Enhancing Communication through Gesture and Naming Therapy

Anna Caute[1]

Tim Pring[1]

Naomi Cocks[1]

Madeline Cruice[1]

Wendy Best[2]

Jane Marshall[1]

[1]The Department of Language and Communication Science, City University, London

[2]Division of Psychology and Language Science, University College London

Corresponding author: Jane Marshall, the Department of Language and Communication Science, City University, London EC1V OHB, UK. J.Marshall@city.ac.uk

Abstract

Purpose: This study investigated whether gesture, naming and strategic treatment improved the communication skills of 14 people with severe aphasia.

Method: All participants received 15 hours of gesture and naming treatment (reported in a companion paper). Half the group received a further 15 hours of strategic therapy, while the remaining seven participants received no further input. The effects of therapy on communication were assessed with two novel measures. These required participants to convey simple messages and narratives to their communication partner. In both assessments a subset of the stimuli featured items that had been targets in gesture or naming treatment.

Results: Performance on the communication measures was stable over two baseline assessments, but improved after gesture and naming treatment. Those who received additional strategic therapy made further gains on the message but not the narrative task. Communication gains were not specific to the stimuli featuring trained items.

Conclusions: This study suggests that gesture and naming treatments can benefit interactive communication. The additional benefits of strategic therapy were less clear cut, but did impact on the transmission of simple messages. Gains seem to reflect the development of general communication skills, rather than the use of trained gestures and/or words.

Introduction

This study explored the impact of aphasia therapy on interactive communication. Two novel evaluation tasks were created, which were sensitive to both verbal and non verbal communication and which could be used with people who have severe aphasia. The effects of different forms of therapy were also explored, with one programme training a vocabulary of gestures and words, and another the application of communication strategies.

Few would quarrel with the assertion that aphasia therapy should change everyday communication, rather than just performance on clinical language tests. Yet demonstrating such a change is difficult. Interview schedules and rating scales may be used, such as the La Trobe Communication Questionnaire (Douglas, O'Flaherty & Snow, 2000), and the Functional Assessment of Communication Skills for Adults (Fratali, Thompson, Holland, Wohl & Ferketic, 1995). However, these typically ask about a range of communication behaviours, only some of which may have been addressed in therapy, and they may not be accessible to people with severe aphasia (Boo & Rose, 2011). Another approach involves sampling and analysing naturally occurring conversations (e.g., Booth & Perkins, 1999; Best, Grassly, Greenwood, Herbert, Hickin & Howard, 2011). However, here variability may mask therapy effects (e.g., see Cunningham & Ward, 2003). Sampling a standard narrative, such as the Cinderella story, reduces variability and allows comparison with control data (Rochon, Saffran, Berndt & Schwartz, 2000). However, this task is not interactive, so may be less informative about everyday communication. As a result, some prefer to analyse the number of propositions conveyed to another person during story retelling (e.g., Hopper, Holland & Rewega, 2002)

Exploring the communicative effects of a gestural therapy introduces new complexities. Here measures must be sensitive not just to changes in speech, but also to changes in the use of gesture. The number of gestures featuring in a person's communication may be counted before and after therapy. However, this requires a stable elicitation method, and may not reveal

whether gestures are effective. The qualitative analysis of gestures used in interactions may capture the latter (e.g., see Goodwin, 2000). However, this does not yield numerical data, so does not lend itself to the experimental investigation of therapy outcomes.

Most previous studies of compensatory gesture therapy employed object or picture based elicitations of gestures as the primary evaluation measure (e.g., Code & Gaunt, 1986; Coelho & Duffy, 1987, 1990; Conlon & McNeil, 1991; Cubelli, Trentini, Montagna, 1991; Daumüller & Goldenberg, 2010). As a result, there is no information about the impact of therapy on communication, or even whether participants used the taught gestures during interactions. Some studies report positive (Cubelli et al., 1991; Rao, 1995) or negative (Coelho & Duffy, 1990) comments from family members about the degree to which taught skills were employed in the home. However, these observations were not systematically elicited or analysed.

Studies that did attempt to assess communication have produced mixed results. Coelho and Duffy (1985) involved family members in message and conversation tasks with their participant, and asked his spouse to keep a daily record of communication attempts. They found that taught gestures could be used in constrained tasks, such as picture description, but not in more naturalistic conditions such as conversation. In Coelho (1991) two participants with moderate and severe aphasia were taught 12 gestures relating to food. These gestures were evaluated not only in the clinic but also in simulated and real restaurant settings, with the finding that transfer was best achieved by the least impaired participant. However, strong conclusions were limited by the small number of gestures trained, and by the fact that no experimental controls were employed. Purdy, Duffy and Coelho (1994) administered multimodality training to 15 participants, involving gesture, a communication board and verbal naming. After therapy, a mean of 49% of trained symbols were used in a structured conversation task, although, interestingly, taught gestures were used significantly less often than taught spoken words (10%

vs 34%). It is also difficult to be sure of the impact of therapy, since baseline data were not collected.

Studies evaluating the effects of gesture therapy on speech have employed a wider range measures. Some purely explored the impact of therapy on picture naming (Pashek, 1998; Raymer &Thompson, 1991), or picture naming and gesturing (Rodriguez, Raymer & Rothi, 2006). However, others looked for changes in conversation (Rose, Douglas & Matyas, 2002; Rose & Sussmilch, 2008; Boo & Rose, 2011) and narrative (Boo & Rose, 2011). One study tailored the sampling procedure to therapy targets, by eliciting narratives related to the taught categories (Rose & Douglas, 2008) and two administered the La Trobe Communication Questionnaire (Rose & Sussmilch, 2008; Boo & Rose, 2011). These studies did not probe the communicative use of gesture, given that their focus was on the facilitation of speech. However, some authors observe that even when naming gains were not achieved there were secondary benefits in gesture production, and while not formally assessed these seemed to have functional consequences for communication (Rodriguez et al., 2006).

To date, the communicative impact of gesture therapy has been minimally explored, particularly for people with severe aphasia who use gesture as a compensatory strategy. Furthermore, existing evaluation tools are poorly adapted to this purpose. It seems that novel measures are required, if the communicative benefits of gesture therapy are to be examined.

In the current study two such measures were developed. In both, the person with aphasia was required to convey information to their communication partner. This information was unknown to the partner, ensuring that a genuine exchange took place, and scoring evaluated the amount of information acquired by the partner. One measure involved simple, two component messages. The other involved video based narratives. Both measures included stimuli that did and did not feature in therapy, enabling us to explore the degree to which communication hinged on trained items.

A further question is what type of therapy best achieves communicative gains? Many previous studies simply trained a set of gestures, presumably in the hope that these would be used when required in everyday communication (e.g., Cubelli et al., 1991). However, as argued by Daumüller and Goldenberg (2010), such a transfer cannot be assumed. Some previous studies included methods that specifically aimed to promote the functional use of gesture, such as communication assignments with other health care professionals (Rao, 1995) or role play simulations (Coelho, 1991). However, the relative contribution of these therapy components was either not evaluated (Rao, 1995), or unclear (Coelho 1991).

Communicative gains are specifically targeted in strategic therapies, such as Conversational Coaching (Holland, 1991; Hopper and Holland, 2002). Such therapies involve the teaching of communication strategies to both the person with aphasia and their communication partner. An initial phase identifies the target strategies, which are then practised in interactive tasks with feedback from the therapist. Such therapies often engage non verbal skills. For example, the person with aphasia might be encouraged to use gesture to convey information, and the partner may be given skills in eliciting this modality.

In the current study all participants received 15 hours of therapy aiming to train a vocabulary of 20 gestures and 20 different words (Therapy A). Half the group additionally received Therapy B. Therapy B did not train specific items. Rather it promoted strategic communication skills using interactive tasks between the person with aphasia and their partner. Of interest was whether this strategic therapy would bring about additional communicative gains, over and above those achieved by training gestures and words.

A companion paper showed that Therapy A resulted in gesture and naming gains, although these were more marked for naming than gesture, and did not generalise beyond treated items (Marshall, et al., in press). Evaluations in that paper were confined to picture based

elicitations of target gestures and names. Here we examine the impact of both Therapy A and B on the novel message and narrative tasks.

The study questions addressed in the current paper can be summarised as follows:

1. Does gesture and naming treatment (Therapy A) improve performance on interactive communication tasks (conveying messages and narratives to a partner)?

2. Are improvements general or confined to messages and narratives that involve treated items?

3. Does strategic therapy (Therapy B) further improve performance on the interactive tasks?

Method

The study received ethical approval from a National Health Service Local Research Ethics Committee, five local NHS Research and Development departments, and the Research Ethics Committee of City University London. All participants provided informed consent before being involved in any of the study procedures.

Participants

Fourteen people with severe aphasia took part. All had severe aphasia, scoring below 20% on spoken and written naming (stimuli taken from the Comprehensive Aphasia Test, CAT, Swinburn, Porter & Howard, 2004). They were at least six months post-stroke and were fluent pre-morbid users of English (established via self report). They had no additional neurological diagnosis, such as dementia, and scored at least 6/10 on an informal object to picture matching screening assessment. All showed evidence of limb and/or oral apraxia on the Apraxia Battery for Adults (Dabul, 2000). Participants received no other speech and language therapy during their involvement in the study.

Table 1 summarises participant details for the whole group and for the seven who received Therapy B. Recruitment to Therapy B was determined by participant choice and

practical considerations, such as timing and availability. T test comparisons on baseline factors (naming, comprehension, semantic memory and apraxia scores) showed no significant differences between the participants who did and did not receive Therapy B. Each participant nominated a communication partner, e.g. a family member or friend. Partners had at least weekly contact with the participant and had no neurological diagnosis (self report). They were involved in our communication assessments and, where relevant, took part in all Therapy B sessions.

Insert Table 1 about here

The Design

The study employed a repeated measures design. Two baseline assessments (time 1 and 2) were conducted, separated by four weeks. All participants then received Therapy A. This was followed by a post therapy (time 3) and follow up (time 4) assessment. Seven participants received Therapy B between time 3 and 4 and seven received no further input. Therapy regimes were typically eight weeks for Therapy A, and 15 weeks for Therapy B. Variations were due to individual circumstances such as ill health. Therapy B commenced immediately after the post therapy assessment at time 3.

Therapy A

This involved 15 hours of domiciliary therapy delivered twice a week. It aimed to train 20 gestures and 20 different words. Time was equally divided between gesture and naming tasks. All tasks and cues followed standard formats, which were specified in a manual. Treatment progressed from recognition tasks (matching gestures or words to pictures) to production tasks, in which a gesture or word had to be repeated from a model or produced in response to a picture (see Marshall et al., in press for further details).

Therapy B

Seven participants additionally received Therapy B between time 3 and 4. This aimed to develop participants' interactive use of gesture, and their ability to integrate gesture with other communication strategies. It also attempted to improve partners' skills in eliciting and interpreting gesture. Therapy comprised 15 one hour, domiciliary sessions conducted with the participant and their partner. Owing to partner availability sessions were conducted once a week. Therapy B was delivered by a single treating therapist (the first author) following procedures specified in a manual.

Therapy B began with a standard goal setting procedure, involving an interview with the participant and their partner about their everyday communication. Topics included: current and pre-stroke communication styles, changes in communication since Therapy A, current uses of gesture, integration of gesture with speech, topics of conversation, examples of communication difficulty, and strategies used to resolve breakdowns. The goal setting discussion identified communication targets for Therapy B. These were not tied to items that were treated in Therapy A.

Much of the goal setting discussion focussed on strategy use. In part, this was stimulated by a strategy checklist completed by partners. Dyads were also asked to make a video of themselves in conversation and to then note the strategies that they either did or could use. From this, target communication strategies for the participant and their partner were identified. So, for example, strategies for the participant might include: gesturing the key item, using an agreed signal for 'yes' and 'no', and using a hand gesture to indicate 'you're close' in response to partner suggestions. Partner strategies might include: asking yes/no questions, prompting gesture e.g. by saying: 'can you show me with your hand', and copying back gestures to confirm that the message has been understood.

Therapy aimed to develop the target strategies, first in constrained and then more open and complex tasks. The tasks were built around each participant's chosen topics, but were selected

from a standard 'menu': 1) Conveying an object photo to the partner; 2) Conveying an action photo to the partner; 3) Conveying a simple message to the partner; 4) Conveying a "problem scenario" to the partner; 5) Describing personal photos to the partner or therapist; 6) Carrying out a scripted conversation with the partner; 7) Telling stories about the previous week to the therapist, with partner support; 8) Recounting news stories or episodes from TV programmes to the partner or therapist; 9) Videoing a conversation during the week and discussing strategies used; 10) Discussing communication problems and successes over the previous week (See further details in Appendix 1).

The strategies themselves were first modelled by the therapist and their value discussed. Once established, the therapist reminded dyads about their target strategies before each task. Where necessary, she also rehearsed skills with the participant, e.g. by priming a target gesture. Finally the therapist aimed to elicit the strategies from the participant and partner before attempting a task.

The structure for providing feedback was similar to that used in Conversational Coaching (Hopper et al., 2002). If a communication breakdown or miscommunication occurred during the task the therapist intervened after two failed attempts. She suggested resolution strategies either for the person with aphasia, their partner, or both, although such interventions were reduced as therapy progressed. After tasks, feedback focussed on the strategies that were or could have been used. In many instances tasks were repeated, to give dyads the opportunity to practise the suggested strategies again.

Assessments

Four assessments were conducted at each time point. Two (the gesture and naming assessments) were reported in our companion paper (Marshall et al., in press). These tested 60 items. Thirty of these were standard to all participants and 30 were personally selected, e.g.

through discussion with the participant and their partner. All standard and personal items could be gestured and pictured, and were mainly drawn from the categories of food, drink, clothes, transport, furniture, personal objects, people and activities. The 60 items were divided into three sets, so that during Therapy A, 20 received gesture treatment, 20 received naming treatment and 20 were untreated. The sets were matched for gesturability, operativity, object familiarity, imageability and word frequency. None was named at a higher than 15% success rate at baseline (gesture scores were similarly low).

The remaining assessments were a message and narrative assessment, and aimed to explore communication gains from therapy.

The Message Assessment

In this assessment participants were asked to convey simple messages to their partner. Messages were constructed from the 60 assessment items described above, i.e. 30 involved standard items and 30 involved personal items. Messages comprised questions, statements or requests. They were written on a card together with the object stimulus picture. So, for the assessment item 'pill', the messages were: 'I take two pills' and 'I need a pill for a headache', and in both cases the written message was accompanied by a picture of a bottle of pills.

Message stimuli were piloted with ten healthy controls (aged > 40), to ensure that they could be conveyed through gesture. Only messages for the standard items were piloted (as the personal item messages differed for each participant). Controls were shown each message card and asked to convey the content to the researcher using gesture. They were further constrained only to use their left hand. Their production was videoed and scored by the first author. Two points were awarded if the total message was conveyed and one if only part was conveyed.

Control scores indicated that the message sets derived from each treatment group of items were well matched. Low standard deviations also signalled consistency of performance.

The mean (S.D) scores were: 1.65 (.28) for the gesture treatment set, 1.62 (.25) for the naming treatment set and 1.65 (.20) for the untreated set.

The assessment procedure for the participants with aphasia was as follows. During the course of the study each participant conveyed 120 messages to their partner, i.e. two derived from each assessment item. At each time point 30 messages were conveyed, with 10 messages from each treatment set (i.e. 10 messages constructed from gesture treatment items, 10 messages from naming treatment items and 10 messages from untreated items). This ensured that each message was conveyed only once, and that partners were always receiving novel information. The message assessment was conducted on a different day from the gesture and naming assessments.

Participants were shown each message card and the content was read aloud. Partners, who were not present, were then called into the room. The person with aphasia attempted to convey the message to their partner using gesture and/or speech. Partners were told that they could ask questions to seek clarification. After 30 seconds partners were required to write down their understanding of the message.

Partner scripts were scored independently by two raters who were blind to the time of assessment. Each message was rated on a 0 – 4 scale, according to the amount of information that had been understood from the original message (see Appendix 2 for scorers instructions). Disagreements between scorers were recorded and resolved by discussion, so that just one agreed score was derived for each message at each time point. A Cohen Kappa analysis yielded a value of K = .796, which is above the .75 'excellent' threshold defined by Fleiss (1981)

The Narrative Assessment

In this assessment participants watched a brief silent video and attempted to convey the contents to their partner.

Six assessment videos were developed for the project, each of which showed an everyday sequence of 10 events. For example one video showed the following sequence: a woman gets out of *bed*, turns on the *shower*, combs her *hair*, puts on her *watch*, polishes her *glasses*, drinks a *cup* of tea, takes a *pill*, eats a *banana*, turns on the TV with the *remote control*, and reads a *letter*. Videos were set in naturalistic settings. For example, the above video was filmed in a house and another, showing a sequence of events in which a woman buys her lunch, was filmed in an office and a shop.

One of the videos ('treated repeated') was constructed from 10 standard items that featured in therapy, five in naming treatment and five in gesture treatment. This is described above with italicised items. Another ('untreated repeated') was constructed from 10 standard items from the untreated set. The remaining four videos ('untreated unrepeated') were not tied to our assessment groups, and care was taken to ensure that they did not feature treated items.

All six assessment videos were piloted with 10 healthy controls, to ensure that the depicted events could be gestured. Controls watched each video and attempted to gesture the content using their left hand. Two constraints were imposed to ensure that they used gesture: speech was forbidden, and no verbal feedback was provided by the researcher. Their production was scored by the first author for the number of events conveyed. Table 2 shows that all videos elicited high mean scores, although one control participant produced low scores with two of the unrepeated videos.

Insert Table 2 about here

The assessment procedure for the participants with aphasia was as follows. Three videos were assessed at each time point. Two were repeated each time (the 'treated repeated' and 'untreated repeated' videos). The third was a novel video (an 'untreated unrepeated' video).

The order of administration of these novel videos was randomised across time and across participants.

The participant watched each video first in its entirety and then in three segments (with each segment containing three or four events). After each segment they were joined by their partner and attempted to convey what they had observed. Partners, who had not seen the video, were told that they could ask questions and seek clarification. After each exchange, partners were asked to write down what they had understood so far. Thus by the end of the assessment the partners had produced a script describing their understanding of the total video narrative.

Partner video scripts were scored by two raters who were blind to the time of assessment. They awarded a score of 0 – 4 for each event in the video, depending on the degree to which it was described in the script (see Appendix 3 for scorers' instructions). Disagreements between scorers were recorded and resolved through discussion. Thus one agreed score (/40) was derived for each video at each time point. A Cohen Kappa analysis indicated that the agreement between scorers was excellent (K = .86).

Results

Findings from two data sets are presented: scores on the message assessment and scores on the narrative assessment. Two ANOVAs were conducted on each set, the first examining scores across all participants, and the second comparing participants who did and did not receive Therapy B between time points 3 and 4. Significance levels were adjusted to .025 to take account of the dual analysis.

Message Assessment

Insert Table 3 about here

Data from the message assessment (see Table 3) were first analysed for all 14 participants (see mean scores in table 3). A three factor within subject ANOVA was conducted. Factors were time (four levels: time 1, 2, 3 & 4); message set (three levels: gesture treatment, naming treatment and no treatment) and item (two levels: personal and standard). In this, and in all subsequent analyses, the Greenhouse-Geisser correction was applied when results failed to meet the sphericity assumption.

The analysis revealed a significant main effect of time ($F (3, 39) = 11.53$, $p < .001$, $\eta^2_p = .47$). Planned comparisons showed no difference between time points 1 and 2 ($p = .913$). There was a highly significant difference between the combined baseline and post therapy scores, or 1 & 2 vs 3 & 4 ($p < .0001$). There was a marginally significant difference between time 3 and 4, although only at $p<.05$.

The messages derived from the gesture treatment items achieved slightly higher scores than the other sets, but the main effect of message set failed to reach significance ($F (1.43, 18.55) = 3.86$, $p = .052$). Message set also did not interact with time ($p = .48$), showing that gains on this task were not specific to the messages based on treated items.

There was no main effect of item ($p = .79$), showing that messages derived from personal items were no more challenging than those derived from standard items. None of the other interactions was significant.

This analysis showed that performance on the message task did not change significantly during the baseline period, but improved after Therapy A. The gains were not related to items that received treatment, but were general across the sets. There was a hint that performance continued to improve at time point 4, when half the participants had received Therapy B.

A second analysis compared participants who did and did not receive Therapy B between time points 3 and 4 (see mean scores in table 4). This was a four factor mixed

ANOVA. The within subject factors were time (two levels: time 3 and time 4), message set (three levels: gesture treatment, naming treatment and no treatment) and item (two levels: standard and personal). The between subject factor was participant group (two levels: those who did and did not receive Therapy B).

Insert Table 4 and Figure 1 about here

There was a main effect of time ($F(1, 12) = 9.48$, $p = .01$, $\eta^2_p = .44$), showing that scores increased at time 4. The main effect of group was not significant ($p = .23$), but there was a significant interaction between time and group ($F(1, 12) = 8.9$, $p = .011$, $\eta^2_p = .43$). Figure 2 illustrates this interaction and shows that group 1, who received Therapy B, improved between time points 3 and 4, while group 2 did not. No other main effects or interactions were significant. Thus the gains made by group 1 occurred equally across all message sets.

As the group analysis masked individual differences, individual gain scores were calculated. The gain score for Therapy A was the difference between participants' scores on the message assessment at time 3 and the average of their performance at time 1 and 2. The gain score for Therapy B was the difference between their scores at time 3 and 4. Total message scores were used, i.e. including treated and untreated items. The scores are reported in Figure 2. Participants who received Therapy B are on the right of the figure.

Insert Figure 2

In line with the ANOVA results, this shows that all bar two participants made gains following Therapy A (between pre-therapy assessments and time point 3). Yet the margin of improvement was often modest, with only 5 participants displaying gains above ten percent. Participants who received Therapy B all made further gains between time points 3 and 4, with five improving by more than ten percent. The performance of the remaining participants was more variable, with four showing a decline during this period.

Narrative Assessment

Scores from the narrative assessment for all 14 participants are presented in table 5. These were analysed using a two factor within subject ANOVA. Factors were time (four levels: time 1, 2, 3 & 4) and video (three levels: treated repeated, untreated repeated and untreated unrepeated).

Insert Table 5 about here

The analysis revealed a main effect of time ($F(1.79, 23.31) = 7.19$, $p = .005$, $\eta^2_p = .37$). Planned comparisons showed no difference between time points 1 and 2 ($p = .9$) or between time points 3 and 4 ($p = .36$). There was, however, a highly significant difference between the combined baseline and post therapy scores, or 1 & 2 vs 3 & 4 ($p = .0001$).

There was also a main effect of video ($F(2, 26) = 6.4$, $p = .005$, $\eta^2_p = .33$), with planned comparisons showing that the untreated unrepeated videos scored lower than the treated repeated (although only at $p < .05$) and untreated repeated ($p < .005$) videos. The repeated videos did not differ. The interaction between time and video was not significant.

This analysis showed that performance on the video task significantly increased after Therapy A, with no change occurring during the baseline. Gains were not specific to the video that featured treated items, given that there was no interaction between time and video. The unrepeated videos proved more challenging than those that were repeated each time. This effect was evident even at the first assessment, and did not change over time.

The lack of change at time point 4 suggests that Therapy B did not influence performance on this task. However, the whole group analysis might mask such an effect. A second analysis therefore compared participants who did and did not receive Therapy B between time points 3 and 4 (see table 6). This was a mixed three factor ANOVA. Within subject factors were time (two levels: time 3 & 4) and video (three levels: treated repeated, untreated

repeated and untreated unrepeated). The between subject factor was participant group (two levels: those who did and did not receive Therapy B).

Insert Table 6 about here

None of the main effects in this analysis was significant. Critically, there was also no interaction between time and group ($p = .53$). There was a marginally significant interaction between time and video ($F (2, 24) = 3.813$, $p = .036$, $\eta^2_p = .24$) and between time, video and group $(2, 24) = 4.029$, $p = .03$, $\eta^2_p = .25$), although these failed to meet the adjusted significance level of .025. This analysis failed to show that Therapy B improved performance on the narrative assessment.

Gain scores on the narrative assessment were calculated for each participant in order to investigate individual differences. Total scores were used, i.e. the combined score from the three videos administered at each time point. The gain score for Therapy A was the difference between the score at time point 3 and the mean scores at time point 1 and 2. The gain score for Therapy B was the difference between time point 3 and 4. The gain scores are reported in Figure 3. Participants on the right of the figure received Therapy B.

Insert Figure 3 about here

The figure shows that all bar one participant made gains on the video assessment following Therapy A. Performance between time points 3 and 4 was more variable, and did not clearly distinguish between those who did and did not receive Therapy B. In both groups, five of the seven participants made further gains. As in the message assessment, gains were often modest. For example, only seven participants improved by over ten percent following Therapy A. The extent and relevance of individual gains are considered further in the concluding discussion.

Discussion

This study developed two novel assessments, and used them to determine whether two distinct programmes of therapy improved communication skills in 14 participants with severe aphasia. This discussion will initially appraise the assessments, with reference to the control data. It will then address each of the study questions, and consider the strength and implications of the findings.

The message assessment required participants to convey a simple request, question or statement to their partner. Stimuli were constructed from items that were included in gesture therapy, naming therapy or which were untreated. Testing with 10 healthy controls confirmed that the stimuli could be successfully conveyed with one handed gesture. Controls achieved an average score of about 1.6 on each item, indicating that they were communicating at least part, and often all, of the message (their production was scored on a 2 point scale). Control scores were also consistent (low SDs) and well matched across groups. Thus messages constructed from the gesture therapy, naming therapy and untreated items seemed to have comparable levels of difficulty.

The narrative assessment required participants to convey a sequence of 10 linked events rather than an isolated message, using six silent videos as stimuli. Here control performance was less reassuring. On two of the 'untreated unrepeated' videos scores ranged from 4 – 10 events conveyed. If a score of 4 is within the normal range, there may be a very narrow margin of improvement for participants with aphasia. However, the range was depressed by just one low scoring control. It should also be noted that there were differences between the test conditions for controls and participants. Unlike the participants with aphasia, the controls were forbidden to use speech on this task (this was essential to elicit gesture from them). Also, to avoid engaging them in an artificial interaction with their partner, they conveyed the videos to the researcher. She did not give feedback or request clarification, but simply made a judgement

about whether they had conveyed the item accurately. Given that test conditions were different for the therapy participants, the concern about the margin for improvement may not apply.

Our first study question asked whether gesture and naming treatment (Therapy A) would improve performance on the two novel tasks. The results suggest that it did, since scores increased following Therapy A, having been stable over the baseline. Furthermore, this pattern was observed in both the message and narrative assessments. Thus it seemed that training a vocabulary of gestures and words increased participants' ability to convey information to their partner. Growing familiarity with the stimuli can be ruled out as a source of change. None of the messages was repeated across assessments; and gains on the narrative task occurred on unrepeated as well as repeated videos.

Our second question asked whether gains would be specific to treated items. In both the message and narrative assessments, a subset of the stimuli involved items that had been treated for gesture or naming during Therapy A. So in the message task, participants were asked to convey 30 messages at each time point. Ten of these were derived from items that were in the gesture treatment set and 10 were derived from items in the naming treatment set. In the narrative task one video was constructed from items featuring in naming and gesture treatment (five from each set), and this video was repeated at each assessment point.

If gains on the message assessment were specific to treated items, there would be an interaction between message set and time. In fact, there was no such interaction. Rather the improvement was general across the sets. Similarly, in the narrative assessment, there was no evidence that the treated repeated video particularly improved.

This finding can be negatively or positively construed. Taking the negative view first, it might be seen to qualify the assertion of a therapy effect. Our companion paper (Marshall et al., in press) showed that results on the gesture and naming assessments improved following

Therapy A, but largely on treated items. It might be expected, therefore, that any communication gain would hinge on those items, which was not the case.

The more positive view argues that gains reflect the development of communication skills that were not item dependent. It should be stressed that participants were not trained on the message or narrative stimuli, and this was true even for the treated sets. Rather they were trained on nouns that featured in the stimuli. So, in the following message (derived from an item in the gesture treatment set) only pill was trained: 'I need a pill for a headache'. Similarly, the treated video included numerous objects and actions that were not included in treatment. Thus the relationship between treatment and the stimuli in our communication assessments was much less strong than for our gesture and naming assessments. Item specific effects, therefore, might be less expected.

What communication skills, therefore, had been gained? Informal observation suggested that participants became more willing to attempt a gesture or word. Even when not successful, such attempts often conveyed elements of a concept. On some occasions they also stimulated further input from the partner, which, collaboratively, lead to the target. Further qualitative analyses of the interactions will be conducted in future, which may illuminate these observations.

Our third question asked whether strategic therapy (Therapy B) would further improve performance on the interactive tasks. Here the message and narrative assessments need to be considered separately. Results from the message assessment suggested that there were indeed increased benefits from Therapy B. Participants who received Therapy B between time points 3 and 4 improved over this period; while participants who received no further input maintained but did not increase their scores. As previously, we cannot be sure how the additional gains were achieved, or even whether they reflect participants' or partners' behaviours. Further qualitative analyses may shed light on this.

A similarly positive story does not emerge from the narrative assessment. Here there was no evidence that the group receiving Therapy B improved more at time 4 than the group that received no further therapy.

The differing message and narrative results may reflect task demands; i.e. conveying constrained messages may be less demanding than conveying a narrative sequence. However, the comparative percentage scores would argue against this. For example, at time 4, group 2, who did not receive Therapy B, achieved a mean narrative score of 62.73% across the videos, compared to a mean of just under 46% for each message set. Relating this back to our scoring criteria (see appendices 2 and 3), this means that participants were successfully conveying at least one element of each event in the videos; but were falling below this level for the messages. It seems, therefore, that conveying the narratives was marginally less challenging than conveying messages for our participants, possibly because the sequential and contextual information in the narrative supported partners' interpretation.

Statistical factors should also be considered. Evaluations of Therapy B compared two groups of just seven participants. We had aimed for a larger comparison (10 per group), but owing to attrition only 14 participants completed the study. The power to detect change was therefore very small, especially given that scores varied quite widely between participants (see SDs in table 4 and 6). A larger study might have revealed more striking effects. Nevertheless, given the very small group numbers, and the adjusted significance level of .025, the significant findings on the message assessment are striking.

Finally, the assessments may differ in their sensitivity to change. Several features of the narrative data suggest that the stimuli were not problem free. As already noted, two of the 'untreated unrepeated' videos elicited wide ranging control scores. Interestingly, the unrepeated videos also elicited significantly lower scores from the participants with aphasia than the repeated videos. If this effect had been confined to the later assessments, the result would be

expected, since both participants and partners could benefit from rehearsal of the repeated videos. Yet, this was not the case. Rather the effect was present at the outset and did not interact with time. The finding also cannot be attributed to one rogue video, as the administration of the four unrepeated videos was randomised over time for each participant. It seems therefore, that the unrepeated videos were intrinsically more difficult than the repeated ones, as was hinted at in the control data. Unstable video scores were also demonstrated by our separate analysis of time points 3 and 4, where there were unexpected interactions between time and video, albeit at a low level of significance. Indeed, Table 6 shows that group 2 who received no therapy in this period, declined on the treated repeated video but improved quite sharply on the novel one.

Thus our data offer, at best, qualified evidence for Therapy B, since one measure improved but the other did not. It is also striking that, on both measures, group 1 (the group that received Therapy B) had lower scores at time point 3 than group 2. Why this occurred is not clear, given that the groups did not differ on the background assessments. It is possible that the continuing problems of group 1 influenced their decision to stay in therapy. This factor suggests that Therapy B was given a stern test, as it was being conducted with the participants who had the most intractable problems. Random assignment of the participants to the different groups at time point 3 would have been desirable. However, as only seven participants were available to continue this was not practical.

The introduction to this paper argued that measuring the communication changes that arise from therapy is problematic, with few previous accounts of gesture therapy addressing this issue. We attempted to do so through two novel assessments. Group analyses showed that scores on both improved, with evidence that this was due to therapy. The improvement at time point 3 suggested that some benefit accrued from Therapy A, which trained specific gestures and names. These benefits were augmented by Therapy B, which focussed on communication

strategies. However, the latter was demonstrated by only one of the novel assessments, possibly reflecting difficulties with our video task. Given the small number of participants in our study it was also possible to inspect individual gains. Unsurprisingly, these followed the group trend. However, the magnitude of gains varied, with many being below ten percent. This begs the question of what constitutes a clinically significant change on our measures and what factors predict the scale of individual improvements. Such questions need to be addressed through a larger study, e.g. looking at baseline predictors of change and the relationships between gain scores across experimental measures.

Acknowledgements

This study was funded by the Stroke Association (TSA 2006/4). We thank our participants, their partners and our healthy controls. Members of Bury Speakeasy served on our Advisory Group, led by Gill Pearl. Ben Roberts, Bernard Camilleri, Jerry McCombie and Lucy Geraldine Main helped in the preparation of assessment material. Andrew Ford helped to create the project website.

References

Best, W., Grassly, J., Greenwood, A., Herbert, R., Hickin, J. & Howard, D. (2011). A controlled study of changes in conversation following aphasia therapy for anomia. *Disability and Rehabilitation, 33, 3*, 229–242

Boo, M. & Rose, M. (2011). The efficacy of repetition, semantic, and gesture treatments for verb retrieval and use in Broca's aphasia. *Aphasiology, 25 (2),* 154-175.

Booth, S. & Perkins, L. (1999). The use of conversational analysis to guide individualised advice to carers and evaluate change in aphasia: A case study. *Aphasiology, 13 (4-5),* 283 – 303.

Code, C. and Gaunt, C. (1986). Treating severe speech and limb apraxia in a case of aphasia. *British Journal of Disorders of Communication, 21 (1),* 11 – 20.

Coelho, C. (1991). Manual sign acquisition and use in two aphasic subjects. In T. Prescott (ed) *Clinical Aphasiology, 19,* 209 – 218. Austin TX: Pro-Ed.

Coelho, C. and Duffy, R. (1985). Communicative use of signs in aphasia: Is acquisition enough? In R. Brookshire (ed) *Clinical Aphasiology Conference Proceedings, 15,* 222 – 228. Mineapolis MN: BRK Publishers

Coelho, C. and Duffy, R. (1990). Sign acquisition in two aphasic subjects with limb apraxia. *Aphasiology, 4 (1),* 1-8.

Conlon, C. and McNeil, M. (1991). The efficacy of treatment for two globally aphasic adults using Visual Action Therapy. In T. Prescott (Ed.), *Clinical aphasiology* (pp. 185-194). Texas: Pro-Ed.

Cubelli, R., Trentini, P. and Montagna, C. (1991). Re-education of gestural communication in a case of chronic global aphasia and limb apraxia. *Cognitive Neuropsychology, 8 (5),* 369 – 380.

Cunningham, R. and Ward, C. (2003). Evaluation of a training programme to facilitate conversation between people with aphasia and their partners. *Aphasiology, 17 (8),* 687-707.

Daumüller, M. and Goldenberg, G. (2010). Therapy to improve gestural expression in aphasia: a controlled clinical trial. *Clinical Rehabilitation, 24 (1),* 55-65

Douglas, J. M., O'Flaherty, C. A., & Snow, P. C. (2000). Measuring perception of communicative ability: The development and evaluation of the La Trobe communication questionnaire. *Aphasiology, 14 (3),* 251–268.

Fleiss, J.L. (1981) *Statistical methods for rates and proportions*. New York: Wiley.

Frattali, C., Thompson, C., Holland, A., Wohl, C. & Ferketic, M. (1995). *Functional Assessment of Communication Skills for Adults (ASHA/FACS).* Rockville, MD: American Speech-Language-Hearing Association.

Goodwin, C. (2000). Gesture, aphasia and interaction. In D. McNeill (Ed.) *Language and Gesture*. Cambridge UK: Cambridge University Press.

Hopper, T. and Holland, A. and Rewega, M. (2002). Conversational coaching: Treatment outcomes and future directions. *Aphasiology, 16 (7),* 745 – 761**.**

Marshall, J., Best, B., Cocks, N., Cruice, M., Pring, T., Bulcock, G., … Caute, A. (in press). Gesture and naming therapy for people with severe aphasia. *Journal of Speech, Language and Hearing Research*.

Pashek, G.V. (1998). Gestural facilitation of noun and verb retrieval in aphasia: A case study. *Brain and Language*, *65*, 177-180.

Purdy, M., Duffy, R. and Coelho, C. (1994). An investigation of the communicative use of trained symbols following multimodality training. *Clinical Aphasiology, 22*, 345 – 356.

Rao, P. (1995). Drawing and gesture as communication options in a person with severe aphasia. *Topics in Stroke Rehabilitation, 2 (1),* 49 – 56.

Raymer, A. and Thompson, C. (1991). Effects of verbal plus gesture treatment in a patient with aphasia and severe apraxia of speech. In T. Prescott (Ed.) *Clinical Aphasiology*, 285 – 297. Austin TX: Pro-Ed.

Rochon, E., Saffran, E., Berndt, R. and Schwartz, M. (2000). Quantitative analysis of aphasic sentence production: further development and new data. *Brain and Language, 72 (3),* 193 – 218.

Rodriguez, A., Raymer, A.M., & Rothi, L.J.G. (2006). Effects of gesture and semantic-phonologic treatments for verb retrieval. *Aphasiology, 20 (2-3-4)*, 286-297.

Rose, M. (2006). The utility of arm and hand gestures in the treatment of aphasia. *Advances in Speech Language Pathology, 8 (2),* 92-109.

Rose, M., Douglas, J., & Matyas, T. (2002). The comparative effectiveness of gesture and verbal treatments for a specific phonologic naming impairment. *Aphasiology, 16 (10-11),* 1001-1030.

Rose, M. & Douglas, J. (2008). Treating semantic deficits in aphasia with gesture and verbal methods. *Aphasiology, 22 (1),* 1-22.

Rose, M. & Sussmilch, G. (2008). The effects of semantic and gesture treatments on verb retrieval and verb use in Broca's aphasia. *Aphasiology, 22 (7-8),* 691-706.

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

Table 1: Participant Details

	Pseudonym	M/F	Age	Months post onset	Neurological Information (all left hemisphere)	L/R handed	Hemiplegia/ hemiparesis	Occupation
1*	Barbara	F	76	49	Subarachnoid haemorrhage	R	Hemiparesis	Cleaner
2	Claire	F	52	43	Ischaemic, MCA	R	Hemiplegia	Health professional
3	David	M	49	42	No further details	R	Hemiplegia	Creative industries
4*	Edwin	M	75	15	Ischaemic, MCA	R	Hemiparesis	Policeman
5	Gail	F	74	58	Haemorrhagic	R	Hemiparesis	Secretary
6*	George	M	83	13	Haemorrhagic, fronto-parietal	R	Hemiparesis	Technician
7	Jack	M	67	67	Ischaemic, carotid artery	R	Hemiplegia	Teacher
8	Jacob	M	66	43	Ischaemic, carotid artery	R	Hemiplegia	Painter
9*	Kathy	F	55	16	Aneursym	R	Hemiplegia	Cashier
10*	Mabel	F	87	48	Ischaemic	L	Hemiplegia	Nurse
11*	Nora	F	55	26	Subarachnoid haemorrhage / intracerebral haematoma secondary to left MCA aneurysm	R	Hemiplegia	Shop assistant
12	Olivia	F	84	180	No further details	R	Hemiplegia	Translator
13	Robert	M	58	53	Ischaemic	R	Hemiplegia	Computing
14*	Terry	M	64	135	Haemorrhagic	R	Hemiplegia	Businessman

*Received Therapy B; Participant numbers are as reported in Marshall et al. (in press)

Table 2: Mean (S.D.) number of events (/10) conveyed from each video by healthy controls

	Mean	Range
Treated repeated	8.5 (1.43)	6 – 10
Untreated repeated	8.8 (1.32)	6 – 10
Untreated unrepeated (1)	7.9 (1.62)	4 – 10
Untreated unrepeated (2)	7.4 (1.51)	6 – 9
Untreated unrepeated (3)	7.6(1.51)	6 – 9
Untreated unrepeated (4)	7.7 (1.82)	4 – 10

Table 3: Mean % scores (SD) across 14 participants for each set of messages at each time point

Message set (N = 10)	Time 1	Time 2	Time 3	Time 4
Gesture treatment	26.66 (17.02)	31.48 (18.02)	39.48 (16.27)	49.75 (18.73)
Naming treatment	29.35 (13.42)	28.79 (18.88)	36.37 (20.86)	41.98 (19.04)
No treatment	28.46 (17.76)	25.76 (16.03)	34.69 (22.74)	39.55 (19.41)

Table 4: Mean % scores (SD) for each set of messages at time points 3 and 4; comparing participants who did and did not receive Therapy B

Group 1: Received Therapy B		
	Time 3	Time 4
Gesture treatment	32.74 (10.71)	47.39 (17.23)
Naming treatment	26.07 (18.98)	38.38 (19.80)
No treatment	25.19 (17.97)	38.92 (19.14)
Group 2: No Therapy B		
Gesture treatment	46.23 (18.80)	52.11 (21.24)
Naming treatment	46.68 (18.39)	45.59 (19.09)
No treatment	44.34 (24.02)	40.19 (21.27)

Table 5: Mean % scores (SD) for each video at each time point for all 14 participants

	Time 1	Time 2	Time 3	Time 4
Treated repeated	41.78 (24.19)	42.32 (33.23)	56.96 (25.29)	58.39 (25.11)
Untreated repeated	44.82 (30.37)	44.28 (27.23)	55.53 (30.30)	52.32 (32.59)
Untreated unrepeated	29.28 (30.26)	30.53 (28.81)	41.78 (31.24)	55.36 (35.02)

Table 6: Mean % scores (SD) for each video at time point 3 and 4; comparing participants who did and did not receive Therapy B

Group 1: Received Therapy B		
	Time 3	Time 4
Treated repeated	47.85 (22.75)	57.85 (23.20)
Untreated repeated	53.92 (32.85)	45.75 (32.49)
Untreated unrepeated	35.00 (27.84)	40.35 (31.10)
Group 2: No Therapy B		
Treated repeated	66.1 (25.97)	58.92 (28.75)
Untreated repeated	57.12 (30.08)	58.92 (33.81)
Untreated unrepeated	48.55 (35.11)	70.35 (34.14)

Figure 1: Mean % scores on the message assessment between time points 3 and 4; comparing participants who did (group 1) and did not (group 2) receive Therapy B

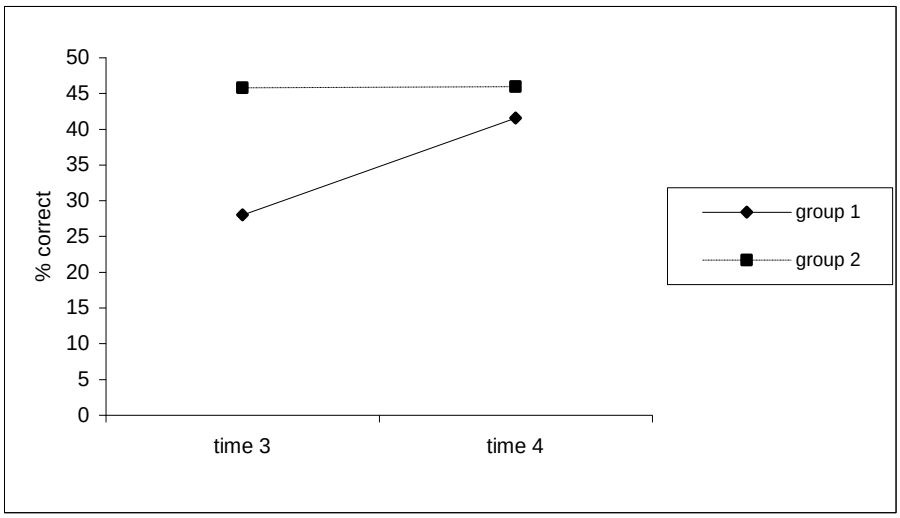


Figure 2: Individual % change scores on the message assessment (Participants 1, 4, 6, 9, 10, 11, and 14 received Therapy B).

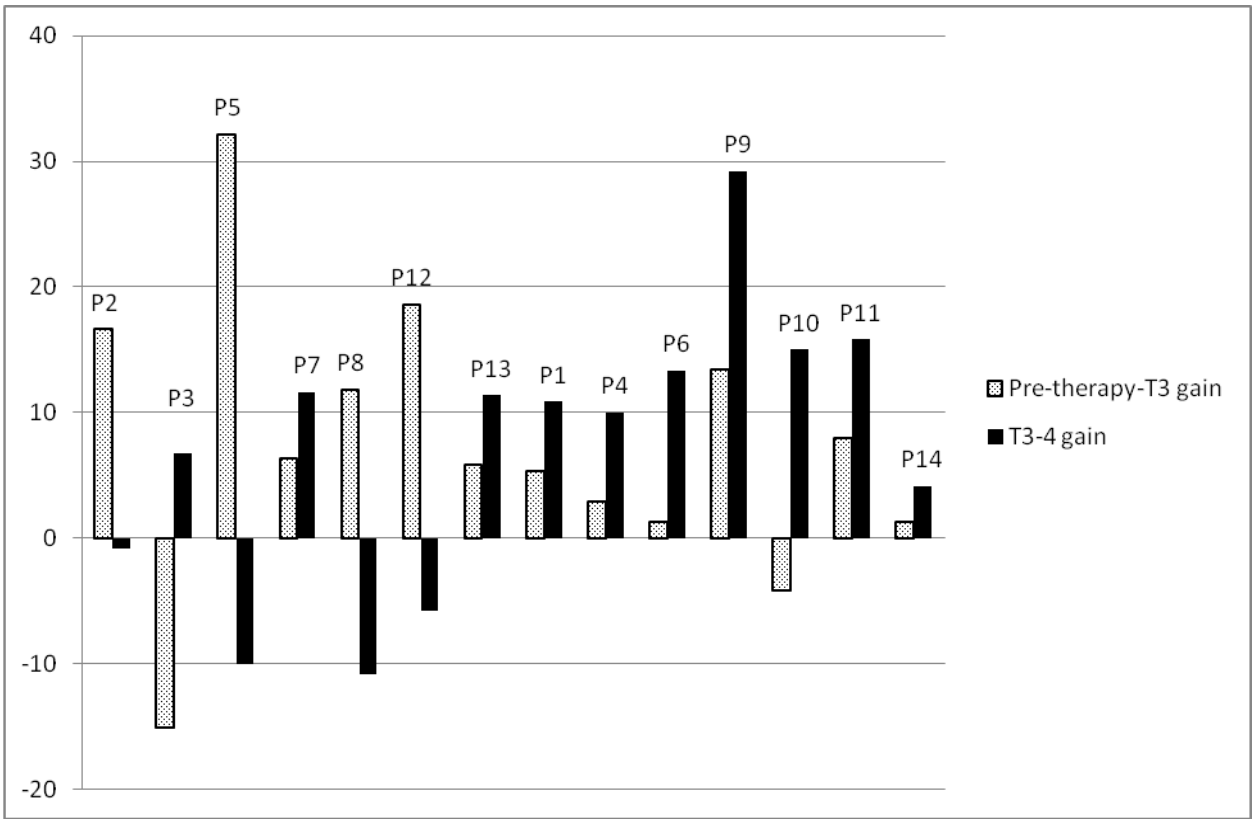
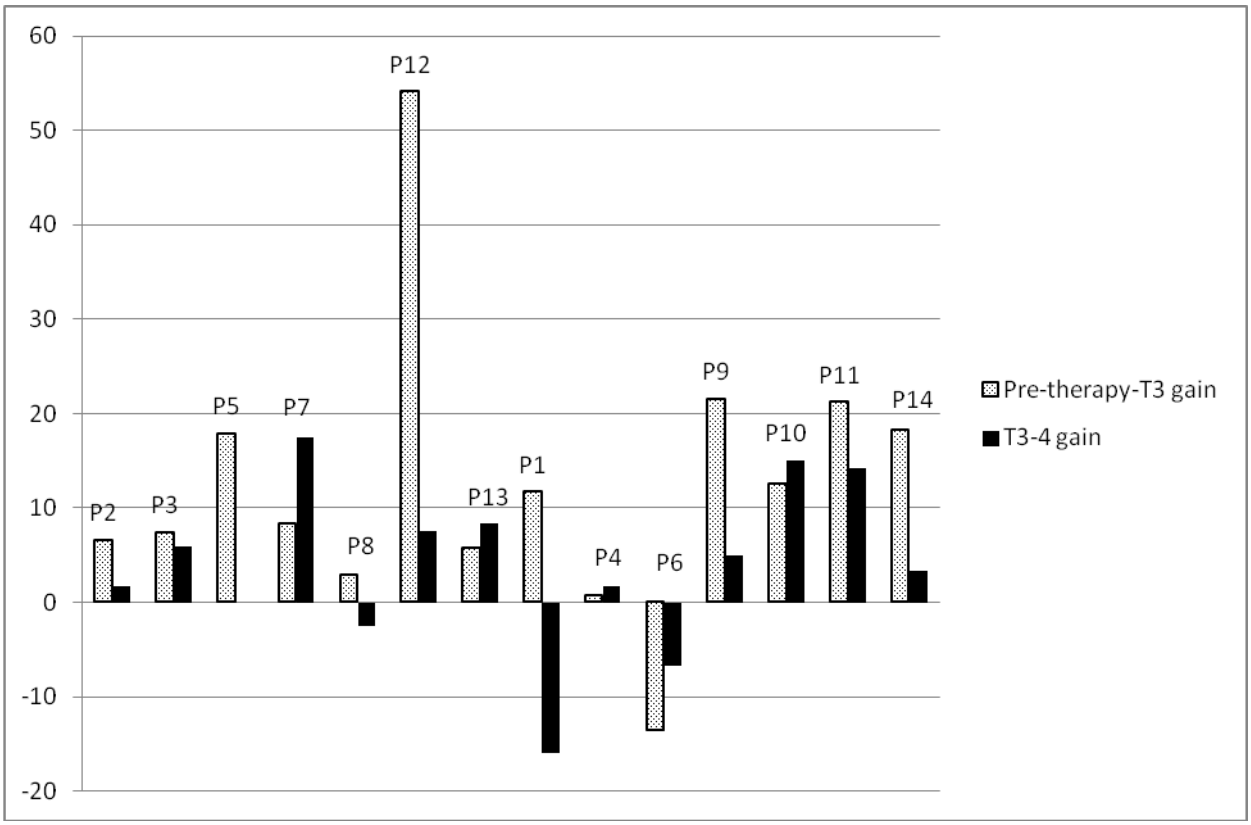


Figure 3: Individual % change scores on the video assessment (Participants 1, 4, 6, 9, 10, 11, and 14 received Therapy B).



Appendix 1: Therapy B task examples

- Conveying an object photo to the partner

The participant was shown an object photo (e.g. of a remote control) and asked to convey it to their partner. If necessary, the therapist practised the target gesture with the participant first, while the partner was not looking.

- Conveying an action photo to the partner

The participant was shown an action picture (e.g. a person ironing) and asked to convey it to their partner. If necessary, the therapist practised the target gesture with the participant first, while the partner was not looking.

- Conveying a simple message to the partner

The message task was explained to the participant and partner and examples of target messages discussed. The message card was shown to the participant who is asked to convey it to the partner. If necessary, facilitation strategies were used, e.g. prompting the participant to focus on the key target item before attempting to convey the message. Stimuli from the message assessment were not used in therapy.

- Conveying a "problem scenario" to the partner

The participant is shown a picture of problem scenario (e.g. a frying pan on fire) and asked to convey it to their partner. If necessary, facilitation strategies were used, e.g. prompting the partner to asked closed questions.

- Describing family photos to the partner or therapist

The participant and partner were asked to find photos of themselves, their family or friends doing different activities, e.g. swimming, eating. A photo was shown to the participant who was asked to convey the person and action to their partner.

- Carrying out a scripted conversation with the partner

The participant and partner were asked to generate ideas of when particular items might be used in conversation. The participant and partner were then prompted to practise these items in a mock, scripted conversation. The partner was given a question to read aloud (e.g. What would you like to eat? How are you feeling?), while a photo, picture or written word was used to prompt the participant to respond using a gesture.

- Telling stories about the previous week to the therapist, with partner support

The therapist asked the participant and partner to think of a story/event from the last week that the participant could recount to the therapist. The therapist used the conversation to demonstrate particular strategies to facilitate conversation.

- Recounting news stories or episodes from TV programmes

The therapist asked the participant to think of a news story that they saw recently and found interesting. The therapist might also use a photo to prompt recollection. The participant is asked to convey the key facts about this story to their partner. Alternatively, the participant chose a TV programme to describe to their partner, usually with the help of pictures.

- Videoing a conversation during the week and discussing strategies used

The participant and partner were provided with a digital video camera and asked to video themselves in conversation when the therapist was not present. The video was watched in the next session. The therapist asked the partner to reflect on what strategies they used, e.g. to help the participant understand or get their message across, and what strategies they could have used. The partner was encouraged to implement these strategies during other tasks. Further video recordings were used to reflect on whether and how these new strategies facilitated communication.

- Discussing communication problems and successes over the previous week

The participant and partner were asked about their communication over the last week. They were asked to give examples of when they used gesture or other strategies successfully. They were also asked about difficulties and how they managed these. If they reported very problematic episodes the therapist helped them to identify strategies that might have helped.

Appendix 2: Scorers instructions in the Message Assessment

- Compare the partner's script to the target message.
- Mark each message according to the following criteria:

0= no element of message successfully conveyed

1= one element conveyed but it is not as specific as the target (e.g. "drink" for *beer*) or is closely related to the target semantically (e.g. "orange" for *apple*). If there is a possible semantic association but this is ambiguous, give "0" (e.g. "card" for *money*).

2= one element of message successfully conveyed. This might be the target item (in bold or in brackets), or relate to the message in which it is embedded. Synonyms should be accepted. For some messages, the target item is not in the message (e.g. for *watch*: What's the time? Or for *car*: "Drive fast"). If the partner has written the target item, it should still be scored as "2" even if that word was not in the message.

3= two elements of message conveyed, but one is semantically inaccurate or not as specific as the target. Or the two elements may have been conveyed but they do not relate to the right person (e.g. "he would like to go to bed" for *Go to bed!*)

4= two elements of the message successfully conveyed. The partner's script does not need to contain all the same words to be scored as "2", but to contain the two main ideas and the right type of speech act (e.g. a question/command/comment).

If in doubt about what score to give, mark it down, i.e. be mean rather than generous.

Examples

Target: **The aeroplane is noisy**

Script: "That's a noisy aeroplane"= 4

Script: "The helicopter is noisy"= 3

Script: "Noisy" = 2

Script; "Loud"= 2

Script: "Plane"= 2

Script: "Ears hurting"= 1

Script: "Bird" = 0

Appendix 3: Scorers instructions in the Narrative Assessment

- Watch the video
- Look at the list of 10 main events in that video
- Compare the partner's script to the list of main events.
- For each of the main events, score:

0= no element of that event is successfully conveyed

1= one element conveyed but it is not as specific as the target (e.g. "fruit" for *she eats an apple*) or is closely related to the target semantically (e.g. "orange" for *apple*).

2= one element of that event is successfully conveyed (e.g. the target object but not the action or vice versa). Synonyms should be accepted.

3= two elements of event conveyed, but one is semantically inaccurate or not as specific as the target (e.g. "she ate an orange/some fruit" for *she eats an apple*.) Or the action and object could be correct but the gender of the person or number of people involved could be wrong (e.g. "he got out of bed" for *she got out of bed* or "the man drank some wine" for *they/the man and the woman drank some wine*).

4= two elements of that event successfully conveyed and the correct number of people and gender. The partner's script does not need to contain all the same words to be scored as "2".

If in doubt about what score to give, mark it down, i.e. be mean rather than generous

Maximum score for each video= 40

Examples

Target: **She eats an apple**

Script: "The lady ate an apple" = 4

Script: "She ate an orange" = 3

Script: "She had an orange" = 3

Script: "He ate an apple" = 3

Script: "Apple" = 2

Script: "She ate something nice" = 2

Script: "Orange" or "fruit" = 1

Script: "Having a drink" = 0

Target: **She puts on her watch**

Script: "She put on her watch"=4

Script: "The lady looks at her watch"= 3

Script: "She tells the time" = 2

Script: "There's a watch"= 2

Script: "She's late" = 1