Playing a team game improves word production in post-stroke aphasia

Romani, Cristina; Thomas, Lucinda; Olson, Andrew; Lander, Louise

DOI:
10.1080/02687038.2018.1548205

License:
Other (please specify with Rights Statement)

Document Version
Peer reviewed version

Citation for published version (Harvard):
Romani, C, Thomas, L, Olson, A & Lander, L 2018, 'Playing a team game improves word production in post-stroke aphasia', *Aphasiology*, vol. 33, no. 3, pp. 253-288. https://doi.org/10.1080/02687038.2018.1548205

Playing a team game
improves word production in post-stroke aphasia

Cristina Romani [1], Lucinda Thomas[1], Andrew Olson [2], & Louise Lander[3]

1. School of Life and Health Sciences, Aston University

2. Department of Psychology, University of Birmingham

3. Moor Green Outpatient Brain Injury Unit-Birmingham Community Healthcare NHS Foundation Trust

Address correspondence to:
Cristina Romani
Aston University
Aston Triangle
B4 7ET, Birmingham, UK
E.mail: C.Romani@Aston.ac.uk

Key words : aphasia, group therapy, game therapy, social games, picture naming

STRUCTURED ABSTRACT

***Background*.** High intensity, one-to-one rehabilitation therapy is effective in the treatment of post-stroke aphasia, but it can put strain on public health providers, as well as lead to high attrition. Working within a group of peers may be efficient for professional speech and language therapists, as well as reduce feelings of isolation and lack of confidence in patients, which can negatively affect progress. Evidence-based, structured group-based approaches, however, are lacking.

Aims: We wanted to assess the feasibility a new group-delivered game-based intervention, designed to provide efficacious word-retrieval rehabilitation, in a cost-effective and motivating environment.

Method and Procedure: Two cohorts of six participants took part. Each was split into two teams to play language games where pictures were named with the help of team members and facilitation from a speech and language therapist. Facilitation was varied in three different cueing conditions: phonemic, gesture+phonemic and semantic+phonemic. Overall 180 words were practiced (90 nouns and 90 verbs). Therapy was delivered three days per week, for 6 weeks (for a total of 54 hours).

Outcomes & Results: Our intervention was equally effective across the three cueing conditions and for nouns and verbs. Gains were demonstrated in naming the pictures used in training, but also in the description of pictured scenes designed to elicit the same words. With these tasks, there were improvements of 25% and 18% from base-line accuracy, which compares well with gains reported in the literature using individually delivered speech and language therapy based on picture naming. Improvements were mostly maintained at both 4-7 weeks and 6-months post-therapy and were significant in all but the two most severely affected participants. There was some generalization of gains to narrative production, but not to other language tasks, nor to untreated words in picture naming. These positive language outcomes were combined with a high level of engagement and satisfaction (with participants stating a preference for games over standard therapy).

Conclusions: Our results support embedding theoretical and empirically-based techniques for aphasia rehabilitation within games with a strong social aspect, which may promote linguistic recovery in a way that is both time and cost efficient and engaging. Future research should explore more formally outcomes in terms of in increased well-being and reduced social isolation, as well as language proficiency.

High intensity, one-to-one rehabilitation therapy is effective in the treatment of post-stroke aphasia, but it can put strain on public health providers, as well as lead to high attrition. This study piloted a new group-delivered game-based intervention, designed to provide efficacious word-retrieval rehabilitation, in a cost-effective and motivating environment. Two cohorts of six participants took part. Each was split into two teams to play language games where pictures were named with the help of team members and facilitation from a speech and language therapist. Facilitation was varied in three different cueing conditions: phonemic, gesture+phonemic and semantic+phonemic. Overall 180 words were practiced (90 nouns and 90 verbs). Therapy was delivered three days per week, for 6 weeks (for a total of 54 hours). The intervention was equally effective across the three cueing conditions and for nouns and verbs. Gains were demonstrated in naming the same pictures used in training, but also in the description of pictured scenes designed to elicit the same words. With these tasks, there were significant improvements from base-line accuracy, respectively of 25% and 18%, which compare well with gains reported in the literature using individually delivered speech and language therapy based on picture naming. Improvements were mostly maintained at both 4-7 weeks and 6-months post-therapy and were significant in all but the two most severely affected participants. There was some generalization of gains to narrative production, but not to other language tasks, nor to untreated words in picture naming. These positive language outcomes were combined with a high level of engagement and satisfaction (with participants stating a preference for games over standard therapy). Our results support embedding theoretical and empirically-based techniques for aphasia rehabilitation within games with a strong social aspect, which may promote linguistic recovery in a way that is both time and cost efficient and engaging. Future research should quantify outcomes in terms of in increased well-being and reduced social isolation, as well as increased language proficiency.

Aphasia is often chronic and life-changing. It reduces quality of life and can hinder education, employment and community integration (e.g., Astrom, Adolfsson, & Asplund, 1993; Hilari, Needle & Harris, 2012; Hilari & Northcott, 2017). Speech and language therapists (SLTs) are uniquely placed to treat people with aphasia (PwA). There is good evidence that speech and language therapy is effective to ameliorate language difficulties, at least when it is delivered with the right intensity (meaning, here, a dose that is high enough in terms in number of hours of therapy; see Bhogal**,** Teasell, & Speechley, 2003; Brady, Kelly, Godwin, et al., 2016; Cherney, 2012; Denes et al., 1996; Hinckley & Carr, 2005). Nevertheless, PwA report unmet needs after they leave in-patient care (McKevitt et al., 2011). This situation is only predicted to worsen. Demand for healthcare is increasing as population grows and people live longer. Many national health services are already strained. To ameliorate this situation, we must devise new ways to deliver aphasia therapy so that it is both effective and cost effective.

Our study assessed the feasibility of a new intervention based on playing a social game with the expectation that this could be, at the same time, efficacious (being based on sound principles of language rehabilitation), fun, motivating, and cost effective, since games can be carried out simultaneously by a group of participants supervised by a single SLT, thus reducing demands on professional time. We focused on post-stroke word production difficulties including both difficulties in word retrieval (Broca's aphasia; anomia) and difficulties in phonological encoding (conduction aphasia; Wernicke's aphasia). Difficulties with word production are one of the most common, debilitating and long-lasting consequences of stroke aphasia (e.g., Goodglass & Wingfield, 1997) affecting a person's ability to communicate (Basso et al., 1990; Herbert, Hickin, Howard, Osborne, & Best, 2008). For this reason, rehabilitation therapies often focus on word production and use confrontation naming as a practicing tool (see Doesborgh et al., 2004; Nickels, 2002; Salter, Teasell, Bhogal, Zettler, Foley, 2009; for reviews see Albert, 2003; Basso, 2005; Bhogal et al., 2003; Wisenburn & Mahoney, 2009). Our study maintains a focus on naming. Studies have shown item specific and non-item specific generalization to connected speech (see Conroy, Sage & Lambon Ralph, 2009; Herbert et al, 2008; Rider, Harris-Wright, Marshall, & Page, 2008) supporting the usefulness of this approach to improve functional communication which is a priority for PwA (Rider et al., 2008).

1.1. Background and Rationale.

One way to reduce therapy costs is to treat patients in a group. Moreover, practicing language in a group offers potential additional advantages because interacting with peers could be less intimidating and more motivating than having face-to-face interactions with a proficient speaker. In addition, participating in a group may reduce social isolation which could be as debilitating as the language impairment itself (Parr, 2007). Aphasia groups are commonly used to help PWA in their pathway to recovery either as the sole form of intervention or as an adjunct to one-to-one SLT (see Elman, 2007a; 2007b). Aphasia groups, however, have been used only very sparingly to deliver structured interventions (see Lanyon, Rose &Worrall, 2013, for a review), with the exception of CIAT/CILT protocols, discussed later, where group size is limited (up to three patients,). Most aphasia groups aim either to provide education and support or to provide a conversational environment for less severe patients (e.g., see Rose, & Attard, 2015).

The social and emotional benefits of participating in a group are clear. Participating in a group normalizes experiences, allows socializing and encourages new friendships (Vickers, 2010), provides much-needed feelings of understanding and acceptance (Northcott et al., 2016; Ross, Winslow, & Marchant, 2006; Vickers, 2010) and reduces depression (Brumfitt & Sheeran, 1997). The language benefits of unstructured conversations, however, are less clear. Only a few studies have assessed these benefits. Two studies assessed gains on general linguistic measures and found positive effects, but results were weakened by possible confounding with spontaneous recovery (Wertz et al., 1981; Elman & Bernstein-Ellis, 1999). Other studies have assessed gains linked more specifically to what was practiced within the group. Drummond and Simmons (1995) examined quality of discourse (in terms of phonology, semantics, syntax and pragmatics) in four PwA while they practiced topics of conversation within group. They found gains in quantity of verbal output, but no improvement in any of the quality measures. Falconer & Antonucci (2012) combined semantic feature analysis with group-based conversation in four PwA and found gains in informativeness and/or efficiency of communication (see also Antonucci, 2009). Two further studies have specifically assessed benefits for word production. The results of these studies provided only weak evidence of benefits. Eales & Pring (1998) carried out a within-

subject study with four PwA. Target words were practiced first with individual therapy and then with group conversations using topics designed to elicit the target words. Performance was assessed at different points with picture naming. Performance improved mostly after individual therapy. It also improved after group conversations, but with no difference between the words practiced in conversation and control words. Nickels, McDonald and Mason (2016) also carried out a within-subject study with four PwA. Participants' lexical retrieval abilities were assessed with both picture naming and structured interviews designed to elicit the target words. Performance was compared for three matched sets of 30 words which were: a) untreated; b) treated with group conversations on associated topics; c) treated with group conversations + home-based confrontation naming exercises. Treated sets, but not untreated sets showed improvements, but only in picture naming. Moreover, gains were confounded by general trends for improvements which occurred both in treatment and no treatment phases of the study, weakening results.

Taken together the studies reviewed above showed limited evidence that non-structured conversion approaches are beneficial. Structured linguistic intervention (following a defined protocol) may work better, especially for patients with moderate to severe impairments who may find conversation too difficult (see also Lanyon, Rose & Worrall, 2013). Structured interventions, however, are mostly delivered one-to-one with the important exception of Constraint-Induced Protocols, which involve small groups of participants (known as CIAT --Constraint-Induced Aphasia Therapy, CILT --Constraint-Induced Language Therapy, or ILAT --Intensive Language Action Therapy; see Pulvermüller et al., 2001; Difrancesco, Pulvermüller & Mohr, 2012; for a review see Balardin & Miotto, 2009; Meinzer, Rodriguez, & Rothi, 2012; Zhang et al., 2017). These protocols share the following defining characteristics : 1) Treatment is delivered in small groups (up to three patients); 2) Practice is strictly focused on a verbal, spoken output with other forms of communication either not practiced or actively discouraged (constrained); 3) Treatment is intensive where intensity refers to the therapy being delivered both with a high-dose and in a compact way (massed rather than distributed practice); 4) Treatment is focused on word production (picture naming); 5) Treatment involves shaping, where word production is practiced repeatedly, with different carrier sentences, and different degrees of facilitation; 6) Naming is promoted in the context of social requests as part of a card game (Go Fish) where participants ask other participants for matching cards. Constraint-Induced Protocols

(from now on CIP) have received a lot of attention because studies have shown benefits for treated words and, occasionally, improvements on standardized tasks (e.g., Carpenter & Cherney, 2016; Pulvermüller et al., 2001; for a review Zhang et al., 2017; but also see for negative results Attard, Rose & Lanyon, 2012; Hameister, Nickels, Abel & Croot, 2017; Kurland, Stanek, Stokes, Li & Andrianopoulos, 2016; Nickels & Osborne, 2016). Which elements are responsible for the succes of CIP, however, remains unclear.

Actively discouraging alternative forms of communication such as gestures does not seem crucial. Relatively unconstrained versions of CIP (where gestures are not prevented) have also been found to be effective (Nickels & Osborne, 2016; Difrancesco et al., 2012; Stahl, Mohr, Dreyer, Lucchese & Pulvermuller, 2016) and as effective as constrained versions (Kurland, Pulvermüller, Silva, Burke & Andrianopoulos, 2012). Moreover, a positive impact of gesture on naming has been noted by some studies (Frick-Horbury & Guttentag, 1998; Morsella & Krauss, 2004; Rose, 2013). While intensity, in terms of therapy dose, may well be important (see Bhogal et al., 2003; Brady, et al., 2016; Hinckley & Carr, 2005; Hinckley & Craig, 1998; Denes et al., 1996), there is no evidence that massed practice is better than distributed practice. If anything, the opposite may be true (Cepeda, Pashler, Vul, Wixted & Rohrer, 2006; Dignam, Rodriguez &, Copland, 2016; Mozeiko, Coelho & Myers, 2016). Therefore, given that many word retrieval treatments are effective (see Boo & Rose, 2011; Boyle, 2004; Coelho, McHugh, & Boyle, 2000; Howard 2000; Maddy, Capilouto, & McComas, 2014; Rider, Wright, Marshall, & Page, 2008) and to a similar degree as CIP (see Zhang et al., 2017), one can ask what makes CIP a desirable form of therapy.

A recent study by Stahl et al. (2016) compared two forms of naming therapy delivered in small groups. One was ILAT, where participants ask for cards in the context of the game 'Go Fish' (the same game used by other CIP). Here, naming is carried out for the purpose of acquiring matching cards (when a matching card is acquired, the pair can be discarded; the player who is left without cards wins). The other was a traditional naming therapy, where participants were asked to name what was depicted on the cards. Eighteen PwA carried out both the ILAT protocol and the confrontation naming protocol in counterbalanced order. Results showed that the ILAT protocol delivered more improvements on subscales of the Aachen Aphasia Test (AAT). The authors interpreted this result as showing the importance of social interaction for therapy results. In particular, they stressed the importance of embedding naming in the context of social requests. Another

possible interpretation of these results, however, is that CIP involves playing a game which could be more motivating than carrying out naming exercises individually.

In our experimental investigation, we wanted to keep a number of elements used in CIP (as well as in other therapies) which we know are effective, such as a focus on spoken word naming, shaping with facilitation techniques based on cueing, and a high dose of therapy. Our protocol, however, also differed from CIP in important respects. We did not focus on speech acts involving requests. We focused on confrontation naming, but we embeded naming in the contex of a social game which allowed more participants to play at once and to play in teams, differentiating it from CIP. We believe that the potential for using social/team games in the treatment of PwA has not been sufficiently exploited. There is evidence that playing games results both in learning and improved mood (e.g., see, for dementia, Dartigues et al., 2013; for motor impairments Vanacken et al., 2010). Embedding language exercises in *team* games played in medium-sized groups may increase motivation and engagement which is a problem with intensive therapy (e.g., Brady et al., 2016) and may provide additional social and emotional benefits, while reducing costs. Finally, we wanted to assess the effect of cueing more systematically, given the importance of cueing facilitation techniques for rehabilitation (see Best et al., 2013; Salter et al., 2009).

There is strong evidence that *phonological* cueing helps with word retrieval, both in control and in aphasic speakers (see Kay & Ellis, 1987; Patterson, Purell, & Morton, 1983; Pease & Goodglass, 1978). It is not clear whether *semantic* cueing significantly helps retrieval at the point when a word is unavailable (see Meteyard & Bose, 2018). However, both naming therapies focused on phonological and semantic cueing have shown to be effective probably because both of them help to strengthen links between phonological and semantic representation in lexical networks (for phonological therapies see Hillis, 1993; 1998; Nickels, 2002; Raymer, Thompson, Jacobs & Le Grand, 1993; for semantic therapies see, Boyle 2004; Coelho, McHugh & Boyle 2000; Nickels, 2002; Raymer et al., 1993; see also Maddy, Capilouto, & McComas, 2014 for a review of the efficacy of semantic feature analysis to improve picture naming). There is also some evidence that practicing picture naming in association with gestures (observed or carried out) is effective, especially for PwA with lexical retrieval difficulties (Boo & Rose, 2011; Kroenke, Kraft, Regenbrecht, & Obrig, 2013; Marangolo et al., 2010; Rose, 2013; Rose & Douglas, 2008; Rose, Douglas, & Matyas, 2002) and that PwA can use gestures to self-cue while naming (Hanlon et al., 1990; Lanyon

& Rose, 2009). Gestures may help naming because of possible relashionships between lexical representations and associatated motor patterns (see embodied cognition; e.g., Jirak, Menz, Buccino, Borghi, & Binkofski, 2010; Pulvermüller, 2005). This may be particularly true for verbs (which are generally the target of gesture facilitation see Boo & Rose, 2011; Marangolo, Cipollari, Fiori,Razzano & Caltagirone, 2013), but it may also apply to concrete nouns which are often associated with actions.

While facilitation approaches are generally effective which one is more successful to improve naming is unclear. When phonological and semantic approaches have been compared, both have been found to be effective (Greenwald, Raymer, Richardson & Rothi, 1995; Stimley & Noll, 1991; Wambaugh, Linebaugh, Doyle, & Martinez, 2001; Wambaugh, 2003) although there is *some* evidence of longer-lasting effects and more generalization with semantic therapies (Howard, Patterson, Franklin, Orchard-Lisle, & Morton, 1985; Holland, Johns, & Woollams, 2018; Neumann, 2018; Lorenz & Ziegler, 2009; for a review see Wisenburn & Mahoney, 2009). Equally, when therapy using gesture has been compared with other approaches, similar efficacy has been reported (Boo & Rose, 2011; Raymer et al., 2007; Rose & Sussmilch, 2008). Comparing different types of cueing with our group-based game-based therapy can provide further evidence about the relative efficacy of different approaches.

In conclusion, our study wanted to assess the feasibility of a new mode of delivering SLT based on playing language games in teams (from now on, game therapy, GT), but incorporating rehabilitation techniques with a strong theoretical and empirical basis. We practiced picture naming and repetition combined with cueing, but in the context of a competitive game where participants worked/played in teams. This approach would be suitable for many patients with aphasia. Picture naming practices word retrieval and benefits participants with a clinical classification of anomia (see Howard, 1994; Maher & Raymer, 2004). Repetition practices phonological encoding and benefits participants who have difficulties in selecting and organizing phonemes for production (Wernicke's aphasia, conduction aphasia, jargon aphasia; see Galluzzi, Bureca, Guariglia & Romani, 2015; Nickels, 2002; Romani & Galluzzi, 2005).

We assessed feasibility in terms of positive outcomes achieved (with gains hopefully being comparable to those reached through one-to-one therapy) and acceptability to participants. More specifically, we assessed efficacy in terms of: a) treatment-specific

gains; b) gains maintained over time and c) gains obtained both in picture naming and in a narrative context, as evidence of generalization to functional communication. In a very preliminary way, we have also compared outcomes of GT with what is currently offered by the NHS and considered interactions with order of administration (ST before GT or vice-versa). We assess acceptability in terms of rate of attrition and responses to a satisfaction questionnaire. Finally, nested within the aim of proving the efficacy GT, we aimed to assess whether different cueing techniques (phonological, semantic or gestural) could be differentially effective. We hoped that a new team-game approach to SLT could be effective and acceptable while, at the same time, bearing the promise of reducing professional cost, and increasing well-being and engagement by making the therapy more fun.

2. Method

2.1. Participants

Twelve participants with stroke-induced aphasia were recruited from an outpatient neurorehabilitation unit (Moor Green Outpatient Brain Injury Unit) and two community services within Birmingham Community Healthcare NHS Foundation Trust. SLTs provided information to their patients and invited participation. Informed consent was obtained using an "aphasia-friendly" information sheet. Recruitment occurred in two phases, each aimed at recruiting a cohort of 6 participants; recruitment stopped as soon as this was achieved.

Inclusion criteria were: moderate to severe word finding difficulties, with performance in the Boston Naming test being < 50% correct and relatively well-preserved comprehension to allow coping with the demands of the game. Exclusion criteria were: a history of alcohol and /or substance abuse, developmental difficulties, and/or any other neurological, psychiatric or degenerative disease that could contribute to language or communication impairment. All participants were fluent English speakers before their stroke. They were either monolingual speakers or bilingual since early childhood with the exception of one participant (P5) who learned English in school in India, but reported to be already fluent in English when he arrived in the UK age 27. All participants were at least three months post onset.

All participants received some standard speech and language therapy (ST) during our study, as well as our experimental game therapy. All participants had also carried out some

SLT prior the beginning of our study. Participants from Cohort 1 (P1-6) received some additional ST *after* Game Therapy (GT). Participants from Cohort 2 (P7-12) received some additional ST after our initial assessment but *before* GT Participants from the two cohorts differed marginally by age and months post onset (cohort 1 included older and more chronic participants), but the two groups did not differ significantly in the amount of additional therapy received, education or baseline measures (see Table 1). Clinical classification was established through discussion with the referring SLT.

Insert Table 1 about here

Standard Speech and Language Therapy (ST) was delivered either at the neuro-rehabilitation outpatient unit (Moor Green) or in the community by NHS Speech and Language Therapists. It was flexibly adapted to the needs of the patient and included a mixture of therapy approaches according to the individual's therapy goals: impairment-based (e.g., picture-naming), functional (e.g., use of a communication book), activity-directed (e.g., practising phone calls) or participation-based (e.g., conversation groups). There was no overlap with the materials used in the game therapy. On average, participants carried out 51 hours of Standard Therapy over four months (15.4 weeks), but there was a lot of variability with patients attending for 7 to 33 weeks and receiving between 7 and 101 hours of ST. This variability was due to different offerings by different NHS services and variable patient needs/goals.

Experimental Game therapy (GT) was carried out at the outpatient neuro-rehabilitation unit. Each game was delivered by a senior SLT (Louise Lander, a member of the research team), assisted by either a trained psychology student or another SLT. Overall, participants carried out 54 hours of Game Therapy over a total period of eight weeks (three periods of therapy with assessment weeks in between).

2.2. Game Therapy Protocol

Each cohort of six participants was split into two teams of three. The purpose of the game was to gain points for one's team by naming pictures. The participant whose turn it

was picked a card from a set and, without showing it to the other players, tried to name it. He/she received facilitating cues if necessary. The other members of the team could also accrue points for the team by helping the participant on call. Once the player on call had produced the target, each member on his/her team would repeat it. This ensured shaping and widened participation on each trial. At the end of each round, the card was placed face down at the bottom of the pile and play passed to the other team. Different numbers of points were gained depending on ease of naming and degree of help by the facilitator. At the end of each (one hour) session, points were tallied and the winning team declared. Participants were encouraged to change teams after each session to ensure that each individual had the opportunity to interact with and against all other individuals. As well as negating potential differences in outcomes due to differing one-to-one interactions, this strategy also helped to maintain interest in the games.

The facilitation techniques used by the SLT during the games, were systematically varied by contrasting phonological, semantic and gestural cueing techniques. These techniques were used with matched sets of nouns and verbs at different phases of therapy, from now on: Game P using phonological cueing, Game PG using phonological and gestural cueing and Game PS using phonological and semantic cueing.

Game P. If the participant could not name the target, the facilitator provided phonemic or syllabic cues, or a model for repetition, as required. For example, for the target word 'umbrella', the following hierarchy of prompts would be used: "what sound does it begin with?" -> "it begins with *uh*" -> "it starts *um*" -> "it's an umbrella".

Game PG. Participants were encouraged to gesture appropriately whilst trying to produce the target. If naming was unsuccessful, the facilitator produced gestural, as well as phonemic cues. For example, for the target word 'umbrella', the following hierarchy of prompts would be used: "can you show me what you do with it?" -> therapist gestures opening an umbrella -> phonemic cueing hierarchy in tandem with gestures.

Game PS. Participants were encouraged to talk around the target by producing similar words, describing its semantic features, or producing a phrase containing the target. If naming was unsuccessful, the facilitator provided semantic, as well as phonemic cues. For example, for the target word 'umbrella', the following hierarchy of prompts would be used: "what do you use it for?" -> "what does it look like?" -> "you need it when it rains" -> "you open it" -> "It's raining, you open your…" -> phonemic cueing hierarchy.

Each game condition was played for three hours per day, split into three separate one hour- sessions, three times a week, for two weeks, totalling 18 hours for each game condition (for a total of 54 hours over six weeks across all game conditions).

For Cohort 1, each item was presented either 16 or 17 times during Games P and PG, and 12 times during Game PS. For Cohort 2, each item was presented 19 or 20 times during Game P, 25 times during Game PG and 15 times during Game PS. The lower number of presentations during the first cohort's therapy reflects the more severely dyspraxic participants in this cohort. These participants often struggled with articulation, taking longer to produce the targets. Fewer presentations during Game PS for both cohorts was due to the additional time needed for semantic elaboration.

2.3. Game Therapy Materials

A set of 60 words was trained in each game condition (30 nouns and 30 verbs) for a total of 180 words. This was deemed acceptable to achieve a reasonable 'therapy dose' for each target, whilst also ensuring that participants remained interested in the protocol and functional gains could be achieved (Cherney, 2012). Set A was trained in Game P, Set B was trained in Game PG and Set C was trained in Game PS. Words in the three sets were carefully matched for frequency, age of acquisition, length and phonological complexity (see Appendix A). Picturable, easy to name verbs are harder to find than equivalent nouns. Thus, across the three sets of words (A, B and C), verbs had significantly higher frequency than nouns and they were shorter (see Appendix A). We included nouns and verbs in our therapy materials because improvement with both types of stimuli are important if functional gains are to be reflected in connected speech and narrative production.

Pictures were black and white line drawings mostly taken from the Object and Action Naming Battery (Druks, 2000) and the International Picture-Naming Project Database (Szekely et al., 2004). A small number were also taken from clipart sources online. All pictures were presented on 8cm^2 white cards. Assessment of trained words was through naming the same pictures used in therapy and through descriptions of pictured scenes which we had previously demonstrated elicit the trained words in control participants.

There were three scenes for each set of words, each designed to elicit 20 target words (see Appendix B). Word set A was probed by scenes depicting: 1. A house interior (with kitchen, study and living room); 2. A beach; 3. A street. Word set B was probed by

scenes depicting: 1. A garden; 2. The interior of a café; 3. A fair at a Castle. Word set C was probed by scenes depicting: 1. Another house interior (2 bedrooms, bathroom and room to be decorated); 2. A countryside scene; 3. A concert. All scenes were black and white drawings; each was presented on an A3 sheet.

The scenes were given to a group of 9 younger control participants and a group of 17 older control participants, all of whom were asked to describe what was happening. The number of target words produced was counted for each participant and each scene. The control results indicated that the pictured scenes were successful in eliciting the production of target words. For the group of older adult controls (N=17), 35.9 (SD 6.9) targets were elicited for Set A, 41.1 (SD 9.1) for Set B and 34.4 (SD 7.3) for Set C (maximum=60 for each set). For the group of younger adult controls (N=9), the figures were 35.4 (SD 8.2) for Set A, 40.1 (SD 9.7) for Set B and 35.8 (SD 8.9) for Set C.

2.4. Design

When participants are few (like in our case where we have 12 participants) and may differ substantially on variables which affect therapy outcome --such as age, severity of lesion, time post-onset, education, etc—between-group comparisons lack power. A better option is offered by multiple baseline designs where the same participants are assessed multiple times with matched materials which have been either treated or untreated (see Nickels, 2002). We used a multiple baseline design in our study. Following a multiple baseline design, we compared performance with trained and untrained word sets at the same point in time as well performance on the same word sets at different times (before and after training).

Our design did not include a direct comparison with an alternative treatment since our main aim was to assess whether our intervention was viable, effective and well-liked by PWA. However, we did want to gather some preliminary results on the relative improvements offered by our group game therapy (from now on GT), and by standard therapy as it is currently offered within the NHS (from now on ST) and on the possible interactions between these treatments based on administration order. Thus, across two groups, we counterbalanced participation in ST, with one group having some ST before GT and another group having GT first and ST afterwards. This allowed us some comparison of general language gains after the two approaches as well as an evaluation of whether GT is

more beneficial when administered before or after some improvements have already been obtained with ST. However, we should note from the start that results can only be considered very preliminary, not only because of the size of our sample, but also because the ST received by our participants was very variable in content, frequency and intensity, mirroring the variability of therapy offered within the NHS which depends on the goals of the individuals, but also on the practice and resources of referring trusts.

A schematic schedule reflecting our design is shown in Table 2. Across times we carried out the following assessments; some were more comprehensive, others more limited (see later for more details) :

Time 1. For cohort 2: Baseline: Comprehensive assessment before ST;

Time 2. For cohort 2: Comprehensive assessment after ST;
for cohort 1: Baseline: Comprehensive assessment before GT;

Time 3. After Game P (both cohorts), word set A (trained) and B (untrained), to compare trained and untrained word sets and performance before and after therapy for set A;

Time 4. After Game PG (both cohorts), word set B (trained) and C (untrained) to compare trained and untrained sets and performance before and after therapy for set B;

Time 5. After Game PS (both cohorts): Comprehensive assessment: word set C (trained) to assess gains compared to baseline; word sets A and B to assess short-term maintenance; language battery to assess general gains;

Time 6. For cohort 1: After ST (and 5-6 months after GT); Comprehensive assessment: all three sets to assess long term maintenance of GT gains; general language assessment to assess any further gains provided by ST.

Insert Table 2 about here

2.5 Assessment

We assessed gains in production of both nouns and verbs with the same materials used in training (picture naming), but also with descriptions of *pictured scenes* which we had previously demonstrated elicited the words used in therapy in control speakers. In addition, we assessed gains in an unrelated narrative task (recount of the Cinderella story) with a number of measures (see later). This will demonstrate whether gains extended beyond the

narrow conditions used in therapy. Finally, we assessed possible improvements in standardized tasks such as the Comprehensive Aphasia Test (CAT) and the Boston Naming Test (BNT) and we probed satisfaction with our protocol using a brief questionnaire and a focus group. More or less comprehensive assessments were carried out at different phases in the therapy.

Limited Assessment. Limited assessments were conducted after each round of game therapy. Potential improvements after each specific game were assessed though production of target words in *picture naming* and *scene description tasks*.

Picture naming. Participants were asked to name the same pictures presented in the therapy, but this time presented in a randomized order on a computer screen. There were no time limits for response. Responses were transcribed and assigned 1 point if correct, 0 points if incorrect, and 0.5 points if produced correctly but after an appreciable delay (more than 5 sec as per CAT) and/or after a self-correction.

Scene descriptions. Participants were presented with each scene in turn and asked "What is happening here?" and, if a particular area needed prompting, "What about here?". Descriptions were recorded and then transcribed verbatim, including hesitations, false starts, fillers (umm..). The number of words trained in therapy which were produced correctly were counted. In addition, the quality of the narrative was scored using total number of words produced, words produced per minute, % of CIU (correct information units), and % of errors (syntactic, morphological, phonological and semantic). The same method was used for the Cinderella Story (described below).

Comprehensive Assessment. A more complete assessment was carried out at three points in time: For cohort 1: before GT, after GT, after ST; for cohort 2: before ST, after ST, after GT. Besides picture naming and scene descriptions, we administered:

The *Boston Naming Test* (BNT; Kaplan, Goodglass, & Weintraub, 1983) this is a standardized measure of picture naming, making it an effective tool for identifying any generalisation of word-retrieval gains to items not directly targeted during the game therapy protocol.

The *Comprehensive Aphasia Test* (CAT; Swinburn, Porter, & Howard, 2004). *CAT* provides a comprehensive assessment of language ability, including 27 language and cognition subtests probing language semantics (semantic memory, word fluency, visual recognition and object use with gestures), repetition (of words, nonwords, digits strings and

sentences), comprehension (of written and spoken words using sentences and paragraphs), spoken production, reading (words, complex words, function words and nonwords) and writing (copying, picture naming, writing to dictation, picture description). We used all but one subtest of the language battery. We excluded CAT 17 (naming objects) since naming was evaluated with the BNT. We calculated an overall standardized score substituting the participant mean for CAT 17. The CAT overall score has a mean of 50 and SD of 10 based on the performance of a large population of PwA. Baseline language assessments with the CAT were used by a trained SLT to classify aphasia type (see Table 1).

The *Cinderella Story Retell* (Saffran, Berndt & Schwartz, 1989). This is a common task used with people with aphasia to probe narrative production. Participants were asked to retell the commonly known story of Cinderella. A picture book with text blocked out was provided prior to the retell task to remind participants of the story. This task provides a way to assess generalization of therapy gains to connected speech (Conroy et al., 2009; Saffran et al., 1989). Narratives were recorded and transcribed verbatim, including hesitations, false starts, and fillers (umm..). We scored the total number of words produced (excluding false starts and fillers), word rate per minute, percentage of meaningful words produced over total number of words (or rate of CIU, correct information units), and percentage of syntactic, morphological, phonological, and semantic errors out of total words produced (see Marini, Andreetta, del Tin & Carlomagno, 2011 and Nicholas & Brookshire, 1993).

Finally, we administered the Disability Questionnaire from the CAT which assesses the impact of the impairment on an individual's life from that individual's perspective with questions such as "what is it like talking to the person closest to you?" and "does it make you feel frustrated?" Questions are answered using a rating scale.

End of Therapy. At the end of the therapy program, participants were invited to provide feedback through a focus group involving five participants and an aphasia-friendly questionnaire asking 12 questions regarding issues such as the suitability of the protocol, whether they enjoyed the protocol, their perceived improvements, and whether they found the therapy tiring.

2.6. Ethical Approval

This study received ethical approval from the NHS Health Research Authority: Coventry and Warwick NRES Committee, REC Reference 15/WM/0210.

3. Results

3.1. Effects of Game Therapy on Trained Words

These effects were assessed for picture naming and scene description.

Overall analyses. Figure 1 shows performance by point in time and word-set (trained vs untrained); results for type of word (nouns and verbs) are collapsed. There are clear interactions between word-set and time with steep improvements in performance after a word set has received training, but not at other times. A number of planned comparisons were run to assess significance of results.

First of all, to assess the overall effect of GT, we carried out within-subjects ANOVAs with rate correct in either picture naming or scene description as the dependent variable and *Time* as the independent variable, contrasting Time 2 (T2, before any GT) with Time 5 (T5, immediately after completion of all GT). Performance was significantly better *after* therapy both for picture naming and scene descriptions (picture naming: $F(1,11)=30.1$, $p<.001$, $\eta p^2=.73$; scene description: $F(1,11)=20.2$, $p=.002$, $\eta p^2=.69$).

Secondly, to establish whether improvements in a word set only occurred after treatment, we ran separate ANOVAs with *Word-set* and *Time* as within-subjects factors. We compared sets A and B at T2 (both untrained) and T3 (A trained, B untrained). We found significant interactions of *Word-set X Time* (picture naming: $F(1,11)=33.9$, $p<.001$, $\eta p^2=.755$; scene description: $F(1,9)=7.4$, $p=.02$, $\eta p^2=.45$). Post-hoc analyses showed that the two sets did not differ at baseline (picture naming: $F(1,11)=.92$, $p=.49$, scene description: $F(1,9)=.51$, $p=.49$) but differed significantly at T3 when only set A was trained (picture naming: $F(1,11)=27.2$, $p<.001$, $\eta p^2=.71$; scene description: $F(1,9)=9.8$, $p=.01$, $\eta p^2=.52$). We also compared and sets B and C at T4 (only set B trained) and T5 (both sets trained). Here, as well, we found a significant interaction of *Word-set X Time* (picture naming: $F(1,11)=17.432$, $p=.002$, $\eta p^2=.613$; scene description: $F(1,9)=5.125$, $p=.05$, $\eta p^2=.363$. The two sets differed at T4 when only set B was trained (picture naming: $F(1,11)=20.7$, $p=.001$, $\eta p^2=.65$; scene description: $F(1,9)=22.6$, $p=.001$, $\eta p^2=.71$), but became more equal at T5, when set C had also been trained (picture naming: $F(1,11)=2.4$, $p=.15$, scene description: $F(1,9)=9.9$, $p=.01$). Finally, planned comparisons showed that each word set improved significantly after training: set A between T2 and T3 (picture naming: $F(1,11)=35.2$, $p<.001$, $\eta p^2=.76$; scene description: $F(1,9)=19.0$, $p=.002$, $\eta p^2=.68$), set B between T3 and T4 (picture naming:

F(1,11)=29.1, p<.001, ηp^2=.73; scene description: F(1,9=30.3, p<.001, ηp^2=.77) and set C between T4 and T5 (picture naming: F(1,11)=17.3, p=.002, ηp^2=.612; scene description: F(1,9)=4.7, p=.058; ηp^2=.34). There were no other significant improvements at any other time. These results provide compelling evidence that improvements were linked to training rather than to general practice, increased motivation or spontaneous recovery (see also Figure 2).

Insert Figure 1 about here

Modulation of outcomes by type of therapy. Effects of therapy on picture naming according to type of game (with phonological cues, phonological + gestural cues, or phonological + semantic cues) and type of word (nouns or verbs) and are shown in Figures 2 and 3 respectively. Effects were statistically analysed with within-subjects ANOVAs containing three within-subjects factors: *Word-class* (nouns vs verbs), *Game-type* (P, PG, PS), and *Therapy-phase* (before GT vs. after game therapy; T2 vs T5). There was no main effect of *Game type* (F(1.33,14.66)= 0.56, p=.52) and no interaction of *Game-type X Therapy-phase* (F(2,22)=0.292, p=.75). The effects of therapy were the same regardless of the type of cueing strategy used in the game. There was also no significant main effect of *Word-class,* (F(1,11)=26.266, p=.18) with similar gains for nouns and verbs, and no interactions: *Word-class X Therapy-phase*, (F(1,11)=0.002, p=.99) or *Game-type X Word-class X Therapy-phase* (F(1,11)=.62, p=.56). One might expect Game PG (stressing gestures) to be particularly beneficial for verbs. However, actions are also closely associated to most concrete nouns. We did not systematically contrast strength and type of association with gestures for nouns and verbs. Instead, we wanted to assess generalized gains across types of words; establishing possible differences between nouns and verbs was beyond the remit of our study.

Insert Figure 2 and 3 about here

Maintenance. We compared therapy gains at three points in time: Immediately after GT (for all three sets), 4 and 7 weeks after completion (for sets B and A respectively, short-term maintenance), and 5-6 months after completion (all three sets for cohort 1; long-term maintenance). Gains in % correct for these three points in time, were respectively: for picture naming: 25%, 18% and 16% and for scene descriptions: 17% 18%, and 13%. All these gains were significant when compared to baseline (picture naming: immediately after GT: $F(1,11)=30.1$, $p<.001$, $\eta p^2=.73$; 4-7 weeks post GT: $F(1,11)=25.6$, $p<.001$, $\eta p^2=.70$; 6 months post-GT: $F(1,15)=9.4$, $p=.03$, $\eta p^2=.65$; scene description: immediately after GT: $F(1,9)=20.2$, $p=.002$, $\eta p^2=.692$; 4-7 weeks post GT: $F(1,9)=12.451$, $p=.006$, $\eta p^2=.58$; 6 months post-GT: $F(1,5)=12.0$, $p=.02$, $\eta p^2=.71$). When performance was compared immediately after GT versus 4-7 weeks later (for set A, performance at T3 vs T5 and, for set B, performance at T4 vs T5) there was a small decrease in picture naming ($F(1,11)=4.9$, $p=.05$, $\eta p^2=.31$), but not in the scene description ($F(1,9)=.07$, $p=0.8$). When performance was compared immediately after GT and 6 months later, there were no significant decreases, but this may be due to lack of power, since fewer participants were tested at this point (only cohort 1; picture naming: $F(1,11)=-2.97$, $p=.15$; scene description: $F(1,5)=.43$, $p=.54$).

***Results by participant*.** Individual participant results are shown in Figure 4. Different panels show outcomes for trained words immediately after therapy, 4-7 weeks after therapy (short-term maintenance) and 5-6 months after therapy (long-term maintenance). Immediately after GT, gains were significant in 9/12 patients in picture naming and in 9/10 patients in scene description. No significant improvements were seen in participants P8 and P9 who had very severe impairments with a floor effect at baseline. A third patient, P1, showed no significant effect in picture naming, but a significant effect in the scene descriptions. P9 and P10 were not tested with scene description because they were unable to complete the task. After 4-7 weeks, 9/12 participants in picture naming and 8/10 participants in scene description showed significant improvement when compared to baseline. After 6 months, 5/6 participants showed significant gains when compared to baseline. Only P4 showed no difference.

Insert Figure 4 about here

Effect of experimental and demographic variables (Cohort, age and time post-onset). To examine a possible effect of cohort we carried out mixed ANOVAs with number of words produced correctly in picture naming and scene descriptions as dependent variables, *Cohort* (cohort 1 with ST after GT vs cohort 2 with ST before GT) as a between-subjects factor and *Therapy-phase* (T2/before GT vs T5/after GT) as a within-subjects factor. There was no significant main effect of Cohort (picture naming: F(1,10)=.117, p=.74; scene description: F(1,8)=0.08, p=.78) and no significant interaction between Cohort and Therapy-phase (picture naming: F(1,10)=.006, p=.94; scene description: F(1,8)=0.5, p=.50). Further studies with well-matched cohorts are needed to properly assess the advantages of delivering ST and GT in different orders. Most importantly, significant improvements were shown across patients. In fact, there was no significant correlation between degree of improvement immediately after GT and either age (Pearson r=.09, p=.79) or months post onset (Pearson r=.32, p=031), although these correlations are based on small samples.

3.2 Other Effects of Game Therapy.

General effects on language functions. Generalization of gains from GT were assessed by comparing performance before and after GT (T2 vs T5) on the CAT, the BNT, untrained words in picture naming and on measures of narrative production. Results are shown in Figure 5. Narrative measures were collapsed across the Cinderella Story and the Scene Descriptions.

Insert Figure 5 about here

There were no significant group differences with the CAT (even considering individual tests) or the naming of untreated words. There was, however, a significant improvement in the BNT when a one-tailed t-test was performed, (t(11)=1.75, p=.05) and significant improvements in measures of narrative production in terms of overall number of words produced (t(9)=2.68, p=.03) and % of CIU (t(9)=2.69, p=.03). Error rate and rate of words per minute did not change (error rate before GT: mean 53.2%, SD 34.6%; after GT: mean 56.1%, SD 28.8%; t(9)=0.466, p=.65; word rate: before GT: mean 136.6, SD 156.7;

after GT: mean 158.9, SD 141.7; t(9)=1.130, p=.29). The presence of significant generalizations from picture naming to connected speech, at the group level, is encouraging.

At the individual level, only P11 showed significant improvement across tasks and measures. He showed gains in producing untrained words in picture naming and in the scene descriptions, as well as improvements in the BNT and in % of CIU in narrative production. Since he was the participant with the most recent stroke (12 weeks post stroke when he entered our study), gains could have been boosted by spontaneous recovery. However, P11 did ST first for three months, and showed no improvement within that period. This lack of improvement does not necessarily indicate that ST was ineffective as he received very little of it (9 hours). Moreover, P11 was initially very anxious and distressed by his condition and this may have affected the assessments. However, the contrast between the lack of gains within the first three months and the significant gains obtained with GT later on indicates that these gains were not simply due to spontaneous recovery. P11 really enjoyed the games and relaxed during the course of GT, therefore, taking full advantage of the practice provided.

Effects of ST vs GT. The effects of ST are shown in Figure 6 which reports performance on experimental words in picture naming and scene description and performance on the BNT and the CAT, before and after ST. Here, we wanted to assess any positive effect of ST, and compare benefits on standardized tasks like the BNT and the CAT with those obtained with GT. Note, however, that our results cannot offer more than a rough indication of outcomes since type and amount of ST was so variable from one participant to another.

Insert Figure 6 about here

Results were analysed using mixed ANOVAs with language performance in different tasks as the dependent variable; *Therapy-phase* (before vs. after ST) as a within-subjects factor; and *Cohort* (1 vs 2) as a between-subjects factor.

There was no main effect of *Therapy-phase* (F(1,11)=.02, p=.9) on production of experimental words, and only a marginal *Therapy-phase X Cohort* interaction (F(1,10)=3.29,

p=.10). There were some marginal gains with Cohort 2, who had ST before GT (F(1,5)=4.727, p=.08), reflecting a general improvement after therapy, but a non-significant decrement with Cohort 1, who had ST after GT (F(1,5)=1.114, p=.34) reflecting some loss during long-term maintenance. With the CAT scores, there was a significant main effect of *Therapy* (F(1,10)=7.17 p=.02, ηp^2=.418) and no interaction with Cohort (F(1,10)=1.82, p=.21). This result contrasts with the lack of any improvement after GT and suggests more generalized language improvement after ST (see General discussion). With the BNT, there was no main effect of *Therapy* (F(1,10)=0.19, p=.68) and no interaction with Cohort (F(1,10)=0.19, p=.67).

Satisfaction with therapy. Our therapy and therapy schedule were very well accepted by participants. Overall 93.6 % (SD=9.9) of scheduled game therapy hours were attended and there was only a minimal loss of attendance over time for the three games (hours attended for Game P= 99.3; Game PG=96.8; Game PS=84.7).

The disability questionnaire of the CAT did not show any differences in self-assessment of disability after either GT or ST. However, the therapy satisfaction questionnaire administered at the end of the protocol reported very positive feedback for GT with participants scores averaging 4.76/5 indicating very strong satisfaction with the therapy. All participants either agreed or strongly agreed that GT increased their confidence and was enjoyable. All participants reported an improvement in their talking and 9/12 expressed a preference for GT over ST, with the remaining three participants not expressing a preference one way or the other.

Participants in the focus groups highlighted how the playing the games was useful, enjoyable and helpful. They also noted how it was good to meet other people with language difficulties and the team work and mutual support during the games made this type of intervention preferable to therapy delivered one-to-one.

4. General Discussion

The aim of our study was to pilot a therapy intervention for post-stroke aphasia which combined, in a novel way, ingredients that we know are effective in therapy rehabilitation. We strived to devise an intervention which:

1. Allowed high-intensity practice, but at reduced professional costs and maintaining high engagement;

2. Focused on treating language impairments, but also improved social interaction and confidence;
3. Incorporated evidence-based rehabilitation techniques;
4. Was suitable for most PwA experiencing moderate to severe difficulties.

Our solution was an intervention focused on word retrieval which used tasks (picture naming and repetition) and cueing techniques (phonological, gestural, semantic) that are of proven efficacy in aphasia rehabilitation but incorporating them within the setting of a team game. This aspect of the therapy was crucial in fulfilling many of the characteristics that we wanted to achieve. It allowed lower costs, since a single SLT could supervise therapy for several PwA at the same time (six in our case). It allowed the therapy to be more enjoyable than in traditional approaches. This, in turn, maintained high motivation throughout the intervention which is especially important in the case of prolonged and high-intensity therapy. Finally, it addressed the need to increase social support and social interaction. Participants playing in teams created more excitement and increased cohesion as participants worked together towards a common goal.

Our approach is not the first attempt to deliver aphasia therapy in the context of a game. CIAT/CILT are popular protocols which adopt a game approach and show benefits (e.g. Zhang et al., 2017). Our intervention, however, has novel aspects. It stresses a social game aspect more than CIAT/CILT by allowing more participants to play at once, split into teams, thus increasing social and motivating aspects of the game. Additionally, it systematically incorporates cueing techniques which have proven efficacy and are commonly used by SLTs in face-to face therapy. These cueing techniques should not only facilitate retrieval, but also strengthen links to phonological representations (through phonological cueing) and semantic representations (though both semantic cues and gestures).

Our results are encouraging. Our intervention was very well tolerated with high rates of attendance. Our intervention was also enjoyed by all participants who often preferred it to the one-to-one standard therapy. Language gains were significant, wide-spread across participants, maintained over time and demonstrable across different tasks, suggesting benefits to functional communication. Gains immediately after Game Therapy (GT) were on average 25% in picture naming and 17% in the scene descriptions, which is close to, or above the level of 20% proposed to be clinically relevant by Ramsberger & Marie (2007).

Across the two tasks, gains were maintained long-term (six months after therapy) with, on average, 14.5% improvement from baseline. All participants, with the exception of two, showed significant gains considering both picture naming and scene descriptions together (10/12 participants). The two participants showing no improvement had very severe impairments, with a floor effect at baseline. They showed no improvement in spite of good engagement with the intervention. These participants may either need more time to show benefits or may not have enough neurological resources left to support recovery (see also Salter et al., 2009 for less or slower recovery in global aphasia).

Importantly, our game intervention produced significant gains with materials other than those directly used in therapy. The trained words were better used in connected speech when our participants were asked to describe pictured scenes constructed to elicit the trained words. Moreover, there were gains in narrative production in terms of both overall number of words produced and rate of meaningful words produced (Correct Information Units or CIU). These behavioural gains align with self-perceived improvements in talking. Our results are consistent with previous studies which have demonstrated significant benefits of practicing picture naming on functional communication (Conroy et al., 2009) and a strong association between the ability to produce words in picture naming and in connected speech (Herbert et al., 2008). In contrast, we found no gains in picture naming for untreated (experimental) words and only marginal gains in the Boston Naming Test. Typically, picture naming therapies do not produce gains in these conditions (e.g., see Best et al., 2013; Nickels et al., 2002; Raymer et al., 2007). Gains for untreated words may be more difficult to demonstrate in conditions where production is very constrained with no leeway in the choice of words. Finally, our game therapy produced no gains in the CAT. This is not surprising. We trained word production and we expected gains to be selective in this domain. In contrast (and pleasingly), gains on the CAT were seen after standard Therapy (ST) where SLTs worked to improve their clients' communication across domains. We are currently developing more articulated group game-based approaches to train more integrated aspects of communication using games where participants practice not only picture naming, but also requests in everyday situations (e.g., at the café, at the doctor, in the post-office etc.) which should booster gains in functional communication.

Our results compare well with gains reported in the literature for other forms of picture naming therapies treating aphasic participants singly or in pairs. We searched the literature

using, in combination, the following key words: aphas* or anomi* AND therap* AND naming or 'word retrieval' or constraint. We reviewed 19 studies and 22 therapy comparisons which reported the number/percentage of words gained after therapy as well as crucial treatment parameters such as number of hours and duration of treatment. Sixteen studies involved a one-to-one intervention, six involved treating participants in pairs and one involved both a one-to-one and a group intervention. On average, studies treated a limited number of participants (N=5.5 per study; SD=4; overall N=122), therapy involved 15 hours (SD=9), lasted on average 25.5 days (SD=15) and treated 57 words (SD=29). Treated words showed a 31% increase in number correct (SD=15), with, on average, 17 words gained after therapy (SD=11). Our study involved more participants (N=12), more therapy hours (N=54), lasted longer (42 days), and treated many more words (N=180); treated words showed a 25% increase in number correct, with 45 words gained after therapy.

To compare the efficacy of different forms of therapy is not straightforward, but two criteria are relevant: number of words gained and effort (number of hours of therapy). Thus, a rough measure of therapy efficacy may be the number of words gained per hour of therapy. With this measure, our reviewed studies returned 1.7 words gained per hour of therapy compared to 0.8 words in our case. Our measure is lower. However, we treated a much larger number of words than most studies since this is important to improve functional communication. It is likely that gains are harder to achieve the larger the number of words treated. Moreover, while for one-to-one therapy the hours engaged by the client and the therapist coincide, this is not the case for group therapy where a single therapist is treating simultaneously several clients (6 in our case; thus, a SLT would spend 1/6 of the time required for one-to-one treatment). Therefore, there are potential cost savings with a group approach. Finally, the enjoyment and social interactions offered by social games, may well produce emotional gains not elicited by one-to-one approaches. Future studies should compare more directly forms of group-game-therapy with matched forms of picture naming therapy delivered one-to-one in terms of language gains, satisfaction with the intervention and emotional gains.

We found no difference in efficacy depending on the type of facilitation cues used during the games. This result is consistent with others from the literature showing similar benefits across types of facilitation techniques (Greenwald et al, 1995; Holland et al., 2018; Neumann, 2018; Stimley & Noll, 1991; Wambaugh et al., 2001; Wambaugh, 2003). This

does not mean that all types of facilitation are equally effective for PwA with different kinds of impairment (although relationships are not always transparent; see Boo & Rose 2011; Kroenke et al., 2013; Lorenz & Ziegler, 2009). In a mixed group, however, it is not surprising to see no differences in average benefits. Our study did not have the power to differentiate between types of impairment. When team language games are applied to a clinical setting, we would favour an inclusive approach where people with different types of impairment are treated together, but facilitation is used flexibly by the game-leader depending on the individual participant. This would be consistent with recent studies which have combined different types of cueing techniques during therapy with positive outcomes (Carragher, Conroy, Sage, & Wilkinson, 2012; Drew & Thompson 1999; Hashimoto 2012; Le Dorze, Boulay, Gaudreau & Brassard 1994; Rose, Attard, Mok, Lanyon, & Foster 2013; for evidence that using multiple cueing techniques results in increased gains see also Greenwald et al., 1995; Lorenz & Ziegler, 2009; Wambaugh, 2003; Wambaugh et al., 2001; Rose, Raymer, Lanyon & Attard, 2013).

Conclusions. We found positive outcomes for a new game-based, group rehabilitation intervention targeting word production difficulties in individuals with post-stroke aphasia. Our results suggest that interventions like ours, which combine theoretically and empirically motivated techniques with the social and motivating aspect of a game are a positive way to supplement one-to-one therapy delivered in resource-stretched national health systems. We specifically targeted word production difficulties, but there is no reason why a similar approach based on social/team games could not be extended to other aspects of language – sentence production, for example-- and, thus, become appropriate for PwA with a wider set of needs. We are not advocating that interventions of the type assessed here should substitute for one-to-one therapy delivered by professional SLTs. However, they can be a valuable means of increasing practice, allowing patients to work in areas of special difficulty, to consolidate gains and to enjoy social interactions in a safe and supportive environment.

References

Albert, M. (2003). Aphasia therapy works. *Stroke 34*(4), 992-993.

Astrom, M., Adolfsson, R. and Asplund, K., (1993). Major depression in stroke patients. A 3-year longitudinal study. *Stroke, 24*, 976-982.

Antonucci, S. M. (2009). Use of semantic feature analysis in group aphasia treatment. *Aphasiology, 23*(7-8), 854-866.

Attard, M., Rose, M., & Lanyon, L. (2012). The comparative effects of multi-modality aphasia therapy and constraint-induced aphasia therapy for severe chronic Broca's aphasia: An in-depth pilot study. *Aphasiology, 27*(1), 80-111.

Balardin, J. B., & Miotto, E. C. (2009). A review of Constraint-Induced Therapy applied to aphasia rehabilitation in stroke patients. *Dementia & Neuropsychologia, 3(4),* 275-282.

Basso, A. (2005). How intensive/prolonged should an intensive/prolonged treatment be?. *Aphasiology, 19*(10-11), 975-984.

Basso, A., Razzano, C., Faglioni, P., & Zanobio, M. E. (1990). Confrontation naming, picture description and action naming in aphasic patients. *Aphasiology, 4*(2), 185-195.

Best, W., Greenwood, A., Grassly, J., Herbert, R., Hickin, J., & Howard, D. (2013). Aphasia rehabilitation: Does generalisation from anomia therapy occur and is it predictable? A case series study. *Cortex, 49*(9), 2345-2357.

Bhogal, S. K., Teasell, R., & Speechley, M. (2003). Intensity of aphasia therapy, impact on recovery. *Stroke, 34*(4), 987-992.

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

Boyle, M. (2004). Semantic feature analysis treatment for anomia in two fluent aphasia syndromes. *American Journal of Speech-Language Pathology, 13*(3), 236-249.

Brady, M. C., Kelly, H., Godwin, J., Enderby, P., & Campbell, P. (2016). Speech and language therapy for aphasia following stroke. *The Cochrane Library.*

Brumfitt, S. M., & Sheeran, P. (1997). An evaluation of short-term group therapy for people with aphasia. *Disability and Rehabilitation, 19*(6), 221-230.

Carpenter, J., & Cherney, L. R. (2016). Increasing aphasia treatment intensity in an acute inpatient rehabilitation programme: a feasibility study. *Aphasiology, 30*(5), 542-565.

Carragher, M., Conroy, P., Sage, K., & Wilkinson, R. (2012). Can impairment-focused therapy change the everyday conversations of people with aphasia? A review of the literature and future directions. *Aphasiology*, *26*(7), 895–916.

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132*(3), 354-380.

Cherney, L. R. (2012). Aphasia treatment: Intensity, dose parameters, and script training. *International Journal of Speech-Language Pathology, 14*(5), 424-431.

Coelho, C. A., McHugh, R. E., & Boyle, M. (2000). Semantic feature analysis as a treatment for aphasic dysnomia: A replication. *Aphasiology, 14*(2), 133-142.

Conroy, P., Sage, K., & Ralph, M. L. (2009). Improved vocabulary production after naming therapy in aphasia: can gains in picture naming generalise to connected speech? *International Journal of Language & Communication Disorders, 44*(6), 1036-1062.

Dartigues, J. F., Foubert-Samier, A., Le Goff, M., Viltard, M., Amieva, H., Orgogozo, J. M., ... & Helmer, C. (2013). Playing board games, cognitive decline and dementia: a French population-based cohort study. *BMJ open, 3*(8), e002998.

Denes, G., Perazzolo, C., Piani, A., & Piccione, F. (1996). Intensive versus regular speech therapy in global aphasia: A controlled study. *Aphasiology, 10*(4), 385-394.

Difrancesco, S., Pulvermüller, F., & Mohr, B. (2012). Intensive language-action therapy (ILAT): The methods. *Aphasiology, 26*(11), 1317-1351.

Dignam, J. K., Rodriguez, A. D., & Copland, D. A. (2016). Evidence for intensive aphasia therapy: consideration of theories from neuroscience and cognitive psychology. *PM&R, 8*(3), 254-267.

Doesborgh, S. J., van de Sandt-Koenderman, M. W., Dippel, D. W., van Harskamp, F., Koudstaal, P. J., & Visch-Brink, E. G. (2004). Effects of semantic treatment on verbal communication and linguistic processing in aphasia after stroke: a randomized controlled trial. *Stroke, 35*(1), 141-146.

Drew, R. L., & Thompson, C. K. (1999). Model-based semantic treatment for naming deficits in aphasia. *Journal of Speech Language and Hearing Research, 42*(4), 972-989.

Druks, J. (2000). *Object and action naming battery*: Psychology Press.

Drummond, S. S., & Simmons, T. P. (1995). Linguistic performance of female aphasic adults during group interaction. *Journal of Neurolinguistics, 9*(1), 47-54.

Eales, C., & Pring, T. (1998). Using individual and group therapy to remediate word finding difficulties. *Aphasiology, 12*(10), 913-918.

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

Elman, R. (Ed.). (2007a). *Group treatment of neurogenic communication disorders: The expert clinician's approach* (2nd ed.). San Diego, CA: Plural Publishing.

Elman, R. (2007b). Introduction to group treatment of neurogenic communication disorders. In R. Elman (Ed.), *Group treatment of neurogenic communication disorders: The expert clinician's approach* (2nd ed., pp. 1–10). San Diego, CA: Plural Publishing.

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

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.

Galluzzi, C., Bureca, I., Guariglia, C., & Romani, C. (2015). Phonological simplifications, apraxia of speech and the interaction between phonological and phonetic processing. *Neuropsychologia, 71*, 64-83.

Goodglass, H., & Wingfield, A. (1997). Word-finding deficits in aphasia: Brain—behavior relations and clinical symptomatology. In *Anomia* (pp. 3-27).

Greenwald, M. L., Raymer, A. M., Richardson, M. E., & Rothi, L. J. G. (1995). Contrasting treatments for severe impairments of picture naming. *Neuropsychological rehabilitation*, *5*(1-2), 17-49.

Hameister, I., Nickels, L., Ca, S., & Croot, K. (2017). "Do you have mowing the lawn?"–improvements in word retrieval and grammar following constraint-induced language therapy in primary progressive aphasia. *Aphasiology, 31*(3), 308-331.

Hashimoto, N. (2012). The use of semantic- and phonological-based feature approaches to treat naming deficits in aphasia. *Clinical Linguistics & Phonetics*, *26*(6), 518–553.

Herbert, R., Hickin, J., Howard, D., Osborne, F., & Best, W. (2008). Do picture-naming tests provide a valid assessment of lexical retrieval in conversation in aphasia? *Aphasiology, 22*(2), 184-203.

Hilari, K. & Northcott, S. (2017). "Struggling to stay connected": comparing the social relationships of heathy older people and people with stroke and aphasia. *Aphasiology*, *31* (6), 674-687.

Hillari, K., Needle, J.J., & Harrison, K.L. (2012). What are the important factors in health-related quality of life in people with aphasia? A systematic review. Archives of Physical Medicine and Rehabilitation, 93, SS86-95.

Hillis, A. E. (1993). The role of models of language processing in rehabilitation of language impairments. *Aphasiology*, *7*(1), 5-26.

Hillis, A. E. (1998). Treatment of naming disorders: New issues regarding old therapies. *Journal of the International Neuropsychological Society*, *4*(6), 648-660.

Hinckley, J., & Carr, T. (2005). Comparing the outcomes of intensive and non-intensive context-based aphasia treatment. *Aphasiology, 19*(10-11), 965-974.

Hinckley, J. J., & Craig, H. K. (1998). Influence of rate of treatment on the naming abilities of adults with chronic aphasia. *Aphasiology, 12*(11), 989-1006.

Holland, R., Johns, S. L., & Woollams, A. M. (2018). The impact of phonological versus semantic repetition training on generalisation in chronic stroke aphasia reflects differences in dorsal pathway connectivity. *Neuropsychological Rehabilitation*, *28*(4), 548–567.

Howard, D. (1994). The treatment of acquired aphasia. *Phil. Trans. R. Soc. Lond. B, 346*(1315), 113-120.

Howard, D. Cognitive neuropsychology and aphasia therapy: the case of word retrieval. Acquired neurogenic communication disorders: A clinical perspective. Edited by: Papathanasiou I. 2000.

Howard, D., Patterson, K., Franklin, S., Orchardlisle, V., & Morton, J. (1985). The Facilitation Of Picture Naming In Aphasia. *Cognitive Neuropsychology, 2*(1), 49-80.

Jirak, D., Menz, M. M., Buccino, G., Borghi, A. M., & Binkofski, F. (2010). Grasping language - A short story on embodiment. *Consciousness and Cognition, 19*(3), 711-720.

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). The Boston naming test. 2nd. *Philadelphia: Lea & Febiger.*

Kroenke, K. M., Kraft, I., Regenbrecht, F., & Obrig, H. (2013). Lexical learning in mild aphasia: Gesture benefit depends on patholinguistic profile and lesion pattern. *Cortex*, *49*(10), 2637-2649.

Kurland, J., Pulvermüller, F., Silva, N., Burke, K., & Andrianopoulos, M. (2012). Constrained versus unconstrained intensive language therapy in two individuals with chronic, moderate-to-severe aphasia and apraxia of speech: behavioral and fMRI outcomes. *American journal of speech-language pathology, 21*(2), S65-S87.

Kurland, J., Stanek, E. J., Stokes, P., Li, M., & Andrianopoulos, M. (2016). Intensive Language Action Therapy in Chronic Aphasia: A Randomized Clinical Trial Examining Guidance by Constraint. *American journal of speech-language pathology, 25*(4S), S798-S812.

Lanyon, L. E., Rose, M. L., & Worrall, L. (2013). The efficacy of outpatient and community-based aphasia group interventions: A systematic review. *International journal of speech-language pathology, 15*(4), 359-374.

Le Dorze, G., Boulay, N., Gaudreau, J. and Brassard, C. 1(994). The contrasting effects of a semantic versus a formal-semantic technique for the facilitation of naming in a case of anomia. *Aphasiology, 8*, 127–141.

Lorenz, A., & Ziegler, W. (2009). Semantic vs. word-form specific techniques in anomia treatment: A multiple single-case study. *Journal of Neurolinguistics, 22*(6), 515-537.

Maddy, K., Capilouto, G., & McComas, K. (2014). The effectiveness of semantic feature analysis: An evidence-based systematic review. *Annals of physical and rehabilitation medicine, 57*(4), 254-267.

Maher, L. M., & Raymer, A. M. (2004). Management of anomia. *Topics in stroke rehabilitation, 11*(1), 10-21.

Marangolo, P., Bonifazi, S., Tomaiuolo, F., Caighero, L., Coccia, M., Altoe, G., Provinciali, L., & Cantgallo, A. (2010). Improving language without words: First evidence from aphasia. *Neuropsychologia*, *48*, 13, 3824-3833.

Marangolo, P., Cipollari, S., Fiori, V., Razzano, C., & Caltagirone, C. (2012). Walking but not barking improves verb recovery: Implications for action observation treatment in aphasia rehabilitation. *Plos one*, *7* (6), e38610

Marini, A., Andreetta, S., Del Tin, S., & Carlomagno, S. (2011). A multi-level approach to the analysis of narrative language in aphasia. *Aphasiology, 25(11),* 1372-1392.

McKevitt, C., Fudge, N., Redfern, J., Sheldenkar, A., Crichton, S., Rudd, A. R., ... & Rothwell, P. M. (2011). Self-reported long-term needs after stroke. *Stroke, 42*(5), 1398-1403.

Meinzer, M., Rodriguez, A. D., & Rothi, L. J. G. (2012). First Decade of Research on Constrained-Induced Treatment Approaches for Aphasia Rehabilitation. *Archives of Physical Medicine and Rehabilitation, 93*(1), S35-S45.

Morsella, E., & Krauss, R. M. (2004). The role of gestures in spatial working memory and speech. *American Journal of Psychology, 117*(3), 411-424.

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

Nicholas, L. E., & Brookshire, R. H. (1993). A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. *Journal of Speech, Language, and Hearing Research, 36(2),* 338-350.

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

Nickels, L., & Osborne, A. (2016). Constraint Induced Aphasia Therapy: Volunteer-led, unconstrained and less intense delivery can be effective. *Neurorehabilitation, 39*(1), 97-109.

Nickels, L., McDonald, B., & Mason, C. (2016). The impact of group therapy on word retrieval in people with chronic aphasia. *Neurorehabilitation, 39*(1), 81-95.

Neumann, Y. (2018). A case series comparison of semantically focused vs. phonologically focused cued naming treatment in aphasia. *Clinical Linguistics & Phonetics*, *32*(1), 1–27.

Northcott, S., Moss, B., Harrison, K., & Hilari, K. (2016). A systematic review of the impact of stroke on social support and social networks: associated factors and patterns of change. *Clinical rehabilitation, 30*(8), 811-831.

Parr, S. (2007). Living with severe aphasia: Tracking social exclusion. *Aphasiology, 21*(1), 98-123.

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

Pulvermüller, F., Neininger, B., Elbert, T., Mohr, B., Rockstroh, B., Koebbel, P., et al. (2001). Constraint-induced therapy of chronic aphasia after stroke. *Stroke, 32*(7), 1621-1626. 64

Ramsberger, G., & Marie, B. (2007). Self-administered cued naming therapy: A single-participant investigation of a computer-based therapy program replicated in four cases. *American Journal of Speech-Language Pathology, 16*(4), 343-358.

Raymer, A. M., Ciampitti, M., Holliway, B., Singletary, F., Blonder, L. X., Ketterson, T., ... & Gonzalez Rothi, L. J. (2007). Semantic-phonologic treatment for noun and verb retrieval impairments in aphasia. *Neuropsychological rehabilitation, 17*(2), 244-270.

Raymer, A. M., Thompson, C. K., Jacobs, B., & Le Grand, H. R. (1993). Phonological treatment of naming deficits in aphasia: Model-based generalization analysis. *Aphasiology*, *7*(1), 27-53.

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

Romani, C. & Galluzzi, C. (2005). Effects of syllabic complexity in predicting accuracy of repetition and direction of errors in patients with articulatory and phonological difficulties. *Cognitive Neuropsychology*. *22*, 7, 817-850.

Rose, M. L. (2013). Releasing the Constraints on Aphasia Therapy: The Positive Impact of Gesture and Multimodality Treatments. *American Journal of Speech-Language Pathology, 22*(2), S227-S239.

Rose, M. L., & Attard, M. C. (2015). Practices and challenges in community aphasia groups in Australia: Results of a national survey. *International journal of speech-language pathology, 17*(3), 241-251.

Rose, M., & Douglas, J. (2008). Treating a semantic word production deficit in aphasia with verbal and gesture methods. *Aphasiology, 22*(1), 20-41.

Rose, M., & Sussmilch, G. (2007). The effects of semantic and gesture treatments on verb retrieval and verb use in aphasia. *Aphasiology, 22*(7-8), 691-706.

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

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. L., Raymer, A. M., Lanyon, L. E., & Attard, M.C. (2013). A systematic review of gesture treatments for post-stroke aphasia. *Aphasiology*, *27*(9), 1090-1127.

Ross, A., Winslow, I., & Marchant, P. (2006). Evaluation of communication, life participation and psychological well-being in chronic aphasia: The influence of group intervention. *Aphasiology, 20*(5), 427-448.

Saffran, E. M., Berndt, R. S., & Schwartz, M. F. (1989). The quantitative analysis of agrammatic production: Procedure and data. *Brain and language, 37*(3), 440-479.

Salter, K., Teasell, R., Bhogal, S., Zettler, L., & Foley, N. (2009). The Evidence-based review of stroke rehabilitation reviews current practices in stroke rehabilitation. *Aphasia, 14,* 1-55.

Stahl, B., Mohr, B., Dreyer, F. R., Lucchese, G., & Pulvermüller, F. (2016). Using language for social interaction: communication mechanisms promote recovery from chronic non-fluent aphasia. *Cortex, 85*, 90-99.

Stimley, M.A. & Noll, J.D. (1991). The effects of semantic and phonemic prestimulation on picture naming in aphasia. *Brain and Language*, *41* (4), 496-509.

Swinburn, K., Porter, G., & Howard, D. (2004). *Comprehensive aphasia test*: Psychology Press.

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., et al. (2004). A new on-line resource for psycholinguistic studies. *Journal of memory and language, 51*(2), 247-250.

Vanacken, L., Notelaers, S., Raymaekers, C., Coninx, K., van den Hoogen, W., Ijsselsteijn, W., Feys, P. (2010). Game-Based Collaborative Training for Arm Rehabilitation of MS Patients: A Proof-of-concept Game. *Proceedings of the GameDays,* pp. 65–75.

Vickers, C. P. (2010). Social networks after the onset of aphasia: The impact of aphasia group attendance. *Aphasiology, 24*(6-8), 902-913.

Wambaugh, J. (2003). A comparison of the relative effects of phonologic and semantic cueing treatments. *Aphasiology*, *17*(5), 433–441

Wambaugh, J. L., Linebaugh, C. W., Doyle, P. J., & Martinez, A. L. (2001). Effects of two cueing treatments on lexical retrieval in aphasic speakers with different levels of deficit. *Aphasiology, 15*(10-11), 933-950.

Wertz, R. T., Collins, M. J., Weiss, D., Kurtzke, J. F., Friden, T., Brookshire, R. H., ... & West, J. A. (1981). Veterans Administration cooperative study on aphasia: A comparison of individual and group treatment. *Journal of Speech, Language, and Hearing Research, 24*(4), 580-594.

Wisenburn, B., & Mahoney, K. (2009). A meta-analysis of word-finding treatments for aphasia. *Aphasiology, 23*(11), 1338-1352.

Zhang, J., Yu, J., Bao, Y., Xie, Q., Xu, Y., Zhang, J., & Wang, P. (2017). Constraint-induced aphasia therapy in post-stroke aphasia rehabilitation: A systematic review and meta-analysis of randomized controlled trials. *PloS one, 12*(8), e0183349.

Table 1: Demographic and clinical information about study participants. BNT: Boston Naming Task; CAT scores are standardized centile scores from an aphasic population; BNT and CAT scores taken at Time 2. *Ppt= participant*; *CVA = Cerebrovascular artefact*, *MCA = Middle cerebral artery*; *EA = expressive aphasia; RA = receptive aphasia; C1= Cohort 1; C2=Cohort 2 (see text); yr=year; m=months; wk=week; hr=hour; // = no formal qualification.*

Ppt	**Sex**	**Age**	**Post onset**	**Classification & severity**	**Education**		**Occupation**	**CVA**	**ST before 1st baseline**		**ST during study**		**BNT**	**CAT**
			ms		***level***	***yrs***			***details***	***wks***	***wks***	***hrs***		
P1	M	78	10	Conduction aphasia with apraxia	BA	16	Sound Engineer (retired)	R MCA infarct*	Stroke association (6 wks), Community (12 wks)	18	15	44	4	55.1
P2	M	67	31	Severe non-fluent EA with apraxia and dysarthria	//	11	Hospital Porter (retired)	L MCA infarct	Early supported discharge (2/3 sessions) + community	15	24	91	29	48.4
P3	M	69	67	Severe non-fluent EA with moderate RA	//	11	Customer Services Advisor (retired)	Bilateral lacunar infarcts	at home (6-8 months)	24	15	51	0	43.4
P4	F	75	8	Moderate anomic aphasia, mild	//	10	Factory Assembly Worker	L MCA infarct	in hospital + at home	7	8	30	17	48.2

				RA			(retired)							
P5	M	76	151	Moderate/ severe anomic aphasia	BA	16	Civil Engineer (retired)	R MCA infarct	in hospital (3 months) + community (4 months)	26	22	68	14	56.8
P6	M	59	27	Moderate EA with apraxia and dysarthria	//	11	Factory Chemical Mixer	L MCA infarct**	Standard Therapy in 2013	28	19	77	10	53.1
P7	F	35	8	Moderate to severe EA, mild RA	//	13	Retail and Catering Trainer	L MCA infarct	Early supported discharge (2/3 sessions)+ community	12	33	79	14	52.1
P8	M	58	11	Severe EA with apraxia	//	9	Report Writer for Drainage Company	MCA infarct	community (twice a wk)	16	8	101	1	47.5
P9	F	50	4	Severe EA with apraxia, moderate RA	O level	11	Sales Role	Grade 1 SAH followed by L PO infarct	in hospital (10 sessions per wk)	5	15	13	0	41.4
P10	M	56	16	Severe EA, moderate RA	//	11	Steel Company Supervisor	Hemorrh agic stroke	in hospital (2/3 sessions per wk)	24	11	46	7	46.6

P11	M	71	3	Moderate EA	A level	13	Police Chief Inspector (retired)	L MCA infarct	Early supported discharge (3/4 sessions per wk)	8	8	9	29	52.5
P12	M	73	4	Moderate to severe anomic aphasia	A level	15	Telephone Engineer (retired)	Ischemic stroke	Early supported discharge (2/3 sessions per wk)	8	7	7	17	54.0
C1		**70.7**	**49**			**12.5**				**19.7**	**17.2**	**60.2**	**12.3**	**50.8**
SD		7	54			3				8	6	23	10	5
C2		**57.2**	**7.7**			**12**				**12.2**	**13.7**	**42.5**	**11.3**	**49.0**
SD		14	5			2				7	10	40	11	4.75
Mean		**63.9**	**28.3**			**12.3**				**15.9**	**15.4**	**51.3**	**11.8**	**49.9**
SD		13	43			2					8	32	10	4.77

**P1 suffered a previous L MCA infarct in 2011, but with no impact on speech or language abilities*

***P6 suffered a previous R MCA infarct in 2007, but with no impact on speech or language abilitie*

Table 2: Schedule of study; ST=standard therapy; WA=word set A; WB=word set B; WC=word set C; GT-P=game therapy with phonemic cueing; GT-S=game therapy with semantic cueing; GT-G=game therapy with gestural cueing; BNT=Boston Naming Test; CAT=Comprehensive Aphasia Test.

	Time 1		**Time 2**		**Time 3**		**Time 4**		**Time 5**		**Time 6**
	1 week	12-33 weeks	1 week	2 weeks	1 week	2 weeks	1 week	2 weeks	1 week	18-24 weeks	1 week
Cohort 2	**Comprehensive Assessment** Picture naming + Scene descriptions (**WA+WB+WC**), CAT, BTN, Cinderella	Therapy ST	**Comprehensive Assessment** Picture naming + Scene descriptions (**WA+WB+WC**), CAT, BTN, Cinderella	Therapy GT-P	**Limited Assessment** Picture naming+ scene descriptions: **WA+WB**	Therapy GT-S	**Limited Assessment** Picture naming+ scene descriptions: **WB+WC**	Therapy GT-G	**Comprehensive Assessment** Picture naming + Scene descriptions (**WA+WB+WC**), CAT, BTN, Cinderella	---	---
			Time 2		**Time 3**		**Time 4**		**Time 5**		**Time 6**
Cohort 1	---	---	**Comprehensive Assessment** Picture naming + Scene descriptions (**WA+WB+WC**), CAT, BTN, Cinderella	Therapy GT-P	**Limited Assessment** Picture naming+ scene descriptions: **WA+WB**	Therapy GT-S	**Limited Assessment** Picture naming+ scene descriptions: **WB+WC**	Therapy GT-G	**Comprehensive Assessment** Picture naming + Scene descriptions (**WA+WB+WC**), CAT, BTN, Cinderella	Therapy ST	**Comprehensive Assessment** Picture naming + Scene descriptions (**WA+WB+WC**), CAT, BTN, Cinderella

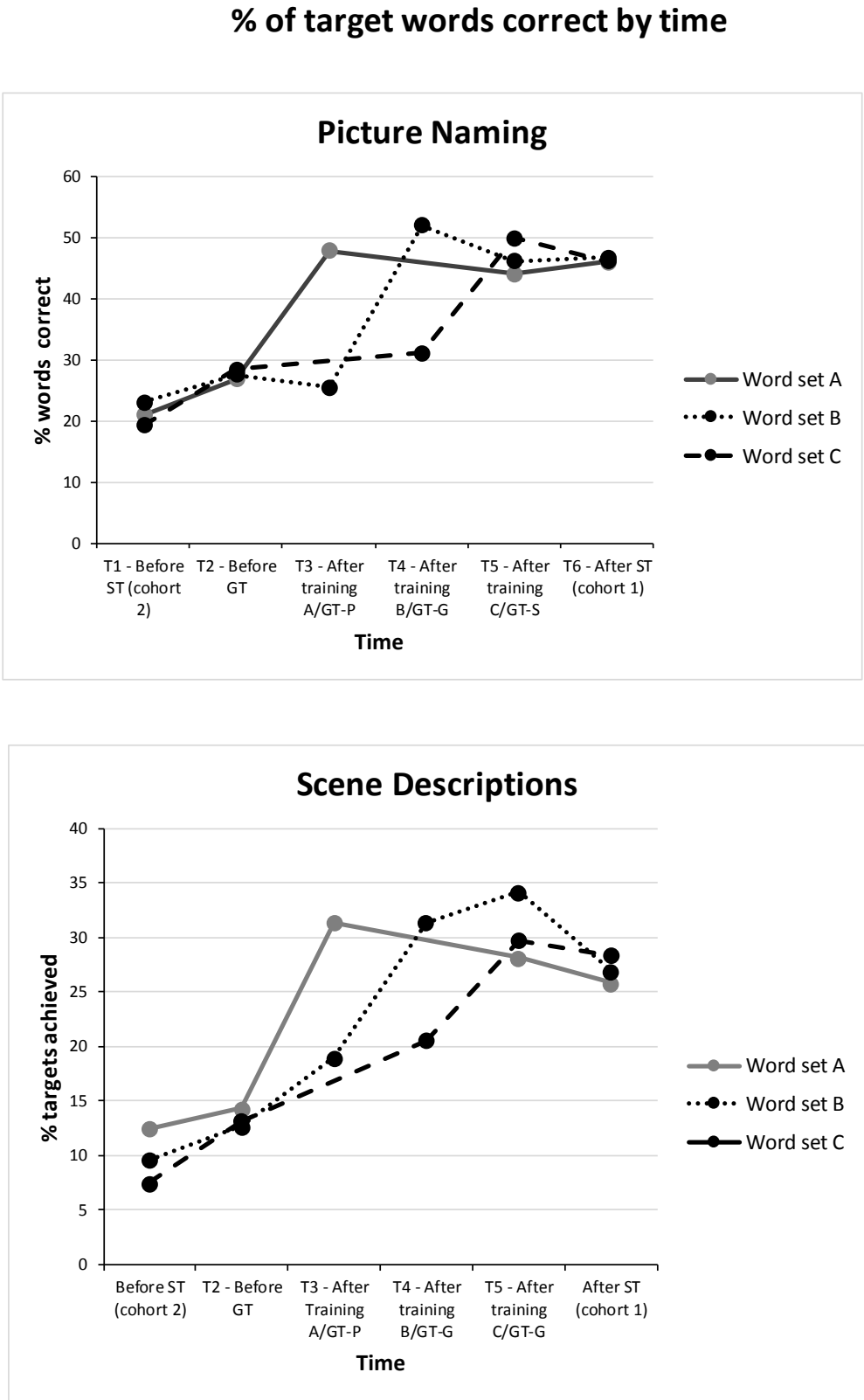


Figure 1: Percent correct of target words by time in therapy protocol and word-set.

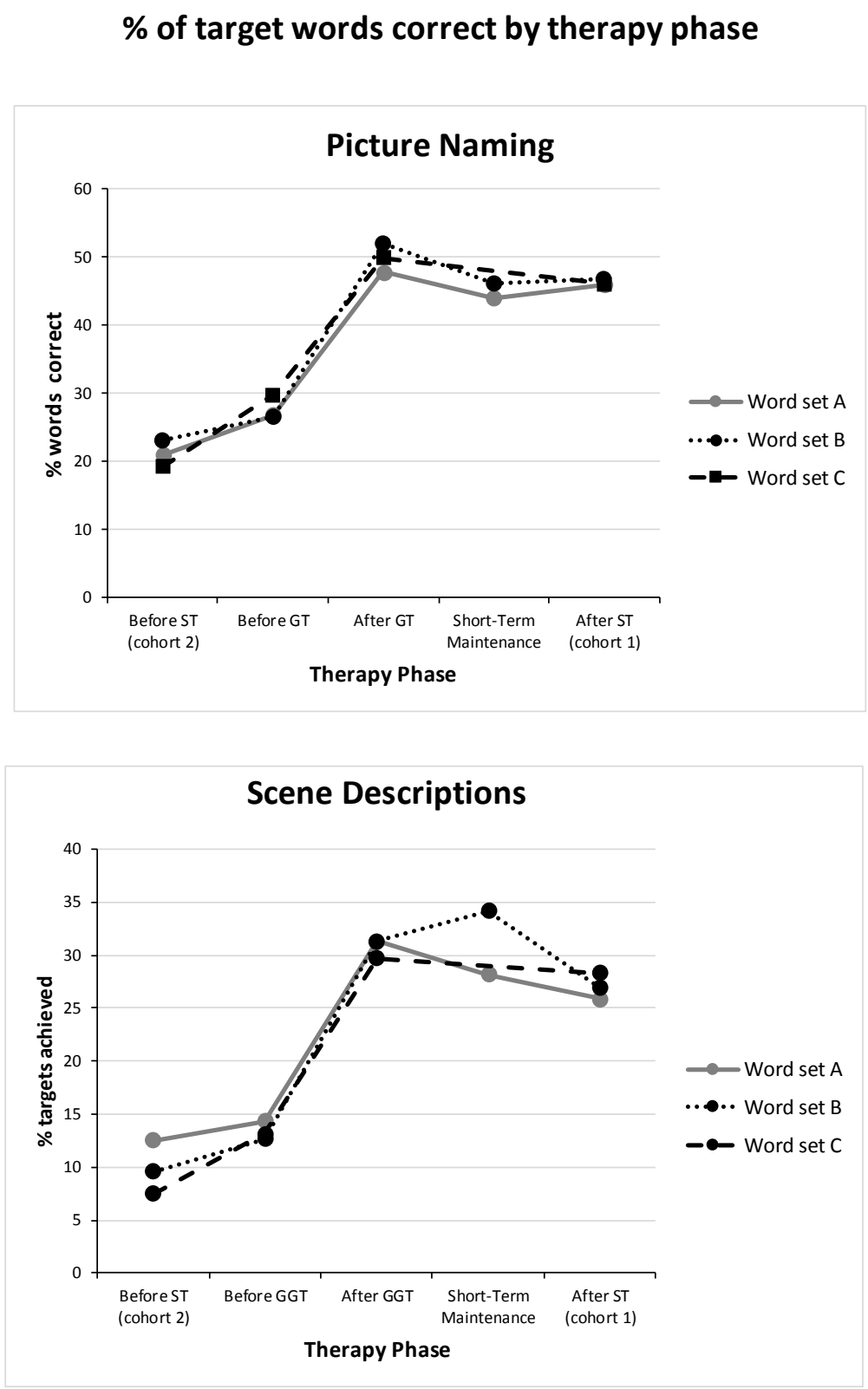


Figure 2: Percent correct of target words by therapy phase and word set.

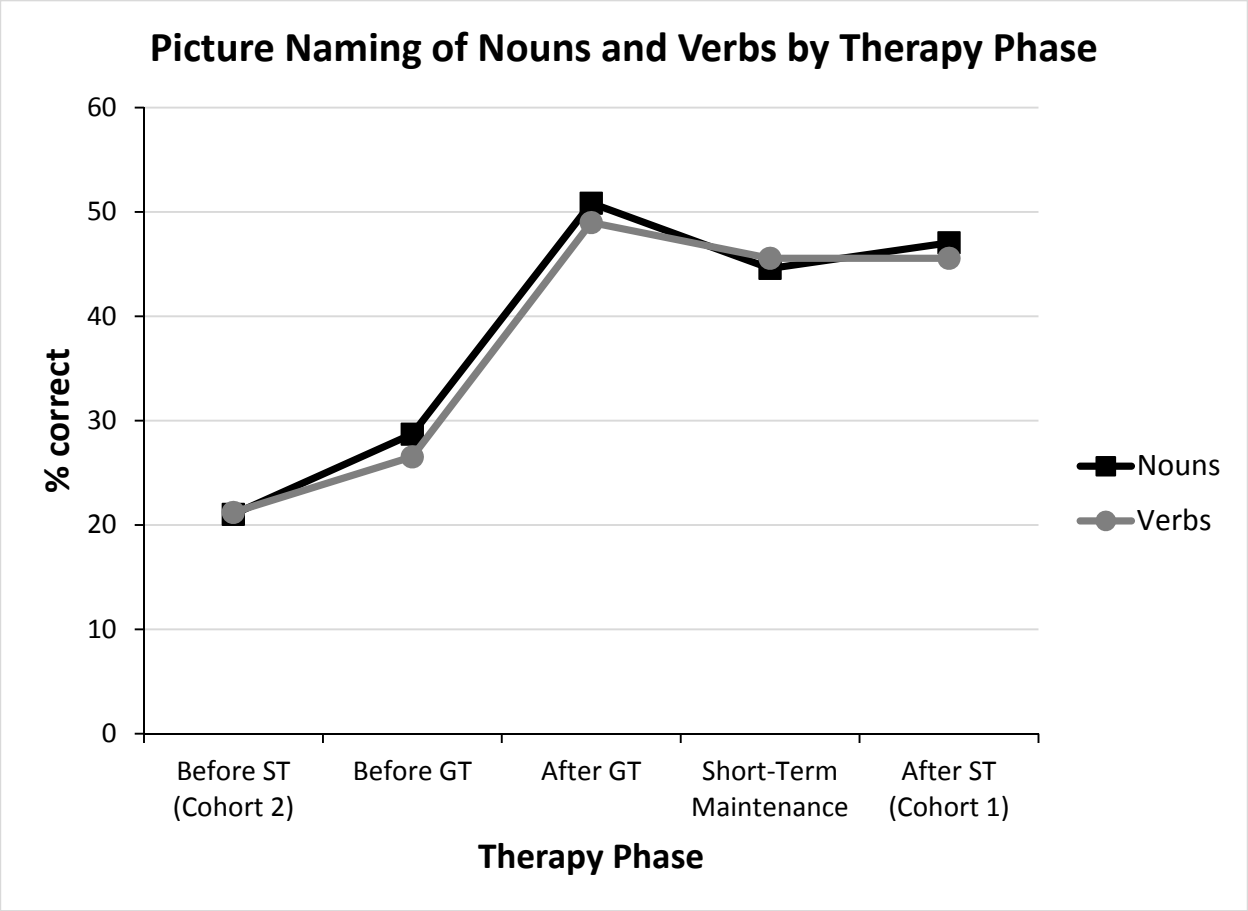
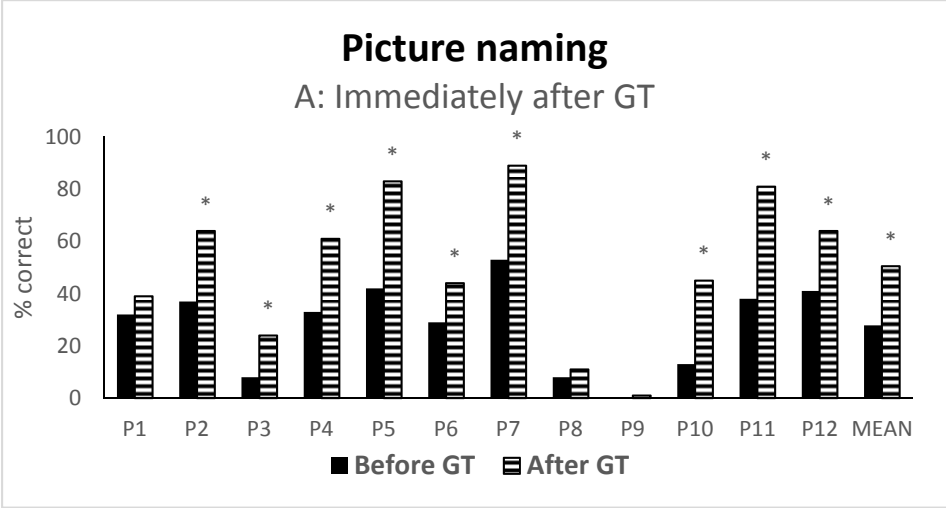
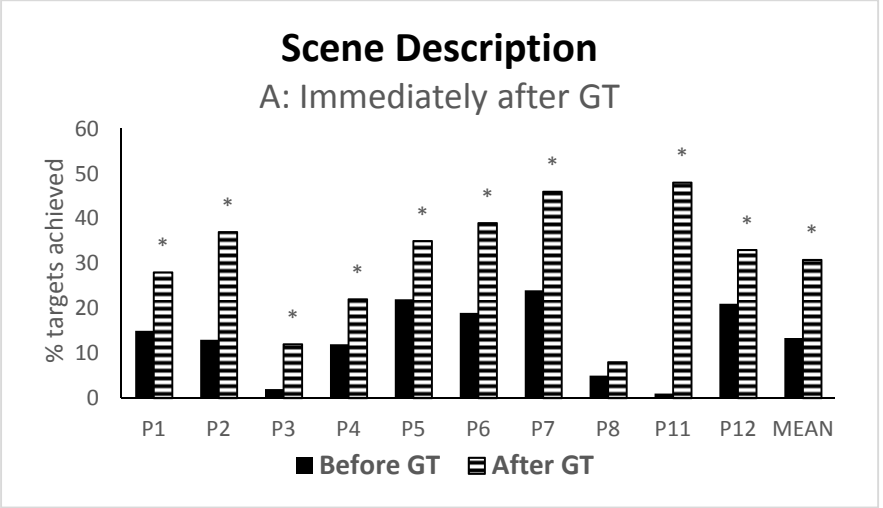
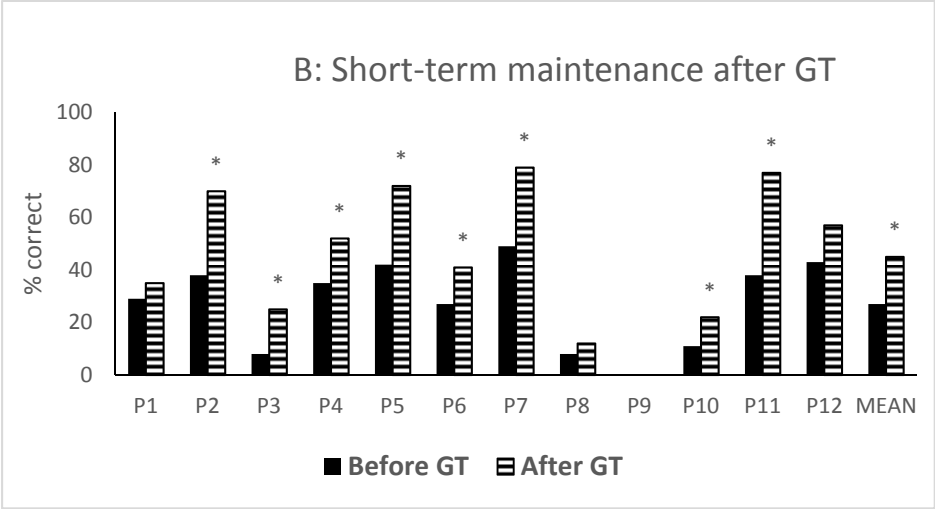
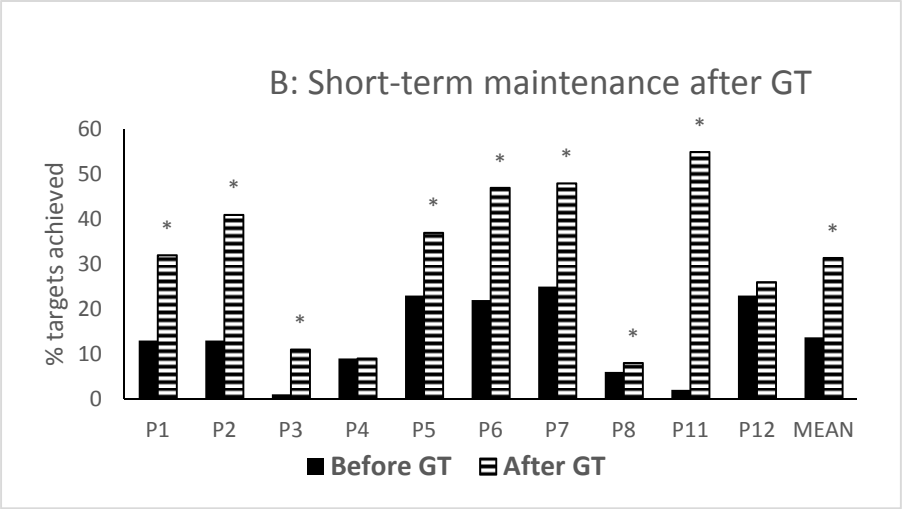


Figure 3: Percent correct for nouns and verbs by therapy phase.


Picture naming
A: Immediately after GT
% correct
100
80
60
40
20
0
P1 P2 P3 P4 P5 P6 P7 P8 P9 P10 P11 P12 MEAN
Before GT
After GT


Scene Description
A: Immediately after GT
% targets achieved
60
50
40
30
20
10
0
P1 P2 P3 P4 P5 P6 P7 P8 P11 P12 MEAN
Before GT
After GT


B: Short-term maintenance after GT
% correct
100
80
60
40
20
0
P1 P2 P3 P4 P5 P6 P7 P8 P9 P10 P11 P12 MEAN
Before GT
After GT


B: Short-term maintenance after GT
% targets achieved
60
50
40
30
20
10
0
P1 P2 P3 P4 P5 P6 P7 P8 P11 P12 MEAN
Before GT
After GT

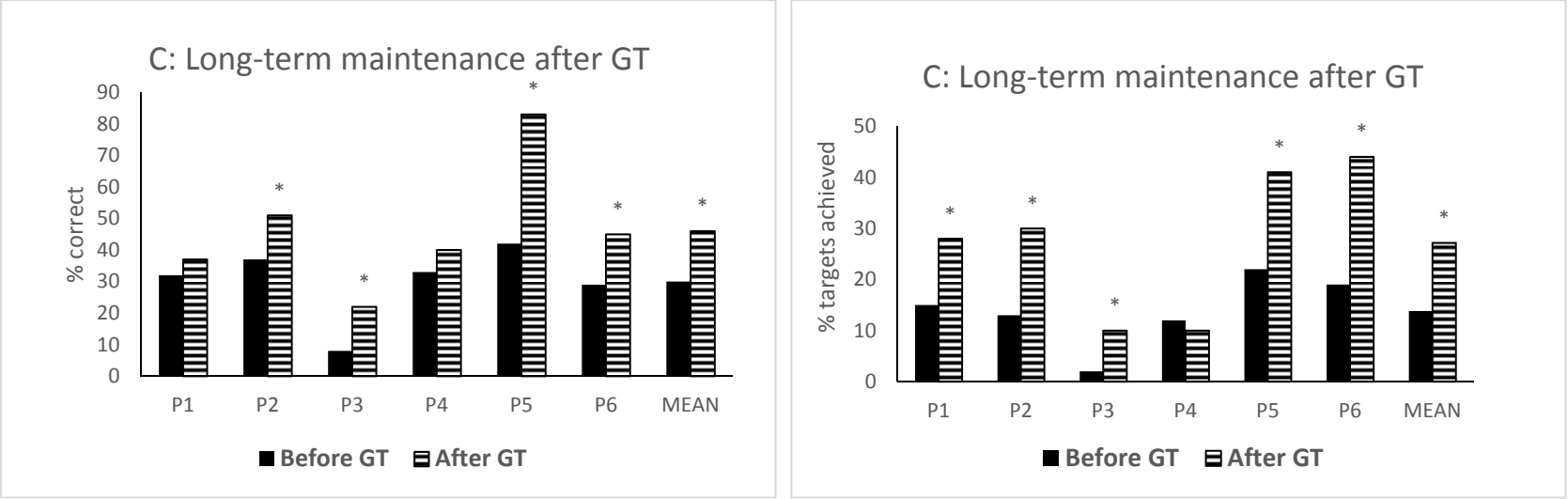
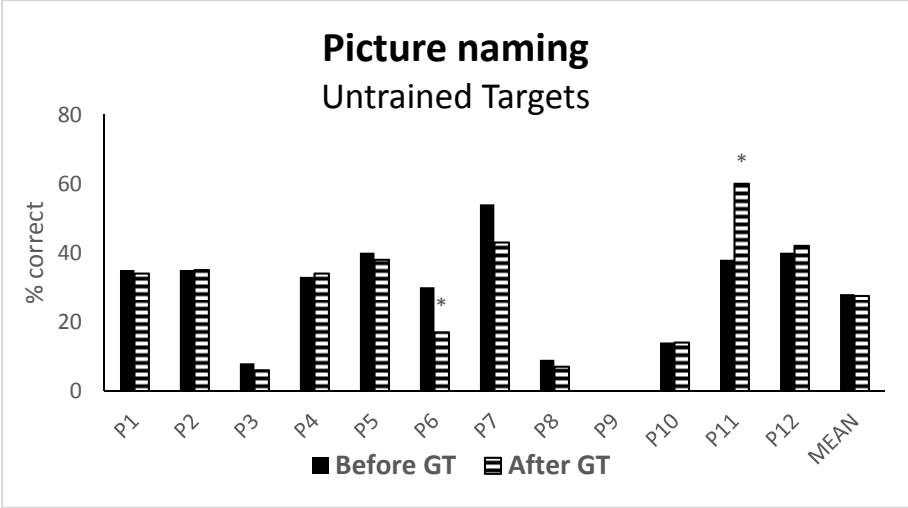
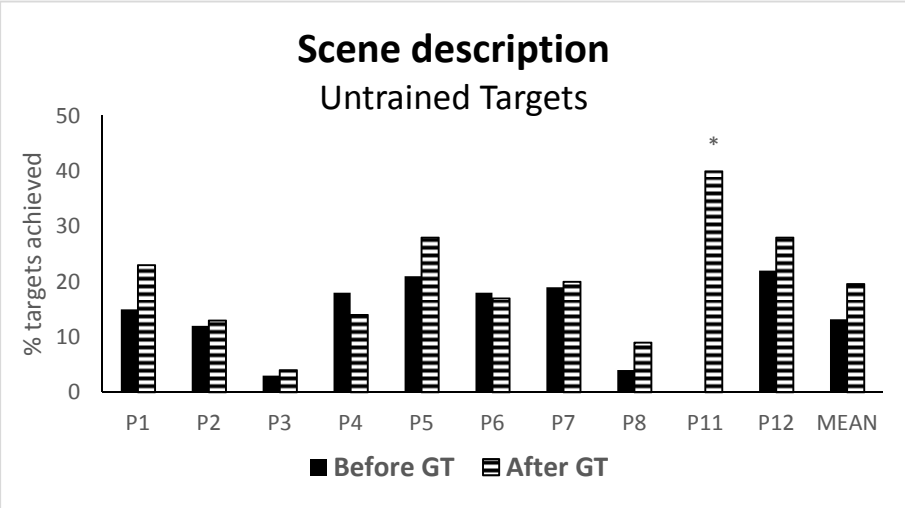
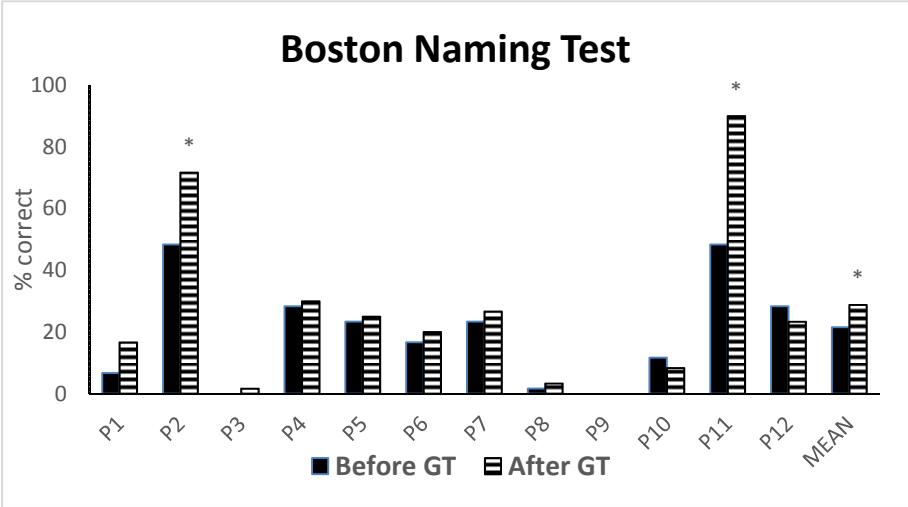
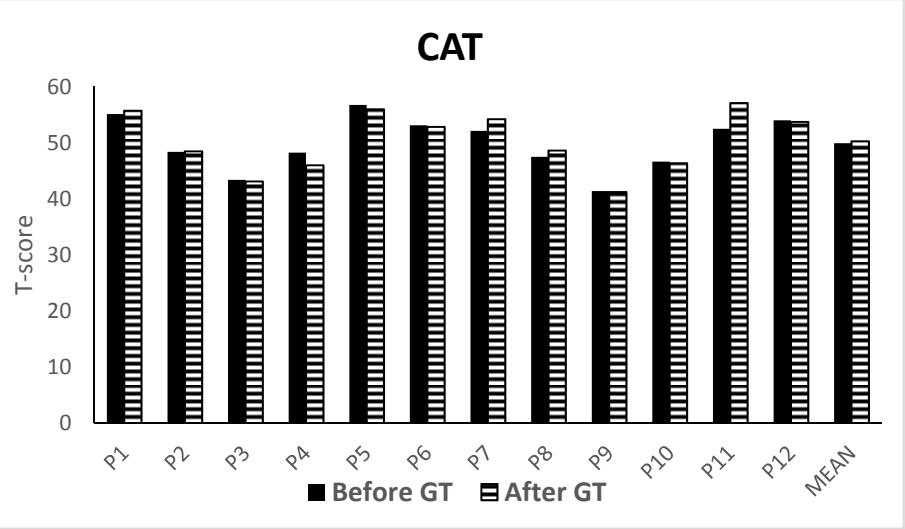


Figure 4. Outcome of Game Therapy for individual participants in picture naming and scene description for trained words. Panel A: immediately after game therapy; Panel B: 2-3 weeks afterward therapy; Panel C: 5-6 months after therapy for cohort 1. Performance in % correct for all three word sets (A, B and C) collapsed (N=180). Asterisks mark significant differences evaluated with chi-square for individual participants and with t-tests for the group (MEAN).


Picture naming
Untrained Targets
% correct
80
60
40
20
0
P1 P2 P3 P4 P5 P6 P7 P8 P9 P10 P11 P12 MEAN
Before GT
After GT


Scene description
Untrained Targets
% targets achieved
50
40
30
20
10
0
P1 P2 P3 P4 P5 P6 P7 P8 P11 P12 MEAN
Before GT
After GT


Boston Naming Test
% correct
100
80
60
40
20
0
P1 P2 P3 P4 P5 P6 P7 P8 P9 P10 P11 P12 MEAN
Before GT
After GT


CAT
T-score
60
50
40
30
20
10
0
P1 P2 P3 P4 P5 P6 P7 P8 P9 P10 P11 P12 MEAN
Before GT
After GT

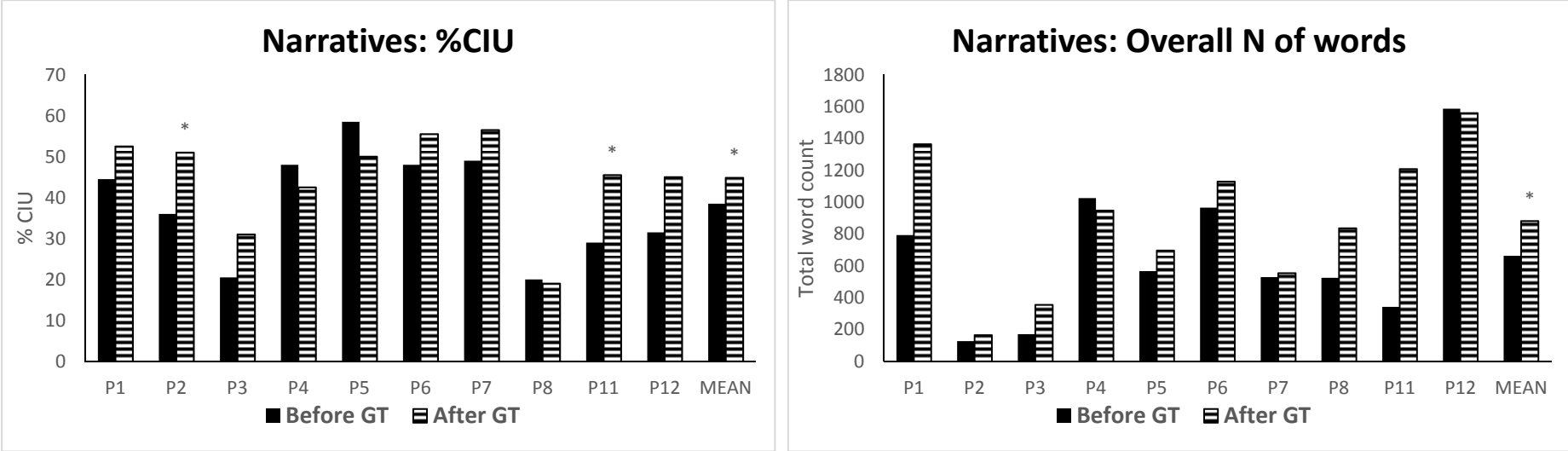
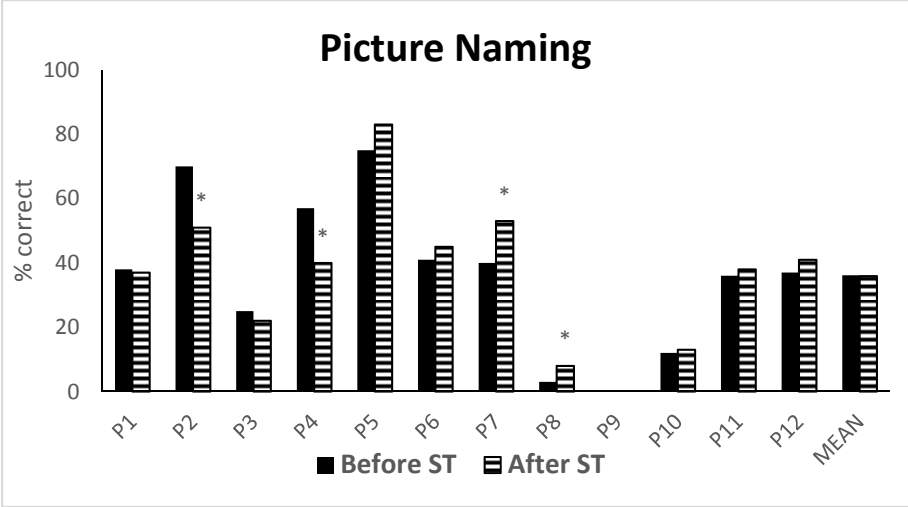
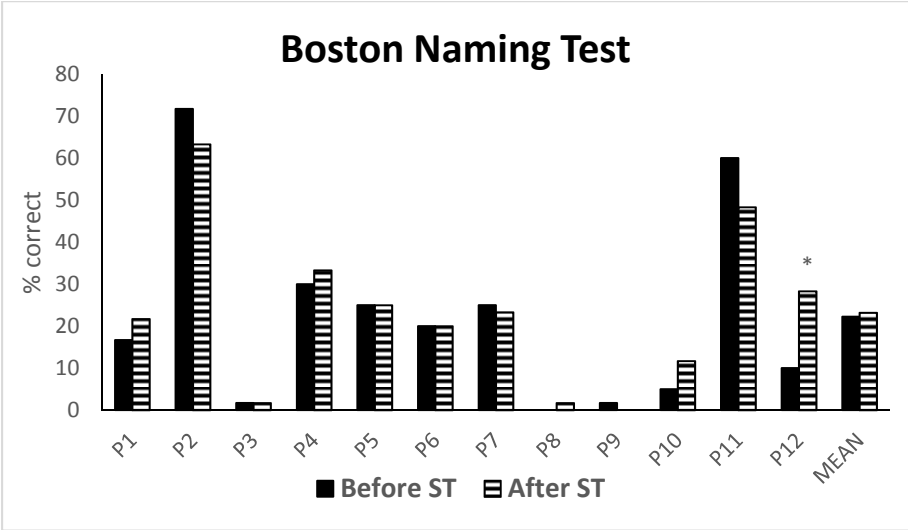


Figure 5 Outcome of Game Therapy for individual participants in measures not targeting trained words. Outcomes compare performance before any game therapy (T2) with performance at the end of all game therapy (T5). Asterisks mark significant differences evaluated with chi-square for individual participants and with paired, one-tailed t-tests for the group (MEAN). As there is no set total for the CAT or word count in narrative production, individual chi-squares could not be performed for these measures.


Picture Naming
% correct
100
80
60
40
20
0
P1 P2 P3 P4 P5 P6 P7 P8 P9 P10 P11 P12 MEAN
Before ST
After ST


Scene Description
% targets achieved
50
40
30
20
10
0
P1 P2 P3 P4 P5 P6 P7 P8 P11 P12 MEAN
Before ST
After ST


Boston Naming Test
% correct
80
70
60
50
40
30
20
10
0
P1 P2 P3 P4 P5 P6 P7 P8 P9 P10 P11 P12 MEAN
Before ST
After ST

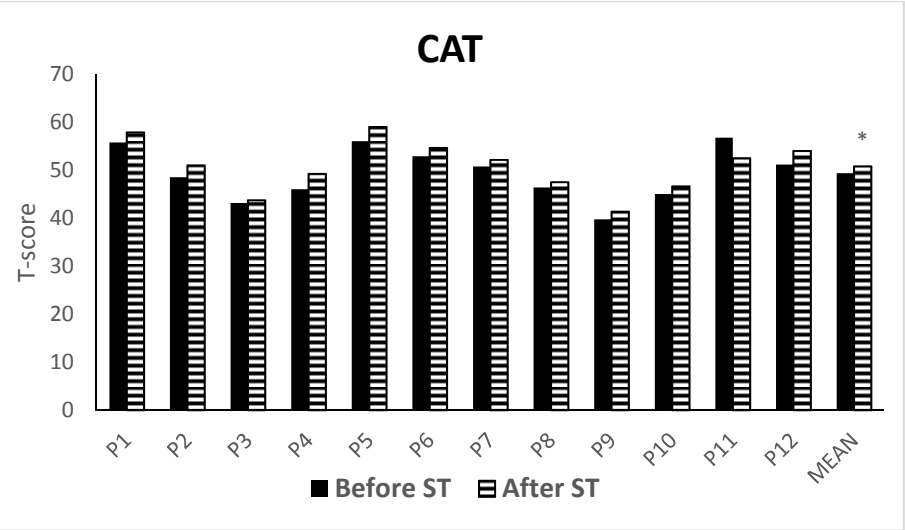

CAT
T-score
70
60
50
40
30
20
10
0
P1 P2 P3 P4 P5 P6 P7 P8 P9 P10 P11 P12 MEAN
Before ST
After ST

Figure 6: Individual outcomes after standard therapy (ST). Asterisks mark significant differences evaluated with chi-square for individual participants and with paired, one-tailed, t-tests for the group (MEAN). As there is no set total for the CAT or word count in narrative production, individual chi-squares could not be performed for these measures.

APPENDIX A: Matching sets of words trained in Game Therapy. Freq = BNC Frequencies from the British National Corpus; (Leech, 1993; http://corpus.byu.edu/bnc/); Subtlex UK Freq = frequencies (Walter et al., 2014; www.psychology.nottingham.ac.uk/subtlex-uk/)

NOUNS - SET A - Game					
Target	**Freq**	**Subtlex UK Freq**	**Phon len**	**Phon comp**	**AoA**
umbrella	432	1547	7	2	5.7
strawberry	307	1952	8	2	4.2
banana	347	3304	6	0	3.8
policeman	228	2322	8	0.5	4.4
sausage	262	3419	5	0	6.1
kitchen	2666	32155	5	0	3.7
dustpan	5	63	7	1	6.5
rubbish	100	10140	5	0	11
window	4050	13791	5	1	4.7
letter	7458	14340	4	0	4.7
jigsaw	53	957	5	0.5	7.6
kettle	68	2105	5	0.5	8.1
picnic	327	2189	6	0.5	5.5
basket	132	2931	6	0.5	5.7
nappy	195	494	4	0	8.4
icecream	28	8	7	2.5	3.3
bucket	374	3295	5	0	5.6
lighthouse	76	1175	3	1.5	6.1
balloon	740	3574	5	0	4.4
seesaw	26	188	4	0	4.8
lemon	112	5806	5	0	4.7
carrot	79	2410	5	0	2.7
handcuffs	12	392	8	2	6.3

NOUNS-SET B-Game pg					
Target	**Freq**	**Subtlex UK Freq**	**Phon len**	**Phon comp**	**AoA**
butterfly	305	2181	7	1.5	3.7
lawnmower	21	257	7	1	6.1
wheelbarrow	26	493	7	1.5	6.8
spaghetti	197	1237	7	1	4.3
garden	1963	38363	5	0.5	5.3
ladder	529	2934	4	0	4.4
spider	290	3520	5	1.5	3.4
flower	563	6389	5	1.5	3.1
hammock	3	269	5	0	7.1
sandwich	576	3556	6	0.5	4.8
wheelchair	552	2061	5	1.5	6.9
wallet	539	1301	5	0	5.9
teapot	47	1211	5	0	6.2
pocket	1585	8690	5	0	4.7
pizza	1006	3596	5	0.5	4.7
waitress	288	478	6	1.5	7.2
candle	312	1230	6	1	5.4
arrow	187	1223	3	1	6.1
helmet	433	2398	6	0.5	5.7
table	7748	25504	5	1	4.4
castle	543	9715	5	0.5	5.8
trophy	280	4771	5	1.5	5.6
feather	166	1518	4	0	4.7

NOUNS-SET C-Game p					
Target	**Freq**	**Subtlex UK Freq**	**Phon len**	**Phon comp**	**AoA**
camera	258	13767	6	0	6
microphone	290	1270	8	2	6.3
piano	1825	4610	5	1.5	5.5
balcony	42	1375	7	0.5	8.1
trousers	1420	3536	7	1.5	7.9
bedroom	2069	20187	6	0.5	3.9
toilet	697	6993	5	0.5	3.5
mirror	1368	5672	4	0	4.9
razor	250	978	4	0.5	7.1
curtains	157	1745	6	1.5	5
slipper	17	314	5	1	5.3
tractor	86	1404	6	1.5	5.5
windmill	30	644	7	1	5.3
rainbow	380	3052	5	1.5	4.3
button	1239	5813	5	0.5	4.8
collar	102	1912	4	0	6.6
jacket	1142	3906	5	0	4
chicken	459	13434	5	0.5	3.3
guitar	1534	4894	4	0	5.3
trumpet	19	1321	7	1.5	6.3
music	792	44883	6	1	3.8
ticket	3418	6464	5	0	5.3
shadow	112	6278	4	0.5	5

chimney	68	1609	5	0.5	6.6	**lightning**	18	2016	6	1	4.8	**bottle**	1426	8898	5	0.5	3.6
drawer	286	1914	3	1	5.4	**branch**	1648	2532	5	2	5.1	**teeth**	1736	10110	3	0	3.6
spoon	282	3970	4	1	2.5	**leaf**	420	3894	3	0	4.6	**shirt**	1629	5976	3	0	3.5
plane	205	7734	4	1.5	5	**glass**	2894	16716	4	1	4.5	**bath**	726	8931	3	0	3.2
cloud	109	11742	4	1.5	3.6	**clock**	1789	15610	4	1	4.4	**duck**	745	6819	3	0	3.5
church	1149	21071	3	0	5.2	**bridge**	2460	10264	4	1	5.6	**scarf**	220	1162	4	1	5.7
rug	36	745	3	0	4.6	**flag**	67	5252	4	1	5.3	**bench**	129	3166	4	1	4.2
mean	**674**	**5245**	**5.1**	**0.7**	**5.4**		**915**	**5973**	**5.1**	**0.8**	**5.2**		**811**	**6650**	**5**	**0.7**	**5**
SD	1537	0.8	7167	1.5	1.7		1501	8385	1.1	0.6	1		815	1.4	0.6	8537	1.3
t-test a/b	0.54	0.72	0.92	0.35	0.7												
t-test a/c							0.67	0.49	0.79	1	0.4						
t-test b/c													0.74	0.76	0.83	0.29	0.49

VERBS - SET A - Game

Target	Freq	Subtlex UK Freq	Phon len	Phon compl	AoA
meditate	58	102	7	0.5	11
deliver	2522	9811	6	0	6.6
open	16507	52274	4	1.5	5
hoover	557	703	4	0	9.2
polish	716	3025	5	0	8.1
relax	1006	6322	6	1	7.1
scold	53	42	5	2.5	8.5
paddle	110	1071	5	0.5	6.6
sunbathe	10	90	6	1	8.5
massage	260	978	5	0	11
marry	1774	5433	4	0	5.8
arrest	3879	4553	5	1.5	7.5

VERBS - SET B - Game p

Target	Freq	Subtlex UK Freq	Phon len	Phon comp	AoA
somersault	12	221	8	1	5.6
celebrate	1852	6869	8	1.5	6.2
whisper	245	946	5	0.5	4.3
water	17769	68324	4	0	2.4
tickle	126	1561	5	0.5	4.4
whistle	442	3543	5	0.5	5.4
prune	34	403	4	1	7.9
propose	907	1259	6	1.5	9
sharpen	86	264	5	0.5	5.6
slip	987	4330	4	1	6.2
juggle	100	466	5	0.5	6.4
balance	4865	10804	6	1.5	9.9

VERBS - SET C - Game p

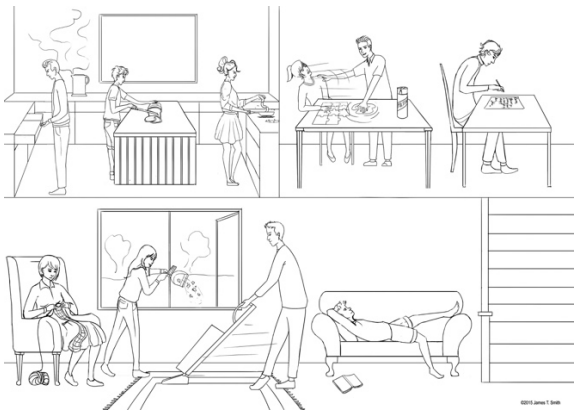
Target	Frequ	Subtlex-UK Freq	Phon len	Phon comp	AoA
decorate	152	1003	7	0.5	7
parachute	253	1037	7	0	7
shower	1029	6660	4	0.5	4.7
buckle	194	816	5	0.5	6.4
shiver	35	375	4	0	7.5
hammer	1314	9185	4	0	5.4
carry	5460	17841	4	0	5.2
follow	7111	18563	4	0.5	5.1
bury	2167	2379	4	0	6.3
conduct	3379	2406	7	1.5	8.6
curtsey	1	61	5	0.5	8.8
listen	14462	35886	5	0.5	5.4

hitchhike	21	27	6	1	9.5	**salute**	399	948	5	0	7.4	**argue**	4278	5551	4	1.5	4.6
cook	2767	21727	3	0	4.2	**jump**	4022	14018	4	1	2.8	**hang**	3453	19529	3	0	6.7
throw	5176	15275	3	1.5	4.1	**catch**	4080	22304	3	0	4.6	**brush**	909	3898	4	1	3.8
write	10561	15768	3	0.5	4.9	**dig**	1208	5499	3	0	4.2	**shave**	217	965	3	0.5	5.9
cut	19622	45221	3	0	4.4	**rake**	88	503	3	0.5	5.3	**sleep**	5549	17715	4	1	2.8
knit	174	875	3	0	7.9	**eat**	7374	32360	2	0.5	2.8	**smile**	2149	10393	4	1.5	3.5
fly	4385	16151	3	1.5	3.1	**pour**	639	3568	2	0	5	**yawn**	74	362	3	0	5.3
reach	6365	13080	3	0	4.9	**fall**	9460	20583	3	0	4.7	**watch**	16770	40471	3	0	4.3
cry	2775	7389	3	1.5	2.8	**laugh**	2790	11958	3	0	3.8	**feed**	2148	9757	3	0	4.2
swim	775	5244	4	1	4.2	**drink**	3620	23199	5	2	3.5	**fish**	3645	31347	3	0	4.1
sweat	737	2339	4	1	7.3	**read**	20764	27812	3	0	4.1	**bark**	229	1677	3	0	5.2
kneel	59	334	3	0.5	7	**miss**	7057	33048	3	0	6.3	**hatch**	1328	1427	3	0	7
sell	7497	33397	3	0	7.1	**climb**	1023	5717	4	1.5	5.3	**sing**	3966	17636	3	0	3.5
push	5108	16712	3	0	4.3	**light**	8885	38200	3	0.5	4.1	**dance**	3328	26748	4	1	4.6
peel	342	2145	3	0.5	6.1	**wave**	2322	7203	3	0.5	4.3	**play**	23231	98534	3	1.5	4.1
drive	6870	19944	4	1.5	5.3	**cross**	4157	17449	4	1	4.7	**clap**	119	1658	4	1	3.4
smoke	3623	6780	4	1.5	4	**win**	14602	62609	3	0	4.2	**shake**	1669	7587	3	0.5	5.3
ring	2987	16766	3	0	4.5	**rain**	3237	24402	3	0.5	3.6	**lean**	689	2268	3	0	7.2
mean	**3577**	**10786**	**4.1**	**0.7**	**6.4**		**4105**	**15012**	**4.1**	**0.6**	**5.1**		**3644**	**13125**	**4**	**0.5**	**5.4**
SD	4794	13241	1.2	0.7	2.2		5402	17837	1.5	0.6	1.7		5406	19738	1.2	0.5	1.5
t-test a/b	0.69	0.3	0.93	0.62	**0**												
t-test a/c							0.96	0.59	0.75	0.19	0.1						
t-test b/c													0.74	0.7	0.71	0.36	0.49
t-test nouns vs verbs	**0**	**0.05**	**0**	0.93	0.1		**0**	**0.01**	**0.01**	0.13	0.8		**0.01**	0.1	**0**	0.19	0.26

APPENDIX B: Pictures for Scene Descriptions. In the boxes examples of target words.

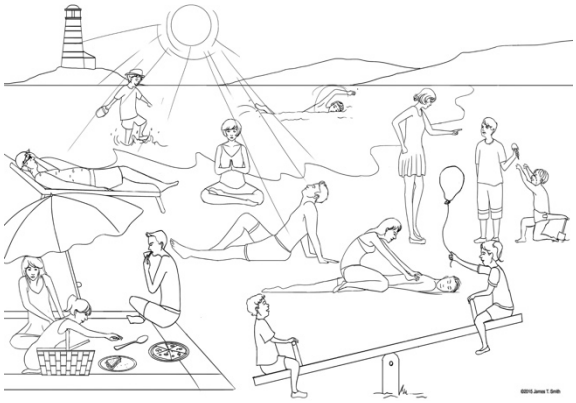
Language Game P: Phonemic Cueing

Scene 1- kitchen and living room



Woman cooking sausage in kitchen. Kettle on. Boy cutting bread. Girl with dustpan, throwing rubbish out of window. Man writing letter. Man hoovering rug. Boy polishing table, pushing girl, whilst girl does jigsaw. Woman knitting. Man opening drawer. Boy relaxing.

Scene 2- beach



Family under umbrella having picnic from basket. Girl reaching for spoon. Baby crying wearing nappy. Mum scolding boy, taken icecream from baby. Boy paddling with bucket. Person swimming, lighthouse on rocks. Children on see saw with balloon. Woman meditating. Man sunbathing, sweating. Woman kneeling, massaging back.

Scene 3- busy street

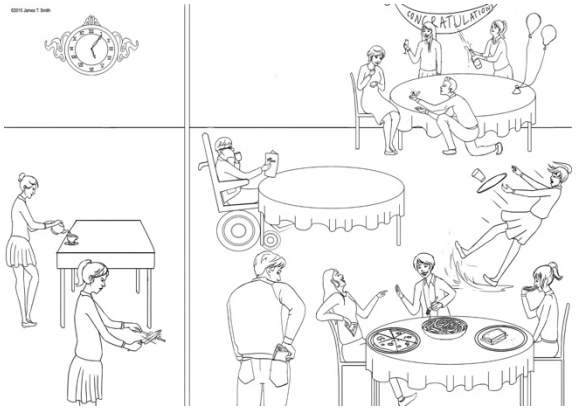
Man selling lemons, strawberries, carrots. Boy peeling banana. Policeman arresting man, got handcuffs. Girl hitchhiking, car drives off. Man delivering parcel. Chimney smoking. People at church getting married, bell ringing. Plane flying in clouds.

Language Game PG: Phonemic+Gestural Cueing

Scene 1- in the garden

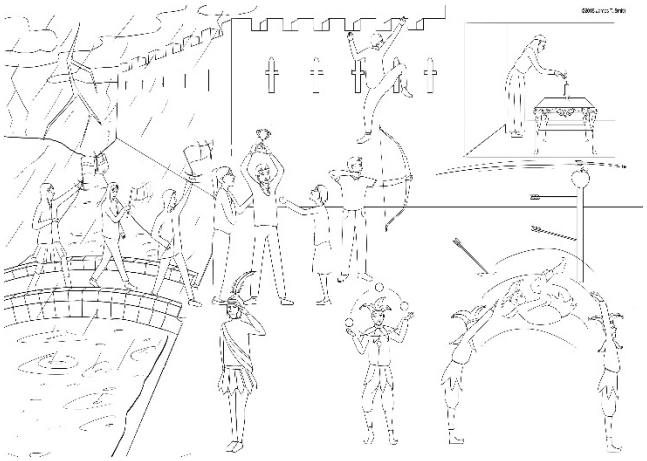
It's a garden, boy catching leaf from branch. Butterfly on flower. Spider on ladder. Children jumping over wall and balancing on wall. Boy whispering to girl in hammock. Girl watering plants, being tickled. Man whistling, pruning. Wheelbarrow and lawnmower against shed, man digging woman raking.

Scene 2- in the café



Man proposing, people celebrating. Eating spaghetti and sandwich, pizza. Man in wheelchair, drinking, reading menu. Friends laughing. Waitress pouring from teapot. Waitress slipping, glass falling. Waitress sharpening knife. Man taking wallet out of pocket. Clock on wall.

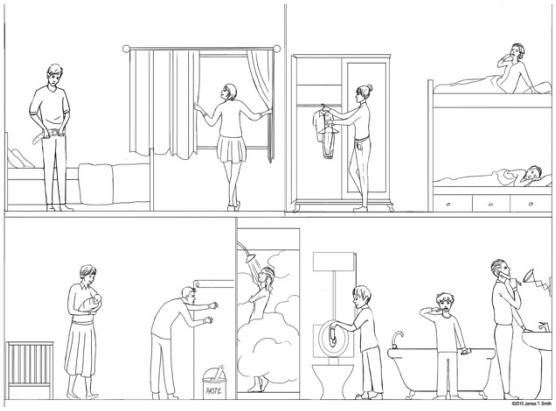
Scene 3- daytrip to the castle



It's a castle, people crossing bridge waving flags. There's lightning, it's raining. Men climbing juggling, somersaulting. Woman lighting candle, on table. Man saluting, feather in helmet. Arrow misses. Man winning trophy.

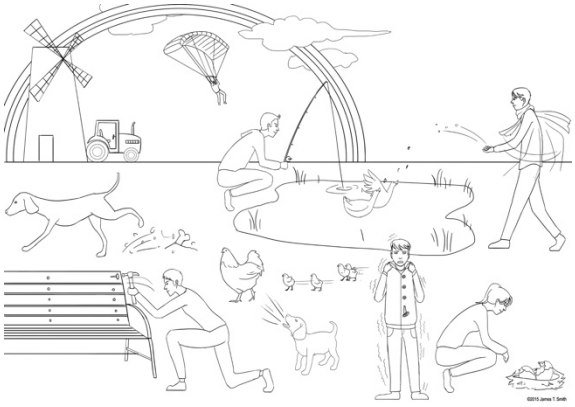
Language Game PS: Phonemic+semantic cueing

Scene1- bedroom and bathroom



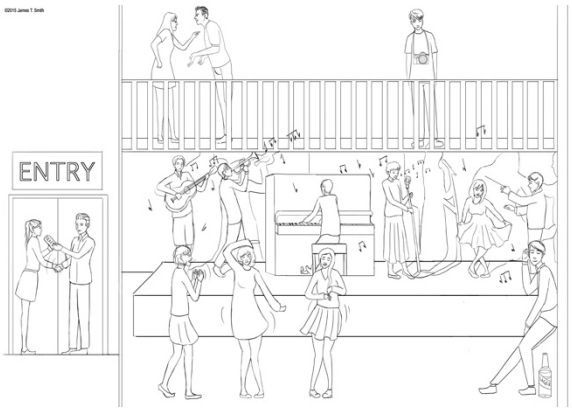
Man in bedroom, buckling belt. Woman opening curtains. Boy sleeping, girl yawning. Woman carrying baby, smiling. Woman hanging shirt and trousers. Man decorating. Boy brushing teeth, in front of bath. Woman showering. Man looking in mirror shaving with razor. Boy dropping slipper in toilet.

Scene 2- in the countryside



There is a tractor by the windmill. Man parachuting under rainbow. Man hammering nail on bench. Boy shivering wearing jacket, turning collar up, button hanging off. Boys feeding ducks and fishing, wearing scarf. Chickens following mother. Girl watching egg hatching. Dog burying bone, dog barking.

Scene 3- at a concert



Band playing guitar trumpet and piano. Woman singing at microphone. Man conducting, woman curtseying. People on balcony arguing, got camera. Listening to music, leaning against wall, got bottle, see shadow. Give ticket, shaking hands. People clapping and dancing.