How many words should we provide in anomia therapy? A meta-analysis and a case series study

Claerwen Snell[a]; Karen Sage[a]; Matthew A. Lambon Ralph[a]

[a] University of Manchester, UK

PLEASE SCROLL DOWN FOR ARTICLE

Psychology Press
Taylor & Francis Group

How many words should we provide in anomia therapy? A meta-analysis and a case series study

Claerwen Snell, Karen Sage, and Matthew A. Lambon Ralph

University of Manchester, UK

Aims: This study investigated whether the number of words provided in naming therapy affects the outcome. A second aim was to investigate whether severity of anomia should be used to determine the number of words provided in therapy.

Methods & Procedures: First a meta-analysis of 21 anomia treatment studies between 1985 and 2006, yielding 109 individual datasets, explored whether the number of items provided and the severity of anomia influenced the success of therapy. The second part was a cross-over case-series study with 13 individuals with aphasia who had varying degrees of anomia. Individuals received two blocks of therapy (each of 10 sessions) where the set size of items to be learned was manipulated: either a small ($n = 20$) or large ($n = 60$) set in each block. Therapy and control sets were matched for baseline naming ability, frequency, phoneme, and syllable length. Therapy consisted of progressive phonemic and orthographic cues until successful naming was achieved. All word sets (small, large and control) were retested immediately after each therapy block finished (within 1 week) and 5 weeks after the end of each block of therapy.

Outcomes & Results: The meta-analysis showed a large variation in the number of items given to participants to learn (from 5 to 120 items) and very different learning outcomes that were not linked to the number of items given. The current literature contained an unexpected bias in that, across studies, more items were given to those with severe aphasia. Consequently, the meta-analysis could not provide a clear answer to how may items should be given in therapy—thus motivating a direct comparison in a new therapy study. We found significant gains in naming accuracy for both the small ($n = 20$) and large ($n = 60$) therapy sets immediately and 5 weeks post therapy. Proportionally, there was no difference between the two sets for the group as a whole, although there was individual variation in the overall therapy effect. If expressed as the raw numbers of words learned after therapy, this means that 12 of the 13 participants learned more words when given the large ($n = 60$) set. Severity of anomia correlated with learning performance but did not interact with set size.

Conclusions: The empicial study suggested that people with anomia could tolerate more items in therapy and that the severity of anomia should not necessarily determine how many words should be given in therapy.

Keywords: Anomia; meta-analysis; CVA; Therapy; Set size.

Anomia is a common feature of aphasia and can occur regardless of site or size of lesion, causing disruption at all levels of communication. Such disruption in communicative flow impacts on social participation, psychological well-being, family relationships, and quality of life (Kagan, 1998; Pound, Parr, Lindsay, & Woolf, 2000). One of the roles of the speech and language therapist (S&LT) is to identify the functional locus of the language difficulty and to provide appropriate and timely remediation (Chapey, 2001). Along the way they make a number of choices about what and how to deliver therapy for the anomia. Some of these choices are backed by research evidence, others are justified by expert clinical opinion, while still more are dictated by the environment and systems within which the therapy is delivered (NHS provision, health insurance package, etc.). There have been a number of studies suggesting which linguistic area to target for anomia therapy (e.g., improving semantics: Coehlo, McHugh, & Boyle, 2000; Francis, Clark, & Humphreys, 2002; or improving phonology: Miceli, Amitrano, Capasso, & Caramazza, 1996; Spencer et al., 2000). Emphasis has been on establishing the most beneficial therapy technique by looking at one particular therapy with one individual (in a single-case study) and measuring improvement in performance (Francis et al., 2002) or by comparing two different types of therapy in one or more individuals. These include, for example, comparisons of: semantic versus phonological cueing (Chin Li & Williams, 1989; Wambaugh, 2003; Wambaugh, Linebaugh, Doyle, & Martinez, 2001); personalised versus phonological cueing (Marshall, Freed, & Karow, 2001); constraint-induced aphasia therapy versus conventional aphasia therapy (Pulvermuller et al., 2001); decreasing versus increasing cues (Abel, Schultz, Radermacher, Willmes, & Huber, 2005; Conroy, Sage, & Lambon Ralph, 2009); errorless versus errorful therapy (Fillingham, Hodgson, Sage, & Lambon Ralph, 2003; Fillingham, Sage, & Lambon Ralph, 2006).

As well as what therapy method to use, other decisions have to be made such as how often, how much, and for how long therapy should be delivered. For these questions there are fewer studies to guide the clinician. Some studies have indicated that an intensive rate of therapy is most effective in treating aphasia (Bhogal, Teasell, & Speechley, 2003; Cherney, Patterson, Raymer, Frymark, & Schooling, 2008; Denes, Perazzolo, Piani, & Piccione, 1996; Pulvermuller et al., 2001). Other studies have also reported good outcomes for therapy provided on more traditional therapy schedules (e.g., Herbert et al., 2003). However, there is a distinct lack of studies on how many words should be given in a therapy programme and, as far as we are aware, no studies that directly explore this important clinical factor. Accordingly, in the first part of this study, we conducted a meta-analysis of the anomia therapy literature with the aim of gathering information on this question from across the existent studies. The second part of this paper reports a study of 13 individuals with anomia, which directly compared the number of items given in therapy. All other variables were held constant in order to allow a comparison of learning outcomes when the list was either short ($n = 20$) or long ($n = 60$).

META-ANALYSIS OF ANOMIA STUDIES (1985–2006)

Purpose

This meta-analysis combined the findings across multiple studies in an attempt to address the question of whether the number of words given in therapy had an effect on the outcome of that therapy and what role severity of anomia had on the results.

In addition, the analysis aimed to show whether these studies reflected clinical practice, i.e., that severity was a key factor when deciding the number of words to give to a patient.

Method

Selection of papers for the meta-analysis. The Web of Science database was used to search for peer-reviewed anomia therapy studies between 1985 and 2006, using key search words (aphasia; anomia; word finding difficulties; naming difficulties; confrontation naming; therapy; remediation; learning). A second-order search method used references cited in the identified articles.

Papers were included in the meta-analysis if the main aim of the study was treatment of word-finding difficulties by picture naming. Any acquired neurological aetiology giving rise to anomia—including cerebral vascular accident (CVA), head injury, dementia etc.—was allowed. Only studies targeting nouns were included while studies using pharmacological agents (e.g., bromocriptine) to improve word finding were excluded. In order to carry out statistical analysis of the grouped data, all studies had to have a control set of untreated items, which allowed for specific effects of therapy to be tested. Studies were excluded if they used as their sole control measure of therapy effect (i.e., for a measure of improvements on untreated items), published tests such as the Boston Naming Test (Goodglass, Kaplan, & Barresi, 2001) or Psycholinguistic Assessment of Language Processing in Aphasia (PALPA; Kay, Lesser, & Coltheart, 1992). Every included study had at least one baseline assessment of all therapy target items and control items so that the exact number of items that a participant could name before therapy began was clear. Where multiple baselines had been carried out, the highest baseline was used in this analysis. Where therapies were carried out concurrently on different word sets, the baselines were collapsed to provide one figure of total baseline naming for the target items. Where therapy occurred to word sets consecutively these were treated as separate data points in the analysis. In order to establish a measure of learning, every study had to indicate individual performance scores post therapy (so that a calculation of learning could be made: the pre-therapy score minus post therapy score). Group studies were included if this information was available for each individual.

A total of 71 studies were highlighted by the search methods described. Once the inclusion criteria had been applied, 21 studies remained (see Appendix 1). Appendix 2 shows the 50 studies that did not meet the criteria for inclusion in the analysis and the reasons why. The commonest reason was that individual changes in naming scores could not be calculated. The second largest exclusion was for studies where no control items were included in the study design. From the 21 included studies, 109 individual results were extracted for the meta-analysis.

Calculating performance in therapy. Because of variation in baseline naming accuracy, individuals (even within the same study) had different numbers of words available to learn. This number ranged from 5 to 120 words. To control for the varying baseline naming scores and prevent those studies with more items showing a greater therapy effect simply because they contained more targets, the proportion gain for each individual was calculated. This allowed a direct comparison of the performance of different individuals across different studies. The calculation took

into account the baseline naming ability for each person; thus the denominator was calculated by subtracting the baseline naming score from the total number of words provided in the therapy package. The numerator (total learned) was calculated by subtracting the baseline naming score from the total number of words correct after therapy. So, for example, in a therapy where the total number of items in the package was 30, 4 of which were named at baseline, the denominator would be 26. If the participant named 28 items correctly after therapy, the numerator would be 24. In this example, the individual would have learnt 0.92 of the items available-to-be-learned in the therapy. The same measure was used in the empirical study (see next section).

Results of the meta-analysis

Figure 1 shows the proportion gain in therapy (y axis) against the number of items in the therapy set (x axis) for the 109 individuals. Each dot on the graph represents one individual in a particular therapy task. The range of therapy items provided in therapy ranged from 5 to 120 (mean = 47). The range of proportion gain was from 0.04 to 1 (mean 0.46). For one case there was a proportional decrease in the number of items named after therapy (–0.67). As can be seen from the linear fit in Figure 1, the results extracted from the selected studies suggest a negative relationship between number of items provided in therapy and the proportion learned; that is, the more words an individual was given, the less they were able to learn proportionately (Spearman $\rho = -0.4$, $df = 107$, $p > .01$).

Figure 1 highlights that the set sizes were not evenly sampled in the literature. There was a cluster of studies that used between 30 and 40 therapy items (predominantly Fillingham, Sage, & Lambon Ralph, 2005a, 2005b, 2006). No set sizes between 73 and 99 were sampled. There was also a second cluster at 120 items
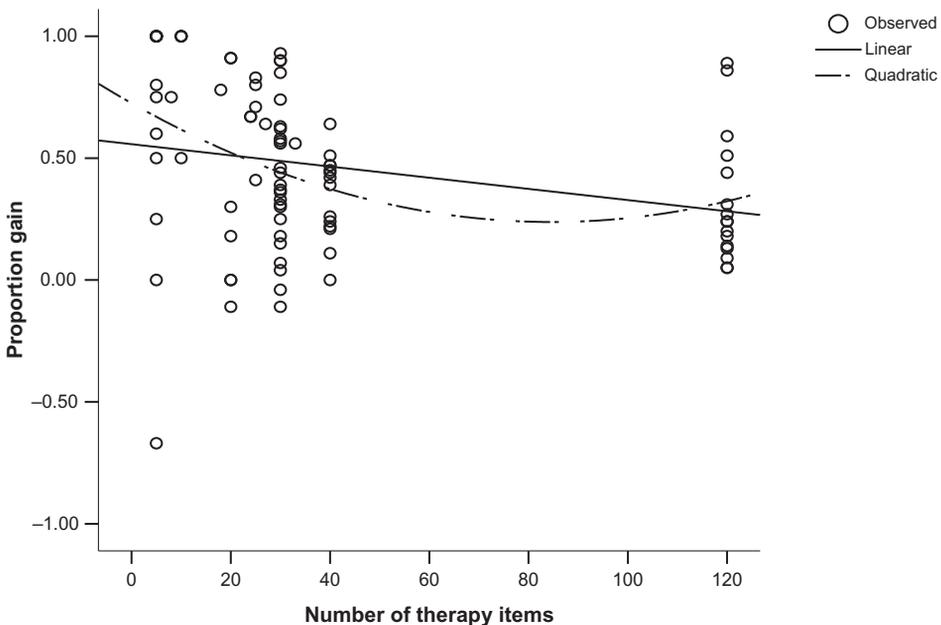

Figure 1. Proportion gain in therapy by set size.

(mainly from Best, Hickin, Herbert, Howard, & Osborne, 2000; Best, Howard, Bruce, & Gatehouse, 1997). As a result the relationship between number of items and therapy gain was best fitted with a higher-order function, producing a U-shaped curve. Specifically, by adding a quadratic term to the regression model it was able to account for 16% of the variance ($r^2 = .16$, $p = .001$) compared to the linear model, which was able to explain 10% of the variance. Of course, any meta-analysis is constrained by the studies contained within it—and although a simple (linear) consideration of the results would suggest that increasing the number of therapy items was detrimental, the positive outcomes for the studies with large set sizes indicate that the relationship is not so simple. In addition, further exploration of this literature unearthed an important confound that, again, calls this conclusion into doubt.

The second question for this meta-analysis was whether severity of anomia was a key factor in determining the number of items provided in therapy. To examine this across the 109 cases, a measure of naming severity common to all of them would be needed to allow comparison of severity and therapeutic gains across all the cases. However, there was no such common measure of naming severity. A wide variety of baseline naming measures were used, for example: The Graded Naming Test (McKenna & Warrington, 1983); The Boston Naming Test (Goodglass et al., 2001); The Western Aphasia Battery (Kertesz, 2006); PALPA naming (Kay et al., 1992); Philadelphia Naming Test (Roach, Schwartz, Martin, Grewel, & Brecher, 1996); and the Test of Adult Word Finding (German, 1990). Since there was no common measure of naming severity used across all studies, the analysis used the most common measure—the Boston Naming Test (Goodglass et al., 2001)—for which there were 40 reported data points. This smaller data set ($n = 40$) showed the same trend as the full data set ($n = 105$) in that the more words an individual was given, the lower the proportional gain but, due to lack of statistical power, the trend did not reach significance (Spearman's $\rho = -0.22$, $df = 39$, $p = .17$). Figure 2 shows the 40 individuals' proportion therapy
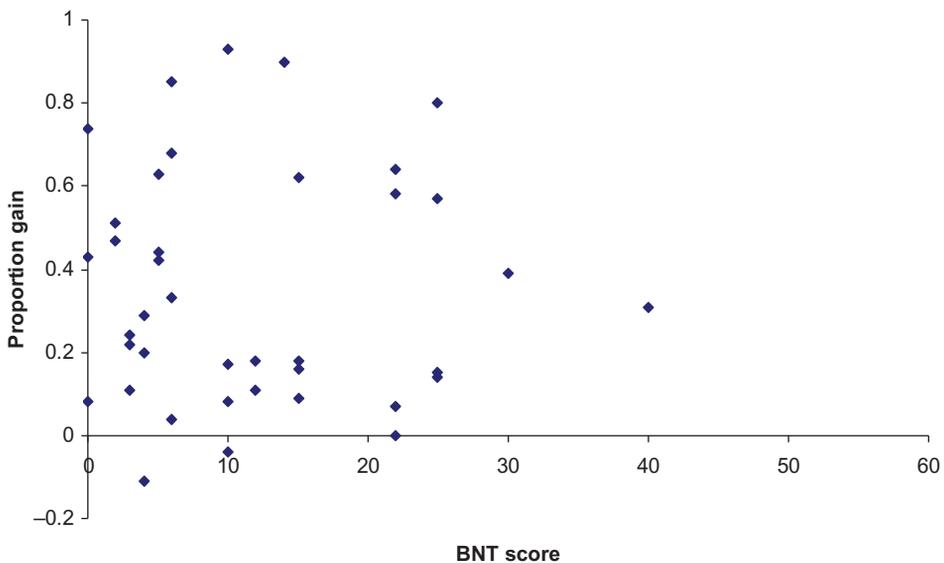

Figure 2. Proportion gain and BNT score.

gain plotted against their BNT score (Goodglass et al., 2001). Correlational analysis did find a relationship between severity and gain, such that those with milder naming impairments showed greater therapy gains (Spearman's $\rho = 0.34$, $df = 39$, $p < .05$).

To examine the effect of anomia severity on list size provided, the therapy gain for the 40 selected studies was correlated against individual BNT (Goodglass et al., 2001) score. The results showed that participants with milder impairments were given fewer items and those with more severe naming impairments were given more items (Spearman's $\rho = -0.36$, $df = 39$, $p < .05$). This slightly odd result, importantly for the previous analysis, indicates that the negative relationship between set size and therapy gain may have been due to this unforeseen bias within the therapy studies found in the literature. Indeed, when baseline naming is controlled in a partial correlation there is no relationship between therapy outcome and the number of items given in therapy ($r = -.23$, $df = 37$, $p = .15$).

Discussion

Using the existing research on anomia therapy, the meta-analysis set out to determine whether the number of words given in therapy affected the outcome of learning and whether naming severity influenced learning. The initial analysis suggested that the more items an individual was given, the less he/she learned, and individuals with milder naming impairments made bigger gains in therapy compared to those with more severe impairments. However, the current literature contained an unexpected bias in that individuals with milder naming impairments were also given fewer items to learn compared to those with more severe impairments. This bias does not, of course, reflect clinical practice.

In conclusion, the existing literature was not able to provide clarity concerning the relationship between the number of therapy items and therapy outcome, nor could it show how severity interacted with the number of words provided in therapy. In order to address these two issues directly we executed an empirical study in which individuals with varying degrees of anomia were given therapy sets of different sizes ($n = 20$ vs $n = 60$ items), thus permitting direct comparisons. A range of anomia severity was represented within the case-series so that the impact of severity on therapy performance could also be determined.

THERAPY STUDY

Method

Participant recruitment

A total of 13 right-handed participants with chronic aphasia following cerebral vascular accident (CVA) (mean post onset 44.77 months) took part in this study. Their anomia severity was measured using the Boston Naming Test (Goodglass et al., 2001) and the Graded Naming Test (McKenna & Warrington, 1983). All participants scored at least 2 standard deviations below the control mean on these naming tests. All participants were recruited from NHS speech and language therapy services and stroke clubs in the north-west of England, and informed consent was obtained for all participants using aphasia-friendly information.

Inclusion criteria

All participants had a left hemisphere CVA at least 6 months prior to the start of the study. All were native speakers of English and literate prior to their CVA. People with additional medical conditions, e.g., those with significant cognitive impairment, Parkinson's disease, or dementia, and people with significant visual agnosia or verbal apraxia were excluded from the study. Those participants who normally wore corrective glasses and/or hearing aids in daily life were encouraged to use them while carrying out the therapy programme.

Participant profiles and background assessment

Table 1 provides participants' age, gender, handedness, occupation, neuroimaging, and time post onset (months). For all figures and tables, participants are displayed in order of the naming severity on the BNT (Goodglass et al., 2001) from most severe (RR) to least severe (SB).

Participants completed comprehensive linguistic and cognitive assessments, the results of which are summarised in Tables 2 and 3 respectively.

Language assessments

Assessment of participants' language skills focused on the domains of naming, input and output phonology, semantics, and sentence processing.

Naming. (a) The Boston Naming Test (Goodglass et al., 2001) contains 60 black and white line drawings of increasing difficulty. Semantic and phonemic cueing were provided when the participant was unable to name within 10 seconds. All participants were at least 2 standard deviations below their age-appropriate controls.

(b) The Graded Naming Test (McKenna & Warrington, 1983) contains 30 black and white line drawings of low-frequency items. Mean control performance is 20.4 (SD = 4.1) (Warrington, 1997).

Phonology. Assessment of auditory discrimination, comprehension of sentences as well as reading and repetition tasks examined participants' abilities in processing phonological input as well as output. The following assessments were included in this section: (a) PALPA 2 (Kay et al., 1992): Auditory minimal word pairs. Participants heard pairs of words that were either the same (e.g., coat – coat) or differed by initial (e.g., bat – mat) or final (e.g., cup – cut) phoneme, or were metathetic (e.g., pack – cap). Lip reading was prevented. Participants were required to judge whether the pair of words they heard were the same or different.

(b) PALPA 31 (Kay et al., 1992): Imageability × frequency word reading. This test comprised 80 words that varied in frequency and imageability, providing four variations: highly imageable, high-frequency words such as "church", "fire", and "letter"; highly imageable, low-frequency words such as "axe", "gravy", and "tractor"; low-imageable, high-frequency words such as "attitude", "moment", and "thought"; and low-imageable, low-frequency words such as "dogma", "analogy", and "mercy".

TABLE 1
Biographical details including age, gender, handedness, occupation, scan results and Boston Naming Test score for all participants

	RR	FT	SM	DB	ER	JM	SS	LC	PG	JA	IH	FL	SB
Age	74	75	48	78	69	58	65	54	62	59	69	67	53
Gender	Female	Female	Male	Female	Male	Male	Female	Male	Male	Male	Male	Male	Female
Handedness	Right	Right	Right	Right	Right	Right	Right	Right	Right	Right	Right	Right	Right
CT/MRI	L temporo-parietal	L fronto-parietal extending into the insular	L temporo-parietal			L fronto-parietal and insular	n/a	L parietal-temporal	n/a	n/a	Subcortical	L frontal extending into insular	
Time post onset (months)	72	13	16	12	55	67	132	7	87	60	11	7	29
BNT ($n = 60$)	7	8	9	15	23	28	29	34	37	38	39	39	43

BNT = Boston Naming Test (Goodglass et al., 2001). n/a = not available. L = left.

TABLE 2
Background language results for all participants

Test	Max	Cut off	RR	FT	SM	DB	ER	JM	SS	LC	PG	JA	IH	FL	SB
GNT	30	12.2+■	0	0	4	1	5	8	11	9	10	13	16	9	10
TROG	20		11	14	1	7	4	_19_	12	9	6	10	15	17	11
P&PT	52	49+	_49_	48	33	37	47	46	42	_50_	38	46	42	44	**50**
PALPA 2	72	65	45	60	56	61	64	_71_	_68_	_65_	_66_	_72_	53	_71_	51
Letter fluency		21.8	1	0	4	7	11	17	9	10	2	8	9	4	5
Category fluency		29.2 and above	7	0	1	9	21	_36_	18	_30_	7	17	20	18	10
Synonyms judgements	96	91	76	74	n/a✳	48	75	66	82	69	77	58	62	77	63
Cookie theft description	wpm	53	_63_	_96_	1.2	19	24.62	_54.8_	_60_	8.6	12	34	36	n/a✚	31
PALPA 31 Reading	80	79–80	11	31	36	63	46	_79_	75	_79_	_79_	74	78	55	59
PALPA 9 Repetition	160	154*	95	69	126	130	49	128	139	124	126	112	115	103	97
PALPA 36 Nonword reading	24	23	0	4	0	8	0	13	12	17	3	0	21	10	6

n/a = not available. GNT = Graded Naming Test (McKenna & Warrington, 1983). TROG = Test for the reception of Grammar (Bishop, 1989). P&PT = Pyramids and Palm Trees (Howard & Patterson, 1992). PALPA = Psycholinguistic Assessment of Language Processing in Aphasia (Kay et al., 1992). Cookie Theft (Goodglass et al., 2001).

Underlined and **emboldened** scores were within the normal range.

✳Unable to understand task; ✚Unintelligible due to dysarthria.

■Normal cut-off for blocked presentation, mixed presentation was administered.

(c) PALPA 36 (Kay et al., 1992): Nonword reading. This test comprised 24 monosyllabic nonwords of varying lengths (3, 4, 5, 6 letters) derived by changing one phoneme of a word (e.g., "kid" – "ked", "smoke" – "smode"). Participants read the nonwords aloud.

(d) PALPA 9 (Kay et al., 1992): Imageability × frequency word and nonword repetition (mixed presentation). The test comprised 160 items. The 80 words were the same as those presented in the reading test (PALPA 31) and so varied in imageability and frequency. There were also 80 nonwords derived in the same way as the nonwords from PALPA 36 (e.g., clee – derived from clue; lutter – derived from letter).

Semantic memory and comprehension. (a) The three picture version of Pyramids and Palm Trees (Howard & Patterson, 1992), assesses semantic knowledge by requiring participants to match one picture to another (from a choice of two) on the basis of semantic relatedness; e.g., for a *pyramid*, the participant chooses between a *palm tree* and a *fir tree*.

(b) A synonym judgement test (Jefferies, Patterson, Jones, & Lambon Ralph, 2009) was used to detect mild semantic impairment. This test required participants to match words (presented simultaneously in written and spoken form) on the basis of semantic relatedness; e.g., for *rogue*, the participant should select *scoundrel*, and not

TABLE 3
Background cognitive assessments for all individuals

Test		Max	Cut off	RR	FT	SM	DB	ER	JM	SS	LC	PG	JA	IH	FL	SB
TEA	Without distraction	7	6	**6**	**6**	4	3	**6**	**7**	**7**	**7**	**7**	**7**	**7**	**7**	**6**
	With distraction	10	5	1	4	0◇	1	0□	**6**	**6**	**5**	**5**	2	2	**6**	4
Digit span	forwards		5	3	0	2	3	2	**5**	4	**5**	4	3	**5**	4	3
	backwards		2	0	0	0	0	2	**4**	**4**	3	2	0	3	**4**	2
Rey figure %ile	copy	36	25–32	31**>16**	262.5	25<1	19<1	15.5<1	32**>16**	22<1	3211	28<1	22<1	19<1	*	30**>16**
%ile	immediate recall	36	4–11	17**79**	6.512	0<1	0<1	41	24**90**	2<1	6<1	4.5<1	15**27**	51	*	114
%ile	delayed recall	36	3–12	18**86**	6.5<1	0<1	0<1	3.5<1	21**76**	1<1	5<1	71	117	6.52	*	5.5<1
RCPM		36		34	20	24	20	24	31	29	30	28	32	27	24	35
WCST	Number of categories completed	6	1	**5**	**3**	●	0	**2**	**4**	**4**	0	**1**	**2**	**3**	**1**	**4**
RAVLT	List A A5	15	13	2	2	0	2	6	2	9	4	4	2	5	2	9
	delayed recall	15	10	2	1	0	0	3	5	0	2	0	0	0	2	4
	recognition of list A	15	13	12	**13**	0	**15**	12	12	**13**	11	6	8	6	8	9

*Unable to do task due to arthritis; ●Did not understand task; ◇Unable to understand task; □Unable to do task due to sensori-neural hearing loss. TEA = Test of Everyday Attention (Roberston et al., 1994); RCPM = Ravens Coloured Progressive Matrices (Raven, 1962); WCST = Wisconsin Card Sort Test (Grant & Berg, 1993); RAVLT = Rey Auditory Verbal Learning Test (Schmidt, 1996).

Underlined and **emboldened** scores were within the normal range.

polka or *gasket*. Probe, target, and foils within each trial were matched for frequency and imageability, and these factors were varied across trials to produce an orthogonal manipulation of the two variables (high vs low frequency; low, medium, and high imageability).

Sentence comprehension. The Test for the Reception of Grammar (TROG) (Bishop, 1989) was administered to measure participants' auditory comprehension and understanding of grammatical contrasts. The participant heard a word or sentence and was required to match it to one of four pictures. For example, the participant heard "the boy has neither hat nor shoes" and saw four pictures: (1) a boy with hat and shoes; (2) a boy with no hat or shoes; (3) a boy with shoes; (4) a boy with a hat. Test performance was evaluated by number of blocks passed (Bishop, 1989).

Verbal fluency. (a) FAS (Benton, Hamsher, & Sivan, 1994). Participants were given 1 minute to name as many words beginning with the letters F, A, and S.
(b) Semantic categories (Mitrushina, Boone, & D'Elia, 1999). Participants were given 1 minute to name as many words from specific categories; animals; school; fruit and vegetables.

Connected speech sample. Participants described the well-known picture (Cookie theft; Goodglass et al., 2001) and from their description the total number of words per minute was calculated. Control mean and standard deviation were taken from Bird, Lambon Ralph, Patterson, and Hodges (2000).

Cognitive assessments

Assessment of participants' cognitive skills included measures of memory, executive and attention skills and auditory verbal learning.

Memory. (a) Forward digit span (Bachrach & Mintz, 1974). Participants listened to sequences of numbers that increased in length (from sequences of 2 up to sequences of 7). They repeated back, in order, the numbers that they had heard.
(b) Backward digit span (Bachrach & Mintz, 1974). Participants listened to sequences of numbers that increased in length (as in forward span testing). This time they repeated the sequence back in the reverse order to that which they had heard. This task required them not only to listen to information but also to manipulate that information.
(c) Complex Figure of Rey (Meyers & Meyers, 1995). This test investigated participants' visuo-spatial ability by requiring them to copy a complex geometric figure. They were then asked to reproduce it from memory 5 minutes later and again 30 minutes later, providing information about their visuo-spatial memory.

Attention. Test of Everyday Attention (Robertson, Ward, Ridgeway, & Nimmo-Smith, 1994). Two subtests of this test were used to look at attention to auditory material: (a) The elevator counting task looked at sustained attention. Participants heard a series of beeps at random time intervals and counted the number of beeps they heard (range 3–14). Written numbers were provided to enable responses from those with verbal number-naming difficulties. Use of their own fingers to indicate the number was also accepted.

(b) The elevator counting task with distraction looked at divided attention. Individuals heard a series of high and low beeps and counted only the low beeps, while ignoring the high beeps. Beeps were presented in a random order at random time intervals (range 2–14). As above, assistance was given for those with verbal number-naming difficulties.

Executive functioning. Wisconsin Card Sort Task (WCST; Grant & Berg, 1993). This test consisted of 128 cards, which varied according to three features: colour (red, yellow, blue, green), shape (circle, triangle, square, cross), and number (one, two three, four). Participants had four reference cards laid out in front of them, were given the pack of cards, and were asked to sort the cards using the reference cards to work out a sorting rule (e.g., sort by colour). The participants were told if their sorting mechanism was right or wrong. Once 10 consecutive cards had been placed correctly the tester changed the sorting rule (e.g., sort by number), without informing the participants of the rule change but indicating the change by providing feedback that their current sorting method was incorrect now. Rule changes were based on colour, shape, or number. The measures of performance examined was the number of categories completed ($n = 6$), which provided an indication of how many rules the participants were able to work out (10 consecutive correct responses).

Rey Auditory Verbal Learning Test (RAVLT; Schmidt, 1996). This test evaluated a participant's ability to learn word lists. A total of 15 words (list A) were read out to participants who then recalled verbally as many as they were able to from the list. This procedure was carried out five times (learning trials A1–A5). The change in the number of words recalled from trials A1–A5 was recorded. Following this, a second list (an interference list B) was read out once and participants recalled words from that list (B). They were then immediately asked to recall items from the first list (A) to provide a measure of learning after interference. Following a 20–45-minute delay participants were then asked to recall the first list (A) again as a measure of how well the participant recalls what was once learned. Finally, in a recognition task participants were provided with a list of 50 words (which included the 15 items on list A, the 15 interference items on list B, plus 20 words that did not appear in either list). Participants indicated those items that appeared in list A, thus giving a measure of how much had been recognised but could not be retrieved explicitly.

Details of participants

RR was a 74-year-old housewife and lived with her husband. She had a left-sided temporo-parietal CVA 72 months prior to the study. RR walked independently around the house and used a walking frame when she went out. Socially, RR spent time with her children and grandchildren. She had good comprehension of everyday conversation but had difficulty with spoken output characterised by word-finding difficulties and phonological errors. She was able to use alternative ways of getting her message across such as gesture, picture books, and getting the actual object for which she was unable to find the word or that would help to get her point across. On formal assessment, RR showed severe word-finding difficulties; her errors were predominantly phonological in nature. Her repetition skills were good when items were highly imageable and highly frequent. She made phonological errors, particularly on low-imageability items and nonwords (which were sometimes lexicalised). On the Pyramids and Palm Trees Test, she performed within normal limits but had difficulty

with more demanding semantic tasks (synonym judgements). Her cookie theft description was fluent with empty speech, word-finding problems, and use of light verbs to convey meaning. Cognitive assessment revealed impairments in divided attention and visuo-spatial and verbal memory, but strengths in sustained attention and executive functioning (WCST).

FT was a 75-year-old retired wages clerk who lived with her husband. In 2005, she had a CVA following a hip replacement. Scanning showed damage to the left fronto-parietal region extending into the insular. FT was able to walk independently and was still able to drive. Socially, FT saw many friends, family, and neighbours, and enjoyed shopping and going to church. She had good understanding of everyday conversation. She had significant difficulty with expression, her speech being charac-terised by numerous phonological errors. She was occasionally able to use gesture and pointing to get her message across. FT's severe anomia was evident on formal testing. Her cookie theft description was fluent with word-finding problems evident but she was able to make use of circumlocutions to convey her meaning. She was sometimes able to access the first letter of the target word. Her errors in naming read-ing and repetition tasks were predominantly phonological. On simple semantic tasks she performed just outside the normal range but had increased difficulty with more demanding semantic tasks. Cognitively, she demonstrated strengths in sustained attention and executive functioning, but had difficulty with divided attention and verbal and visuo-spatial memory.

SM was a 48-year-old man who lived in a residential home and who had worked as a radio sports commentator prior to his CVA. In 2005, he had a left temporo-pari-etal CVA leaving him with right-sided hemiplegia. He used a walking stick to mobi-lise. Socially, SM had become withdrawn; he occasionally socialised with other residents at the home but often spent time alone in his room. He did, however, go out regularly to the pub with relatives. On formal language tests, SM presented with sig-nificant auditory comprehension difficulties, which were alleviated to some degree when a total communication approach was employed. For the synonym judgement task and the WCST, he was unable to understand the task instructions, even when a total communication approach was used. Significant impairments in semantic and phonological processing were evident. He made semantic errors on naming and read-ing of which he was sometimes aware. He had limited spontaneous speech (occasion-ally producing single words). He was able to repeat words well but had difficulty in nonword repetition making lexicalisation errors. On the cookie theft picture, he sup-plied no verbal output but did make gestures and the sound "whoosh" while pointing to the water. Cognitive assessment indicated significant impairments in both sustained and divided attention, and verbal and visuo-spatial memory.

DB was a 78-year-old retired accountant who lived on her own. She had a left-sided CVA 1 year prior to participation in the study. She had severe physical disability with right-sided hemiplegia, and required carers to help with personal care and meals. Socially, DB had many friends and neighbours who visited, and she spoke to friends and family on the phone. She had good comprehension of everyday conversation provided background noise was kept to a minimum and her conversation partner reduced their rate of speech. On formal language tests, DB presented with moderate-severe word-finding difficulties; her errors were mainly omissions but she was sometimes able to gesture when she could not produce the target word. She also made a number of semantic and perseverative errors. She had relative strengths in word repetition and reading words aloud, although this was affected by imageability

and frequency. She produced lexicalisations when reading nonwords. Her cookie theft sample was non-fluent, with some light verbs and evidence of word-finding and some phonological errors. DB was impaired on all the cognitive tasks.

ER was a 69-year-old widower and retired sheet metal worker who lived on his own. He had a left-sided CVA with a temporo-parietal lesion in 2002. He presented with right-sided hemiparesis and limb apraxia. He used a stick to mobilise and had carers to help with his meals. Socially, he saw his sons at the weekend and his neighbours occasionally. He infrequently phoned friends and family. He had a noise-induced, high-frequency, sensori-neural hearing loss. His understanding of everyday conversation was generally good but some repetition was required. His speech was characterised by word-finding difficulties and he was sometimes able to cue himself in by writing the first letter or letters of the word. On formal language tests, he was unable to repeat or read nonwords, making lexicalisation errors. He made semantic and phonological errors in all output tasks including the cookie theft sample, which was fluent. Cognitive assessment revealed verbal and visuo-spatial memory difficulties. It was not possible to assess ER's divided attention using the elevator counting task with distraction (Robertson et al., 1994) because his sensori-neural hearing loss prevented him hearing the distractor tone.

JM was a 58-year-old man who lived with his wife. Prior to his stroke he worked for a computer company and was a football coach for his local team. In 2001, he had a left fronto-parietal stroke. JM presented with right hemiparesis and walked independently. Socially, he saw his children and grandchildren, he also attended the local stroke club and participated in various courses, e.g., computer training. He gave a talk each year for speech and language therapy students at a local university. He remained involved with the football team as an assistant coach. He had excellent comprehension of everyday conversation and fluent speech with some word-finding difficulties. He was able to use strategies such as gesture or circumlocution to overcome these difficulties. On formal language assessment, he demonstrated moderate anomia, where his errors were either omissions or semantic. He responded to phonemic cues. He had difficulty reading and repeating nonwords, producing lexicalisation and phonological errors. In contrast, his reading and repeating of words were excellent. He performed within normal limits on auditory sentence comprehension (TROG; Bishop, 1989). His cookie theft description was short but fluent and grammatically correct. He made one semantic error, calling the mother "lad". He showed a mild semantic impairment on the Pyramids and Palm Trees test (Howard & Patterson, 1992) and greater impairment on more demanding tests of semantics. On cognitive tasks, he showed excellent skills in attention (TEA), the Rey Figure, and Raven's Progressive matrices. His forward and backwards digit span were within normal range. JM had the strongest cognitive profile on testing.

SS was a 65-year-old widow and retired school secretary who lived on her own. She had a left-sided CVA in 1995. She was independently mobile but since her stroke was no longer able to drive. Socially, SS saw her friends and family frequently, she also attended the local stroke club and church regularly. Her understanding of everyday conversation was good but she became confused when too much information was given at once. Her speech was fluent with occasional word-finding difficulties; she was able to use circumlocution effectively as a strategy. On formal language tests, she presented with moderate anomia; she used circumlocution as a strategy to facilitate production of the target word. Her errors were semantic in nature; she did not make phonological errors. She responded well to phonological cues. She was able to carry

out the repetition and reading tests when these involved words, but when nonwords were used she produced lexicalisations and phonological errors. Her cookie theft sample was fluent with evidence of both word-finding difficulties and semantic errors. On cognitive testing, she showed difficulties in visuo-spatial memory and verbal recall memory but her recognition memory was within normal limits (RAVLT; (Schmidt, 1996).

LC was a 54-year-old retired railway maintenance worker who lived with his wife and daughter. He had a left parieto-temporal CVA 7 months prior to the start of the study. Socially, LC spent time at the local rugby club with his wife. His understanding of everyday conversation was good but was hindered by noisy environments. On formal language tests, he had difficulty with semantic processing when the task was more demanding (e.g., synonym judgement) and showed an imageability effect. He presented with a moderate anomia, which responded to phonemic cues. His errors on naming were both semantic and phonological. He performed within normal limits for word reading and repetition but showed moderate difficulty with nonword reading, making phonological and lexicalisation errors. His cookie theft sample was non-fluent agrammatic with little information provided and a preference for naming the items present. Cognitively, he performed within normal limits for both sustained and divided attention but had marked difficulty with tests of executive function (WCST; Grant & Berg, 1993).

PG was a 62-year-old retired architect who lived with his wife. He had a sub-arachnoid haemorrhage in 2005 after which he remained independently mobile with a right-sided hemiparesis. Socially, he saw his friends regularly and attended various art (drawing, painting) and gardening classes. He had good understanding of everyday conversation but occasionally required some repetition. His speech was non-fluent and characterised by word-finding difficulties. His cookie theft description reflected this pattern. He was able to use alternative means to get his message across (gesture, using objects in the environment). On formal language tests, he presented with a moderate anomia and his errors were predominantly semantic in nature. He responded well to phonemic cueing. His reading of words was within normal limits but he produced lexicalisations on nonword reading. His word repetition was relatively good but was outside the normal range for low-imageable, low-frequency words and he had marked difficulty with nonword repetition, making numerous lexicalisations. Cognitive assessment revealed performance within normal limits for sustained and divided attention and executive function. He had particular difficulty with verbal memory (RAVLT; Schmidt, 1996).

JA was a 59-year-old retired bricklayer who lived with his wife. He had a left-sided CVA in 2001. He presented with a right-sided hemiparesis and mobilised around the house with a stick. He used an electric scooter to go out. Socially, JA saw his family regularly and attended the local stroke club. He enjoyed painting and doing jigsaw puzzles. His understanding of everyday communication was good although he occasionally required repetition. If he found a task difficult he would become distracted. On formal language tests he presented with a moderate anomia but was facilitated by phonemic cues. His errors were mainly omissions but he also produced some semantic errors. His cookie theft description was non-fluent, with reduced content and phonemic and some unintelligible speech errors. He was able to read and repeat words but produced phonological errors on nonword repetition and nonword reading. He would either use the nonword to form the first part of a real word, e.g., "nar" = narrow boat, "lat" = lattice, or he would

spell out the nonword but be unable to say it. Cognitively, he had good sustained attention but difficulty with divided attention. He also had marked difficulty with verbal memory.

IH was a 69-year-old retired processing manager who lived with his wife. He had a sub-cortical CVA in 2005. He had a right-sided hemiparesis and was mobile with a stick over short distances. He had good understanding of everyday conversation but had more difficulty in situations with increased background noise. His speech was fluent with word-finding difficulties. This pattern was reflected in his cookie theft description, which was fluent with correct but reduced content. On formal language tests, he presented with moderate anomia and produced semantic errors and responded well to phonemic cueing. He had good reading and repetition of words but some difficulty with nonwords, making lexicalisation and phonological errors. His nonword reading was good, producing only three errors all of which were lexicalisations. Cognitively, he performed well on tests of executive function and sustained attention but had marked difficulty with divided attention.

FL was a 67-year-old widower and retired lorry driver. He lived alone in warden-controlled accommodation. In 2006, he had a left-sided frontal CVA, which extended into the insular. Socially, he saw his children and grandchildren regularly. His understanding of everyday conversation was good and he was an avid listener of BBC Radio 2. His speech was fluent but dysarthric, becoming more so in connected speech. His cookie theft description reflected this speech difficulty in that its content was not decipherable. He was able to employ strategies to increase intelligibility such as reducing rate of speech, and taking regular breaths. On formal language tests, he presented with a moderate anomia and produced phonological errors; he would occasionally respond to phonemic cues. He had mild-moderate difficulties on semantic tests, with an imageability effect on more demanding tasks. His auditory discrimination skills were within normal limits. He had difficulty with both word and nonword repetition and reading, which was not attributable to his dysarthria. FL showed good attention skills but had more difficulty with the Wisconsin Card Sorting task. He was unable to complete the Rey copy because of severe arthritis. He had a reduced digit span.

SB was a 53-year-old retired shop assistant who lived with her husband. She had a left-sided CVA 29 months prior to the study. Socially, SB regularly saw her children and grandchildren; she attended the local stroke club and participated in various courses (sewing, computers, art, gardening). Her understanding of everyday conversation was good but she had increased difficulty when background noise was present. Her speech was fluent with some dyspraxia errors. This was reflected in her cookie theft description, which was fluent but revealed some word-finding problems, as well as phonological and speech errors. On formal language tests, she presented with a mild-moderate anomia with semantic errors on naming. She was helped by phonemic cues. Her semantic processing was within normal limits for less-demanding tasks but a semantic impairment was evident as task demands increased. Difficulty with phonological processing was apparent, with impaired performance on auditory discrimination, word and nonword repetition, and reading. Her errors on these tasks were phonological. Cognitively, she performed within normal limits for sustained attention and executive functioning but had difficulty with divided attention and verbal and visuo-spatial memory.

Therapy method

Item selection for therapy. In order to make up individualised therapy and control sets, each participant was asked to name on three separate occasions a set of 400 pictures. They were compiled from the Snodgrass pictures (*n* = 236; Snodgrass & Vanderwart, 1980), the Category Specific Naming Test (*n* = 105; McKenna, 1998), and pictures from the Internet selected to meet specific frequency requirements (*n* = 59 pictures). Written frequency ratings were obtained from the MRC psycholinguistic database (Coltheart, 1981). If the participant succeeded in naming a picture on all three occasions, that picture was eliminated as a potential therapy item. If they were unable to name the item at all, or named it in only once out of the three opportunities, this item became available for therapy. From this cohort of pictures, a set of 120 items were selected for each participant.

Therapy sets. For each participant, the 120 selected items were divided into four sets; two experimental sets (*n* = 60, *n* = 20) and two control sets (each *n* = 20). The four sets were matched for Celex combined spoken and written frequency (Baayen, Piepenbrock, & Rijn, 1993), number of phonemes, number of syllables, and baseline naming ability (whether the item had been named once or not at all over the three baselines). For each participant, no set was significantly different on these variables. Allocation for which therapy set was treated first was decided using random number generation. Those participants who were assigned an even number completed the small set first and those assigned an odd number started with the large set. Therapy to each set was delivered sequentially. Order of item presentation within the sets was random but, once the order of pictures was decided, this order was maintained throughout the therapy sessions.

Method of therapy. All the pictures in the therapy set (either 20 or 60) were presented three times at every session for naming. In both conditions there were 10 sessions, and session duration ranged from 30 to 90 minutes. If there was no response within 10 seconds or the response was incorrect, progressive phonemic and orthographic cues were provided. If the target item was not produced after 5 seconds, the next phoneme and corresponding letter(s) was/were provided. Words were divided up into a maximum of four cues. For example: flag = /fl/ /flæ/ /flæg/, gorilla = /gə/ /gər/ /gərIl/ /gərIlla/. Cueing continued until the target item was produced or until the whole target word had been presented in spoken and written format by the therapist. If the whole word was provided via cueing, then the participant was encouraged to repeat/read the target item. The next picture was then presented. A record of performance across all three naming attempts during the therapy sessions was recorded.

Post-therapy testing. The therapy set and its corresponding control set were tested immediately after the therapy (within 1 week) and again 5 weeks after the therapy, in order to examine the immediate effects of therapy to the treated items, to ascertain whether there was any effect on untreated items and to find out how well words were retained following a period without specific therapy. The therapy and control items were mixed together and presented in random order at these two testing times.

Results

The results are separated into two sections to address the two research questions. First, the question of whether the number of words given in therapy affected how many words were learned was addressed; second, the role severity played in participants' performance was examined. The data were analysed at both group and participant level. Group results are presented first followed by any participant results that differed from the group performance. Since there were three baseline scores for each picture, the highest baseline score was used in the analysis of therapy and retention effects, i.e., when comparing performance between baseline and immediate and 5-week assessments. Participant results are presented using their scores on BNT (Goodglass et al., 2001) to order the participants, with the most severe on the far left. The group mean is shown on the far right of all figures. To allow for direct comparison between the two therapy sets, the proportion gain for each participant was calculated. As well as making the different-sized therapy sets directly comparable, this also allowed each participant's baseline naming score to be taken into account. To do this, the same calculation as that used in the meta-analysis in Part 1 was used.

Comparison of performance between the small ($n = 20$) and large ($n = 60$) therapy sets

To test for a specific effect of therapy an ANOVA was carried out that included time (before, immediately after, and at follow-up) and set type (treated versus untreated items). As expected the interaction between time and set type was significant, $F(2, 26) = 44.6$, $p < .001$, reflecting considerable improvement in naming accuracy over time for the treated but not the untreated items. Treated and untreated items were then explored separately; see below for details.

To investigate whether the number of words provided in therapy affected how many words were learned during therapy, performance across the small ($n = 20$) and large ($n = 60$) therapy sets was compared. The number of words given in therapy did not affect the proportion of words learned at either the immediate or follow-up assessments, $F(2, 22) < 1$ (see Figure 3). Remember that this value is expressed as a proportion to enable direct comparisons between sets, such that, to obtain the same proportion in the two sets, the participants actually learnt more items in the larger set. To confirm this formally we repeated the analyses with the raw scores. The ANOVA confirmed a significant effect of therapy: improved accuracy over time, $F(2, 22) = 105.56$, $p < .001$. There was a significant effect of set size—20 vs 60 items: $F(2, 22) = 57.5$, $p < .001$—and a significant effect of set type—therapy vs control items: $F(2, 22) = 97.43$, $p < .001$. All interactions were significant: time × set size: $F(2, 22) = 26.37$, $p < .001$; time × set type: $F(2, 22) = 71.28$, $p < .001$; set size × set type: $F(2, 22) = 47.09$, $p < .001$; time × set size × set type: $F(2, 22) = 52.24$, $p < .001$. To establish where the significant effects were coming from, ANOVAs were carried out on the therapy and control items separately. Analysis of the therapy items showed that, in terms of the raw numbers, participants learned more words in the large (mean number learned = 39, $SD = 14.73$) than the small therapy set (mean number learned = 15, $SD = 5.03$). The ANOVA confirmed a significant effect of therapy—accuracy over time: $F(2, 22) = 96.86$, $p < .001$—set size—60 > 20 items: $F(2, 22) = 56.77$, $p < .001$—and a significant interaction—time × set size: $F(2, 22) = 41.92$,
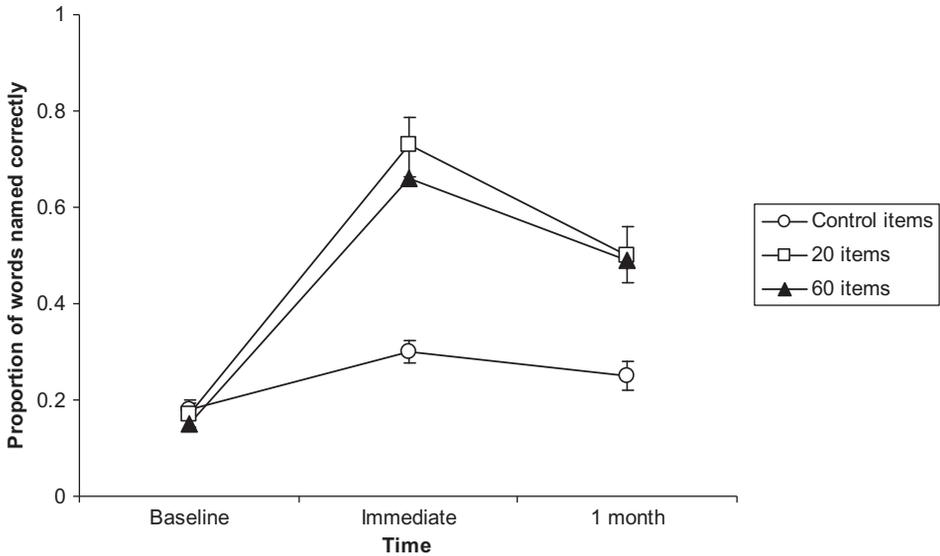
Figure 3. Group means and standard errors for therapy and control items at baseline, immediate, and 1-month follow-up.

$p < .001$. Post hoc analyses confirmed significant differences between the two sets (60 > 20 sets) at all time points—baseline: $t(12) = 4.69$, $p < .001$; immediate: $t(12) = 7.8$, $p < .001$; 5 weeks: $t(12) = 7.34$, $p < .001$. There were minimal changes to the control set. The ANOVA showed a main effect of time, $F(2, 22) = 6.63$, $p < .005$, no main effect of set size was found, $F(2, 22) < 1$. The interaction between time and set size was approaching significance, $F(2, 22) = 3.29$, $p = .053$.

Although all participants improved across both sets in their ability to name, there was some individual variation (see Figures 4 & 5). For example, immediately post therapy three participants performed significantly better on one of the therapy sets compared to the other. SS and IH performed significantly better (proportionally) on the 20-item set (SS: $\chi^2 = 4.75$, $df = 1$, $p < .05$; IH: $\chi^2 = 7.45$, $df = 1$, $p < = .01$). In contrast, FL performed significantly better on the 60-item set ($\chi^2 = 9.94$, $df = 1$, $p < .01$). However, in terms of raw numbers, it was still the case that all three participants learned a greater number of items with the 60-item set (SS named 20 from the small set compared with 45 on the large set, IH named 18 on the small set and 31 on the large set, and FL named 10 on the small set and 36 on the large set).

At the 5-week follow-up two participants showed variation from the group result. RR was significantly better on the 20-item set than the 60-item set ($\chi^2 = 3.82$, $df = 1$, $p < .05$), while JA was significantly better on the 60-item set than the 20-item set ($\chi^2 = 5.1$, $df = 1$, $p = .02$). RR retained the same number of words regardless of the set size (8 words in each). JA named substantially more items with the large set (11 items from the small set and 52 from the large set).

McNemar tests (one-tailed; summarised in Table 4, raw scores are given in Table 5) were carried out to determine whether participants were learning and retaining therapy items. For the small and large therapy sets individual performance was compared by examining changes in accuracy across baseline and immediately post-therapy, baseline and 5 weeks post therapy, and across the two post-therapy assessments

(a)
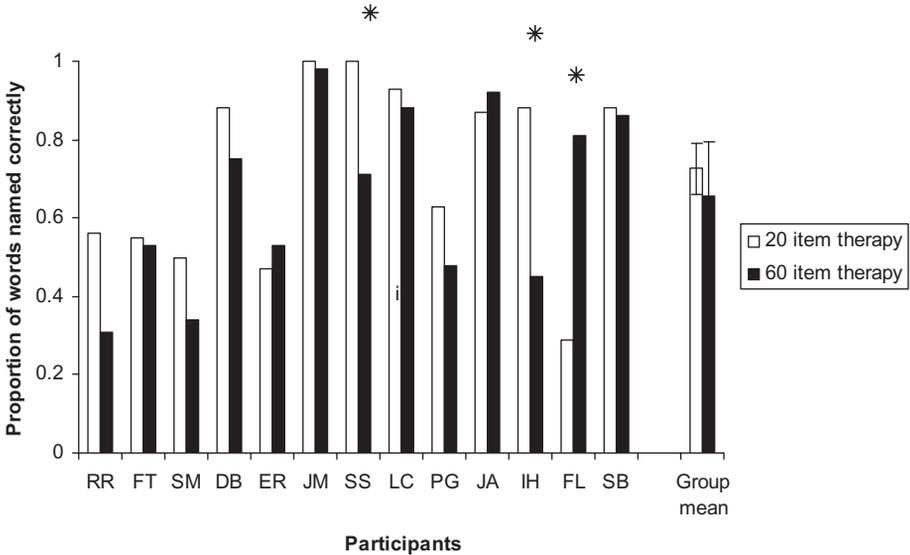
∗ = significant difference between set sizes

(b)
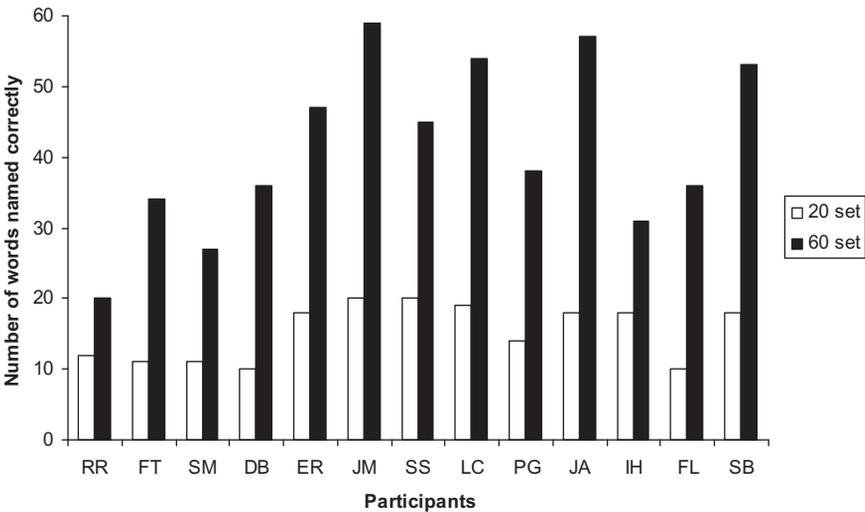
Figure 4. Therapy study: individual performance immediately post therapy proportion scores (a) and raw scores (b).

(immediate and 5 weeks). On the small therapy set all participants were able to learn items after therapy (that is, they were all able to name significantly more items after the therapy than they could at the start). This learning was significant for 11 of the 13 participants. For the small set 10 of the participants maintained this learning 5 weeks after therapy. RR and FL's performance dropped back to their baseline naming level

(a)

∗ = significant difference between set sizes

(b)
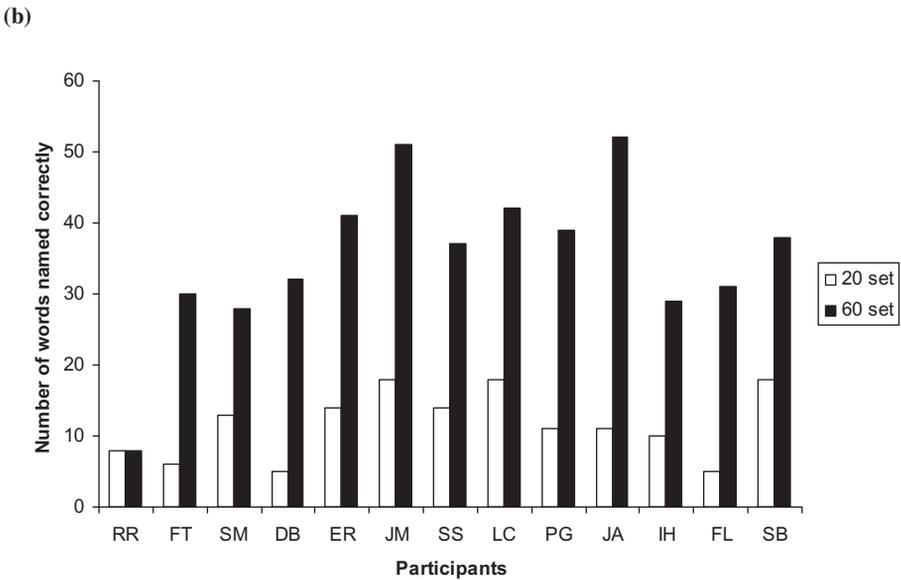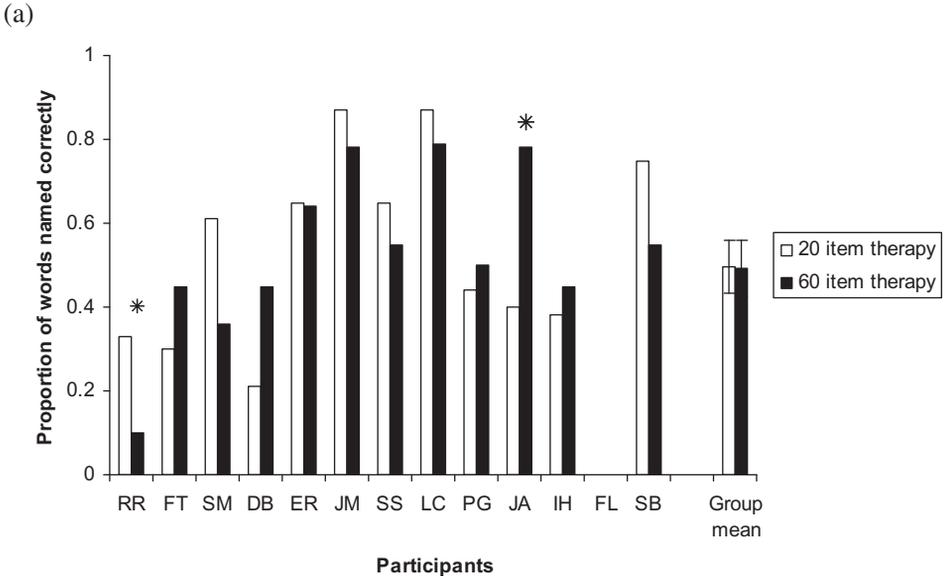
Figure 5. Therapy performance: individual performance 5 weeks post therapy proportion score (a) and raw scores (b).

immediately post therapy (baseline naming RR = 7 items, FL = 6 items; 5 weeks post therapy RR = 8 items, FL = 5 items) whereas DB was naming more items than she had been at baseline (1 vs 5 items) but this was not significant. For the large therapy items all participants learned significantly and all maintained their learning five weeks after therapy. When accuracy between the two post therapy assessments was assessed, for the small therapy set four (SS – McNemar one-tailed, $p = .01$; JA – McNemar $p < .01$; IH – McNemar one-tailed, $p < .01$; FL – McNemar one-tailed,

TABLE 4
McNemar *p* values

| | 20 set | | | | | | 60 set | | | | | |
| | *Therapy* | | | *Control* | | | *Therapy* | | | *Control* | | |
	Baseline-immediate	Baseline-1 month	Immediate-1 month	Baseline-immediate	Baseline-1 month	Immediate-1 month	Baseline-immediate	Baseline-1 month	Immediate-1 month	Baseline-immediate	Baseline-1 month	Immediate-1 month
RR	ns	*	+	ns	ns	ns	**	*	*	ns	ns	ns
FT	**	*	+	*	ns	**	**	**	ns	ns	ns	ns
SM	**	**	ns	*	ns	ns	**	**	ns	ns	+	+
DB	**	+	+	ns	ns	ns	**	**	ns	ns	ns	ns
ER	**	**	+	+	ns	+	**	**	+	ns	ns	ns
JM	**	**	ns	**	+	ns	**	**	**	ns	ns	ns
SS	**	**	*	ns	ns	ns	**	**	*	+	**	ns
LC	**	**	ns	ns	ns	ns	**	**	**	ns	ns	ns
PG	**	*	ns	ns	ns	ns	**	**	ns	ns	+	ns
JA	**	*	*	ns	ns	ns	**	**	+	ns	ns	ns
IH	**	**	**	ns	ns	ns	**	**	ns	ns	+	**
FL	**	*	*	ns	ns	ns	**	**	ns	ns	ns	ns
SB	**	**	ns	ns	ns	ns	**	**	**	ns	ns	ns

Therapy items = one tailed; control items = two tailed) results for individual performance on small ($n = 20$) and large ($n = 60$) therapy sets between different time points (baseline and immediate assessment; baseline and 5-week assessment; immediate and five week assessments).

ns = not significant; $+.10\ p < .05$; $*p < .05$; $**p < .01$.

TABLE 5
Raw scores for all participants (therapy and control items) at baseline, immediate, and 5-week
assessments

Participant	RR	FT	SM	DB	ER	JM	SS	LC	PG	JA	IH	FL	SB
20 items													
Baseline													
Therapy	2	0	2	1	3	5	3	5	4	5	4	6	4
Control	0	0	1	2	4	7	5	6	4	4	5	2	6
Immediate													
Therapy	12	11	11	10	18	20	20	19	14	18	18	10	18
Control	2	7	6	3	9	12	4	6	6	6	3	5	8
Follow-up													
Therapy	8	6	13	5	14	18	14	18	11	11	10	5	18
Control	3	0	4	4	5	12	5	7	6	7	3	2	8
60 items													
Baseline													
Therapy	2	5	10	9	7	19	9	12	18	24	4	9	11
Control	1	1	4	2	1	5	6	1	5	6	6	4	4
Immediate													
Therapy	20	34	27	36	47	59	45	54	38	57	31	36	53
Control	1	3	5	5	2	7	2	1	7	8	4	6	7
Follow-up													
Therapy	8	30	28	32	41	51	37	42	39	52	29	31	38
Control	2	0	9	3	2	3	8	2	10	7	11	6	6

$p < .05$) out of the thirteen participants showed a significant drop in naming accuracy, while in the large therapy set five (RR – McNemar one-tailed, $p < .01$, JM – McNemar one-tailed, $p < .01$; SS – McNemar one-tailed, $p < .05$; LC – McNemar one-tailed, $p < .01$; SB - McNemar one-tailed, p < .01) showed a significant drop in naming accuracy.

For the control items individual performance was again compared by examining changes in accuracy across baseline and immediately post-therapy, baseline and 5 weeks post therapy, and across the two post-therapy assessments (immediate and 5 weeks). On the small set, three of the participants significantly improved (FT – McNemar one-tailed, $p < .01$; SM – McNemar one-tailed, $p < .05$; JM – McNemar one-tailed, $p < .01$) but these gains were not maintained at the 5-week follow-up. When looking at performance at the two post therapy assessments only one participant demonstrated a significantly different performance: this was FT (McNemar one-tailed, $p < .01$) whose naming of the control items dropped back to zero at the 5-week assessment. On the large therapy set none of the participants improved on naming accuracy of the control items immediately after the therapy. However, at the 5-week assessment one participant (SS: McNemar one-tailed, $p < .01$) showed a significant improvement in naming accuracy when compared to baseline that was not present immediately after therapy. When performances on the two post-therapy assessments were compared, one participant (IH: McNemar one-tailed, $p < .05$) showed significant gains in naming accuracy (baseline = 6, immediate assessment = 4, 5 weeks = 11).

Severity

The second aim of this study was to investigate the effect of naming severity on therapy, in particular whether a participant's severity predicted whether they performed

better with a small ($n = 20$) or large ($n = 60$) set of words. The measure of severity used in the analysis was BNT score. Proportion gain immediately post therapy for the small ($n = 20$) and large ($n = 60$) sets were used as measure of learning performance. An analysis of covariance indicated that—while severity was an important factor in participants' ability to learn, $F(1, 11) = 7.37$, $p = .02$, in that the participants with more severe anomia learned less overall than participants with milder anomia—naming severity did not interact with set size, $F(1, 11) < 1$.

GENERAL DISCUSSION

This study investigated (1) whether the number of words in a therapy set influenced the number of words learned in aphasia therapy and (2) whether list size interacts with the severity of the anomia. There were two components: a meta-analysis of the literature and an empirical study of word learning in a diverse group of anomic participants. The initial meta-analysis suggested that the more items an individual was given, the less he/she learned, and individuals with milder naming impairments made larger gains in therapy compared to those with more severe impairments. However, the current literature contained an unexpected bias in that individuals with milder naming impairments were also given fewer items to learn compared to those with more severe impairments. Thus, the existing literature is not able to provide clarity concerning the relationship between the number of therapy items and therapy outcome, nor can it show how severity interacts with the number of words provided in therapy.

In order to address these two issues directly we executed an empirical study in which individuals with varying degrees of anomia were given therapy sets of different sizes ($n = 20$ vs $n = 60$ items), thus permitting direct comparisons. A range of anomia severity was represented within the case-series so that the impact of severity on therapy performance could also be determined. Learning was assessed twice: once immediately post therapy and again after a period of 5 weeks in which no therapy occurred. For the group as a whole, set size had no impact on naming accuracy, with equivalent proportional improvements on both the small ($n = 20$) and large ($n = 60$) sets at both post-therapy assessments. This is especially striking given that, when converted back to raw numbers, all participants relearned many more picture names in the large than small sets. These results indicate that many participants may tolerate, and benefit from, much larger set sizes in therapy.

Examination of individual participants' performance revealed some individual differences, both immediately and 5 weeks post therapy. Immediately post therapy two participants—SS and IH—performed better proportionately on the small set of therapy items. Neither participant differed from the rest of the group in terms of age, adjustment to aphasia, or social life, although SS was the participant who was the longest time post onset (132 months). In terms of naming ability neither participant differed from the rest of the group. However, both participants showed the largest drop between trial A5 and the delayed recall on the RAVLT, suggestive of a fast forgetting rate. It might be, therefore, that for such participants a smaller learning set was tolerated better as it made fewer demands on their memory. Studies comparing errorless and errorful learning in people with anomia following CVA (Fillingham et al., 2006) found that learning was greater and was retained for longer in those participants who had better recognition memory and executive/problem-solving skill. These results raise the question of whether other cognitive factors may be relevant when considering what size list to provide for a client, over and above the characteristics or severity of the anomia itself.

At follow-up only JA performed (proportionately) better on the larger therapy set than the small. His profile did not differ from the group in terms of age, time post onset, adjustment to aphasia, or social life, or in terms of anomia measures, but he was vocal about the strategies he employed that might have helped him retain individual items. For example, on one page the target item began with "C" and the next item happened to begin with "D"—he used the fact that these letters were adjacent in the alphabet to cue himself in with "D". He would also tell stories involving the target item. Within the group he had one of the milder anomias, but he was not an outlier in terms of any of the assessment results. These two cases (RR and JA) underlie the necessity not only to consider the anomia within the language disorder but also to examine cognitive skills and learning strategies shown by each participant when planning anomia therapy.

The study also asked whether the severity of anomia interacted with the set size given to a participant in therapy. This study found that naming severity predicted how much each participant could learn, with participants with milder impairments learning more than those with more severe impairment. However, severity did not interact with the number of words given in therapy. The study demonstrated that, irrespective of anomia severity, when more words were given, more words were learned. There was only one exception to this rule, participant RR with the severest anomia, who learned eight items no matter what the set size. Overall, these results suggest that in clinical practice, speech and language therapists might be able to provide far more items than is traditionally done irrespective of naming severity. However, this finding also needs to be weighed up against practical considerations such as the increased time needed to provide and work with more items. This might be offset by reducing the amount of times an item is presented in therapy, although this might reduce the therapy effect. However, in the study by Hickin, Best, Herbert, Howard, and Osborne (2002), participants were able to show significant increases in naming after seeing each item once only per session (over eight sessions).

Conclusion

Participants post CVA, regardless of severity, were able to learn large numbers of words when given the opportunity. This appears to be true not only immediately after therapy but also after a follow-up period (without any kind of active maintenance programme). Two factors were found that may have impacted on learning: additional cognitive impairments and individual differences including characteristics of the anomic disorder and the ability to employ strategies (Fillingham et al., 2006; Marshall, Pound, White-Thomson, & Pring, 1990; Warrington, 1997). These results offer a challenge to clinical practice to consider providing larger numbers of items in a therapy programme.

REFERENCES

Abel, S., Schultz, A., Radermacher, I., Willmes, K., & Huber, W. (2005). Decreasing and increasing cues in naming therapy for aphasia. *Aphasiology*, *19*(9), 831–848.
Annoni, J. M., Khateb, A., Custodi, M. C., Debeauvais, V., Michel, C. M., & Landis, T. (1998). Advantage of semantic language therapy in chronic aphasia: A study of three cases. *Aphasiology*, *12*(12), 1093–1105.

Baayen, R. H., Piepenbrock, R., & Rijn, V. (1993). *The CELEX lexical database* [CD-ROM]. Philadelphia: University of Pennsylvania, Linguistic Data Consortium.

Bachrach, H., & Mintz, J. (1974). Wechsler Memory Scale as a tool for detection of mild cerebral dysfunction. *Journal of Clinical Psychology*, *30*(1), 58–60.

Baddeley, A., & Wilson, B. A. (1994). When implicit learning fails – Amnesia and the problem of error elimination. *Neuropsychologia*, *32*(1), 53–68.

Bastiaanse, R., Bosje, M., & Franssen, M. (1996). Deficit oriented treatment of word finding problems: Another replication. *Aphasiology*, *10*(4), 363–383.

Benton, A. L., Hamsher, K., & Sivan, A. B. (1994). *Multilingual Aphasia Examination* (3rd ed.). Iowa City: AJA.

Berthier, M. L., Green, C., Higueras, C., Fernandez, I., Hinojosa, J., & Martin, M. C. (2006). A randomized, placebo-controlled study of donepezil in poststroke aphasia. *Neurology*, *67*(9), 1687–1689.

Best, W., Hickin, J., Herbert, R., Howard, D., & Osborne, F. (2000). Phonological facilitation of aphasic naming and predicting the outcome of treatment for anomia. *Brain and Language*, *74*(3), 435–438.

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

Best, W., & Nickels, L. (2000). From theory to therapy in aphasia: Where are we now and where to next? *Neuropsychological Rehabilitation*, *10*(3), 231–247.

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

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.

Bishop, D. M. V. (1989). *Test for the Reception of Grammar* (2nd ed.). Abingdon, UK: Thomas Leach Ltd.

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

Chapey, R. (2001). *Language intervention strategies in aphasia and related neurogenic communication disorders* (4th ed.). Baltimore, MD: Lippincott, Williams and Wilkins.

Cherney, L. R., Patterson, J. P., Raymer, A., Frymark, T., & Schooling, T. (2008). Evidence-based systematic review: Effects of intensity of treatment and constraint-induced language therapy for individuals with stroke-induced aphasia. *Journal of Speech, Language and Hearing Research*, *51*(5), 1282–1299.

Chin Li, E., Kitselman, K., Dusatko, D., & Spinelli, C. (1988). The efficacy of PACE in the remediation of naming deficits. *Journal of Communication Disorders*, *21*(6), 491–503.

Chin Li, E., & Williams, S. (1989). The efficacy of two types of cues in aphasic patients. *Aphasiology*, *3*(7), 619–626.

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

Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology*, *33A*, 497–505.

Conroy, P., Sage, K., & Lambon Ralph, M. A. (2009). 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.

Davis, C. H., Harrington, G., & Baynes, K. (2006). Intensive semantic intervention in fluent aphasia: A pilot study with fMRI. *Aphasiology*, *20*(1), 59–83.

DeDe, G., Parris, D., & Waters, G. (2003). Teaching self-cues: A treatment approach for verbal naming. *Aphasiology*, *17*(5), 465–480.

Deloche, G., Dordain, M., & Kremin, H. (1992). Confrontation naming rehabilitation in aphasics: A computerised written technique. *Neuropsychological Rehabilitation*, *2*(2), 117–124.

Deloche, G., Dordain, M., & Kremin, H. (1993). Rehabilitation of confrontation naming in aphasia – Relations between oral and written modalities. *Aphasiology*, *7*(2), 201–216.

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

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

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

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

Fillingham, J. K., Hodgson, C., Sage, K., & Lambon Ralph, M. A. (2003). The application of errorless learning to aphasic disorders: A review of theory and practice. *Neuropsychological Rehabilitation*, *13*(3), 337–363.

Fillingham, J. K., Sage, K., & Lambon Ralph, M. A. (2005a). Further explorations and an overview of errorless and errorful therapy for aphasic word-finding difficulties: The number of naming attempts during therapy affects outcome. *Aphasiology*, *19*(7), 597–614.

Fillingham, J. K., Sage, K., & Lambon Ralph, M. A. (2005b). Treatment of anomia using errorless versus errorful learning: Are frontal executive skills and feedback important? *International Journal of Language & Communication Disorders*, *40*(4), 505–523.

Fillingham, J. K., Sage, K., & Lambon Ralph, M. A. (2006). The treatment of anomia using errorless learning. *Neuropsychological Rehabilitation*, *16*(2), 129–154.

Francis, D. R., Clark, N., Humphreys, G.W. (2002). Circumlocution-induced naming (CIN): A treatment for effecting generalisation in anomia? *Aphasiology*, *16*(3), 243–259.

Fridriksson, J., Holland, A. L., Beeson, P., & Morrow, L. (2005). Spaced retrieval treatment of anomia. *Aphasiology*, *19*(2), 99–109.

German, D. (1990). *Test of Adolescent/Adult Word Finding*. Allen, TX: DLM Teaching Resources.

Gold, M., Van Dam, D., & Silliman, E. R. (2000). An open-label trial of bromocriptine in nonfluent aphasia: A qualitative analysis of word storage and retrieval. *Brain and Language*, *74*(2), 141–156.

Goodglass, H., Kaplan, E., & Barresi, B. (2001). *Boston Diagnostic Aphasia Examination* (3rd ed.). Philadelphia: Lippincott, Williams & Wilkins.

Graham, K. S., Patterson, K., Pratt, K. H., & Hodges, J. R. (2001). Can repeated exposure to "forgotten" vocabulary help alleviate word-finding difficulties in semantic dementia? An illustrative case study. *Neuropsychological Rehabilitation*, *11*(3–4), 429–454.

Grant, D. A., & Berg, E. A. (1993). *The Wisconsin Card Sort Task*. San Antonio, TX: Psychological Assessment Resources, Inc.

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

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

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

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

Hillis, A. E. (1991). *Effects of separate treatments for distinct impairments within the naming process*. Austin, TX: Pro:Ed.

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

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

Holland, A. (1998). A strategy for improving oral naming in an individual with a phonological access impairment. In N. Helms-Esterbrooks (Ed.), *Approaches to the treatment of aphasia*. London: Singular Publishing Group, Inc.

Howard, D., Best, W., Clarke, P., Hickin, J., & T., R. (2000). Facilitation of word rehearsal in aphasia revisited. *Brain and Language*, *74*(3), 441–444.

Howard, D., & Patterson, K. (1992). *Pyramids and Palm Trees: A test of semantic access from pictures and words*. Bury St Edmunds, UK: Thames Valley Test Company.

Howard, D., Patterson, K., Franklin, S., Orchard-Lisle, V., & Morton, J. (1985). Treatment of word retrieval deficits in aphasia. A comparison two therapy methods *Brain*, *108*, 817–829.

Jacobs, D. H., Shuren, J., Gold, M., Adair, J. C., Bowers, D., Williamson, D. J. G., et al. (1996). Physostigmine pharmacotherapy for anomia. *Neurocase*, *2*(2), 83–91.

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

Kagan, A. (1998). Supported conversation for adults with aphasia: Methods and resources for training conversation partners. *Aphasiology*, *12*(9), 816–830.

Kay, J., Lesser, R., & Coltheart, M. (1992). *Psycholinguistic Assessments of Language Processing in Aphasia*. Hove, UK: Lawrence Erlbaum Associates Ltd.

Kertesz, A. (2006). *The Western Aphasia Battery – Revised*. New York: Psychological Corporation.

Kiran, S., Thompson, C. K., & Hashimoto, N. (2001). Training grapheme to phoneme conversion in patients with oral reading and naming deficits: A model-based approach. *Aphasiology*, *15*(9), 855–876.

Kiviniemi, K., Laine, M., Tarkiainen, A., Jarvensivu, T., Martin, N., & Salmelin, R. (2000). Anomia treatment modifies naming-related cortical activation: Evidence from an MEG study. *Brain and Language*, *74*(3), 433–435.

Laganaro, M., Di Pietro, M., & Schnider, A. (2006). What does recovery from anomia tell us about the underlying impairment: The case of similar anomic patterns and different recovery. *Neuropsychologia*, *44*(4), 534–545.

LeDorze, G., Boulay, N., Gaudreau, J., & Brassard, C. (1994). The contrasting effects of a semantic versus a formal semantic technique for the facilitation of naming in a case of anomia. *Aphasiology*, *8*(2), 127–141.

Marshall, J., Pound, C., White-Thomson, M., & Pring, T. (1990). The use of picture/word matching tasks to assist word retrieval in aphasic patients. *Aphasiology*, *4*(2), 167–184.

Marshall, R. C., Freed, D. B., & Karow, C. M. (2001). Learning of subordinate category names by aphasic subjects: A comparison of deep and surface-level training methods. *Aphasiology*, *15*(6), 585–598.

Marshall, R. C., Karow, C. M., Freed, D. B., & Babcock, P. (2000). Learning of subordinate category names by aphasic subjects: Phonological versus personalized cueing. *Brain and Language*, *74*(3), 438–441.

Marshall, R. C., Karow, C. M., Freed, D. B., & Babcock, P. (2002). Effects of personalised cue form on the learning of subordinate category names by aphasic and non-brain-damaged subjects. *Aphasiology*, *16*(7), 763–771.

Marshall, R. C., Neuburger, S. I., & Phillips, D. S. (1992). Effects of facilitation and cueing on labeling of novel stimuli by aphasic subjects. *Aphasiology*, *6*(6), 567–583.

Martin, N., Fink, R., & Laine, M. (2004). Treatment of word retrieval deficits with contextual priming. *Aphasiology*, *18*(5–7), 457–471.

Martin, N., Fink, R., Laine, M., & Ayala, J. (2004). Immediate and short-term effects of contextual priming on word retrieval in aphasia. *Aphasiology*, *18*(10), 867–898.

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.

McKenna, P. (1998). *The Category Specific Names Test*. Hove, UK: Psychology Press.

McKenna, P., & Warrington, E. K. (1983). *The Graded Naming Test*. Windsor, UK: NFER-Nelson.

McNeil, M. R., Doyle, P. J., Spencer, K., Goda, A. J., Flores, D., & Small, S. L. (1998). Effects of training multiple form classes on acquisition, generalization and maintenance of word retrieval in a single subject. *Aphasiology*, *12*(7–8), 575–585.

Meyers, J. E., & Meyers, K. R. (1995). *Rey Complex Figure Test and Recognition Trial*. New York: Psychological Assessment Resources, Inc.

Miceli, G., Amitrano, A., Capasso, R., & Caramazza, A. (1996). The treatment of anomia resulting from output lexical damage: Analysis of two cases. *Brain and Language*, *52*(1), 150–174.

Milders, M. (1998). Learning people's names following severe closed-head injury. *Journal of Clinical and Experimental Neuropsychology*, *20*(2), 237–244.

Mitrushina, M. N., Boone, K. B., & D'Elia, L. F. (1999). *Handbook of normative data for neuropsychological assessment*. New York: Oxford University Press.

Nettleton, J., & Lesser, R. (1991). Therapy for naming difficulties in aphasia: Application of a cognitive neuropsychological model. *Journal of Neurolinguistics*, *6*(2), 139–157.

Nickels, L. (1992). The autocue – Self-generated phonemic cues in the treatment of a disorder of reading and naming. *Cognitive Neuropsychology*, *9*(2), 155–182.

Nickels, L. (2002). Improving word finding: Practice makes (closer to) perfect? *Aphasiology*, *16*(10–11), 1047–1060.

Nickels, L., & Best, W. (1996). Therapy for naming disorders 2: Specifics, surprises and suggestions. *Aphasiology*, *10*(2), 109–136.

Parkin, A. J., Hunkin, N. M., & Squires, E. J. (1998). Unlearning John Major: The use of errorless learning in the reacquisition of proper names following herpes simplex encephalitis. *Cognitive Neuropsychology*, *15*(4), 361–375.

Patterson, K., Purrell, C., & Morton, J. (1989). Facilitation of word retrieval in aphasia. In C. Code & D. J. Muller (Ed.), *Aphasia therapy: Studies in disorders of communication* (2nd ed.)*.* London: Code & Whurr.

Pound, C., Parr, S., Lindsay, J., & Woolf, C. (2000). *Beyond aphasia: Therapies for living with communication disability*. Bicester, UK: Winslow.

Pring, T., Hamilton, A., Harwood, A., & Macbride, L. (1993). Generalization of naming after picture word matching tasks – only items appearing in therapy benefit. *Aphasiology*, *7*(4), 383–394.

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

Raven, J. C. (1962). *Coloured Progressive Matrices*. London: H.K. Lewis.

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

Reilly, J., Martin, N., & Grossman, M. (2005). Verbal learning in semantic dementia: Is repetition priming a useful strategy? *Aphasiology*, *19*(3–5), 329–339.

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

Roach, A., Schwartz, M. F., Martin, N., Grewel, R. S., & Brecher, A. (1996). The Philadelphia Naming Test. *Clinical Aphasiology*, *24*, 121–133.

Robertson, I. H., Ward, T., Ridgeway, V., & Nimmo-Smith, I. (1994). *The Test of Everyday Attention (TEA)*. Bury St Edmunds, UK: Thames Valley Test Company.

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

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

Sbisa, S., D'Andrea, J., Semenza, C., & Tabossi, P. (2001). Rehabilitating anomia: A case study. *Brain and Language*, *79*(1), 46–47.

Schmidt, M. (1996). *Rey Auditory and Verbal Learning Test: A handbook*. Los Angeles: Western Psychological Services.

Snodgrass, J. G., & Vanderwart, M. (1980). Standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology-Human Learning and Memory*, *6*(2), 174–215.

Spencer, K. A., Doyle, P. J., McNeil, M. R., Wambaugh, J. L., Park, G., & Carroll, B. (2000). Examining the facilitative effects of rhyme in a patient with output lexicon damage. *Aphasiology*, *14*(5–6), 567–584.

Wambaugh, J., Cameron, R., Kalinyak-Fliszar, M., Nessler, C., & Wright, S. (2004). Retrieval of action names in aphasia: Effects of two cueing treatments. *Aphasiology*, *18*(11), 979–1004.

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

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

Warrington, E. K. (1997). The Graded Naming Test: A restandardisation. *Neuropsychological Rehabilitation*, *7*(2), 143–146.

White-Thomson, M. (1999). Naming therapy for an aphasic person with fluent empty speech. In S. Byng, K. Swinburn, & C. Pound (Eds.), *The aphasia therapy file* (pp. 93–100). Hove, UK: Psychology Press.

APPENDIX 1

Studies included in the meta-analysis, number of participants per study, range of medical diagnoses, and therapy type

Study	Number of participants	Medical diagnoses	Therapy type
(Best, Howard, Bruce, & Gatehouse, 1997)	1	CVA	Semantic and lexical tasks, phonological cueing
(Bruce & Howard, 1987)	5	CVA	Computer generated phonemic cueing
(Davis, Harrington, & Baynes, 2006)	1	CVA	Semantic
(Deloche, Dordain, & Kremin, 1993)	2	Meningeal haemorrhage	Written naming
(Fillingham, Sage, & Lambon Ralph, 2005b)	7	6 CVA, 1 acute haematoma and infarction secondary to previous haemorrhage	Errorless and errorful learning
(Fillingham, Sage, & Lambon Ralph, 2005a)	7	CVA	Errorless and errorful learning

APPENDIX 1
(*Continued*)

Study	Number of participants	Medical diagnoses	Therapy type
(Fillingham, Sage, & Lambon Ralph, 2006)	11	10 CVA, 1 acute haematoma and infarction secondary to previous haemorrhage	Errorless and errorful learning
(Francis, Clark, & Humphreys, 2002)	1	CVA and bilateral aneurysms of internal carotids	Circumlocute until target produced. Phonemic cueing provided in exceptional circumstances
(Greenwald, Raymer, Richardson, & Rothi, 1995)	2	CVA	Phonological cueing, semantic cueing, semantic associate picture naming, oral word reading
(Herbert, Best, Hickin, Howard, & Osborne, 2003)	6	CVA	Progressive orthographic and phonological cues with distractors
(Hickin, Best, Herbert, Howard, & Osborne, 2002)	8	CVA	Orthographic and phonological cues with distractors
(Marshall, Pound, White-Thomson, & Pring, 1990)	10	CVA	Picture and word matching tasks
(Martin, Fink, & Laine, 2004)	2	CVA	Spoken word to picture matching, repetition and naming
(Miceli, Amitrano, Capasso, & Caramazza, 1996)	2	1 ruptured AV malformation, 1 CVA	Reading, repetition, picture/word, phonemic cueing
(Nickels & Best, 1996)	3	CVA	Semantic function judgements, relatedness judgements, written word to picture matching
(Parkin, Hunkin, & Squires, 1998)	1	HSE	Errorless learning
(Pring, Hamilton, Harwood, & Macbride, 1993)	5	CVA	Word to picture matching
(Robson, Marshall, Pring, & Chiat, 1998)	1	CVA	Phonological judgements
(Rose, Douglas, & Matyas, 2002)	1	SAH	Gesture, visualisation, phonological
(Sbisa, D'Andrea, Semenza, & Tabossi, 2001)	1	CVA	Semantic and phonological
(Best & Nickels, 2000)	4	CVA	Phonological cueing

APPENDIX 2

Studies excluded from the meta-analysis and primary reason

Reason	Author
No control	Chin Li & Williams, 1989; Deloche, Dordain, & Kremin, 1992; Marshall, Karow, Freed, & Babcock, 2002; McCall, Cox, Shelton, & Weinrich, 1997; LeDorze, Boulay, Gaudreau, & Brassard, 1994; Annoni et al., 1998; Bastiaanse, Bosje, & Franssen, 1996; Fridriksson, Holland, Beeson, & Morrow, 2005; Marshall, Karow, Freed, & Babcock, 2000; Nickels, 2002; Wambaugh, Cameron, Kalinyak-Fliszar, Nessler, & Wright, 2004.
No baseline	Marshall et al., 2002; McCall et al., 1997.
No statistics	Chin Li, Kitselman, Dusatko, & Spinelli, 1988; Nettleton & Lesser, 1991; Wambaugh, 2003; Hillis, 1991; Wambaugh et al., 2004.
Individual data cannot be extracted	Abel et al., 2005; Best, Hickin, Herbert, Howard, & Osborne, 2000; Doesborgh et al., 2004; Howard, Best, Clarke, Hickin, & T, 2000; Howard, Patterson, Franklin, Orchard-Lisle, & Morton, 1985; Laganaro, Di Pietro, & Schnider, 2006; DeDe, Parris, & Waters, 2003; Howard, Patterson, Franklin, Orchard-Lisle, & Morton, 1985; Nickels, 1992; Raymer, Thompson, Jacobs, & Legrand, 1993; Spencer et al., 2000; Wambaugh et al., 2001; Drew & Thompson, 1999; Coehlo, McHugh, & Boyle, 2000; Hillis, 1989, 1998; Marshall et al., 2001; White-Thomson, 1999; McNeil et al., 1998; Eales & Pring, 1998; Holland, 1998; Kiviniemi et al., 2000; Marshall, Neuburger, & Phillips, 1992; Martin, Fink, Laine, & Ayala, 2004; Milders, 1998; Patterson, Purrell, & Morton, 1989; Renvall, Laine, Laakso, & Martin, 2003.
Drug study	Baddeley & Wilson, 1994; Berthier et al., 2006; Gold, Van Dam, & Silliman, 2000; Jacobs et al., 1996.
Therapy not picture naming	Graham, Patterson, Pratt, & Hodges, 2001; Nettleton & Lesser, 1991; Reilly, Martin, & Grossman, 2005; Hinckley & Carr, 2005; Kiran, Thompson, & Hashimoto, 2001.
Targets other than nouns	Wambaugh et al., 2004.