INT. J. LANG. COMM. DIS., SEPTEMBER–OCTOBER 2010,
VOL. 45, NO. 5, 586–599

informa
healthcare

Research Report

Narrative-based intervention for word-finding difficulties: a case study

Ian Marks† and Stephanie F. Stokes‡

†National Health Service, Speech Pathology, Albion Road Resource Centre, North Shields, Newcastle upon Tyne, UK
‡School of Psychology, Curtin University, Perth, WA, Australia

(*Received April 2009; accepted August 2009*)

Abstract

Background: Children with word-finding difficulties manifest a high frequency of word-finding characteristics in narrative, yet word-finding interventions have concentrated on single-word treatments and outcome measures.
Aims: This study measured the effectiveness of a narrative-based intervention in improving single-word picture-naming and word-finding characteristics in narrative in a case study.
Methods & Procedures: A case study, quasi-experimental design was employed. The participant was tested on picture naming and spoken word to picture matching on control and treatment words at pre-, mid-, and post-therapy and an 8-month maintenance point. Narrative samples at pre- and post-therapy were analysed for word-finding characteristics and language production. A narrative-based language intervention for word-finding difficulties (NBLI-WF) was carried out for eight sessions, over 3 weeks. The data were subjected to a repeated-measures trend analysis for dichotomous data.
Outcomes & Results: Significant improvement occurred for naming accuracy of treatment, but not for control words. The pattern of word-finding characteristics in narrative changed, but the frequency did not reduce.
Conclusions & Implications: NBLI-WF was effective in improving naming accuracy in this single case, but there were limitations to the research. Further research is required to assess the changes that may occur in language production and word-finding characteristics in narrative. Community clinicians are encouraged to refine clinical practice to ensure clinical research meets quality indicators.

Keywords: word-finding difficulties, narrative-based intervention.

What this paper adds
This paper demonstrates the value of using narrative-based intervention for word-finding difficulties in a school aged child. Clinicians are encouraged to replicate this study with similar children, being sure to meet the quality indicators for single subject research. A repeated measures analysis for dichotomous data, a method that is not yet widely known, is used to analyse the change in naming scores over time.

Introduction

This study investigates the nature of word-finding difficulties (WFDs) in the narratives of a boy with WFD and explores the usefulness of a narrative-based word-finding therapy. The classical description of WFDs is that of reduced word-naming ability relative to intact word comprehension. WFDs occur in around 23% of children with developmental language disorders (Dockrell *et al.* 1998). German and colleagues (German 1987, 1994, 2000, 2002, German and Simon 1991, German and Newman 2004) have described the characteristics of children's WFDs in detail over a number of years, and German and Newman's (2004) succinct summary neatly encapsulates the problem as one of 'delayed or inaccurate responses with a high incidence of repetitions, reformulations, word substitutions, insertions, time fillers, and empty words' (p. 624) in confrontation naming tasks and discourse. WFDs may have a negative effect on oral communication, academic performance (German 2002), and literacy development (Wolf and Obregon 1992).

Address correspondence to: Stephanie F. Stokes, School of Psychology, Curtin University, GPO Box U1987, Perth, WA 6845, Australia; e-mail: s.stokes@curtin.edu.au

International Journal of Language & Communication Disorders
ISSN 1368-2822 print/ISSN 1460-6984 online © 2010 Royal College of Speech & Language Therapists
http://www.informahealthcare.com
DOI: 10.3109/13682820903277951

A hallmark of the condition is difficulty with oral lexical retrieval in comparison with comprehension. This may be characterized as age-appropriate performance on assessments of comprehension, but poor performance on naming. However, for some children both comprehension and naming may be below the level expected for their age (Messer and Dockrell 2006). Before outlining the case, we briefly review theoretical accounts of WFDs and how these have translated into intervention approaches.

Accounts of word-finding errors

German (1994, 2000) proposed that children with WFDs fall into one of three subtypes: retrieval difficulties, comprehension difficulties, or comprehension and retrieval difficulties. Children in subtype 1 fulfil the classic definition of WFDs, having age-appropriate comprehension but difficulty in reliably retrieving words. Children in subtype 2 have little difficulty accessing and producing words that are stable and fully represented in their lexicon, but have difficulty in storing and understanding unfamiliar words, which in turn leads to difficulty with production. Subtype 3 represents underlying problems in learning word meanings combined with retrieval difficulties for words, even when they are stable in the child's lexicon. German's subtypes are derived from an adapted model of spoken word production (German 2000) based on the lexical architectural model proposed by Levelt (1989, 1991). German (2000) suggested that breakdown in word retrieval (resulting in a word-finding error) may occur at one of three levels in the retrieval process. The first is when the child's conceptual structure for a word does not access the target lemma. Such breakdowns are likely to manifest as semantic substitutions. The second level occurs in the access of phonological content from the lemma, possibly resulting in a 'no response' or 'don't know' response. The third level also concerns access to the phonological content, but German (2000) suggested a level of breakdown where only partial access occurs. Such a breakdown may lead to errors involving phonemic approximations of the target word.

Ultimately, no single model of speech production has been developed that takes into account the developmental processes of language acquisition (Messer and Dockrell 2006). The current models available for considering word retrieval are not only based on an adult 'end state' (for example, Goldrick and Rapp 2002, and Levelt 1989, 1991; for alternate models, see German 2002), they also predominantly focus on the production of single words. Situating accounts of WFDs within models of single-word production has resulted in intervention programmes that target the possible levels of breakdown described above, with a focus on single-word production.

Single-word intervention for word-finding difficulties

Intervention for WFDs has largely consisted of either semantic (Casby 1992, McGregor and Leonard 1989) or phonological approaches (German 2002, McGregor 1994), or a comparison of these approaches (Wing 1990). Both these approaches have led to reductions in word-finding errors and improvements in naming performance as measured by tests of single-word retrieval. Best (2005) drew on research into WFDs in adults with aphasia and demonstrated the effectiveness of a computerized aid with 6–10-year-old children that generates phonemic cues, adding to the paradigms in clinical use. However, research has not yet clearly demonstrated an approach that is more effective than any other. Wing (1990) compared phonological and semantic treatment approaches and concluded that phonological intervention was more effective for 6-year-olds with WFDs. Recent research has suggested the need to go beyond single-word naming performance of children with WFDs to explore their word-finding abilities in discourse/narrative (Tingley *et al.* 2003).

Word-finding characteristics in narrative

German has carried out the most comprehensive investigation of word-finding characteristics in narrative (German 1987, German and Simon 1991). Her research considered the types of word-finding characteristics demonstrated by children with WFDs in spontaneous narrative activities. The results demonstrated that children with WFDs evidenced the same *types* (repetitions, reformulations, substitutions, delays, empty words, time-fillers and insertions) and distribution of word-finding characteristics as chronological age-matched peers, but that they tended to use significantly *more* in their narratives (German and Simon 1991). As narrative ability has been associated with academic and literacy achievement (Soto *et al.* 2006) and WFDs have been found to impact negatively on oral narrative performance (German 1987, German and Simon 1991), researchers have begun to explore the usefulness of narrative-based intervention for children with WFDs.

Narrative-based intervention

Several research teams have demonstrated that novel or new words may be learnt from narrative contexts. For example, Brackenbury and Fey (2003) found that 4-year-olds were able to learn verbs following a few exposures in an oral story. This type of indirect learning

of words occurs from an early age (Brackenbury and Fey 2003). Oetting *et al.* (1995) reported rapid word learning from a videotaped story for typically developing 6–8-year-old-children. A group of children with specific language impairment (also 6–8 years of age) demonstrated some ability to learn words in this context, but learned significantly fewer words than their typically developing peers. Nash and Snowling (2006) compared word learning in two different tasks, one of which was learning by word definitions, the other was learning from a written narrative. Their 24 (7–8-year-old) children with WFDs learned in both conditions, but better maintenance of learning was seen in the narrative condition. Thus, a story-based intervention approach may be an effective way to stimulate word learning in children with WFDs. While a narrative-based approach may result in gains in vocabulary-naming scores, little is known about whether or not concomitant changes in word-finding characteristics can be achieved. Intervention studies have not sought to investigate changes in word-finding characteristics. This study investigates the nature of WFD in the narratives of a boy with WFD and explores the usefulness of a new narrative-based word-finding therapy. Before outlining the research, we briefly consider issues of treatment efficacy.

Measuring outcomes

The aim was to measure the outcomes at single-word and narrative levels, investigating changes in naming accuracy, word-finding characteristics and language productivity (research questions one to three below). Prior intervention research indicated that while there should be gains on treated words, generalization to an untreated control set of words could not be expected (for example, German 2002, and Casby 1992). In this intervention we explored learning of treated and untreated (control) word sets. In addition to exploring different linguistic levels (that is, words versus narrative), in this study we demonstrate the application of a new statistical analysis for single-subject intervention using a repeated-measures trend analysis for dichotomous data (Howard 2007). The latter is described in the Methods section.

Hypotheses

The literature reviewed above suggests that an NBLI-WF might be effective for improving picture labelling for treated words (for example, Nash and Snowling 2006), and that generalization to untreated (control) words should not be expected. It is not known whether or not therapy would reduce the incidence, type, or distribution of word-finding characteristics in connected speech, but we were optimistic that it would. Our hypotheses were as follows:

- NBLI-WF would increase AB's accuracy of naming treated but not untreated words.
- NBLI-WF would reduce the incidence of WFCs in connected speech.
- NBLI-WF would increase the productivity of AB's narrative.
- NBLI-WF may result in a re-distribution of WFCs in connected speech.

Research questions

- Does narrative-based intervention improve the accuracy of naming of treated vocabulary?
- Does narrative-based intervention reduce word-finding characteristics (repetitions, reformulations, substitutions, delays, empty words, time-fillers and insertions) in narrative?
- Does narrative-based intervention increase language productivity in narrative as measured by number of words/number of T-units in a story sample?

Method

Participant

AB (not his real initials) was a boy aged 8;01, without a previous history of speech and language difficulties, who was referred to speech pathology services by his family doctor after his parents expressed concerns about his apparent dysfluency. A National Health Services (NHS) specialist in fluency disorders assessed AB and concluded that his difficulties lay in expressive and receptive language impairment in the absence of any identified cognitive impairment. She recommended that AB be assessed for possible WFD and enrolled in a WFD intervention programme if found to meet the criteria for a child with WFD. Subsequently, AB was assessed by the first author (the results are reported below) and enrolled into the intervention programme described in this report. As this was an evaluation of practice, NHS ethics approval was not required, but ethical approval for the study was granted by the Research Ethics Committee of the School of Education, Communication and Language Sciences of Newcastle University. In addition, AB's parents signed an informed consent slip for reporting of these results.

AB was an immigrant to England at the age of 4 years, having been born in Malaysia. Both parents use English and Malay at home. AB was reported to be a late talker who does not use or understand Malay, speaking only in English. His developmental history was reported to be normal.

Cognition, receptive and expressive language

AB scored at just above the 50th percentile on the Raven's Coloured Progressive Matrices (Raven *et al.* 1998). He scored at the 13th percentile on the British

Picture Vocabulary Scale—2nd Edition (Dunn *et al.* 1997). At sentence level, AB scored at the 5th percentile for both the Test for Reception of Grammar—Version 2 (Bishop 2003) and the Concepts and Directions subtest of the Clinical Evaluation of Language Functioning (CELF)-4 (Semel *et al.* 2003). These three results indicated poor receptive vocabulary and receptive grammar. AB scored at the 2nd percentile on both the recalling sentences subtest and the formulated sentences subtest of the CELF-4 (Semel *et al.* 2003) suggesting significant problems with expressive grammar.

Morphology

AB scored at the 37th percentile on the word structure subtest of the CELF-4 (Semel *et al.* 2003) and immature morphological forms were noted in his productions, for example, 'falled' and 'eated'.

Word finding

AB was tested on word finding using the Test of Word Finding—2nd Edition (TWF-2) (German 2000). This assessment seeks to identify discrepancies between comprehension and expression of vocabulary at single-word and sentence level, which conforms to the classical definition of WFDs (Messer and Dockrell 2006). AB scored at the 13th percentile indicating 'below average' word-finding abilities, which reflects the discrepancy between AB's raw scores of 43/70 for picture naming and sentence completion compared with 64/70 for comprehension of the same items.

Narrative

AB scored at a level appropriate for a 4–5-year-old child on Renfrew's *The Bus Story* (1969) (percentile scores are not available). Narrative performance is described in more detail as part of the intervention programme.

Fluency

During formal assessment of narrative, it was noted that AB's speech was very slow and stilted but did not contain dysfluencies.

Summary

AB exhibited an overall delay in expressive and receptive language. This did not appear to be commensurate with his non-verbal development, as no concerns had been expressed regarding his cognitive, motor or social development. AB demonstrated difficulty with expressive and receptive language at single-word, sentence and discourse levels. Note, though, that on TWF-2, AB's performance reflected the classical definition of a WFD, that of reduced word-naming ability relative to intact word comprehension, possibly indicative of German's (1994, 2000) subtype 1 (retrieval difficulties).

Design and setting

A single-subject, ABA design was used. Intervention sessions were carried out at AB's school. All sessions took place in a small resource room, which seated the child and therapist at a small desk with two chairs. Sessions were audio-recorded and lasted for 50–60 minutes. Two core areas of performance were examined: connected speech performance (productivity and word-finding characteristics within a narrative) and single-word performance (comprehension and production of picture names). The narrative task is described first, followed by the single-word task.

The narrative task

A narrative sample was elicited using a 24-picture storybook without words, *Frog, Where Are You?* (Mayer 1969) following the procedure set out by Berman and Slobin (1994). AB was presented with the story book and given the following instructions:

> Here is a book. This book tells a story about a boy [point to picture on cover], a dog [point] and a frog [point]. First, I want you to look at all the pictures. Pay attention to each picture that you see and afterwards you will tell the story. (Berman and Slobin 1994: 22)

Prompts were kept to a minimum and, where used, followed Berman and Slobin's hierarchy for prompts:

(1) Silence/nod.
(2) 'uh huh', 'OK', 'Yes'.
(3) 'Anything else?'
(4) 'And?'
(5) 'Go on.'

AB's narrative sample was recorded using an Olympus WS-300M Digital Voice Recorder and transcribed orthographically. Transcriptions of the narrative samples were analysed using the method outlined by German and Simon (1991). This analysis seeks to represent the profile of word-finding characteristics displayed in discourse. Two indices are considered: language productivity and incidence of word-finding characteristics. The samples were segmented into T-units (the shortest unit into which a linguistic utterance can be divided without leaving a fragment; German and Simon 1991).

Language productivity

Total words were counted including repetitions, empty words, reformulations and substitutions. Words/utter-

ances that were excluded from the total word count were time fillers ('um', 'er'), initial sound repetitions ('p, p, postman'), initial conjunctions, starting statements ('I am going to tell a story'), and finishing sentences ('That's all'). Total T-units was a frequency count of the total number of T-units in a narrative sample.

Word-finding characteristics index

The global word-finding characteristics index represented overall occurrence of word-finding characteristics within a narrative sample. It was calculated as the percentage of T-units containing one or more word-finding characteristics. The frequency of occurrence of each type of WFC was calculated as the number of times each WFC occurred, divided by the total number of WFC, pre- and post-therapy. The categories of word-finding characteristics under investigation are outlined and defined below:

(i) *Repetitions* were words within a T-unit that were unnecessarily repeated ('He, he picked up the dog').
(ii) *Word reformulations* were defined as words within a T-unit that have been changed or replaced by a revision ('He found a frog at, in the forest').
(iii) *Empty words* were words that did not add content to the T-unit ('They saw a, you know, big bird').
(iv) *Insertions* were words or utterances that comment on the language process ('They went in to the ... not sure what that's called').
(v) *Substitutions* were words within a T-unit that are target word substitutions ('The bottle [*jar*] smashed').
(vi) *Delays* were pauses of six seconds or longer without any verbalization within a T-unit.
(vii) *Time fillers* were sounds or syllables ('um', 'er') that filled time while the child was attempting to retrieve a word. As time fillers are a common occurrence in normal spoken productions, they were only registered when three or more occurred within a T-unit.

Language productivity and the word-finding characteristics index were both measured pre- and post-therapy.

Single-word task

Two sets (control and treatment items) of 30 nouns were chosen from words that AB recognized (scored as correct on the receptive vocabulary measures) but not produced (scored as incorrect for expressive vocabulary). The two sets were matched (see appendix A) for mean reaction time, number of syllables, frequency and age of acquisition from the psycholinguistic database of the International Picture Naming Project (Szekely *et al.* 2003). Control and treatment items were matched on psycholinguistic features as research has shown that naming and word-finding is affected by factors such as age of acquisition and frequency (German and Newman 2004). It is less common for studies to match words for reaction time, but this was included, as many researchers have considered slow reaction time in naming as a feature of WFDs (for example, German 2000). Pictures of all 60 words were produced as black-and-white images. Each item was produced as a 7×7 centimetre, laminated picture card.

Production and comprehension of labels

The control and treatment items were presented one at a time in a random order. AB was instructed to look at each picture and name it. Responses were recorded as either correct or incorrect. Incorrect responses were orthographically transcribed. Second, a spoken word to the picture-matching test of the treatment and control items was administered (receptive vocabulary test). The pictures were presented at random in four sets of 15 pictures, displayed in a 3×5 matrix (Casby 1992). AB was required to point to the correct picture when named by the investigator in a random order.

Responses to the picture naming and spoken word to picture-matching assessments were coded as one for correct and zero for incorrect. Correct responses for picture naming were recorded when AB's response matched the target word with his first reaction to the picture stimuli, within 4 seconds (German 2000). Correct responses for spoken word to picture matching were also recorded when the target item was selected within 4 seconds. Performance on naming and receptive knowledge for the two sets of 30 words was measured pre-therapy, after four sessions (mid-therapy) post-therapy (after eight sessions), and at maintenance (8 months post-therapy). Note that although all initial assessments were conducted by the fluency specialist, the pre- and post-therapy assessments (picture-naming, picture-recognition and narrative tasks) were conducted by the first author. This shortcoming is raised in the Discussion.

Intervention

Content

Treatment used stories and sentences (see appendix B for examples). The treatment set of 30 words was split into four groups (2×7 and 2×8 words). A simple story was composed for each group of words. Each story consisted of 150–250 words and contained each of the words at least once. The aim was for each story to be engaging, illustrative of the word meanings through context and of a level of complexity accessible by AB. Two sentences were composed for each treatment word, one *definitional* sentence and one *contextual* sentence. The *definitional* sentences contained one or two key

features of the treatment words. These were either visual or functional features, depending on what was considered to be the most salient for each item. The *contextual* sentences contained a person or character relating to or using the target item in a way that would illustrate the word meaning. In total there were four stories and sixty sentences (30 definitional, 30 contextual).

Structure

There were eight therapy sessions, each lasting between 50 and 60 minutes, over a 3-week period, an intensity that would usually be reserved for high-priority clients. This model of service delivery was required for the present child as the study was conducted as part of a student clinical experience for the first author. Each session followed the same procedure and contained three activities. The four groups of words and their associated stories and sentences were used twice each, allowing for approximately equal exposure to each treatment word. Various games were used to maintain motivation in the activities.

The procedure for the intervention was based upon the programme of Narrative Based Language Intervention (NBLI) outlined by Swanson *et al.* (2005). The original NBLI programme targeted narrative structure (for example, story grammar) and complex syntax through narrative activities (Finestack *et al.* 2006, Swanson *et al.* 2005). The focus for AB was neither of these. Although the NBLI model was followed, with inclusion of the components of story retell, sentence imitation, story generation and post-story drawing of the story structure, the therapy focus was word learning. In order to help AB develop a logical story, he was encouraged to follow the narrative guidelines used by his school (who, what, where, when, and what happened). This procedure will be labelled as NBLI-WF and is outlined in detail below.

Story retell-imitation task. Pictures representing each story were placed in front of AB. Each story was read by the therapist and AB was asked to point to the appropriate picture when he heard a target word in the story. The story was then retold one story component at a time. The term 'story component' was taken from the NBLI programme (Swanson *et al.* 2005). However, the researchers do not define the term. Swanson *et al.* (2005) refer to Hoffman *et al.* (1990), but this research does not include or define the term 'story component'. For our purposes, a story component was defined as one or two sentences containing a story 'event'. AB was asked to retell the same story component by component, again referring to pictures of the treatment words. The focus of the retell was not for AB to reproduce the story word for word but to show understanding of the main story ideas and correctly produce the treatment words.

Sentence imitation task. AB was asked to imitate two sentences for each treatment word. One sentence was a 'definitional' sentence and one was a 'contextual' sentence. For example, if the treatment word was 'binoculars', the definitional sentence was 'Binoculars are for looking at far away things' and the contextual sentence was 'Kamal used his binoculars to see a little insect'. AB was encouraged to listen carefully to each sentence and try to repeat it exactly. If he made errors, the sentence was broken down into smaller parts. If errors continued, AB was allowed to give his best attempt before moving on to the next sentence.

Story generation task. AB was asked to generate his own story using three or four of the treatment words for that session. Each of the words was used in one story over the total eight sessions. AB was asked to generate a 'who', 'where', 'when', and 'what happened/problem—resolution' element before starting to tell his story (this was in line with the school's approach to narrative teaching). AB then told his story and the therapist gave feedback on the use of the target vocabulary and offered prompts for naming where necessary. AB was then given an opportunity to draw a picture for his story to help conclude this activity.

Results

Picture naming and spoken word to picture matching

AB was tested on expression (picture naming) and reception (spoken word to picture matching—SWPM) for the treatment and control sets at pre-therapy,

Table 1. Expressive and receptive vocabulary scores for pre-, mid- and post-therapy and maintenance measurements

	Pre-therapy		Mid-therapy		Post-therapy		Maintenance	
Set	Expressive	Receptive	Expressive	Receptive	Expressive	Receptive	Expressive	Receptive
[a]Treatment	8	23	19	30	27	30	17	30
[a]Control	8	23	7	26	10	27	12	26

Note: [a]The maximum possible score is 30 for each set.

mid-therapy (after four sessions), post-therapy (after eight sessions), and maintenance (after 8 months) time points. The results are given in Table 1. Visual inspection of the data suggests greater improvement in picture naming for the treatment set compared with the control set. This was the case at mid-therapy and post-therapy. At the maintenance point there was less of a difference in treated and control sets with treated items returning to about the same level of accuracy as at the mid-therapy point, and the control items continuing to improve in accuracy.

Treatment set

There was a significant improvement in picture naming from pre- to post-therapy (McNemar's test exact $p < 0.001$, one-tailed), which supports the hypothesis that words treated with NBLI-WF would show naming improvement. This improvement is displayed over the therapy block with significant improvement in naming of treated words from pre- to mid-therapy (McNemar's test exact $p < 0.001$, one-tailed) and from mid- to post-therapy (McNemar's test exact $p < 0.001$, one-tailed). This indicates that AB made significant progress in expressive knowledge of treated items after just four sessions, with each item used in just one of those four treatment sessions. Although expressive scores dropped at the maintenance phase, scores were still significantly higher than pre-therapy (McNemar's test exact $p = 0.003$, one-tailed).

There was a significant improvement for reception of treated items from pre- to post-therapy (McNemar's test exact $p = 0.016$, two-tailed) and from pre-therapy to maintenance (McNemar's test exact $p = 0.016$, two-tailed). AB was able to respond correctly in SWPM for all treated items after four treatment sessions (mid-therapy).

Control set

There was no significant difference in naming accuracy for the control set for any of the time comparisons (McNemar's test exact $p = 1.00$, 0.63, and 0.13, two-tailed for pre-therapy versus mid-therapy, pre-therapy versus post-therapy, and pre-therapy versus maintenance), despite the gradual increase in naming accuracy from eight to twelve.

Comparison of treatment and control sets

There was a significant difference between the scores of the treated and control sets for expression at both mid-therapy (Fisher exact $p = 0.002$, one-tailed) and post-therapy (Fisher exact $p < 0.001$, one-tailed) but there was no significant difference at follow-up (Fisher exact $p = 0.301$, one-tailed), due to the continued increase in accuracy of the control set and the drop in the treated set to approximately mid-therapy scores. There were small differences between the sets for reception, but these did not reach significance at mid-therapy (Fisher exact $p = 0.112$, two-tailed) or at post-therapy (Fisher exact $p = 0.237$, two-tailed). The treated receptive items reached 100% accuracy at the mid-therapy point and this was maintained even until 8 months post-therapy. However, expression, not comprehension, was the focus of therapy.

Measurement of improvement over time

Howard's (2007) method for evaluating treatment effects was used. In this modified Marascuilo and McSweeney (1977) statistical analysis, serial dependence of dichotomous scores across time is taken into account. This procedure calculates an *S*-statistic for each word (observation). To achieve this, it is first necessary to calculate the number of later attempts (observations) that have a higher value than the observation, minus the number of previous attempts that have a higher value than the observation, and sum this across trials (pre-, mid-, and post-therapy, and maintenance). This yields an *S*-statistic for each word. For example, if the child's scores for the word 'telescope' are 0, 1, 1, 1, then the *S*-score for 'telescope' is 3; if the scores for the word 'pirate' are 0, 0, 1, 0, then the *S*-score for 'pirate' is 1; and if the scores for the word 'astronaut' are 0, 1, 0, 0, then the *S*-score for 'astronaut' is −1. In this way, 30 *S*-values (for 30 words) are generated for each time point.

These values were subjected to two (expressive and receptive) paired sample *t*-tests to determine whether or not the rate of learning was different between the treatment and control sets. For expressive vocabulary there was a significant difference between control and treated words ($F(1, 29) = 28.68$, $p < 0.001$, partial $\eta^2 = 0.50$), with mean slope values of 0.13 and 1.23 for control and treated words, respectively. For receptive vocabulary there was no significant difference between control and treated words ($F(1, 29) = 0.81$, $p = 0.38$, partial $\eta^2 = 0.03$), with mean slope values of 0.27 and 0.47 for control and treatment lists, respectively. When the maintenance phase was included, there was still a significant difference between control and treated words for expressive vocabulary ($F(1, 29) = 5.32$, $p = 0.028$, partial $\eta^2 = 0.16$), with mean slope values of 0.50 and 1.17 for control and treated words, respectively, although the difference was not as high as it had been immediately after therapy.

Word finding in narrative

A comparison was made between the word-finding characteristics in a pre- and a post-therapy narrative sample. Narrative was not reassessed in the maintenance

phase due to clinical time constraints. The identified word-finding characteristics have been compared by incidence of types of word-finding characteristics.

Language productivity

The pre-therapy sample contained 291 words and the post-therapy sample contained 290 words, so productivity was almost identical for the two samples. The pre-therapy sample had 46 T-units, at an average of 6.3 words per T-unit. Post-therapy 40 T-units were observed, at an average of 7.3 words per T-unit. This indicates that AB produced meaningful units that were on average one word longer in his post-therapy sample.

Incidence of word-finding characteristics

There was little difference between the two samples on the global measures of word finding. The pre-therapy sample contained 33 occasions of word-finding characteristics and the post-therapy sample contained 25. The Global Word Finding Characteristics Index (the percentage of T-units with word-finding characteristics) was 48% pre-therapy (22/46) and 45% post-therapy (18/40). Three of the specific word-finding characteristic categories were not observed in either sample (*empty words*, *insertions*, *time fillers*) and one category (*delays*) had just one incidence in the pre-therapy sample (Table 2). The incidence of *substitutions* was similar between the pre- and post-therapy samples (15% and 24% of all WFCs, respectively). Differences between the two samples can be observed for the word-finding characteristics *repetitions* and *word reformulations*. Pre-therapy, 70% of WFCs were repetitions and 12% were word reformulations. Post-therapy, only 32% were repetitions and 44% were word reformulations.

Discussion

The outcomes of the intervention, using a single-case design, were measured by picture naming, language productivity in narrative, and analysis of word-finding characteristics in narrative. The narrative-based intervention program was effective in improving the accuracy of picture naming for treated vocabulary items. There was a significant difference between control and treated expressive words at all phases, except the maintenance phase, where control items scored higher than at any point during the intervention, and the treated items returned to the mid-therapy accuracy rate. However, measurement of the rate of learning across time (slopes) showed a significantly steeper learning curve for treated versus control items, even with the drop in treated items at the maintenance phase. It should be noted that the intention was to determine whether or not this type of narrative therapy could have a positive effect on a child's WFD. It could be argued that although statistical significance of the results have been demonstrated, the clinical significance of the results is questionable, as the naming accuracy of 30 words might not be indicative of an important change for the child. However, the intention was to demonstrate the potential of the method and further research is required. There does not appear to be an effect of the intervention programme on language productivity, but changes are evident in the pattern of word-finding characteristics.

Table 2. Incidence of word-finding characteristics (WFCs) in pre- and post-therapy narrative samples

	Pre-therapy	Post-therapy
Number of T-units	46	40
Number of WFCs	33	25
Number of T-units with a WFCs	22	18
Percentage of all WFCs		
Repetitions	70	32
Word reformulations	12	44
Substitutions	15	24
Delay	3	0

Picture naming

Previous investigations have demonstrated improvement on confrontation naming tasks as a result of phonological (German 2002, McGregor 1994) and semantic/elaboration (Casby 1992, McGregor and Leonard 1989) intervention approaches. The results of this investigation add to the range of approaches that may be considered effective in improving picture-naming accuracy. Improvement was restricted to treated vocabulary items, which indicates that the improvement was as a result of the intervention. Previous intervention studies for WFDs have similarly shown improvement restricted to treated items (for example, German 2002, and Casby 1992). Some researchers have argued that effective intervention for WFDs should provide children with strategies that they can employ and generalize to untrained words (Wing 1990). However, intervention studies that have targeted generalization to untrained words through the teaching of word-finding strategies have not always achieved their aims (for example, Best 2005). Therefore, an intervention may be considered effective if improvement on treated vocabulary may be achieved in a clinically feasible period of time (German 2002). It is necessary to rule out improvement due to a rehearsal effect. Whilst there was more exposure to treated items than control items, control items were seen on four occasions (pre-, mid- and post-therapy, and at maintenance) without significant changes in naming accuracy. Furthermore,

the exposure rate of each item (word) was limited to two out of eight sessions, as only seven to eight treatment items were used per session, which is considered to be a relatively limited level of exposure. Casby (1992) found that there was a minimal effect of rehearsal across their multiple-baseline tests, and German (2002) used rehearsal as part of a phonological intervention approach but stated that rehearsal alone did not lead to improvement for control items.

It is proposed that the NBLI-WF activities may have acted to strengthen and establish stronger links between the treated words and other entries in AB's semantic memory, leading to improved naming accuracy. However, other processes may also have been involved in improving recall of treatment words. The importance of providing a clear context for word meaning was posited by Nash and Snowling (2006). They proposed that learning of vocabulary through context may be enhanced by providing a format through which information about word meaning may be accessed and by ensuring that materials and activities are engaging. NBLI-WF gave structured and repeated examples of the treatment items in contexts that illustrated a simplified definition and provided a clear context. AB was able to engage with the vocabulary through both active exposure to examples of the items in sentence and narrative contexts, as well as being given the opportunity to demonstrate his learning through creative narrative activities.

The fact that significant improvement in naming accuracy has been demonstrated through NBLI-WF appears to support a semantic/storage deficit account of WFDs (for example, Dockrell *et al.* 2003). If the process of NBLI-WF works by enhancing knowledge of word meaning and use, as well as establishing new and improved semantic links for vocabulary items, this suggests that the source of WFDs lies in impaired lexical storage and knowledge (Casby 1992). However, German's (2000) model of word retrieval proposes a potential breakdown between the conceptual structure of a target word and access to the target word lemma. It could be argued that NBLI-WF causes improvement at the conceptual level by providing accessible information on the functions and perceptual attributes of treated items. It may also lead to improvement of knowledge at the lemma level by providing examples of semantic and syntactic word features and strengthening the links between the conceptual and lemma levels by providing opportunities for practice in using treatment words through sentence repetition and story repetition activities. The outcomes of this investigation do not provide adequate data to argue strongly in favour of either theoretical account of WFDs.

Language productivity

German and Simon (1991) compared the total number of words produced in narratives samples for children with WFDs and typically developing children aged 7–12 years. They reported no significant differences between the groups, although the group means for number of words were 282 and 363, respectively. As AB's story lengths were of 291 and 292 words. It appears that productivity may not have been an issue for AB and thus it might not have been an appropriate target for intervention. It may be more appropriate to target language productivity for those children with WFDs who show reduced productivity in narrative (that is, those children who scored at least below 282 words in similar narrative contexts).

Word-finding characteristics in narrative

AB's incidence of word-finding characteristics in the narrative samples was supportive of the diagnosis of WFDs. The Global Word Finding Characteristic Index (GWFCI) (the percentage of total T-units containing one or more word-finding characteristics) for the pre-therapy sample was 48% and 45% post-therapy. German and Simon (1991) reported that the narrative samples of 7–12-year-old children with WFDs had a mean GWFCI of 34% (standard deviation (SD) = 15.08), and so AB's incidence of WFCs could be described as 'high'. There was no change in the global incidence of word-finding characteristics for AB's narratives following intervention. This suggests that NBLI-WF was not effective in reducing global word-finding characteristics in narrative for this child.

The most salient characteristic of AB's narrative samples is the shift from pre- to post-therapy in the incidence and distribution of certain, specific word-finding characteristics. His narrative samples contained no *empty words*, *insertions* or *time fillers*, unlike the children with WFDs in German and Simon's (1991) cohort. AB's use of *substitutions* did not change significantly from pre- to post-therapy and was very close to the level of incidence found by German and Simon. The static nature of *substitutions* further reflects the fact that NBLI-WF led to improvement for treatment words only. The lexical items in the story that AB used *substitutions* for were the same at pre- and post-therapy (for example, 'bottle' for *jar*). If improvement in picture naming had generalized to control items, it might have been expected also to generalize to words in the story that had previously caused WFDs.

The pre-therapy sample contains a high number of *repetitions*, which is more than 1 SD above the mean in German and Simon's (1991) data. German and Simon suggest that a high incidence of repetitions may be a dual

indicator of both WFDs and difficulties with fluency. This is particularly significant for AB, as his referral to speech and language therapy was for dysfluency. However, there is a body of research that has consistently identified repetitions as indicators of WFDs (Wiig and Semel 1984, German 1987, German and Simon 1991). Furthermore, the post-therapy sample contains approximately one-third as many *repetitions*, which represents an incidence that is matched closely to German and Simon's (1991) data. There is no clear link between the intervention programme and the reduction of *repetitions*. It is possible that there was a rehearsal effect in narrative. Practising narratives may have contributed to reduced repetitions by introducing familiarity into the discourse structure. Another contributing factor may have been anxiety at the time of the pre-therapy narrative sample. However, AB did not appear anxious at the time, and was already familiar with the therapist conducting the sample. An alternative explanation may be related to the increase in the incidence of *word reformulations* in the post-therapy sample. It is possible that AB had begun to adapt to his WFDs with an alternative behaviour as represented by an increase in word reformulations.

The post-therapy narrative sample shows an increase in *word reformulations*. The incidence of *word reformulations* in the pre-therapy sample falls in the lower range of that reported by German and Simon (1991) but increases to approximately 1.5 SDs above the mean level expected for children with WFDs. German (1987) purports that *word reformulations* are manifestations of a searching or scanning process where the child is attempting to find a word, express a thought or correct a grammatical error. The behaviour also indicates the child's dissatisfaction with their lexical choices, and may be unique to narrative/spontaneous language tasks, particularly where children are producing longer narratives (German 1987). It would appear that, for AB, the increase in *word reformulations* represents a heightened awareness of WFDs and an increased ability to anticipate and recognize errors in language production. Both typically developing children and children with WFDs display a similar distribution of *repetitions* and *word reformulations* in narrative (German and Simon 1991). AB shifts from a pre-therapy preponderance of *repetitions* to a higher percentage of *word reformulations* post-therapy. The distribution in the post-therapy samples matches more closely the mean distributions of word-finding characteristics in German and Simon's data, which may indicate that NBLI-WF has effected a change in AB's narrative profile so that it more closely approximates the average profile found in children with WFDs.

Limitations of the investigation

The limitations in case study research can be considered within a framework of ideal conditions for single-subject research (Horner *et al.* 2005). Horner *et al.* provide a checklist of quality indicators for single-subject research, which includes 21 items across seven categories: (1) participant and settings (characteristics, selection, location), (2) dependent variable (operationally defined, quantifiable, valid, measured repeatedly, inter-rater reliability), (3) independent variable (replicable, manipulated, fidelity of implementation), (4) baseline (repeated-measures, replicable), (5) experimental control/internal validity (repeated-measurement, eliminate common threats to validity, experimental controls), (6) external validity (replicated across participants), and (7) social validity (importance of outcome, social significance of degree of change, therapy is practical and cost effective, community-based therapy). As this model could not be applied to the standard clinical practice that was implemented for this case study, the framework has not been described in detail here. Rather, it is useful to consider how standard clinical practice could be improved to ensure that clinical research meets quality indicators in order to contribute to the evidence base of treatment efficacy for speech–language therapists. Some aspects of the Horner *et al.* checklist may have been met in the current case study, but the checklist was not systematically applied. Here we reflect on the quality indicators that were obviously not met.

Selection of treatment and control words

The selection of treatment and control words was led by the goal of creating matched sets of words that had a high likelihood of being present in AB's receptive lexicon but that he would show difficulty accessing. It has been demonstrated that psycholinguistic variables are a significant factor in word retrieval (for example, the number of syllables, frequency, the age of acquisition; German and Newman 2004). However, researchers have also argued that where generalization to untreated items is not predicted, the selection of treatment words should be based upon the child's school curriculum (German 2002, Parsons *et al.* 2005). A further consideration is the selection of nouns for treatment. The majority of intervention studies for WFDs have focused on nouns but research has suggested that verbs are as likely as nouns to cause WFDs (Dockrell *et al.* 2003). It could be argued that vocabulary from AB's school curriculum and verbs should have been included in the word sets.

Maintenance

Although we were able to collect maintenance data 8 months post-therapy for the vocabulary items, it was

not possible to reassess narrative performance due to the constraints of clinical service provision. Research that has compared treatment approaches have evaluated effectiveness based upon the superior maintenance of effects in a particular approach (for example, Nash and Snowling 2006), which demonstrates the necessity of including maintenance data wherever possible.

Reliability

Research using narrative samples typically includes inter-scorer reliability data to ensure consistency in transcription and analysis (for example, German and Simon 1991) and this would have been desirable in terms of providing evidence of best practice (Horner *et al.* 2005). This research was conducted in a real clinical environment, albeit part of a student experience, in a busy NHS clinic, and the data should have been subjected to reliability analysis. Although the narrative samples were recorded in order to achieve the highest possible levels of accuracy in transcription by the experimenter (the first author), recordings were unfortunately destroyed shortly after transcription. This is one limitation that must be overcome in further research, and practice managers who wish to strengthen the evidence base for the profession are encouraged to ensure that reliability of measurement is built into practice.

Baselines

In this investigation, a single baseline measure was taken for picture-naming and word-finding characteristics in narrative. This was useful in assessing the effectiveness of treatment. However, it would have been preferable to have obtained at least two baseline assessments (Horner *et al.* 2005). This would have taken into account the potential for inconsistency in naming or word-finding characteristics across time. Further, in order to calculate effect sizes following treatment, it has been suggested that two baseline measurements are required (Beeson and Robey 2006). The premise of Howard's (2007) approach is that trend analysis obviates the need for multiple baselines before intervention.

Extensions to treatment

Horner *et al.* (2005) suggested that in order to establish fully external validity of the treatment approach, treatment should be replicated across different participants, or in different settings or conditions. This was the first measurement of this novel treatment approach, and replication across participants is highly recommended.

Blinding

Although the senior fluency expert conducted the initial assessment and recommended assessment for a possible WFD, the first author conducted pre- and post-therapy assessments, except the maintenance probes, which were conducted by the senior fluency expert. The first author also carried out the intervention. Clearly, blinding was not applied and this is not ideal in clinical research.

Further assessment

Given AB's profile, with hindsight it would have been useful to assess his ability to do a non-word or sentence-repetition task to measure his short-term memory abilities, especially as sentence repetition was used in therapy. Also, an oral motor examination may have been useful given his hesitancies in connected speech.

Clinical implications

In spite of the limitations of the current investigation, the improvement shown in AB's naming accuracy suggests that NBLI-WF may be a useful addition to the range of techniques available to clinicians in treatment of WFDs. Of the available therapy approaches, 65% of clinicians employ interventions that target semantic representations for children with WFDs (Dockrell *et al.* 1998), so the rationale for NBLI-WF will be familiar to most clinicians. Furthermore, recent research has led to the recognition of narrative as a key area for both assessment and intervention of children's language development (Justice *et al.* 2006). NBLI-WF resulted in significant improvement in naming accuracy following just four sessions. This suggests that, with careful selection of treatment items, the approach could be adopted with a reduced duration and still achieve clinically significant improvement. However, the drop in naming accuracy 8 months post-therapy suggests that a distributed learning model might be useful in shoring up therapy gains. Research is needed to investigate the effects of top-up therapy, for example offered at 1-month intervals for 6 months, on maintenance.

Directions for future research

There is still a great deal of research required to identify the most effective approaches to intervention for children with WFDs. Studies with small sample sizes have shown the effectiveness of phonological (German 2002, McGregor 1994), semantic (Casby 1992, McGregor and Leonard 1989), and phonemic cueing (Best 2005) approaches. All these approaches require testing on larger samples of children, in-group versus individual designs, across age ranges, with children with

varying presentations in terms of concomitant language difficulties, and with variation in intervention agents (for example, speech therapy assistants). There has been very limited examination of the generalization of intervention approaches from single-word to sentence or narrative/discourse levels (German 2002). The current investigation is illustrative of one way in which this generalization may be evaluated and future research should consider the discourse/narrative level impact of WFDs, as well as testing the generalization of treated vocabulary to sentence or narrative/discourse contexts (Tingley *et al.* 2003). NBLI-WF requires more stringent evaluation using single-case or case series designs (Best 2005), which address the limitations of the current investigation, in order to determine whether the effects found for AB can be replicated with other children with WFDs.

We are cognisant of the realistic time pressures of a community clinic, but clinical practice needs to include provision for reliability measures, repeated baselines assessments, blinding of intervention agents, and replication across participants in order to allow clinicians to employ best practice in conducting clinical research to build a stronger evidence base for the profession. We intend to apply Horner *et al.*'s (2005) checklist in future research and encourage clinicians and clinical researchers to ensure that quality indicators are met.

Acknowledgements

The authors are grateful to the child and his family for their agreement and willing participation. Our thanks are extended to Lucy Patterson, senior speech therapist, for her support and professional assistance.

References

BEESON, P. M. and ROBEY, R. R., 2006, Evaluating single-subject treatment research: lessons learned from the aphasia literature. *Neuropsychology Review*, **16**, 161–169.

BERMAN, R. A. and SLOBIN, D. I., 1994, *Relating Events in Narrative: A Crosslinguistic Developmental Study* (Hillsdale, NJ: Erlbaum).

BEST, W., 2005, Investigation of a new intervention for children with word-finding problems. *International Journal of Language and Communication Disorders*, **40**, 279–318.

BISHOP, D. V. M., 2003, *The Test for Reception of Grammar, Version 2 (TROG-2)* (London: Psychological Corporation).

BRACKENBURY, T. and FEY, M. E., 2003, Quick incidental verb learning in 4-year-olds: identification and generalization. *Journal of Speech, Language, and Hearing Research*, **46**, 313–327.

CASBY, M. W., 1992, An intervention approach for naming problems in children. *American Journal of Speech–Language Pathology*, **1**, 35–42.

DOCKRELL, J. E., MESSER, D., GEORGE, R. and RALLI, A., 2003, Beyond naming patterns in children with WFDs—definitions for nouns and verbs. *Journal of Neurolinguistics*, **16**, 191–211.

DOCKRELL, J. E., MESSER, D., GEORGE, R. and WILSON, G., 1998, Notes and discussion: children with word-finding difficulties—prevalence, presentation and naming problems. *International Journal of Language and Communication Disorders*, **33**, 445–454.

DUNN, L. M., DUNN, L. M., WHETTON, C. and BURLEY, J., 1997, *British Picture Vocabulary Scale—2nd Edition (BPVS-II)* (London: NFER-Nelson).

FINESTACK, L. H., FEY, M. E., SOKOL, S. B., AMBROSE, S. and SWANSON, L. A., 2006, Fostering narrative and grammatical skills with 'Syntax Stories'. In A. V. Kleeck (ed.), *Sharing Books and Stories to Promote Language and Literacy* (San Diego, CA: Plural), pp. 319–346.

GERMAN, D. J., 1987, Spontaneous language profiles of children with word-finding problems. *Language, Speech, and Hearing Services in Schools*, **18**, 217–230.

GERMAN, D. J., 1994, Word-finding difficulties in children and adolescents. In G. P. Wallach and K. G. Butler (eds), *Language Learning Disabilities in School-Age Children and Adolescents: Some Principles and Applications* (New York, NY: Macmillan), pp. 323–347.

GERMAN, D. J., 2000, *Test of Word Finding—Second Edition (TWF-2)* (Austin, TX: PRO-ED).

GERMAN, D. J., 2002, A phonologically based strategy to improve word-finding abilities in children. *Communication Disorders Quarterly*, **23**, 177–190.

GERMAN, D. J. and NEWMAN, R. S., 2004, The impact of lexical factors on children's word-finding errors. *Journal of Speech, Language, and Hearing Research*, **47**, 624–636.

GERMAN, D. J. and SIMON, E., 1991, Analysis of children's word-finding skills in discourse. *Journal of Speech and Hearing Research*, **34**, 309–316.

GOLDRICK, M. and RAPP, B., 2002, A restricted interaction account (RIA) of spoken word production: the best of both worlds. *Aphasiology*, **16**, 20–55.

HOFFMAN, P. R., NORRIS, J. A. and MONJURE, J., 1990, Comparison of process targeting and whole language treatments for phonologically delayed preschool children. *Language, Speech, and Hearing Services in Schools*, **21**, 102–109.

HORNER, R. H., CARR, E. G., HALLE, J., MCGEE, G., ODOM, S. and WOLERY, M., 2005, The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, **71**, 165–179.

HOWARD, D., 2007, A new statistical test for trends: establishing the properties of a test for repeated binomial observations on a set of items, Paper presented at the Clinical Aphasiology Conference Scotsdale, Arizona, May, 2007.

JUSTICE, L. M., BOWLES, R. P., KADERAVEK, J. N., UKRAINETZ, T. A., EISENBERG, S. L. and GILLAM, R. B., 2006, The index of narrative microstructure: a clinical tool for analyzing school-age children's narrative performances. *American Journal of Speech–Language Pathology*, **15**, 177–191.

LEVELT, W. J. M., 1989, *Speaking: From Intention to Articulation* (Cambridge, MA: MIT Press).

LEVELT, W. J. M., 1991, *Lexical Access in Speech Production* (Cambridge, MA: Blackwell).

MARASCUILO, L. A. and MCSWEENEY, M., 1977, *Nonparametric and Distribution-Free Methods for the Social Sciences* (Monterey, CA: Brooks-Cole).

MAYER, M., 1969, *Frog, Where Are You?* (New York, NY: Dial).

MCGREGOR, K. K., 1994, Use of phonological information in a word-finding treatment for children. *Journal of Speech and Hearing Research*, **37**, 1381–1393.

McGregor, K. K. and Leonard, L. B., 1989, Facilitating word-finding skills of language-impaired children. *Journal of Speech and Hearing Disorders*, **54**, 141–147.

Messer, D. and Dockrell, J. E., 2006, Children's naming and word-finding difficulties: descriptions and explanations. *Journal of Speech, Language, and Hearing Research*, **49**, 309–324.

Nash, H. and Snowling, M., 2006, Teaching new words to children with poor existing vocabulary knowledge: a controlled evaluation of the definition and context methods. *International Journal of Language and Communication Disorders*, **41**, 335–354.

Oetting, J. B., Rice, M. L. and Swank, L. K., 1995, Quick incidental learning (QUIL) of words by school-age children with and without SLI. *Journal of Speech and Hearing Research*, **38**, 434–445.

Parsons, S., Law, J. and Gascoigne, M., 2005, Teaching receptive vocabulary to children with specific language impairment: a curriculum-based approach. *Child Language Teaching and Therapy*, **21**, 39–59.

Raven, J., Raven, J. C. and Court, J. H., 1998, *Coloured Progressive Matrices* (Oxford: Oxford Psychologists Press).

Renfrew, C., 1969, *The Bus Story: A Test of Continuous Speech* (Oxford: author).

Semel, E., Wiig, E. H. and Secord, W. A., 2003, *Clinical Evaluation of Language Fundamentals—Fourth Edition (CELF-4)* (San Antonio, TX: Psychological Corporation).

Soto, G., Hartmann, E. and Wilkins, D. P., 2006, Exploring the elements of narrative that emerge in the interactions between an 8-year-old child who uses an AAC device and her teacher. *Augmentative and Alternative Communication*, **22**, 231–241.

Swanson, L. A., Fey, M. E., Mills, C. E. and Hood, L. S., 2005, Use of narrative-based language intervention with children who have specific language impairment. *American Journal of Speech–Language Pathology*, **14**, 131–143.

Szekely, A., D'Amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., Jacobsen, T. and Bates, E., 2003, Timed picture naming: extended norms and validation against previous studies. *Behavior Research Methods, Instruments, and Computers*, **35**, 621–633.

Tingley, S. J., Kyte, C. S., Johnson, C. J. and Beitchman, J. H., 2003, Single-word and conversational measures of word-finding proficiency. *American Journal of Speech–Language Pathology*, **12**, 359–368.

Wiig, E. H. and Semel, E. M., 1984, *Language Assessment and Intervention for the Learning Disabled*, revd edn (Columbus, OH: Merrill).

Wing, C. S., 1990, A preliminary investigation of generalization to untrained words following two treatments of children's word-finding problems. *Language, Speech, and Hearing Services in Schools*, **21**, 151–156.

Wolf, M. and Obregon, M., 1992, Early naming deficits, developmental dyslexia, and a specific deficit hypothesis. *Brain and Language*, **42**, 219–247.

Appendix A: Psycholinguistic data for treatment and control items

Table A1. Treatment set

Word	Mean reaction time (ms)	Syllables	Frequency	AoA
Racket	963	2	0	3
Eskimo	1206	3	0	3
Screwdriver	1179	3	1.38	3
Stethoscope	1209	3	0.69	3
Broom	1092	1	4.55	3
Watering can	1577	4	0	3
Light switch	966	2	0	3
Parachute	1437	3	1.61	3
Magnet	1189	2	1.39	3
Anchor	951	2	1.95	3
Pumpkin	909	2	1.10	2
Ostrich	1337	2	1.39	3
Cactus	933	2	1.39	3
Drill	1311	1	2.20	3
Saddle	1019	2	2.40	3
Dolphin	894	2	1.39	3
Envelope	794	3	3.22	3
Spanner	1331	2	1.39	3
Binoculars	1055	4	4.649	3
Knight	1318	1	2.64	3
Flute	1402	1	1.39	3
Skeleton	817	3	2.57	3
Volcano	1063	3	1.95	3
Giraffe	783	2	1.10	1
Wheelbarrow	1207	3	0.690	3
Chimney	1169	2	2.40	3
Clothes peg	1589	2	0	3
Grasshopper	1234	3	1.39	3
Beaver	1395	2	1.39	3
Handcuffs	1113	2	1.10	3
Mean	1148.07	2.3	1.58	2.9

Table A2. Control set

Word	Mean reaction time (ms)	Syllables	Frequency	AoA
Porcupine	1291	3	0.69	3
Hammock	1378	2	0.69	3
Helmet	921	2	2.64	3
Celery	1362	3	1.39	3
Hoof	1088	1	2.20	3
Mosquito	1436	3	1.79	3
Telescope	1011	3	2.20	3
Globe	883	1	2.49	3
Piggybank	965	3	0	3
Eagle	1213	2	2.30	3
Palm tree	908	2	0	3
Fountain	966	2	2.57	3
Funnel	1243	2	1.10	3
Microscope	1212	3	2.20	3
Windmill	1226	2	2.30	3
Ladybird	1164	3	0	3
Lawnmower	1166	3	0	2
Leopard	1194	2	2.20	3
Igloo	963	2	0.69	3
Medal	1197	2	2.49	3
Mousetrap	1193	2	0.69	3
Paperclip	1262	3	0	3
Pelican	1102	3	1.10	3
Screw	1176	1	2.40	3
Pyramid	987	3	2.08	3
Corkscrew	1509	2	0.69	3
Trophy	1452	2	1.61	3
Saxophone	1061	3	0.69	3
Bench	896	1	3.18	2
Branch	1092	1	4.55	3
Mean	1150.57	2.23	1.57	2.93

Appendix B: Sample story and sentences for the NBLI-WF session

Vocabulary ($n = 7$)—*knight*, *saddle*, *anchor*, *volcano*, *parachute*, *pumpkin*, *skeleton*.

Story

Long ago, in a far away land, there was a brave *knight*. He had fought many battles and done many great deeds and he was looking forward to retiring and enjoying his days in peace.

The *knight* went to the queen and said, 'I wish to retire'. The queen replied, 'You have served me well but I have one last request. Find me something I have never seen before and you may retire.' 'I will do it!' exclaimed the *knight*.

The next morning the *knight* put the *saddle* on his horse and rode out to find something the queen had never seen. First, at the foot of the great *volcano*, he found the bones of a man but he knew the queen had seen a *skeleton* in her schooldays.

Next he rode to the coast and found a great, abandoned *anchor* lying on the beach. But he remembered the queen was born at sea. Then suddenly, the strangest thing happened. In the distance the knight saw a giant *pumpkin* floating downwards in the sky.

He rode towards it and caught the great *pumpkin* in his arms as it hung from a huge *parachute*. The *knight* hurried back to the castle and showed the *pumpkin* to the queen. 'There, you have something I have never seen before', said the queen with a smile. The *knight* bowed and left the castle to live the rest of his days in peace.

Sentences

(1) a) Knights fight on horseback.
b) The brave knight won the duel.

(2) a) Saddles are usually made of leather.
b) At the end of the day, the rider took the saddle off his horse.

(3) a) Volcanoes may erupt rock and lava.
b) The villagers escaped just before the volcano erupted.

(4) a) All of our bones make up a skeleton.
b) Jimmy was scared by the creepy skeleton.

(5) a) An anchor keeps a ship in place.
b) The captain dropped the anchor near the shore.

(6) a) Pumpkins are large, orange fruits.
b) At Halloween we make lanterns out of pumpkins.

(7) a) Parachutes slow you down when you are falling.
b) The parachute jump was scary but fun.