The Effectiveness of Semantic Therapy for Word Finding Difficulties in Pupils with Persistent Language Impairments: A Randomised Control Trial

Susan H. Ebbels[1,2], Hilary Nicoll[1], Becky Clark[1], Beth Eachus[1], Aoife L. Gallagher[1], Karen Horniman[1], Mary Jennings[1], Kate Montgomery[1], Liz Nimmo[1], Gail Turner[1]

[1]Moor House School, Surrey, UK

[2]Division of Psychology and Language Sciences, University College London, UK

Beth Eachus is now at Fairley House School, London; Karen Horniman is now at Greenshaw High School, Sutton, Surrey; Kate Montgomery is now a private speech and language therapist in Sydney, Australia; Liz Nimmo is now at Hebron School, Ootacamund, Tamil Nadu, South India.

Correspondence concerning this article should be addressed to Susan Ebbels, Moor House School, Mill Lane, Hurst Green, Oxted, Surrey, RH8 9AQ, UK. ebbelss@moorhouseschool.co.uk.

Keywords: intervention, randomised control study, word finding difficulties, language impairment, semantic therapy, adolescents

Abstract

Background: Word finding difficulties (WFDs) in children have been hypothesised to be caused at least partly by poor semantic knowledge. Therefore, improving semantic knowledge should decrease word finding errors. Previous studies of semantic therapy for WFDs are inconclusive.

Aim: To investigate the effectiveness of semantic therapy for secondary school-aged pupils with WFDs using a randomised control trial with blind assessment.

Methods & Procedures: 15 pupils with language impairments and WFDs (aged 9;11 to 15;11) were randomly assigned to a therapy versus waiting control group. In Phase 1, the therapy group received two 15 minute semantic therapy sessions per week for eight weeks with their usual speech and language therapist. Therapy for each child targeted words from one of three semantic categories (animals, food, clothes).

All participants were tested pre- and post-phase1 therapy on the brief version of the Test of Adolescent Word Finding (TAWF), semantic fluency, and the Test of Word Finding in Discourse (TWFD). In Phase 2, the waiting control group received the same therapy as the original group, who received therapy targeted at other language areas. Testing after Phase 2 aimed to establish whether the waiting control group made similar progress to the original therapy group and whether the original group maintained any gains.

Outcomes & Results: The original therapy group made significant progress in standard scores on the TAWF (d=0.94) which was maintained five months later. However, they made no progress on the semantic fluency or discourse tests. Participants in the waiting control group did not make significant progress on the TAWF in Phase 1 when they received no word-finding therapy. However, after Phase

2, when they received the therapy, they also made significant progress (*d*=0.81). The combined effect of therapy over the two groups was *d*=1.2. The mean standard scores on the TAWF were 67 pre-therapy and 77 post-therapy.

Conclusions & Implications: Four hours of semantic therapy on discrete semantic categories led to significant gains on a general standardised test of word finding, enabling the participants to begin to close the gap between their performance and those of their typically developing peers. These gains were maintained after 5 months. A small amount of therapy can lead to significant gains even with secondary-aged pupils with severe language difficulties. However, further studies are needed to find ways of improving word finding abilities in discourse.

What this paper adds

What is already known on this subject

Many children supported by speech and language therapy services have word finding difficulties. The majority of word finding errors are semantic in nature and are hypothesised to arise predominantly from semantic difficulties. According to this theory, semantic therapy should be effective. Previous intervention studies show that phonological therapy improves targeted words, but does not usually generalize to other words, while studies using semantic therapy have varying results.

What this study adds

This randomised control trial shows that semantic therapy targeted at one discrete semantic category can lead to significant gains on a general test of word finding. Progress was maintained, but did not generalise to discourse. Therefore

studies of intervention methods to transfer these improvements to general discourse are needed.

Introduction

Word-finding difficulties (WFDs) are characterized by hesitations, false starts, fillers (such as 'um', 'er'), empty words (such as 'thingy'), circumlocutions (where the child describes the word without accessing it), phonological distortions and semantic or phonological substitutions. Twenty-three percent of children supported by speech and language therapy services have WFDs (Dockrell, Messer, George, & Wilson, 1998), but evidence of the effectiveness of therapy is limited.

In confrontation naming tasks, children with WFDs are both slower and less accurate at naming pictures than typically developing (TD) children (Dockrell & Messer, 2007; Lahey & Edwards, 1996; Seiger-Gardner & Brooks, 2008). In terms of accuracy, three main types of errors have been identified in the literature: semantic substitutions, phonological errors (including both substitutions and distortions) and 'don't know' errors. Children with WFDs make more errors overall than TD children, but both groups make many more semantic than phonological errors (Dockrell, Messer, & George, 2001; Faust, Dimitrovsky, & Davidi, 1997; McGregor, 1997). In addition to overall error rates, children with WFDs differ from TD children in other ways; they produce a higher proportion of 'don't know' responses than TD children (Faust et al., 1997; McGregor & Windsor, 1996) and a higher proportion of phonological errors (Dockrell et al., 2001; Faust et al., 1997). However, in a study which looked at WFDs in children with different profiles of language impairments (Lahey & Edwards, 1999), differing patterns were found for the different profiles. Children with expressive language impairments only (with good receptive language) made a higher proportion of phonological errors than TD children. Meanwhile, children with both expressive and receptive impairments produced a higher proportion of semantic errors. We now consider each type of error separately.

Semantic errors

The most common semantic errors are coordinates (e.g., producing "cat" for "dog", McGregor, 1997) and these errors are associated with less detailed semantic representations. The detail of semantic representations was assessed by McGregor and colleagues by asking children to draw pictures and provide definitions of words they named correctly and those where they made errors. Both TD children (McGregor, Friedman, Reilly, & Newman, 2002) and children with Specific Language Impairment (SLI) (McGregor & Appel, 2002; McGregor, Newman, Reilly, & Capone, 2002) produced less detailed drawings and definitions for items where they made semantic naming errors than for those they named correctly. This suggested they had less detailed semantic representations for those words they had difficulty naming. Further indications of a link between the ability to name an item and the robustness of its semantic representation are given by findings that the proportion of items children with WFDs are able to name is similar to the proportion they are able to define (Dockrell & Messer, 2007). Also, while they are able to produce as many definitions as their age-matched peers, their definitions differ: they describe perceptual features of objects, rather than semantic categories (Dockrell, Messer, George, & Ralli, 2003). In addition, children with WFDs name letters and numbers (with minimal semantic content) as quickly and accurately as chronological age and language controls, but are worse at naming pictures (Dockrell et al., 2001).

These studies indicate that the underlying semantic representation of a lexical item influences the ability to name it correctly and that ‘sparse’ or ‘fragile’ representations could lead to semantic errors. Further indirect evidence of a link between naming and language representation or comprehension is given by the finding (Dockrell & Messer, 2007) that receptive language scores account for a

significant proportion of the variance in naming. This may account for Lahey and Edwards' (1999) finding that children with receptive language difficulties make a higher proportion of semantic errors while those with expressive difficulties but no receptive language difficulties make a higher proportion of phonological errors.

Semantic errors could arise for several reasons, as discussed by McGregor, Newman et al. (2002). The children could have a lexical gap where they do not know the item and therefore name another item which has similar physical features to those in the picture stimulus. Alternatively, they could fail to access the correct phonological representation, despite a good semantic representation ('tip-of-the-tongue' state). In these cases, children would have good drawings and definitions, but McGregor, Newman et al. (2002) found these cases were rare. Finally, the children's semantic representations could be 'sparse' or fragile. In these cases, they may access an organised semantic category (McGregor & Waxman, 1998), but instead of the target, retrieve a coordinate. This could be because there is not enough information in the semantic representations of either the target or its coordinate to be able to differentiate between them (Lahey & Edwards, 1999), therefore either could be produced. However, the coordinate is more likely than the target to be accessed if the coordinate is more frequent, earlier acquired or has more phonological neighbours (i.e., has a similar phonological form to many other words) than the target as it may then have a relatively better developed access path (German & Newman, 2004).

Semantic errors are reduced by semantic priming (hearing a sentence ending in a semantically related prime before seeing a picture of the target), but less in children with WFDs than TD children (McGregor & Windsor, 1996). The reduced effect of semantic priming in children with WFDs could be due to fewer links between lexical items (Sheng & McGregor, 2010) meaning that priming is less effective at increasing

the activation of the target or because poorly differentiated semantic coordinates are primed as well as the target (McGregor & Windsor, 1996).

Phonological errors

Phonological errors seem to differ both in frequency and in cause from semantic errors. In addition to being much less common than semantic errors for all children (Dockrell et al., 2001; Faust et al., 1997; McGregor, 1997), they are also associated with good drawings and definitions, therefore semantic information does not appear to be 'sparse' (McGregor & Appel, 2002). Phonological errors imply that the correct semantic representation has been accessed, but the phonological representation is either poorly defined or difficult to access (Constable, Stackhouse, & Wells, 1997). Phonological errors are predicted by a word's frequency and the frequency of its phonological neighbours (words with similar phonological forms, German & Newman, 2004). The word frequency effect could be due to higher activation thresholds of the phonological representations of low frequency words, making them harder to access. The effect of frequency of a word's phonological neighbours could be because phonological representations containing less common phoneme combinations are more difficult to learn and store and thus the phonological representations of these words may be more poorly defined and hence phonological distortions may result. Although children with WFDs produce a small number of phonological errors, they do produce a higher proportion than TD children (Dockrell et al., 2001; Faust et al., 1997), although this may be restricted to children with expressive language impairments and good receptive language abilities (Lahey & Edwards, 1999).

'Don't know' errors

'Don't know' errors are more common in children with WFDs than in TD children, but are still far less common than semantic errors (McGregor & Windsor, 1996). 'Don't know' errors are more common for words from sparse phonological neighbourhoods (i.e., words which have uncommon phonemes and combinations of phonemes), therefore the children may not be accessing the appropriate region of the phonological lexicon (German & Newman, 2004). These errors are not reduced by semantic priming (McGregor & Waxman, 1996), and therefore are more likely to be due to a phonological difficulty. However, McGregor, Newman et al., (2002) found that, like semantic errors, 'don't know' errors are also associated with less detailed drawings and definitions. Therefore, these errors may have many sources. In some cases they could be caused by sparse semantic knowledge and in some by difficulties accessing the phonological representation from an accurate semantic representation (e.g., Constable et al., 1997), or children may simply respond with "don't know" when they are unsure, cautious or forgetful (McGregor et al., 2002).

Summary

Semantic errors are the most common naming errors in all children, but especially in children with mixed SLI (Lahey & Edwards, 1999). Semantic errors are likely to be due to 'sparse' or fragile semantic representations and/or fewer or less robust semantic links between words. Therefore, therapy focused on elaborating semantic representations and semantic links may decrease semantic errors and hence improve naming in general. Phonological and 'don't know' errors are less common in all children, but children with expressive SLI show a slightly higher proportion of these errors (Lahey & Edwards, 1999). Phonological errors are possibly due to difficulties accessing an accurate phonological representation or to 'sparse' or fragile

phonological representations (Constable et al., 1997). These types of errors may be improved by therapy aimed at improving the detail of phonological representations or using phonological neighbours as a cue to access the appropriate region of the phonological lexicon. 'Don't know' errors may arise for a wide range of reasons, including sparse phonological or semantic representations and hence may be reduced by either phonological or semantic therapy. Thus, WFDs could be caused by difficulties with semantic and/or phonological representations. The relative difficulties with semantic versus phonological representations probably varies between individual children and some children may have difficulties with both. This may be related to their receptive abilities (Lahey & Edwards, 1999). Thus, both semantic and phonological therapy may be effective, but to different degrees in different children.

Intervention for WFDs

Intervention studies for WFDs vary in whether they focus on semantics, phonology or both. Semantic therapy tends to focus on developing knowledge of categorisation, attributes, definitions and associations between words, while phonological therapy tends to focus on identifying or counting syllables and phonemes and matching pictures or objects which rhyme or begin with particular phonemes.

Some intervention studies have focused both on phonology and semantics (Easton, Sheach, & Easton, 1997; McGregor & Leonard, 1989; Wittman, 1996; Wright, 1993) and found naming of targeted words improved with therapy. In three of these (Easton et al., 1997; McGregor & Leonard, 1989; Wittman, 1996), progress was maintained. In two (Wittman, 1996; Wright, 1993) progress generalised to control words, while in another (Easton et al., 1997) control words improved for a few

participants but this was not maintained at follow-up testing. These studies do not establish whether working on phonology or semantics contributed most to success.

Intervention studies focusing only on phonology found that naming of targeted words improved with therapy and this effect was maintained after therapy ceased, but the effects did not generalise to other words (Best, 2005; German, 2002; McGregor, 1994). However, Best (2005) did find that when analysed as a group, the children's naming of control items did improve, but this could be due to general maturation. She also found generalisation to discourse for two of the five children. The therapy methods varied between studies: McGregor (1994) encouraged the children to identify the initial sound and number of syllables of target words and German (2002) focused on identifying syllables and phonological neighbours of target words. Best (2005) used a computer to provide letter cues: when a child couldn't retrieve the target word, the computer provided an array of possible first letters, the child then chose the correct letter from the array and the computer converted it to a phoneme cue.

Intervention studies focusing predominately on semantics have also found that the words targeted in therapy improved. Casby (1992) focused on "deep processing" of stimulus words, asking the child to say something about the picture and use it in a sentence (which uses both semantic and syntactic processing). This led to faster and more accurate naming of target items post-therapy. Therapy focused on semantics embedded in a narrative approach (Marks & Stokes, 2010) improved naming of targeted, but not control words, and progress was maintained. However, they found no reduction of WFDs in discourse. In contrast, Stiegler and Hoffman (2001) showed that after therapy targeted specifically at WFDs in discourse focusing mainly on semantics (but also involving requests for clarification, restructuring and phonemic

cueing if necessary) their three participants had fewer WFDs in discourse. However this could have been due to non-specific effects as no effective control was included.

Some studies have aimed to directly compare interventions focusing on semantics versus phonology (Wing, 1990; Wright, Gorrie, Haynes, & Shipman, 1993). Wright et al., (1993) found children receiving semantic therapy improved significantly more at naming words not targeted in therapy than their controls. In contrast, children receiving phonological therapy did not improve more than their controls. Wing (1990) compared semantic with phonological / perceptual therapy (the perceptual tasks involved the children looking at a picture, closing their eyes, seeing the picture and hearing its name in their mind, naming the picture and locating the picture from an array of six). She found naming on the Test of Word Finding (German, 1986) improved after the phonological / perceptual therapy. The children receiving semantic therapy did not as a group show improved naming on the Test of Word Finding. However, there were only five children in each group and four out of five of the semantic group made progress. The results may have been skewed by the one child who achieved a lower score post-therapy than pre-therapy. If this child is removed from the analysis, the semantic therapy group also made significant progress.

In summary, the therapy studies to date seem to show that phonological therapy improves only the words targeted in therapy and does not usually generalise to other words. Wing (1990) is an exception, but this could be due to the addition of the perceptual cues. There is some evidence (Marks & Stokes, 2010; Wright et al., 1993) for the effectiveness of semantic therapy at improving the naming of pictures, but this is not conclusive (Wing, 1990). However, since the majority of word finding errors are semantic and are thought to be caused by ‘sparse’ semantic representations, we hypothesise that a semantic approach should be effective, particularly for children

who make the most semantic errors; those children with mixed receptive and expressive language impairments (Lahey & Edwards, 1999). We wished to investigate the effectiveness of semantic therapy for such children. Therefore, we selected a subgroup of school-aged children with both receptive and expressive language difficulties who also had WFDs and investigated whether those receiving semantic therapy made more progress than those who received no therapy targeted at WFDs.

The majority of the participants in this study were secondary-aged (over 11 years). This is a neglected group in terms of intervention research. Of the studies of WFD therapy discussed above, the majority are with primary-aged children (5-11 years). Only one (Wright et al., 1993) involved secondary-aged young people with language impairments and/or WFDs. The lack of intervention research with this age group is also reflected in the intervention literature for other areas of language, such as syntax and morphology, where only four studies (Bishop, Adams, & Rosen, 2006; Ebbels & van der Lely, 2001; Ebbels, 2007; Ebbels, van der Lely, & Dockrell, 2007) have, to the authors' knowledge, been published investigating the effectiveness of intervention for secondary-aged young people with language impairments. Clinical services to this age group are also limited (Dockrell, Lindsay, Letchford, & Mackie, 2006). This is despite evidence that language impairments persist into early adolescence (Beitchman, Wilson, Brownlie, Walters, & Lancee, 1996; Botting, Faragher, Simkin, Knox, & Conti-Ramsden, 2001) and beyond into late adolescence and adulthood (Clegg, Hollis, Mawhood, & Rutter, 2005; Conti-Ramsden, 2008; Mawhood, Howlin, & Rutter, 2000) and have negative effects on children's educational achievements (Conti-Ramsden, 2008; Dockrell & Lindsay, 2008; Mawhood et al., 2000) and social adjustment (Clegg et al., 2005; Conti-Ramsden, 2008; Howlin, Mawhood, & Rutter, 2000). The current study will add to the limited

evidence base regarding the effectiveness of therapy for secondary-aged children with language impairments and/or WFDs. As a randomised control trial, this study may provide stronger evidence than previously published studies on intervention for WFDs with this age group (e.g., Wright et al., 1993).

Method

Participants

This study was carried out at a specialist residential school for pupils with severe language impairments. Speech and language therapists (SLTs) referred into the study pupils who they judged required therapy for WFDs. These pupils were tested with the Test of Adolescent/Adult Word Finding (TAWF, German, 1990) to confirm WFDs (standard scores <85); one pupil was excluded as his score was too high on the TAWF. Fifteen participants remained, mean age 13;5 (9;11-15;11), 11 males and 4 females, a ratio of 3:1. These participants were also tested on other standardised language tests: the Clinical Evaluation of Language Fundamentals – 3 UK (CELF-3 UK, Semel, Wiig, & Secord, 1995) and the British Picture Vocabulary Scales-II (BPVS-II, Dunn, Dunn, Whetton, & Burley, 1997). Mean pre-therapy standard scores on the specific WFD measures used and other standardised tests are shown in Table 1. Individual data on all tests at all timepoints are shown in Appendices A and B.

TABLE 1 ABOUT HERE

The standard scores on the TAWF confirmed that the pupils had WFDs[1]. However on the Phonological Awareness Battery (PhAB, Frederickson, Frith, &

[1] One participant (T5 in Appendix A) achieved a score adjusted for comprehension above 85 (see below for description of how this adjusted score is calculated), but his unadjusted score was 76. Thus it is possible that some of his WFDs were a reflection of poor vocabulary levels rather than WFDs per se, but given that his adjusted score was only just above the cut-off (85) and his SLT considered he had WFDs, he was included in the study.

Reason, 1997) Semantic Fluency test, the standard scores varied widely, indeed eight participants scored within the normal range. The general language tests showed all the participants except one (WC3 in Appendix B, whose general language scores are excluded from the right-hand pair of columns in Table 1, as their inclusion distorts the general picture) had general receptive and expressive language difficulties. For this participant, despite language scores in the normal range, he gained a TAWF standard score of 69 showing that he did have WFDs and was thus included in the study[2].

Most of the participants were outside the age range for calculating standard scores on the Test of Word Finding in Discourse (TWFD, German, 1991). Their scores in terms of the percentage of T-units containing at least one WFD (as defined in the test manual and discussed below) ranged from 21% to 58% with an average of 38%, showing that their WFDs occurred frequently in discourse.

Measurements

Test of Adolescent/Adult Word Finding (TAWF, German, 1990)

The TAWF is a standardised test of word finding for adolescents and adults from age 12 upwards. It requires participants to name pictures (nouns and verbs), complete sentences with missing words, name items from descriptions and name categories on hearing a list of members. A brief version is available, which was used in this study. The manual (p63) also describes a method of rescoring for low comprehension. This is recommended for any individual who scores below 90% on comprehension of the target items. This prorated accuracy score represents the individual's naming only on those words he or she knew (as indicated in the comprehension section). In order to do this, the percentage of known words named accurately is calculated and a prorated accuracy standard score can then be found in

[2] He also had pragmatic language difficulties.

the tables provided[3]. We used the TAWF to identify participants with WFDs at the start of the project and to measure progress with therapy in general WF abilities.

Test of Word Finding in Discourse (TWFD, German, 1991)

The TWFD is a standardised test of word finding in discourse for children aged 6;6-12;11. It consists of 3 composite pictures. The participants first have to describe the picture (e.g., park scene) and then say how it would be different if….(e.g., it were snowing). The tester gives minimal prompts and does not discourage any deviation in topics, as long as the variation from the original is self-imposed. The minimum length is 21 'T-Units'[4] and the maximum analysed is 60 T-Units. The participants' responses are recorded and transcribed later. All WFDs are noted and the raw score calculated is the percentage of 'T-units' with at least one WFD. Standard scores were not available for the age range of the majority of the participants in this study, thus we only used raw scores. We included this test to determine whether any progress in word finding generalised to discourse.

Phonological Awareness Battery (PhAB, Frederickson et al., 1997) Semantic Fluency test and additional semantic fluency testing

We asked the participants to list as many animals, food and clothes (the three topic areas used in the therapy) as possible in one minute. Their responses were transcribed on-line and recorded for later checking and scoring. Each participant was assigned two types of scores. The first type was the number of (different) items listed

[3] Because the prorated standard score depends on the number of items correctly comprehended, changes in the raw score and prorated standard score need not necessarily go in the same direction. For example, participant T5 had a raw score of 26 pre-therapy falling to 23 post-phase1 therapy. However, his comprehension dropped by a greater amount from 35 known words (out of 40) to 29. Thus, his prorated standard score actually increased from 86 to 92.

[4] "A T-unit (Hunt, 1965) is the shortest unit into which a linguistic utterance can be divided without leaving a remaining fragment; it consists of a main clause plus the subordinate clauses that are attached to the main clause." (German & Simon, 1991, p311)

within each category within the minute and the mean of these three scores. We used one minute (rather than 30s as used in the PhAB) as we assumed a longer period would be more likely to show change with therapy. We used these to measure progress on the semantic areas targeted in therapy. The first 30 seconds of food and animals can be converted into a standard score using the PhAB (a standardised test for children aged 6;0-14;11). The one minute recordings were used to note the number of items listed in each of these two categories in the first 30 seconds. These raw scores are combined in order to obtain a standard score for semantic fluency. This second score was used to establish whether the participants receiving therapy made progress in semantic fluency relative to their TD peers.

Study Design

This study is a randomised control trial with blind assessment. The fifteen participants were randomly assigned to two groups (Therapy vs. Waiting Control) by the first author using the random number function in Excel to sort the participants into a random sequence. The order of assignment of the random sequence to therapy group and phase had been pre-determined (i.e., the first eight participants to the therapy group and the next seven to the waiting control group). Within each group, the participants were then randomly assigned one of three semantic categories for intervention work: food, animals or clothes. The therapy group received therapy in the Summer Term (Phase 1) and the waiting controls in the Autumn Term (Phase 2).

All participants in the study were assessed pre-therapy (in April), immediately after Phase 1 (in July) and immediately after Phase 2 (in December). All post-therapy testing was carried out by visiting speech and language therapy students who were blind to the participants' group assignment. The majority of the pre-therapy testing was carried out by a volunteer (a recently qualified SLT) who was also blind to group

assignment. However, the volunteer was unable to complete the pre-therapy testing, so for five participants, pre-therapy testing on the TAWF and TWFD was carried out by their own SLT.

During periods when participants were not receiving WFD therapy (Phase 1 for waiting controls and Phase 2 for the original therapy group), they continued with their normal therapy package (however, word finding was not targeted during this time). Therefore both groups received their normal amount of therapy at all times, but during their WFD therapy phase, had therapy targeted at word finding for 30 minutes per week while the other group had therapy on other language areas (at the discretion of their therapist).

Analyses of the pre-therapy scores for the two groups showed there were no differences between the two groups in gender distribution, $\chi^2(1)=0.02$, $p=1.0$, age, $t(13)=0.97$, $p=.35$, $d=0.5$, standard scores on the CELF-3 Expressive Language[5], $W=56.0$, $n_1=7$, $n_2=8$, $p=1.0$, CELF-3 Receptive Language, $W=51.5$, $n_1=7$, $n_2=8$, $p=.62$, BPVS, $t(12)=.087$, $p=.40$, $d=0.47$, TAWF, $t(12)=0.83$, $p=.42$, $d=0.45$, or PhAB, $t(13)=0.32$, $p=.75$, $d=0.17$, or raw scores on the TWFD, $t(12)=0.39$, $p=.70$, $d=0.21$.

Therapy method

Participants were seen twice per week, for 15 minutes, for 8 weeks (4 hours in total). The therapy was usually provided by the participants' own SLT as part of their normal therapy package. There were four exceptions to this; one participant (T7) was seen entirely by the first author (who was not their usual SLT), and three (T6, WC5, WC6) who were seen by their own SLT and a speech and language therapy assistant. In these latter cases, the assistant and SLT ran alternate sessions but the SLT assistant

[5] Wilcoxon signed ranks test was used for non-normally distributed data

also sat in on the sessions run by the SLT. The structure of the programme (which was jointly planned, recorded and followed by all the SLTs participating in the project) meant that the assistant repeated the session she had observed from the previous week (see the schedule in Appendices C and D).

Each participant was taught using photo cards (Franklin, 1994) of one particular category (animals, clothes or food)[6]. Within each category the SLTs chose which cards to use for each participant to allow flexibility for their differing levels of ability (partly due to the wide age range). The main activities consisted of sorting pictures by semantic categories (broad, then narrower), discussing the semantic attributes of the pictures, comparing pictures in terms of these attributes and using these attributes and categories as cues in games. In the second session of each week a new step was introduced, followed by practice of the previous steps. The full details of the therapy are given in Appendix C and the schedule of steps in Appendix D.

Attendance and Treatment Fidelity

Attendance and treatment fidelity were assessed indirectly by the first author who interviewed the SLTs after the completion of the therapy programme and checked their case notes, which are written within 24 hours of each therapy session. In the interviews, SLTs were asked about the amount of therapy received by each participant and the actual content and timing of sessions.

No participant withdrew from the study at any point. However, two participants in the waiting control group did not complete the therapy programme in Phase 2. One (WC6) did not receive the last two 15 minute sessions because he was unavailable due to other school activities. One (WC2) only received 10 of the 16

[6] the brief version of the TAWF only includes two items, a starfish and seahorse, which could have been included in the therapy on animals. The words *paw* and *seed*, also appear in the test and could possibly have been mentioned in the therapy on animals and food respectively, but not targeted.

sessions. This was due to her absence on college visits and requiring time in therapy sessions to discuss other issues unrelated to the word finding therapy. Her data were not included when analysing the progress of the waiting controls in Phase 2, but were included during Phase 1 when she did not receive any therapy.

The majority of the participants received the therapy as planned in sixteen 15 minute slots, delivered twice per week over eight weeks. One of the Phase 1 therapy group (T8) initially had fortnightly 30 minute sessions (covering the content of two 15 minute sessions) due to SLT illness and the need to spend other sessions discussing emotional issues unrelated to the word finding therapy. After this period, the first author helped deliver the remainder of the sessions, which were delivered in 15 minute slots, 3-4 times weekly in order to complete the therapy programme before the re-testing period. This participant therefore received the same total amount of therapy time as the others but unevenly distributed. One participant from the Phase 1 therapy group (T2) had 16 sessions which were twice as long as planned because a severe stammer meant his SLT could not complete the planned programme for each session within 15 minutes. He therefore received twice as much therapy time as the other participants, but covered the same content in the same order.

According to the SLTs' notes and their reports when interviewed, they all followed the basic plan of the therapy. Several added games to increase motivation. Two (KH & KM) added a game of "Connect 4" (Hasbro, 2004) at the end of the session while two (BE & KM) added games during the sessions, whereby the participant / therapist could have a turn at the game after each step of therapy (e.g., after describing the attributes of one picture, or correctly guessing a picture). Another (SE) added a points system to Steps 9 and 10 of the schedule (see Appendix C), whereby the person giving cues got a mark for the total number of cues given before

the 'guesser' correctly identified the word. The person with the least marks (i.e., who gave the best cues) was the winner.

The WFD therapy was provided as part of the participants' normal therapy package. The normal package at the school includes joint planning and teaching of English lessons and support in some other lessons, although this varies with the age of the pupils and their needs. All pupils also receive one Social and Interaction Skills group (1 hour) per week and at least one other group (1 hour) per week targeting specific areas. The focus of the groups attended by the participants during the study varied greatly and included: vocabulary, phonological awareness, past tense, conjunctions, text comprehension and inferencing, social and pragmatic skills. The older participants also had at least one additional group, focused more on functional skills (e.g., making phone calls and giving verbal presentations), but had less individual therapy than the younger participants. The total amount of individual therapy varied with age and with the individual participant's profile of difficulties; the amount and topic of the therapy was at the discretion of the individual SLT (except for during the WFD therapy phase when they had to follow the study protocol). All of the participants received individual therapy while they were in the control phase of the study (Phase 1 for waiting controls and Phase 2 for original therapy group), this focused on a wide range of areas (articulation/phonology, phonological awareness, vocabulary, past tense, conjunctions, inferencing/prediction, stammering and proof-reading), but not WFDs. Some participants also received extra individual therapy during their WFD therapy phase, this focused on articulation, phonological awareness, conjunctions, past tense, narratives and inferencing.

In order to test whether the amount of therapy differed between the two groups in either of the two phases of the study, we carried out independent sample t-tests

comparing the amount of therapy received by the two groups in each phase. We compared the amount of individual therapy, group therapy and the total amount of therapy combined. None of these analyses showed any significant differences between the groups (Phase 1: individual therapy, $t(13)=1.0$, $p=.36$, $d=0.5$; group therapy, $t(13)=1.1$, $p=.29$, $d=0.57$; total therapy $t(13)=0.44$, $p=.65$, $d=0.23$; Phase 2: individual therapy, $t(12)=0.003$, $p=1.0$, $d<0.01$; group therapy, $t(12)=0.2$, $p=.87$, $d=0.09$; total therapy $t(12)=0.1$, $p=.89$, $d=0.08$).

Results – Phase 1

Test of Adolescent/Adult Word Finding (TAWF)

The standard score on the TAWF can be adjusted for the participants' comprehension of the target words, thereby measuring word finding and not vocabulary levels (see above for an explanation of how this is done). We used these prorated standard scores throughout the project. Two participants were too young to calculate standard scores (one from each group) – the participant from the Phase 1 therapy group (T5) was only two months too young at the start of the study and therefore we used standard score for 12 year olds; the participant from the waiting control group (WC7) was two years too young and therefore standard scores were not calculated for him, but his raw score on this test was used to analyse progress.

The boxplot in Figure 1 shows the pre- and post-phase1 therapy prorated standard scores (adjusted for comprehension levels) on the TAWF. This shows that the therapy group improved, while the waiting control group did not. A comparison of the change in prorated standard score (post-phase1 therapy minus pre-therapy) between the two groups using a one-tailed[7] *t*-test revealed a significant difference

[7] We used one-tailed tests for all analyses as we were interested only in whether therapy improved performance, therefore we were looking for significant change in one direction only.

between the two groups, $t(12)=1.93$, $p=.04$, $d=1.00$, where the therapy group showed more progress than the waiting controls.

FIGURE 1 ABOUT HERE

In order to test whether progress made by either group was significantly better than zero, we looked at the change in raw score for each group (standard scores were not used due to difficulties with regression to the mean). These results are shown in the boxplot in Figure 2. One-tailed, one-sample *t*-tests on the change in raw score for each group showed that the therapy group made progress which was significantly greater than zero, $t(7)=2.6$, $p=.02$, $d=0.94$, while the waiting controls did not, $t(6)=1.2$, $p=.14$, $d=0.45$.

FIGURE 2 ABOUT HERE

Test of Word Finding in Discourse (TWFD)

Due to equipment failure, pre-therapy data for the TWFD was not available for one of the waiting controls (WC3) so his data from all time points were excluded from all tables and analyses. Table 2 shows three different scores for the TWFD: pre- and post-phase1 therapy scores for the two groups and also the change in their scores. The first score is the percentage of T-units containing one or more WFDs. This is the score recommended in the manual of the TWFD. However, we felt that this might not capture progress sufficiently. For example, if a participant had an average of four WFDs per T-unit and this reduced to an average of two, they may still have the same percentage of T-units with at least one WFD and thus their progress would not be captured. Therefore, we calculated two additional scores which we felt might be more sensitive: the mean number of WFDs per T-unit and the percentage of WFDs relative to the total number of words (as used by Stiegler & Hoffman, 2001).

TABLE 2 ABOUT HERE

Both groups appear to show little change on all measures. Indeed on the percentage of T-units containing at least one WFD, the mean score for both groups increased, indicating an increase in WFDs in discourse. However, the scores on the other two more sensitive measures both show a decrease for the therapy group, showing a change in the desired direction (although the waiting controls also showed a decrease on one of these measures). One-tailed *t*-tests (or Wilcoxon signed ranks tests for the more sensitive measures due to non-normal distributions) comparing the amount of progress made by each group (post-phase1 therapy score minus pre-therapy score) showed the two groups did not differ significantly on any measure: % T-units containing WFDs, $t(12)=1.4$, $p=.09$, $d=0.76$; WFDs per T-unit, $W=57.0$, $n_1=8$, $n_2=6$, $p=.38$; WFDs / total words, $W=56.0$, $n_1=8$, $n_2=6$, $p=.33$. Thus there is no evidence of progress on this test with therapy.

Semantic fluency

Table 3 shows the mean number of items listed in the three categories in one minute pre-therapy, post-phase1 therapy and the change in score. More detailed, individual data are in Appendices A and B. Both groups appear to show little change. A one-tailed *t*-test comparing the amount of progress made by the two groups (post-phase1 therapy score minus pre-therapy score) showed they did not differ significantly, $t(13)=0.13$, $p=.36$, $d=0.18$.

TABLE 3 ABOUT HERE

The first 30 seconds of food and animals can be converted into a standard score using the PhAB. This was not available for one of the waiting controls (WC3) pre-therapy and one of the waiting controls (WC6) post-therapy due to difficulties with the audio recordings. Their results are therefore excluded at all time points. The

data for the remaining participants are also shown in Table 3. Three participants (WC2, T4, T2) were too old for the standardisation range of the PhAB, but we used the scores for the oldest age range listed (up to 14;11), one year younger than the oldest participant.

Comparisons of the change scores on the PhAB of the two groups using one-tailed *t*-tests showed no significant difference between the groups, $t(11)=0.33$, $p=.37$, $d=0.20$. If the scores for the participants who were older than the standardisation range are removed from the analyses, the differences between the groups were still not significant $t(8)=0.72$, $p=.25$, $d=0.46$. However, it should also be noted that eight participants (T1, T3, T4, T5, T6, WC1, WC2, WC6) scored within the normal range (SS>85) on this test pre-therapy and therefore it may have been unrealistic to expect positive change.

Summary of Phase 1

The therapy group improved significantly more than the waiting controls on the TAWF, but not on the semantic fluency tests or TWFD. The second phase of the study aimed to establish whether the therapy group maintained the progress they had made on the TAWF with no further therapy on WFDs and whether the waiting controls made a similar amount of progress when they too received the therapy. Given the null results of the semantic fluency and TWFD tests, these were not repeated. The only measure used in Phase 2 was the TAWF.

Results - Phase 2

Test of Adolescent Word Finding (TAWF)

In order to establish whether progress was maintained for the original therapy group and whether the waiting controls made a similar amount of progress when they

also received therapy, the same analyses were carried out as before, but this time using changes from pre-therapy to post-phase2 therapy. The data from the waiting control who did not complete the therapy in Phase 2 (WC2) are excluded from all graphs and analyses in this section. The boxplot in Figure 3 shows the pre- and post-phase2 therapy prorated standard scores on the TAWF. This shows that both groups have now improved. A comparison of the change in prorated standard score (from pre-therapy to post-phase2 therapy) between the two groups using a two-tailed *t*-test showed the previous difference (post-phase1 therapy) between the two groups has disappeared, $t(12)=0.35$, $p=.74$, $d=0.19$.

FIGURE 3 ABOUT HERE

In order to test whether progress made by either group over Phase 1 and Phase 2 combined was significantly better than zero, we looked at the change in raw score for each group (standard scores were not used due to difficulties with regression to the mean). The results are shown in Figure 4. One-tailed, one-sample *t*-tests on the change in raw score for each group showed that the original therapy group made progress over the whole study (pre-therapy to post-phase2 therapy) which was significantly greater than zero, $t(7)=2.8$, $p=.01$, $d=0.99$, as did the waiting controls, $t(5)=3.9$, $p=.006$, $d=0.81$. In order to get an overall effect size for progress made by both groups over the whole study, we combined the two groups and compared their progress (from pre-therapy to post-phase2 therapy) to zero and found a significant effect: $t(13)=4.6$, $p<.001$, $d=1.22$. The combined effect for just the period when enrolled in the therapy (progress over Phase 1 for the original therapy group and over Phase 2 for the original waiting controls) was $t(13)=4.6$, $p<.001$, $d=1.20$.

FIGURE 4 ABOUT HERE

Possible factors influencing progress on the TAWF

In order to investigate the possible factors which could have influenced progress on the TAWF, we carried out correlations between change in TAWF raw score with therapy (over Phase 1 for the original therapy group and over Phase 2 for the original waiting controls) and pre-therapy standardised test scores. We used Pearson's correlations for the TAWF, BPVS and PhAB semantic fluency tests and Spearman's non-parametric rank correlations for the CELF receptive and expressive language scores (due to a floor effect on these tests). Progress on the TAWF was not significantly correlated with pre-therapy raw scores on the TAWF, r=-.15, p=.61 or any pre-therapy standard score: TAWF, r= -.32, p=.27; BPVS, r= -.08, p=.79; PhAB Semantic Fluency, r=.05, p=.87; CELF Receptive Language, r=0.19, p=0.52; CELF Expressive Language, r=0.45, p=0.88.

The TAWF includes a method for categorizing participants as 'fast' or 'slow' namers (see Appendices A and B for individual categorisations). We compared whether the amount of progress differed for children assigned to these two groups. A non-parametric Wilcoxon's signed ranks test showed no difference in the amount of progress between participants who were fast versus slow namers pre-therapy, W=27.5, n_1=4, n_2=9, p=.97.

Discussion

Four hours of therapy focused on semantics significantly improved word finding ability on a standardised test of word finding (TAWF) among secondary-aged pupils with receptive and expressive language impairments and WFDs. On average (the original therapy group over Phase 1 and waiting controls over Phase 2 combined), the participants increased from a standard score (adjusted for comprehension) immediately pre-therapy of 67 to immediately post-therapy of 77.

The waiting controls made little progress during their baseline period (Phase 1) but made progress with a similar effect size (d>0.8) to the original therapy group when they too had received the therapy (after Phase 2). Progress could not have been due to a placebo effect as both groups received their normal therapy package at all times, but during their WFD therapy phase, had therapy targeted at word finding for 30 minutes per week while the other group had therapy on other language areas. The original therapy group maintained progress for 5 months after their therapy had ceased. Unfortunately however, progress did not generalise to discourse and there was no progress on semantic fluency. The amount of progress made on the TAWF was not correlated with pre-therapy performance on word-finding or general language measures.

Our finding that the participants improved on a general test of word finding, when their therapy was only on one specific category, strengthens the findings of Wright et al., (1993), but using a stronger RCT design, that semantic therapy can improve naming of words not targeted in therapy. In addition we showed that these effects can be maintained over a period of 5 months. This raises the question of how the therapy worked. In the introduction we discussed the hypothesis that sparse or fragile semantic representations could lead to WFDs (McGregor, Newman, Reilly, & Capone, 2002), where coordinates are accessed in preference to the target (especially when the coordinate is higher frequency and/or has more phonological neighbours, German & Newman, 2004). Fewer and weaker semantic links could also be at fault (Sheng & McGregor, 2010). The therapy could have worked by increasing and strengthening semantic links between words and by increasing the detail in the semantic representations of targeted words, so that they are less sparse. However, the therapy could not have worked just by improving the semantic representations of

targeted words as the effects generalised to other words. Therefore the effect must have been general. Perhaps by improving awareness of the rich semantic links between words and the level of detail possible in semantic representations, the participants think in more detail about the semantics of all pictures to be named. This would enable them to distinguish between the target and competing coordinates with broadly similar semantic representations and thus encourage retrieval of the target. Older children may be more able to take advantage of this more meta-cognitive approach which could go some way to explaining why semantic therapy appears to be more effective in this study and that of Wright et al. (1993) than in Wing (1990) which involved younger children, but this requires further investigation.

Our analyses of the possible factors underlying the amount of progress made on the TAWF did not reveal any factors which correlated with progress. Thus, pre-therapy word finding ability and general language ability had no effect on the amount of progress. This indicates that the semantic therapy used in this study was effective regardless of the level of language impairment or WFDs. However, our participants were relatively homogeneous with respect to general language (dis)abilities, thus the lack of variance in their language scores may contribute to the lack of correlation between these and progress with therapy. Recruiting participants with a wider range of language abilities, including those with receptive or general language abilities in the normal range, may reveal some influence of pre-therapy language abilities on progress.

Our study found no change on the semantic fluency task and the associated PhAB score, despite good progress on the TAWF. Differences between the tests could account for these findings. Indeed, Dockrell and Messer (2007) found semantic fluency did not correlate with Test of Word Finding (TWF, German, 1986) for

children with WFD, although they were impaired on both. On the semantic fluency tasks (including the PhAB), within category semantic errors are not possible to identify because the tester does not know the target the participant is aiming for. Therefore coordinate 'errors' would be scored. Also, half of our participants already scored in the normal range pre-therapy as measured by the PhAB, so had less potential for progress.

We also found no change on the TWFD. This is similar to the Marks and Stokes (2010) study which found no decrease in errors on this test using the score of percentage of T-units containing at least one WFD. Best (2005) found three of her five participants did not improve their word finding in discourse, while two did. Stiegler and Hoffman (2001) showed a decrease in WFDs in discourse using a measure of percentage of WFDs relative to total words produced. On average, their three participants showed a decrease from 9.4% to 7.2% (a drop of 2.2%) which they concluded showed the therapy was effective. Our therapy group showed a decrease from 12.3 to 8.8 (a drop of 3.5%). However, our analyses showed that this decrease was not significantly different from the control group (who showed a greater drop). Thus, the findings of Stiegler and Hoffman (2001) must be treated with caution as their study provided no effective control. Thus, only one study (Best, 2005) has been able to document a statistically significant change in WFDs in discourse with therapy, but this was only for two out of five children. The greater linguistic demands of discourse could restrict progress, particularly among our participants who all had language impairments in addition to their WFDs. Further research is urgently needed in this area to find methods and amounts of therapy which will improve word finding in discourse.

Limitations and future directions

The RCT presented here was small in scale and based in one school. Thus, it cannot be assumed that the results will generalise to other groups of young people with WFDs and/or language impairments of different ages or in different settings. Future work could aim to establish whether similar results to our study are obtained for the same semantic therapy but with different age groups or children with "pure" WFDs with no other apparent language difficulties. Different methods of delivery could also be investigated. Indeed, we are currently investigating whether this therapy method is effective when delivered in groups and pairs, when delivered by experienced SLTs, newly qualified SLTs and SLT students (Nicoll & Ebbels, in prep.)

The precise aspects of the semantic therapy which are responsible for progress could be investigated and also the amount of therapy which is needed to make significant or optimal gains. Future work could also compare phonological therapy directly with semantic therapy, aiming to establish whether one method is more effective in general for children with WFD (with or without additional language impairments) or whether children with particular profiles of difficulties respond differently to the different types of therapy. In addition, future studies urgently need to investigate different methods of therapy to find a method which will improve word finding in discourse.

Conclusions

School-aged participants (mostly secondary aged) receiving four hours of individual therapy focused on semantics, made significant progress in their general word finding abilities as shown by their scores on a standardised test of word finding. Waiting controls made no progress until they also received the therapy, when they made similar gains. Progress was maintained for five months, but did not generalise to

discourse. Thus, the therapy was effective in improving word finding abilities in confrontation naming tasks, but further work is required to establish the best ways to achieve generalisation to conversational speech.

This study contributes further to existing evidence (Ebbels & van der Lely, 2001; Ebbels, 2007; Ebbels et al., 2007; Wright et al., 1993) that therapy for adolescents with WFDs and/or language impairments can be effective. Thus, this age group should not be neglected in the provision of therapy services.

Acknowledgements

We would like to thank the following Speech and Language Therapists and Assistants who helped carry out the therapy: Hannah Leniston, Nataša Marić, Sarah Parkes and Susan Pope. We would like to thank the following student and volunteer Speech and Language Therapists who carried out the 'blind' assessment: Anna Dickens, Nicola Robinson, Ashna Ganatra, Alex Baxter, Ruth Dehaney, Agnes Osborn, Kerry Seabrook and Na'Ama Cohen. Thank you also to Julie Dockrell and David Messer for their helpful comments on earlier drafts of this paper and to all the participants in the study.

Declaration of interest

This study was funded by Moor House School, where all the authors were employed at the time of the study.

Appendix A – Individual data for participants in original therapy group (receiving therapy in Phase 1)

	Participant code	T1	T2	T3	T4	T5	T6	T7	T8
	Gender	male	male	male	male	male	male	female	female
	Age at start of study	14;3	15;2	12;11	15;3	11;10	13;7	13;8	14;2
	SLT(A)(s) delivering WFD therapy	GT	KH	MJ	KH	BE	KM&HL	SE	SPa&SE
	Category used in therapy	animals	animals	animals	clothes	clothes	food	food	animals
Pre-phase 1 therapy	CELF-3 UK expressive language	71	64	64	64	77	75	64	64
	CELF-3 UK receptive language	64	64	76	69	85	86	64	64
	BPVS-II	69	65	82	70	78	80	76	74
	TAWF RS	19	15	18	21	26	25	12	26
	TAWF Prorated SS	66	44	58	57	86	74	44	73
	TAWF naming type	fast	slow	slow	fast	fast	fast	slow	fast
	PhAB semantic fluency SS	94	69	109	103	107	88	69	77
	% T-units with >0 WFDs	27	58	50	53	26	33	45	22
	WFDs/T-unit	0.52	3.29	1.02	0.73	0.38	0.57	0.85	0.22
	% WFDs/ total words	6	43	10	7	4	7	17	3
	clothes in 1 min	20	6	15	16	18	16	10	13
	food in 1 min	25	15	22	19	21	17	7	13
	animals in 1 min	20	12	26	22	21	19	12	21
	mean of 3 categories in 1 min	22	11	21	19	20	17	10	16
Post-phase 1 therapy	TAWF RS	22	24	24	25	23	30	13	28
	TAWF Prorated SS	76	74	71	79	92	86	46	78
	PhAB semantic fluency SS	113	69	99	82	121	94	69	82
	% T-units with >0 WFDs	30	54	43	40	40	53	60	34
	WFDs/T-unit	0.56	0.92	0.66	0.45	0.65	1.00	0.75	0.50
	% WFDs/ total words	4	12	7	4	18	7	14	5
	clothes in 1 min	16	8	14	16	22	16	6	18
	food in 1 min	26	13	19	22	20	15	6	14
	animals in 1 min	24	11	25	22	30	21	10	17
	mean of 3 categories	22	11	19	20	24	17	7	16
Post-phase 2 therapy	TAWF RS	24	22	31	25	24	29	13	29
	TAWF Prorated SS	78	69	89	66	92	81	44	81

Appendix B – Individual data for participants in original waiting control group (receiving therapy in Phase 2)

	Participant code	WC1	WC2	WC3	WC4	WC5	WC6	WC7
	Gender	female	female	male	male	male	male	male
	Age at start of study	11;11	15;11	12;8	13;4	13;8	13;8	9;11
	SLT(A)(s) delivering WFD therapy	MJ	KH	LN	LN	NM&SPo	NM&SPo	HN
	Category used in therapy	clothes	clothes	food	animals	animals	clothes	food
Pre-phase 1 therapy	CELF-3 UK expressive language	64	64	90	77	64	64	64
	CELF-3 UK receptive language	64	64	97	76	64	64	64
	BPVS-II	81	too old for SS	97	89	63	60	85
	TAWF RS	16	28	23	28	21	16	20
	TAWF Prorated SS	54	74	69	81	71	62	too young
	TAWF naming type	slow	fast	equipment failure	fast	data not collected	slow	slow
	PhAB semantic fluency SS	91	94	equipment failure	69	85	105	82
	% T-units with >0 WFDs	44	54		21	29	43	30
	WFDs/T-unit	0.59	1.07	equipment failure	0.36	0.74	1.17	0.37
	% WFDs/ total words	7	10		7	10	39	7
	clothes in 1 min	10	20	9	10	12	13	14
	food in 1 min	14	20	26	16	18	22	16
	animals in 1 min	16	24	26	10	21	9	11
	mean of 3 categories in 1 min	13	21	20	12	17	15	14
Post-phase 1 therapy	TAWF RS	18	26	31	31	18	20	20
	TAWF Prorated SS	58	69	89	89	58	69	too young
	PhAB semantic fluency SS	93	88	not re-done due to previous equipment failure	69	99	equipment failure	69
	% T-units with >0 WFDs	53	55		35	53	53	50
	WFDs/T-unit	0.84	1.04		0.43	0.92	0.83	0.71
	% WFDs/ total words	10	9		5	7	10	11
	clothes in 1 min	11	13	13	9	14	19	17
	food in 1 min	16	24	15	11	15	22	10
	animals in 1 min	21	26	12	12	28	13	9
	mean of 3 categories	16	21	13	11	19	18	12
Post-phase 2 therapy	TAWF RS	19	therapy not completed	32	33	28	22	20
	TAWF Prorated SS	60		92	93	84	71	too young

Appendix C – Therapy steps

Task	Step	Category		
		animals	**clothes**	**food**
categorise pictures	1	by broad semantic categories		
	2	by more detailed semantic categories		
discuss attributes of pictures	3	category		
	4	what does it look like?		
	5	what noise does it make?	how does it feel/what made from?	how does it taste/feel?
	6	where does it live?	who wears it?	where do you find it?
	7	what does it eat?	when do you wear it?	when do you eat it?
compare pictures: similarities and differences	8	in terms of all of the above		
20 questions	9	participant gives semantic cues, SLT guesses picture		
	10	SLT gives semantic cues, participant guesses		
strategy game	11	Throw a dice with semantic/phonological strategies on it, pick up card and use named strategy, if can do this, get a token to place on noughts and crosses board		

Appendix D – Therapy Schedule

Week and session	Steps	Week and session	Steps	Week and session	Steps	Week and session	Steps
1a	1 3	3a	4 3&4 9&10	5a	6 3,4,5,6&7 9&10	7a	8 11
1b	1&2 3 9&10	3b	5 3,4,5 9&10	5b	7 3,4,5,6&7 9&10	7b	8 9&10
2a	1&2 3 9&10	4a	5 3,4,5 9&10	6a	7 3,4,5,6&7 9&10	8a	9&10 11
2b	4 3&4 9&10	4b	6 3,4,5,6 9&10	6b	8 11	8b	9&10 11

Reference List

Beitchman, J. H., Wilson, B., Brownlie, E. B., Walters, H., & Lancee, W. (1996). Long-term consistency in speech/language profiles: 1. developmental and academic outcomes. *Journal of the American Academy of Child and Adolescent Psychiatry, 35,* 804-814.

Best, W. (2005). Investigation of a new intervention for children with word-finding problems. *International Journal of Language & Communication Disorders, 40,* 279-318.

Bishop, D. V. M., Adams, C. V., & Rosen, S. (2006). Resistance of grammatical impairment to computerized comprehension training in children specific and non-specific language impairments. *International Journal of Language & Communication Disorders, 41,* 19-40.

Botting, N., Faragher, B., Simkin, Z., Knox, E., & Conti-Ramsden, G. (2001). Predicting pathways of specific language impairment: What differentiates good and poor outcome? *Journal of Child Psychology and Psychiatry and Allied Disciplines, 42,* 1013-1020.

Casby, M. W. (1992). An intervention approach for naming problems in children. *American Journal of Speech-Language Pathology, 1,* 35-42.

Clegg, J., Hollis, C., Mawhood, L., & Rutter, M. (2005). Developmental language disorders - a follow-up in later adult life. Cognitive, language and psychosocial outcomes. *Journal of Child Psychology and Psychiatry, 46,* 128-149.

Constable, A., Stackhouse, J., & Wells, B. (1997). Developmental word finding difficulties and phonological processing: the case of the missing handcuffs. *Applied Psycholinguistics, 18,* 507-536.

Conti-Ramsden, G. (2008). Heterogeneity of Specific Language Impairment (SLI): outcomes in adolescence. In C.F.Norbury, B. Tomblin, & D. V. M. Bishop (Eds.), *Understanding Developmental Language Disorders.* Hove: Psychology Press.

Dockrell, J. E. & Lindsay, G. (2008). Inclusion versus specialist provision: ideology versus evidence based practice for children with langauge and communication difficulties. In C.F.Norbury, B. Tomblin, & D. V. M. Bishop (Eds.), *Understanding Developmental Language Disorders.* Hove: Psychology Press.

Dockrell, J. E., Lindsay, G., Letchford, B., & Mackie, C. (2006). Educational provision for children with specific speech and language difficulties: Perspectives of speech and language therapist managers. *International Journal of Language and Communication Disorders, 41,* 423-440.

Dockrell, J. E. & Messer, D. (2007). Language profiles and naming in children with word finding difficulties. *Folia Phoniatrica et Logopaedica, 59,* 318-323.

Dockrell, J. E., Messer, D., & George, R. (2001). Patterns of naming objects and actions in children with word finding difficulties. *Language and Cognitive Processes, 16,* 261-286.

Dockrell, J. E., Messer, D., George, R., & 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., & Wilson, G. (1998). Children with word-finding difficulties - prevalence, presentation and naming problems. *International Journal of Language & Communication Disorders, 33,* 445-454.

Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). The British Picture Vocabulary Scale, Second Edition. BPVS II. Windsor, NFER-NELSON.

Easton, C., Sheach, S., & Easton, S. (1997). Teaching vocabulary to children with wordfinding difficulties using a combined semantic and phonological approach: an efficacy study. *Child Language Teaching and Therapy, 13,* 125-142.

Ebbels, S. & van der Lely, H. (2001). Meta-syntactic therapy using visual coding for children with severe persistent SLI. *International Journal of Language & Communication Disorders, 36,* 345-350.

Ebbels, S. H. (2007). Teaching grammar to school-aged children with Specific Language Impairment using Shape Coding. *Child Language Teaching and Therapy, 23,* 67-93.

Ebbels, S. H., van der Lely, H. K. J., & Dockrell, J. E. (2007). Intervention for verb argument structure in children with persistent SLI: a randomized control trial. *Journal of Speech Language and Hearing Research, 50,* 1330-1349.

Faust, M., Dimitrovsky, L., & Davidi, S. (1997). Naming difficulties in language-disabled children: Preliminary findings with the application of the tip-of-the-tongue paradigm. *Journal of Speech Language and Hearing Research, 40,* 1026-1036.

Franklin, I. (1994). ColorLibrary COLOURCARDS. Speechmark Ltd.

Frederickson, N., Frith, U., & Reason, R. (1997). The Phonological Awareness Battery. Windsor, UK, NFER-Nelson.

German, D. I. & 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. (1986). Test of Word Finding - TWF. DLM Teaching Resources.

German, D. J. (1990). Test of Adolescent/Adult Word Finding - TAWF. DLM Teaching Resources.

German, D. J. (1991). Test of Word Finding in Discourse - TWFD. DLM Teaching Resources.

German, D. J. (2002). A phonologically based strategy to improve word-finding abilities in children. *Communication Disorders Quarterly, 23,* 117-190.

German, D. J. & Simon, E. (1991). Analysis of Childrens Word-Finding Skills in Discourse. *Journal of Speech and Hearing Research, 34,* 309-316.

Hasbro. (2004). Connect 4. MB Games.

Howlin, P., Mawhood, L., & Rutter, M. (2000). Autism and Developmental Receptive Language Disorder: a Follow-up Comparison in Early Adult Life. II: Social, Behavioural, and Psychiatric Outcomes. *Journal of Child Psychology and Psychiatry, 41,* 561-578.

Lahey, M. & Edwards, J. (1999). Naming errors of children with specific language impairment. *Journal of Speech Language and Hearing Research, 42,* 195-205.

Marks, I. & Stokes, S. F. (2010). Narrative-based intervention for word-finding difficulties: a case study. *International Journal of Language & Communication Disorders, 45,* 586-599.

Mawhood, L., Howlin, P., & Rutter, M. (2000). Autism and developmental receptive language disorder - a comparative follow-up in early adult life. I: cognitive and language outcomes. *Journal of Child Psychology and Psychiatry, 41,* 547-559.

McGregor, K. K., Friedman, R. M., Reilly, R. M., & Newman, R. M. (2002). Semantic representation and naming in young children. *Journal of Speech, Language and Hearing Research, 45,* 332-346.

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. (1997). The nature of word-finding errors of preschoolers with and without word-finding deficits. *Journal of Speech Language and Hearing Research, 40,* 1232-1244.

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

McGregor, K. K., Newman, R. M., Reilly, R. M., & Capone, N. C. (2002). Semantic representation and naming in children with specific language impairment. *Journal of Speech Language and Hearing Research, 45,* 998-1015.

McGregor, K. K. & Waxman, S. R. (1998). Object naming at multiple hierarchical levels: a comparison of preschoolers with and without word-finding deficits. *Journal of Child Language, 25,* 419-430.

McGregor, K. K. & Windsor, J. (1996). Effects of priming on the naming accuracy of preschoolers with word-finding deficits. *Journal of Speech and Hearing Research, 39,* 1048-1058.

Nicoll, H. & Ebbels, S. (in prep, 2011). The effectiveness of different methods of delivery of semantic therapy for word-finding difficulties.

Semel, E., Wiig, E. H., & Secord, W. A. (1995). Clinical Evaluation of Language Fundamentals-3. San Antonio, The Psychological Corporation.

Sheng, L. & McGregor, K. K. (2010). Lexical-Semantic Organization in Children With Specific Language Impairment. *Journal of Speech Language and Hearing Research, 53,* 146-159.

Stiegler, L. N. & Hoffman, P. R. (2001). Discourse-based intervention for word finding in children. *Journal of Communication Disorders, 34,* 277-303.

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 in Schools, 21,* 151-156.

Wittman, S. (1996). A case study in word finding. *Child Language Teaching & Therapy, 12,* 300-313.

Wright, S. H. (1993). Teaching word-finding strategies to severely language-impaired children. *European Journal of Disorders of Communication, 28,* 165-175.

Wright, S. H., Gorrie, B., Haynes, C., & Shipman, A. (1993). What's in a name? Comparative therapy for word-finding difficulties using semantic and phonological approaches. *Child Language Teaching and Therapy, 9,* 214-229.

Table 1: *pre-therapy standard scores on standardised tests*

	Mean (n=15)	Range	Mean (n=14)[a]	Range [a]
CELF-3 expressive language	69	64-90	67	64-77
CELF-3 receptive language	71	64-97	69	64-86
BPVS	76	60-97	75	60-89
TAWF SS	63	44-81		
TAWF SS (prorated for comprehension)	66	44-86		
PhAB semantic fluency	87	69-109		
TWFD (%T-units with at least 1 WFD)	38%	21%-58%		

[a]= excluding participant with much higher language scores

Table 2: *Mean (SD) scores on Test of Word Finding in Discourse*

		Therapy group (n=8)	Waiting Controls (n=6)
% T-units containing one or more WFD	Pre-therapy	39 *(14)*	37 *(12)*
	Post-phase1 therapy	44 *(10)*	50 *(7)*
	Change	5 *(12)*	13 (*8)*
mean number of WFDs per T-unit	Pre-therapy	0.95 (*0.98)*	0.72 *(0.34)*
	Post-phase1 therapy	0.69 *(0.19)*	0.80 *(0.21)*
	Change	-0.26 *(0.90)*	0.08 *(0.24)*
% WFDs / total words	Pre-therapy	12.26 *(13.06)*	13.11 *(12.81)*
	Post-phase1 therapy	8.85 *(5.12)*	8.57 *(2.43)*
	Change	-3.4 *(12.5)*	-4.5 *(12.3)*

Table 3: *Mean (SD) scores on Semantic Fluency test*

		Therapy group	Waiting Controls
		(n=8)	(n=7)
Mean number (across 3 categories) of words listed in 1 minute	Pre-therapy	16.9 *(4.5)*	16.0 *(3.6)*
	Post-phase1 therapy	17.1 *(5.6)*	15.7 *(3.9)*
	Change	0.2 *(1.9)*	-0.3 *(3.5)*
		(n=8)	(n=5)
Standard Score on PhAB semantic fluency subtest (all participants)	Pre-therapy	89.5 *(16.4)*	84.2 *(9.7)*
	Post-phase1 therapy	91.1 *(19.2)*	83.6 *(13.9)*
	Change	1.6 *(12.7)*	-0.6 *(10.0)*
		(n=6)	(n=4)
Standard Score on PhAB semantic fluency subtest (excluding oldest participants)	Pre-therapy	90.7 *(16.0)*	81.8 *(9.3)*
	Post-phase1 therapy	96.3 *(19.2)*	82.5 *(15.8)*
	Change	5.7 *(10.3)*	0.8 *(11.1)*

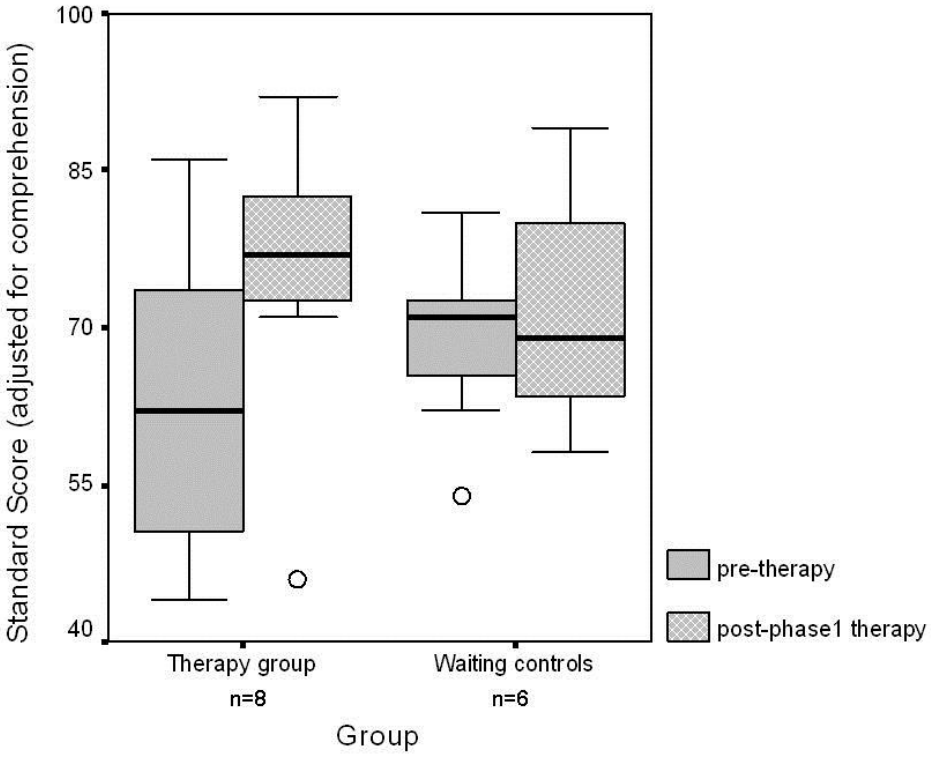

Figure 1: TAWF Standard Score (adjusted for comprehension, as per test manual) pre- and post-phase1 therapy

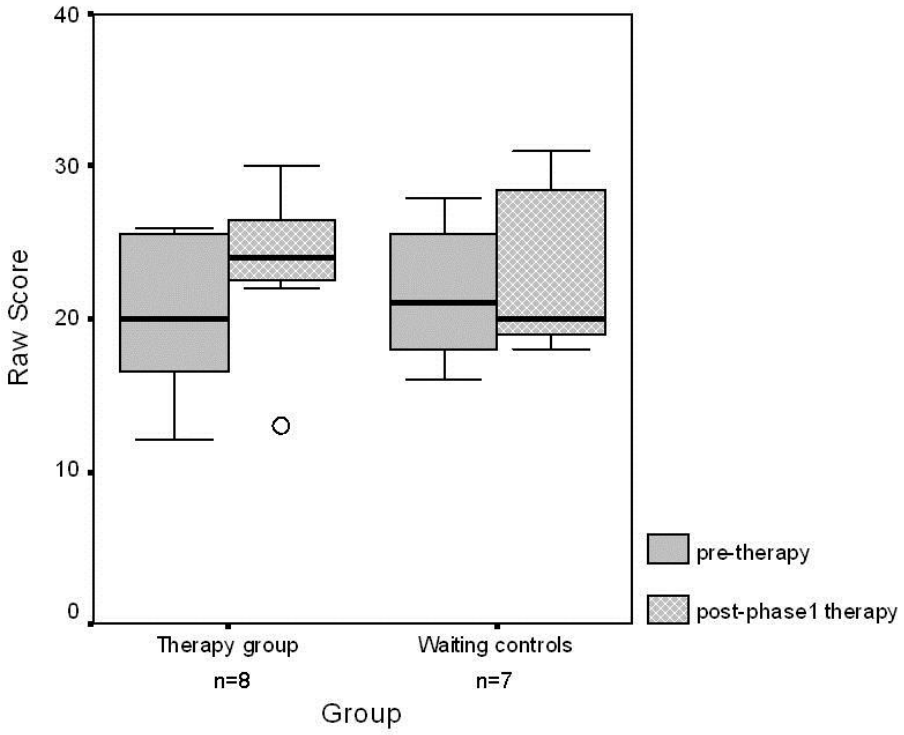

Figure 2: TAWF Raw Score pre- and post-phase1 therapy

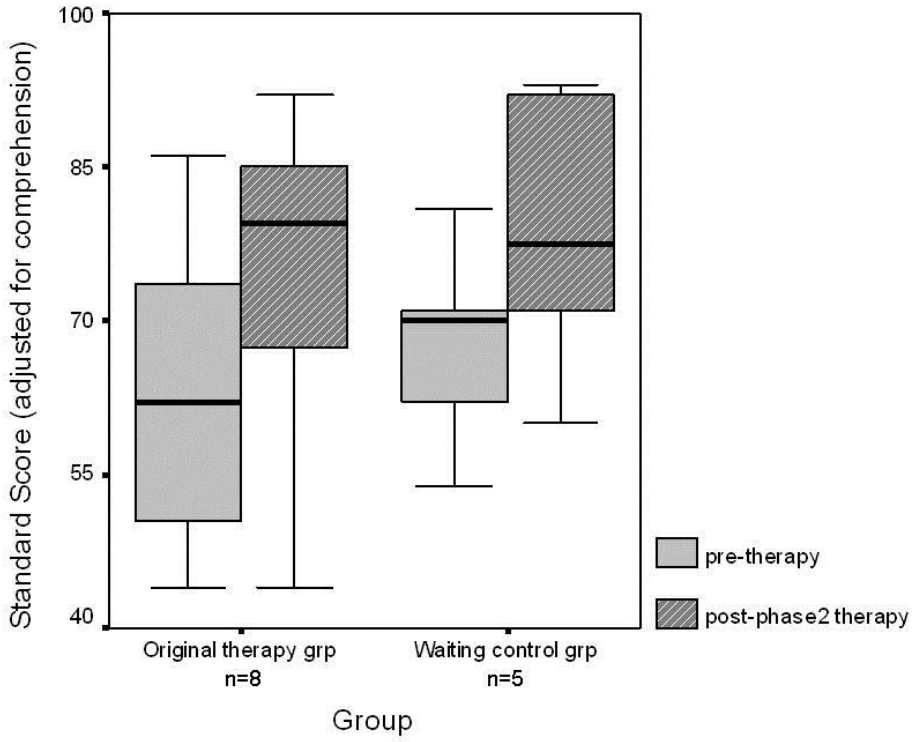

Figure 3: TAWF Standard Score (adjusted for comprehension, as per test manual) pre- and post-phase2 therapy

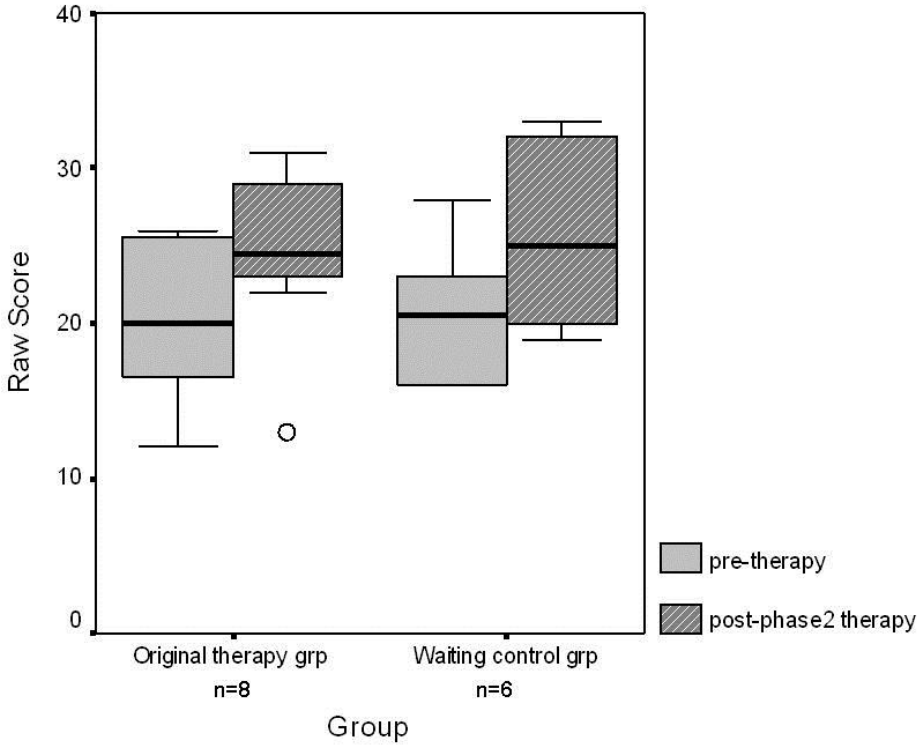

Figure 4: TAWF Raw Score pre- and post-phase2 therapy