survey of UK speech-language pathologists (SLPs) suggests that, even when recommended as part of a published therapy program, video feedback on conversation may not be routinely employed in clinical practice (Beckley et al., 2016).

This article reports on findings from a UK Stroke Association-funded project to evaluate a conversation-focused intervention designed to train a CP and a PWA to use self-selected strategies to facilitate conversation, and to decrease behaviors that act as barriers to conversation. Video feedback is a key feature of the intervention, Better Conversations with Aphasia (BCA; Beeke et al., 2013), which has its theoretical roots in CA (for an overview of CA-informed aphasia therapy see Wilkinson, 2014). The impact of the intervention on interaction is evaluated using videos of everyday conversation.

Participants in this study have agrammatic aphasia, classically characterized as non-fluent with associated "telegraphic" language output and word order and morphological errors, alongside relatively spared comprehension (Caramazza and Zurif, 1976; Bastiaanse and Edwards, 2004). Interventions for agrammatism, once only targeted at surface grammar, now demonstrate the benefits of rigorous theoretically motivated therapy aimed at underlying syntax (see for example, Thompson and Shapiro, 2007). Whilst there is evidence that grammatical ability in the clinical setting can be improved in this way, it has proved hard to detect carryover to everyday conversation; indeed there are few studies where this is explicitly evaluated. It may be that generalization occurs but much research has failed to capture it (Carragher et al., 2015). In the field of anomia therapy, there is evidence that work on retrieving single words can influence connected speech tasks and conversation (Conroy et al., 2009; Best et al., 2011). However, given what we understand about conversation in aphasia (for a review see Beeke, 2012), it is possible that carryover from grammatical therapies is not supported because interventions are based on utterances stripped of natural interactional context. CA research has demonstrated that utterances produced by agrammatic speakers in everyday conversation with a family member differ significantly from utterances elicited by decontextualized assessment and therapy tasks (Beeke et al., 2003, 2007; Heeschen and Schegloff, 2003).

The BCA intervention targets conversation, attempting to change behaviors known to influence the flow of conversations when a speaker has aphasia (Wilkinson, 2014). Specifically, it aims to reduce conversation barriers, such as "test" questions, asked despite the CP already knowing the answer, and to increase conversation facilitators, such as the use of gesture or writing within the turn of a PWA. In their review of conversation focused interventions, Wilkinson and Wielaert (2012) concluded there was evidence that interventions involving feedback informed by CA resulted in change in the conversational behaviors of the PWA and/or their CP, both in terms of a reduction in barriers for CPs and an increase in facilitators for both PWAs and CPs. The authors point out that for some studies the evidence was qualitative, while others combined quantitative and qualitative analyses of change; no group studies were included. Therefore this study now reports findings at group and case series level; see Thompson (2006) and Howard et al. (2015) for discussion of case and case series methodology.

When targeting conversation directly, an interesting and under researched question arises as to whether one would predict changes on formal language testing as a result. A study

that speaks to this is Wilkinson et al. (2010). In this single case study there was no reported change on linguistic and cognitive tasks after conversation focused intervention. However, the participant with chronic aphasia reportedly produced more turns containing "sentences" after therapy (no further detail is supplied). More recently, Wilkinson et al. (2011) trained a person with chronic aphasia to use topic alerters such as "by the way" as a method of establishing topic. The authors reported significant improvement in the PWA's ability to name and read aloud real words after therapy, although they do not account for this aspect of their findings. Thus there does appear to be some evidence for change on tasks tapping language impairment as a result of intervention targeting conversation, but it is very limited. One might hypothesize a mechanism whereby language tasks linked to conversation strategy use might improve, for example written picture naming in a PWA who chooses to practise the strategy of writing during conversational turns. However, with the focus of intervention squarely on interaction and not language processing, changes in reading aloud as found by Wilkinson et al. (2011) are harder to predict or explain. We selected digit span as a sensitive control task. Historically, aphasia has been viewed as an impairment of memory and it is well established that there is a correlation between performance on verbal short term memory tasks such as digit span and performance on language tasks in PWA (Salis et al., 2015). While assessment of short term memory is not without issues, digit span is widely used in research and practice, with some tests of this showing strong construct validity and predictive and discriminant validity (Murray et al., 2016). A further consideration in task selection was that the participants would be likely to be neither consistently at floor or ceiling on the task, allowing scope for change. Finally, there is a small but growing literature on short term memory intervention with PWA; several studies demonstrate improved digit span post-intervention (for an overview, see Salis et al., 2015). The knowledge that performance on this task can change with targeted treatment and also that there is no evidence, to our knowledge, of conversation intervention altering auditory-verbal short term memory, makes it a suitable control task for intervention focusing on conversation.

There is increasing momentum in applying conversation-based approaches in aphasia rehabilitation. However, the issue of measuring conversational outcomes is far from straightforward and efforts to produce quantitative outcome measures have had mixed success. Ramsberger and Rende (2002) and Ramsberger and Menn (2003) devised a reliable and valid method of analysis, but chose to focus on structured dialog rather than everyday conversation, and on information transfer without consideration of the interactional features of conversation. Kagan et al. (2004) report high inter-rater reliability (IRR) for the Measure of Skill in Supported Conversation (MSC), used to rate a CP, and the Measure of Participation in Conversation (MCP), used to rate a PWA. However, these tools are not designed to measure outcomes for therapies that target the adoption of dyad-specific trained strategies, as is the case for those underpinned by CA. Wilkinson et al. (2010) used a combination of global rating

and frequency counts of individual conversational behaviors, asking a group of 15 SLPs naive to the intervention to rate conversation samples; IRR was not examined. Herbert et al. (2013) produced the "Profile of word retrieval in speech" (POWERS), also based on CA principles, which provides a reliable measure of word-retrieval in conversation (Herbert et al., 2008 explore intra- and inter-rater reliability) but the profile does not quantify strategy use. Wielaert (2016) devised a measure based on SLP judgement of paired video samples using broad CA concepts of turn-taking, repair, and conversation balance, but reported poor IRR. This literature on outcomes combined with the considerable variability found in conversation across occasions (Perkins et al., 1999) lead us to a design with multiple (six) conversations analyzed before and after intervention. Furthermore, a new conversation measure was developed using insights from POWERS and aiming to target aspects of conversation beyond word-finding. IRR for the new measure was evaluated as part of the current study.

The BCA intervention program was adapted from SPPARC (Lock et al., 2001), and aims to raise insight into the effects of aphasia on conversation, and teach strategies to allow: (i) a PWA to produce more complete, and thus successful, turns at talk, thereby increasing the likelihood of mutual understanding; and (ii) a CP to modify their responses to PWA turns, and thus enhance their partner's chance of communicating more effectively. After video-based reflection on barriers and facilitators to their conversations, both the PWA and CP (henceforth a dyad) chose strategies to practice in coached conversations with the SLP, and in home tasks between sessions. Published findings to date are detailed single (and double) cases reporting qualitative findings alongside quantitative changes in conversation behavior (Beeke et al., 2014, 2015), and qualitative investigations of conversation change (Beeke et al., 2011), and of the therapeutic learning process, including reflections on video feedback (Beckley et al., 2013). Heilemann et al. (2014) report an investigation of treatment fidelity (TF). The findings from these studies have resulted in a free e-learning resource, also called BCA, for SLPs and for PWA and their families, which aims to increase understanding of conversations with aphasia, and promote access to conversation therapy (Beeke et al., 2013). The team has also explored the mechanisms of behavior change and "active ingredients" of BCA, using theory and tools from behavior change research in health psychology (Johnson, 2015; Johnson et al., in press). This article is the first time the data from all dyads have been compiled and analyzed at the level of the group and for the case series. While this approach does not allow for the detailed consideration afforded by single cases, which include mixed qualitative and quantitative data, it is important in providing robust group-level outcomes for conversation therapy, as called for by Wilkinson and Wielaert (2012) and Simmons-Mackie et al. (2014). We do not explore outcome in relation to background assessments or patterns in conversation. Although intervention targets differed for different dyads, in this article, we step back and examine conversational change for barriers and facilitators as a whole.

The goal of the current study is to determine the effectiveness of BCA intervention for people with agrammatic aphasia and their regular CPs at a group and case series level. Additionally the study examines the conversation strategies that change (barriers and facilitators) and draws out implications for practice and for joint goal setting.

The following questions are asked in the study: (1) Does performance on language tasks and digit span change after therapy? From the literature it was hypothesized that language tasks might show change but that digit span would remain stable; (2) Does the number of conversation facilitators before intervention differ from that after intervention?; and (3) Does the number of conversation barriers before intervention differ from that post intervention? The study also explores intervention fidelity and IRR for the video-based measure of conversation.

## MATERIALS AND METHODS

### Participants

The study was granted multi-site UK National Health Service (NHS) ethical approval from Cambridgeshire 1 Research Ethics Committee (project ID 08/H0304/40) and written informed consent was obtained from all participants. Participants were people with agrammatic aphasia who had: (i) a left CVA at least 6 months prior to involvement in the study; (ii) no significant hearing loss, cognitive or psychiatric disorder which would affect participation; (iii) a spouse, family member or friend who communicated with them daily and was keen to participate; and who were (iv) not having speech and language intervention at the time of the study. Referrals came from the geographical areas of Greater London, the South, and South East of England, and were made by NHS SLPs, and by the coordinators of university aphasia clinics and stroke support groups run by charities. A diagnosis of agrammatism was made by the project team from an audio recorded picture description task, and free conversation with the research SLP.

### Intervention

#### Description of Intervention

Weekly therapy sessions of approximately 1.5 h each took place at participants' homes for 8 weeks. The PWA and CP were present for all sessions, which were designed to actively engage them both in discussion and problem solving focused on strategy use in their conversations, using video feedback alongside written materials. Session 1 introduced conversation and the impact of agrammatism, session 2 explored turns and turn construction, and session 3 repair. The next two sessions facilitated the selection (from a set of suggestions) of up to three strategies for the PWA (session 4) and the CP (session 5) to practice. Facilitation techniques included reflection on short video clips of their own conversations, to help a dyad identify positive points and challenges, and consider what they could have done differently (i.e., encourage a focus on strategy use). Session 6 explored topic, including who initiates topics and how. The final sessions (7 and 8), focused on active practice of each

person's chosen strategies in structured conversation activities. Where possible, these were videotaped to allow immediate reflection and feedback on the consequences of conversational behaviors. An overview of sessions grouped by main aim is provided in the Supplementary Material (Appendix 1). The free BCA electronic resource[1] provides further detail about the intervention, and video-recorded examples of therapy sessions.

All therapy for all dyads was carried out by the fourth author (FB) of the study. Prior to therapy, the project team met to view baseline videos recorded by a dyad, in order to identify key barriers and facilitators to conversation. The barriers and facilitators used by each dyad are listed in the Supplementary Material (Appendix 2). These determined the choice of video clips that were prepared for therapy, which formed the basis of discussion and dyad self-reflection. While the project team guided a dyad's reflection in this way, the final decision on which strategies to practice in therapy rested with the PWA and CP.

The intervention was tailored via the use of individual dyads' own conversation videos, and by giving each dyad a choice concerning strategies for practice. However, therapy plans for each session detailing generic activities congruent with underlying session aims were followed throughout. Thus, once the study had begun there was no modification to the intervention schedule or principles (TIDieR checklist item 9, Hoffmann et al., 2014).

### Fidelity of Intervention

In order to investigate how closely the actual delivery of the therapy corresponded to that which had been planned (Tate et al., 2013), aspects of TF were explored using data from seven dyads (see Heilemann et al., 2014). A rater independent of the SLP (the third author, CH) observed 23% of the 56 videotaped therapy sessions and used a BCA-specific fidelity tool in order to quantify TF. The main focus of this investigation related to two components typically associated with the concept of TF: adherence to therapy components (i.e., fidelity to the BCA session plans) and therapy delivery (i.e., therapist behavior desirable for carrying out the BCA therapy). The fidelity tool reflected these concepts in two sections: (a) procedural and (b) qualitative. In (a), the items were rated on a scale of 1, 0.5 or 0, corresponding to therapy content that was judged as being "fully", "partly" or "not delivered" in a session, and in (b), the items were rated 1, 0.5 or 0 corresponding to statements about desired therapist behavior judged to be observable "most of the time", "occasionally", or "not at all".

### Experimental Design and Processes

The design of the study is illustrated in **Figure 1**. Justification is provided under rationale for design, below, and a full explanation of the conversation sampling can be found under outcome measurement. In addition to conversations, as shown on the Figure, formal assessments were carried out both pre and post intervention: sentence comprehension (CAT; Swinburn et al., 2004), written naming (Kay et al., 1992) and digit span forward (CAT; Swinburn et al., 2004).

---

[1]https://extendstore.ucl.ac.uk/product?catalog=UCLXBCA

| | Pre-therapy Baseline Assessments (8 weeks) | | | | | | | | Therapy (8 weeks) | | | | | | | | Post-therapy Follow-up Assessments (8 weeks) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Week** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
| **Conversation** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | | | | 10 | | | | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
| **Assessment** | Pre-therapy 1 | | | Pre-therapy 2 | | | Pre-therapy 3 | | | | | | | | | | Post-therapy 1 | | | | Post-therapy 2 | | | |

**FIGURE 1 | Design of intervention study.** A total of 18 conversation samples were collected over the course of the study (eight pre-therapy, two during therapy and eight post-therapy). The conversations highlighted in the figure, six pre-therapy and six post-therapy samples, were analyzed. The figure illustrates the timing of pre-therapy baselines 1–3 and post-therapy assessments 1 and 2.

## Rationale for Design

The study employed a pre-post design with multiple conversations and assessments both pre and post intervention. Use of an experimentally stronger crossover design was considered inappropriate for two reasons. First, participants were involved in the study for 6 months on average and commitments beyond this were deemed too much to ask, particularly given the requirement for the involvement throughout of CPs, several of whom were working. Second, the reliance on the goals and strategy choices of the participants themselves and the expected generalization to everyday conversation meant it was not possible to focus first on one aspect of conversation and then cross over to another aspect, as this design would require. Impairment focused intervention was not considered a suitable control as it would have been likely to have resulted in aims contrary to those of the conversation intervention in, for example, drawing attention to aspects of language such as word-finding or sentence production rather than optimizing conversation. There was no randomization to different conditions. Experimental control (Howard et al., 2015) was exerted in the following ways. First, there was ongoing contact during the pre-therapy phase which included time with the SLP, a range of language activities and questions about aphasia and conversation. This contact provided control for any Hawthorn effect (aka SLP "charm"), as change due to general SLP contact would be observed over the baseline phase rather than after therapy. Second, assessment of digit span was carried out before and after therapy. This measure provided control for any generalized effect of intervention (for a review see Murray et al., 2016). If this was occurring, change would be predicted to occur on this sensitive task, which was not at floor or ceiling for any participant. Third, all conversations were transcribed, analyzed and scored by raters blind to point of data collection avoiding any bias that could occur if pre-post therapy status of the samples was revealed.

## Outcome Measurement

All language assessments were carried out by the fourth author (FB) of the article (**Table 1** provides details).

The variability inherent in conversation data (Perkins et al., 1999) means it is necessary to take multiple measures to reveal patterns present over and above noise, which may be occurring for reasons including participant variables (such as fatigue) and conversation variables (such as topic). Conversation assessment entailed video recording eight pre- and eight post-therapy samples and two conversations during therapy (18 in total, henceforth labeled C1–C18). Each conversation sample lasted up to 20 min. Each dyad was trained in how to operate a digital video camera. The research SLP was not present when recordings were made. Dyads were asked to record their conversations at a time when they would normally sit down and talk, for example, to catch up on events and news. Specific topics for conversation were not proposed in an attempt to make samples as ecologically valid as possible.

Counts were made of barrier and facilitator behaviors in 5-min video samples from a subset of the 18 conversations recorded. We discarded C1, the first conversation, and C18, the final post therapy sample in order to avoid any effects of the dyad settling in to recording conversations or being aware that the study was ending. Of the remaining 14 samples, 12 were rated, six pre therapy (C2–4 and C6–8) and six post therapy (C11–14 and C16–17). C5 and C15 were only used to replace missing pre-/post-therapy samples, respectively. Where a sample was 10+ min in length, sampling began from 5 min into the recording. If a conversation was less than 10 min in length, the final 5 min were used.

The counts result from the work of 17 student SLPs (15 graduate and 2 undergraduate) at University College London between 2009 and 2013 and one PhD student (the fifth author (FJ), an experienced SLP), 18 students in total. Each year, between two and six students performed ratings as part of a thesis supervised by core members of the project team (the first (WB), second (JM) and final authors (SB)). All were trained to identify conversational behaviors in the video samples. Depending on the focus and scope of each thesis, students rated between 2 and 10 samples (with an equal number taken from pre- and post-therapy data) for between two and four dyads. For each year-group of students, training led by the final author was spread over 9 months and consisted of: a 3 h introduction to CA; a 2 h transcription training session including practice with language disordered video data; a 1 h introduction to the project; a 1.5 h introduction to the outcome measure, with discussion of written guidelines

**TABLE 1 | Participant data, background information and language profiles.**

| Dyad number | Age at recruitment (gender) | Months since onset of aphasia (at time of 1st session) | Previous employment | CP relation to PWA | Pyramids and Palm Trees (out of 52)[1] | Spoken word to picture matching (out of 40)[2] | Minimal pairs discrimination (out of 40)[3] | Naming objects (averaged across 3 time points; out of 10)[4] | Naming actions (averaged across 3 time points; out of 10)[5] | Severity: Quartiles[6] |
|---|---|---|---|---|---|---|---|---|---|---|
| Dyad 1 | 49 (f) | 33 | Jazz singer | Twin | 50 | 38 | 40 | 8.33 | 2.67 | 3 (48%) |
| Dyad 2 | 39 (m) | 30 | Own business | Wife | 52 | 36 | 40 | 8.33 | 5.67 | 2 (66%) |
| Dyad 3 | 55 (m) | 59 | Senior sales manager | Wife | 51 | 39 | 40 | 9.00 | 2.33 | 3 (37%) |
| Dyad 4 | 63 (m) | 60 | Team manager NHS | Partner | 48 | 35 | 29 | 3.33 | 2.00 | 3 (31%) |
| Dyad 5 | 57 (f) | 39 | Cashier at bookmakers | Son | 51 | 24 | 30 | 5.33 | 2.00 | 3 (42%) |
| Dyad 6 | 60 (m) | 17 | Gardener/book illustrator | Wife | 50 | 39 | 38 | 4.00 | 5.00 | 2 (53%) |
| Dyad 7 | 71 (f) | 40 | Deputy head teacher | Daughter | 49 | 35 | 34 | 4.00 | 3.67 | 3 (28%) |
| Dyad 8 | 57 (m) | 10 | Self-employed van driver | Wife | 44 | 33 | 31 | 4.33 | 1.33 | 4 (24%) |

*[1]Pyramids and Palm Trees (Howard and Patterson, 1992); [2]PALPA 47 (Kay et al., 1992); [3]PALPA 4 (Kay et al., 1992); [4]and [5]Object and Action Naming Battery (OANB; Druks and Masterson, 2000); [6]The severity measure is derived from five language subtests, carried out during the first pre-assessment and chosen to represent known areas of difficulty for people with agrammatism: OANB (Druks and Masterson, 2000), Comprehension of written sentences (CAT, Swinburn et al., 2004), Writing single words to dictation (PALPA 53, Kay et al., 1992), Spoken sentence comprehension (VAST; Bastiaanse et al., 2002), Tests and subtests which showed ceiling (e.g., Pyramids and Palm Trees) and floor effects (e.g., VAST sentence production) were excluded. The severity rating is computed in quartiles, with 4 representing most severe. Further details of performance of these tests are available from the authors. f = female, m = male. For the communication partners (CPs), gender is transparent, except dyad 1 (female) and dyad 4 (male).*

giving definitions and examples of conversational behaviors, and practice identifying turns in a short video clip of aphasic conversation; a homework task to rate the same video clip for a sample of conversation behaviors; a 2 h follow-up to give feedback on their homework ratings and to practice with a second video clip, and two support workshops (totaling 3 h). In addition, students had access to a group email address permitting them to post queries to the project team and their fellow student raters as and when they wished. During the training and rating process, the written rating guidelines were modified as necessary in response to student requests for clarification, and a revised version was emailed to all students. The rating process was as follows: student raters viewed their 5-min video samples repeatedly, and then produced a turn-by-turn orthographic transcription of each sample. They then rated the sample for core conversation behaviors using the video recording alongside the transcript, and with the support of the guidelines. Importantly at the time of rating, all students were "blind" as to all sample collection dates. Overall for each barrier and facilitator total there were 96 counts (eight dyads, 12 conversations) with between two and six barriers and 7–14 facilitators contributing to each of these counts. These variations arose as a result of the individualized choices made by each dyad.

Of the 96 conversations rated, 40 (41.7%) were analyzed by more than one student; for each of the eight dyads, at least four conversation samples were analyzed by a pair of students. In these cases, after independently rating the same sample, the student pairs agreed counts for individual behaviors. Agreement was defined as both students applying the same rating to a specific behavior; consensus agreement techniques were not used. These behaviors have no natural denominator and so percentage agreement was applied (see Hartmann, 1977). IRR was measured by averaging the IRR for each behavior and then, in turn, averaging across all barrier and all facilitator items. This measure of agreement includes counts where a behavior was agreed to occur and counts where a behavior was agreed to be absent.

## Statistical Methods

Data from student projects were compiled by the third author (CH) under the supervision of the first (WB) and final authors (SB). Where difficulties in collating information from the student projects arose, they were resolved by discussion with the wider team (e.g., we agreed to use a mean count if there was more than one count available from different student projects for a specific behavior). The question of whether there was change with therapy was evaluated statistically for the group by taking the mean of pre-therapy scores and mean of post-therapy scores on each measure and comparing these employing a paired-sample t-test. This method was chosen in consultation with a statistician and in order to match the comparison made for the structured assessment tasks with that for the conversation variables. For digit span and language assessments, raw scores were entered. For conversation variables the data were first transformed by taking the square root of raw scores in order to normalize

the distribution. Effect sizes were calculated using G Power version 3.1[2].

For the case series the conversation data were analyzed using Poisson Trend Tests suitable for observations occurring in a Poisson distribution. All pre-therapy conversations were weighted the same as one another, as were all post therapy conversations. The outcomes for different participants were investigated using a Test for Homogeneity, significance demonstrated that the effects for different dyads cannot be centered on mean with variability only by chance, i.e., that there are real differences across participants. Where the Test for Homogeneity was significant the effect for different dyads was calculated by employing the Holm-Bonferroni procedure[3]. Thus the statistical procedures did not investigate trends over time but examined whether there was a difference between pre- and post-therapy scores.

## RESULTS

### Participants

We successfully recruited nine dyads to the study (henceforth D1–D9, or PWA1–PWA9 where referring solely to the participant with aphasia). Eight dyads completed the full intervention program; the ninth withdrew half way through therapy because they felt it was not suitable for them. **Table 1** provides an overview of the participant background information and scores on formal assessments. The findings are presented for the dyads in the order in which they were recruited.

PWA and their CPs varied in age and in their relationship to one another. All participants with aphasia were more than 17 months post onset at the start of their involvement in the study. On formal language assessments participants showed a range of abilities congruent with a diagnosis of agrammatic aphasia. Semantic processing of pictures as assessed by the three picture version of Pyramids and Palm Trees was near ceiling for all but PWA8. A wider range of scores was obtained on the input tasks tapping single word comprehension (word-to picture matching) and auditory discrimination (minimal pairs). Naming of objects and actions was impaired for all PWA and, for some, notably PWA1, PWA4 and PWA5, was further impaired by elements of verbal dyspraxia (based on clinical judgement; dyspraxia was not formally assessed). With regard to severity, no PWA had a score in the first or last quartile. Examples of narrative and sentence production are provided in the Supplementary Material (Appendix 3), in the form of transcriptions of the Dinner Party cartoon strip narrative (Fletcher and Birt, 1983).

---

[2]http://www.gpower.hhu.de/
[3]The Holm-Bonferroni procedure works as follows: with all eight dyads, consider the $p$-values (arising from the Poisson Trend Test) from smallest to largest. The smallest is compared with 0.05/8 and if less than this it is significant. If it significant then assess the next at $0.05/(n-1)$, if that is significant assess the next at $0.05/(n-2)$. The process was repeated until the result for the next dyad was not significant.

## Fidelity of Treatment

To investigate IRR of the fidelity tool, 20% of the sessions rated by the first rater (the third author) were observed and rated by a second rater (a trained student SLP). There was an overall percentage agreement of 86.8% and 87.5% for the procedural and the qualitative sections of the fidelity tool, respectively.

In terms of delivery of the planned components (section (a) of the fidelity tool), results indicate a high overall adherence of the SLP to the components of BCA (fidelity score: 91.9%, SD = 3.9; based on 227 observations). For section (b) of the tool, qualitative aspects of therapy delivery, a high degree of desired therapist behavior was achieved during BCA sessions (96.7%, SD = 4.1). For a detailed description of dyad-specific and overall fidelity results see Heilemann et al. (2014).

## Inter-Rater Agreement on the Conversation Measure

The overall percentage IRR for facilitators was 69% (SD = 21) and for barriers 64% (SD = 27).

## Performance on Repeated Formal Assessments

**Figure 2** illustrates the mean scores for the group of PWA on the repeated language assessments and digit span test. Full details of scores for each participant at each assessment point are provided in the Supplementary Material (Appendix 4).

Visual analysis suggests relative stability in short term memory but gradual change over the course of the study for the group, with a tendency for greater change over the therapy phase in written naming. Statistical analysis, carried out on pre vs. post scores to align the comparison with that for the conversations, shows no significant change in digit span from pre to post therapy (mean pre 4.42 (SD 3.13), post 4.25 (SD 3.28), paired-sample $t$-test $t_{(7)} = 0.247$, n.s., two-tailed, effect size Cohen's $d = 0.05$). In contrast there was a small but significant
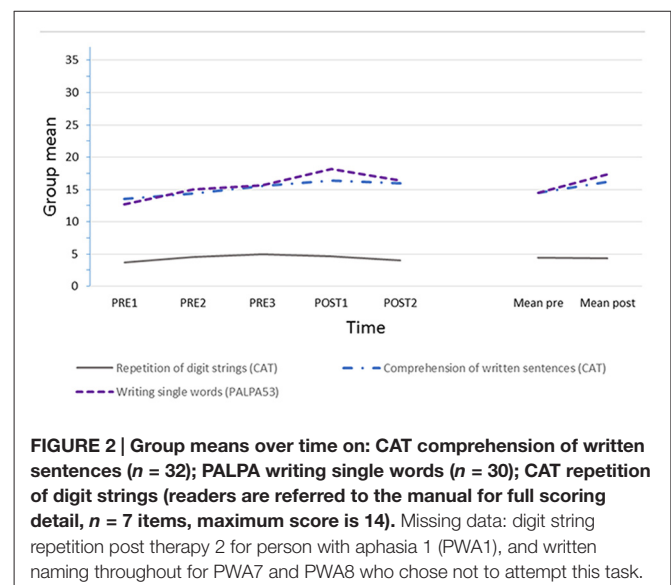
**FIGURE 2 | Group means over time on: CAT comprehension of written sentences ($n = 32$); PALPA writing single words ($n = 30$); CAT repetition of digit strings (readers are referred to the manual for full scoring detail, $n = 7$ items, maximum score is 14).** Missing data: digit string repetition post therapy 2 for person with aphasia 1 (PWA1), and written naming throughout for PWA7 and PWA8 who chose not to attempt this task.
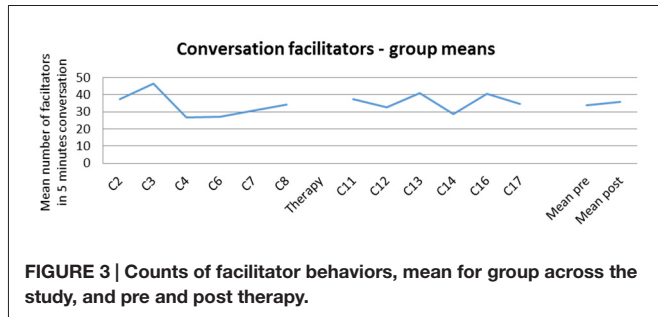
**FIGURE 3 | Counts of facilitator behaviors, mean for group across the study, and pre and post therapy.**
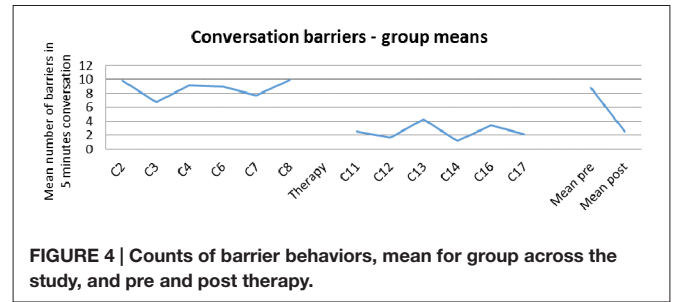


**FIGURE 4 | Counts of barrier behaviors, mean for group across the study, and pre and post therapy.**

change from pre to post therapy in average performance in written naming (pre 14.44 (SD 6.02), post 17.50 (SD 4.81), paired-sample $t$-test, $t_{(5)} = 3.051$, $p = 0.028$, two-tailed, Cohen's $d = 0.55$), and also in comprehension of written sentences (pre 14.46 (SD 4.33), post 16.63 (SD 4.27), $t_{(7)} = 2.573$, $p = 0.037$, two-tailed, Cohen's $d = 0.50$).

## Change in Use of Conversation Facilitators

The change in use of conversation facilitators for the group over the course of the study is illustrated in **Figure 3**. These facilitators include, for example, multi-modal turns (mime, gesture, drawing and writing/sky-writing) for PWA, and turns where the CP responds without the need for repair, or lets the conversation continue so the PWA can use strategies. There was no significant change in facilitators between pre and post therapy for the group (mean pre 33.72, post 35.70, paired-sample $t$-test on square-rooted values, $t = 0.73$ (7), two-tailed $p = 0.492$, n.s., Cohen's $d$ on raw values = 0.14, minimal).

The counts for conversation facilitators for each of the dyads separately are provided in **Table 2**.

The raw figures demonstrate considerable variability in facilitator behaviors both across occasions for the same dyad, and between different dyads. The dyads also show large differences in pre-post scores. This was supported by a significant test for homogeneity ($H = 60.52$, df = 7, $p < 0.0001$) meaning the $z$ scores for change calculated using the Poisson Trend

Test are not centered on a mean score with variability only by chance, but that there are real differences in change between the dyads. Follow-up analysis, exploring this and employing the Holm-Bonferroni procedure, showed statistically significant change with an increase in facilitator behaviors for D2 ($z = 4.36$, $p < 0.00,001$) and D1 ($z = 3.76$, $p < 0.001$) with D5 showing a trend to significance ($z = 1.98$, $p < 0.047$), but $p < 0.0125$ is necessary for significance once the Holm-Bonferroni procedure is applied. Dyad 7 showed a statistically significant decrease in facilitator behaviors (D7 $z = -4.09$, $p < 0.00,001$) and Dyad 8 a trend in the same direction ($z = -2.45$, $p < 0.025$, but $p < 0.01$ is required when the Holm-Bonferroni procedure is applied). D3, D4 and D6 did not show significant change between pre- and post-therapy samples in either direction.

## Change in Use of Conversation Barriers

**Figure 4** illustrates the marked and significant decrease in the number of barriers for the group (mean pre 8.73, post 2.52, $t$ = test on square-rooted values, $t = 2.71$ (7), two-tailed $p = 0.015$, sig., Cohen's $d = 0.73$, medium). The change reflects a reduction in number of barriers to less than a third of the number present before therapy.

The counts for conversation barriers for each of the dyads separately are provided in **Table 3**.

The case series data shows that, as for facilitators, considerable variability is evident in barrier behaviors and that dyads show very large differences between one another in mean

**TABLE 2 | Counts of facilitator behaviors for each dyad at each conversation.**

|  | C2 | C3 | C4 | C6 | C7 | C8 | Tx | C11 | C12 | C13 | C14 | C16 | C17 | Mean pre | Mean post |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| D1 | 44.5 | *44.5* | 49.5 | 26 | 38.5 | 32 |  | 44.5 | *44.5* | 60 | 86 | *44.5* | *44.5* | 39.17 | 54.00 |
| D2 | 45 | 42.5 | 18 | 17.5 | 5.5 | 31 |  | 52.5 | 56.5 | 15 | 26 | 64 | 33.5 | 26.58 | 41.25 |
| D3 | 42 | 93 | 42 | 56 | 51 | 52 |  | 56 | 39 | 68.5 | 27 | 39 | 85.5 | 56.00 | 52.50 |
| D4 | 23.5 | 18 | 16 | 32 | 41.5 | 57 |  | 42 | 40 | 49 | 18 | 46 | 22 | 31.33 | 36.17 |
| D5 | 19.5 | 25 | 17 | 10 | 10 | 29 |  | *19.5* | 12 | 42 | 30 | 26 | 12.5 | 18.42 | 23.67 |
| D6 | 7.5 | 15 | 31 | 9 | 2 | 18 |  | 29 | 4 | 28 | 15 | 18 | 6.5 | 13.75 | 16.75 |
| D7 | 80.5 | 74 | 14 | 40 | 62 | 32 |  | 30 | 46 | 21 | 17 | 51 | 45 | 50.42 | 35.00 |
| D8 | 36.5 | 61 | 26 | 26 | 33 | 22 |  | 24 | 17.8 | 43 | 11 | 36 | 26 | 34.08 | 26.31 |
| **Mean** | 37.38 | 46.63 | 26.69 | 27.06 | 30.44 | 34.13 |  | 37.19 | 32.48 | 40.81 | 28.75 | 40.56 | 34.44 | 33.72 | 35.70 |
| **SD** | 22.02 | 27.81 | 13.19 | 15.77 | 22.24 | 13.60 |  | 13.47 | 18.70 | 18.64 | 24.05 | 14.38 | 24.81 |  |  |

*Tx—intervention; D1–8 is dyads 1–8; C2-C17 is conversations 2–17. Numbers in italic orange reflect missing data. In each case the median score for this dyad for all the conversations that were available has been used. Note that for D8, pre-therapy conversation samples analyzed were C2, C3, C4, C5, C6 and C7; the dyad failed to record C8.*

**TABLE 3 | Counts of barrier behaviors for each dyad at each conversation.**

|      | C2    | C3   | C4    | C6   | C7    | C8   | Tx | C11  | C12  | C13  | C14  | C16  | C17  | Mean pre | Mean post |
|------|-------|------|-------|------|-------|------|----|------|------|------|------|------|------|----------|-----------|
| D1   | 5     | *5*  | 5     | 7    | 14.5  | 4    |    | 4    | *5*  | 5.5  | 2    | *5*  | *5*  | 6.75     | 4.42      |
| D2   | 0     | 1    | 1     | 2    | 1     | 0    |    | 1    | 0    | 3    | 3    | 3    | 1    | 0.83     | 1.83      |
| D3   | 8     | 0    | 12    | 0    | 5     | 21   |    | 0    | 1    | 0    | 0    | 0    | 0    | 7.67     | 0.17      |
| D4   | 10.5  | 1    | 0     | 11   | 0     | 20   |    | 2    | 1    | 0    | 0    | 7    | 0.5  | 7.08     | 1.75      |
| D5   | 12.5  | 15   | 8     | 21   | 5.5   | 11   |    | *5.5*| 1    | 0    | 0    | 0    | 1    | 12.17    | 1.25      |
| D6   | 0     | 4    | 0     | 9    | 2     | 13   |    | 1    | 0    | 0    | 0    | 0    | 5.5  | 4.67     | 1.08      |
| D7   | 0.5   | 0    | 0     | 0    | 0     | 0    |    | 0    | 1    | 0    | 0    | 1    | 0    | 0.08     | 0.33      |
| D8   | 42    | 28   | 47    | 22   | 34    | 10.5 |    | 7    | 4    | 26   | 4    | 11   | 4    | 30.58    | 9.33      |
| **Mean** | 9.81 | 6.75 | 9.13 | 9 | 7.75 | 9.94 | | 2.56 | 1.63 | 4.31 | 1.13 | 3.38 | 2.13 | 8.73 | 2.52 |
| **SD**   | 13.88 | 9.91 | 15.93 | 8.72 | 11.63 | 8.17 | | 2.64 | 1.85 | 9.00 | 1.64 | 4.03 | 2.31 | | |

*Tx—intervention; D1–8 is dyads 1–8; C2-C17 is conversations 2–17. Numbers in italic orange reflect missing data. In each case the median score for this dyad for all the conversations that were available has been used. Note that for D8, pre-therapy conversation samples analyzed were C2, C3, C4, C5, C6 and C7; the dyad failed to record C8.*

pre-post scores. This was supported by a significant test for homogeneity ($H = 95.39$, df $= 7$, $p < 0.0001$) demonstrating the $z$ scores for change calculated using the Poisson Trend Test are not centered on a mean score with variability only by chance, but that there are real differences in change between the dyads. Follow-up analysis employing the Holm-Bonferroni procedure, showed statistically significant change with a decrease in barrier behaviors for, ordered by $z$ score, the following dyads: D8, D5, D3, D4 and D6 (D8 $z = -8.24$, $p < 0.01$, D5 $z = -7.30$, $p < 0.01$, D3 $z = -6.65$, $p < 0.01$, D4 $z = -4.4$, $p < 0.025$, D6 $z = -3.66$, $p < 0.025$). D1, D2 and D7 did not show significant decrease or increase in barrier use in conversation.

## DISCUSSION

This article aimed to investigate the answers to three key questions. First, we explored whether performance on language tasks and digit span changed after conversation intervention using video feedback. The group demonstrated fairly stable performance on digit span throughout the study, with no statistically significant change between pre- and post-therapy scores. This suggests there was no change in auditory-verbal short term memory. If there had been progress, this may have reflected continued recovery or an effect of therapy and in this case the task would not have provided experimental control (Howard et al., 2015). Participants were not consistently at floor or ceiling on this task, which was selected with care, so there was scope for improvement. The stability, in spite of this, suggests the group were not within a phase of spontaneous recovery and that the intervention did not influence their short term memory/phonological processing. In this sense, it provided good experimental control, although future studies could employ multiple measures of short-term memory as advised by Murray et al. (2016).

In contrast to the stability in digit span, the group showed significantly better written naming after intervention than before, and also better comprehension of written sentences. In both cases the changes, while statistically significant, were numerically small (correct written names were provided for

three additional items and sentence comprehension improved by two items, on average) and it was only for written naming that there looked to be a tendency for change to occur over the intervention period. The result might be due to the intervention as the facilitatory behaviors targeted included use of writing or "sky-writing" (i.e., tracing letters in the air). Thus, a tentative conclusion would be that conversation intervention of the type used in this study may also result in improvement in language processing on tasks that tap into processing routes employed by the strategies targeted in conversation (see Beeke et al., 2014, 2015; for a more detailed mixed methods exploration of the writing strategy for three of the dyads in this study: D6, and D4 and D8, respectively). This finding could be explored in future research by using a range of language tasks and making specific predictions as to change according to the nature of the conversation intervention for each dyad.

Second, we asked whether the number of conversation facilitators after intervention differed from that before intervention. There was no significant change in the number of facilitators for the group (minimal numerical increase from 33.7 to 35.7 on average in 5 min of conversation). This is a new finding and it stands alone as, while there are some well controlled case studies in the literature (Carragher et al., 2015), there is no research examining change at a group level in conversation behaviors after an intervention with a direct focus on conversation. The differing findings for the experimentally controlled case series are important in highlighting the variability between dyads making up the group. The significant test for homogeneity reflects the fact that there are genuine differences in change. When employing rigorous statistical analysis with an inbuilt Holm-Bonferroni procedure to adjust for multiple dyads, we found that there was a significant increase in facilitators for two of the dyads (D1 and D2). The changes for D2 are considered using a conversation analytic method in Beeke et al. (2011). D7 showed a statistically significant *decrease* in facilitator behaviors. This is important as it shows that intervention may not always operate in the desired direction. Further qualitative analysis is necessary to shed light on the nature

of this statistically significant change. The study is novel in exploring the effects for the group and separately for the case series, and in demonstrating that findings for the participants may differ. Given the heterogeneity present in aphasia profiles, conversations (Perkins et al., 1999) and response to intervention, it is important for future studies of conversation-based interventions to present and analyze findings at both levels.

Finally, the study investigated whether the number of conversation barriers after intervention differed from that before intervention. For the group there was a significant and dramatic reduction from more than eight to less than three barrier behaviors on average in 5 min of conversation. This is a considerable change and one which, scaled up to the conversations between regular conversation partners throughout the day, is likely to have a large impact on everyday communication. The variability between the dyads is again reflected in the data from the case series. Here, for five dyads (D3, D4, D5, D6 and D8) there was a statistically significant reduction in barrier behaviors. The learning process that may underlie this change for PWA3 is considered in Beckley et al. (2013), and the changes for D4 and D8 are considered in detail in Beeke et al. (2015), and for D6 in Beeke et al. (2014). The remaining three dyads showed no change in barrier use in either direction. Thus, for facilitators as for barriers, the finding for the group is not the same as that for the separate dyads, although in this case there was no significant increase in barrier use (i.e., no change in the opposite direction to that targeted in intervention).

The contrast in findings for facilitators and barriers is important, and suggests a link with the effects of video feedback. It may be that participants found it easier, or more salient, to reduce or terminate barrier behaviors than to develop new or increased uses of trained facilitators. In a qualitative investigation of participants' experience of BCA therapy and conversation change, Johnson (2015) shows that realizations about the use and impact of barrier behaviors, prompted by video feedback, resonate very powerfully with participants and are often the thing they remember most. Alternatively it may be that the therapy process used in BCA to target barriers was more effective than that for facilitators. An investigation by Johnson et al. (in press) of behavioral change in BCA demonstrates that the mechanisms reportedly supporting change to facilitators are more numerous than those reportedly supporting change to barriers. Furthermore it appears that the mechanisms of change associated with facilitators may be used in more varied combinations during therapy. However, the most effective combination of mechanisms to support change in conversation barriers and facilitators is unknown at this stage. Clearly, changing conversation behaviors through conversation therapy is a complex process that will need further investigation to ensure it is optimally effective.

It is also possible that the process of measuring change in facilitators and barriers by using frequency counts influenced the contrast in findings for facilitators and barriers. Whereas barriers by their nature are expected to decrease or disappear, facilitators

*may* become more frequent, but they may instead be used more effectively while not increasing significantly in numbers. For example, a PWA may use writing as a strategy both before and after intervention; the effectiveness of this strategy within conversation may alter as a result of therapy, but not be reflected in counts of writing behavior. In addition, given that facilitators represent a range of behaviors already in use in conversation, it is unclear how much of an increase needs to happen for therapy to be considered successful. To counter this difficulty, we have attempted (in other publications, see for example Beeke et al., 2014, 2015) to deploy measures of qualitative change in the use of facilitators alongside quantitative methods, but further work in this challenging area of mixed methods is warranted.

In this study video was used with both PWA and CPs to raise awareness, promote reflection on the consequences of their conversational behaviors, and to permit them to reflect immediately on the successes and failures of strategy use in practice conversations. This promoted in-depth self-reflection and, by the very nature of video feedback, rendered the task concrete and accessible to PWA, not just their CPs. Arguably it would not have been possible to discuss abstract concepts such as turns and repair without using video-playback techniques, since speakers are generally unaware of the conversational rules that govern their interactions. In addition, speakers with aphasia are likely to experience significant difficulties with expressing views on conversation behaviors in the absence of concrete examples to scaffold such discussions. As Simmons-Mackie et al. (2014) point out, more conversation therapy directed at PWA is warranted. Video feedback would seem to be a crucial element to ensure the likely success of this form of strategy training for PWA. Indeed, as the case series findings show, two PWA significantly increased their use of facilitators in conversation post therapy. Given that generalization of learnt strategies to everyday interaction in aphasia rehabilitation is an ever-present challenge, understanding the role of video feedback in conversation therapy is a priority for systematic investigation in future research.

The inclusion of IRR is relatively new in the field and brings a level of confidence to the results. However, as the IRR results for our conversation outcome measure reveal, reliably rating conversation behaviors is not straightforward. One factor that influenced IRR was the natural variability of behaviors across conversation samples and dyads, a feature of conversation first systematically noted in relation to attempts at quantification by Perkins et al. (1999). Thus some behaviors were present for all dyads whereas some were specific to one dyad only, and some behaviors occurred throughout the conversation samples whereas others occurred only sporadically. Also, there is the issue of interpretation of IRR levels for conversation data. While there is no formally recognized acceptable level of agreement for measures of naturally occurring conversation, for the application of a measure of communicative informativeness and efficiency (the Correct Information Unit or CIU analysis of Nicholas and Brookshire, 1993) to conversation, Oelschlaeger and Thorne (1999) state that 80% and above should be considered acceptable. They remark that the 90% agreement levels obtained by other CIU studies are too high for an application to naturally occurring

language phenomena. In the absence of other work in this area, the 80% agreement figure is still our best benchmark. However, the Oelschlaeger and Thorne (1999) study ultimately found that reliable CIU measures could not be obtained for everyday conversation. This is not surprising given the analytic focus on transaction (of information) to the detriment of interactional features that permit the establishment and maintenance of interpersonal relationships. This calls into question whether it is sensible to take 80% as the level of acceptable agreement when rating conversation, and points to the urgent need for further investigation of the process of interpreting IRR for such data.

This study includes several "firsts" for the field of conversation therapy for aphasia. As the largest group study in the field which deploys a measure of targeted behaviors, it presents group and case series analyses of pre- and post-therapy samples of everyday conversation, the expected site of communication behavior change following this type of intervention. While 32 dyads completed the protocol for the Wielaert et al. (2014) implementation study, low IRR of a pilot outcome measure based on SLP judgement of paired video samples made it difficult to draw conclusions regarding conversation change. As Wilkinson and Wielaert (2012) and Simmons-Mackie et al. (2014) conclude, there is a dearth of larger studies providing quantitative outcome data for conversation-based rehabilitation programs. Group results often hide the improvement that can be made after intervention; in this study a case series design reveals how individual dyads can make significant gains after conversation therapy. A strength of the study is the collection and analysis of multiple pre and post-therapy conversations; this enabled patterns to emerge over and above the noise inherent in behavioral data from everyday interaction. An analysis of TF found the delivery of therapy was in line with the intervention protocol, suggesting that BCA can be delivered as intended. Furthermore, the program appears to be acceptable to PWA and CPs of different ages, educational backgrounds, dyadic relationships and severity of aphasia. Of the nine dyads referred to the study, only one did not complete it; they reported that they would prefer an impairment based intervention, specifically targeting speech output.

Results underline the need for SLPs to tailor conversation skills training, and to consider the fact that video feedback may be crucial for generalization, although it is accepted that further evidence is needed before this can be concluded to be a vital part of such interventions. The results also show that intervention can assist communication in the absence of clear and clinically meaningful improvement on standard language tests. This is important, as a key aim of work on language in clinical practice and in some research is to improve everyday communication. The finding that everyday conversations change after video intervention work that targets conversation directly, provides clear confirmation that this is another approach that can be added to those already in use by SLPs. Finally, the study reveals the possibility of improving the everyday communication skills of some PWA many months beyond the period of spontaneous recovery. Seven of the eight dyads showed significant change at case series level and these included PWA at 59 and 60 months post stroke.

We now turn to the limitations of the study and implications for future research. From the perspective of the hierarchy of evidence, the fact that the study did not have a control group is a major concern. However, it is not clear what would be a suitable control, particularly given the heterogeneous nature of both aphasia and conversation. Furthermore, matching on conversation variables would be crucial for such a design and as shown in the data in **Tables 2**, **3**, no two dyads show similar profiles at the level of total conversation barriers and facilitators. This is also the case for individual behaviors such as use of test questions, and gesture. Given this variability, a design in which participants act as their own control is the most likely to inform understanding. Future research might employ a cross-over design comparing work on conversation with a very different intervention (e.g., training working memory, Salis et al., 2015). As cautioned in the "Introduction" Section, work on language would not be a suitable contrasting condition for research because of the potential for confusion over goals. While this limitation holds for research, in practice, when therapy is flexible, it may well be appropriate to work directly on a language skill (e.g., using cueing to aid word retrieval) and then on transferring this ability into everyday interactions (see for example Herbert et al., 2003).

While the study is one of the largest in the field, inclusion of more dyads would strengthen the research and would enable further exploration of candidacy issues (Turner and Whitworth, 2006; Eriksson et al., 2016) and of the links between patterns in pre-therapy conversations and outcome. Future studies, with more participants, could also explore in more depth the finding that it may be easier to reduce barrier behaviors than increase facilitator behaviors. This is particularly important given the clinical implications. A tentative suggestion would be that in selecting goals for conversation intervention, SLPs should facilitate PWA and CPs to select at least one barrier behavior each, rather than a set limited to facilitator behaviors. While the inclusion of additional dyads in research would enable exploration of such issues, there is a tension present. Due to inevitable resource limitations, including more participants is likely to entail a corresponding reduction in the depth and richness of data collected and analyzed for each dyad.

The study was carried out by a University based team and, while assessment and intervention took place in participants' homes, replication of the research in the clinical setting is important, particularly as issues of implementation in practice are likely to arise and will warrant exploration. Future research could also explore issues of dosage (bearing in mind adequate time is likely to be necessary between sessions for homework and consolidation) and intensity.

The study's findings are applicable to people with agrammatic aphasia and their conversation partners. Use of BCA principles and particularly use of video feedback in conversation should be explored in other populations including those with fluent aphasias, and primary progressive aphasia (see Volkmer, 2013) as a starting point. The related approach of Parent Child Interaction (PCI) has been shown to be effective at changing the conversations between pre-school children with language needs and their parents (Falkus et al., 2016). Future research

with school age children and teenagers with language disorders combining use of video feedback with principles of BCA and PCI is also warranted, particularly as video technology is now a routine part of many young peoples' everyday lives.

## CONCLUSION

In conclusion, this article adds to the evidence base for use of conversation approaches with PWA and their CPs. The dramatic decrease in barrier behaviors after eight sessions of intervention suggests change can be expected within a clinically realistic timescale. The experimentally controlled case series (with careful correction of multiple comparisons) demonstrates that findings for the group (for example the lack of a significant increase in facilitators) are not necessarily the same as those for individual dyads. This highlights the importance of reporting findings beyond the group level. Further exploration of use of video technology in conversation therapy should be a research priority.

## AUTHOR CONTRIBUTIONS

SB was principal investigator (PI) on the study, which was conceived and designed by SB, JM and WB and based in part on SB's doctoral research. This core team with FB created the BCA intervention and with SIE determined criteria for inclusion and assessments to be used. WB led on the quantification of aspects of conversation, the writing of this article and carried out the analysis of conversation facilitators and barriers for the group and case series. JM led the collection and compilation of connected speech data and the IRR analysis. CH led on fidelity and under the supervision of SB and WB assembled all the data from the student projects and compiled Appendices 2 and 3. CH and JM analyzed the data from the formal assessment tasks, and CH contributed substantially to the IRR analysis. FB recruited all participants to the study, assessed them and carried out the intervention. FJ analyzed

conversation data from two of the dyads, contributed to the BCA online resource and carried out the qualitative analysis discussed here in relation to this study's quantitative findings. SIE led on inclusion criteria and facilitated recruitment to the study. DH supplied the Poisson Trend Test and provided advice on statistical analysis. SB was PI for the follow-on study which funded development of the BCA online resource for rehabilitation.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fnhum. 2016.00562/full#supplementary-material

## REFERENCES

Bastiaanse, R., and Edwards, S. (2004). Word order and finiteness in Dutch and English Broca's and Wernicke's aphasia. *Brain Lang.* 89, 91–107. doi: 10. 1016/s0093-934x(03)00306-7

Bastiaanse, R., Edwards, S., and Rispens, J. (2002). *The Verb and Sentence Test.* Bury St. Edmunds, UK: Thames Valley Test Company.

Beckley, F., Best, W., and Beeke, S. (2016). Delivering communication strategy training for people with aphasia: what is current clinical practice? *Int. J. Lang. Commun. Disord.* doi: 10.1111/1460-6984.12265 [Epub ahead of print].

Beckley, F., Best, W., Johnson, F., Edwards, S., Maxim, J., and Beeke, S. (2013). Conversation therapy for agrammatism: exploring the therapeutic process of engagement and learning by a person with aphasia. *Int. J. Lang. Commun. Disord.* 48, 220–239. doi: 10.1111/j.1460-6984.2012. 00204.x

Beeke, S. (2012). "Aphasia: the pragmatics of everyday conversation," in *Cognitive Pragmatics [Handbook of Pragmatics, Vol. 4]*, ed. H.-J. Schmid (Berlin: Mouton de Gruyter), 345–371.

Beeke, S., Sirman, N., Beckley, F., Maxim, J., Edwards, S., Swinburn, K., et al. (2013). Better conversations with aphasia: an e-learning resource. Available online at: https://extend.ucl.ac.uk/

Beeke, S., Beckley, F., Johnson, F., Heilemann, C., Edwards, S., Maxim, J., et al. (2015). Conversation focused aphasia therapy: investigating the adoption of strategies by people with agrammatism. *Aphasiology* 29, 355–377. doi: 10. 1080/02687038.2014.881459

Beeke, S., Johnson, F., Beckley, F., Heilemann, C., Edwards, S., Maxim, J., et al. (2014). Enabling better conversations between a man with aphasia and his conversation partner: incorporating writing into turn-taking. *Res. Lang. Soc. Interact.* 47, 292–305. doi: 10.1080/08351813.2014.925667

Beeke, S., Maxim, J., Best, W., and Cooper, F. (2011). Redesigning therapy for agrammatism: initial findings from the ongoing evaluation of a conversation-based intervention study. *J. Neurolinguistics* 24, 222–236. doi: 10. 1016/j.jneuroling.2010.03.002

Beeke, S., Wilkinson, R., and Maxim, J. (2003). Exploring aphasic grammar 2: do language testing and conversation tell a similar story? *Clin. Linguist. Phon.* 17, 109–134. doi: 10.1080/0269920031000061786

Beeke, S., Wilkinson, R., and Maxim, J. (2007). Individual variation in agrammatism: a single case study of the influence of interaction. *Int. J. Lang. Commun. Disord.* 42, 629–647. doi: 10.1080/13682820601160087

Best, W., Grassly, J., Greenwood, A., Herbert, R., Hickin, J., and Howard, D. (2011). A controlled study of changes in conversation following aphasia therapy for anomia. *Disabil. Rehabil.* 33, 229–242. doi: 10.3109/09638288.2010. 534230

Bilda, K. (2011). Video-based conversational script training for aphasia: a therapy study. *Aphasiology* 25, 191–201. doi: 10.1080/0268703100379 8254

Caramazza, A., and Zurif, E. B. (1976). Dissociation of algorithmic and heuristic processes in language comprehension: evidence from aphasia. *Brain Lang.* 3, 572–582. doi: 10.1016/0093-934x(76)90048-1

Carragher, M., Sage, K., and Conroy, P. (2015). Outcomes of treatment targeting syntax production in people with Broca's-type aphasia: evidence from psycholinguistic assessment tasks and everyday conversation. *Int. J. Lang. Commun. Disord.* 50, 322–336. doi: 10.1111/1460-6984.12135

Conroy, P., Sage, K., and Ralph, M. L. (2009). Improved vocabulary production after naming therapy in aphasia: can gains in picture naming generalize to connected speech? *Int. J. Lang. Commun. Disord.* 44, 1036–1062. doi: 10.3109/13682820802585975

Des Roches, C. A., Balachandran, I., Ascenso, E. M., Tripodis, Y., and Kiran, S. (2015). Effectiveness of an impairment-based individualized rehabilitation program using an iPad-based software platform. *Front. Hum. Neurosci.* 8:1015. doi: 10.3389/fnhum.2014.01015

Druks, J., and Masterson, J. (2000). *An Object and Action Naming Battery.* Hove: Psychology Press.

Eriksson, K., Hartelius, L., and Saldert, C. (2016). On the diverse outcome of communication partner training of significant others of people with aphasia: an experimental study of six cases. *Int. J. Lang. Commun. Disord.* 51, 402–414. doi: 10.1111/1460-6984.12216

Falkus, G., Tilley, C., Thomas, C., Hockey, H., Kennedy, A., Arnold, T., et al. (2016). Assessing the effectiveness of parent–child interaction therapy with language delayed children: a clinical investigation. *Child Lang. Teach. Ther.* 32, 7–17. doi: 10.1177/0265659015574918

Fletcher, M., and Birt, D. (1983). *Storylines: Picture Sequences for Language Practice.* London: Longman.

Hartmann, D. P. (1977). Considerations in the choice of interobserver reliability estimates. *J. Appl. Behav. Anal.* 10, 103–116. doi: 10.1901/jaba.1977.10-103

Heeschen, C., and Schegloff, E. (2003). "Aphasic agrammatism as interactional artifact and achievement," in *Conversation and Brain Damage*, ed. C. Goodwin (New York, NY: Oxford University Press), 231–282.

Heilemann, C., Best, W., Johnson, F., Beckley, F., Edwards, S., Maxim, J., et al. (2014). Investigating treatment fidelity in a conversation-based aphasia therapy. *Aphasie und Verwandte Gebiete* 2, 14–26. Available online at: http://www.aphasie.org/de/3-fachpersonen/3.1-fachzeitschriften/fachzeitschrift-no2–2014/aphasie-u.-v.-gebiete-no-2_2014.-gesamt.pdf [article in English]

Herbert, R., Best, W., Hickin, J., Howard, D., and Osborne, F. (2003). Combining lexical and interactional approaches to the treatment of word-finding deficits in aphasia. *Aphasiology* 17, 1163–1186. doi: 10.1080/02687030344 000454

Herbert, R., Best, W., Hickin, J., Howard, D., and Osborne, F. (2013). *Profile of Word Errors and Retrieval in Speech (POWERS).* North Guilford: JR Press.

Herbert, R., Hickin, J., Howard, D., Osborne, F., and Best, W. (2008). Do picture naming tests provide a valid assessment of lexical retrieval in conversation in aphasia? *Aphasiology* 22, 184–203. doi: 10.1080/026870307012 62613

Hoffmann, T. C., Glasziou, P. P., Boutron, I., Milne, R., Perera, R., Moher, D., et al. (2014). Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 348:g1687. doi: 10.1136/bmj. g1687

Howard, D., Best, W., and Nickels, L. (2015). Optimising the design of intervention studies: critiques and ways forward. *Aphasiology* 29, 526–562. doi: 10.1080/02687038.2014.985884

Howard, D., and Patterson, K. E. (1992). *The Pyramids and Palm Trees Test: A Test of Semantic Access from Words and Pictures.* Bury St. Edmunds, UK: Thames Valley Test Company.

Johnson, F. M. (2015). What works in conversation therapy for aphasia and how? Searching mechanisms of change and active ingredients using tools and theory from behaviour change research. [Doctoral thesis]. London: UK: University College London. Available online at: http://discovery.ucl.ac.uk/1472090/

Johnson, F., Best, W., Beckley, F., Maxim, J., and Beeke, S. (in press). Identifying mechanisms of change in a conversation therapy for aphasia, using behaviour change theory and qualitative methods. *International Journal of Language and Communication Disorders*

Kagan, A., Winckel, J., Black, S., Duchan, J. F., Simmons-Mackie, N., and Square, P. (2004). A set of observational measures for rating support and participation in conversation between adults with aphasia and their conversation partners. *Top. Stroke Rehabil.* 11, 67–83. doi: 10.1310/cl3v-a94a-de5c-cvbe

Kay, J., Lesser, R., and Coltheart, M. (1992). *Psycholinguistic Assessments of Language Processing in Aphasia (PALPA).* Hove: Lawrence Erlbaum Associates.

Lock, S., Wilkinson, R., and Bryan, K. (2001). *Supporting Partners of People with Aphasia in Relationships and Conversation (SPPARC): A Resource Pack.* Bicester: Speechmark.

Marshall, J., Roper, A., Galliers, J., Wilson, S., Cocks, N., Muscroft, S., et al. (2013). Computer delivery of gesture therapy for people with severe aphasia. *Aphasiology* 27, 1128–1146. doi: 10.1080/02687038.2013.786803

Murray, L., Salis, C., Martin, N., and Dralle, J. (2016). The use of standardised short-term and working memory tests in aphasia research: a systematic review. *Neuropsychol. Rehabil.* doi: 10.1080/09602011.2016.1174718 [Epub ahead of print].

Nicholas, L. E., and Brookshire, R. H. (1993). A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. *J. Speech Hear. Res.* 36, 338–350. doi: 10.1044/jshr.3602.338

Nouwens, F., Dippel, D. W. J., de Jong-Hagelstein, M., Visch-Brink, E. G., Koudstaal, P. J., and de Lau, L. M. L. (2013). Rotterdam Aphasia Therapy Study (RATS)–3: "The efficacy of intensive cognitive-linguistic therapy in the acute stage of aphasia"; design of a randomised controlled trial. *Trials* 14:24. doi: 10.1186/isrctn67723958

Oelschlaeger, M. L., and Thorne, J. C. (1999). Application of the correct information unit analysis to the naturally occurring conversation of a person with aphasia. *J. Speech Lang. Hear. Res.* 42, 636–648. doi: 10.1044/jslhr.4203.636

Palmer, R., Cooper, C., Enderby, P., Brady, M., Juliou, S., Bowen, A., et al. (2015). Clinical and cost effectiveness of computer treatment for aphasia post stroke (Big CACTUS): study protocol for a randomised controlled trial. *Trials* 16:18. doi: 10.1186/s13063-014-0527-7

Perkins, L., Crisp, J., and Walshaw, D. (1999). Exploring conversation analysis as an assessment tool for aphasia: the issue of reliability. *Aphasiology* 13, 259–281. doi: 10.1080/026870399402091

Ramsberger, G., and Menn, L. (2003). "Co-constructing Lucy: adding a social perspective to the assessment of communicative success in aphasia," in *Conversation and Brain Damage*, ed. C. Goodwin (Oxford: Oxford University Press), 283–303.

Ramsberger, G., and Rende, B. (2002). Measuring transactional success in the conversation of people with aphasia. *Aphasiology* 16, 337–353. doi: 10.1080/02687040143000636

Rose, T., Worrall, L., Hickson, L., and Hoffmann, T. (2010). Do people with aphasia want written stroke and aphasia information? A verbal survey exploring preferences for when and how to provide stroke and aphasia information. *Top. Stroke Rehabil.* 17, 79–98. doi: 10.1310/tsr1702-79

Salis, C., Kelly, H., and Code, C. (2015). Assessment and treatment of short-term and working memory impairments in stroke aphasia: a practical tutorial. *Int. J. Lang. Commun. Disord.* 50, 721–736. doi: 10.1111/1460-6984. 12172

Schmidt, J., Fleming, J., Ownsworth, T., and Lannin, N. A. (2012). Video feedback on functional task performance improves self-awareness after traumatic brain injury: randomized controlled trial. *Neurorehabil. Neural Repair* 27, 316–324. doi: 10.1177/1545968312469838

Schmidt, J., Lannin, N., Fleming, J., and Ownsworth, T. (2011). Feedback interventions for impaired self-awareness following a brain injury: a systematic review. *J. Rehabil. Med.* 43, 673–680. doi: 10.2340/16501977-0846

Simmons-Mackie, N., Raymer, A., and Cherney, L. R. (2016). Communication partner training in aphasia: an updated systematic review. *Arch. Phys. Med. Rehabil.* doi: 10.1016/j.apmr.2016.03.023 [Epub ahead of print].

Simmons-Mackie, N., Savage, M. C., and Worrall, L. (2014). Conversation therapy for aphasia: a qualitative review of the literature. *Int. J. Lang. Commun. Disord.* 49, 511–526. doi: 10.1111/1460-6984.12097

Swinburn, K., Porter, G., and Howard, D. (2004). *Comprehensive Aphasia Test (CAT).* Hove, UK: Psychology Press.

Tate, R. L., Perdices, M., Rosenkoetter, U., Wakim, D., Godbee, K., Togher, L., et al. (2013). Revision of a method quality rating scale for single-case experimental designs and *n*-of-1 trials: the 15-item Risk of Bias in *N*-of-1 Trials (RoBiNT)

Scale. *Neuropsychol. Rehabil.* 23, 619–638. doi: 10.1080/09602011.2013.824383

Thompson, C. K. (2006). Single subject controlled experiments in aphasia: the science and the state of the science. *J. Commun. Disord.* 39, 266–291. doi: 10.1016/j.jcomdis.2006.02.003

Thompson, C. K., and Shapiro, L. P. (2007). Complexity in treatment of syntactic deficits. *Am. J. Speech Lang. Pathol.* 16, 30–42. doi: 10.1044/1058-0360(2007/005)

Turner, S., and Whitworth, A. (2006). Clinicians' perceptions of candidacy for conversation partner training in aphasia: how do we select candidates for therapy and do we get it right? *Aphasiology* 20, 616–643. doi: 10.1080/02687030600687860

van de Sandt-Koenderman, W. M. E. (2011). Aphasia rehabilitation and the role of computer technology: can we keep up with modern times? *Int. J. Speech Lang. Pathol.* 13, 21–27. doi: 10.3109/17549507.2010.502973

Volkmer, A. (2013). *Assessment and Therapy for Language and Cognitive Communication Difficulties in Dementia and Other Progressive Diseases.* North Guilford: JR Press.

Wielaert, S. M. (2016). 'Conversation is a two-way process': introducing conversation partner training in aphasia in Dutch rehabilitation practice. [Doctoral thesis] University of the West of England. Available online at: http://eprints.uwe.ac.uk/26904

Wielaert, S. M., van de Sandt-Koenderman, W. M. E., Dammers, N., and Sage, K. E. (2014). Making an ImPACT. Implementing conversation partner training in Dutch rehabilitation. *Arch. Phys. Med. Rehabil.* 95, e26–e27. doi: 10.1016/j.apmr.2014.07.065

Wilkinson, R. (2014). Intervening with conversation analysis in speech and language therapy: improving aphasic conversation. *Res. Lang. Soc. Interact.* 47, 219–238. doi: 10.1080/08351813.2014.925659

Wilkinson, R., Bryan, K., Lock, S., and Sage, K. (2010). Implementing and evaluating aphasia therapy targeted at couples' conversations: a single case study. *Aphasiology* 6, 869–886. doi: 10.1080/02687030903501958

Wilkinson, R., Lock, S., Bryan, K., and Sage, K. (2011). Interaction-focused intervention for acquired language disorders: facilitating mutual adaptation in couples where one partner has aphasia. *Int. J. Speech Lang. Pathol.* 13, 74–87. doi: 10.3109/17549507.2011.551140

Wilkinson, R., and Wielaert, S. M. (2012). Rehabilitation for aphasic conversation: can we change the everyday talk of people with aphasia and their significant others? *Arch. Phys. Med. Rehabil.* 93, S70–S76. doi: 10.1016/j.apmr.2011.07.206

Wilson, S., Roper, A., Marshall, J., Galliers, J., Devane, N., Booth, T., et al. (2015). Codesign for people with aphasia through tangible design languages. *CoDesign* 11, 21–34. doi: 10.1080/15710882.2014.997744

Woolf, C., Caute, A., Haigh, Z., Galliers, J., Wilson, S., Kessie, A., et al. (2016). A comparison of remote therapy, face to face therapy and an attention control intervention for people with aphasia: a quasi-randomised controlled feasibility study. *Clin. Rehabil.* 30, 359–373. doi: 10.1177/0269215515582074