
Article:
Conroy, P, Sotiropoulou Drosopoulou, C orcid.org/0000-0002-7237-3514, Humphreys, GF et al. (2 more authors) (2018) Time for a quick word? The striking benefits of training speed and accuracy of word retrieval in post-stroke aphasia. Brain, 141 (6). pp. 1815-1827. ISSN 0006-8950

https://doi.org/10.1093/brain/awy087

OXFORD UNIVERSITY PRESS | Brain

Time for a quick word? The striking benefits of training speed and accuracy of word retrieval in post-stroke aphasia.

Journal:	*Brain*
Manuscript ID	Draft
Manuscript Type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Conroy, Paul; University of Manchester, Neuroscience and Aphasia Research Unit (NARU) Sotiropoulou Drosopoulou, Christina; University of Manchester, Neuroscience and Aphasia Research Unit (NARU) Humphreys, Gina; University of Manchester, Neuroscience and Aphasia Research Unit (NARU) Halai, Ajay; University of Manchester, Neuroscience and Aphasia Research Unit, School of Biological Sciences Lambon Ralph, Matthew; University of Manchester,
Subject category:	CNS injury and stroke
To search keyword list, use whole or part words followed by an *:	Aphasia < CNS INJURY AND STROKE, Stroke: rehabilitation < CNS INJURY AND STROKE, Stroke: imaging < CNS INJURY AND STROKE, Neuropsychology < NEUROPSYCHIATRY, Clinical practice < SYSTEMS/DEVELOPMENT/PHYSIOLOGY

Time for a quick word? The striking benefits of training speed and accuracy of word retrieval in post-stroke aphasia.

Paul CONROY✉ *, Christina SOTIROPOULOU DROSOPOULOU*, Gina F. HUMPHREYS, Ajay D. HALAI, & Matthew A. LAMBON RALPH✉

Neuroscience & Aphasia Research Unit,
Division of Neuroscience & Experimental Psychology,
School of Biological Sciences
University of Manchester UK

✉Address correspondence to: Dr. Paul Conroy or Prof. Matt Lambon Ralph, Neuroscience and Aphasia Research Unit (NARU), School of Biological Sciences, University of Manchester, Oxford Road, Manchester M13 9PL, UK. E-mail: Paul.Conroy@manchester.ac.uk or matt.lambon-ralph@manchester.ac.uk

* - Joint first authors

ORCiD:
Conroy: 0000-0003-2189-4149
Sotiropoulou Drosopoulou: 0000-0002-7237-3514
Humphreys: 0000-0002-4157-3409
Halai: 0000-0003-1725-7948
Lambon Ralph: 0000-0001-5907-2488

Running title: speed + accuracy of word retrieval in aphasia

Acknowledgements: We would like to thank the participants with aphasia who took part in this study. This research was supported by grants from the Rosetrees Foundation, the MRC (MR/J004146/1) and ERC (GAP: 670428 - BRAIN2MIND_NEUROCOMP).

Abstract

One third of stroke survivors experience deficits in word-retrieval as a core characteristic of their aphasia, which is frustrating, socially-limiting and disabling for their professional and everyday lives. The, as yet undiscovered, "holy grail" of clinical practice is to establish a treatment that not only improves item naming but also generalizes to patients' connected speech. Speech production in healthy participants is a remarkable feat of neurocognitive engineering being both rapid (at least 120 words per minute) and accurate (~one error per 1000 words). Accordingly, we tested the hypothesis that word-finding treatment will only be successful and generalize to connected speech if patients' word retrieval is both accurate and quick.

This study compared a novel combined speed- and accuracy-focused intervention (RISP: 'repeated, increasingly-speeded production') to standard accuracy-focused treatment. Both treatments were evaluated for naming and connected speech outcomes, and related to the patients' neuropsychological and lesion profiles. Twenty participants with post-stroke chronic aphasia of varying severity and subtype took part in 12 computer-based treatment sessions over 6 weeks. Four carefully-matched word-sets were randomly allocated either to RISP, standard accuracy-only treatment, or untreated (two control sets). In the standard treatment, sound-based naming cues facilitated naming accuracy. RISP encouraged naming to become gradually quicker, aiming towards the naming time of age-matched controls.

RISP was significantly more effective in improving and maintaining picture naming accuracy and speed (reduced latencies). Generalization of treated vocabulary to connected speech was significantly increased for all items relative to the baseline. RISP generated substantial and significantly greater deployment of targeted items in connected speech. These gains were maintained at one month post intervention. There was a significant negative correlation for RISP between the patients' phonological scores and the magnitude of the therapy effect, which may have reflected the fact that the substantial, beneficial effect of RISP generated a ceiling effect in the milder patients. Maintenance of the RISP effect correlated positively with executive skills. The neural correlate analyses revealed that participants with the greatest damage to the posterior superior temporal gyrus extending into the white matter of the

inferior longitudinal fasciculus, showed the greatest RISP benefit. RISP was well tolerated by participants across the range of severity and aphasia subtype, indicating that this type of intervention has considerable clinical utility and broad applicability.

Keywords: aphasia, word retrieval, speed, naming, treatment, stroke.

Introduction

Fluent speech production requires rapid and errorless retrieval of vocabulary, which occurs at a rate of at least two words per second and less than one error per 1000 words (Bird, Lambon Ralph, Patterson, & Hodges, 2000; Levelt, 1989). Aphasia, an acquired language disorder, occurs in at least one third of stroke survivors (The Stroke Association (UK), 2016). Failures, errors or delays in word retrieval (anomia) are the most pervasive aphasic symptoms (Laine & Martin, 2006). They occur across the range of aphasic severity; in severe aphasia, simple, everyday words cannot be produced easily or quickly; in milder presentations, there may be delay in retrieval of infrequently accessed, more sophisticated vocabulary.

Both the assessment and treatment of aphasia typically involve confrontation naming. This approach makes the assumption that performance in picture naming tasks will reflect, at least in broad terms, the ease and reliability of word retrieval in connected speech and everyday communication (Herbert, Hickin, Howard, Osborne, & Best, 2008). There remains, however, a lack of clarity and paucity of data in the aphasia treatment literature as to what extent the treatment of naming benefits vocabulary use in expressive language (a form of post-treatment generalization, which can be differentiated from therapy generalization from treated to untreated items: (Best et al., 2013; Conroy, Sage, & Lambon Ralph, 2009a). Generally, there is a strong clinical belief that there is a lack of generalization to connected speech for standard naming therapies (Nickels, 2002; Wisenburn & Mahoney, 2009), yet typically studies (a) have lacked a systematic method for assessing generalization and (b) have been underpowered.

One variable which may be critical in determining whether words retrieved in isolation can also be used in fluent speech is naming latency (Conroy, Sage, & Lambon Ralph, 2009b; Crerar, 2004). As noted above, connected speech is highly demanding in terms of speed and accuracy. As such, we hypothesise that retrained vocabulary will only generalise if it can be retrieved within the demanding time window required by connected speech. Indeed, this hypothesis aligns with the broader observations that (a) naming speed is an important variable, not only within assessment but also in treatment tasks (McCall, Cox, Shelton, & Weinrich, 1997) and (b) for many people with mild aphasia, their expressive vocabulary may be largely recovered except for delayed naming latencies (Crerar, 2004). In fact, it has been noted that the recovery patterns of aphasic individuals often end with identical accuracy to

neurotypical participants (Kertesz & McCabe, 1977), yet the time needed to complete the tasks may be considerably longer than that of healthy controls (Neto & Santos, 2012).

In order to tackle this critical clinical need, the current study contained various novel developments and hypotheses. We developed a novel treatment to reduce speed and increase picture naming accuracy, simultaneously ('repeated, increasingly-speeded production'; RISP). This intervention was directly compared to (a) a standard treatment that targeted accuracy alone and (b) no treatment. We hypothesized that (i) RISP would generate greater improvements in both speed and accuracy of confrontation naming and (ii) that speedier naming would significantly increase the likelihood of producing treated words in post-treatment connected speech samples (evaluated through a newly-developed, systematic method). Finally, we related the patients' variable therapy outcomes to both their background neuropsychological profiles and the distributions of the underlying lesions.

Materials and methods

Participants

Twenty participants (11 males, 9 females; mean age 65.8 years, SD = 12.34) with a clinical diagnosis of chronic aphasia following cerebrovascular accident (CVA) took part in the study. All were recruited from aphasia support groups and similar services in Greater Manchester and North-West England. Participants had aphasia of varying degrees of severity and subtype. All were right handed, native English speakers, who had sustained one left hemisphere stroke (ischaemic or haemorrhagic) at least one year prior to the recruitment to the study and had no contradistinctions for MRI scanning (i.e. no pacemakers, metal implants, claustrophobia, etc.). Neuroimaging data from a healthy age and education matched control group (8 female, 11 male) was used in order to determine abnormal voxels using the automated lesion identification procedure (Seghier et al., 2008). All participants gave written informed consent with approval from the local ethics committee.

Prerequisites for participating in the study were to have normal or corrected-to-normal hearing and vision, as well as minimal repetition skills, where the latter was judged to be above 40% on an immediate word repetition test (PALPA 9: (Kay, Lesser, & Coltheart, 1996). Potential participants with co-existing neurological impairments (e.g. dementia or

multiple sclerosis), global aphasia, severe perceptual problems, or with very severe naming difficulties (below 8% or 5/60 on the Boston Naming Test: (Goodglass, Kaplan, & Barresi, 2000), were excluded from the study. All other levels and types of aphasic participants were included because we wanted to obtain a broad clinical sample so that the newly-developed therapy could be trialled in a range of patients (severe – moderate – mild anomia).

Table 1 about here

Demographic details of the participants are given in Table 1 together with baseline picture naming accuracy and speed. Scores representing each participant's accuracy in producing the same vocabulary items in composite picture descriptions are also reported in Table 1 (with participants ordered according to their baseline naming accuracy).

Background assessments

Before taking part in this study, participants also completed extensive linguistic and cognitive assessment. The results are summarised in Table 2. These measures were used to relate the participants' therapy outcomes to their background neuropsychological and aphasiological profiles.

Table 2 about here

The background assessment battery included the following specific tests. The Boston Naming Test (BNT) (Goodglass et al., 2000) was used to assess the wide range of word retrieval skills across the participants. Four repetition tasks were used (Psycholinguistic Assessments of Language Processing in Aphasia, PALPA 9: Kay et al., 1996): (a) word repetition immediate; (b) word repetition delayed; (c) non-word repetition immediate; (d) non-word repetition delayed. Two other phonological tasks included word minimal pairs and non-word minimal pairs (PALPA 2 & PALPA 1 respectively: Kay et al., 1996). Participants also completed six tests of comprehension and semantic memory: (a) spoken sentence comprehension from the Comprehensive Aphasia Test (CAT: (Swinburn, Porter, & Howard, 2004); (b) Synonym Judgement Test (Jefferies, Patterson, Jones, & Lambon Ralph, 2009); and from the Cambridge Semantic Battery (Bozeat, Lambon Ralph, Patterson, Garrard, & Hodges, 2000):

(c) picture naming; (d) spoken word-to-picture matching; (d) written word-to-picture matching; and (e) the picture-version of the Camel and Cactus Test (CCT) of semantic association knowledge. To test short-term memory skills, the forward memory span and the backward memory span assessments were administered (Wechsler, 1945). Two standard cognitive-executive tests were also completed: (a) Brixton Spatial Rule Anticipation Test (Burgess & Shallice, 1997) and Raven's Coloured Progressive Matrices (Raven, 1962). Following previous studies, we utilised principal component analysis to express the underling dimensions of variation in performance across these background behavioural measures (Butler, Lambon Ralph, & Woollams, 2014; Halai, Woollams, & Lambon Ralph, 2017). A PCA with varimax rotation was calculated for these behavioural measures for our full N=70 chronic aphasia patient dataset of which these N=20 therapy patients were a subset. Four principal components with an eigenvalue>1 were extracted; these corresponded to phonological, semantic, executive and speech quanta dimensions (see Halai et al., 2017 for the details of these principal components and their lesion correlates). The factor scores on these four dimensions for the subset of 20 therapy patients were used in the correlation analyses in order to relate therapy outcome to the patients' background assessment profiles.

Table 3 about here

Acquisition of Neuroimaging data

High resolution structural T1-weighted Magnetic Resonance Imaging (MRI) scans were acquired on a 3.0 Tesla Philips Achieva scanner (Philips Healthcare, Best, The Netherlands) using an 8-element SENSE head coil. A T1-weighted inversion recovery sequence with 3D acquisition was employed, with the following parameters: TR (repetition time) = 9.0 ms, TE (echo time) = 3.93 ms, flip angle = 8°, 150 contiguous slices, slice thickness = 1 mm, acquired voxel size 1.0 x 1.0 x 1.0 mm^3, matrix size 256 x 256, FOV = 256 mm, TI (inversion time) = 1150 ms, SENSE acceleration factor 2.5, total scan acquisition time =575 s.

Analysis of Neuroimaging data

Structural MRI scans were pre-processed with Statistical Parametric Mapping software (SPM8: Wellcome Trust Centre for Neuroimaging, http://www.fil.ion.ucl.ac.uk/spm/). The images were normalised into standard Montreal Neurological Institute (MNI) space using a

modified unified segmentation-normalisation procedure optimised for focal lesioned brains (Seghier, Ramlackhansingh, Crinion, Leff, & Price, 2008). Data from all participants with stroke aphasia and all healthy controls were entered into the segmentation-normalisation. This procedure combines segmentation, bias correction and spatial normalisation through the inversion of a single unified model (Ashburner & Friston, 2005). In brief, the unified model combines tissue class (with an additional tissue class for abnormal voxels), intensity bias and non-linear warping into the same probabilistic models that are assumed to generate subject-specific images. The lesion of each patient was automatically identified using an outlier detection algorithm, compared to a group of healthy controls, based on fuzzy clustering. The default parameters were used apart from the lesion definition 'U-threshold', which was set to 0.5 to create a binary lesion image. We modified the U-threshold from 0.3 to 0.5 after comparing the results obtained from a sample of patients to what would be nominated as lesioned tissue by an expert neurologist. The images generated for each patient were individually checked and visually inspected with respect to the original scan, and were used to create the lesion overlap map in Figure 1 ($2mm^3$ MNI voxel size). We should note here, explicitly, that although commonly referred to as an automated 'lesion' segmentation method, the technique detects areas of unexpected tissue class – and, thus, identifies missing grey and white matter but also areas of augmented CSF space. We used the T1-weighted images with continuous signal intensity values across the whole brain and correlated these with magnitude of the RISP effect using a voxel-based correlational methodology (VBCM) (Tyler, Marslen-Wilson, & Stamatakis, 2005), a variant of voxel symptom lesion mapping (VSLM) (Bates et al., 2003). VBCM does not require a binary classification of the intact/lesioned brain to be marked, as in the case of VSLM, as both the behaviour and tissue concentration measures are treated as continuous variables (conducted in SPM8). All anatomical labels were based on the Harvard-Oxford atlas in MNI space.

Figure 1 about here

Therapy Methods

Stimuli

As noted in the Introduction, one reason for the lack of information in the literature with regard to generalization from naming therapy to connected speech, is the lack of a systematic assessment method (Maendl, 1998). In order to measure word retrieval for target items in

both picture naming and a connected speech task, various picture stimuli were adapted for the study. Four very detailed composite multi-event pictures were selected (from the "Where's Wally/Waldo?" publications). These contained detailed depictions of hundreds of items and events (e.g., animals, objects and events at a busy zoo or fairground) from which a small minority of target items were selected. To assess and treat confrontational naming for these targets, we selected new unrelated pictures of the same exemplars (the pictures presented each exemplar singly and without any background context). Piloting with both younger and older control participants, without brain injury, was undertaken to assess the control participants' likelihood of word retrieval for the target vocabulary in the single and composite pictures. The target stimuli (120 items from across the four composite pictures) were all nouns, selected to meet the following criteria: (a) the items tended to be named spontaneously in the control participants' picture descriptions – specifically, the items' names had to be produced spontaneously by more than 3/10 participants (note – when describing each picture, participants and patients were asked to describe freely "what is going on in the picture" and were not directed to any areas or items within the scene in order to assess "spontaneous" connected speech); (b) items from the composite scene could be depicted as an individual, new picture of the item with 100% name agreement (thus pictureable nouns such as *bench* rather than *water* or actions were selected); (c) selected items had no synonyms or no equally frequent synonyms, which were used alternatively in the picture description task (e.g., *dodgems* and *bumper cars*).

From these 120 nouns, four matched sets of 20 items were selected: two sets were allocated to the treatment conditions, i.e., 20 items for standard word-cue treatment (standard production – 'SP') and 20 items for combined speed+accuracy treatment (repeated, increasingly-speeded production – 'RISP'). The remaining two sets served as untreated control items (thus controlling for any non-specific effects, including the small boost in performance that can result from repeated assessment (Nickels, 2002a). One treatment set and its paired untreated set related to a subset of the vocabulary appearing in two of the four composite pictures. Likewise, the other treatment set and its paired control related to the remaining two composite pictures. This allowed us to separate the effects of each therapy by avoiding target vocabulary for the two treatments appearing in the same composite picture. The allocation of picture sets to the two treatments was counterbalanced across participants.

The word sets were matched (using the Match program: (Van Casteren & Davis, 2007) for (a) the likelihood of retrieval in the spontaneous picture descriptions by the older age- and education-matched control participants; (b) for frequency from the British National Corpus (BNC Consortium, 2007); and (c) phoneme length.

Baseline and post-therapy assessment

Baseline performance for the four composite picture descriptions and 80-item confrontational naming test were assessed twice before therapy commenced (across four separate assessment sessions with composite description assessed before the confrontational naming). There was no significant change in performance across the two baseline assessments (confirming a stable baseline) and thus we used the first assessment results to compare the post-therapy results to. Post-therapy performance was assessed for composite picture description then picture naming ability at one-week and again at one-month to establish the longer-term benefits of the therapy (no maintenance or practice regimes were used post therapy). An additional, fifth composite picture description was also assessed before and after therapy. No target vocabulary from this fifth picture was included in the therapy as treated or untreated items. As such, the fifth picture acted as a control for the target composite pictures in order to control for non-specific improvements that might arise in connected speech simply from repeated assessment.

For the picture naming assessment, participants were presented with all 80 items in random order. Pictures were presented using E-Prime software (Schneider, Eschman, & Zuccolotto, 2002) such that each picture was presented simultaneously with an auditory beep and remained on the screen for a maximum of ten seconds. Audacity software was used to measure naming latencies by measuring the time elapsed from the beep to the onset of the participant's correct response. When no correct name was produced, the reaction time for that trial was treated as missing data.

To elicit connected speech samples, participants were informed that they were going to see four 'busy' pictures, one at a time on a computer screen. They were asked to describe what they saw in each picture in as much detail as they could for about 5-10 minutes. Participants' responses were digitally audio-recorded (Sony Digital IC Recorder). The order of

presentation of each picture was randomised across participants, thus counterbalancing any effect of relative difficulty in the pictures described first.

Treatments

Figure 2 about here

The treatments were delivered in two phases (see Figure 2), each phase containing two sessions per week for three weeks (six treatment sessions per phase). In the first phase, only standard therapy was administered for all items (n=40) in order to boost naming accuracy before introducing any speed requirement. In the second phase, standard (accuracy-only) treatment was continued for one set, whilst the other was treated with RISP (see below). In both phases, stimuli in each set were randomised and the order of sets was counterbalanced across sessions.

***Standard treatment*:** This was a standard increasing cues, naming therapy, which aimed to improve participants' picture naming accuracy only. Participants were asked to name each picture, presented on a computer screen, in 10 seconds without support, i.e. with no cues. After each naming attempt, feedback was provided both verbally by the experimenter and presented in writing on the screen. Initially, minimal cues were provided (e.g. the initial consonant and vowel of the target word, e.g. "wi" for 'window') but the cues were increased if naming was not achieved (next most of word, e.g. "wind" for 'window', and then the whole word 'window'). Participants worked through all therapy items three times per session.

Repeated, Increasingly-Speeded Production (RISP treatment): This treatment was a hybrid intervention that combined cued naming with the deadline naming method used in experimental psycholinguistics research to assess the effect of speeded naming on speech production (Hodgson & Lambon Ralph, 2008; Vitkovitch & Humphreys, 1991). Participants were instructed that the computer would present the picture for a limited amount of time and their task was to try to name the picture before the beep at the end of the stimulus presentation. In each therapy session the presentation duration/time-to-the-beep was reduced (see below). Specifically, during each trial, the target picture was presented on the computer screen for a fixed time. At the end of the allotted time, the picture disappeared and a beep

sound was produced by the computer. A blank screen was then displayed for 1000msec. Participants were then presented with the written target word on the screen as feedback and the correct spoken name of the picture was played by the computer. In the case of an incorrect response, the participant was asked to repeat the correct name after the computer/experimenter three times. Participants cycled through all therapy items three times per session.

The naming deadline was shortened in a controlled fashion across the six RISP sessions. The initial picture exposure time was set to the mean of all participants' baseline picture naming speed (3 seconds). This ensured that each participant's first 'speeded' naming attempt would feel reasonably natural. The ultimate target deadline in the 6th RISP session was 1 second, which matched the mean naming speed of elderly neurotypical participants (mean naming time: 1002 msec). Following the first RISP session, the target naming speed was reduced until the goal of 1 second was reached. This was implemented in a systematic way with relatively larger time reductions in early sessions and smaller time reductions in later sessions. Specifically, the beep interval for session 1 of RISP was 3 seconds, session 2 = 2.5 seconds, session 3 = 2 seconds, session 4 = 1.6 seconds, session 5 = 1.3 seconds, and session 6 = 1 second. The same target naming speed was used for the three cycles of each session and only reduced on the start of the next session.

Scoring

Participant's performance was scored based on their first response for all picture naming. Self-corrections were considered correct if the correct name was produced immediately (i.e. in less than two seconds) after the first response. To measure picture naming latencies, the time between the first standard beep sound and participant's correct response was calculated using Audacity software.

Results

Picture naming accuracy

A global 3×4 ANOVA indicated that there was a main effect of time ($F(2,38) = 55.6$, $p < .0005$), a main effect of treatments ($F(3,57) = 35.7$, $p < .0005$), and a significant interaction between time and treatments ($F(6,114) = 18.0$, $p < .0005$; see Figure 3a) – indicating very different effects of therapy on the treated and untreated items. In comparing each treatment condition to the control sets, a 2×3 ANOVA showed that both therapies generated significantly improved accuracy scores relative to their control sets (significant interaction: $p < .0005$ for both therapies). In directly comparing the two treatments, a 2×3 ANOVA indicated that there was a trend towards a borderline interaction between time and treatment: $F(2,38) = 2.3$, $p = .117$. Both therapies significantly increased picture naming accuracy between the baseline and post-treatment assessments ($p < .0005$). Additional pairwise analyses showed that the RISP therapy effect was significantly greater than SP not only at the 1week post-treatment assessment ($p < .0005$), but also at the follow-up (1 month) assessment ($p = .001$).

Figures 3 about here

Picture naming speed

The naming speed for correctly named items was analysed with a 3×4 ANOVA (Figure 3b). There was a main effect of 'Time Point': $F(2,36) = 21.1$, $p < .0005$, no main effect of 'Treatment' factor [$F(3,54) = 1.7$, $p = .174$], but a significant interaction between 'Time Point' and 'Treatment' [$F(6,108) = 5.7$, $p < 0.0005$] – indicating significantly different changes in naming speed for the treated vs. untreated sets. Follow-up 2×3 ANOVAs confirmed that the effect of each therapy was significantly different from its control [Time Point × Set interactions were significant: RISP $F(2,36)=8.6$, $p=0.001$; SP $F(2,36)=3.9$, $p=0.03$]. A 2×3 ANOVA comparing the two treated sets indicated that there was a significant interaction between 'Time Point' and 'Treatment' [$F(2,36) = 3.2$, $p = .05$]. Whilst both treatments significantly reduced picture naming latencies between the baseline and both post-treatment assessments (1 week and 1 month), additional pairwise analyses showed that there was a trend for the RISP treatment to reduce RTs more than SP from baseline to the immediate assessment (1w, $p = .101$) and, most strikingly, RISP was significantly more

effective in maintaining the treatment effect in terms of quicker naming responses at the one month follow-up assessment (p = .001). In comparing the two untreated conditions, only the main effect of the 'Time Point' factor was significant (F(2,36) = 3.23, p = .05) – reflecting a small reduction in naming latencies across repeated assessments (presumably reflecting repetition priming). The main effect of 'Set' was not significant (F(1,18) < 1), nor was the interaction between 'Time Point' and 'Set' (F(2,36) = 1.3, p = 0.28).

Generalisation to connected speech: Word retrieval in composite picture descriptions

A 3×4 ANOVA indicated that there was a significant effect of the 'Time Point' factor [F(2,38) = 87.8, p < .0005], a main effect of 'Treatment' factor [F(3,57) = 43.7, p < .0005] and a highly significant interaction between 'Time Point' and 'Treatment' [F(6,114) = 19.9, p < .0005; (Figure 3c)] – indicating very different production of the target vs. untreated vocabulary in the patients' narratives before and after therapy. Directly comparing the two treatments, a 2×3 ANOVA indicated that there was a highly significant interaction between 'Time Point' and 'Treatment' [F(2,38) = 19.6, p < .0005]. Additional pairwise analyses showed that the RISP effect on connected speech production was significantly stronger than SP both at the 1 week and 1 month post-treatment assessments (both p < .0005). Comparing each treatment to its control set, separately, we found significant 'Time Point' × 'Set' interactions for the RISP and SP sets [F(2,38)=19.6, p<0.0005; F(2,38)=5.2, p=0.01, respectively]. Thus, although there is a general clinical belief that standard therapy does not induce generalisation to connected speech, our newly-developed assessment was able to demonstrate that this is incorrect – there is, in fact, a small but significant generalisation to connected speech for SP both at one week and one month (though the effect was significantly smaller than for the RISP therapy – see above). Finally, the two untreated conditions were compared. The main effect of 'Time Point' was significant [F(2,36) = 3.2, p = .05], indicating a small improvement in target vocabulary production simply through repeated assessment, but neither the main effect of 'Set' [F(1,18) < 1] nor the interaction between 'Time Point' and 'Set' were significant [F(2,36) = 1.3, p = 0.28).

Content analysis of the connected speech samples

As well as exploring the generalization of trained vocabulary to the connected speech samples, it is also important to investigate the connected speech samples more generally. It is

possible, for example, that improved vocabulary promotes connected speech more generally or that the improvement on the trained items comes at the cost of reduced performance on the untrained vocabulary. To do so, we examined the connected speech samples in terms of the total number of nouns produced (tokens), the number of unique nouns produced (types), nouns per minute, the type/token ratio, average word frequency and average imageability for the treated and untreated pictures.

In summary, the overall secondary effects on the patients' connected speech samples were entirely positive. Specifically, for the treated pictures, the speech samples including all items showed that significantly more unique items were produced after therapy compared to baseline (mean at 1 week = 103.6, mean at baseline = 85.4; t(18) = -2.30, p = 0.03). There was also a significant decrease in the average word frequency of the nouns used (mean at 1 week = 1.40, mean at baseline = 1.55; $t(18) = 4.21$, $p < 0.001$). There were no significant changes found in terms of the number of nouns produced per minute, the type/token ratio, and average imageability rating. Importantly, there were no significant effects found in similar analyses for the untreated picture indicating that the improved connected speech samples did not reflect a non-specific effect of repeated assessment.

This first analysis included all items, including the target therapy items. Accordingly, we repeated the analysis to remove these items from consideration. In this second analysis, the increase in unique items from baseline to post therapy was no longer significant (mean at 1 week = 84.8, mean at baseline = 77.8; t(18) = -0.95, p = 0.3). The reduction in mean word frequency, however, was still significant (mean at 1 week = 1.48, mean at baseline = 1.58; t(18) = 2.86, p < 0.01).

Correlations with individual's background neuropsychological profile

Although there were significant and reliable therapy effects at the group level, the effect varied across individual patients. We performed correlations between the background neuropsychological profile (with respect to four principal neuropsychological components (see Methods and Halai et al., 2017): phonological, semantic, executive, and verbal fluency ability) and the magnitude of the therapy effect (1 week vs. baseline performance, and 1 month (maintenance) vs. baseline performance) in order to reveal which aspects of the patients' profile were related to the therapy outcome. Overall, no correlations were found

between any of the factors and the outcome on the standard therapy. For the RISP therapy, however, a significant negative correlation was found between the patients' phonological factor scores and the magnitude of their therapy effect, at both 1 week (r = -.55, p < .01) and 1 month (r = -.61, p < .005). This demonstrates that patients with the poorest phonological abilities showed the largest RISP benefit. As can be seen across the case-series (Figure 4), this negative correlation seems to reflect the fact that the RISP therapy was particularly beneficial leading to a ceiling effect for many of the milder patients (note that if patient JS with poor phonological abilities but a large therapy effect is removed, then the correlation is still significant).

Figure 4 about here

It was also possible to determine how each factor correlated with the maintenance of the therapy effect (i.e., 1 month vs. 1 week performance). In this analysis, the maintenance of the RISP effect was found to correlate positively with performance on the executive tasks (r = .53, p < .01). Thus, the patients with better executive abilities exhibited the best therapy maintenance.

Neural correlates of RISP

In order to determine the neural correlates of the RISP effect we correlated each patient's therapy effect (1 week vs. baseline performance, and 1 month vs. baseline performance) with their T1-weighted MRI using voxel-based correlational methodology (VBCM: Tyler, Marslen-Wilson, & Stamatakis, 2005). This analysis revealed that patients with the greatest damage to the posterior superior temporal gyrus extending into the white matter of the inferior longitudinal fasciculus, showed the greatest RISP benefit both at 1 week and 1 month (height threshold p < .001, cluster corrected using FWE p < .05). This region is known to play an important role in phonological performance, as illustrated in Figure 5 whereby the RISP effect overlaps closely with the area related to the lesion correlate for the patients' phonological skill factor found previously by Halai et al. (2017) and thus aligns with the behavioural correlation between phonological ability and therapy effect noted above. It appears, therefore, that the RISP effect may relate particularly to the patients' phonological abilities. Finally, no voxels were found to correlate significantly with the RISP maintenance effect (1 month vs. 1 week performance).

Figure 5 about here

Discussion

Anomia is an immensely frustrating and disabling feature of aphasia, which is a common disorder post stroke (around 1/3 cases) and in other neurological conditions. Accordingly, it is important to establish effective interventions for remediating word-finding skills and generalising these improvements to patients' connected speech. Given the observation that fluent speech requires both quick and accurate word retrieval, we investigated and confirmed the novel hypothesis that a behavioural treatment, focussing on both speed and accuracy rather than accuracy alone (as is the case in standard methods), would generate greater improvements in both confrontation naming and also generalisation of this improved vocabulary to connected speech. A second key, novel feature of this study was that the interventions were not examined in isolation but we also investigated the neuropsychological and lesion correlates of treatment responsiveness. Although such analyses are a rarity in the literature to date (Abel, Weiller, Huber, Willmes, & Specht, 2015), increasing our understanding about both the neuropsychological and lesion correlates of variable therapy success will be a critical step towards future neuroscience-led stratification of patients and choice of clinical pathways.

To address these questions, we developed a novel naming treatment that focussed on both speed and accuracy (RISP), which we compared to a standard accuracy-only treatment (SP). As expected, both treatments increased picture naming accuracy with long lasting effects, i.e. participants were considerably and significantly more accurate in naming all treated items at the end of the treatment (SP and RISP; assessed one week following the end of the intervention). This increase in picture naming accuracy was also largely retained in the follow-up (one month) assessment even without maintenance practice. RISP was, however, significantly more effective than SP, in promoting increased accuracy in both post-treatment naming assessments, particularly at the important long-term follow-up assessment. The same pattern was found in naming speed – thus, as intended, RISP was much more effective in speeding successful name retrieval and maintaining these improvements at follow-up assessment. Perhaps most importantly, we found that RISP's combined, long-term improvements in both speed and accuracy, generalised from naming individual target items into the patients' connected speech – a "holy grail" for speech and language therapy.

With regard to neuropsychological and neural correlates of therapy effects, we found a significant negative correlation for the RISP therapy between the patients' degree of phonological impairment and the magnitude of their therapy effect, both immediately after therapy and at follow up assessment. This initially somewhat counter-intuitive finding probably reflects that RISP appears to be an especially beneficial treatment, such that milder patients show a resultant ceiling effect in their speech production assessment whereas the more severe patients can exhibit a much more dramatic improvement on the target items. This finding may also be consistent with the observation from Best and colleagues' (2013) meta-analysis that better treatment responsiveness was evident in participants classified as having relatively less semantic difficulties and greater phonological output deficits (note, our use of principal component analysis to extract the pattern of underlying language-cognitive deficits means that, over and above phonology per se, the potential additional influence of semantic, skills, speech fluency and cognitive-executive factors were already partialled out: see Butler et al., 2014; Halai et al., 2017). This behavioural correlate for the RISP therapy was also mirrored directly in the lesion correlate analysis: the RISP benefit was most evident in participants with the greatest damage to the posterior superior temporal gyrus extending into the white matter of the inferior longitudinal fasciculus. This region has been implicated in auditory-phonological processing not only through neuropsychological studies (Baldo, Katseff, & Dronkers, 2012; Robson, Grube, Lambon Ralph, Griffiths, & Sage, 2013; Robson, Sage, & Ralph, 2012) but also in fMRI explorations of healthy function (Hickok & Poeppel, 2004; Rauschecker & Scott, 2009; Warren & Griffiths, 2003). Finally, with regard to the long-term maintenance of the RISP treatment, follow-up performance correlated positively with cognitive-executive skills. Again, this finding has been observed in previous neuropsychological and neuroimaging studies, and consequently, it has been suggested that both patients' degree of language impairment and remaining executive skill may be critical in recovery of function and therapy outcome (Brownsett et al., 2014; Lambon Ralph, Snell, Fillingham, Conroy, & Sage, 2010; Sharp, Turkheimer, Bose, Scott, & Wise, 2010).

Two different hypotheses can be made about the mechanisms underlying the speeded treatment effect, which can be tested in future investigations. The first, language-specific hypothesis is related to the aim of the RISP treatment to target both accuracy and speed. For optimally easy and efficient word retrieval, the language system requires precise representations that allow the target meaning to be converted to phonological and motor-

speech representations (Lupker et al., 1997). Computational models of speech production and reading have repeatedly shown that as these representations and mappings are refined through learning, performance of models becomes both more accurate and more efficient (Ellis & Lambon Ralph, 2000; Plaut, McClelland, Seidenberg, & Patterson, 1996). Accordingly, because the RISP treatment deliberately aims beyond accuracy to improve speed as well, the language representations and mappings may have been pressured not only to reform but also to be 'sharpened up' to become more precise. Indeed, this hypothesis might also explain why, aside from speed, RISP led to significantly better naming accuracy than the accuracy-only focussed SP (following the fact that both speed and accuracy reflect the precision of the underlying language representations). Another possible hypothesis accounting for the RISP effect is related to a domain-general, cognitive-executive mechanism (Geranmayeh, Brownsett, & Wise, 2014; Lambon Ralph, Snell, Fillingham, Conroy, & Sage, 2010). Not only did we find that the degree of treatment maintenance was related to the patients' cognitive-executive skills, but all participants reported that RISP was especially engaging and motivating (indeed, all participants, irrespective of severity, reported positive engagement in the RISP task). Accordingly, RISP may have been much better than SP in engaging the patients' executive and attentional skills, in addition to the speech production system, resulting in improved learning and retention. From a neurobiological perspective, increased motivation and reward-seeking behaviour has been being strongly associated with dopamine release (Fiorillo, 2013; Morita et al., 2013; Foerde et al., 2015) and dopamine has been associated with improved learning and therapy effects (Berthier & Pulvermüller, 2011; Gill & Leff, 2012). This observation speaks to the wider potential of 'gamification', that is utilising the dynamic and engaging aspects of commercial gaming software to ramp up the engagement required for rehabilitation tasks. Such principles are well established in physical therapies after stroke, such as for upper limb function or balance (Ferreira, Guimarães, Santos, & Sousa, 2014), but have yet to influence the language/cognitive rehabilitation domains.

References:

Abel, S., Weiller, C., Huber, W., Willmes, K., & Specht, K. (2015). Therapy-induced brain reorganization patterns in aphasia. *Brain*, *138*(4), 1097–1112. https://doi.org/10.1093/brain/awv022

Ashburner, J., & Friston, K. J. (2005). Unified segmentation. *NeuroImage*, *26*(3), 839–851. https://doi.org/10.1016/j.neuroimage.2005.02.018

Baldo, J. V., Katseff, S., & Dronkers, N. F. (2012). Brain Regions Underlying Repetition and Auditory-Verbal Short-term Memory Deficits in Aphasia: Evidence from Voxel-based Lesion Symptom Mapping. *Aphasiology*, *26*(3–4), 338–354. https://doi.org/10.1080/02687038.2011.602391

Bates, E., Wilson, S. M., Saygin, A. P., Dick, F., Sereno, M. I., Knight, R. T., & Dronkers, N. F. (2003). Voxel-based lesion-symptom mapping. *Nature Neuroscience*, *6*(5), 448–450. https://doi.org/10.1038/nn1050

Berthier, M. L., & Pulvermüller, F. (2011). Neuroscience insights improve neurorehabilitation of poststroke aphasia. *Nature Reviews. Neurology*, *7*(2), 86–97. https://doi.org/10.1038/nrneurol.2010.201

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. https://doi.org/10.1016/j.cortex.2013.01.005

Bird, H., Lambon Ralph, M. A., Patterson, K., & Hodges, J. R. (2000). The rise and fall of frequency and imageability: noun and verb production in semantic dementia. *Brain and Language*, *73*(1), 17–49. https://doi.org/10.1006/brln.2000.2293

BNC Consortium. (2007). *The British National Corpus, version 3 (BNC XML Edition)*. Oxford, England: Oxford University Computing Services.

Bozeat, S., Lambon Ralph, M. A., Patterson, K., Garrard, P., & Hodges, J. R. (2000). Non-verbal semantic impairment in semantic dementia. *Neuropsychologia*, *38*(9), 1207–1215. https://doi.org/10.1016/S0028-3932(00)00034-8

Burgess, P., & Shallice, T. (1997). *Hayling and Brixton Tests*. Oxford, England: Pearson.

Butler, R. A., Lambon Ralph, M. A., & Woollams, A. M. (2014). Capturing multidimensionality in stroke aphasia: mapping principal behavioural components to neural structures. *Brain: A Journal of Neurology*, *137*(Pt 12), 3248–3266. https://doi.org/10.1093/brain/awu286

Conroy, P., Sage, K., & Lambon Ralph, M. (2009a). 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. https://doi.org/10.1080/13682820802585975

Conroy, P., Sage, K., & Lambon Ralph, M. A. L. (2009b). The effects of decreasing and increasing cue therapy on improving naming speed and accuracy for verbs and nouns in aphasia. *Aphasiology*, *23*(6), 707–730. https://doi.org/10.1080/02687030802165574

Crerar, M. A. (2004). Aphasia rehabilitation and the strange neglect of speed. *Neuropsychological Rehabilitation*, *14*(1–2), 173–206. https://doi.org/10.1080/09602010343000174

Ellis, A., & Lambon Ralph, M. A. (2000). Age of Acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems: Insights from connectionist networks. *JEP:LMC*, *26*(5), 1103–1123.

Ferreira, C., Guimarães, V., Santos, A., & Sousa, I. (2014). Gamification of Stroke Rehabilitation Exercises Using a Smartphone. In *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare* (pp. 282–285). ICST, Brussels, Belgium, Belgium: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering). https://doi.org/10.4108/icst.pervasivehealth.2014.255326

Geranmayeh, F., Brownsett, S. L. E., & Wise, R. J. S. (2014). Task-induced brain activity in aphasic stroke patients: what is driving recovery? *Brain*, *137*(10), 2632–2648. https://doi.org/10.1093/brain/awu163

Gill, S. K., & Leff, A. P. (2012). Dopaminergic therapy in aphasia. *Aphasiology*, *28*(2), 155–170. https://doi.org/10.1080/02687038.2013.802286

Goodglass, H., Kaplan, E., & Barresi, B. (2000). *Boston Diagnostic Aphasia Examination - Third Edition (BDAE-3)*. Pearsons.

Halai, A. D., Woollams, A. M., & Lambon Ralph, M. A. (2017). Using principal component analysis to capture individual differences within a unified neuropsychological model of chronic post-stroke aphasia: Revealing the unique neural correlates of speech fluency, phonology and semantics. *Cortex*, *86*, 275–289. https://doi.org/10.1016/j.cortex.2016.04.016

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. https://doi.org/10.1080/02687030701262613

Hodgson, C., & Lambon Ralph, M. (2008). Mimicking aphasic semantic errors in normal speech production: Evidence from a novel experimental paradigm. *Brain and Language*, *104*(1), 89–101.

Jefferies, E., Patterson, K., Jones, R. W., & Lambon Ralph, M. A. (2009). Comprehension of concrete and abstract words in semantic dementia. *Neuropsychology*, *23*(4), 492–499. https://doi.org/10.1037/a0015452

Kay, J., Lesser, R., & Coltheart, M. (1996). Psycholinguistic assessments of language processing in aphasia (PALPA): An introduction. *Aphasiology*, *10*(2), 159–180. https://doi.org/10.1080/02687039608248403

Kertesz, A., & McCabe, P. (1977). Recovery patterns and prognosis in aphasia. *Brain: A Journal of Neurology*, *100 Pt 1*, 1–18.

Laine, M., & Martin, N. (2006). *Anomia: Theoretical and clinical aspects.* Hove, England: Psychology Press.

Lambon Ralph, M. A., Snell, C., Fillingham, J. K., Conroy, P., & Sage, K. (2010). Predicting the outcome of anomia therapy for people with aphasia post CVA: both language and cognitive status are key predictors. *Neuropsychological Rehabilitation*, *20*(2), 289–305. https://doi.org/10.1080/09602010903237875

Levelt, W. (1989). *Speaking: from intention to articulation*. Cambridge, Mass.: MIT Press.

Maendl, L. (1998). *Word finding difficulties in aphasia.* University of York.

McCall, D., Cox, D. M., Shelton, J. R., & Weinrich, M. (1997). The influence of syntactic and semantic information on picture-naming performance in aphasic patients. *Aphasiology*, *11*(6), 581–600. https://doi.org/10.1080/02687039708248491

Neto, B., & Santos, M. E. (2012). Language after aphasia: Only a matter of speed processing? *Aphasiology*, *26*(11), 1352–1361. https://doi.org/10.1080/02687038.2012.672023

Nickels, L. (2002). Therapy for naming disorders: Revisiting, revising, and reviewing. *Aphasiology*, *16*(10–11), 935–979. https://doi.org/10.1080/02687030244000563

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103*, 56–115.

Raven, J. (1962). *Raven's Coloured Progressive Matrices (CPM)*. Oxford, England: Pearson.

Robson, H., Grube, M., Lambon Ralph, M. A., Griffiths, T. D., & Sage, K. (2013). Fundamental deficits of auditory perception in Wernicke's aphasia. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *49*(7), 1808–1822. https://doi.org/10.1016/j.cortex.2012.11.012

Robson, H., Sage, K., & Ralph, M. A. L. (2012). Wernicke's aphasia reflects a combination of acoustic-phonological and semantic control deficits: a case-series comparison of Wernicke's aphasia, semantic dementia and semantic aphasia. *Neuropsychologia*, *50*(2), 266–275. https://doi.org/10.1016/j.neuropsychologia.2011.11.021

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime Computer Software and Manual*. Pittsburg, PA: Psychology Software Tools Inc.

Seghier, M. L., Ramlackhansingh, A., Crinion, J., Leff, A. P., & Price, C. J. (2008). Lesion identification using unified segmentation-normalisation models and fuzzy clustering. *Neuroimage*, *41*(4–3), 1253–1266. https://doi.org/10.1016/j.neuroimage.2008.03.028

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

The Stroke Association (UK). (2016). State of the Nation Stroke statistics January 2016.

Tyler, L. K., Marslen-Wilson, W., & Stamatakis, E. A. (2005). Dissociating neuro-cognitive component processes: voxel-based correlational methodology. *Neuropsychologia*, *43*(5), 771–778. https://doi.org/10.1016/j.neuropsychologia.2004.07.020

Van Casteren, M., & Davis, M. H. (2007). Match: a program to assist in matching the conditions of factorial experiments. *Behavior Research Methods*, *39*(4), 973–978.

Vitkovitch, M., & Humphreys, G. W. (1991). Perseverant responding in speeded naming of pictures: It's in the links. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(4), 664–680. https://doi.org/10.1037/0278-7393.17.4.664

Wechsler, D. (1945). A Standardized Memory Scale for Clinical Use. *The Journal of Psychology*, *19*(1), 87–95. https://doi.org/10.1080/00223980.1945.9917223

Wisenburn, B., & Mahoney, K. (2009). A meta-analysis of word-finding treatments for aphasia. *Aphasiology*, *23*(11), 1338–1352. https://doi.org/10.1080/02687030902732745

Table 1 Participant Demographic Details including baseline scores and aphasia subtypes

Participant	Age (years)	Gender	Handedness	Education (years)	Time post-stroke (years)	Baseline Naming Accuracy (max=80)	Baseline Naming Speed (in msec.)	Baseline Description Accuracy (max=80)	BDAE Classification
JBo	79	male	right	13	4	9	3.862	5	Broca
KS	63	male	right	12	5	12	3.411	6	TSA
MD	74	male	right	11	3	12	2.553	6	MN
AD	77	female	right	11	5	15	2.103	4	Broca
AB	52	male	right	13	7	28	4.475	14	Anomia
JW	83	male	right	10	2	28	2.620	3	Broca
JSc	79	male	right	12	10	29	2.274	5	Broca
JS	70	female	right	19	4	31	2.625	13	Anomia
KAd	69	male	right	11	3	33	2.376	14	Anomia
DF	52	female	right	11	6	35	3.581	16	Anomia
GP	60	male	right	11	4	38	3.350	10	Anomia
PR	73	female	right	11	4	39	2.989	3	Broca
EB	53	female	right	11	4	43	2.373	9	Anomia
JSo	60	male	right	11	14	45	3.495	11	TMA
RL	84	male	right	11	3	45	2.675	24	Anomia
CH	44	female	right	13	3	49	2.342	17	TMA
JBr	69	female	right	16	3	51	2.679	19	Anomia
DCS	49	female	right	12	5	52	1.911	13	Broca
DM	52	male	right	17	7	55	2.986	25	Broca
JM	74	female	right	11	10	58	1.945	23	Anomia

TSA = transcortical sensory aphasia, MN = mixed non-fluent, TMA = transcortical motor aphasia

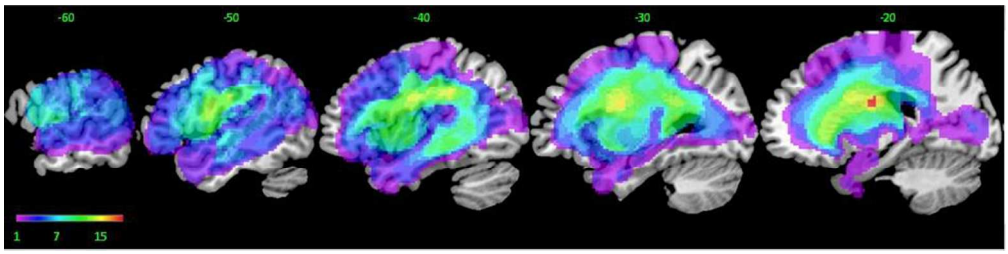
Table 2 Participants' performance on language, semantic and cognitive assessments

Participant	Boston Naming Test	Word Repetition: Immediate	Word Repetition: Delayed	Non-word Repetition: Immediate	Non-word Repetition: Delayed	Word Minimal Pairs	Non-word Minimal Pairs	CAT Spoken Sentence Comprehension	CBU 64 Item Naming
KS	**11.67**	**91.25**	93.75	73.33	80.00	95.83	94.44	**84.38**	**26.56**
JBo	**13.33**	**43.75**	**35.00**	**23.33**	**6.67**	93.06	95.83	**34.38**	**32.81**
AD	**18.33**	**57.50**	**55.00**	**23.33**	**13.33**	93.06	95.83	**68.75**	**37.50**
MD	**26.67**	**50.00**	**61.25**	**26.67**	**16.67**	98.61	98.61	**12.50**	**50.00**
AB	**28.33**	**86.25**	**63.75**	**26.67**	**13.33**	87.50	80.56	**75.00**	**54.69**
EB	**38.33**	**81.25**	**78.75**	**66.67**	**36.67**	95.83	93.06	**75.00**	**76.56**
JW	**40.00**	**65.00**	**66.25**	**33.33**	**16.67**	81.94	86.11	90.63	**64.06**
DCS	**43.33**	**70.00**	**68.75**	**40.00**	56.67	97.22	97.22	93.75	**68.75**
JS	**45.00**	**90.00**	88.75	**50.00**	**46.67**	93.06	**75.00**	**81.25**	**81.25**
JSo	**45.00**	93.75	x	80.00	x	97.22	97.22	**81.25**	x
DF	**46.67**	93.75	**41.25**	**53.33**	**10.00**	95.83	90.28	**62.50**	**78.13**
PR	**46.67**	**85.00**	91.25	**56.67**	**43.33**	94.44	80.56	87.50	**60.94**
JSc	**51.66**	**90.00**	91.25	**36.67**	63.33	86.11	**75.00**	**75.00**	**73.44**
GP	**55.00**	95.00	95.00	**43.33**	**46.67**	95.83	98.61	**78.13**	**64.06**
KAd	**55.00**	96.25	96.25	83.33	70.00	95.83	95.83	**78.13**	**76.56**
CH	**56.67**	**92.50**	88.75	**60.00**	**40.00**	97.22	95.83	**84.38**	**81.25**
RL	**63.33**	**61.25**	**53.75**	**13.33**	**16.67**	**59.72**	**56.94**	**62.50**	**84.38**
DM	**70.00**	**73.75**	**68.75**	**53.33**	**10.00**	93.06	80.56	**56.25**	**73.44**
JBr	81.67	97.50	95.00	90.00	80.00	97.22	90.28	96.88	**89.06**
JM	81.67	96.25	96.25	93.33	63.33	95.83	90.28	**71.88**	96.88

Table 3 Participants' performance on further language, semantic and cognitive assessments

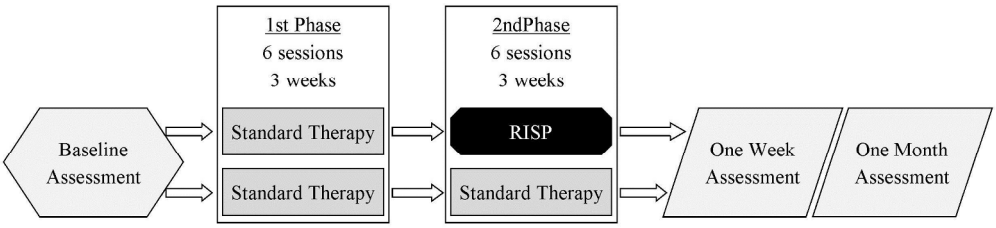
Participant	Spoken Word to Picture Matching	Written Word to Picture Matching	96 Synonym Judgement	Camel & Cactus Test: Pictures	Forward Digit Span	Backward Digit Span	Brixton Spatial Rule Anticipation Test	Raven's Coloured Progressive Matrices
KS	**71.88**	**67.19**	**84.38**	**68.75**	100.00	50.00	52.73	86.11
JBo	100.00	**96.88**	**65.63**	**73.44**	62.50	**37.50**	60.00	77.78
AD	96.88	98.44	**83.33**	**85.94**	75.00	**37.50**	**30.91**	63.89
MD	96.88	**93.75**	**57.29**	**59.38**	**37.50**	**25.00**	58.18	38.89
AB	**95.31**	98.44	**75.00**	**79.69**	**37.50**	**25.00**	88.89	88.89
EB	100.00	98.44	**83.33**	90.63	**50.00**	**25.00**	47.27	66.67
JW	96.88	98.44	**85.42**	**81.25**	87.50	**25.00**	61.82	88.89
DCS	100.00	98.44	91.67	95.31	62.50	**49.99**	81.82	100.00
JS	100.00	100.00	96.88	95.31	**50.00**	50.00	65.45	100.00
JSo	100.00	X	92.71	x	**50.00**	**25.00**	50.91	91.67
DF	100.00	**96.88**	**78.13**	92.19	**37.50**	**25.00**	**43.64**	88.89
PR	100.00	100.00	**83.33**	**84.38**	75.00	**0.00**	50.91	80.56
JSc	98.44	98.44	**76.04**	**82.81**	62.5	**42.86**	**43.64**	77.78
GP	98.44	100.00	**89.58**	93.75	**37.50**	**25.00**	78.18	97.22
KAd	100.00	100.00	**79.17**	**84.38**	87.50	**37.50**	65.45	77.78
CH	100.00	100.00	**84.38**	90.63	**50.00**	**25.00**	76.36	91.67
RL	96.88	98.44	93.75	95.31	62.50	**42.86**	72.73	80.56
DM	98.44	98.44	95.83	98.44	**37.50**	**0.00**	50.91	91.67
JBr	100.00	100.00	93.75	93.75	75.00	62.50	67.27	97.22
JM	100.00	100.00	91.67	93.75	62.50	50.00	50.91	83.33

Cases are ordered according performance of healthy control participants (Butler et al., 2014). [a]Cut-off based on published norms. [b]No cut-off available. Authors' details for published assessments are provided in the text. to BNT severity. Scores are given in percentages. Missing data for one participant (JSo) are indicated with 'x'. Scores marked in bold fall below the cut-off for normal performance. The cut-off was calculated as 2 SD below the mean



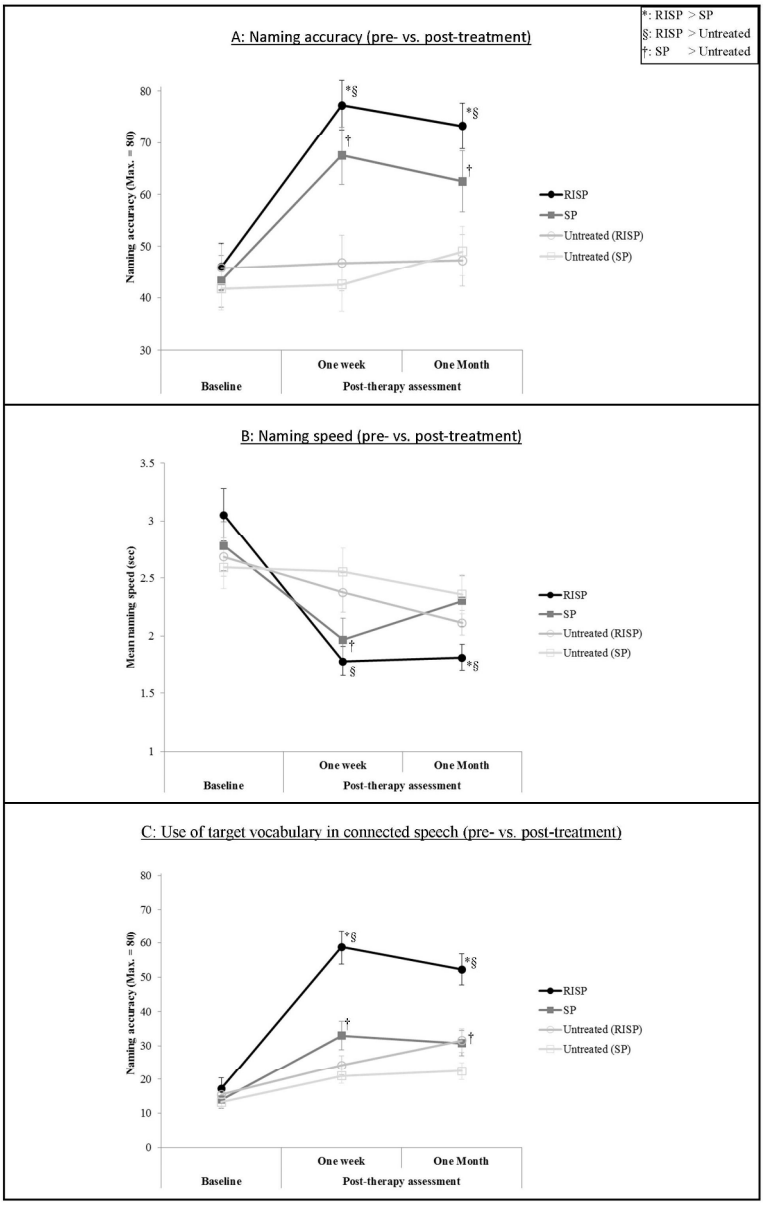
Lesion overlay map

252x61mm (96 x 96 DPI)



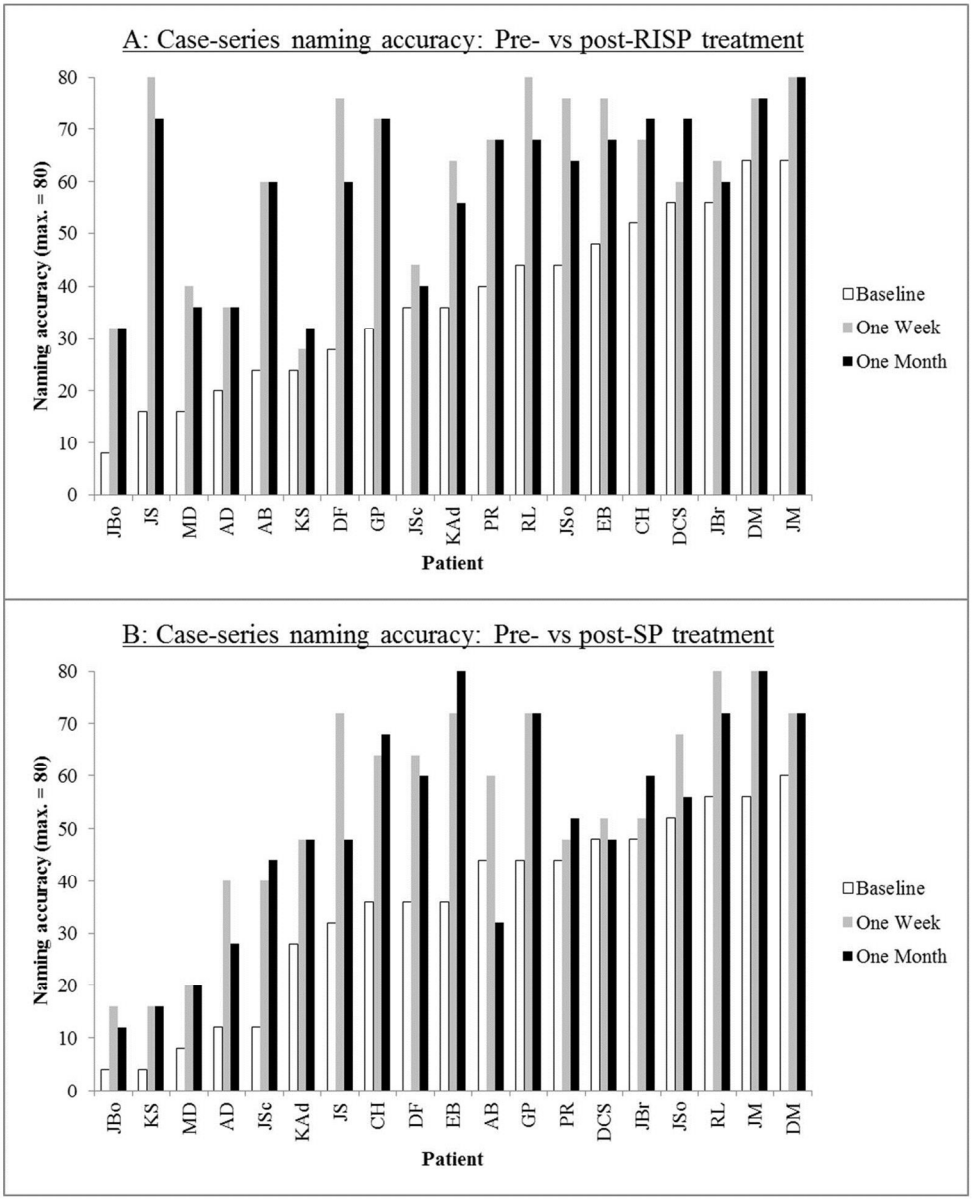
Treatment Overview

209x47mm (300 x 300 DPI)



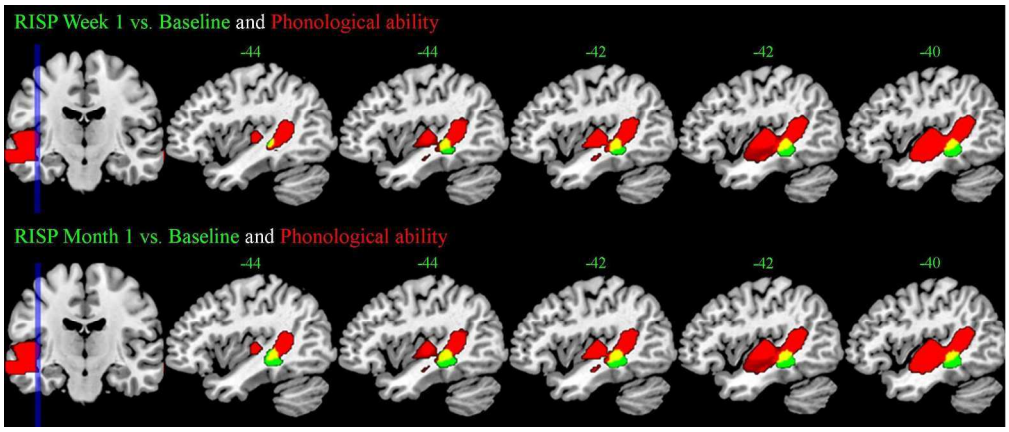
Treatment phases

163x258mm (300 x 300 DPI)



Naming accuracy across the case series

134x165mm (300 x 300 DPI)



Neural correlates of RISP

203x85mm (300 x 300 DPI)