Gesture and naming therapy for people with severe aphasia:

A group study

Jane Marshall[1]

Wendy Best[2]

Naomi Cocks[1]

Madeline Cruice[1]

Tim Pring[1]

Gemma Bulcock[1]

Gemma Creek[1]

Nancy Eales[1]

Alice Lockhart Mummery[1]

Niina Matthews[1]

Anna Caute[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 a group of people with severe aphasia could learn a vocabulary of pantomime gestures through therapy, and compared their learning of gestures with their learning of words. It also examined whether gesture therapy cued word production and whether naming therapy cued gestures.

Method: Fourteen people with severe aphasia received 15 hours of gesture and naming treatment. Evaluations comprised repeated measures of gesture and word production, comparing treated and untreated items.

Results: Baseline measures were stable, but improved significantly following therapy. Across the group, improvements in naming were greater than improvements in gesture. This trend was evident in most individuals' results, although three made better progress in gesture. Gains were item specific and there was no evidence of cross modality cuing. Items that received gesture therapy did not improve in naming, and items that received naming therapy did not improve in gesture.

Conclusions: Results show that people with severe aphasia can respond to gesture and naming therapy. Given the unequal gains, naming may be a more productive therapy target than gesture for many (although not all) individuals with severe aphasia. The communicative benefits of therapy were not examined, but are addressed in a follow up paper.

Introduction

When formal language is impaired by aphasia gesture seems an obvious alternative. Yet, while some people turn to this medium spontaneously, others do not, so aphasia therapy often targets gesture (Rose, 2006). This paper reports an evaluation of such therapy and compares the outcomes with those achieved by naming therapy.

Gestures are not simply employed by those with impaired language, but are ubiquitous in human communication (Kita, 2009). Definitions of gesture draw a distinction between those that accompany speech (often termed co-speech gestures) and those that stand alone (McNeill, 2005). The latter include pantomimes, the type of gesture focussed in this study. These have been described as a ‘dumb show’ which can convey a complete idea or be linked in sequence to convey a narrative (McNeill, 1992). They can include facial and even vocal elements, in addition to hand movements.

Many argue that a primary function of gesture is to communicate (e.g., see Beattie & Shovelton, 2006). This is the case even for co-speech gestures, with close analysis showing that they supplement, rather than simply reflect what is being said (Kendon, 2000; Melinger & Levelt, 2004). Furthermore, it has been shown that listeners pay attention to gestures and derive information from them (Beattie & Shovelton, 1999 a & b; Cocks, Sautin, Kita, Morgan & Zlotowitz, 2009).

Although gesture production is greatest in face to face conditions, speakers continue to gesture even when they cannot be seen (Alibali, Heath & Myers, 2001). This suggests that gestures perform an additional facilitory role for the speaker (Krauss, Chen & Gottesman, 2000). In line with this, there is evidence that gesturing increases when production is demanding (Melinger & Kita, 2007), or when speech is unrehearsed

(Chawla & Krauss, 1994). There is also evidence that gesture suppression impacts negatively on speakers, e.g., by inducing non fluency (Rauscher, Krauss & Chen, 1996) and Tip of the Tongue states (Frick-Horbury & Guttentag, 1998; although see Beattie & Coughlan, 1999 for counter evidence). The proposal that gesture facilitates speech is also consistent with evidence of neural connections between language and action. For example, we know that hearing face or leg action terms (such as 'lick' and 'kick') stimulates cortical activation in the relevant motor areas (Pulvermuller, 2005); and Transcranial Magnetic Stimulation applied to motor areas speeds lexical decision on related action terms (Pulvermuller, Hauk, Nilulin & Ilmoniemi, 2005).

Theories about the connections between language and gesture have informed models of gesture and speech production. In de Ruiter's Sketch Model, gesture and speech originate from a common conceptualizer (de Ruiter, 2000). Thus they share the same communicative intention and collaborate in conveying that intention. According to this model any gestural facilitation of speech arises at the conceptual level, e.g., by stimulating access to mental imagery (de Ruiter, 1998). The model of Krauss and colleagues (2000) includes a link between the motoric level of gesture production and phonological encoding, so enabling gesture to facilitate access to word forms. Such facilitation is only available from what the authors term 'lexical gestures', i.e., spontaneous gestures that accompany and bear a meaningful relationship to speech.

The centrality of gesture in human communication has important implications for people with aphasia. Most obviously, they may be able to exploit its communicative function to compensate for their impaired language. They may also benefit from its facilitative role. For example, gestures produced during word finding blocks may stimulate access to word forms (Lanyon & Rose, 2009). Stand alone pantomimes might also facilitate conceptual processing, with eventual benefits for naming.

There are several documented cases of people with aphasia who made good use of gesture (e.g., Kemmerer, Chandrasekaran & Tranel, 2007; Marshall, Atkinson, Smulovitch & Thacker, 2004; Wilkinson, Beeke & Maxim, 2010). However, there is also longstanding evidence of aphasic gesture impairments (e.g., Duffy & Duffy, 1981; Duffy, Watt & Duffy, 1994). These may be due to other stroke related disorders of movement (Borod, Fitzpatrick, Helm-Estabrooks & Goodglass, 1989) or executive skills (Purdy & Koch, 2006), or may reflect an impairment in symbolic thinking (Goldenberg, Hartmann & Schlott, 2003). Whatever the reason, it seems that many individuals, particularly those with severe aphasia, need therapeutic input to help them exploit gesture.

Whether gesture can be enhanced by therapy has been investigated in a number of studies. Rose (2006) reviewed 18 that promoted the compensatory use of gesture. Although all reported positive results, the quality of evidence was variable. Most accounts were of single cases or very small groups, and many lacked an experimental design. A recent study (Daumuller & Goldenberg, 2010) attempted to address these concerns. This entered 25 people with severe aphasia into the treatment group (although only 9 completed all phases) and compared their outcomes to untreated controls. Results showed that repeated testing, as experienced by the control group, did not significantly improve gesture production, whereas therapy did. Gains were greatest on gestures that were practised in therapy, but were also evident on unpractised ones, albeit to a much lesser extent. The rate of learning was also explored with the finding that about three hours of therapy was needed for the acquisition of each gesture.

In the above study, gesture was promoted as a compensatory strategy. Others have explored its potential to facilitate lexical access. Rose and Douglas (2001) found that making an iconic gesture for an item significantly improved immediate naming for half (three) of their participants. Furthermore other cues, such as pointing or visualising the use of the object, had no such effect (see also Rose, Douglas & Matyas, 2002).

If ‘one off’ gesture cues can stimulate naming, the effects of gesture therapy may be even more marked. Two single case studies conducted by Pashek (1997, 1998) supported this view. Both compared naming therapy with treatment that combined naming and gesture. Results showed superior outcomes for the verbal plus gesture therapy, although this was not demonstrated statistically.

Further studies have shown that naming therapies with a gestural element can significantly improve the production of nouns (Raymer, Singletary, Rodriguez, Ciampitti, Heilman & Rothi, 2006; Rose & Douglas, 2008; Rose et al, 2002) and verbs (Boo & Rose, 2011; Marangolo et al, 2010; Raymer et al, 2006; Rodriguez, Raymer, & Rothi, 2006; Rose & Sussmilch, 2008). While encouraging, these findings do not confirm the facilitatory role of gesture. In all studies the gestural therapy included an element of verbal practice, such as repeated naming (Marangolo et al, 2010; Raymer et al, 2006; Rodriguez et al, 2006; Rose & Douglas, 2008) or semantic feature analysis (Boo & Rose, 2011). None of the studies replicated Pashek’s finding, or showed that treatments involving gesture were more effective than purely verbal approaches. It should also be noted that not all participants benefited, at least in terms of their naming (e.g., see Rodriguez et al, 2006).

Although questions remain, studies to date suggest that gesture can be acquired as a compensatory strategy in aphasia, and may support the rehabilitation of language. However, not everyone advocates its use. Practitioners of Constraint Induced Aphasia Therapy (CIAT) argue that strategies like gesture should be actively discouraged (or constrained), because they promote the learnt non use of speech (Pulvermuller, Neininger, Elbert, Mohr, Rockstroh, Koebbel & Taub, 2001). Although Constraint Induced therapies have achieved positive outcomes (e.g., Berthier et al, 2009; Meinzer, Djundja, Barthel, Elbert & Rockstroh, 2005; Pulvermuller et al, 2001) this is not a reason to banish gesture from aphasia therapy. Firstly, it may be intensity of practice, rather than constraint, that is the key ingredient in these therapies (e.g. see arguments in Cherney, Patterson, Raymer, Frymark & Schooling, 2010). Secondly, for people with a limited prognosis for speech recovery alternative communication strategies, like gesture, may be the only option.

This study aimed to offer more evidence about the role of gesture in the rehabilitation of aphasia. Our immediate concern was to examine whether a group of people with severe aphasia could acquire a vocabulary of pantomime gestures through therapy. By treating both gesture and naming we aimed to discover whether, for some participants, therapy for gesture offers a better prospect for communication than therapy for spoken or written words. We also explored whether learning of both gestures and words was item specific or extended to untrained items, and whether learning gestures cued the equivalent words (and vice versa). In so doing, we ensured that treatments were modality specific; e.g., no naming practice occurred during gesture treatment, and items treated in one modality were not treated in the other. Thus we contribute to the debate about the potential of gesture to cue word retrieval. The study questions and hypotheses can be summarised as follows:

1. Can people with severe aphasia acquire a vocabulary of gestures?
2. Will learning be specific to treated items, or generalise beyond these?
3. Can people with severe aphasia acquire a vocabulary of spoken or written words?
4. Will learning be specific to treated items or generalise beyond these?
5. How does learning of gestures compare to learning of words?
6. Will gesture therapy cue word production?
7. Will naming therapy cue gesture production?

Given the results of previous studies (e.g., Daumuller and Goldenberg, 2011) we hypothesised that gesture acquisition would be achieved (question 1), with learning largely confined to treated items (question 2). We similarly hypothesised that naming gains would occur (question 3), although again restricted to treated items (question 4). A number of studies suggest that word acquisition is limited in the context of severe impairments (e.g., Marangolo et al, 2010). We therefore hypothesised that gesture gains would outstrip naming gains (question 5). It was difficult to predict the effects of cross modality cueing (questions 6 and 7), since previous studies have not explored the cueing effects of gesture alone, or the impact of naming treatment on gesture production. In line with the Sketch Model (de Ruiter, 2000) we hypothesised that pantomime gestures might prime conceptual processing, with potential gains for word production.

As noted by Daumuller and Goldenberg (2010) studies of gesture therapy should also explore whether acquired skills impact upon communication. This question is addressed in a companion paper (Caute, Pring, Cocks, Cruice, Best & Marshall, submitted). In this we examine whether gesture and/or naming therapy improved performance on communicative tasks (conveying messages and narratives to a partner), and whether gains were enhanced by follow up therapy focussing on interactive skills.

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.

Participants

Twenty four participants were recruited via NHS and independent speech and language therapy services, community groups and self referral. Ten failed to complete the study, mainly because of ill health, so data is reported on fourteen. Analyses of the available data indicate that the groups that completed and failed to complete the study did not differ on screening and background test scores. Table 1 reports details for the 14 participants who completed the study.

All participants had severe aphasia, scoring below 20% on spoken and written naming (assessed by the naming subtest of 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 diagnosed cognitive impairment, such as dementia, and could match objects to photos and drawings with at least 60% accuracy (established via a 10 item screening test). They received no other speech and language therapy during the study.

We also recruited an Advisory Group of four people with aphasia, who were not otherwise involved in the study. The group met three times during the project, to advise us on: (i) participants' information and consent materials, (ii) the test stimuli and (iii) the content and administration of therapy. The group was led by an independent facilitator and the discussion was recorded by students of speech and language therapy. The group recommended several revisions to the project procedures. For example the cueing hierarchy used in therapy was revised in response to their feedback.

Background Assessments

Tests of semantic memory, recognition memory and written and spoken word comprehension (CAT, Swinburn et al, 2004) were carried out during the baseline period to develop a profile of participants' language and cognitive abilities. The oral and limb subtests from the Apraxia Battery for Adults (ABA, Dabul, 2000) were also administered. These required participants to imitate ten oral gestures (e.g., 'bite your lower lip') and ten hand/arm gestures (e.g., 'wave goodbye'). Accurate imitation of the gesture was scored as 2. Inaccurate or 'crude' gestures were scored 1, and a complete inability to perform the gesture was scored 0. As many participants had impaired comprehension, stages involving purely verbal instructions were omitted. Finally, we screened for picture and gesture recognition.

Scores on the background assessments are presented in Table 2, together with control data from Swinburn et al (2004). Control data were not available for the ABA, given the modification to the test procedures. Control data were also not collected on our screening tests. However, as these simply involved matching a picture to an object or iconic gesture, controls would be expected to perform at ceiling.

Design

This study had a repeated measures experimental design. Before therapy, two baseline assessments were conducted (Time 1 & 2), separated by a four-week gap. The assessments were repeated immediately after the first phase of therapy (Time 3) and again at least six weeks later (Time 4). Treated and untreated items were assessed at each time point. The design therefore enabled us to compare change during treatment and no treatment phases as well as on treated and untreated items. All participants received Therapy A between Time 2 and 3, which aimed to train 20 gestures and 20 different words. Seven participants took part in a second phase of therapy (Therapy B). This did not work on the treated vocabulary, but targeted communication strategies. This second phase took place between Time 3 and 4. Therapy B is reported in a companion paper.

Assessment and Therapy Stimuli

Sixty items were included in our experimental assessments. Thirty of these were standard across all participants, while 30 were personal, or selected by each individual. The personal items aimed to increase motivation and took account of the fact that learning is often item specific (e.g., Daumuller & Goldenberg, 2010; Nickels, 2002).

The 30 standard items were concrete, picturable nouns from the categories of food, drink, clothes, transport, furniture and household objects. Gender-specific items such as *razor* were avoided, and all could be gestured using one hand. Piloting with 10 healthy controls (aged >40) confirmed that this was possible. The 30 personal items were chosen for their relevance to the individual and had to differ from the standard items. They had to be picturable and gesturable so that a stranger could understand.

Composition of treatment groups. The 60 items were divided into three groups of 20. One group received gesture treatment, one received naming treatment and one was untreated. Each group included an equal number of standard and personal items.

Table 3 shows that the standard items in the three treatment groups were well matched for gesturability, operativity, familiarity and imageability. Although frequency is less well matched, the group means do not differ significantly ($F(2, 57) = 1.668$, df 2, $p>.2$). The gesturability and operativity ratings were obtained for the study from 31 students of speech and language therapy. For gesturability, the students were asked to rate each item on a five point scale according to how easy or difficult it would be to gesture with one hand. For operativity they were asked to rate each item on a seven point scale, using the criteria defined by Gardner (1973). It was not feasible to gather ratings for the personal items, given that these differed for each participant. These were therefore randomly allocated to treatment and no treatment groups.

Experimental Assessments

At each assessment point, four tasks were carried out. Two explored the impact of therapy on communication, so are described in a companion paper. Here we report the gesture and naming assessments. These employed colour photos and drawings of the 60 items, approximately 15 x 12 cm in size. Images excluded text or brand names.

Gesture assessment. Target gestures for each item in the gesture assessment were modeled to and agreed with participants in a separate session before assessment began. At each time point (Time 1 - 4) participants were assessed on their ability to produce the gestures for the 60 items. They were shown each stimulus picture and told: 'show me with your hands and face what this is'.

Gesture production was videoed and scored by student assessors, using the following procedure. The 240 gestures produced by each participant across the whole study were edited into four separate sets. Each set contained all 60 items, with an equal distribution of gestures produced at each time point. The order of appearance of these gestures was randomised. One scorer was allocated to each set. They scored each gesture in two conditions, with materials presented to them on a power point presentation. In the first ‘blind’ condition, they watched the video clip of the gesture and wrote down their understanding of the target. They were then shown a second slide which presented four written options for selection (the ‘select’ condition). These were the target (e.g., bed), a semantic distractor (e.g., chair), a gesture distractor (e.g., telephone) and an unrelated distractor (e.g., computer). Gesture distractors were items that might elicit a similar gesture to the target. So in the given example, bed was gestured by holding a flat hand against the side of the face and tipping the head to that side. Telephone is also gestured at the side of the face, but with a different handshape. The unrelated distracters were semantically related to the gesture distractors. This ensured that scores did not simply select from two related options. Student assessors were told not to amend their first ‘blind’ response after the options were presented. Throughout, they were unaware of which gestures were produced at which time point.

This procedure yielded two recognition scores for participants’ gesture production at each time point. The ‘blind’ score was the number of gestures that could be recognised without any context. The assessors’ ‘blind’ responses were recorded as correct if they named the target, a synonym of the target or if they produced a phrase containing the target, such as ‘he’s in bed’. The ‘select’ score was the number of gestures that could be recognised against four written options, with one point awarded for each target selection. This dual scoring aimed to be maximally sensitive to change. For example, a small improvement in gesture production might increase select but not blind scores.

To assess inter-rater reliability, eight sets of gestures (14% of the data) were scored by a second student assessor. Each set was taken from a different participant. Correlations showed good levels of reliability (blind scoring $r=.641$, $p<0.05$; select scoring $r=.877$, $p<0.001$).

Naming assessment. Naming of the 60 items was assessed at each time point. Participants were shown each stimulus picture and asked to name it using either speech or writing, depending on their chosen modality. Spoken responses were transcribed by the first author. Responses were scored correct if they named the target or a synonym of the target. Responses that deviated from the target by one phoneme or letter were recorded as errors. 960 responses (28% of the data) were scored independently by two members of the research team to check reliability. Percentage agreement between scorers was 94.4%.

At each time point the naming and gesturing of the 60 items was assessed over two sessions, and in each session half the items were assessed in naming and half in gesture. The order of test administration was counter balanced, e.g., to ensure that that gesture was not always assessed before naming.

Therapy Procedure

Therapy comprised 15 one-hour sessions. The planned regime was two sessions per week, although unforeseen circumstances, such as ill health, varied this for some participants. Half of each session was devoted to gesture treatment and half to naming and the order of treatment (gesture vs naming) alternated across sessions. Therapy followed the same tasks and procedures for all participants (albeit with different personal items) and these were prescribed in a manual. Naming and gesture tasks were applied in blocks of five items, to ensure that items in each treatment group received the same level of exposure. The pictures used in treatment differed from those used in assessment.

Gesture treatment. Each block began with a recognition task. The participant was shown the five target pictures and had to match each one to a gesture produced by the therapist. If they made errors the number of pictures was reduced. When they succeeded with all five, semantically related distractors were introduced.

The focus of therapy moved onto production when participants achieved 100% in the recognition tasks, or after three trials on each block. There were three levels of production task. The first involved producing the gesture from a picture with the therapist's support. The second and third involved conveying a hidden picture to the therapist, so that it could be selected from unrelated (level 2) and then semantically related (level 3) options. Gesture production was supported by a hierarchy of cues, ranging from moulding (in which the therapist made hand on hand contact with the participant), simultaneous copying, delayed copying, providing the first handshape, and giving a verbal cue, such as 'imagine that you are taking a photo' for camera. The maximal level of cue needed by each person was established at the start of the production phase and gradually reduced.

Naming treatment. Naming was very challenging for most participants. To promote some level of success and therefore engagement, naming treatment was conducted in the participant's chosen modality, i.e., spoken or written. All naming assessments were conducted in the corresponding modality.

As in the gesture treatment, each block began with a recognition task. Participants were shown the five target pictures and were required to match them to a spoken or written name. As above, the recognition task was adjusted by reducing the number of pictures or introducing semantic distractors. The criteria for progression was the same as in gesture treatment (100% correct or after three trials). The first production task involved naming pictures with the therapist's support. The second and third involved naming a hidden picture so that it could be identified from unrelated (level 2) and then related (level 3) distractors. Naming was supported by an increasing hierarchy of cues: a verbal lead in ('that's a …'), a semantic definition (such as 'that's a piece of furniture that you sleep on'), a semantic closure (such as 'on Sundays they stay in ..' *bed*), a minimal phonological cue (first phoneme), a maximal phonological cue (first syllable or initial consonant plus vowel), and provision of the target for the person to repeat.

Results

Gesture Assessment. This analysis addressed the first two study questions: Can people with severe aphasia acquire a vocabulary of gestures; and will learning be specific to treated items, or generalise beyond these? It also examined whether items that received naming treatment showed benefits in gesture production (question 7), and whether results differed across scoring methods and for standard and personal items.

The results of the gesture assessment (see Table 4) were subjected to a four factor within subject analysis of variance (ANOVA). Factors were time (4 levels: time 1, 2, 3 & 4), scoring condition (2 levels: blind and select), type of item (2 levels: personal and standard) and treatment (3 levels: gesture treatment, naming treatment and untreated).

There was a significant main effect of time ($F(3,39) = 3.188$, $p < .05$; $\eta^2_p = 0.197$). Planned comparisons showed no difference between the two baselines (1 vs 2) or between the post therapy and follow up assessments (3 vs 4). The comparison between the two baseline and the two post therapy assessments (1 & 2 vs 3 & 4) was significant

($F(1, 39) = 8.48$, $p < .01$; $\eta^2_p = 0.24$). These results show that participants' gesture production improved significantly over the course of the study, and that gains were contingent upon the receipt of therapy.

There was also a main effect of treatment ($F(2, 26) = 5.298$, $p < .01$; $\eta^2_p = 0.29$). Items that received gesture treatment scored more highly than items receiving naming and no treatment. Figure 1 suggests that these effects arose primarily from the improved performance on the items treated by gesture at time point 3. However, the interaction between time and treatment was not significant ($p = .11$).

The final significant main effect was for the scoring condition ($F(1, 13) = 451.7$, $p <.001$; $\eta^2_p = 0.97$). Unsurprisingly, gestures were scored more highly in the select than the blind condition. This, however, did not interact with time; i.e., gains over time were no greater in the select than the blind scoring condition.

There was no significant main effect of item ($p = .14$), although personal items achieved marginally lower scores than standard items. There was, however, an interaction between item and condition ($F(1, 13) = 7.381$, $p < .05$; $\eta^2_p = 0.36$). Figure 2 shows that personal items were particularly disadvantaged in the more stringent blind scoring condition. All other interactions were not significant.

Naming Assessment. This analysis addressed the 3rd and 4th study questions: Can people with severe aphasia acquire a vocabulary of spoken or written words; and will learning be specific to treated items or generalise beyond these? It also examined whether items that received gesture treatment showed benefits in naming (question 6) and whether gains differed across standard or personal items.

The results of the naming assessment (see Table 5) were subjected to a three factor within subject ANOVA. Factors were time (4 levels: time 1, 2, 3 & 4); type of item (2 levels: personal and standard) and treatment (3 levels: gesture treatment, naming treatment and untreated). Where the data failed the sphericity assumption, the Greenhouse-Geiser correction was applied.

The analysis revealed a main effect of time ($F\ (1.17, 15.21) = 10.93$, $p < .001$; $\eta^2_p = 0.46$). Planned comparisons found no difference between times 1 and 2 (before therapy), or times 3 and 4 (after therapy). However, the comparison of the two baseline with the two post therapy assessments (1 & 2 vs 3 & 4) was highly significant ($F\ (1, 39) = 31.65$, $p < .001$; $\eta^2_p = 0.61$). Thus participants' naming improved over the course of the study, with gains being contingent on the receipt of therapy.

There was also a main effect of treatment ($F\ (1.11, 14.47) = 5.23$, $p < .05$; $\eta^2_p = 0.29$). Items that received naming therapy achieved higher scores than items receiving gesture or no treatment.

Figure 3 shows a similar pattern to Figure 1. As in the gesture assessment, gains occurred mainly on items that received the relevant treatment (in this case naming). Here, however, the interaction of treatment with time was highly significant ($F\ (3.28, 42.61) = 7.63$, $p < .001$).

The final main effect of item was also significant ($F\ (1, 13) = 5.15$, $p < .05$; $\eta^2_p = 0.28$), with standard items scoring more highly than personal. Puzzlingly, this interacted with treatment ($F\ (1.4, 18.24) = 4.39$, $p < .05$; $\eta^2_p = 0.02$). The interaction was derived mainly from the untreated group, with standard items being named surprisingly well, and personal items named very poorly. None of the other interactions was significant.

Gesture vs Naming. The final analyses compared the learning of gestures with the learning of words (study question 5). The analyses so far showed that gains on the gesture and naming assessments were largely confined to items that were treated in the relevant modality. So gesture gains arose mainly on items that received gesture treatment, and naming gains arose mainly on items that received naming treatment. The treatments were therefore compared by analysing the within modality responses over time for each treatment group. Thus gesture responses to the gesture treatment items were compared to naming responses to the naming treatment items. Blind gesture scores were used for the analysis.

A three factor within subject ANOVA was conducted. The factors were time (4 levels: time 1, 2, 3 & 4), response (2 levels: gesture and naming) and item (personal and standard). Where results failed the sphericity assumption, the Greenhouse-Geiser correction was applied. The only significant main effect was for time ($F (1.61, 20.9) = 22.66$, $p <.001$; $\eta^2_p = 0.63$), confirming the overall improvement in responses during the course of the study. Although there was no significant main effect of response, the interaction between time and response was significant ($F (1.71, 22.24) = 3.99$, $p < .05$; $\eta^2_p = 0.23$). This interaction is illustrated in Figure 4, showing that, while responses in both modalities improved, gains in naming outstripped gains in gesture. No other interactions were significant.

The ANOVA showed that, across all participants, naming was more responsive to treatment than gesture. However, this might conceal important individual differences. Individual gain scores in gesture and naming were therefore computed. These were the differences between scores at time point 3 and the mean of scores at time point 1 and 2. Total gesture and naming scores were used (i.e., /60). Figure 5 shows the gain scores of each participant. Most individuals followed the general trend, some dramatically so. However, participants 4, 8, and 10 profited more in gesture than naming.

Discussion

This section will evaluate the study questions against the results, and consider their clinical implications. Our first questions asked whether participants could learn a vocabulary of gestures and whether learning would be confined to treated items. These questions can be answered by a qualified 'yes'. Responses on the gesture assessment were stable over the baseline period, but improved following therapy, and the gain was maintained at time point 4. Results also showed that gains occurred mainly on items that received gesture treatment. There was an effect of treatment group, favoring the gesture items; and although there was no interaction between treatment and time, Figure 1 indicates a strong trend in this direction.

We similarly asked whether participants could learn a vocabulary of words, and whether here too learning would be item specific. Again results were positive. Naming improved after therapy but not before, and although scores fell at time point 4, they did not do so significantly. As with gesture, there was an effect of treatment group, which in this analysis interacted with time. Thus, naming gains were highest in the group that received naming treatment.

The fifth study question addressed the comparative learning of gestures and words. Results were unequivocal. Improvements in naming clearly outstripped improvements in gesture, both at the group level and in most individual gain scores. This finding was surprising, given the severity of participants' aphasia, and ran counter to our hypothesis. As an explanation, inequalities in the treatment regimes can be dismissed, since words and gestures were given the same level of exposure and were practised in very similar tasks. Another possibility is that gestures were disadvantaged by our assessment process. However, the inclusion of the select condition aimed to maximize scoring sensitivity. It was also striking that, although blind scores were much lower than select, the gains achieved in both were similar. A third possibility is that participants were more motivated to learn words than gestures, although this was not evident during therapy. We therefore offer three alternative explanations.

Our first proposal is that gesture and naming impose unequal learning demands. Naming treatment aims to restore access to lexical forms that were laid down prior to the stroke. The lack of such pre-established representations for gestures means that novel forms have to be acquired, potentially making them more challenging. Employing gesture also requires participants to switch from speech into a non habitual response mode, which may increase the challenge (Purdy and Koch, 2006).

The second, related proposal is that pantomime gestures are particularly demanding to acquire. It has been suggested that pantomimes are 'special gestures' that impose heavy cognitive demands, e.g., of working memory (Bartolo, Cubelli, Della Sala & Drei, 2003). As a result, they may be particularly challenging for people with severe post stroke impairments.

The final possibility is that baseline factors adversely affected the learning of gesture. For example, it is striking that all participants displayed a degree of limb apraxia (see table 2). The negative impact of apraxia on gesture production is disputed. Research has shown that it may not inhibit gesture use in natural conversations (Rose & Douglas, 2003), and even people with severe limb apraxia can participate successfully in gestural therapy for word retrieval (Raymer et al, 2006). In the current study, limb apraxia scores were not predictive of gestural gains. For example, one of the three individuals who improved more in gesture than naming (participant 8) had the lowest limb apraxia score, and participants with similar apraxia scores achieved very different outcomes in gesture (e.g. compare participant 1 and 10). Nevertheless, the small participant numbers make our findings inconclusive. A larger study could explore the role of apraxia further by correlating baseline apraxia assessment scores with treatment gains.

The final study question concerned cross modality generalization, or the degree to which each treatment cued responses in the other modality. Such generalization would be signaled by improved naming of the items that received gesture treatment, and improved gesturing of the items that received naming treatment. In fact, Figures 1 and 3 provide no evidence that this occurred. Although naming of the gesture treatment items improved marginally at time 3, the gain was less than on the items that received no treatment. Gesturing of the naming treatment items remained virtually unchanged throughout the study.

The lack of a cuing effect for gesture runs counter to previous findings (e.g., Rose and Douglas, 2008; Marangolo et al, 2010), so requires explanation. Differences between participant groups may be a factor. Participants in the current study had profound naming impairments, with 10 showing evidence of impaired semantic memory (see table 2). Their prognosis for gestural cuing, therefore, may have been poor. Indeed several previous studies suggest that people with semantic level impairments respond least well to gestural naming therapy (Marangolo et al, 2010; Rose and Douglas, 2001; Rodriguez et al, 2006; Rose & Sussmilch, 2008). Another difference may lie in the therapy. Previous studies employed therapies that integrated gesture with naming practice, e.g., by encouraging participants to gesture prior to word production (Rose & Douglas, 2008). The current study deliberately separated the modalities during treatment. This gives rise to two possibilities. Previous studies may have overestimated the contribution of gesture to their results, with facilitation arising mainly from the verbal component of therapy. Alternatively facilitation may occur only when gesture is employed with speech. This is argued by Krauss and colleagues (2000), although not with respect to the pantomime forms of gesture that were taught in this therapy. We hypothesized that stand alone pantomime gestures might prime conceptual processing, with potential gains for either spoken or written naming. Our results do not support this hypothesis.

The study produced one further, unexpected finding. The assessment and therapy stimuli comprised 30 standard and 30 personal items, the latter being chosen by each participant. An effect of item was not hypothesized, although any prediction would probably favor personal items, given that participants were presumably highly motivated to work on these. Yet, in both the gesture and naming results, personal items scored below standard. This was a general feature for naming, and specific to the blind condition for gesture. It was not possible to match personal and standard items on psycholinguistic variables, given that participant chose their own sets. We conducted a post hoc analysis of the available frequency, familiarity and imageabily values for all personal items chosen across the group (N = 216). This indicated that the personal sets were marginally less imageable and familiar than the standard sets, and markedly less frequent (mean imageability = 572.68 (59.25); mean familiarity = 567.61 (38.51); mean frequency = 8.99 (10. 35); see table 3 for standard item values). It seems, therefore, that our participants opted to work on a vocabulary that was more challenging than the one set by the study team, possibly because they were targeting items that they knew to be problematic. In this respect they were in line with recent proposals calling for complex, rather than simple therapy stimuli (Kiran, 2007; 2008).

Results of this study show that improvements in gesture and naming can be achieved by people with severe and chronic aphasia, and in response to a limited therapy dose. As in Daumuller and Goldenberg (2010) gains in gesture were modest, with an average of just under two new gestures acquired from seven and a half hours of gesture treatment. The rate of word learning was greater, with an average of 8 words acquired.

It might be argued that the results of this study do not strongly advocate for the use of gesture in therapy for people with severe aphasia. Gestural gains were modest, and well exceeded by the naming gains. There was also no indication that gesture therapy cued naming. However, here the individual results are important. Although most participants followed the group trend, three demonstrated improvements in gesture that were not matched by naming, indeed in all three instances, naming completely failed to benefit from therapy. It seems that for some individuals gesture is more responsive to treatment than conventional language. A larger study might reveal the baseline factors that identify such individuals. Another unresolved question relates to communication, or whether or not gesture and naming treatment benefits interactions. This is addressed in our companion paper.

Acknowledgments

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. We thank Carol Sacchett and members of the adult speech and language therapy teams at Whipps Cross University Hospital and Ealing Primary Care Trust, who gave us feedback on our therapy manual. Many members of our student body supported this study, e.g., by scoring the gesture data and trialling our therapy manual. Andrew Ford helped to create the project web site.

References

Alibali, M., Heath, D., & Myers, H. (2001). Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. *Journal of Memory and Language*, 44, 169 – 188.

Bartolo, A., Cubelli, R., Della Sala, S. & Drei, S. (2003). Pantomimes are special gestures which rely on working memory. *Brain and Cognition*, 53, 483 – 494,

Beattie G. & Coughlan J. (1999). An experimental investigation of the role of iconic gestures in lexical access using the tip-of-the-tongue phenomenon. *British Journal of Psychology*, 90, 1, 35-56.

Beattie, G. & Shovelton, H. (1999a). Do iconic hand gestures really contribute anything to the semantic information conveyed by speech? An experimental investigation. *Semiotica*, 123, 1-30.

Beattie, G. & Shovelton, H. (1999b). Mapping the range of information contained in the iconic hand gestures that accompany spontaneous speech. *Journal of Language and Social Psychology*, 18, 438 – 462.

Beattie, G. & Shovelton, H. (2006). A critical appraisal of the relationship between speech and gesture and its implications for the treatment of aphasia. *Advances in Speech-Language Pathology*, 8(2), 134-139.

Berthier, M. L., Green, C., Lara, J. P., Higueras, C., Barbancho, M. A., Dávila, G., & Pulvermüller, F. (2009). Memantine and constraint-induced aphasia therapy in chronic post-stroke aphasia. *Annals of Neurology,* 65 (5), 577 – 585.

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

Borod, J., Fitzpatrick, P., Helm-Estabrooks, N., & Goodglass, H. (1989). The relationship between limb apraxia and the spontaneous use of communicative gesture in aphasia. *Brain and Cognition*, 10, 121-131.

Caute, A., Pring, T., Cocks, N., Cruice, C. & Marshall, J. (submitted for publication). Enhancing Communication through Gesture and Naming Therapy.

Chawla, P., & Krauss, R. M. (1994). Gesture and speech in spontaneous and rehearsed narratives. *Journal of Experimental Social Psychology* , **30**, 580-601.

Cherney, L., Patterson, J., Raymer, A., Frymark, T. & Schooling T. (2010). Updated Evidenced-based systematic review: Effects of intensity of treatment and constraint induced language therapy for individuals with stroke-induced aphasia. ASHA.

Cocks, N., Sautin, L., Kita, S., Morgan, G., & Zlotowitz, S. (2009). Gesture and speech integration: An exploratory study of a man with aphasia. *International Journal of Language and Communication Disorders*, 44, 795-804.

Dabul, B. (2000) Apraxia Battery for Adults, Second ed. Austin, Tx: Pro-Ed.

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

De Ruiter, J.P. (1998). *Gesture and speech production.* Doctoral dissertation at Catholic University of Nijmegen, Netherlands.

De Ruiter, J.P. (2000). *The production of gesture and speech.* In McNeill (Ed.) Language and Gesture. Cambridge: Cambridge University Press

Duffy, R. & Duffy, J. (1981). Three studies of deficits in pantomimic expression and pantomimic recognition in aphasia. *Journal of Speech and Hearing Research*, 46, 70-84.

Duffy, R., Watt, J. & Duffy, J. (1994). Testing causal theories of pantomime deficits in aphasia using path analysis. *Aphasiology*, 8, 361-379.

Frick-Horbury, D., & Guttentag, R. E. (1998). The effects of restricting hand gesture production on lexical retrieval and free recall. *American Journal of Psychology* 111, 1, 43 – 62.

Goldenberg, G., Hartmann, K. & Schlott, I. (2003). Defective pantomime of object use in left brain damage: Apraxia or asymbolia? *Neuropsychologia*, 41, 1565-1573

Kemmerer, D., Chandrasekaran, B., & Tranel, D. (2007). A case of impaired verbalization but preserved gesticulation of motion events. *Cognitive Neuropsychology*, 24 (1), 70–114.

Kendon, A. (2000). Language and gesture: Unity or duality? In D. McNeill (ed) *Language and Gesture.* Cambridge: Cambridge University Press.

Kiran, S. (2007). Semantic complexity in the treatment of naming deficits. *American Journal of Speech Language Pathology*, 16 (Feb), 1 – 12.

Kiran, S. (2008). Typicality of inanimate category exemplars in aphasia: Further evidence for semantic complexity. *Journal of Speech, Language and Hearing Research*, 51, 1550 – 1568.

Kita, S. (2009). Cross-cultural variation of speech-accompanying gesture: A review. *Language and Cognitive Processes*, 24(2), 145-167.

Krauss, R., Chen, Y. & Gottesman, R. (2000). Lexical gestures and lexical access: a process model. In D. McNeill (ed) *Language and Gesture*, 261 - 283. Cambridge: Cambridge University Press.

Lanyon, L., & Rose, M. (2009). Do the hands have it? The facilitation effects of arm and hand gesture on word retrieval in aphasia. *Aphasiology,* 23, 809-822

Marangolo, P., Bonifazi, S., Tomaiuolo, F., Craighero, L., Coccia, M., Aleoe, G., Provinciali, L. & Cantagallo, A. (2010). Improving language without words: First evidence from aphasia. *Neuropsychologia*, 48, 3824 – 3833.

Marshall, J., Atkinson, J., Smulovitch, E., Thacker, A. & Woll, B. (2004). Aphasia in a user of British Sign Language: Dissociation between sign and gesture. *Cognitive Neuropsychology*, 21 (5), 537 - 554.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.

McNeill, D. (2005). *Gesture and Thought*. Chicago: University of Chicago Press.

Meinzer, M., Djundja, D., Barthel, G., Elbert, T. & Rockstroh, B. (2005). Long-Term Stability of improved language functions in chronic aphasia after Constraint-Induced Aphasia Therapy. *Stroke,* 36, 1462-1466.

Melinger, A., & Levelt, W. (2004). Gesture and the communicative intention of the speaker, *Gesture*, 4, 2, 119 – 141.

Melinger, A. & Kita, S. (2007). Conceptualisation load triggers gesture production. *Language and Cognitive Processes*, 22, 4, 473 – 500.

Nickels, L. (2002). Therapy for naming disorders: Revisiting, revising, and reviewing. *Aphasiology*, 16 (10/11), 935–979.

Pashek, G. (1997). A case study of gesturally cued naming in aphasia: Dominant versus nondominant hand training. *Journal of Communication Disorders*, 30, 349-366.

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

Pulvermuller, F., Neininger, B., Elbert, T.E., Mohr, B., Rockstroh, B., Koebbel, P. & Taub, E. (2001). Constraint-induced therapy of chronic aphasia following stroke. *Stroke*, 32, 1621-1626.

Pulvermuller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience,* 6 (7), 576 – 582.

Pulvermuller, F., Hauk, O., Nilulin, V., & Ilmoniemi R. (2005). Functional links between motor and language systems. *European Journal of Neuroscience*, 21, 793 – 797.

Purdy, M. & Koch, A. (2006). Prediction of strategy usage by adults with aphasia. *Aphasiology*, 20 (2-4), 337-348.

Rauscher, F. B., Krauss, R. M., & Chen, Y. (1996). Gesture, speech and lexical access: The role of lexical movements in speech production. *Psychological Science*, **7**, 226-231.

Raymer, A.M., Singletary, F., Rodriguez, A., Ciampitti, M., Heilman, K.M., & Rothi, L.J.G. (2006). Gesture training effects for noun and verb retrieval in aphasia. *Journal of the International Neuropsychological Society*, *12*, 867-882.

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

Rose, M. (2006). The utility of gesture treatments in aphasia. *Advances in Speech Language Pathology*, 8 (2), 92-109.

Rose, M. & Douglas, J. (2003). Limb apraxia, pantomine, and lexical gesture in aphasic speakers: Preliminary findings. *Aphasiology*, 17(5), 453-464.

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

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

Rose, M. & Douglas, J. (2001). The differential facilitatory effects of gesture and visualisation processes on object naming in aphasia. *Aphasiology,* 15, 977-990.

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.

Wilkinson, R., Beeke, S. & Maxim, J. (2010). Formulating actions and events with limited linguistic resources: Enactment and iconicity in agrammatic aphasic talk. *Research on Language and Social Interaction*, 43(1), 57–84.

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

Table 2: Background test results

	Pseudonym	Spoken naming	Written naming	Comp spoken	Comp written	Semantic memory	Recognition memory	Limb apraxia	Oral apraxia	Object/ picture matching	Gesture/ picture matching
1	Barbara	0/48	0/48	2/30	7/30	2/10	3/10	14/20	14/20	8/10	7/10
2	Claire	0/48	9/48	28/30	27/30	10/10	9/10	15/20	11/20	10/10	10/10
3	David	0/48	0/48	11/30	0/30	8/10	9/10	14/20	12/20	9/10	10/10
4	Edwin	0/48	0/48	14/30	6/30	4/10	2/10	13/20	12/20	6/10	4/10
5	Gail	9/48	0/48	30/30	27/30	10/10	10/10	12/20	20/20	10/10	10/10
6	George	0/48	0/48	7/30	6/30	4/10	7/10	13/20	13/20	8/10	4/10
7	Jack	0/48	2/48	20/30	26/30	7/10	5/10	15/20	10/20	10/10	10/10
8	Jacob	2/48	0/48	16/30	4/30	2/10	2/10	8/20	10/20	7/10	7/10
9	Kathy	0/48	2/48	18/30	16/30	10/10	4/10	13/20	18/20	10/10	9/10
10	Mabel	0/48	0/48	21/30	22/30	8/10	10/10	13/20	7/20	9/10	7/10
11	Nora	0/48	2/48	25/30	0/30	2/10	9/10	10/20	3/20	10/10	9/10
12	Olivia	4/48	0/48	20/30	26/30	8/10	10/10	10/20	10/20	9/10	8/10
13	Robert	2/48	0/48	24/30	22/30	10/10	10/10	14/20	18/20	10/10	10/10
14	Terry	4/48	0/48	18/30	15/30	6/10	9/10	17/20	16/20	9/10	10/10
	Control Mean (SD)[1]	46.37 (1.6)		29.15 (1.35)	29.63 (.79)	9.81 (.40)	9.7 (.54)				

[1] Data from Swinburn et al, 2004

Table 3: Mean Values (S.D.) for the Standard Items in Each Treatment Group

	Gesture Treatment	Naming Treatment	No Treatment
Gesturability	2.43 (.8)	2.8 (.75)	2.5 (.6)
Operativity	5.91 (.5)	5.69 (.6)	5.97 (.4)
Familiarity[1]	580.6 (34.8)	580.9 (41)	608.3 (24.4)
Imageability[1]	579 (29.3)	602.8 (26.4)	614.7 (25.9)
Frequency[1]	42.6 (37.2)	53.6 (61.1)	99.3 (105.5)

[1] Values drawn from the MRC Data Base

Table 4: Mean % correct blind and select gesture scores (SD) at each time point for each group of items

	Time 1		Time 2		Time 3		Time 4	
Treatment group (N = 20)	blind	select	blind	select	blind	select	blind	select
Gesture treatment	12.50 (12.82)	58.96 (15.12)	13.93 (8.59)	55.36 (16.59)	21.07 (14.70)	67.14 (15.53)	22.50 (14.11)	66.07 (16.66)
Naming treatment	13.21 (13.95)	51.78 (20.81)	13.92 (10.03)	55.75 (20.70)	12.14 (11.04)	56.43 (21.07)	13.57 (10.46)	60.35 (20.89)
No treatment	15.71 (14.39)	52.14 (24.39)	15.71 (12.17)	53.92 (22.28)	17.86 (16.73)	54.28 (20.74)	16.78 (16.24)	59.28 (18.38)

Table 5: Mean % correct naming scores (SD) at each time point for each group of items

Treatment group (N = 20)	Time 1	Time 2	Time 3	Time 4
Gesture treatment	9.28 (16.62)	11.80 (16.83)	17.14 (22.42)	16.78 (18.77)
Naming treatment	7.85 (12.51)	12.85 (21.27)	36.07 (32.35)	30.71 (31.43)
No treatment	15.40 (21.12)	15.35 (17.70)	23.21 (24.69)	22.50 (25.02)

Figure 1: The mean percentage of correct gestures (across scoring conditions) at each time point for each group of items

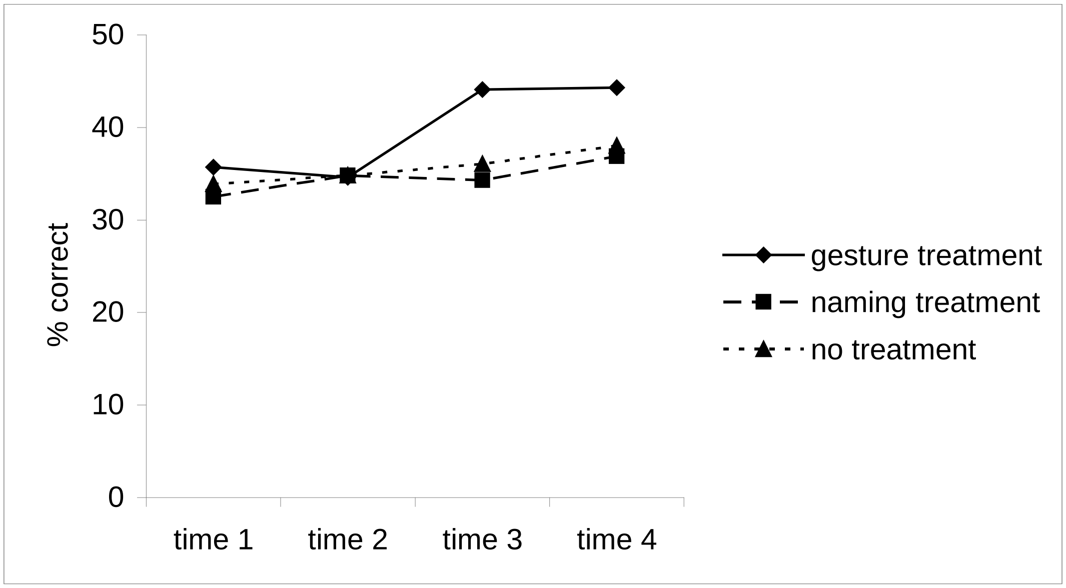

Figure 2: Mean percentage correct for personal and standard items in the two gesture scoring conditions

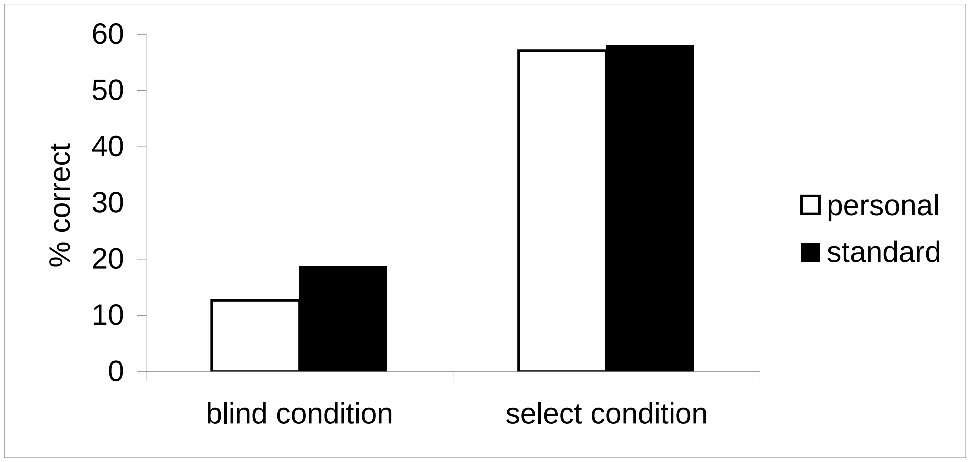

Figure 3: The mean percentage of correct names at each time point for each group of items

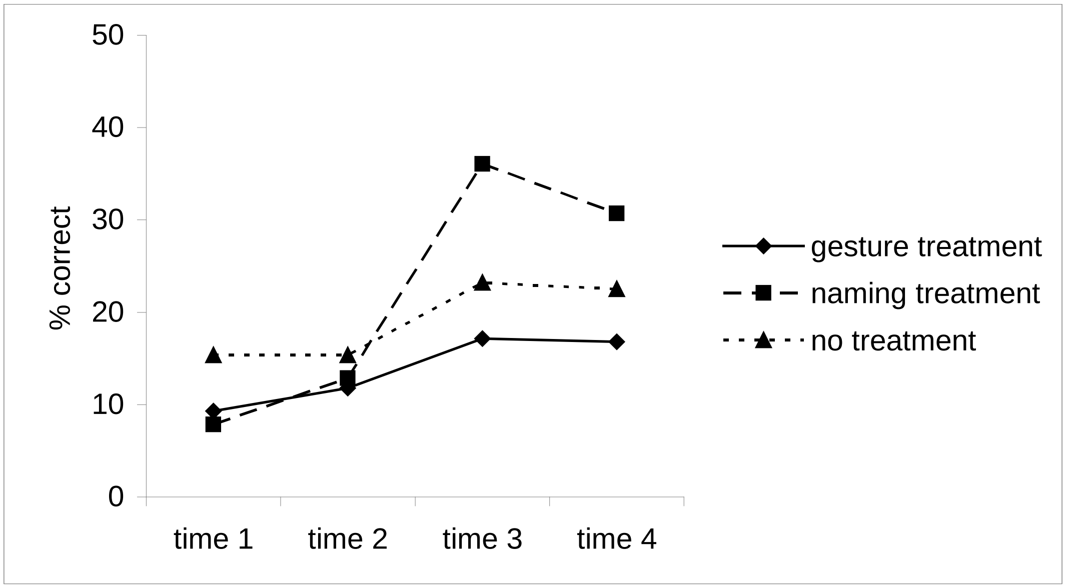

Figure 4: The mean percentage of correct gesture and naming responses over time (gesture and naming treatment groups only)

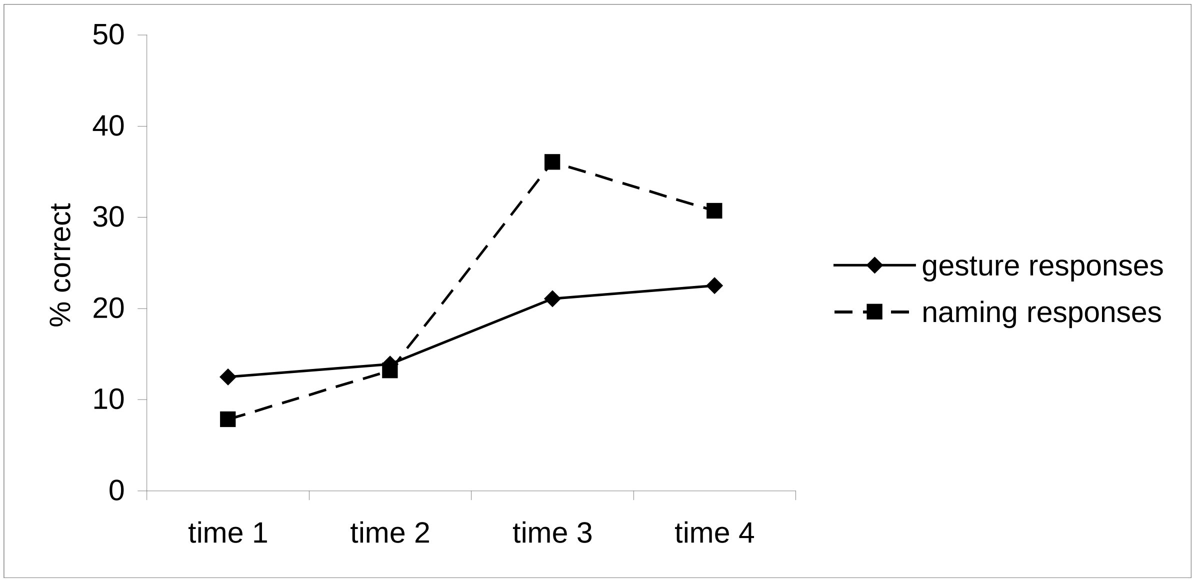

Figure 5: Individual gain scores in gesture and naming