Understanding differing outcomes from semantic and phonological interventions with children with word-finding difficulties: a group and case series study

Wendy Best[*1], Lucy Hughes[a], Jackie Masterson[b], Michael S. C. Thomas[c], David Howard[d], Anna Kapikian[b] & Kate Shobbrook[a]

[a] Division of Psychology and Language Sciences, University College London, London, UK
w.best@ucl.ac.uk
[b] Department of Psychology and Human Development, UCL Institute of Education, London, UK
j.masterson@ucl.ac.uk
[c] Department of Psychological Sciences, Birkbeck College London, London, UK
m.thomas@bbk@ac.uk
[d] School for Education, Communication and Language Sciences, Newcastle University, London, UK
david.howard@ncl.ac.uk

* Corresponding author: Division of Psychology and Language Sciences, University College London, Chandler House, 2 Wakefield Street, London WC1N 1PF, UK. Email address: w.best@ucl.ac.uk

Abstract

Developmental Language Disorder occurs in up to 10% of children and many of these children have difficulty retrieving words in their receptive vocabulary. Such word-finding difficulties (WFD) can impact social development and educational outcomes. This research aims to develop the evidence-base for supporting children with WFD and inform the design and analysis of intervention studies.

We included 20 children (age 6 to 8) with WFD each of whom participated in two interventions one targeting semantic attributes and the other phonological attributes of target words. The interventions, employing word-webs, were carefully constructed to facilitate direct comparison of outcome which was analysed at both group and case-series level.

The study used a robust crossover design with pre-intervention baseline, between–intervention wash-out and post-intervention follow-up testing. We incorporated: matching of item sets on individual performance at baseline, independent randomisation of order of intervention and items to condition, blinding of assessor, evaluation of fidelity and control items. The interventions were clinically feasible, with weekly sessions over six weeks.

Intervention improved children's word-finding abilities with statistically significant change only during treatment phases of the study and not over baseline, wash-out or follow-up phases. For the group the semantic intervention resulted in a gain of almost twice as many items as the phonological intervention, a significant difference. However, children differed in their response to intervention. Importantly, case-series analysis revealed outcomes predictable on the basis of children's theoretically driven language profiles. Taking account of individual profiles in determining choice of intervention would enable more children to benefit.

The study provides new evidence to inform and refine clinical practice with this population. Future studies should be designed such that results can be analysed at both group and case series levels to extend theoretical understanding and optimise use of appropriate interventions.

Keywords Word-finding, intervention, semantic, phonological, developmental language disorder

Abbreviations

WFD – word finding difficulties, DLD – developmental language disorder, Pre – assessments carried out before intervention, Post – assessments after intervention.

Funding sources

This work was supported by the ESRC [grant number RES-062-23-2721, UKRI grant ref. ES/H046925/1].

1. Introduction

The aim of this paper is to report findings from a study in which children with WFD participated in two different interventions, and also to explore the value of presenting and analysing data from a case series as well as at the group level. We first highlight the need and rationale for the intervention study (1.1) and then consider the research design (1.2). Finally, we provide a theoretical model in the domain of the study, word production, within which findings are interpreted and clinical implications determined (1.3) and outline the resulting research questions (1.4).

1.1 Intervention study

1.1.1 Prevalence and impact

Low language (≥1.25 SD below the mean) is found in up to a fifth of 7 year olds and is associated with concurrent literacy and school difficulties (McKean et al., 2017). If a child has a language difficulty on school entry it is likely to persist (Stothard & Snowling, 1998). An international consensus study agreed on the term 'developmental language disorder' (DLD) for children "who are likely to have language problems enduring into middle childhood and beyond, with a significant impact on everyday social interactions or educational progress" (p.1070, Bishop, Snowling, Thompson, & Greenhalgh, 2016). Children with DLD have heterogeneous profiles, yet within this variety sub-groups with particular areas of language strength and need can be identified. One such group is children with word-finding difficulties (WFD), which occur when a child is unable to produce words despite having an understanding of their meaning. WFDs occur in up to a quarter of children attending language support services (Dockrell, Messer, George, & Wilson, 1998) and have implications for education and social development (e.g. Messer, Dockrell, & Murphy, 2004). Thus, language difficulties are prevalent in childhood and word-finding difficulties (WFD) occur in many children with DLD. Given the impact of WFD on development it is crucial to investigate interventions for children with this difficulty.

1.1.2 Interventions with children with WFD

There is a growing evidence base demonstrating that interventions for children with WFD can be effective in improving retrieval for treated words. Research studies tend to provide intervention focusing either on semantics, exploring the meaning of treated items (e.g. Ebbels et al., 2012), or on phonology/orthography, exploring an item's spoken or written form (e.g. McGregor, 1994). A range of activities have been used, including categorisation tasks, odd one out tasks and word-webs where features of meaning or form are requested or provided (e.g. semantic category, first sound, etc.).

Due to heterogeneity in the nature of WFD (see section 1.3), it is important to determine which approach works best for which child/ren. There are very few studies which compare interventions directly, but Bragard and colleagues' 2012 case series study with four participants with WFD is an exception. The results showed long-lasting effects of intervention with a pattern of response that could be related to the nature of the children's difficulties. Clinical practice is likely to be informed by such studies. In the meantime, in our experience (which combines to over half a century of working in the UK with children with DLD and training students to do the same) clinical practice is eclectic, with many practitioners using approaches including word-webs (the word-webs used in this study are described in detail in section 2.4) which entail information about both meaning and form of targeted items. In sum, whilst the research evidence is encouraging there is very little that contrasts different approaches directly or which relates outcome to the nature of children's language profiles.

1.2 Design

Research designs in the field of intervention for DLD vary. They include experimentally controlled single case and case-series studies, and randomised controlled trials (RCTs). In

each case there are debates over the optimum way of exercising experimental control and analysing the research findings (Gold et al., 2017; Tate, Rosenkoetter, Wakim, Sigmundsdottir, Doubleday, Togher & Perdices, 2015). Given the lack of consensus, it is important to compare findings from different designs and to explore the value of the different approaches.

The group RCT is held to be the 'gold standard' in medically related research, and increasingly in education research. This has a profound influence on the likelihood of obtaining research funding. Group studies with randomisation to different treatment and control conditions have many strengths, including generalisation of findings to others meeting the inclusion criteria and enabling combination of effect sizes across studies in systematic reviews.

Systematic reviews and 'n-of-1 trials' share the 'top-spot' in the Oxford Centre for Evidence Based Medicine's hierarchy of evidence in order to answer the question 'does this intervention help?' 'N of 1 trials' refer to a variety of designs which investigate the effect of treatment with one individual (Shamseer et al., 2015). A key point highlighted by these authors is that findings for a group may not be true for all the individuals within the group. In order for us to determine progress over time compared with progress with an intervention at an individual level, more than one (and preferably several) pre-intervention baselines are required (Howard, Best & Nickels, 2015).

There are differences in what can be concluded from these contrasting research designs. This is important because discrepancies in the interpretation of findings from group and single case studies may result in different theoretical and practical implications. Investigating outcome at these different levels *within the same intervention study* is novel and likely to have important implications for the design of future studies.

1.3 Theoretical background

1.3.1 Models of production and stages of processing

Within the vast literature on language production there is agreement that there are at least three levels of processing necessary for single word production: semantic, lexical and phonological (e.g. Dell, Schwartz, Martin, Saffran & Gagnon, 1997). This theoretical framework is discussed in detail in relation to WFD in children in Best (2005) and (Friedmann, Biran & Dotan, 2013). During the widely used task of picture naming, for example, there will be activation of semantic representations. This will, in turn, activate representations in a phonological output lexicon. There is then a post-lexical stage of processing during which phonological forms are assembled for production. In the present research, we employ the Restricted Interaction Account (RIA, Goldrick & Rapp, 2002) in which activation is largely feed-forward but there is also feedback from phonological representations to the output lexicon. In summary, within a child's developing language system there are several levels of processing necessary for word-retrieval and production, each of which may have developed typically or atypically.

1.3.2 Sub-types of word-finding difficulty (WFD)?

The results of the two different interventions employed in the present study are interpreted in relation to three potential sub-types of difficulties in word-retrieval, with the recognition that individual children may experience difficulty with multiple levels of processing; here we aim to identify the principle level of difficulty for each child. It is important at this point to note that children were not recruited to the current study on the basis of having different profiles; instead these were identified and used in analyses after the intervention study in order to look for potential explanations of the different outcomes. We consider three sub-types: (i) classic WFD, (ii) semantic difficulties and (iii) phonological difficulties (in the context of strong semantic processing).

Classic WFD (sub-type i) occur when children have identified difficulties in word-finding despite relatively good semantic processing and good phonological processing. In this case the difficulty is in accessing the phonological form for production (see, e.g. Best, 2005).

WFDs may reflect less elaborate semantic representations (sub-type ii). For example, McGregor and colleagues investigated the drawings and definitions provided by children with DLD and found less detail was provided for items for which children made semantic errors in naming (McGregor & Appel, 2002; McGregor, Newman, Reilly & Capone 2002). In line with this, some children with WFDs show relative strengths in phonological awareness and decoding whilst demonstrating difficulty with reading comprehension and semantic fluency tasks (Messer & Dockrell, 2013; Messer et al., 2004).

Finally, for other children, WFDs appear to stem from limitations with phonological processing, i.e., at a later stage within the language production process (sub-type iii). For example, Constable, Stackhouse and Wells (1997) used a psycholinguistic framework approach and present a careful single case study of a seven-year-old who had no observable semantic difficulties but pervasive difficulties with tasks requiring phonological processing. German and Newman (2004) found the naming errors of children with WFDs were influenced by the post-semantic psycholinguistic variables of word frequency and phonological neighbourhood density, providing further support for a later-stage account.

1.3.3 Predictions from different interventions according to level of difficulty

While there is an increasing evidence base for interventions for WFDs and this spans semantic (e.g. Ebbels et al., 2012) and phonological/orthographic approaches (e.g. Best, 2005; McGregor, 1994), we remain unclear about which approach is appropriate for an individual child. This is particularly important in the context of limited resources and the impact WFD can have on everyday communication and education. Group studies have produced conflicting results (Wing, 1990; Wright, Gorrie, Haynes & Shipman, 1993).

Bragard and colleagues (2012) were able to provide some evidence that children may respond differently according to their background language profile. They reported that, of the four children in their study, the two with more semantically based difficulties responded better to phonological intervention, whereas one participant with apparently more phonologically based WFDs responded more favourably to semantic intervention. However, this study involved only four children, one of whom did not show a positive response to intervention of either type. We consequently have little direct evidence on the optimal intervention approach for a child given their language profile.

We can use models of language production to support the development of theoretically motivated predictions as to the outcome of intervention given the strengths and difficulties of an individual child. The current study, termed 'WORD' (WOrd Retrieval and Development), is the first to include a large number of children with WFD, each of whom participated in two types of intervention. It therefore provides us with a unique opportunity to explore how the outcome of different interventions relates to language profiles.

1.3.4 Order of interventions

If a child receives multiple interventions, there is the possibility that the order in which the interventions are deployed will have an effect. There is limited evidence regarding whether the order in which semantic and phonological interventions occur can influence outcome. Zens, Gillon and Moran (2009) reported that for new *word-learning*, children with language impairment who participated in a phonological awareness intervention followed by semantic intervention showed improved word production of the new forms while the reverse order did not result in improved performance at post-test, and neither order impacted on the comprehension of new words (Zens et al., 2009). Since models of word production describe the process as involving the use of semantic information to access phonological forms, it would seem that there is sense in predicting the opposite result i.e. that semantic intervention preceding phonological intervention may be more effective than the reverse. In

conclusion, while order effects may be present, there is no strong evidence on the direction these may take. If effects are present there will be clinical implications for the order of treatment approaches.

1.4 Research questions

The overarching research question addressed by the study is whether for children with WFDs, naming skills will improve with WORD intervention. The intervention is described in detail in 2.4. A second key question is whether, when the results are considered at different levels (group and case series), the conclusions differ.

Hypotheses are divided into sections: hypotheses addressed by data from the group (1.4.1) and case series (1.4.2), hypotheses addressed by considering the children in theoretically informed sub-groups (1.4.3) and a hypothesis pertaining to the design and level of analysis (1.4.4).

1.4.1 Hypotheses for group

i) children's naming will improve as a result of the WORD intervention, specifically there will be significant improvement over the intervention phases of the study and not over the non-intervention phases

ii) there will be a difference between the effect of the two interventions (semantic and phonological)

iii) the order in which the semantic and phonological interventions occur will influence the outcome

iv) the children will differ from one another in their response to intervention

1.4.2 Hypotheses for case series

i) individual children will show greater change during therapy phases for treated items (n=50)

ii) individual children will show greater change during the semantic therapy phase of the study for the sub-sets of items given semantic therapy (n=25) .

iii) individual children will show greater change during the phonological therapy phase of the study for the subsets of items given phonological therapy (n=25).

1.4.3 Hypotheses with children grouped according to background assessments

Predictions for the outcomes for children with different language profiles will be outlined in turn and are summarised in Figure 1.

(i) classic WFD

Children with this profile are predicted to benefit from the semantic or phonological intervention or both. By definition their difficulty is in accessing phonology from meaning and providing and processing either additional semantic or phonological information may aid their word retrieval.

(ii) semantic difficulties

Children with this profile are predicted to benefit from semantic intervention but not to benefit from phonological intervention. This is because they may not have adequate information about an item's meaning to drive retrieval, therefore, the provision and processing of semantic information is likely to aid retrieval of the word form. In contrast, information about phonology will not help if this later stage of processing is not activated for hard to retrieve items.

(iii) phonological difficulties (in the context of strong semantic processing)

Children whose profiles show only phonological difficulties are predicted to benefit from phonological intervention but not to benefit from semantic intervention. This is because, by definition, they have relatively good knowledge of word-meanings, in their case benefit is likely to derive from providing and processing additional phonological information.

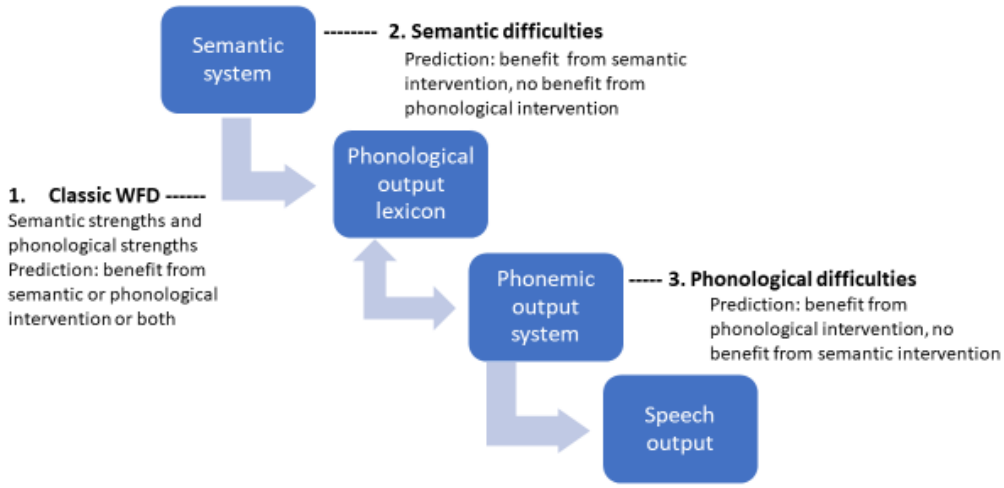


Figure 1: Predictions of response to intervention in relation to a model of spoken word production (modified from Friedmann et al, 2013). Note that all children have naming difficulties relative to their comprehension of the same items.

1.4.4 Hypotheses for analysing outcomes for different levels of design

(i) The results from the group analysis will provide an understanding of outcomes on average which may be generalised to other children with WFD meeting the inclusion criteria.

(ii) Consideration of heterogeneity in individual children's profiles and outcomes may enable more specific theoretically and clinically informative conclusions.

2. Materials and methods

We report all measures in the study. how we determined our sample size, all inclusion/exclusion criteria (which were established prior to data analysis) and all manipulations.

2.1 Materials

A range of standardised assessments were used to obtain a broad picture of children's word-finding ability (Test of Word Finding 2, German, 2000), in relation to their scores on verbal (Clinical Evaluation of Language Fundamentals 4, 4th Edition, Semel, Wiig, & Secord, 2003; Test for the Reception of Grammar, Bishop, 1996) and non-verbal tasks (British Ability Scales Edition II pattern construction sub-test, Elliot, Smith, & McCullouch, 1996). Legal copyright restrictions prevent public archiving of the above assessments used in this research, the materials can be obtained from the copyright holders in the cited references

The primary outcome measure for the intervention was picture naming of a set of 100 experimental items. These were pictures that had been previously employed in the children's picture naming studies of Funnell, Hughes and Woodcock (2006) and Masterson, Druks and Galliene (2000). The stimuli from Funnell et al. were obtained by personal communication, those from Masterson et al., (2000) were items published in the Object and Action Naming Battery (Druks and Masterson, 2000). The items had been devised in the two studies to be familiar to young children and had been used with children aged three to eleven years. Example items were: scorpion, courgette, whisk, sledge and submarine. They were presented as line drawings and shown individually on a 15" screen laptop computer using the software DMDX (Forster & Forster, 2003). Four fixed randomised orders were rotated across children for administration. Trials began with a fixation point for 500 msecs then a picture appeared on the screen for a maximum of 10000 msecs. If the child named the picture within the timeout period the tester moved to the next trial with a key press. Naming responses were recorded using an external microphone. Children were asked to provide a single word for each picture. Responses were noted at the time of testing and were checked later from the recording.

Following the naming assessment, four subsets of 25 pictures each were devised for each child. Items were matched across the four subsets for pre-intervention naming accuracy and, as closely as possible for the psycholinguistic variables spoken frequency, imageability, visual complexity, number of phonological neighbours and word-length in phonemes. The sub-sets were created following the naming assessment for each child by the fifth co-author at a separate institution (see section 2.2 for randomisation). The sub-sets were allocated on a random basis to four conditions: (i) items treated with semantic intervention, (ii) items treated with phonological intervention, (iii) untreated named items (these items were presented for naming at the start of each intervention session), (iv) untreated unseen items (these were named only on the occasions when the full 100-item picture set was assessed and they were not seen during the intervention sessions).

Additionally, in order to maximise functional impact, each child had a set of personally chosen items that were named immediately before and after the end of both interventions. The results for these personally chosen items fall outside the scope of this report which focuses on fully experimentally controlled items (but see Best et al., 2017).

The results from two further assessments were used in order to interpret the outcome of intervention in relation to sub-types of WFD. The first of these, the Children's Test of Nonword Repetition (Gathercole & Baddeley, 1996), provides a measure of children's phonological processing skills. The second task assessed semantic processing using picture judgements (PJs). In this task children selected one of two pictures of co-ordinate items from the same semantic category (for example *bed* and *chair*) to go with a picture of a semantically associated (e.g. *pyjamas*, full details of the task are provided in Best et al., 2015 and section 3.3. below details the method by which sub-groups were created using these tasks.).

2.2 Design and randomisation

The study was designed in order to exert adequate experimental control for meaningful outcomes from both a RCT and a case-series analysis (Figure 2). Each child participated in both the semantic and the phonological intervention. The RCT entailed a crossover design with wash-out between interventions and follow-up post-intervention. Each phase of the study lasted approximately six weeks (half a school term). In order for the findings to be meaningful at a case-series level, each child was also assessed three times before the start of intervention. Prior to the study reported here the same children had been randomly assigned to 'wait' versus 'intervention' groups and the results are reported in Best et al. (2017). In the current paper we report the effectiveness of the two WORD interventions (one focusing on semantics and the other on phonology) and include the data from all the children gathered over the full duration of the study[2].

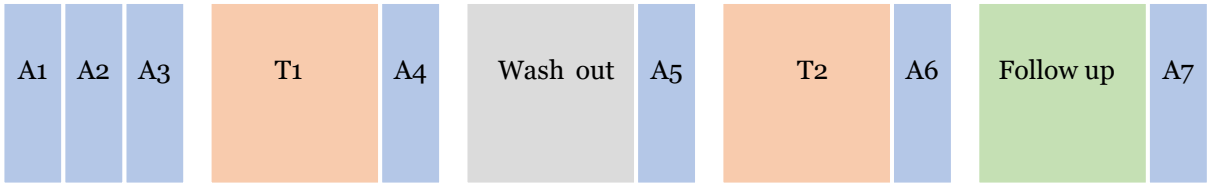


Figure 2: study design

Key: A: Assessment; T: Therapy phase

Random assignment was carried out by author DH, based at a different institution from those carrying out the intervention and assessments with the children. The randomisation was revealed only to the first two authors. The order of the two interventions was random as was the allocation of each subset of naming items to the four conditions described under Materials (2.1), that is, items treated with semantic intervention, items treated with phonological intervention, untreated named items, and untreated unseen items. No part of the study procedure or analysis was pre-registered. In order to avoid difficulties with

[2] Here the data for those assigned to wait at the start of the previously reported study (Best et al., 2017) is aligned with that from those assigned to intervention immediately so there are three pre-therapy assessments for all children.

interpretation that can occur due to regression to the mean (e.g. see Howard et al., 2015), we did not select items that were difficult to name.

2.3 Participants

Twenty children between 6 and 8 years of age participated in the study. They were all attending mainstream schools and were referred as having word-finding difficulties. An N of 20 was chosen based on an *a priori* power analysis (G*Power 3.1; Faul et al., 2007) assuming a within-participant effect size of the naming intervention of at least d=0.7 based on previous studies involving children with word-finding difficulties (Best, 2005; Bragard et al., 2012; Campbell, Nicoll & Ebbels, 2019; Ebbels et al., 2012; Wilson et al., 2015; Zens, et al., 2009), with an alpha level of 0.05, and power of 0.8 (2-tailed).The first two children to complete the intervention are described in a paper which uses computational modelling to predict the outcome of intervention (Best et al., 2015). Inclusion criteria, which were retained throughout the study, were as follows: (i) Word-finding quotient of below 90, and comprehension within the normal range (63+), as measured by the Test of Word-finding 2 (TWF2, German, 2000), to ensure WFD (ii) T score above 35 for BAS Pattern Construction Task (BAS II, Elliot, Smith & McCullouch, 1996) (iii) score on TAPS Word Discrimination subtest (TAPS3, Martin & Brownell, 2005) falls above scaled score of 6. This task was used to ensure children did not have difficulties with auditory discrimination which might impact their ability to respond to intervention. All children had either been exposed to English at home from birth OR had been in an English-speaking nursery since age 3 and were currently exposed to English at home. The children had no significant developmental diagnosis (i.e. autism, dyspraxia, severe stammer), sensory or behavioural difficulties. Finally, once the children had met the criteria above, recordings of their everyday conversation were considered by the first and second authors (both experienced SLPs) in order to establish whether there was clear evidence of word-finding difficulties in discourse.

Individual scores on language and non-verbal assessment are shown in Table 1. The group had average performance on the BAS pattern construction task but performance below average the TROG and CELF-4. Unsurprisingly the group performed poorly on the TWF2. When the children are considered individually, it is clear that their language profiles vary considerably. Some appear to have relatively isolated WFD while others show difficulties across the language assessments employed in the study reflecting the anticipated heterogeneity of WFD.

Child	**Gender**	**Age (months)**	**TWF-2 naming Standard Score**	**TWF-2 Comp Raw Score**	**CELF-4 Core Lang. Standard Score**	**TROG Standard Score**	**BAS non-verbal T score**	***CN Rep Standard Score***	***PJs z score***
1	M	82	62	66	75	85	47	*72*	*0.52*
2	F	81	62	64	56	82	40	*†*	*2.48*
3	M	104	75	68	67	80	47	*<46*	*0.45*
4	M	99	60	63	81	80	50	*75*	*2.98*
5	F	89	68	64	72	77	57	*<50*	*0.82*
6	M	76	68	64	96	87	45	*62*	*1.46*
7	F	92	79	69	85	80	44	*63*	*-0.52*
8	F	96	75	67	84	90	36	*70*	*1.91*
9	M	88	62	67	81	71	49	*83*	*-0.27*
10	F	86	86	65	88	87	47	*78*	*0.83*
11	F	93	60	68	81	94	42	*51*	*0.75*
12	F	93	60	65	60	86	44	*52*	*1.58*
13	M	90	60	61	56	72	44	*70*	*0.79*
14	M	86	82	67	97	111	53	*78*	*0.38*
15	M	89	75	69	97	104	53	*83*	*0.09*
16	M	78	65	64	67	85	41	*75*	*1.94*
17	M	96	78	68	75	68	54	*<46*	*0.42*
18	M	87	70	64	93	104	46	*83*	*0.62*
19	F	78	74	68	78	91	47	*78*	*0.20*
20	M	84	62	64	88	90	51	*73*	*0.60*
Mean	-	88.35	69.15	65.75	78.85	86.2	46.85	*71.63*	*0.90*
SD	-	7.36	8.31	2.24	12.89	11.11	5.19	*10.06*	*0.90*

Table 1: Participant data

Key: F: female, M: male, CELF-4: Clinical Evaluation of Language Fundamentals, 4th Edition (Semel, Wiig, & Secord, 2003), BAS Non-Verbal: pattern construction subtest from the British Ability Scales Edition II (Elliot, Smith, & McCullouch, 1996), TWF-2: Test of Word Finding (German, 2000), TROG: Test for the Reception of Grammar (Bishop, 1996). *Scores on tests in italics in the two right-hand columns were used later in the study to*

sub-group the children by type of WFD in order to understand the outcome of intervention CN Rep: Children's Test of Nonword Repetition (Gathercole & Baddeley, 1996); PJs: Picture Judgment task (score relative to typically developing children, see section 3.3). Children are ordered by date of referral to WORD to enable transparency on whether, for example, changes emerged in the profile of those recruited over the course of the study. *† Child did not complete this task.*

2.4 Intervention

The intervention, using word-webs and named '*WO*rd *R*etrieval and *D*evelopment' (WORD) was carried out on a 1:1 basis. The rationale for the intervention and full protocol are available at: https://www.ucl.ac.uk/pals/speech-language-intervention-research/resources/word-therapy-guide. Sessions took place once a week for 45 minutes (approximately 10 minutes of assessment, five minutes activity whilst the therapist selected un-named items, and up to 30 minutes intervention).

The semantic and phonological interventions share many core components. In each case a word-web was used to help the child retrieve the item. A child was asked for features of the word which were added to hexagons around a central picture of the item in a set order. If the child was unable to generate features a choice was provided. In the semantic intervention children were asked to generate semantic features: category, appearance, location, use, action and linked words (features were included only when appropriate to the target e.g. animals may not have a 'use'). In the phonological intervention children were asked to generate phonological (and one orthographic) features: first sound, things that start with the same sound, number of syllables, things that rhyme, first letter and ways to break down the word. Once all features had been generated or chosen they were reviewed and the child was asked to say the word or, if necessary, encouraged to repeat it.

A key aim of the intervention was to develop metacognitive awareness by encouraging the child to reflect on which aspects of word webs were most helpful to them. Later intervention sessions incorporated games to and a review of the most useful strategies learnt during

therapy. Children were encouraged to take these forward themselves after the project with, for example, self-generated cue cards to remind themselves of the features most helpful in aiding their own word-retrieval.

A fidelity check was carried out on 10% of the intervention sessions selected at random from the 240 available (20 children, 6 sessions, 2 interventions) by an undergraduate student independent of the study.

3. Theory/calculation

3.1 Group analysis

A mixed ANOVA was carried out with one within-subjects variable, TIME, with 7 levels: assessment from A1 to A7, and one between-subjects variable, GROUP, with 2 levels according to which of the two interventions happened first. Significant main effects were followed up with pairwise comparisons. A homogeneity test was employed to determine whether the individual children responded in the same way to the interventions.

3.2 Case series analysis

Analysis of the case series data employed established item-by-item statistics in which autocorrelation is avoided by weighting each data point according to hypotheses and summing to zero (Howard et al., 2015). We investigated whether there was greater change during intervention than non-intervention phases of the study for each child for all items and separately for the sub-sets of items given each type of intervention. Furthermore, we investigated whether there was significant change for the sub-set of items given each intervention by simply comparing performance before and after each intervention for each sub-set using the McNemar test.

3.3 Sub-grouping from assessments

After the interventions, in order to relate the outcome of the two different interventions to individual children's background profiles, performance on a semantic task and a phonological task was used to divide the 20 children into three sub-types according to the children's primary area of difficulty. To do this a z-score was calculated for the PJs (semantic picture judgement) task, taking into account both accuracy and RT in relation to scores in the PJs task for 100 typically developing children (see Best et al., 2015 for details[3]). For phonological abilities, z-scores for accuracy in the CNRep were used. In the CNRep task a child has to process, hold and output novel phonological forms. A cut-off of -1.25 standard deviations from the mean on each task was employed. This resulted in three theoretically driven sub-groups which align with those in Figure 1: (i) classic WFD, children diagnosed with WFD but without major difficulties in either semantic or phonological processing, (ii) semantic difficulties and (iii) phonological difficulties (in the context of strong semantic processing).

4. Results

4.1 Fidelity

A fidelity checklist included number of sessions, length of intervention period for each type of therapy (semantic or phonological), session length and key aspects of the protocol. Twenty-four intervention sessions were selected and examined according to a pre-determined template. Nineteen children of the twenty included (95%) received exactly six intervention sessions for each type of intervention, occurring over an approximate six-week period (mean length of intervention period for each therapy type = 5.38 weeks). The average length for

[3] Raw data on these tasks from typically developing participants and children with WFD are available via UK Dataservice:
https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=851771

the intervention session, excluding assessment, set-up and clear-up time, ranged from 21 to 28 minutes per child with a mean of 25 minutes.

In relation to the therapy protocol, for 100% of cases therapists followed the protocol accurately. This entailed two aspects (i) using the word webs and (ii) asking the child to provide the appropriate features for a target. The latter necessitated maintaining a clear distinction between the semantic and phonological phases of intervention (i.e. only giving semantic cues in the semantic phase of the intervention and vice versa).

4.2 Intervention outcomes for the group

We hypothesised that children's naming would improve with WORD interventions, specifically there would be more improvement over the intervention than non-intervention phases of the study.

The naming scores over the course of the study are provided in Appendix 1. Mean number of items gained for the whole group was approximately 12 (11.9) items over the two phases of intervention and a single item (1.1) over the non-intervention phases (baseline, wash-out and follow up) of the study. The mixed ANOVA showed no significant effect of GROUP i.e. order of intervention ($F=0.135$, $p=0.718$, partial eta sq=0.007), a significant effect of TIME ($F=36.393$, $p<0.001$, partial eta sq=0.669) and no significant interaction ($F=0.570$, $p=0.754$, partial eta sq=0.031). The lack of a significant interaction indicates that the effect of TIME on performance was not significantly different across the two GROUPs (i.e. the different orders of intervention). Pairwise comparisons for naming over time (each assessment compared with the next from A1 to A7) did not reach significance with the exception of change between assessments 3 and 4 i.e. over the first therapy phase (mean difference=6.25, se=1.205, $t=5.187$, Cohen's $d=1.16$, $p=0.001$) and assessments 5 and 6 i.e. over the second therapy phase (mean difference=5.65, se=1.388, $t=4.069$, Cohen's $d=0.91$, $p=0.025$). The results for all 100 items are illustrated in Figure 2a.

In summary, there is a clear effect of intervention over and above development and hypothesis 1.4.1 (i) was not rejected.

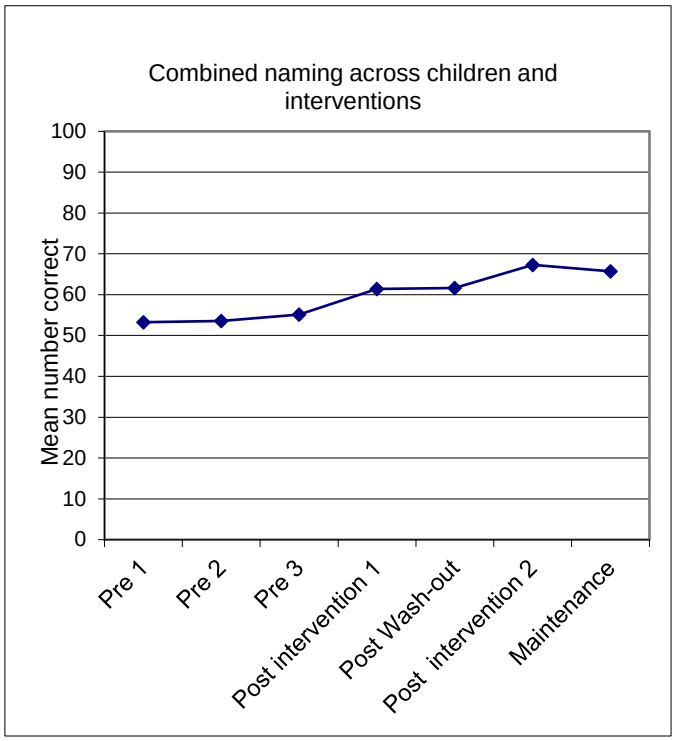
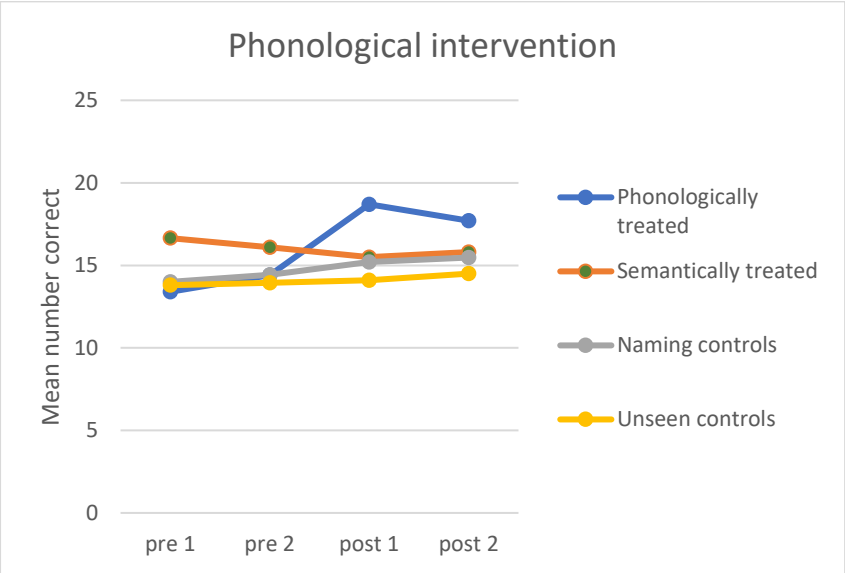


Figure 2a: Naming accuracy for the 100 items over the course of the study. Note that the first intervention was semantic for 10 children and phonological for 10 children, thus Figure 2a cannot be used to compare progress with the two interventions.

We turn next to hypothesis 1.4.1.ii regarding gain from different sets of items. Gain on the four 25 items sub-sets is illustrated in Figure 2b with the data from the two different orders of intervention combined. The change at the point of intervention is evident, but only for sets that are treated.



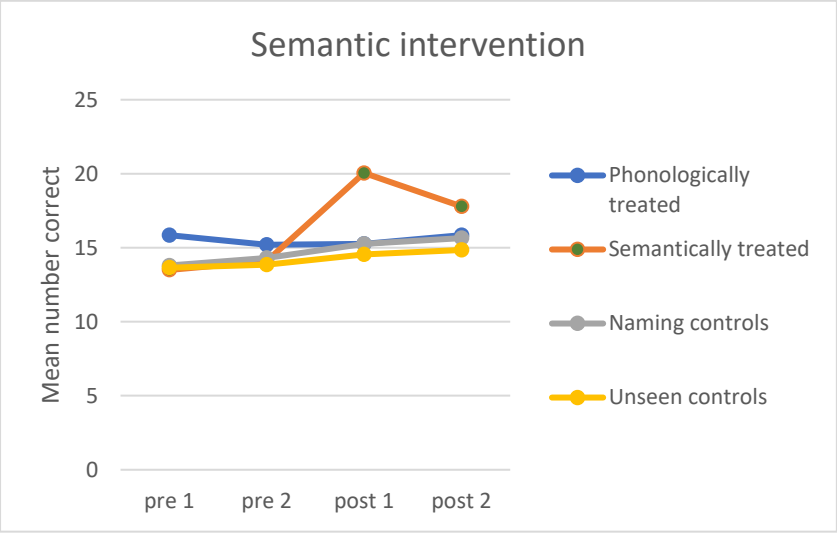


Figure 2b: Number of items correct over the course of the study for the four sub-sets of items for the group. The figure ignores order of intervention and has two pre-intervention data points because this is the maximum number available for sets treated second (see Figure 2, intervention design). In this figure these sessions are A2 and A3 for the first intervention and A4 and A5 for the second intervention. The post 1 and 2 are sessions A4 and A5 for the first intervention and A6 and A7 for the second intervention.

With regard to the two different interventions, the mean number of items gained from the semantic intervention was 7.65 while the mean number of items gained from the phonological intervention was 4.25. Thus, for the group as a whole, the semantic

intervention resulted in a gain of almost twice as many items as that from the phonological intervention. This difference is statistically significant (paired samples t-test, t= 2.187, df 19, p=0.014, 2-tailed, Cohen's d=0.489) and is illustrated in Figure 2c. Thus hypothesis 1.4.1 ii) that there will be a difference in the effect of the two interventions is not rejected.

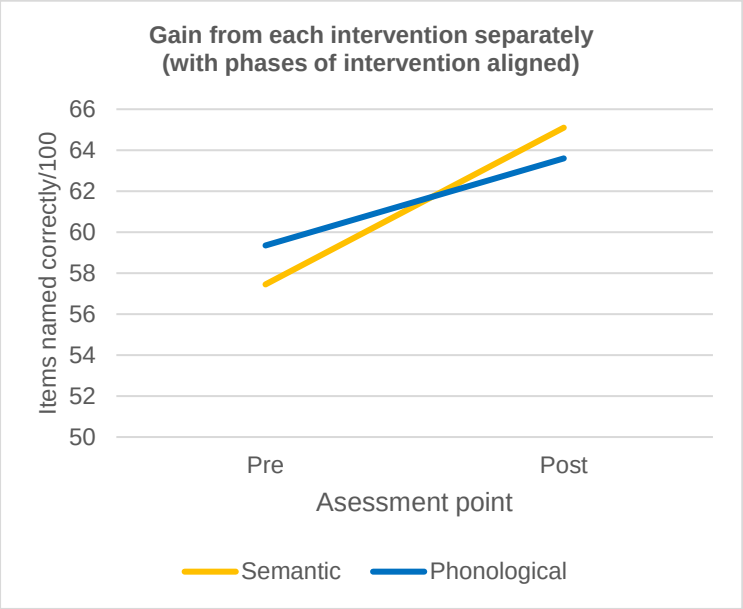


Figure 2c: Naming accuracy on 100 items, showing group data from immediately pre and immediately post each intervention separately. Each child contributes to each gain. Note the group performance is not matched at pre-intervention as the data is taken from different points in the study for the children who participated in the two orders of intervention.

With regard to the order of the interventions (to which children were assigned randomly) the mean number of items gained (/100) over both interventions was 11.5 for those who participated first in the phonological intervention (n=10 children) and 12.8 for those who participated first in the semantic intervention (remaining 10 children). The non-significant effect of GROUP (i.e. order of interventions) in the mixed ANOVA reported above means hypothesis 1.4.1 (iii) that there would be an effect of intervention order was not supported.

Tests for homogeneity were carried out to establish whether the change in naming over the study differed for different children. The first asked whether there was a difference in the trend to change from A1 to A7 for different children. The test for homogeneity was significant

(z=14.96, p=0.000). The second test asked whether the extent to which change during the intervention phases of the study was greater than change over the non-intervention phases (baseline, wash-out and follow-up) differed across children. This was also statistically significant (z=5.85, p=0.000). Thus hypothesis 1.4.1 (iv), that the children would differ in their response to intervention, was supported.

A scatterplot illustrating change with the two different interventions is provided in Figure 2d. It is clear that it is not the case that there are simply some children who respond to therapy and others who do not. Instead there is *no* correlation between the amount of benefit from one intervention and benefit from the other (r=0.062, df 18, n.s.).

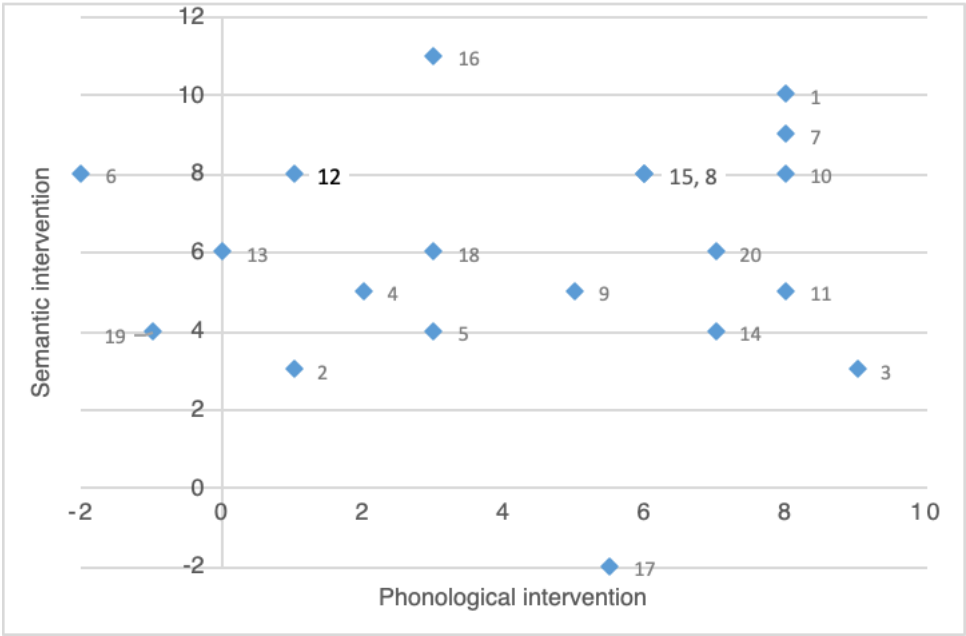


Figure 2d: Items gained by each child on the two different interventions (with individual participant numbers shown). Data shown is out of the 25 items treated with each intervention (these sets differed for each child).

In summary, the children differ from one another in the change in naming over the course of the study, in whether they changed more over intervention than non-intervention phases and in their response to the different interventions. The group data does not therefore reveal the

full story for the individual children. A key question is whether we can predict the response to intervention from the children's language profiles.

4.3 Intervention outcomes for the case series

Results were analysed at the case series level with each child acting as their own control.

Child*	**Trend over study n=100**		**Greater change during therapy phases treated items n=50**		**Gain on set treated with semantic word-webs n=25**	**McNemar test**	**Gain on set treated with phonological. word-webs n=25**	**McNemar test**
	t	**p**	**t**	**p**	**Raw number**		**Raw number**	
1	3.75	0.000	4.72	0.000	10	sig	8	sig
3	4.98	0.000	3.91	0.000	3	ns	9	sig
7	3.87	0.000	4.69	0.000	9	sig	8	sig
9	5.62	0.000	3.93	0.000	5	sig	5	ns
10	4.24	0.000	5.04	0.000	8	sig	8	sig
13	1.66	0.050	1.41	0.083	6	sig	0	ns
14	0.13	0.448	3.22	0.001	4	ns	7	sig
15	6.02	0.000	3.69	0.000	8	sig	6	sig
18	4.40	0.000	1.96	0.028	6	sig	3	ns
19	2.40	0.009	1.35	0.092	4	ns	-1	ns
20	4.85	0.000	2.71	0.005	6	sig	7	sig
2	2.87	0.002	0.44	0.330	3	ns	1	ns
4	2.44	0.008	2.30	0.013	5	ns	2	ns
6	0.53	0.299	2.48	0.008	8	sig	-2	ns
8	3.99	0.000	2.81	0.004	8	sig	6	ns (0.055)
12	3.68	0.000	2.63	0.006	8	sig	1	ns
16	4.30	0.000	2.42	0.010	11	sig	3	ns

5	3.28	0.001	2.04	0.023	4	ns	3	ns
11	4.31	0.000	2.56	0.007	5	ns	8	sig
17	2.63	0.005	0.95	0.174	-2	ns	6	sig
Number of children for whom significant		18		16		12		9

Table 2: The results of case series analyses for all items and treated sub-sets.

*In this table, children have been re-ordered by clinical sub-groups. See section 4.4, below.

The results of the case series analysis are presented in Table 2. The left hand side of the table employs weighted statistics (Howard et al., 2015). There was a trend to better naming over the study which was statistically significant for 18 of the 20 children. There was significantly greater change during intervention phases than other phases (baseline, wash-out and follow-up) of the study for 16 of the 20 children.

Taking the individual patterns in turn, the right-hand columns show that for 12 of the 20 children there was significant change on the sub-set of items given semantic therapy when comparing immediately before and after intervention (of the 12, 7 did not show significant benefit from phonological intervention). For 9 of the 20 children there was significant change on the sub-set of items given phonological therapy comparing scores immediately before and after intervention (of the 9, 4 did not show significant benefit from the semantic intervention). We also carried out weighted statistics taking account of data from across the course of the study. The results taking into account all assessment points, were more generous with regard to statistical significance, showing benefit for 15 children from semantic intervention and for 11 from the phonological intervention. Here we use the more conservative findings from the McNemar tests which simply examined whether there was significant change on naming the set of items treated at the point of intervention

4.4 Intervention outcomes in relation to sub-types from background testing

We hypothesised that the patterns of improvement would differ according to sub-grouping. Specifically, we predicted that outcome would vary according to children's conceptual semantic processing (tapped by judgements in the PJs task) and phonological processing (tapped by non-word repetition task performance).

The outcome of the two interventions in relation to background profile is shown in Figure 3. The sub-types which emerged are also shown in Appendix 1.

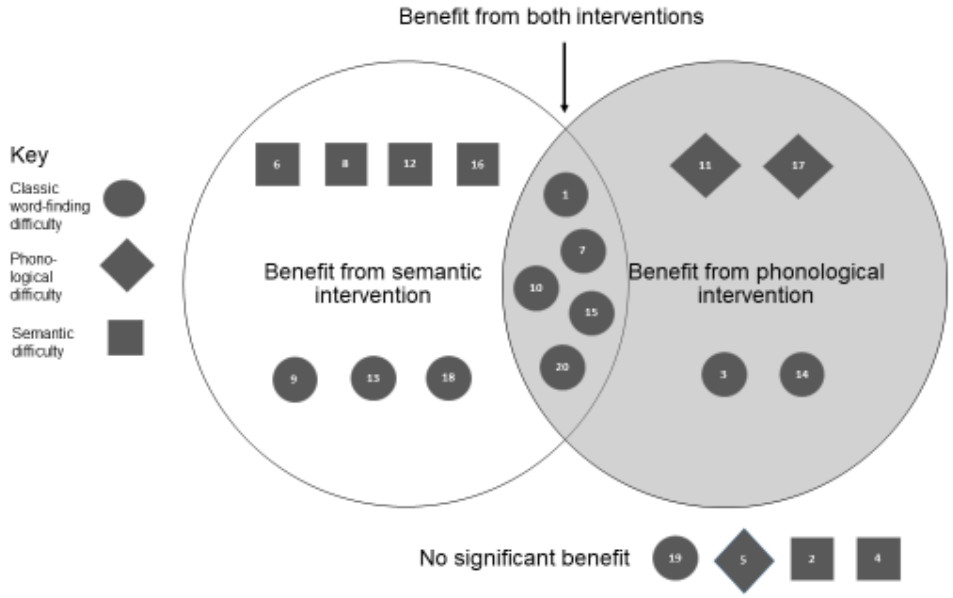


Figure 3: Venn diagram illustrating outcome of intervention per participant (participant numbers as per Tables 1 and 2) as assessed by significant change on the set treated with each intervention in relation to sub-group.

Outcome in relation to predictions for sub-groups.

(i) classic WFD

Children with this profile were predicted to benefit from the semantic or phonological interventions or both. Eleven children fell into the category of classic WFD. Three showed significant benefit only from semantic intervention, two only from phonological intervention and five from both interventions. Thus, all the children except one benefitted from one or both types of word-web.

(ii) semantic difficulties

Six children performed relatively poorly (more than 1.5 sd below the mean for typically developing children). Those with this profile were predicted to benefit from semantic intervention but not from phonological intervention. This was also largely supported as four of the children benefitted from the semantic approach and none from the phonological word-webs. Neither of the two children in this sub-group (child 2, child 4) who did not benefit significantly from either intervention performed above chance level of accuracy on the semantic picture judgements, both obtained a raw score of 11/20 (the other 18 children obtained a raw score of 14/20 or above). One of these two children, child 4, did gain 5 items from the semantic intervention, but this did not reach statistical significance due to variability in his naming.

(iii) phonological difficulties (in the context of strong semantic processing)

Three children performed well on the semantic task but struggled with non-word repetition. Such children with phonological difficulties were predicted to benefit from phonological intervention but to show no benefit from semantic intervention. This prediction was largely supported as two of the three showed significant improvement from the phonological word-webs and none benefitted significantly from the semantic intervention.

Overall, the outcome fully aligned with the predictions made according to sub-typing for 16 of the 20 children. For the remaining four an intervention predicted to benefit did not result in significant change. However, for two of these children (children 4 and 5) there was

significantly greater change in naming during the therapy phases of the study. In no case did a child demonstrate change from an intervention predicted not to benefit their word-retrieval.

4.5 Outcomes according to hypotheses for analysing outcomes for different levels of design

(i) The results from the group analysis showed that there was benefit from the intervention and that the semantic word-webs produced significantly more benefit in naming than the phonological word-webs. The logic of deriving implications from RCTs is that this outcome can be generalised to other children with WFD meeting the inclusion criteria. A reasonable decision made on the basis of this level of evidence alone would be that resources should be directed towards semantic rather than phonological intervention with children with WFD. However, this conclusion is discussed in detail as it does not follow for children identified as having phonological difficulties.

(ii) We predicted, and found, that the sub-grouping would be possible given heterogeneity in children's profiles. Furthermore, analysis of the outcomes for individual children in relation to sub-groups revealed different outcomes according to sub-grouping providing a basis for theoretically informed and clinically useful conclusions.

5. Discussion

5.1 Overview of findings in relation to predictions

The findings from 20 children with WFD, each of whom participated in two types of intervention, provide us with new insights into outcomes at group and case-series level. The intervention improved children's word-finding abilities as measured by picture naming with statistically significant change occurring only during treatment phases of the study. At the group level, significantly more benefit resulted from the semantic than from the phonological word-webs. This is in line with the success of other approaches focusing on semantic processing (e.g. Ebbels et al., 2012). At the case series level for the 50 items treated (sub-sets of 25 each in the phonological and semantic interventions) 16 of the 20 children made

significantly greater gains during treatment phases of the study than non-treatment phases. This fits well with the wider literature demonstrating the effectiveness of WF interventions (Bragard et al., 2012; German, 1992; German, Schwanke, & Ravid, 2012; McGregor, 1994) and with the phase of this work reported previously in which children assigned to intervention immediately gained four times more items than those assigned to a waiting control group (Best et al., 2017)

The children's profiles were heterogeneous and they responded differently to the two interventions in the study which provides us with an opportunity to determine whether outcomes were predictable when children were sub-divided by background profiles. Specifically, the differential response (see scatterplot figure 2d) is consistent with the two interventions operating in different ways and suggests the effects are arising at different loci within language production. The results broadly align with our predictions:

Children with classic WFD benefitted from either or both interventions. These children were most likely to benefit with 11/12 showing significant gains from one or both interventions. This fits well with their difficulty being in accessing the form of lexical items. This may be supported by an additional semantic boost or by additional phonological information feeding up from phonological representations to the output lexicon (Goldrick & Rapp, 2007).

Those with relatively more semantic difficulties (regardless of performance on phonological tasks) tended to benefit from semantic but not phonological therapy. In this case there may not be adequate information about word meanings feeding forward into production and in enriching this, word-retrieval is facilitated. This was the case for four out of six children with this profile, the remaining two did not show significant benefit from either intervention.

Children with relatively more phonological processing difficulties and semantic strengths benefitted from phonological intervention but did not benefit from intervention with semantic word-webs. For these children semantic knowledge may already be available and

information about the phonological form can aid naming. This was the case for two of the three children with this profile, the remaining child did not benefit significantly from either intervention.

Finally, in relation to group and case series analyses, we hypothesised (1.4.4. (i)) that the results from the group analysis would provide an understanding of outcomes on average which may be generalised to other children with WFD meeting the inclusion criteria. The superiority of the semantic over the phonological approach demonstrated for the group *could* be used to suggest that WFD generally arise due to semantic difficulties and be taken as evidence that interventions should all entail semantic rather than phonological word-webs. Much of 'Evidence Based Practice' relies on this approach.

Importantly, however, the case series analysis, possible because of baseline testing, enabled us to demonstrate that, as in other studies (Smith Lock et al., 2015), the results for some individuals differ from those for the group. Because of this we know that the 'one size fits all' semantic approach will not be the optimum intervention for some children with WFD. We hypothesised (1.4.4. (ii)) that heterogeneity in children's profiles would result in different and predictable outcomes enabling more specific theoretical and clinically informative conclusions.

If the findings from the group analysis are applied, by providing semantic therapy for all, then 12 of the 20 children benefit. If the findings from the case series are applied, by tailoring intervention, an additional four, that is 16 children overall, show significant improvement. Taking account of the individual profiles would thus enable a third more children to benefit. We therefore place significant emphasis on the findings from the case series analyses. We also recognise that not every prediction was supported fully at an individual level and that the sub-groups that emerged differ in size.

5.2 Strengths of the study

The research is novel in the inclusion of 20 children with WFD each of whom participated in two interventions. While a greater number of children would have increased power, the study is the largest in the field and is unusual in both the length of involvement and the amount of data collected (for example,100 items were named on seven occasions by each child). The intervention approaches were carefully constructed to facilitate direct comparison of outcome. Further strengths include blinding of the assessor to intervention, evaluation of fidelity, inclusion of control items named at the start of each session and follow-up testing half a term after the end of the second intervention. The finding that the effects maintain is important as it is these persisting gains that will re-inforce future retrieval of treated words. The lack of change on the set of untreated items which were presented for naming at the start of each intervention session (the naming controls), indicates that the improvement in treated items results from something other than simply practice in trying to retrieve the target words.

The experimental design was strong enough to enable analysis at group and case series levels. Randomisation was built into the design in several ways and was carried out by the 5th author at a separate institution. Children were randomly assigned to which intervention came first. Sets, which differed for each child and were matched for psycholinguistic variables and for baseline naming, were randomly assigned to condition: semantic or phonological intervention, naming control or unseen control. The study did not entail selecting only 'hard to name' items, as in that case apparent treatment effects could have reflected regression to the mean (Howard et al., 2015). Finally, we employed the more conservative McNemar test (which results in fewer children demonstrating significant change), rather than weighted statistics across the whole study, to determine whether there was significant change.

Clinical strengths of the research were inclusion of children with WFD attending mainstream schools. Previous WF intervention studies have involved a smaller number of children (e.g.

Bragard et al., 2012, Best, 2005, Ebbels et al., 2012). Because of the heterogeneous response to the different interventions we know that it is not simply the case that some children are more amenable to this type of intervention. Rather, it is possible to use theoretically driven sub-grouping to predict outcome and to obtain statistically significant improvement in line with predictions.

The intervention was clinically realistic as it reflects the half-termly blocks of six sessions of therapy often offered to children with DLD in the UK. While additional intervention might produce further gains, we have demonstrated significant benefit from this dose and intensity. We also included a set of items unique to the child, following their interests, items provided by parents and/or new topics in the school curriculum (the outcome for these items after the first phase of therapy is described in Best et al., 2017).

5.3 Potential weaknesses

We turn now to the limitations of the study. Firstly, despite the focus on meta-linguistic processing and on using features to aid word-retrieval in communicative contexts (barrier games), the changes found were very largely limited to treated items. This is in line with the wider literature. However, Wilson et al. (2015) found some evidence for generalisation to semantically related items from their WF intervention. In closely related interventions with adults with anomia as part of their aphasia, generalisation to untreated items is also rarely obtained. Future studies could manipulate semantic and phonological overlap between treated and control items to investigate this further. In depth consideration of this issue leads us to suggest that it could well be that there is improvement on untreated items, *just not on those that happened to be re-tested after intervention*. This result also highlights the importance of selecting items to work on that are relevant to the individual children according to both their own interests and for the school curriculum.

With regard to the specific clinical predictions according to sub-group it is perhaps generous to say the prediction was fully supported for 16/20 children as this relies on 'either' of the two interventions being effective for those with 'classic WFD'. Whether this sub-group would benefit most from word webs containing *both* semantic and phonological features, rather than one or the other as in this study, remains to be investigated in future research. In an ideal world the outcome could be completely predictable for each child. Although this was not achieved, in no case did a child predicted *NOT* to benefit from an intervention demonstrate significant change. In practice SLPs commonly use word-webs incorporating both semantic and phonological features and will modify the nature of the intervention according to each child's response.

A final, important, concern is that two participants (child 2, child 19) spent time in the research study but did not demonstrate significant measurable change in word retrieval on formal naming assessment either on treated sets or during therapy phases relative to non-therapy phases of the study (baseline, wash-out and follow-up). Both children, like those who made gains, enjoyed participation in the study. After the study child 2, whose profile showed semantic difficulties, was reported as being more talkative and confident according to her family and school, but this may have occurred in any case. Child 19, despite having relative strengths in semantic and phonological processing, i.e. with classic WFD, did not benefit from either approach. It may be that a combined semantic and phonological word-webs could have benefitted her naming. Certainly, further research investigating the success of a combined approach is warranted. It is also important to acknowledge that approaches other than word-webs may be more suitable for some children with WFD, perhaps particularly for those children at an earlier stage developmentally and not yet able to manage semantic picture-based tasks or perhaps unable to exploit meta-linguistic strategies employed in the word-web interventions.

5.4 Implications for research

For optimal understanding of results in intervention research in this field, and others, it is beneficial to analyse findings for both the group and case series. Here, the conclusion of the former would be that semantic therapy for WFD results in significantly greater gains than phonological therapy. Indeed children, on average, gained almost twice as many items from the former approach than the latter. However, in this project, the inclusion of all children in both interventions enabled analysis of outcome for each individual. The results highlight the value of well controlled case series and have implications for the design and analysis of intervention studies.

One approach in future could be to recruit according to the sub-typing that helped us interpret the findings from this study. This would need to be a very large-scale project, however, as the sub-groups would need to be matched for variables such as initial naming accuracy and it is not yet clear which other factors may be relevant in predicting outcome of intervention for children with WFD. It is also important to caution against large over-powered studies which may produce statistically, but not clinically, significant results.

While there remains no fully articulated developmental model of lexical representation and retrieval, the findings can be understood in relation to models of language production with feedback from phonological representations to lexical items during retrieval (e.g. Goldrick and Rapp, 2002). As parallels between language systems of children, particularly those of school age, and adults are acknowledged (Bishop, Nation, & Patterson, 2014) and further models develop (Friedmann et al., 2013) including computational accounts (Thomas et al., 2019), this should become more fruitful.

5.5 Implications for clinical practice

RCTs carry weight in planning service delivery. Sound studies with appropriate control conditions can result in useful conclusions generalisable to the wider population meeting the inclusion criteria, in this case 6-8 year old children with WFD as assessed by the TWF. The

group results show that, if taking a blanket approach, recommending semantic word webs would benefit more children than recommending the use of phonological word webs for intervention.

It is important to recognise, however, that RCTs can obscure variability in outcome. The optimum choice of intervention for more children can be obtained by considering results from the case series. This focus on the individual is in line with the wider movement towards individualised medicine and a broader understanding of evidence based practice. Greenhalgh et al., (2014) call for *real* Evidence Based Medicine, for example, which demands meaningful individualised evidence to inform clinical decision making.

We recognise that SLPs will use expert judgement and tailor intervention according to the child, outcome and clinical context (for example, in the UK this reflects guidance from the professional body: Royal College of Speech and Language Therapists, 2018). This study enhances the evidence base by suggesting:

i) Children with relative strengths in both domains (i.e. difficulty in accessing word forms) may benefit from either approach – or integrated word-webs not used in this study.

ii) Those with relatively more semantic difficulties (regardless of performance on phonological tasks) are likely to benefit from semantic but not phonological intervention.

iii) Children with relatively more phonological difficulties and semantic strengths are likely to benefit from phonological but not semantic intervention

iv) Some children will respond to neither. These children may be those who perform poorly on semantic picture judgement tasks.

v) Given the lack of generalisation to untrained items that were tested the selection of functionally useful items is important for maximising the impact of intervention.

These recommendations can be also be applied to programmes used with parents or in schools. As with all guidance this should be interpreted in the light of clinical context, shared decision making and with responsiveness of the child to the fore.

In clinical practice, SLPs will also routinely aim to ensure changes in children's ability to retrieve single words in isolation are carried over to their use in discourse. The optimum ways of achieving this remain to be determined. Furthermore, the most effective interventions will be encouraging children to develop and extend their own word-retrieval strategies according to what is most useful individually.

Finally, as demonstrated by one parent's views on the effect of involvement in the study:

'I have seen a difference from last year, he is more talkative, he is better at picking up new words'. 'He will ask me what a word means and I will try and explain it and then when he hears it again he will ask me again'

Although it was not the focus of the research, children participating in the study were also reported to have become more aware of *word-learning* and to use more strategies after the intervention. It may be the case that as new words are learned with enriched representations children's word-finding is supported on a longer-term basis.

6. Conclusions

The study provides a theoretical framework and evidence-based recommendations that aim to optimise the time spent on key lexical features to help children find the words with which they are struggling. Generalisation of word learning and retrieval beyond trained items continues to be a key challenge for researchers and clinicians. Analysis of case series data can reveal patterns beyond those that emerge from analysis at group level and can produce elevated rates of successful outcomes.

Acknowledgements:

We are grateful to the children involved in the project and to their parents and teachers. Several speech and language therapists and researchers contributed to data collection during the wider study: Vivien Gibson, Liory Fern-Pollak, Jo Piper, Silvia Roncoli and Donna Lynn Shepherd. We are especially grateful to our expert advisory panel Kathleen Cavin and Susan Ebbels and our wider advisory group. Colleagues who have contributed to the wider WORD project of which this forms a part include Hala Alireza, Carol Amos, Claudia Bruns, Chris Donlan, Anna Fedor, Victoria Fleming, Katie Kinnear, Caroline Newton, Merle Mahon, Jo Piper, Rachel Rees, Sarah Simpson, Amelia Wong.

References:

Best, W., Fedor, A., Hughes, L., Kapikian, A., Masterson, J., Roncoli, S., … Thomas, M. S. C. (2015). Intervening to alleviate word-finding difficulties in children: case series data and a computational modelling foundation. *Cognitive Neuropsychology*, *32*(3–4), 133–168. https://doi.org/10.1080/02643294.2014.1003204

Best, Wendy. (2005). Investigation of a new intervention for children with word-finding problems. *International Journal of Language and Communication Disorders*, *40*(3), 279–318. https://doi.org/10.1080/13682820410001734154

Best, Wendy, Hughes, L. M., Masterson, J., Thomas, M., Fedor, A., Roncoli, S., … Kapikian, A. (2017). Intervention for children with word-finding difficulties: a parallel group randomised control trial. *International Journal of Speech-Language Pathology*, pp. 1–12. https://doi.org/10.1080/17549507.2017.1348541

Bishop, D.V. M. (1996). *Test for the Reception of Grammar* (2nd ed.). Manchester, England: University of Manchester: Age and Cognitive Performance Research

Centre.

Bishop, Dorothy V.M., Nation, K., & Patterson, K. (2014). When words fail us: Insights into language processing from developmental and acquired disorders. *Philosophical Transactions of the Royal Society B: Biological Sciences*. https://doi.org/10.1098/rstb.2012.0403

Bishop, D. V, Snowling, M. J., Thompson, P. A., & Greenhalgh, T. (2016). CATALISE: a multinational and multidisciplinary Delphi consensus study of problems with language development. Phase 2. Terminology. *PeerJPreprints*, 1–31. https://doi.org/10.7287/PEERJ.PREPRINTS.2484V1

Bragard, A., Schelstraete, M.-A., Snyers, P., & James, D. G. H. (2012). Word-Finding Intervention for Children With Specific Language Impairment: A Multiple Single-Case Study. *Language Speech and Hearing Services in Schools*, *43*(2), 222. https://doi.org/10.1044/0161-1461(2011/10-0090)

Campbell, L., Nicoll, H., & Ebbels, S. H. (2019). The effectiveness of semantic intervention for word-finding difficulties in college-aged students (16–19 years) with persistent Language Disorder. *Autism & Developmental Language Impairments*. https://doi.org/10.1177/2396941519870784

Constable, A., Stackhouse, J., & Wells, B. (1997). Developmental word-finding difficulties and phonological processing: The case of the missing handcuffs. *Applied Psycholinguistics*, *18*(04), 507–536. https://doi.org/10.1017/S0142716400010961

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical Access in Aphasic and Nonaphasic Speakers. *Psychological Review*, *104*(4), 801–838. https://doi.org/10.1037/0033-295X.104.4.801

Dockrell, J. E., Messer, D., George, R., & Wilson, G. (1998). Notes and discussion: Children with word-finding difficulties prevalence, presentation and naming problems. *International Journal of Language and Communication Disorders*, *33*(4), 445–454. https://doi.org/10.1080/136828298247721

Druks, J & Masterson, J. (2000). *Object and Action Naming Battery*. Hove: Psychology Press.

Ebbels, S. H., Nicoll, H., Clark, B., Eachus, B., Gallagher, A. L., Horniman, K., … Turner, G. (2012). Effectiveness of semantic therapy for word-finding difficulties in pupils with persistent language impairments: A randomized control trial. *International Journal of Language and Communication Disorders*, *47*(1), 35–51. https://doi.org/10.1111/j.1460-6984.2011.00073.x

Elliot, C.D., Smith, P. & McCullouch, K. (1996). *British Ability Scales* (2nd ed.). Windsor: NFER-Nelson.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods, 39, 175-191.

Forster, K. I. and Forster, J. C. (2003) DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, 35 (1), 116-124.

Friedmann, N., Biran, M., & Dotan, D. (2013). Lexical retrieval and its breakdown in aphasia and developmental language impairment. In C. Boeckx & K. K. Grohmann (Eds.), *The Cambridge Handbook of Biolinguistics* (pp. 350–374). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511980435.021

Funnell, E., Hughes, D., & Woodcock, J. (2006). Age of acquisition for naming and knowing: A new hypothesis. *Quarterly Journal of Experimental Psychology*, *59*(2), 268–295. https://doi.org/10.1080/02724980443000674

Gathercole, S. & Baddeley, A. (1996). *Children's Test of Nonword Repetition*. London: Psychological Corporation.

German, D. (2000). *Test of Word Finding (TWF-2)* (2nd ed.). Austin, TX: Pre-Ed.

German, D. J. (1992). Word-finding intervention for children and adolescents. *Topics in Language Disorders*, *13*(1), 33–50. https://doi.org/10.1097/00011363-199211000-00006

German, D. J., & Newman, R. S. (2004). The impact of lexical factors on children's word-finding errors. *Journal of Speech, Language, and Hearing Research : JSLHR*, *47*(3), 624–636. https://doi.org/10.1044/1092-4388(2004/048)

German, D. J., Schwanke, J. H., & Ravid, R. (2012). Word finding difficulties: Differentiated vocabulary instruction in the speech and language room. *Communication Disorders Quarterly*, *33*(3), 146–156. https://doi.org/10.1177/1525740111405840

Gold, S. M., Enck, P., Hasselmann, H., Friede, T., Hegerl, U., Mohr, D. C., & Otte, C. (2017). Control conditions for randomised trials of behavioural interventions in psychiatry: a decision framework. *The Lancet Psychiatry*, *4*(9), 725–732. https://doi.org/10.1016/S2215-0366(17)30153-0

Goldrick, M., & Rapp, B. (2002). A restricted interaction account (RIA) of spoken word production: The best of both worlds. *Aphasiology*. https://doi.org/10.1080/02687040143000203

Goldrick, M., & Rapp, B. (2007). Lexical and post-lexical phonological

representations in spoken production. *Cognition*, *102*(2), 219–260. https://doi.org/10.1016/j.cognition.2005.12.010

Greenhalgh T., Howick J., Maskrey N. (2014) Evidence based medicine: a movement in crisis? *BMJ* ; 348 :g3725

Howard, D., Best, W., & Nickels, L. (2015). Optimising the design of intervention studies: critiques and ways forward. *Aphasiology*. https://doi.org/10.1080/02687038.2014.985884

Masterson, J., Druks, J., & Gallienne, D. (2008). Object and action picture naming in three- and five-year-old children. *Journal of Child Language*, 35, 373–402. http://doi.org/10.1017/S0305000907008549

McGregor, K K. (1994). Use of phonological information in a word-finding treatment for children. *Journal of Speech and Hearing Research*, *37*(6), 1381–1393.

McGregor, Karla K., & Appel, A. (2002). On the relation between mental representation and naming in a child with specific language impairment. *Clinical Linguistics & Phonetics*, *16*(1), 1–20. https://doi.org/10.1080/02699200110085034

McGregor, Karla 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*(3), 998–1014. https://doi.org/10.1044/1092-4388(2002/081)

McKean, C., Reilly, S., Bavin, E. L., Bretherton, L., Cini, E., Conway, L., … Mensah, F. (2017). Language Outcomes at 7 Years: Early Predictors and Co-Occurring Difficulties. *Pediatrics*, *139*(3), e20161684. https://doi.org/10.1542/peds.2016-1684

Messer, D., & Dockrell, J. E. (2013). Children with word finding difficulties: Continuities and profiles of abilities. *First Language*, *33*(5), 433–448. https://doi.org/10.1177/0142723713493345

Messer, D., Dockrell, J. E., & Murphy, N. (2004). Relation between naming and literacy in children with word-finding difficulties. *Journal of Educational Psychology*, *96*(3), 462–470. https://doi.org/10.1037/0022-0663.96.3.462

Royal College of Speech and Language Therapists (2018). Placing children and young people at the heart of delivering quality speech and language therapy: Guidance on principles, activities and outcomes Available at: https://www.rcslt.org/speech-and-language-therapy/where-slts-work/childrens-services#section-5 Accessed: 3 July, 2020.

Semel, E., Wiig, E., & Secord, W. (2003). *Clinical evaluation of language fundamentals, fourth edition (CELF-4)*. *Pearson Clinical Assessment*.

Shamseer, L., Sampson, M., Bukutu, C., Schmid, C. H., Nikles, J., Tate, R., … Vohra, S. (2015). CONSORT extension for reporting N-of-1 trials (CENT) 2015: Explanation and elaboration. *BMJ (Clinical Research Ed.)*, *350*, h1793. https://doi.org/10.1136/bmj.h1793

Smith-Lock, K. M., Leitão, S., Prior, P., & Nickels, L. (2015). The effectiveness of two grammar treatment procedures for children with SLI: A randomized clinical trial. *Language, Speech, and Hearing Services in Schools*, 46(4), 312–324.

Stothard, S., & Snowling, M. (1998). Language-impaired preschoolersA follow-up into adolescence. *Journal of Speech, …*, *41*(2), 407–418. https://doi.org/10.1044/jslhr.4102.407

Tate, R.L., Rosenkoetter, U., Wakim, D., Sigmundsdottir, L., Doubleday, J., Togher,

L., Perdices, M. (2015). *The Risk of Bias in N-of-1 Trials (RoBiNT) Scale An expanded manual for the critical appraisal of single-case reports*. Sydney: Australia.

Thomas, M. S. C., Fedor, A., Davis, R., Yang, J., Alireza, H., Charman, T., … Best, W. (2019). Computational modeling of interventions for developmental disorders. *Psychological Review*. https://doi.org/10.1037/rev0000151

Wilson, J., Aldersley, A., Dobson, C., Edgar, S., Harding, C., Luckins, J., & Wiseman, F., (2015). The effectiveness of semantic therapy for the word finding difficulties of children with severe and complex speech, language and communication needs. *Child Language Teaching and Therapy*, 31(1), 7–17. doi: 10.1177/0265659014523299

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*(3), 151–156. https://doi.org/10.1044/0161-1461.2103.151

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*(3), 214–229. https://doi.org/10.1177/026565909300900304

Zens, N. K., Gillon, G. T., & Moran, C. (2009). Effects of phonological awareness and semantic intervention on word-learning in children with SLI. *International Journal of Speech-Language Pathology*, *11*(6), 509–524. https://doi.org/10.3109/17549500902926881

Appendices

Appendix 1: Naming accuracy on 100 items for each child.

Participant / sub-group	Intn. order	A1 Pre 1	A2 Pre 2	A3 Pre 3	A4 Post intn. 1	A5 Post 'wash out'	A6 Post intn. 2	A7 Post 'follow-up'
1	PS	49	51	54	65	65	68	62
3	PS	47	41	42	47	50	57	59
7	SP	48	46	50	55	54	53	56
9	PS	61	57	60	65	57	66	57
10	PS	62	61	68	75	73	85	88
13	PS	55	62	59	61	74	68	63
14	SP	52	58	54	67	62	69	68
15	SP	55	55	56	70	62	72	71
18	PS	59	59	64	61	66	68	66
19	PS	44	54	54	58	63	67	66
20	SP	64	64	65	76	77	83	80
2	SP	69	67	62	77	78	82	83
4	SP	43*	54	52	52	55	58	55
6	SP	57	56	60	66	65	79	74
8	SP	51	50	56	57	59	64	63
12	PS	42	39	46	57	50	65	58
16	SP	42	42	40	51	57	62	64

5	PS	53	51	54	52	48	58	59
11	SP	53	47	52	60	62	53	49
17	PS	49	58	55	56	56	69	74

Order of intervention: P – phonological, S –Semantic, Intn. – intervention

Sub-group is indicated by colour of text. Those with classic WFD 1, 3, 7, 9, 10, 13, 14, 15, 18, 19, 20 are shown in green. Children with semantic difficulties, children 2, 4, 6, 8, 12, 16 are in blue. Those with phonological difficulties 5, 11, 17 are shown in orange.