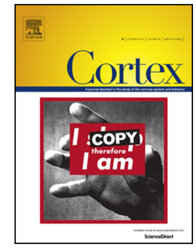
CORTEX 49 (2013) 2345–2357

Available online at www.sciencedirect.com

SciVerse ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex



Research report

Aphasia rehabilitation: Does generalisation from anomia therapy occur and is it predictable? A case series study

Wendy Best [a,*], *Alison Greenwood* [b], *Jennie Grassly* [a,b], *Ruth Herbert* [c], *Julie Hickin* [d] and *David Howard* [e]

[a] *Psychology & Language Sciences, UCL, London, UK*
[b] *Amersham Hospital, Buckinghamshire, UK*
[c] *Department of Human Communication Sciences, University of Sheffield, UK*
[d] *School of Health and Human Services, Central Queensland University, Rockhampton, Australia*
[e] *School of Education, Communication and Language Sciences, University of Newcastle upon Tyne, Newcastle upon Tyne, UK*

ARTICLE INFO

Article history:
Received 17 May 2012
Reviewed 17 July 2012
Revised 31 October 2012
Accepted 3 December 2012
Action editor Roberto Cubelli
Published online 4 February 2013

Keywords:
Aphasia
Anomia
Rehabilitation
Therapy
Generalisation

ABSTRACT

Introduction: The majority of adults with acquired aphasia have anomia which can respond to rehabilitation with cues. However, the literature and clinical consensus suggest change is usually limited to treated items. We investigated the effect of an experimentally controlled intervention using progressive cues in the rehabilitation of noun retrieval/production in 16 participants with chronic aphasia.
Method: Participants were sub-divided relative to the group according to performance on semantic tasks (spoken/written word to picture matching) and phonological output processing (presence/absence of word length effect and proportion of phonological errors in picture naming) in order to investigate outcome in relation to language profile. Cueing therapy took place weekly for 8 weeks.
Results: Intervention resulted in significant improvement on naming treated items for 15/16 participants, with stable performance on control tasks. Change occurred at the point of intervention and not during pre-therapy assessments. We predicted particular patterns of generalisation which were upheld. Only participants classified as having relatively less of a semantic difficulty *and* more of a phonological output deficit demonstrated generalisation to untreated items. Outcome did not relate to traditional aphasia classification.
Conclusion: A cueing hierarchy can improve word retrieval/production for adults with aphasia. In some cases generalisation to untreated items also occurs. The study demonstrates that the results of behavioural testing can be used to guide predictions of recovery with intervention.

* *Corresponding author*. UCL Division of Psychology and Language Sciences, Chandler House, 2 Wakefield Street, London WC1, United Kingdom.
E-mail address: w.best@ucl.ac.uk (W. Best).

0010-9452/$ – see front matter
http://dx.doi.org/10.1016/j.cortex.2013.01.005

1. Introduction

The majority of people with aphasia have difficulty in finding or producing words and this can be a significant cause of breakdown in conversation (e.g., Perkins et al., 1999). There is a large and growing body of evidence demonstrating that intervention can help improve word retrieval or word production (see Nickels, 2002 for a review). However, the majority of interventions result in change primarily on treated items (e.g., Abel et al., 2005; Fillingham et al., 2006; Laganaro et al., 2003; Wisenburn and Mahoney, 2009).

Given these fairly consistent findings a key question of both clinical and theoretical importance arises: what pattern(/s) of strengths and difficulties leads to generalisation to untreated items? The answer to this question may inform clinical practice and our understanding of how intervention is altering word retrieval/production.

1.1. *Models and levels of impairment*

There are several models of 'speech production', more recently and accurately termed 'language production' ranging from classic 'box and arrow' models (Ellis and Young, 1988; Kay et al., 1992) to connectionist models (Dell et al., 1997; Goldrick and Rapp, 2002; Levelt et al., 1999). While the models vary considerably in their specification, in relation to retrieving single words for production, all require the following three stages:

(1) Lexical-semantic processing or accessing word meaning (sometimes termed 'lexical semantics' and usually distinguished from 'conceptual semantics')
(2) Accessing abstract phonological word form (the 'phonological output lexicon' in box and arrow models; the 'phonological level' in Dell's account)
(3) Phonological encoding (or 'phonological assembly' in box and arrow models, commonly also termed 'post-lexical' processing).

In this paper 'word (or, for connected speech, language) production' will be used to refer to all three stages of processing. Thus, 'word production' incorporates retrieving the word's meaning and form and abstract phonological encoding. 'Word production' is more general than specific difficulties with word finding or word retrieval, sometimes used to refer exclusively to stage (2) above. All these occur prior to motor programming for speech (Ziegler, 2002).

Detailed single case studies link aphasic individuals' patterns of language strengths and weaknesses to difficulties with a particular level of processing. For example, E.E. (Howard, 1995) was held to have a deficit within the phonological output lexicon: he was consistent in the items he was unable to retrieve and was not helped by phonological cues. Howard suggests items were lost from his lexicon. Franklin et al. (2002) describe M.B. whose output included many phonological errors and whose performance was better on short than long words. M.B.'s difficulty was in assembling phonemes for production.

There is a confound in much of the research to date between the level of deficit and the target of intervention. This study employs the same intervention with participants with different levels of deficit enabling us to investigate the relationship between the level of impairment and outcome, in particular any generalisation to untreated items.

1.2. *Linking outcome to background findings*

In a seminal study, Hillis (1989) investigated a cueing therapy designed to improve written naming in two participants with severe aphasia. The participant with more lexical-semantic difficulty (stage 1 on the model above and common to accessing both written and spoken forms for production) improved and the change generalised to untreated items (and spoken naming). The second participant, with written naming difficulties arising from an orthographic equivalent to level 2, improved only on written naming of treated items. Hillis argued it is important to determine the source of an individual's naming difficulty in order to predict the outcome of intervention.

However, more recently, Lorenz and Ziegler (2009) did not find a direct relationship between the nature of the deficit and treatment approach. Participants with post-semantic anomia (stages 2 or 3 above) benefited from semantic intervention and also participants with semantic anomia (stage 1 on the model outlined above) benefitted from phonological/orthographic (word form) approach. Neither of these findings would be predicted from a straightforward link between intervention approach and breakdown in level of word production.

Fillingham et al. (2006) compared errorless learning with errorful learning. All participants completed a detailed language and neuropsychological assessment battery prior to intervention. Fillingham et al. found strong relationships between response to therapy and underlying neuropsychological profiles, with participants who responded better overall to both types of therapy having better recognition memory, executive/problem solving skills and monitoring ability. Strikingly, however, there was no clear relationship between language skill and therapy outcome.

What might be the reasons for the difficulty in relating language profile to the outcome of intervention? Firstly, people with aphasia rarely have a single clearly identifiable level of impairment in language production. For example, the same individual often makes both semantic and phonological errors in word retrieval. Furthermore, individuals' word production is often influenced by variables held to reflect different levels of processing. Secondly, almost all interventions involve participants in producing the target word thereby strengthening links from word meaning to word form (Howard, 2000) and potentially benefiting everyone with difficulty at some stage(s) in word production.

1.3. *Generalisation in word production interventions*

The findings from therapy studies for spoken word-production deficits are somewhat mixed with regards to the extent of the effect of treatment.

Limited or no generalisation to untreated items is the result across the majority of intervention studies including those investigating: errorless learning (Fillingham et al., 2006), production of nouns and verbs (Raymer et al., 2007), a cueing hierarchy (Thompson et al., 2006) and contextual priming (Renvall et al., 2007).

There are a few exceptions to this pattern. Interventions focused on process, particularly those with a semantic component (Renvall et al., 2003; Coelho et al., 2000; Boyle, 2004) are held to influence production of untreated items to some extent. Phonological Feature Analysis (Leonard et al., 2008) also resulted in generalisation to untreated items for 3/10 participants. Generalisation to homophones of targets has been found from intervention with a cueing hierarchy (Biedermann and Nickels, 2008) but not to phonologically or semantically related control items.

The distinction between therapy for semantic deficits (which targets this level) and semantic therapy for word production is important. In the former, 'semantic' tasks such as categorisation or semantic feature judgements are employed with the aim of improving a person's semantic processing; this should influence comprehension and production. In the latter, while meaning is involved in the task, e.g., through pictures, the intervention facilitates word production rather than semantic processing itself. An example is the study by, Howard et al. (2006) who demonstrated that manipulating the 'depth' of semantic processing did not influence naming outcome. Participants that benefited the most from semantic therapy for word production had a deficit in the links between word meaning and form (stage 2 on the model of word production outlined above). These results combined suggest this intervention is not actually operating at a semantic level but rather strengthening links between meaning and form.

Thus, there is consensus that repeatedly activating the links between an item's meaning and form [stages (1) and (2) above] often results in item specific improvement in naming (Howard, 2000), and this is the likely focus for change in a large number of therapy studies. However, the picture may not be as bleak as it first appears.

In a review of therapy for naming disorders, Nickels (2002) makes a distinction between approaches involving 'repair' and those that involve 'strategy'.

In the first case there is held to be a change in the individual's impairment. When the studies with methodological weaknesses were excluded, then 11 of the 44 people given phonological or orthographic information showed some generalisation to untreated items. Thus, around a quarter of participants in these studies improved on untreated as well as treated items. Findings from approaches involving 'strategy' and aimed at re-organising processes, such as orthographic self-cueing, were even more encouraging. Thirteen of nineteen cases showed some generalisation. Such approaches are, however, suitable for only some individuals with particular strengths (e.g., in retrieving orthographic knowledge). Interestingly, in a case series intervention using written cues, sixteen of eighteen participants improved on written naming, and four of these showed transfer to untreated items (Deloche et al., 1997; see also Carlomagno et al., 2001). This mirrors Nickels' review in suggesting around one quarter may demonstrate generalisation in word production.

There are several experimentally controlled single case studies with participants with deficits in post-lexical processing where intervention resulted in improvement on both treated and untreated items (Fisher et al., 2009; Franklin et al., 2002; Robson et al., 1998) For example, Fisher et al. (2009) worked with a man with 'mild phonological encoding impairment'. He showed significant generalisation to untreated items from an intervention which involved attempting to name pictures with unrelated names or with shared phonology (magnet, mattress, macaroni). In contrast, Waldron et al. (2011) found no generalisation to untreated items, despite employing a previously successful intervention (Franklin et al., 2002). The participants in Waldron's study had a combination of lexical (stage 2) and post-lexical (stage 3) impairments. Raymer et al. (2012), in a study investigating errorless naming treatment and gestural facilitation of naming did not obtain generalisation to untrained items for the three participants with semantic anomia, but obtained some generalisation in naming for three of five participants with phonological anomia. Finally, studies using orthographic cueing aids demonstrate convincing generalisation to untreated items (Best et al., 1997; Bruce and Howard, 1987; Howard and Harding, 1998).

1.4. *Aims*

We aimed to explore the effects of a cueing hierarchy, especially generalisation to untreated items, and to relate the outcome to level of breakdown in naming.

Specifically, we ask:

(i) Can a cueing therapy improve word production (i.e., retrieval of meaning and form and phonological encoding) in participants with aphasia?
(ii) Do some participants show improvement on untreated items?
(iii) Can any generalisation to untreated items be related to the participants' language profiles?

From previous studies we predicted:

(a) those with a post-semantic deficit, stage 2, with relative strengths in semantic and phonological output processing and a specific deficit in retrieving lexical forms will show item specific changes in naming (following e.g., Howard et al., 2006; Raymer et al., 2007)
(b) those with a post-lexical deficit, stage 3, with relative strengths in semantic processing and weakness in phonological output processing will show effects of intervention which generalise to untreated items (following e.g., Franklin et al., 2002; Fisher et al., 2009).

1.5. *Value*

The study is of theoretical importance. Evidence for a link between the nature of the impairment and change with intervention can inform our understanding of improvement mechanisms. In rehabilitation for word production, any intervention which involves pictures and producing spoken words will necessarily activate all the representations and levels of processing in the model outlined above. The question is whether therapy can operate at different levels and whether generalisation reflects the level at which change in the system is occurring.

This investigation is also of clinical importance. Those people who show generalised improvement to untreated items are likely to be benefiting more than those who show changes limited to treated items, although item specific changes may also impact on everyday life (e.g., Best et al., 2008; Raymer et al., 2007). For those who improve only on treated items, selection of these items to be of maximum functional benefit to each individual is crucial. Finally, the study is of clinical relevance because we include 'all comers'. Rather than including only those with clearly identifiable impairments at a single level, we included everyone referred to the study who met the general criteria.

Prognosis in aphasia is generally linked to stroke related variables (initial aphasia severity, nature of lesion, e.g., Saur et al., 2010) rather than patient related variables (gender, handedness, education, e.g., Plowman et al., 2011). Pederson et al. (2004) found language outcome was related to aphasia severity but not type of aphasia. Thus, from both the detailed single case cognitive neuropsychological and the broader prognosis literature, our hypothesis is that generalisation to untreated items may not be predicted by participants' traditional aphasia classification, but rather by language scores from behavioural testing.

2. Method

2.1. Participants

Sixteen participants with varying profiles and severity of aphasia were recruited. Criteria for inclusion were minimised in order for participants to better reflect the clinical population rather than, for example, selecting those most likely to benefit from rehabilitation (e.g., highly motivated participants). All those who met the criteria were included; all had word finding difficulties as a significant part of aphasia and were more than a year post-onset. All participants had aphasia due to a single left cerebrovascular accident (CVA). Participants gave informed consent via an aphasia friendly form and process (Osborne et al., 1998). Results from two intervention studies were combined to provide the data for this investigation.

Participants ranged from one to eight years post-onset at the time of the study and from 42 to 77 years. Participants' aphasia type was agreed by the research clinicians, all of whom are experienced speech and language therapists; there was complete agreement as to the categorisation of participants as fluent or non-fluent. Where a traditional aphasia subtype is shown in Table 1 there was also agreement as to the category as determined by background language profiles and connected speech. Eight participants had fluent aphasia and eight had non-fluent aphasia.

2.2. Background assessments

Naming was assessed using a set of 200 black and white line drawings (for which there is 95% name agreement from older control participants). The influence of psycholinguistic variables on naming was investigated and the nature of participants' errors was coded. A phonological error was counted where the attempt was a word or non-word for which 50% or more of the target phonemes were in the response or 50% or more of the phonemes in the response were in the target. Participants' comprehension of single words was assessed using spoken and written word to picture matching from the Comprehensive Aphasia Test (CAT; Swinburn et al., 2004). Single word reading and repetition were assessed using the same set of 152 items.

2.3. Intervention

The data from this study come from two separate but strongly related projects: the Tavistock study and the Buckinghamshire study. The Tavistock study used phonological and orthographic cues in the treatment of word finding difficulties in aphasia (Best et al., 2002; Hickin et al., 2002; Herbert et al., 2003). In this study the eight participants were provided with a choice of phonological cues or a choice of orthographic cues in treatment. The Buckinghamshire study was a collaborative project with therapists working in NHS and academic settings

Table 1 – Participant details, at time of study. The first eight participants were in the Tavistock Study (University based) and the final eight were in the Buckinghamshire (Health Service based) replication.

Participant	Gender	Years post-onset	Age	Aphasia type	Occupation at time of CVA
H.M.	M	6	45	NF Broca's	Cabinet maker
P.H.	F	3	77	F Anomic	Homemaker
S.C.	M	5	65	F Mixed/Wernicke's	Retired
D.C.	F	5	70	F Anomic	Retired
O.L.	F	2	65	F Anomic	Retired
N.K.	M	3	52	F Anomic	Accountant
I.K.	M	3	68	NF Broca's	Retired, ran a business
K.R.	F	8	38	NF Broca's	Homemaker
T.E.	M	1	69	F Anomic	Ran building business
F.A.	F	2	64	NF some apraxia	Personal assistant
G.B.	M	3	71	NF	Retired florist
C.M.	M	5	52	NF	Plumber
C.V.	F	2	56	NF	Florist/gardener
D.J.	F	1	65	F	Volunteer
P.P.	F	2	75	F Wernicke's	Homemaker
L.M.	F	7	42	NF Broca's	Homemaker

and was based in the Health Service. Thus, the study investigated the effectiveness of this approach in the clinical setting, rather than the efficacy of the intervention under optimum conditions (Pring, 2005). The Buckinghamshire study compared single cues with a choice of cues however in this study all cues were provided in both phonological and orthographic form (see Appendix 1 for examples) and investigated maintenance of effects and the eight participants' views of intervention and change (Best et al., 2008; Greenwood et al., 2010).

The two projects designs and the cues used are summarised in Appendix 2. There are very strong similarities which enable us to ask questions about generalisation combining data across the two studies.

Design aspects common to both studies:

(i) Baseline

There was an 8-week pre-therapy baseline with regular contact with the therapists, matching the contact during the therapy phase. This allowed us to look for change over baseline and to control for possible 'charm' effects. During this time, a range of background assessments were used to provide a profile of each participant's strengths and impairments in language processing.

(ii) Stimuli

The same 200 pictures were named at the start and end of the baseline phase. One hundred items were selected for inclusion in therapy and these were matched with control items for baseline naming for each participant. We deliberately did NOT select items which participants failed to name as this could lead to regression to the mean in naming post-therapy i.e., apparent treatment effects could appear simply from inherent variability in aphasic naming. Within the 200 items we constructed a sub-set of items varying on length and matched for other psycholinguistic variables (imageability, age of acquisition, frequency and familiarity, details in Appendix 3).

(iii) Intervention

Intervention took place once a week for 8 weeks, sessions lasted around 1 h. Participants were provided with the treated items to name. If unable to name the pictures after 5 sec, participants were given cues. The first cue was a single phoneme plus schwa and/or single grapheme. The second cue was the first syllable of the word or C.V. if the target was monosyllabic. Cues were provided approximately 5 sec apart. If the progressive cues did not aid naming participants were given the word to repeat in the presence of the picture.

(iv) Primary outcome

Naming of treated and unseen, untreated items was reassessed immediately after therapy. This was the primary outcome measure for the intervention. Naming assessments were recorded and, to investigate inter-rater agreement, a sub-set was scored directly from audio recordings by an independent rater blind to data collection point. The results were compared with the in vivo scoring.

(v) Control tasks

Data from two control tasks were collected each time naming was assessed in order to investigate whether any changes were limited to word retrieval/production or evident in an untreated task. Verbal short term memory span was selected as avoiding floor and ceiling effects for this population. Participants heard a set of picture names and pointed to these in the stated order, the task thus avoiding the need for language production. Written sentence comprehension was selected as a further verbal task, impaired in most people with aphasia and not involving components of language processing targeted in the cueing intervention.

3. Results and discussion

The findings from the background assessments are reported, followed by the results of the cueing intervention for the treated items. Thereafter, change on untreated items is presented and related to the findings from the background psycholinguistic assessments.

3.1. *Results from background assessments*

All participants performed well above chance (25% correct) on spoken and written word to picture matching with scores ranging from 67% to 100% correct (Table 2). Picture naming scores varied considerably. Errors ranged between 10% and 56% semantic and between 0 and 48% phonological. There was also a wide range of performance on word repetition (36–100% correct) and single word reading aloud (28–97% correct). The considerable variety in participants' scores enabled us to divide them into sub-groups according to the nature of their relative language processing strengths and difficulties.

3.2. *Classification into sub-groups*

Participants were classified according to their performance on the tasks tapping semantic processing. We did not include the proportion of semantic errors in naming in this process as such errors may reflect semantic difficulties but may also reflect difficulty in retrieving phonological forms (Nickels and Howard, 1994). For non-fluent participants, single word semantic errors may be curtailed circumlocutions produced when a response is required. Instead we used the better of the two word to picture matching tests for each individual to calculate a z-score. Thus, for the three participants scoring the same with spoken and written input, this score was used. However, for the 13 participants with a discrepancy between spoken and written word to picture matching (due to impairments processing either spoken or written input) the lower score was ignored and the score from the other modality is used. This is most likely to reflect semantic processing ability. The method is not foolproof as some participants may have difficulty with processing *both* written and spoken input. However, from the data available, the z-score provides the best measure of semantic processing.[1] Those with a

[1] In previous studies we have combined z-scores across semantic tasks (Hickin et al., 2002), but this approach results in scores that are influenced by impaired input processing specific to one modality. In future studies we would recommend using more discriminating semantic tasks. We return to this point in the discussion.

Table 2 – The results of background assessments.
Semantic tests: SWPM: Spoken word to picture matching test: percentage correct (CAT; $n = 30$), WWPM: Written word to picture matching test: percentage correct (CAT; $n = 30$), PN SE: Picture naming: semantic errors as a proportion of total errors. Phonological tests: Rep Wd: Repetition of words: percentage correct ($n = 152$), Read Wd: Reading words aloud: percentage correct ($n = 152$), PN PE: Picture naming: phonological errors as a proportion of total errors.

Participant	SWPM	WWPM	PN SE	Rep Wd	Read Wd	PN PE
H.M.	1.00	.87	.52	.73	.70	.20
P.H.	.93	.97	.25	.97	.97	.05
S.C.	.87	.77	.28	.57	.15	.02
D.C.	1.00	.97	.50	.95	.97	.11
O.L.	.97	.93	.16	.99	.91	.00
N.K.	.93	.97	.33	.99	.92	.00
I.K.	.93	.80	.18	.52	.31	.22
K.R.	.93	.90	.16	.90	.64	.02
T.E.	1.00	1.00	.17	.87	.88	.48
F.A.	.87	.90	.17	.36	.20	.33
G.B.	.87	.90	.10	.36	.33	.20
C.M.	.83	.90	.28	.70	.35	.29
C.V.	.67	.73	.25	.89	.78	.00
D.J.	.97	.97	.56	.45	.60	.12
P.P.	.87	.97	.17	.57	.28	.28
L.M.	.97	1.00	.20	1.00	.96	.05

negative score (i.e., worse than mean for the group) are marked ‘Y’ in Table 3. They are classified as having relatively more of a semantic deficit. Those with a positive score (i.e., better than mean for the group) are marked ‘N’ as having relatively less of a semantic deficit. The same sub-grouping is obtained by using the better word to picture matching test and splitting at the median score.

With regard to phonological processing, we classified participants according to the proportion of phonological errors made in picture naming and according to whether there was a significant influence of length on their picture naming ability using the matched sub-sets of 1, 2 & 3 syllable items (Appendix 3). In order to be classified as having a phonological production deficit/post-lexical difficulty in production (i.e., stage 3 on the model) participants needed a positive z-score for phonological errors, and for word length to influence their naming with significantly worse performance on the long than short words (the Jonckheere Trend Test was used to determine the statistical significance of the effect of number of syllables; $p < .05$, one-tailed). Table 3 (3rd and 4th columns) shows that 15 of the 16 participants would have been entered into the same group regardless of which of these measures was used for classification (there was a discrepancy only for P.H.).

This resulted in four sub-groups according to whether participants had relatively better or worse semantic processing (column 2 of Table 3) and relatively better or worse phonological output processing (column 5 of Table 3).

Table 3 – Categorising participants according to focus of word production difficulty.

Participant	Semantic deficit? (i.e., z-score on better w–p matching test is negative)	Phonological (z-score on proportion phonological errors is positive)	Length effect on picture naming	Phonological deficit? (i.e., both high proportion of phonological errors and length effect)
H.M.	N	Y	Y	Y
P.H.	N	N	Y	N
S.C.	Y	N	N	N
D.C.	N	N	N	N
O.L.	N	N	N	N
N.K.	N	N	N	N
I.K.	Y	Y	Y	Y
K.R.	Y	N	N	N
T.E.	N	Y	Y	Y
F.A.	Y	Y	Y	Y
G.B.	Y	Y	Y	Y
C.M.	Y	Y	Y	Y
C.V.	Y	N	N	N
D.J.	N	N	N	N
P.P.	N	Y	Y	Y
L.M.	N	N	N	N

3.3. Inter-rater agreement on scoring naming

An independent rater scored naming accuracy on 10 assessments directly from audio recordings. The overall agreement with in vivo ratings was 91% ($n = 1598$ items, Kappa .812, $p < .001$). Inter-rater agreement was substantial for both pre- and post-therapy assessments.

3.4. Results of intervention for treated items

All participants made a numerical improvement in naming treated items (Fig. 1). The change was statistically significant for 15 participants (Wilcoxon matched samples, one-tailed test, $p < .05$), with S.C. in the Tavistock study showing no significant change in naming treated items (further details in Hickin et al., 2002). A comparison between the mean pre-intervention score [43.5, standard deviation (SD) 18.12] and the mean post-intervention score (62, SD 22.85) for treated items reveals the large effect size for the group (Cohen's d of .897).

3.5. Results of intervention for untreated items

The findings for untreated items are shown in Fig. 2. The change shown is proportional as there were different numbers of unseen items in the two projects (Tavistock study 100; Buckinghamshire study 50). A comparison between the mean pre-intervention raw score (33.84, SD 17.61) and the mean post-intervention score (36.31, SD 19.17) for untreated items reveals an effect size (Cohen's d) of .134. While this should be interpreted with care due to the different number of items in the different studies, it is clear the effect size for the group is minimal.

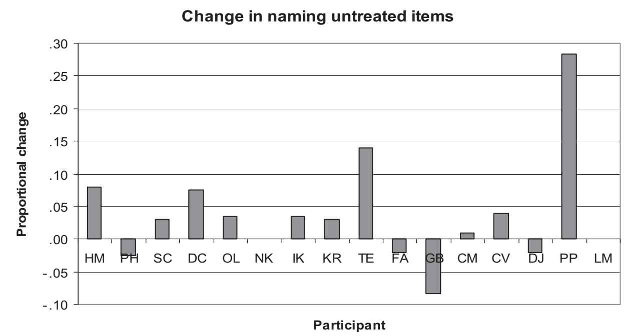


Fig. 2 – Proportional change in picture naming for untreated items for all participants (A3-mean A1A2: $n = 100$ Tavistock study; $n = 50$ Buckinghamshire study).

3.6. Results on control tasks

Table 4 shows that there was stability in the control tasks across occasions (raw scores for each participant are provided in Appendix 4). A One way Repeated Measures Analysis of Variance (ANOVA) demonstrated no significant difference between the mean scores at different time points on either task [short term memory (STM) pointing span, $F(2, 22) = .12$, $p = .88$; Sentence comprehension $F(2, 22) = .94$, $p = .40$].

3.7. Relating change in naming to profiles from background assessments

The following section relates the categories to which we allocated participants on the basis of background language testing to the change in picture naming with therapy.

Table 5 provides mean change on treated items for the four sub-groups with relatively stronger and poorer semantic and phonological output processing (naming of the whole 200 items is provided in Appendix 5).

The sub-groups change on treated items ranges from 14 to 22%, with those having relatively better semantic processing and better phonological output processing making slightly more change on average, although none of the sub-groups stands out. This was confirmed by a 2 × 2 between subjects ANOVA [$F(1, 12) < 1$, n.s. for effect of semantic impairment, effect of phonological impairment and interaction].

Fig. 3 shows mean change on untreated items for the four sub-groups.

The three participants (H.M., T.E., P.P.) with relatively less of a semantic difficulty *and* more of a phonological output deficit (stage 3) show a pattern of generalisation to untreated items. A 2 × 2 between subjects ANOVA on the untreated items shows: an effect of semantic impairment $F(1, 12) = 7.73$, $p = .017$; no effect of phonological impairment $F(1, 12) = 3.58$, $p = .083$; and a highly significant interaction $F(1, 12) = 12.74$, $p = .004$.

Interestingly, the three participants who generalised differ according to traditional aphasia classification (H.M.,

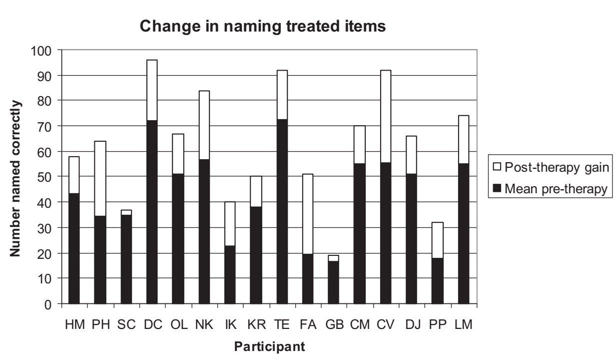


Fig. 1 – Illustrates the change in picture naming for treated items after intervention. The black section of the bars illustrates the mean naming across the two baseline assessments for 100 treated items (except for G.B. and P.P. for whom 60 items were treated). The white section shows the additional items named after the intervention. For the first eight participants, those in the Tavistock Study, improvement in naming is from items treated with a choice of spoken or choice of written cues. For the remaining eight participants, those in the Buckinghamshire Study, improvement in naming is from single combined (spoken and written) cue or choice of combined (spoken and written) cues. Raw data for all participants naming all items on all three occasions are provided in Appendix 4.

Table 4 – Control tasks: (1) Verbal short term memory – picture pointing span; (2) Written sentence comprehension (chance performance is 4/16).

Participant	Assessment 1 (pre-therapy 1)	Assessment 2 (pre-therapy 2)	Assessment 3 (post-therapy)
Verbal short term memory, pointing span			
Mean (SD)	3.32 (.79)	3.27 (.72)	3.35 (.95)
Written sentence comprehension ($n = 16$)			
Mean (SD)	10.54 (2.86)	11.50 (2.44)	11.01 (2.62)

Broca's aphasia; T.E., Anomic aphasia; P.P., Wernicke's aphasia).

The only participant to show more than 4% change on untreated items (see Fig. 2) and not to fall into the sub-group with better semantic processing and impaired phonological processing was D.C. She did have relatively good semantic processing but made 11% phonological errors so was on the border of being classified as having a phonological output impairment with respect to picture naming errors. Furthermore, while she did not demonstrate a significant effect of length on picture naming overall (Jonckheere Trend Test, $z = 1.20$, $p = .11$, one-tailed), she did show a dip in performance for naming three syllable items (1 syll. .71, 2 syll. .74, 3 syll. .63). Thus, D.C.'s pattern of performance is not out of line with the general statement that those with relatively less of a lexical-semantic deficit and more of a phonological encoding deficit may show some generalisation to untreated items. In using predetermined cut-offs to assign participants to different theoretically motivated cells the detail of her performance has been obscured.[2]

4. General discussion

4.1. *Overview of findings*

The study posed three research questions:

(i) Can a cueing therapy improve word production (i.e., retrieval of meaning and form and phonological encoding) in a series of participants with aphasia?

In line with previous research we can answer yes. In this study, 15 of 16 participants showed significant change on naming treated items. The stability on the control tasks, along with all participants being well out of the phase of spontaneous recovery at the start of the study, point to the changes resulting from the intervention.

(ii) Do some participants show improvement on untreated items?

While the change was limited to treated items for the majority of participants, there were several for whom there was also change on the untreated items.

(iii) Can the outcome and, in particular, any generalisation to untreated items be related to the participants' language profiles?

Specifically, we predicted

(a) those with a post-semantic deficit, stage 2, with relative strengths in semantic and phonological output processing, will show item specific changes in naming
(b) those with a post-lexical deficit, stage 3, with relative strengths in semantic processing and weakness in phonological output processing, will show effects of intervention which generalise to untreated items.

The answer to question (iii) is considered below in sections on: sub-grouping, outcomes in relation to this and more traditional aphasia classification, and generalisation in relation to sub-groups. Finally, we discuss the clinical and research implications of the findings.

4.2. *Sub-grouping participants*

While our method of comparison relative to the group enabled classification of participants into four theoretically motivated sub-groups to achieve the aims of this study, further consideration is necessary before such methods are used in future research or clinical practice. Classifying this set of participants using z-scores on word to picture matching resulted in participants with a score of .93 or less being scored as having more of a semantic deficit, and .97 or more as having relatively less of a semantic deficit. Thus, for participants in this study, a cut-off score for degree of semantic impairment could be set at around .95. However, clinically, this should be used with caution. The cut-off warrants verification from further research and more discriminating tasks e.g., word picture verification with reaction times could be employed in future studies and in clinic. We would continue to advocate taking the better of the spoken or written tasks as a measure of semantic processing.

All but one (15/16) participants were classified into the same group for phonological production deficit from either proportion of phonological errors or from the presence/absence of a length effect in naming. This suggests that either a length effect on naming *or* the presence of a high proportion of phonological errors may be taken as indicating a deficit at stage 3 on the model.

In considering the proportion of phonological errors, although Table 2 shows half the participants made 11% or fewer phonological errors while half made 12% or over, we would not suggest using a number between these as the exact cut-off score. Further research investigating nature of difficulty and outcome of intervention is necessary. From this study we suggest those with a small percentage of phonological errors (up to and including 5%) are not likely to have a phonological production deficit that results in generalised therapy effect. Those for whom 20% or more of errors are phonological are

[2] We are grateful to an anonymous reviewer for requesting further clarification on D.C.'s pattern of performance.

Table 5 – Mean change on *treated* items (SD in parentheses) for participants in sub-groups categorised by background assessment.

	Phonological output processing better	Phonological output processing impaired
Semantic processing better	PH OL NK DC LM DJ .22 (.06)	HM TE PP .19 (.04)
Semantic processing impaired	SC KR CV .14 (.14)	IK FA CM GB .17 (.12)

likely to have such a deficit. All participants except D.C., discussed above, and D.J. fall into one of these two groups.

4.3. *Improvement on treated items and aphasia classification*

The results for treated items replicate previous research which has shown intervention involving cues can aid naming in adults with aphasia (Nickels, 2002). The study shows that change can occur from intervention once a week for 8 weeks. The outcomes do not relate straightforwardly to traditional aphasia classification. For example, from Fig. 1, it is clear that, of the two participants who made least change in naming, one had fluent aphasia (S.C.) and the other had non-fluent aphasia (G.B.). Likewise, the participant in the first study who named the most extra items (P.H.) had anomic aphasia; in contrast, the participant in the Health Service based study who named the most extra items (F.A.) had non-fluent aphasia. Thus, the results do not relate to traditional aphasia classification or even the distinction between fluent and non-fluent aphasia. It is, therefore, unlikely that the extent of improvement in picture naming of treated items would relate to lesion site, although this remains to be explored. This disassociation between outcome and traditional aphasia classification is also in line with other studies treating written and spoken naming (e.g., Carlomagno et al., 2001; Leonard et al., 2008).

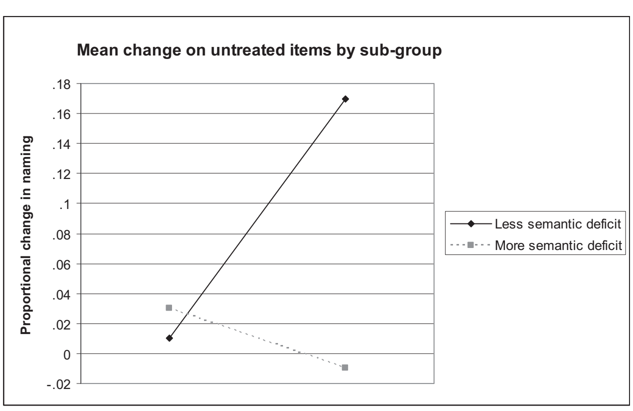


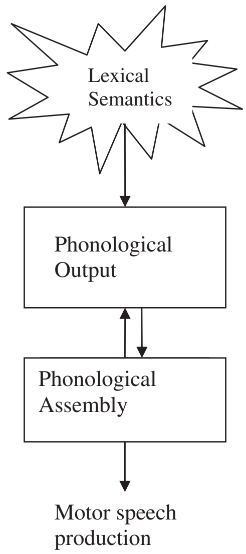
Fig. 3 – The outcome for untreated items in relation to the four sub-groups derived from background assessments. Mean change (SD): less semantic deficit, less phonological deficit .01 (.04), more semantic deficit, less phonological deficit .03 (.01), less semantic deficit, more phonological deficit .17 (.10), more semantic deficit, more phonological deficit −.01 (.05).

4.4. *Relating generalisation to background profiles*

The introduction outlined three stages of processing in spoken language production. We return to these and relate them to findings from other studies which have investigated levels of deficit in relation to outcome and to the data from this study. Stages 1–3, outlined in the introduction, are illustrated to the left of Fig. 4 which displays assessment findings and not the nature of intervention provided.[3] The figure includes only studies where detailed background assessment enables the link between level of deficit and outcome of intervention to be explored.

The participants with anomia with a deficit at stage 1 (accessing word meaning) or stage 2 (accessing word form) do *not* show generalisation to untreated items from therapy directed at their anomia. With deficits at the first stage of production it may be that intervention which targets semantic processing directly (therapy for semantics) can produce generalised effects (e.g., Renvall et al., 2003; Coelho et al., 2000; Boyle, 2004). However, there is very little evidence for generalised treatment effects with participants with a deficit at stage 2 i.e., in accessing the phonological

[3] The findings of Deloche et al. (1997) are not included in the table as their study focuses on the relationship between oral and written naming, and background assessment which would allow classification is not provided.



Level of impairment	Franklin et al., 2002	Lorenz & Ziegler 2009	Present study	Generalisation obtained from level of deficit?
Lexical semantic processing		BR, ZU, BF	SC, KR, CV,	No
Accessing phonological form		GE, GG, HW	PH, OL, NK, DC, LM, DJ	No
Phonological encoding/ assembly (post-lexical processing)	MB		HM, TE, PP	Yes

Fig. 4 – Levels of processing deficit in language production in relation to studies investigating generalisation to untreated items. The participants from the current study with relatively impaired semantic *and* phonological processing (I.K., F.A., C.M., G.B.) are not shown in the figure. P.H., O.L., N.K., D.C., L.M., D.J. are shown as having a deficit in accessing the phonological form as they demonstrate word finding problems in the context of relatively good semantic and phonological output processing. In the Lorenz and Ziegler (2009) study J.K. is excluded because, although he did show generalised changes, there were also changes during baseline suggesting an effect of spontaneous recovery rather than the intervention. Finally, participants in their study with mixed anomia (semantic and phonological: E.B., M.H., R.A.) are also not shown as the focus of their deficit is unclear.

form. This is the case whether the intervention is semantic (e.g., Howard et al., 2006; Lorenz and Ziegler, 2009) or involves cueing as in the present study. The lack of generalisation found for those with a naming deficit arising at stage 2 (i.e., participants with naming difficulties but nevertheless relatively good lexical-semantic processing and good phonological encoding: P.H., O.L., N.K., D.C., L.M., D.J.) aligns with prediction (a) (Section 1.5).

The partial generalisation from Phonological Feature Analysis (Leonard et al., 2008) remains to be further explored in relation to level of anomic deficit. In their study, three of 10 participants improved in naming treated and untreated items (P2, P3, P4). Two of these show high proportions of phonologically related errors (formal or nonword) with the third, P4, making mainly errors of omission, which may suggest good self-monitoring. In common with most studies in the field, the effect of word length in picture naming is not investigated. Further data in line with the claims arising from the present paper come from the fact that two (P2 & P4) of the three participants who showed generalised effects also show less of a semantic deficit relative to their study participants (taking the better of the spoken and written word to picture matching scores; Leonard et al., 2008, Table 2).

In the studies with participants where the focus of the deficit appears to be in phonological encoding (M.B. Franklin et al., 2002; H.M., T.E., P.P. present study; see also T.V. Fisher et al., 2009) there was generalisation to untreated items. This is in line with our second prediction (b) (Section 1.5).

However, not all those who make a high proportion of phonological errors in picture naming show generalisation to untreated items; those with a co-occurring semantic deficit (I.K., F.A., C.M. & G.B. in present study) did not demonstrate change on untreated items. A possible explanation for this outcome is that due to the lexical-semantic deficit, during word retrieval there is insufficient activation feeding through to the level of phonological encoding; the level at which the generalisation to untreated items is occurring. It is only when lexical-semantic processing remains relatively well preserved, which enables partial activation at the level of phonological encoding, that the intervention can produce generalised changes.

The outcomes also relate to the more general question of whether intervention should target relative strengths or weaknesses in individuals' language processing. In relation to the model of word production, our findings suggest participants with relatively good semantic and phonological processing but impaired access to phonological word forms (stage 2) can show item specific benefits, whereas those with relatively good semantic processing but impaired phonological encoding (stage 3) can show generalised benefit from the cueing hierarchy intervention.

There are several possible accounts of how the generalisation to untreated items is occurring. This has been explored in detail in two of the single case experimental studies (M.B. Franklin et al., 2002; and, from this research, T.E. Greenwood et al., 2010). The authors claim that their intervention improved phoneme retrieval for M.B. and strengthened bidirectional connections between words and phonemes for T.E. In models in which each phoneme feeds back to multiple lexical items (Dell et al., 1997; Goldrick and Rapp, 2002) improvement in untreated words arises directly from either account of the mechanism of change.

Our findings concur with the claim that it is possible to use background language assessments to predict the outcome from cueing therapy (Hillis, 1989). Abel et al. (2007) delivered therapy according to predictions made about participants' underlying language profiles and also conclude that models can be informative when making decisions about which therapy to use. Interestingly, in their 2005 study no participants improved with vanishing cues only, but several showed positive effects with increasing cues alone (as in the present study) or with both increasing and vanishing cues.

4.5. *Implications*

The results of this inceptive study demonstrate that generalised improvement to untreated items *can* result from cueing therapy. Although the majority of participants made item specific improvements, which can be of functional benefit, our results corroborate the findings of Nickels' review (2002) in which around a quarter of participants also improved on untreated items following this type of intervention.

The ability to predict those people who might show generalisation to untreated items is of clinical and theoretical importance. Participants who display relatively good semantic processing and poor phonological encoding are more likely to improve in naming untreated items. We suggest this underlying profile may be more important in guiding our predictions of recovery than traditional aphasia classification.

4.6. *Future directions*

Tate et al. (2008) list criteria for sound single case/case series experimental studies. The work presented in this paper met the majority of the criteria with an exception being that reassessment was not carried out by an independent investigator blind to the stage of assessment. The high inter-rater agreement obtained for naming when comparing in vivo scoring by the therapist with scoring from recordings (where the rater was blind to stage of study) goes some way to alleviate concern over bias. However, we would advocate blind reassessment in future studies.

Employing a case series approach, with enough detail from each participant to allow sub-grouping, has been crucial in relating background profile to therapy outcome and we would strongly recommend this approach in future intervention studies (Carlomagno et al., 2001; Schwartz and Dell, 2012) alongside detailed single cases and computational modelling allowing the mechanisms of change to be fully explored. Furthermore, future studies could include exploration of the relationship between memory/executive skills and therapy outcome (Fillingham et al., 2006) and investigation of maintenance without the further phase of connected speech therapy included in the present study (see Appendix 2 and Herbert et al., 2003).

The present study also highlights the need for further research which carefully relates nature of a person with aphasia's difficulty and strengths to the outcome of intervention. In particular, studies comparing multiple interventions, particularly semantic versus phonological approaches, are necessary. Studies should consider the following: (i) using case series designs with three or more baseline assessments, (ii) measuring outcome beyond picture naming, including participants' views of intervention and outcome and (iii) the outcome of approaches directed at different levels of communication (e.g., single words *vs* conversation).

5. Overview

In this experimentally controlled case series study, 15/16 participants improved significantly in naming treated items. There are several lines of evidence that demonstrate the change resulted from the specific intervention:

(i) the change was specific to treated items for most participants
(ii) all were out of the phase of spontaneous recovery
(iii) participants were stable on control language tasks
(iv) change occurred at the point of intervention and not during baseline.

The generalisation to untreated items for a minority of participants relates to their language production profiles in line with our predictions. While the pattern of findings warrant further exploration, our intervention involving cues did not produce generalisation to untreated items in those with relatively greater semantic deficits or difficulty in accessing the form for production. Rather, it occurred in all of those with post-lexical speech production deficits where these co-occurred with relatively intact semantic processing.

Funding

This work was supported by The Tavistock Trust for Aphasia (to W.B. & D.H.), The Stroke Association (to W.B. & J.H.), Wycombe PCT (to A.G., J.G.), UCL and Birkbeck College. The writing was completed while W.B. was in receipt of an ESRC Grant and D.H. an NIHR Grant.

Acknowledgements

We are very grateful to the 16 participants with aphasia for enthusiastic participation in the study. Jenny Sugden, NHS manager, supported the part of the study based in Buckinghamshire. Emma Prince provided the inter-rater reliability from audio recordings.

Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.cortex.2013.01.005.

REFERENCES

Abel S, Schultz A, Radermacher I, Willmes K, and Huber W. Decreasing and increasing cues in naming therapy for aphasia. *Aphasiology*, 19(9): 831–848, 2005.

Abel S, Willmes K, and Huber W. Model-oriented naming therapy: Testing predictions of a connectionist model. *Aphasiology*, 21(5): 411–447, 2007.

Best W, Greenwood A, Grassly J, and Hickin J. Bridging the gap: Can impairment-based therapy for anomia have an impact at the psycho-social level? *International Journal of Language & Communication Disorders*, 43: 390–407, 2008.

Best W, Herbert R, Hickin J, Osborne F, and Howard D. Phonological and orthographic facilitation of word-retrieval in aphasia: Immediate and delayed effects. *Aphasiology*, 16(1–2): 151–168, 2002.

Best W, Howard D, Bruce C, and Gatehouse C. Cueing the words: A single case study of treatments for anomia. *Neuropsychological Rehabilitation*, 7(2): 105–141, 1997.

Biedermann B and Nickels L. The representation of homophones: More evidence from the remediation of anomia. *Cortex*, 44(3): 276–293, 2008.

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

Bruce C and Howard D. Computer-generated phonemic cues – An effective aid for naming in aphasia. *British Journal of Disorders of Communication*, 22(3): 191–201, 1987.

Carlomagno S, Pandolfi M, Labruna L, Colombo A, and Razzano C. Recovery from moderate aphasia in the first year post stroke: Effect of type of therapy. *Archives of Physical Medicine and Rehabilitation*, 82: 1073–1080, 2001.

Coelho CA, McHugh RE, and Boyle M. Semantic feature analysis as a treatment for aphasic dysnomia: A replication. *Aphasiology*, 14(2): 133–142, 2000.

Dell GS, Schwartz MF, Martin N, Saffran EM, and Gagnon DA. Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, 104(4): 801–838, 1997.

Deloche G, Hannequin D, Dordain M, Metz-Lutz M-N, Kremin H, Tessier C, et al. *Journal of Communication Disorders*, 30: 11–22, 1997.

Ellis AW and Young AW. *Human Cognitive Neuropsychology*. Hove and London, UK: Lawrence Erlbaum Associates, 1988.

Fillingham JK, Sage K, and Lambon Ralph MA. The treatment of anomia using errorless learning. *Neuropsychological Rehabilitation*, 16(2): 129–154, 2006.

Fisher CA, Wilshire CE, and Ponsford JL. Word discrimination therapy: A new technique for the treatment of a phonologically based word-finding impairment. *Aphasiology*, 23(6): 676–693, 2009.

Franklin S, Buerk F, and Howard D. Generalised improvement in speech production for a subject with reproduction conduction aphasia. *Aphasiology*, 16(10–11): 1087–1114, 2002.

Goldrick M and Rapp B. A restricted interaction account (RIA) of spoken word production: The best of both worlds. *Aphasiology*, 16(1–2): 20–55, 2002.

Greenwood A, Grassly J, Hickin J, and Best W. Phonological and orthographic cueing therapy: A case of generalised improvement. *Aphasiology*, 24(9): 991–1016, 2010.

Herbert R, Best W, Hickin J, Howard D, and Osborne F. Combining lexical and interactional approaches to therapy for word finding deficits in aphasia. *Aphasiology*, 17(12): 1163–1186, 2003.

Hickin J, Best W, Herbert R, Howard D, and Osborne F. Phonological therapy for word-finding difficulties: A re-evaluation. *Aphasiology*, 16(10–11): 981–999, 2002.

Hillis AE. Efficacy and generalization of treatment for aphasic naming errors. *Archives of Physical Medicine and Rehabilitation*, 70(8): 632–636, 1989.

Howard D. Cognitive neuropsychology and aphasia therapy: The case of word retrieval. In Papathanasiou I (Ed), *Acquired Neurogenic Communication Disorders: A Clinical Perspective*. London: Whurr, 2000: 76–99.

Howard D. Lexical anomia: Or the case of the missing lexical entries. *Quarterly Journal of Experimental Psychology, Section A: Human Experimental Psychology*, 48(4): 999–1023, 1995.

Howard D and Harding D. Self-cueing of word retrieval by a woman with aphasia: Why a letter board works. *Aphasiology*, 12(4–5): 399–420, 1998.

Howard D, Hickin J, Redmond T, Clark P, and Best W. Re-visiting "semantic facilitation" of word retrieval for people with aphasia: Facilitation yes but semantic no. *Cortex*, 42(6): 946–962, 2006.

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

Laganaro M, Di Pietro M, and Schnider A. Computerised treatment of anomia in chronic and acute aphasia: An exploratory study. *Aphasiology*, 17(8): 709–721, 2003.

Leonard C, Rochon E, and Laird L. Treating naming impairments in aphasia: Findings from a phonological components analysis treatment. *Aphasiology*, 22(9): 923–947, 2008.

Levelt WJM, Roelofs A, and Meyer AS. Multiple perspectives on word production. *Behavioral and Brain Sciences*, 22(1): 61–75, 1999.

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

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

Nickels L and Howard DA. Frequent occurrence? Factors affecting the production of semantic errors in aphasic naming. *Cognitive Neuropsychology*, 11(3): 289–320, 1994.

Osborne F, Hickin J, Best W, and Howard D. Treating word-finding difficulties – Beyond picture naming. *International Journal of Language & Communication Disorders/Royal College of Speech & Language Therapists*, 33: 208–213, 1998.

Perkins L, Crisp J, and Walshaw D. Exploring conversation analysis as an assessment tool for aphasia: The issue of reliability. *Aphasiology*, 13(4–5): 259–281, 1999.

Plowman E, Hentz B, and Ellis C. Post-stroke aphasia prognosis: A review of patient-related and stroke-related factors. *Journal of Evaluation in Clinical Practice*, 18(3): 689–694, 2011.

Pederson P, Vinter K, and Olsen TS. Aphasia after stroke: Type, severity and prognosis. *Cerebrovascular Diseases*, 17(1): 35–43, 2004.

Pring T. *Research Methods in Communication Disorders*. London & Philadelphia: Whurr, 2005.

Raymer AM, McHose B, Smith KG, Iman L, Ambrose A, and Casselton C. Contrasting effects of errorless naming treatment and gestural facilitation for word retrieval in aphasia. *Neuropsychological Rehabilitation*, 22(2): 235–266, 2012.

Raymer AM, Ciampitti M, Holliway B, Singletary F, and Blonder LX. Semantic-phonologic treatment for noun and verb retrieval impairments in aphasia. *Neuropsychological Rehabilitation*, 17(3): 244–270, 2007.

Renvall K, Laine M, Laakso M, and Martin N. Anomia treatment with contextual priming: A case study. *Aphasiology*, 17(3): 305–328, 2003.

Renvall K, Laine M, and Martin N. Treatment of anomia with contextual priming: Exploration of a modified procedure with additional semantic and phonological tasks. *Aphasiology*, 21(5): 499–527, 2007.

Robson J, Marshall J, Pring T, and Chiat S. Phonological naming therapy in jargon aphasia: Positive but paradoxical effects. *Journal of the International Neuropsychological Society*, 4(6): 675–686, 1998.

Saur D, Ronneberger O, Kuemmerer D, Mader I, and Weiller C. Early functional magnetic resonance imaging activations predict language outcome after stroke. *Brain*, 133: 1252–1264, 2010.

Schwartz MF and Dell GS. Case series investigations in cognitive neuropsychology. *Cognitive Neuropsychology*, 27(6): 477–494, 2012.

Swinburn K, Porter G, and Howard D. *The Comprehensive Aphasia Test*. Hove: Psychology Press, 2004.

Tate RL, McDonald S, Perdices M, Togher L, and Schultz R. Rating the methodological quality of single-subject designs and n-of-1 trials: Introducing the Single-Case Experimental Design (SCED) Scale. *Neuropsychological Rehabilitation*, 18(4): 385–401, 2008.

Thompson CK, Kearns KP, and Edmonds LA. An experimental analysis of acquisition, generalisation, and maintenance of naming behaviour in a patient with anomia. *Aphasiology*, 20(12): 1226–1244, 2006.

Waldron H, Whitworth A, and Howard D. Therapy for phonological assembly difficulties: A case series. *Aphasiology*, 25(4): 434–455, 2011.

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

Ziegler W. Psycholinguistic and motor theories of apraxia of speech. *Seminars in Speech and Language*, 23(4): 231–244, 2002.