

Title	The impact of intervention dose form on oral language outcomes for children with developmental language disorder
Authors	Frizelle, Pauline; Tolonen, Anna-Kaisa; Tulip, Josie; Murphy, Carol-Anne; Saldana, David; McKean, Cristina
Publication date	2021-07-02
Original Citation	Frizelle, P., Tolonen, A.-K., Tulip, J., Murphy, C.-A., Saldana, D. and McKean, C. (2021) 'The Impact of Intervention Dose Form on Oral Language Outcomes for Children With Developmental Language Disorder', Journal of Speech, Language, and Hearing Research, doi: 10.1044/2021_JSLHR-20-00734
Type of publication	Article (peer-reviewed)
Link to publisher's version	https://pubs.asha.org/doi/abs/10.1044/2021_JSLHR-20-00734 - 10.1044/2021_JSLHR-20-00734
Rights	© 2021 American Speech-Language-Hearing Association - https://creativecommons.org/licenses/by-nc-nd/4.0/
Download date	2024-04-19 22:02:13
Item downloaded from	https://hdl.handle.net/10468/11635

1 **The impact of intervention *dose form* on oral language outcomes for children**
2 **with Developmental Language Disorder**

3

4 Pauline Frizelle,^{1,6} Anna-Kaisa Tolonen² Josie Tulip³, Carol-Anne Murphy⁴, David Saldana
5 ⁵, and Cristina McKean³,

6

7 ¹ Pauline Frizelle (corresponding author), Department of Speech and Hearing Sciences,
8 University College Cork, Republic of Ireland. Email: p.frizelle@ucc.ie

9 ² Research Unit of Logopedics, Faculty of Humanities, University of Oulu, Finland

10 ³ Department of Speech and Language Sciences, School of Education, Communication &
11 Language Sciences, Newcastle University, UK

12 ⁴ School of Allied Health, Faculty of Education and Health Sciences, University of Limerick,
13 Republic of Ireland

14 ⁵ Departamento de Psicología Evolutiva y de la Educación, University of Seville, Spain

15

16 Conflict of Interest

17 The authors declare no conflicts of interest.

18

19 Funding Statement

20 This research paper was part of the work of the Cost
21 Action IS1406 network entitled: *Enhancing children's oral language skills across Europe and*
22 *beyond - a collaboration focusing on interventions for children with difficulties learning their*
23 *first language*. The network was supported by COST (European Cooperation in Science and
24 Technology) funded by the European Union [Grant: COST 106/14]. This work was also

25 supported in part by an NHMRC fellowship to C.McK. (Centre of Research Excellence in
26 Child Language 1023493).

27

28 RUNNING HEAD: Dose form manipulation in language interventions for DLD

29

30

Abstract

31 *Purpose:* To extract key learning from intervention studies in which qualitative aspects of
32 dosage, *dose form*, have been examined for children with developmental language disorder
33 (DLD) - in vocabulary, morphosyntax and phonology domains. This research paper emerged
34 from a pair of systematic reviews, aiming to synthesise available evidence regarding qualitative
35 and quantitative aspects of dosage respectively. Whilst quantitative aspects had been
36 experimentally manipulated, the available evidence for *dose form* (tasks or activities within
37 which teaching episodes are delivered) was less definitive. Despite this, the review uncovered
38 insights of value to DLD research.

39 *Method:* A pre-registered systematic review (PROSPERO ID=CRD42017076663) adhering to
40 PRISMA guidelines was completed. Included papers were: Quasi-experimental, RCT or cohort
41 analytic studies, published in any language between January 2006 and May 2019; oral language
42 interventions with vocabulary, morpho-syntax or phonology outcomes; and participants with
43 DLD (M=3-18 years). The intention was to include papers in which dose form was
44 experimentally manipulated or statistically analysed, while quantitative dosage aspects were
45 controlled, such that definitive conclusions about optimal dose form could be drawn, and gaps
46 in the evidence identified.

47 *Results:* 224 papers met the above inclusion criteria; 27 focused on *dose form*. No study
48 controlled for all quantitative aspects of dosage such that we could effectively address our

49 original research questions. Despite this, key points of learning emerged with implications for
50 future research

51 *Conclusions:* There is tentative evidence of advantages for explicit over implicit instruction,
52 and of the benefits of variability in input, elicited production and gestural and other visual
53 supports. With careful design of dose form, there is potential to design more efficient
54 interventions. SLP research would benefit from an agreed taxonomy of dose form components
55 and standardised reporting of intervention studies to enable cross-study comparisons and a
56 systematic accrual of knowledge to identify optimal dose form for clinical application.

57

58 **Introduction**

59 Over the past 20 years, there has been an increased need for accountability in SLP practice
60 both in educational and health settings. Increasingly, the focus of accountability is not only on
61 effectiveness, but also efficiency of practice: an issue inextricably linked to intervention
62 dosage. Warren et al. (2007) put forward five dosage characteristics to describe intervention
63 intensity. Dose form is the qualitative dosage component of their taxonomy and refers to the
64 tasks or activities within which the teaching episodes are delivered. It seems intuitive that
65 differing dose forms will be more or less effective and therefore require more or less
66 quantitative dosage to effect change. However, there is little guidance available for practice in
67 this regard. This paper examines and synthesises current evidence regarding intervention dose
68 form, with respect to children with Developmental Language Disorder (DLD), in the domains
69 of vocabulary, morpho-syntax and phonology. We report findings from a systematic review
70 which aimed to identify studies, which allow valid conclusions to be drawn regarding the
71 *optimal and most efficient* dose forms used in interventions. That is, studies which compare the
72 relative efficacy of differing dose forms, whilst controlling for quantitative aspects of dose. In
73 the following we a) define and describe components of dose form in detail; b) describe the

74 range of dose forms used in interventions for children with DLD, reported in previous reviews
75 (e.g. Cirrin & Gillam 2008; Marulis & Neuman, 2010; Proctor-Williams, 2009; Wren et al.,
76 2018), and what is known about their effectiveness; and c) outline the aims and approach of
77 the current study.

78 **Defining dose form**

79 Over time, the construct of dose form has been developed and refined. In 2009, Proctor-
80 Williams built on the work of Warren and colleagues (2007) in her description of the
81 components of dose form to include ‘the commonly used techniques, procedures, and
82 intervention contexts that constitute teaching episodes’ (p. 295). In this new definition
83 *techniques* are the specific actions/ teaching behaviors thought to have benefit and *procedures*
84 refer to how these techniques are combined. *Intervention contexts* are described as falling on a
85 continuum from highly child-centered, to hybrid, to clinician-directed (Fey, 1986).

86 We further specify and extend the *intervention contexts* component of dose form to
87 capture potential other ‘active ingredients’, we judged as missing from the Proctor-Williams
88 (2009) taxonomy: the *activity* in which the technique/ teaching behavior is being delivered and
89 *the degree of variability* in the linguistic input, or materials used. We also add an additional
90 component: the *method of instruction*, with specific reference to explicit versus implicit
91 approaches. See Table 1.

92 Insert Table 1 about here

93 **What do we know about techniques, procedures, methods of instruction and intervention** 94 **contexts for children with DLD?**

95 ***Techniques***

96 **Vocabulary.** Techniques typically used in vocabulary interventions include exposing
97 children to target words in varied contexts; using visual supports such as iconic gestures;
98 development of meta-linguistic strategies during reading; use of stress/ slower speech rate; and

99 elaboration designed to give the child a richer knowledge of target words', semantics or
100 phonology. A previous review by Cirrin & Gillam (2008), in relation to school aged children
101 with oral language disorders, reported positive effects for slowed rate; interactional
102 conversational reading strategies; and iconic gestures. However, few studies of vocabulary
103 interventions have systematically compared the effectiveness of one technique relative to
104 another. Additionally, prior to the Warren et al. (2007) paper, most studies did not fully report
105 or consider the potential impact of quantitative aspects of dosage in the interpretation of their
106 findings (Zeng et al., 2012).

107 **Morphosyntax.** Techniques most typically used in morphosyntax interventions are
108 imitation; priming; modelling; recasting /enhanced conversational recasting; expansions; and
109 elicitation prompting. While empirical studies show that these techniques facilitate
110 morphosyntactic development in children with DLD (see Proctor-Williams, 2009), lack of
111 information on quantitative dose again makes comparisons of techniques, problematic. There
112 are substantial differences in how techniques, such as imitation training and modelling, are
113 implemented (Eisenberg et al., 2020) and inconsistencies with respect to how these techniques
114 are combined (e.g. Smith-Locke et al., 2015 (modelling recasting and cueing); Owen Van
115 Horne et al. 2018 (drills, imitation, modelling and recasting)). Consequently, it is difficult to
116 tease apart their relative effects. Contradictory findings regarding the relative impact of one
117 technique over another are therefore not surprising. For example, with respect to a group of
118 children with DLD, Courtright and Courtright (1979) found modelling resulted in higher target
119 morpheme use than imitation, whereas Connell and Stone (1992) found imitation to be more
120 effective than modelling.

121 **Phonology.** There is a wealth of different techniques in the field of phonological
122 interventions, including the use of minimal and maximal sound contrasts, stimulability, meta-
123 phonology, and auditory bombardment. However, few studies include children with DLD, and

124 focus instead on those with Speech Sound Disorders without DLD. In a recent review of
125 published studies by Wren and colleagues (2018) 11 different ‘procedures’ were identified
126 falling into five categories – environmental, auditory-perceptual, cognitive-linguistic,
127 production or ‘integrated’ approaches. The authors conclude that it is currently not possible to
128 determine which are most effective.

129 *Method of Instruction – Explicit plus Implicit versus Implicit only*

130 Interventions can also vary depending on whether they can be implemented explicitly,
131 where the child is given information with respect to the rule underlying the teaching target, or
132 implicitly, where they are required to induce the rule/ pattern (see Finestack (2018) for a
133 detailed explanation of each).

134 **Vocabulary.** Explicit vocabulary interventions include providing detailed definitions
135 and examples before, during, or after a book reading activity, with follow up discussions
136 reviewing target vocabulary. An implicit approach on the other hand, would involve a story
137 reading activity, including aspects such as slowed speech, emphasis of target words and
138 gestures, without stopping to reflect on target word meanings or giving target word definitions.
139 Techniques such as those that use analogical thinking/ verbal analogy (highlighting similarities
140 between pairs of words), are thought to require explicit instruction, to be effective with children
141 with language impairment (Masterson & Perrey, 1999). In contrast, results from a meta-
142 analysis completed by Marulis and Neuman (2010) on young typically developing children and
143 those ‘at risk’, suggest the highest effect sizes for interventions using a combination of methods
144 of instruction.

145 **Morphosyntax.** With respect to morphosyntax, modelling, recasting and priming have
146 been most frequently implemented implicitly (Camarata et al., 1994; Leonard et al., 2006), and
147 have been criticized for achieving only moderate intervention effects (Finestack, 2018). For
148 this reason, there has been a move towards using more explicit approaches such as the Shape

149 Coding system (Ebbels, 2014). However, although primarily an explicit approach, which
150 visually codes grammatical rules to make them explicit to the child, the Shape Coding system
151 can also be combined with techniques usually used implicitly, such as modelling, elicitation
152 and recasting. Prior to those papers included in the review reported here, only one study in
153 which quantitative dose was controlled, systematically compared an implicit and explicit
154 intervention approach. The treatment target was a novel noun morpheme, and while overall
155 response to treatment was low, generalization (among those that did respond) was greater in
156 the implicit condition (Swisher et al., 1995).

157 **Phonology.** Explicit phonological interventions used with children with DLD include
158 phonological awareness training, with positive effects reported, albeit in small scale quasi-
159 experimental studies (Gillon, 2000), whilst implicit approaches such as phonological auditory
160 ‘bombardment’, and minimal meaningful contrasts have not been tested, to our knowledge, in
161 children with DLD, and often contain elements of explicit interventions alongside implicit
162 approaches (Baker & McLeod, 2004; Hodson, 2015). Without systematically manipulating the
163 methods of instruction within a given technique or intervention context, it is difficult to
164 ascertain their respective impacts.

165 ***Intervention contexts***

166 Lastly, techniques can be used in a range of intervention *contexts* that vary with respect
167 to degree of structure and naturalness. In more recent years, therapy has moved from the use
168 of highly structured/ decontextualized drills, to embedding targets in meaningful activities
169 within a social context. Most recently, there has been increased interest in the impact of the
170 *linguistic* context and, in particular, the variability of the input.

171 **Vocabulary.** Interactive book reading is an example of a meaningful activity in which
172 targets can be embedded with respect to vocabulary. In relation to the linguistic context and
173 variability of the input, most work has been conducted with typically developing children and

174 findings are mixed. Some has reported positive word learning effects using high referent
175 variability with young typically developing children (Perry et al., 2010), other studies suggest
176 that too much variability (across a number of features) can have a negative effect (Twomey et
177 al., 2013). The impact of variability is also thought to be target dependant such that noun
178 learning might be particularly enhanced by increased variability, but verb learning less so
179 (Gómez, 2002; Maguire et al., 2008). Few studies have investigated the effect of variability in
180 relation to word learning in children with DLD. Alt et al. (2014) found input variability to
181 enhance vocabulary learning in a group of late talking toddlers.

182 **Morphosyntax.** Recasting is an example of a technique that can be embedded in a
183 meaningful activity. Regarding the variability of the input, it is hypothesized that increasing
184 input variability allows children to extract and generalize abstract grammatical rules more
185 easily, informed in part by studies of artificial grammar learning. Work by Plante and
186 colleagues (2014) is one example of the translation of this learning into a therapeutic context
187 and is described later in this review. However, overall, surprisingly few studies have integrated
188 increased variability into treatments for children with DLD.

189 **Phonology:** On the child-centred, clinician-directed continuum, phonological
190 interventions tend to take a more clinician-directed approach. Categories of intervention put
191 forward by Wren et al. (2018) also include an environmental approach, which encompasses
192 phonological interventions that are embedded in everyday interactions, rather than directed
193 activities, that focus solely on change in a child's speech-sound system. Again, the focus here
194 is on those with speech sound disorder rather than DLD.

195

196 **Current study**

197 Systematic reviews of interventions for children with DLD have tended to focus on the
198 overall effectiveness of different treatments and delivery models without looking specifically

199 at dose form (see Cirrin et al., 2010; Cirrin & Gillam, 2008; Law et al., 2004). Dose form has
200 been the point of discussion in two ‘clinical forum’ papers (Eisenberg, 2014; Kamhi, 2014)
201 and was reviewed by Proctor-Williams (2009) solely in the morphosyntax domain. Since then,
202 several informative papers, across domains, have been published. As a result of these
203 significant papers, bringing the importance of dosage to the attention of researchers and
204 clinicians, recently published intervention studies have begun to provide more detailed
205 information on the quantitative aspects of dosage. This development we believed, would allow
206 for a more valid assessment of the *relative efficacy* of different dose forms than had previously
207 been possible. Our aim was to capitalise on this opportunity and conduct a systematic review
208 and narrative synthesis of intervention studies for children with DLD in which aspects of oral
209 language intervention dose form were experimentally manipulated, or retrospectively
210 statistically analysed. The review was to be the second of a pair completed with similar
211 methodology: the first focussing on the quantitative aspects of dosage (see Frizelle et al., 2021).
212 However, the process was not straightforward in relation to dose form.

213 To allow for direct comparisons of different dose forms, we planned to include only
214 studies in which the efficacy of one intervention was compared to an alternative and studies in
215 which dose (number of teaching episodes per session), dose frequency and total intervention
216 duration were either controlled or reported in sufficient detail to allow valid conclusions
217 regarding the effects of dose form to be drawn. However, of those papers that did provide
218 information on the quantitative aspects of dose, there was significant variation in the levels of
219 detail reported, with some papers only providing information on planned dose and others giving
220 information on the discrepancy between what was planned versus what was received. It was
221 also problematic that some papers controlled for one quantitative aspect of dosage (such as
222 dose) but not another (such as dose frequency). As we progressed through the process, it
223 became increasingly evident that there were no papers that purely manipulated dose form

224 (while controlling for all quantitative aspects of dose) and therefore we could not definitively
225 answer the research questions we planned to address.

226 However, although the quality of studies was such that it was difficult to draw definitive
227 conclusions about the relative efficacy of varying dose forms, the process uncovered several
228 key insights with the potential to be of value to the discipline of intervention research for
229 children with DLD. Therefore, the aims of this research paper are to document and disseminate
230 those insights. We present our learning from the papers reviewed under each of the
231 subcategories of dose form described above. We then synthesise our findings and make
232 recommendations regarding key next steps for the systematic accrual of knowledge, necessary
233 in the field of DLD research, to identify optimal dose form for clinical application.

234 This paper is unique in that it systematically addresses the impact of dose form across domains,
235 using the most up to date papers, which increasingly give information on quantitative dose,
236 therefore bringing us closer to uncovering the unique contribution of dose form on intervention
237 effectiveness for children with DLD.

238 **Method**

239 The review was registered with PROSPERO (ID=CRD42017076663) and our methods
240 adhered to PRISMA guidelines for systematic reviews (Moher, 2009). The method is reported
241 in more detail in our review of the quantitative aspects of dosage (Frizelle et al., 2021). More
242 detailed definitions of our included research designs and our definition of intervention are
243 given in the PROSPERO pre-registration (McKean et al., 2017) (ID=CRD42017076663).

244 **Search Procedures**

245 Searches were conducted to identify empirical peer reviewed articles, in any language,
246 relating to oral language interventions with children with DLD, published January 2006 - May

247 2020. The study commenced in 2016 as part of a COST Action 1406 ¹ and so the previous 10
248 years were targeted, to provide a comprehensive but manageable sample of papers to review.
249 The review continued into 2020 and so the searches were updated to include a further 4 years.
250 Seven electronic databases were used: Web of Science (Including Medline, SSCI),
251 MEDLINE(PubMed), ERIC, PsycINFO, Cochrane Library, Scopus, and LLBA. (see Appendix
252 A for search string). Reference lists of the final list of papers included and relevant systematic
253 reviews were also hand searched for additional papers. Inclusion / Exclusion criteria are
254 reported in supplementary material and in Frizelle et al. (2021) with one addition - Included
255 papers experimentally manipulated or statistically analysed dose form whilst either a) keeping
256 other quantitative variables such as dose constant or b) explicitly quantifying dose, dose
257 frequency, total intervention duration and consequently cumulative intervention intensity, to a
258 sufficient degree such that insights about the potential impact of dose form could be inferred.
259 The review focused on interventions in which there were outcomes in the domains of
260 vocabulary, morpho-syntax and phonology.

261 **Paper Selection and Reliability of Search Procedures**

262 Stage 1: The initial search formed the basis of several COST Action IS1406 reviews with
263 differing foci. The aim was to identify papers evaluating interventions for children with DLD
264 across all language domains. These papers were initially screened on title and abstract for
265 inclusion/ exclusion based on the criteria of date, target group, level of evidence (whether there
266 was an element of control included in the study design) or evaluation of an intervention.
267 Twenty percent were double screened by two independent reviewers, using specialist software
268 supporting systematic reviews (EPPI – Reviewer 4). Overall agreement was 96%. All non-
269 English papers at this and subsequent stages were considered by either author AKT (who is

¹ Action 1406 focussed on understanding intervention and service delivery for children with DLD across Europe and a number of partner countries.

270 fluent in a number of languages) or by a native speaker of the relevant language in the COST
271 Action and relevant criteria discussed with the first author after translation. This stage yielded
272 1198 papers.

273 Stage 2: 100% of papers included after stage 1 were screened on title and abstract by two
274 independent reviewers (PF and AKT). However, we did not include pragmatics domain as this
275 was the focus of a different review led by other members of the COST Action. Agreement was
276 93%. This yielded 698 papers. At this stage and each of the subsequent stages, disagreements
277 were discussed and a consensus was reached as per PRISMA guidelines.

278 Stage 3: Full text screening was completed by the same two independent reviewers. Agreement
279 rate was 94%. This resulted in 244 papers.

280 Stage 4: Full text screening was completed to include a) papers with a specific focus on dosage
281 characteristics, which were experimentally manipulated or statistically analysed, b) research
282 design levels 1, 2, or 3 (the Oxford Centre for Evidence Based Medicine Hierarchy of
283 evidence). Agreement rate at this stage was 97% and yielded 39 papers.

284 Stage 5: Full text screening was completed on the above 39 papers and only those that focused
285 specifically on dose form with the necessary controls identified above were included (n=27).

286 See Figure 1 for PRISMA flowchart. Insert Figure 1 about here.

287 **Data Extraction**

288 The data were extracted from the included papers by the first author (PF) and are
289 presented in Table 2. Note that throughout this review we use the term DLD, however the
290 terminology used in Table 2 reflects that which was used in each included paper. All papers
291 were also reviewed by the last author (CMK). For the most part, we used the intention of the
292 authors to categorise the manuscripts. However, in a few incidences we needed to infer the
293 information, for example when categorizing Owen van Horne et al. (2018) (under procedures);
294 Fey et al. 2017 (under intervention context) Haebig et al., 2019 (under intervention context).

295 There were no disagreements with respect to the coding of dose form components or aspects
296 of dose frequency. Insert Table 2 about here.

297 **Risk of Bias**

298 Studies were appraised by the first and last author, using the Cochrane Risk of bias tool
299 for RCTs (Higgins et al., 2011). We assigned risk of bias ratings of high, low or unclear. Both
300 reviewers rated each article independently and disagreements were resolved by consensus. The
301 risk of bias assessment for each paper is shown in Figure 1 in supplementary materials.

302

303

Results

304 Thirty-three studies were identified in which dosage was experimentally or statistically
305 manipulated, 27 of which manipulated dose form. (Table 2). The majority of studies identified
306 children as having DLD² using the following criteria a) a composite score of below 1 standard
307 deviation on a standardized language measure such as the Clinical Evaluation of Language
308 Fundamentals (CELF 4) or the Structured Photographic Expressive Language Test (SPELT-3)
309 b) non-verbal IQ scores within 1 standard deviation of the norm on a test of cognitive
310 functioning (standard score > 85) c) hearing within the normal range (shown by passing a pure
311 tone hearing screening) or no known sensory impairments and d) no known neurological,
312 social-emotional or psychiatric disorders. Interestingly, there was only 1 intervention study
313 (4%) for children with DLD with phonological outcomes, in which dose form was directly
314 manipulated. In contrast there were 12 studies (44%) specific to vocabulary, and 14 (52%)
315 specific to morphosyntax. Thirty-three percent (n = 9) were RCTs, level 2 in the hierarchy of
316 evidence (OCEBM, 2011), 59% (n = 16) were quasi-experimental, (level 3) studies and two
317 studies (7%) were a cohort analytical design (level 3). RCT participant numbers were small,

² Note this is not the term used in all papers, as many papers preceded the move from SLI to DLD as per the CATALISE consensus (Bishop et al. 2016)

318 ranging from six to 30 per group, which raises concerns regarding the interpretation of null
319 findings. The aim of RCTs is to control for differences across groups but with such small
320 sample sizes randomization does not ensure group equity, and is likely to result in biases. There
321 was considerable variation in the risk of bias between papers, with performance bias blinding
322 often impossible to achieve as is the case for most SL interventions. There was a notable
323 increase in controlling for detection bias (i.e. blind outcome measurement) in more recently
324 published studies. Overall, attrition bias and selective reporting were least evident across all
325 papers.

326 **Summary of exploratory findings in relation to dose form**

327 Table 3 presents the dose form comparisons that have been completed across each
328 domain. Insert Table 3 about here.

329 The first finding of note is the distinct lack of intervention studies in which dose form has been
330 manipulated with phonological outcomes for children with DLD, (see Heikkilä et al., 2018). In
331 general studies of Speech Sound Disorder (SSD) tend to exclude children with DLD and vice
332 versa. Given the high comorbidity in these groups and the more negative prognosis in children
333 with a history of both disorders, in particular with respect to literacy (Pennington & Bishop,
334 2009; Eadie et al., 2014), this is an important gap in the necessary evidence to inform clinical
335 practice.

336 There is an inherent difficulty in the interpretation and synthesis of the available
337 evidence due to the high degree in variability of research design across several key
338 methodological choices. Across the discipline of intervention research for children with DLD,
339 there has not been a systematic approach to the examination of the effect of dose form
340 components and how they may interact with age, language level, and other dosage
341 characteristics. In addition, factors such as the nature and timing of outcome measures or the
342 definition of mastery or generalization of a given linguistic skill, have not been methodically

343 examined. In the following we draw out the key learning we feel *can* be identified within this
344 challenging context and discuss lessons which must be learned if future research is to answer
345 key questions regarding optimal dose form.

346 ***Techniques – (The specific actions/ teaching behaviors thought to have benefit.)***

347 **Vocabulary.**

348 ***Semantic and phonological support.*** With respect to vocabulary techniques (the
349 specific actions/ teaching behaviors thought to have benefit), studies have reported on a
350 semantic teaching focus, phonological focus or a combination of the two (e.g. Korat et al.,
351 (2019); Steele et al., (2013)). Overall, the findings suggest that with equivalent doses in a single
352 session, a semantic approach is more beneficial than a phonological or combined approach
353 when learning new vocabulary (Steele et al., 2013). However, in this study learning was
354 measured through tasks which were much more similar to the semantic than the phonological
355 teaching; outcomes were measured immediately post intervention and did not include
356 generalization items; overall dose was low and sample size was small. Therefore, we do not
357 know how treatment effects change with higher doses; if the benefit is maintained over time or
358 is generalizable to other related items; or if the findings would be the same with a larger sample.
359 Within a semantic focus dose forms have included varying levels of support (e.g. dictionary
360 support, giving explanations in the context of the story), the relative effectiveness of which
361 appears to be dependent on the outcome measure used i.e. the clinical goal. While dictionary
362 support appeared to be most effective with respect to word use, providing explanations in the
363 context of the story was more effective in relation to word definition outcomes (Korat et al.,
364 2019). An interaction between semantic supports and children’s language level was also
365 reported in that a combined approach incorporating both definition and context was most
366 effective for those with higher levels of language.

367 **Supplementary gesture manipulation.** There is some evidence to suggest benefits to
368 the use of supplemental gesture with dose held constant. Using comprehension probes as the
369 outcome, administered pre-, during and immediately post intervention, van Berkel-van Hoof
370 et al. (2019) found that children with DLD learned more pseudo-words in a signed than
371 unsigned condition. In addition, Vogt and Kaushke (2017a; 2017b) report on the learning
372 advantage of iconic compared to attention getting gestures across comprehension, naming and
373 word definition outcomes. However, findings are based on target rather than generalization
374 items, and similar to the van Berkel-van Hoof study, outcomes were measured immediately
375 post intervention. Interestingly, the finding of a supplemental gesture advantage was not
376 replicated (one-week post intervention) in an intervention of similar duration (Lüke et al.,
377 2011) but where the dose was lower and the iconicity of the gestures was not specified. The
378 potential effect of the outcome measure timing is particularly pertinent in the context of recent
379 work carried out by Storkel et al. (2019) showing a 40% drop in word learning 5/6 days post
380 intervention.

381 **Morphosyntax.**

382 **Recasts and elicited expression.** With respect to morphosyntax, studies have compared
383 various combinations of techniques (e.g. enhanced conversational recasting versus recasting;
384 cueing versus recasting; prompted elicitation with either recasting or modelling versus
385 recasting alone; and recasting and modelling versus recasting alone). Overall, they suggest that
386 apparently quite subtle differences in the cues and recasts provided could affect intervention
387 efficacy. Hence there are a number of techniques which offer potential for manipulation to
388 improve efficiency. However, across studies the impact of expressive practice has rarely been
389 controlled and so is a potential confound meaning further research is needed for definitive
390 conclusions to be drawn.

391 Eidsvåg and colleagues (2019) manipulated dose form technique by comparing
392 individual morphological error treatment (where children are only exposed to their own target)
393 to treatment carried out in a pair (where children are additionally exposed to their partner's
394 target). Despite hearing their partner's target morpheme modelled during treatment, children
395 in the paired condition showed no significant gains in their ability to produce them. The authors
396 interpreted their findings as an indication that recasts were only effective when directed to a
397 specific child, required the child's active attention and included attempts at morpheme
398 production (i.e. when they were enhanced). The authors acknowledge that there was an
399 element of expressive practice involved for each child when working on their own morpheme,
400 which was absent in the unenhanced recasting condition (in the presence of the paired child's
401 treatment). The authors also counsel caution against over-interpretation due to the study's small
402 sample size. Similarly, Smith-Lock and colleagues (2015) found a cueing hierarchy (that also
403 included recasting) to be more effective than recasting alone in improving grammar production
404 in children with DLD. However, the average age of the cueing group was almost twice that of
405 the recasting group; the cueing condition involved actively eliciting productions whereas the
406 recasting only condition did not; the cueing condition had explicit methods of instruction for
407 some morphemes but not others; and although the authors report on dose, it was extrapolated
408 from a single 15 minute activity from 8 hours of treatment, meaning we cannot rule out the
409 possibility of dose production differences between groups. Findings from Yoder et al. (2011)
410 support the use of prompted elicitation followed by either recasts or models over the use of
411 recasts alone. They found that for children with an MLU of 1.84 (Brown's Late Stage I) the
412 prompted elicitation with modelling/recasting was more effective than recasting alone. Again,
413 it is worth noting that the former approach involves actively eliciting productions whereas the
414 latter does not. It is also interesting to note that although the recasting dose was higher in the
415 recasting only condition, the benefits of expressive practice outweighed the benefits of this

416 increased dose. The potential impact of expressive practice is also evident in the Proctor-
417 Williams and Fey (2007) study, where accuracy of verb production did not differ between the
418 two treatment conditions: recasting alone versus recasting and modelling. It is significant that
419 children produced the target verbs in both dose form conditions. However, the low dose (given
420 at a rate typical of conversation) may also have prevented differences from emerging. Finally,
421 Hassink and Leonard (2010) examined differences in types of conversational recasting, with
422 preschool children with DLD through a retrospective analysis of transcripts of therapeutic
423 recasting sessions. Findings suggest that non-corrective recasts are associated with better
424 morphological outcomes, than those that are corrective (particularly in the use of third-person
425 singular) and subject-less recasts are associated with poorer outcomes than those in which the
426 subject is included. Findings therefore indicate that the quality of clinician recasts, and the
427 relationship between child and clinician utterances, impacted learning.

428 **Phonology.**

429 *Visual supports.* The addition of visual supports (in the form of video) was reported in
430 the only study in which dose form was manipulated with respect to phonological training
431 (Heikkilä et al., 2018). Despite measuring numerous outcomes, the only advantage found was
432 in relation to non-word repetition. It is difficult to draw any firm conclusions based on one
433 study (with small numbers) and while we acknowledge the work that has been done in this area
434 with children with speech sound disorders, much more work is required for children with DLD.

435 *Procedure (the combination and order of treatment delivery)*

436 **Morphosyntax.**

437 *Order effects.* There is emerging evidence of the importance of order effects suggesting
438 they do affect the treatment outcome. Plante et al. investigated whether modelling (in the form
439 of auditory bombardment) given before or after enhanced conversational recast treatment,
440 would result in a greater morphological treatment effect. While there were no differences at

441 group level, individually an increased number of children (86%) were considered ‘treatment
442 responders’ in the auditory bombardment *after* recast condition than the before condition.
443 Comparisons to a previous study with similar dose, Plante et al. 2014, where only 56% of
444 children were treatment responders support the finding that modelling *after* recasting serves to
445 augment its effects, consolidating children’s internal morpheme representations and producing
446 better results than recasting alone.

447 Van Horne and colleagues (2017; 2018) also investigated order effects in their
448 complexity-based approach, which appears to enhance morphosyntactic treatment effects.
449 They found that those who began treatment with harder to inflect verbs first, made greater gains
450 in past tense accuracy on both target and generalisation verbs, than those who were initially
451 treated with verbs that were easier to inflect. Although the ‘hard first’ group advantage was no
452 longer evident at follow up on structured probes, when assessed using spontaneous language
453 samples, they continued to show greater gains both immediately and delayed post treatment.
454 Although the age difference was not statistically significant, it may be worth noting the ‘hard
455 first’ group was on average 9-months older than the easy-first group. The authors conclude that
456 the developmental model (which would seem most intuitive) may not be optimal to achieve
457 generalisation of newly acquired morphosyntactic rules.

458 ***Method of Instruction (manner in which techniques are delivered)***

459 **Morphosyntax.**

460 ***Explicit plus Implicit versus Implicit only.*** Findings from both papers addressing this
461 comparison indicate a learning advantage for explicit instructions in children with DLD with
462 an average age of 7 years. Finestack and Fey (2009) compared a deductive to an inductive
463 approach. We have interpreted the deductive approach as explicit + implicit and the inductive
464 as implicit only but refer the readers directly to the papers for minor differences in how the
465 terms are defined. Findings showed that when compared to the implicit only groups, children

466 in their explicit+ groups used the novel target morpheme across all probes, and did this more
467 quickly with less intervention. What is particularly noteworthy in the context of the current
468 review is that because the implicit only group were slower to learn the patterns than the
469 explicit+ group, they were exposed to a higher recasting dose. Despite this, the explicit
470 instruction appeared to outweigh the benefits of the increased input dose. In the second, more
471 recent paper, while keeping dose constant, Finestack (2018) again compared explicit + implicit
472 versus implicit only, Again, there was a learning advantage for the addition of explicit rule
473 instruction for children with DLD in this age range.

474 *Intervention Contexts*

475 **Vocabulary.**

476 *Activity within which the technique is being delivered.* Dose form manipulation has
477 also been explored with respect to the context / activity within which techniques are being
478 delivered, for example in spoken/ sung format or using static or video images. The findings
479 point up the potential for contexts to interfere as well as facilitate learning. Smeets et al. (2012)
480 carried out two experiments where video and static stories were presented with and without
481 music and sounds. While they found that video and static stories were equally effective, the
482 presence of music and sounds interfered with children's learning in both contexts, such that the
483 interference was greater for children whose DLD was more severe. In contrast, Kouri and Winn
484 (2006) reported that when words and melody were presented as a single unit (as is the case in
485 a song), music appeared not to be detrimental to quick incidental word learning.

486 *Variability in the linguistic input or materials used.* High input variability appears to
487 be advantageous in the vocabulary domain, for children with DLD. The advantage of increased
488 variability is shown in referent exemplars (how an object is represented), language input and
489 changes in context. Findings suggest that increasing exemplar variability in treatment has the
490 potential to improve children's ability to generalize their lexical knowledge and to increase the

491 efficacy of word learning interventions (Aguilar et al., 2018). Aguilar and colleagues (2018)
492 found that with high variability in the referent, preschool children with DLD had the ability to
493 learn three new words having been exposed to them 18 times over three sessions, and asked to
494 name the items once per session. The findings were maintained 6 weeks post intervention. This
495 was in considerable contrast to previous studies where typically only one object exemplar is
496 given (e.g. Alt & Plante, 2006) and where many more presentations were required for learning
497 to occur (e.g. 40 non-word exposures to learn eight non-words (Gray, Pittman & Weinhold,
498 2014) or 36 exposures to learn ~ three words (Storkel et al., 2017), (albeit with different
499 outcome measures). Giving children the opportunity to retrieve word names, also appears to
500 aid word learning with respect to nouns (Leonard et al., 2019a) and adjectives (Leonard et al.,
501 2019b), although overall learning was not high in either of these studies. Interestingly, although
502 the receptive exposure dose was constant, the design of both studies was such that there were
503 expressive dose differences between conditions in each study. The findings therefore suggest
504 that given the same receptive exposure dose children will learn words more easily, if given the
505 opportunity to use rather than just hear them. Word retrieval exercises, in which intervening
506 words are presented (and therefore spaced), also assist word learning and retention more than
507 when the context does not change (Haebig et al., 2019). However, because the intervening
508 material also serves to create a more spaced learning condition, it makes it difficult to tease
509 apart the facilitating factors. The implications of massed versus distributed/ spaced learning is
510 discussed in more detail in our review (Frizelle et al., 2021) on the quantitative aspects of
511 dosage.

512 **Morphosyntax.**

513 *Variability / specificity in the linguistic input.* With the exception of Riches and
514 colleagues (2006), who found that variability of the input did not influence children's
515 morpheme learning (based on only two verbs), other studies suggest that highly variable input

516 in the therapeutic dose form facilitates grammatical morpheme learning, in children with DLD.
517 The theory posited, as to why variability in the input aids learning, is that it helps children
518 extract the morpho-syntactic rules. It is also in keeping with recent morpho-syntactic priming
519 work with younger children with typical language skills (Krok & Leonard, 2018). The positive
520 effect of variability has been shown in the study by Plante et al. (2014), in which a high (24
521 unique verbs once) and low (12 unique verbs twice) variability group were compared. The high
522 variability group was the only group to show a treatment effect (albeit modest) and also showed
523 better generalisation of morpheme use. Building on this work Krzemien and colleagues (2020)
524 suggest that when learning to generalise constructions, gradually increasing variability in the
525 input (through progressive alignment) may be more beneficial for children with DLD, than
526 using maximum variability at the outset. There is also an inherent link between complexity and
527 variability in different dose forms. Children are exposed to many examples of easy verbs in
528 their ambient language input. These core verbs are heard and used frequently and are therefore
529 in keeping with a low variability approach. It is thought that they serve as an initial platform
530 from which children may begin to observe a morphological rule, but that it is only when
531 children are exposed to less common verb exemplars that their representations strengthen, and
532 this helps to generalise the rule to a wider range of contexts. Exposure to less common verbs
533 (many of which meet Van Horne's (2018) definition of hard verbs), automatically increases
534 the variability of the input the child hears across intervention and other contexts, therefore
535 making it difficult to tease apart whether the enhanced treatment effect is driven more by the
536 variability or complexity of the input. Given the number of aspects of verb complexity to
537 consider (relative frequency, phonological complexity and telicity - the completeness of the
538 event described by the verb), we agree with Owen Van Horne and colleagues (2018), who
539 suggest that from a clinical perspective, variability may be easier to operationalise than
540 complexity.

541 One final study, which we found difficult to align with other studies, was carried out
542 by Fey and colleagues (2017). In this study, the linguistic input was specifically manipulated
543 to test the competing sources of input (CSI) hypothesis. The hypothesis proposes that when
544 children use main verbs, not marked for tense, it is because they are treating certain sentence
545 types in their input, as models for declaratives. The authors compared two treatments, one
546 based on the CSI principles and one which did not adhere to these principles (TRAD). A
547 number of techniques were used in both treatments, but it was the specific nature of the
548 linguistic input that was being compared. The CSI group showed greater improvements for one
549 of the three morphemes examined. We refer the reader directly to the paper, for a more
550 comprehensive description of how the input was manipulated.

551

552

Discussion

553 Literature shows that children with DLD can learn from many of the dose form components
554 discussed in this research paper, however our knowledge with respect to how dose form
555 interacts with the quantitative aspects of dosage is in its infancy. We do have some evidence
556 that with careful design of dose form context, changes can occur with lower levels of
557 quantitative dose (see Aguilar et al., 2018). However, this research is just beginning. Hence, it
558 is not yet possible to conclude which are the *optimal and most efficient* dose forms used in
559 interventions. Nevertheless, it is possible to identify some key learning from the review above.
560 With respect to vocabulary, the relative effectiveness of teaching with a phonological versus a
561 semantic focus has not been established and the relative effects of different semantic supports
562 appear to be dependent on the outcome measure. The use of gesture may be advantageous in
563 the short term but we do not know if these effects are maintained. High variability, in relation
564 to how an object is represented, as well as changes in the language input, and context also

565 appear to enhance word learning and may allow for learning to take place with fewer
566 exemplars.

567 Research comparing different dose forms in morphosyntax has been more extensive. However,
568 with respect to techniques, studies have not been sufficiently similar to make any kind of
569 summary statements about their relative effects. It seems that whatever the technique, whether
570 the child gets an opportunity to produce the target has a role to play, in improving outcomes.
571 However, opportunities for expressive practice are rarely tightly controlled and become a
572 confound in a number of studies. The order of techniques (auditory bombardment post
573 recasting) or linguistic targets (hard verbs first) also affect the treatment outcome but have been
574 relatively under-researched. In relation to method of instruction, explicit instruction appears to
575 rarely occur on its own but is beneficial to morphosyntactic learning in children with DLD,
576 when added to an implicit approach. Finally, variability in the linguistic input appears to
577 facilitate grammatical morpheme learning in children with DLD.

578 It is clear from this review that we are a long way from being able to make definitive
579 recommendations, regarding optimal dose forms for interventions for children with DLD. We
580 are yet further away from understanding whether these vary with respect to the child's age,
581 intervention target, or severity or profile of language difficulties. In the following we describe
582 the approach to research we believe is necessary if we are to answer these important questions
583 of relevance to clinical practice, commissioning and funding of services and best outcomes for
584 children with DLD.

585

586 **Recommendations for future research**

587 *Systematic Programme of research*

588 A systematic program of research is required, first, to the manipulation and
589 measurement of dose form, so that effects are measured in isolation. Although clinicians and

590 researchers often combine techniques in the belief that it will yield better outcomes, combining
591 techniques makes it difficult to tease apart their relative effects. Differences in how techniques
592 are operationalised also causes contradictory findings regarding the relative impact of one
593 technique over another. Taking a more systematic approach, we can begin by measuring dose
594 form *techniques* in isolation and directly comparing one technique with another, we can then
595 gradually build knowledge through measuring the effect of *combining techniques* in varying
596 *procedures* and *contexts* and with different *methods of instruction*.

597 A systematic approach is also needed to explore interactions between variability in dose
598 form and quantitative dose. Preliminary findings that dose form variability in both vocabulary
599 and morphosyntax could reduce the dose required to achieve an effect (Aguilar et al., 2018;
600 Plante et al., 2014), have important implications for cost-effectiveness and efficiency and
601 suggest that this would be a worthwhile avenue for future research.

602 This program of research must also examine the impact of different dose form
603 components, across a broader age range, and with differing severity and profiles of language
604 difficulties. Even though children with DLD continue to have difficulties into adulthood, the
605 majority of studies included in this review (with respect to both vocabulary and morphosyntax)
606 focussed on children between 4;0 and 6;0 years. It would also be beneficial to explore any
607 potential interaction between age or language level and dose form. While an interaction
608 between dose form and language level was reported by Korat et al. (2019), Smeets et al. (2012),
609 and Yoder et al. (2011), overall the relationship between the two has been significantly under-
610 researched. A final consideration is whether differing dose form components are more or less
611 effective at different stages in the intervention process. Perhaps explicit methods of instruction
612 are more beneficial at the earlier stages of learning, whilst implicit are more beneficial later,
613 during generalisation and consolidation (Ebbels, 2014).

614 To complete such a systematic and comprehensive program, including all aspects of
615 dose form, it is clear that a large number of studies are required, each building incrementally
616 on the findings and approaches of prior studies. This must be a cross-laboratory endeavour if
617 we are to gain these insights in a timely manner. Our review suggests there are significant
618 inconsistencies across research teams in how specific dose form components are defined,
619 making the collaboration or synthesis across studies described above highly problematic.

620

621 *Operationalizing aspects of dose form*

622 It is essential that a consistent approach to the reporting of dose form components be
623 developed and adopted. Interventions that are often similarly labelled, use different techniques
624 or procedures. For example, Proctor Williams (2009) specifies that modelling involves “the
625 presentation of target syntactic forms, without an opportunity for child production” (p. 297).
626 However, this is not how modelling is used by Yoder and colleagues (2011) where prompting
627 to elicit a target structure was followed by either a model or a recast. More recently Eisenberg
628 et al. (2020) suggest that the term model refers “to any exemplar of a target form presented by
629 the clinician, regardless of whether or not the child is asked to attempt production of the target
630 after hearing the model” (p. 206). Terms such as prompting and cueing are also commonly
631 used and can incorporate a range of techniques (e.g. imitation, questioning, commenting),
632 which are not always specified. There is no agreed regulated practice for each dose form
633 technique and the level of detail in describing techniques is often not sufficient to allow faithful
634 replication. For example, in relation to vocabulary studies, where supplemental gesture is used,
635 the iconicity of the gestures is not always described (see Lüke, et al., 2011). In the context of
636 Open Science this is hugely problematic, as without increased clarity in describing each aspect
637 of dose form, replication will remain limited or flawed. Additionally, it is impossible for
638 clinicians to know that they are implementing the approach with sufficient fidelity to achieve

639 results. In the absence of faithful replications, clinicians and researchers are frequently
640 interpreting results regarding the effectiveness or efficiency of different dose forms, based on
641 a single study and very often with small sample sizes. This is problematic as with such small
642 sample sizes only very large effects can be detected.

643 *A new taxonomy and reporting guidelines*

644 We recommend the development of an agreed taxonomy of dose form components
645 across different domains and the adoption of reporting guidelines across the discipline to aid
646 comparison and application of available evidence. The development of a taxonomy has been
647 carried out in phonological interventions for children with Speech Sound Disorder (Baker et
648 al., 2018) and lessons could be learned from this process for the domains of vocabulary and
649 morphosyntax for children with DLD. With regard to reporting guidelines, for journal articles
650 in which an experimental manipulation or intervention is reported, the JARS (The Journal
651 Article Reporting Standards) (American Psychological Association, 2010) specify eight topics
652 that should be included in an intervention description: intervention content, method of delivery,
653 deliverer, setting, exposure quantity and duration, time span, and activities aimed at increasing
654 compliance or adherence. In addition, the Template for Intervention Description and
655 Replication (TIDieR) checklist and guide has been developed to improve the completeness of
656 reporting and consequently the replicability of interventions (Hoffmann et al., 2014). While
657 these generic guidelines are helpful, given the variation in levels of detail in current reporting
658 on dose form (and indeed other aspects of dosage), it is likely that we need more discipline
659 specific reporting guidelines, to ensure sufficient detail of dose form characteristics are
660 included to move research and practice forward, with regard to optimal dosage characteristics.
661 It is also the case that without an agreed taxonomy detailing each characteristic of dose form,
662 inconsistencies in how different aspects of dose form are implemented will continue to be

663 problematic, and use of reporting guidelines alone would only go a small way towards
664 addressing the problems in the literature.

665 **Quantitative aspects of dosage.** Reporting guidelines are also needed which specify
666 the level of detail required in each of the five dosage characteristics, outlined by Warren et al.
667 (2007). There are four quantitative characteristics - *dose*, *dose frequency*, and *total intervention*
668 *duration*, which are then combined to quantify *cumulative intervention intensity*. Within *dose*
669 (the number of properly administered teaching episodes during a single intervention session)
670 both the receptive (input) and expressive (output) dose need to be specified. Each of the dose
671 subcomponents also need to be addressed a) the average rate of teaching episodes per unit of
672 time b) the length of the intervention session, and c) the distribution/ density of episodes over
673 the session. In addition, planned versus actual received dose should be specified along with a
674 minimum threshold of sessions from which dose can be extrapolated, if data from all sessions
675 is not available. Finally, it is only with this level of detail on quantitative dosage characteristics
676 that we can really begin to compare intervention dose form mechanisms.

677 **Outcome measures - definition and timing.** Reporting guidelines would also be
678 helpful in creating consistency in features of outcome measurement, such as how outcomes are
679 defined, when outcomes are measured, and whether both target and generalisation items are
680 included. In relation to defining outcomes, at a very fundamental level, more debate is required
681 within the discipline as to what constitutes ‘knowing a word’, due to the substantial
682 heterogeneity in outcome measurement for vocabulary outcomes. Results are likely to be
683 dramatically different if ‘knowing a word’ is measured by defining versus naming that word.
684 Choices regarding timing of outcome measurement are also not trivial. Immediate versus
685 delayed testing are measuring two different types of learning: the former measuring children’s
686 ability to encode new information, the latter tapping into the level of decay or consolidation
687 that has occurred. Storkel and colleagues (2019) highlighted this, when reporting a 40% drop

688 in word learning a mere 5 to 6 days post intervention. Other studies, reporting a learning
689 advantage for one dose form over another during an intervention, find no such advantage when
690 follow up measures are taken (e.g. Lüke et al., 2011). Contrary findings have also been
691 reported, where initial encoding for children with DLD appeared weaker than their typically
692 developing peers immediately post intervention, but their retention scores were similar one
693 week later (Leonard et al., 2019a).

694 Testing on target probes versus generalisation items is also significantly different in
695 relation to what children have learned and there are examples in the literature, where no
696 learning advantage is shown, for example, on target items, but is shown when measuring
697 generalisation outcomes, particularly in relation to variability (e.g. Aguilar et al., 2018; Owen
698 Van Horne et al., 2018).

699

700 **Conclusion**

701 Clearly, continuing at the current pace of discovery, and using the small samples
702 represented in much of the work reviewed here, we will not complete the program of research
703 necessary to move the field forward, any time soon. To accelerate Scientific progress, we
704 recommend cross-laboratory co-operation and the adoption of internationally recognised
705 standardised reporting of research methods and intervention taxonomies. In addition, we
706 recommend a culture of open science, where interventions are pre-registered and clearly
707 described; accompanying manuals are made available; and where trial data is shared through
708 open access repositories. In this way study comparison, meta-analysis and data-pooling would
709 be enabled; it would serve to accelerate the rate of discovery within the field, and would
710 maximise the potential for learning. Our paper further develops existing operationalisations of
711 dose form and identifies key aspects that we believe should be encoded in such a taxonomy.
712 The benefits of large collaborative research teams have been shown in other areas of science

713 such as in the field of genetics. We advocate that it is time to change research practice and
714 reporting in the field of speech and language pathology, so that we can expedite the delivery
715 of benefits to the lives of those with DLD, through more effective and efficient intervention
716 delivery.

717

718

References

719 Aguilar, J. M., Plante, E., & Sandoval, M. (2018). Exemplar Variability Facilitates Retention of
720 Word Learning by Children With Specific Language Impairment. *Language, Speech and*
721 *Hearing Services in Schools*, 49(1), 72–84. [https://doi.org/10.1044/2017_LSHSS-17-](https://doi.org/10.1044/2017_LSHSS-17-0031)
722 [0031](https://doi.org/10.1044/2017_LSHSS-17-0031)

723 Alt, M., Meyers, C., Oglivie, T., Nicholas, K., & Arizmendi, G. (2014). Cross-situational
724 statistically based word learning intervention for late-talking toddlers. *Journal of*
725 *Communication Disorders*, 52, 207-220. <https://doi.org/10.1016/j.jcomdis.2014.07.002>

726 American Psychological Association, *APA Style Journal Article Reporting Standards (APA*
727 *Style JARS)*. (2010).

728 <https://apastyle.apa.org/jars#:~:text=APA%20Style%20Journal%20Article%20Reportin>
729 [g%20Standards%20\(APA%20Style%20JARS\)%20are](https://apastyle.apa.org/jars#:~:text=APA%20Style%20Journal%20Article%20Reportin) accessed 18 December 2020.

730 Baker, E., & McLeod, S. (2004). Evidence-based management of phonological impairment in
731 children. *Child Language Teaching and Therapy*, 20(3), 261-285.
732 <https://doi.org/10.1191/0265659004ct275oa>

733 Baker, E., Williams, A. L., McLeod, S., & McCauley, R. (2018). Elements of Phonological
734 Interventions for Children With Speech Sound Disorders: The Development of a
735 Taxonomy. *American Journal of Speech-Language Pathology*, 27(3), 906–935.
736 http://doi.org/10.1044/2018_AJSLP-17-0127

737 Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & CATALISE Consortium.
738 (2016). CATALISE: A multinational and multidisciplinary Delphi consensus study.
739 Identifying language impairments in children. *PLOS ONE*, *11*(7), e0158753.
740 <https://doi.org/10.1371/journal.pone.0158753>

741 Camarata, S. M., Nelson, K. E., & Camarata, M. N. (1994). Comparison of Conversational-
742 Recasting and Imitative Procedures for Training Grammatical Structures in Children With
743 Specific Language Impairment. *Journal of Speech, Language, and Hearing*
744 *Research*, *37*(6), 1414–1423. <https://doi.org/10.1044/jshr.3706.1414>

745 Cirrin, F., & Gillam, R. (2008). Language Intervention Practices for School-Age Children With
746 Spoken Language Disorders: A Systematic Review. *Language, Speech, and Hearing*
747 *Services in Schools*, *39*(1). [https://doi.org/10.1044/0161-1461\(2008/012\)](https://doi.org/10.1044/0161-1461(2008/012))

748 Cirrin, F. M., Schooling, T. L., Nelson, N. W., Diehl, S. F., Flynn, P. F., Staskowski, M.,
749 Torrey, T. Z., & Adamczyk, D. F. (2010). Evidence-Based Systematic Review: Effects
750 of Different Service Delivery Models on Communication Outcomes for Elementary
751 School–Age Children. *Language, Speech, and Hearing Services in Schools*, *41*(3), 233–
752 264. [https://doi.org/10.1044/0161-1461\(2009/08-0128\)](https://doi.org/10.1044/0161-1461(2009/08-0128))

753 Connell, P. J., & Stone, C. A. (1992). Morpheme learning of children with specific language
754 impairment under controlled instructional conditions. *Journal of Speech & Hearing*
755 *Research*, *35*(4), 844–852. <https://doi.org/10.1044/jshr.3504.844>

756 Courtright, J. A., & Courtright, I. C. (1979). Imitative modeling as a language intervention
757 strategy: The effects of two mediating variables. *Journal of Speech & Hearing*
758 *Research*, *22*(2), 389–402. <https://doi.org/10.1044/jshr.2202.389>

759 Eadie, P., Morgan, A., Ukoumunne, O., Ttofari Eecen, K., Wake, M., & Reilly, S. (2014). Speech
760 sound disorder at 4 years: prevalence, comorbidities, and predictors in a community

761 cohort of children. *Developmental Medicine & Child Neurology*, 57(6), 578-584.
762 <https://doi.org/10.1111/dmcn.12635>

763 Ebbels, S. (2014). Effectiveness of intervention for grammar in school-aged children with
764 primary language impairments: A review of the evidence. *Child Language Teaching and*
765 *Therapy*, 30(1), 7-40. <https://doi.org/10.1177/0265659013512321>

766 Eidsvåg, S. S., Plante, E., Oglivie, T., Privette, C., & Mailend, M.-L. (2019). Individual Versus
767 Small Group Treatment of Morphological Errors for Children With Developmental
768 Language Disorder. *Language, Speech, and Hearing Services in Schools*, 50(2), 237–252.
769 https://doi.org/10.1044/2018_lshss-18-0033

770 Eisenberg, S. (2014). What Works in Therapy: Further Thoughts on Improving Clinical Practice
771 for Children With Language Disorders. *Language, Speech, And Hearing Services In*
772 *Schools*, 45(2), 117-126. https://doi.org/10.1044/2014_lshss-14-0021

773 Eisenberg, S. L., Bredin-Oja, S. L., & Crumrine, K. (2020). Use of Imitation Training for
774 Targeting Grammar: A Narrative Review. *Language, Speech and Hearing Services in*
775 *Schools*, 51(2), 205–225. http://doi.org/10.1044/2019_LSHSS-19-00024

776 Fey, M. E. (1986). *Language intervention with young children*. College Hill Press.

777 Finestack, L. H. (2018). Evaluation of an Explicit Intervention to Teach Novel Grammatical
778 Forms to Children With Developmental Language Disorder. *Journal of Speech Language*
779 *and Hearing Research*, 61(8), 2062–2075. [https://doi.org/10.1044/2018_JSLHR-L-17-](https://doi.org/10.1044/2018_JSLHR-L-17-0339)
780 [0339](https://doi.org/10.1044/2018_JSLHR-L-17-0339)

781 Finestack, L. H., & Fey, M. H. (2009). Evaluation of a deductive procedure to teach grammatical
782 inflections to children with language impairment. *American Journal of Speech-Language*
783 *Pathology*, 18(3), 289–302. [https://doi.org/10.1044/1058-0360\(2009/08-0041\)](https://doi.org/10.1044/1058-0360(2009/08-0041))

784 Frizelle, P., Tolonen, A.-K., Tulip, J., Murphy, C.-A., Saldana, D., & McKean, C. (2021). The
785 Influence of Quantitative Intervention Dosage on Oral Language Outcomes for Children

786 With Developmental Language Disorder: A Systematic Review and Narrative Synthesis.
787 *Language, Speech, and Hearing Services in Schools*, 1–17.
788 https://doi.org/10.1044/2020_lshss-20-00058

789 Gillon, G. (2000). The Efficacy of Phonological Awareness Intervention for Children With
790 Spoken Language Impairment. *Language, Speech, And Hearing Services In*
791 *Schools*, 31(2), 126-141. <https://doi.org/10.1044/0161-1461.3102.126>

792 Gómez, R. L. (2002). Variability and Detection of Invariant Structure. *Psychological Science*,
793 13(5), 431–436. <https://doi.org/10.1111/1467-9280.00476>

794 Haebig, E., Leonard, L. B., Deevy, P., Karpicke, J., Christ, S. L., Usler, E., Kueser, J. B., Souto,
795 S., Krok, W., & Weber, C. (2019). Retrieval-Based Word Learning in Young Typically
796 Developing Children and Children With Development Language Disorder II: A
797 Comparison of Retrieval Schedules. *Journal of Speech, Language, and Hearing Research*,
798 62(4), 944–964. https://doi.org/10.1044/2018_jslhr-1-18-0071

799 Hassink, J. M., & Leonard, L. B. (2010). Within-treatment factors as predictors of outcomes
800 following conversational recasting. *American Journal of Speech-Language Pathology*,
801 19(3), 213–224. [https://doi.org/10.1044/1058-0360\(2010/09-0083\)](https://doi.org/10.1044/1058-0360(2010/09-0083))

802 Heikkilä, J., Lonka, E., Meronen, A., Tuovinen, S., Eronen, R., Leppänen, P. H., Richardson, U.,
803 Ahonen, T., & Tiippana, K. (2018). The effect of audiovisual speech training on the
804 phonological skills of children with specific language impairment (DLD). *Child*
805 *Language Teaching and Therapy*, 34(3), 269–287.
806 <https://doi.org/10.1177/0265659018793697>

807 Higgins, J. P. T., Altman, D. G., Gotzsche, P. C., Juni, P., Moher, D., Oxman, A. D., Savovic, J.,
808 Schulz, K. F., Weeks, L., & Sterne, J. A. C. (2011). The Cochrane Collaboration’s tool
809 for assessing risk of bias in randomised trials. *BMJ*, 343(oct18 2), d5928–d5928.
810 <https://doi.org/10.1136/bmj.d5928>

811 Hodson, B. W. (2015). Cycles phonological patterns approach. In C. Bowen (Ed.), *Children's*
812 *speech sound disorders* (2nd ed, pp. 36-40). Wiley-Blackwell.

813 Hoffmann, T. C., Glasziou, P. P., Boutron, I., Milne, R., Perera, R., Moher, D., Altman, D. G.,
814 Barbour, V., Macdonald, H., Johnston, M., Lamb, S. E., Dixon-Woods, M., McCulloch,
815 P., Wyatt, J. C., Chan, A.-W. & Michie, S. (2014). Better reporting of interventions:
816 template for intervention description and replication (TIDieR) checklist and guide. *BMJ*,
817 *348*(mar07 3), g1687–g1687. <https://doi.org/10.1136/bmj.g1687>

818 Kamhi, A. (2014). Improving Clinical Practices for Children With Language and Learning
819 Disorders. *Language, Speech, and Hearing Services in Schools*, *45*(2), 92-103.
820 https://doi.org/10.1044/2014_lshss-13-0063

821 Korat, O., Graister, T., & Altman, C. (2019). Contribution of reading an e-book with a dictionary
822 to word learning: Comparison between kindergarteners with and without DLD. *Journal*
823 *of Communication Disorders*, *79*, 90–102. <https://doi.org/10.1016/j.jcomdis.2019.03.004>

824 Kouri, T. A., & Winn, J. (2006). Lexical learning in sung and spoken story script contexts.
825 *Child Language Teaching and Therapy*, *22*(3), 293–313.
826 <https://doi.org/10.1191/0265659006ct309xx>

827 Krzemien, M., Thibaut, J.-P., Jemel, B., Levaux, E., & Maillart, C. (2020). How do children
828 with developmental language disorder extend novel nouns? *Journal of Experimental*
829 *Child Psychology*, *202*, 105010. <https://doi.org/10.1016/j.jecp.2020.105010>

830 Krok, W. C., & Leonard, L. B. (2018). Verb Variability and Morphosyntactic Priming With
831 Typically Developing 2- and 3-Year-Olds. *Journal of Speech Language and Hearing*
832 *Research*, *61*(12), 2996–3009. https://doi.org/10.1044/2018_JSLHR-L-17-0410

833 Law, J., Garrett, Z., & Nye, C. (2004). The Efficacy of Treatment for Children With
834 Developmental Speech and Language Delay/Disorder. *Journal of Speech, Language, and*
835 *Hearing Research*, *47*(4), 924-943. [https://doi.org/10.1044/1092-4388\(2004\)069](https://doi.org/10.1044/1092-4388(2004)069)

836 Leonard, L. B., Camarata, S. M., Pawlowska, M., Brown, B., & Camarata, M. N. (2006). Tense
837 and agreement morphemes in the speech of children with specific language impairment
838 during intervention: Phase 2. *Journal of Speech, Language, and Hearing Research, 49*,
839 749–770. [https://doi.org/10.1044/1092-4388\(2006/054\)](https://doi.org/10.1044/1092-4388(2006/054))

840 Leonard, L. B., Karpicke, J., Deevy, P., Weber, C., Christ, S., Haebig, E., Souto, S., Kueser, J.
841 B., & Krok, W. (2019a). Retrieval-Based Word Learning in Young Typically
842 Developing Children and Children With Developmental Language Disorder I: The
843 Benefits of Repeated Retrieval. *Journal of Speech, Language, and Hearing*
844 *Research, 62*(4), 932–943. https://doi.org/10.1044/2018_jslhr-1-18-0070

845 Leonard, L. B., Deevy, P., Karpicke, J. D., Christ, S., Weber, C., Kueser, J. B., & Haebig, E.
846 (2019b). Adjective Learning in Young Typically Developing Children and Children
847 With Developmental Language Disorder: A Retrieval-Based Approach. *Journal of*
848 *Speech, Language, and Hearing Research, 62*(12), 4433–4449.
849 https://doi.org/10.1044/2019_jslhr-1-19-0221

850 Lüke, C., Rohlfing, K., & Stenneken, P. (2011). Gebärden und kommunikative Mitteilung bei
851 Kindern mit umschriebener Sprachentwicklungsstörung. *Sprache · Stimme · Gehör*,
852 35(04), e149–e157. <https://doi.org/10.1055/s-0031-1287816>

853 Maguire, M. J., Hirsh-Pasek, K., Golinkoff, R. M., & Brandone, A. C. (2008). Focusing on the
854 relation: fewer exemplars facilitate children’s initial verb learning and extension.
855 *Developmental Science, 11*(4), 628–634. [https://doi.org/10.1111/j.1467-](https://doi.org/10.1111/j.1467-7687.2008.00707.x)
856 [7687.2008.00707.x](https://doi.org/10.1111/j.1467-7687.2008.00707.x)

857 Marulis, L., & Neuman, S. (2010). The Effects of Vocabulary Intervention on Young Children’s
858 Word Learning. *Review of Educational Research, 80*(3), 300-335.
859 <https://doi.org/10.3102/0034654310377087>

860 Masterson, J. J., & Perrey, C. D. (1999). Training Analogical Reasoning Skills in Children With
861 Language Disorders. *American Journal of Speech-Language Pathology*, 8(1), 53–61.
862 <https://doi.org/10.1044/1058-0360.0801.53>

863 McKean, C., Frizelle, P., Tolonen, AK., Kunnari, S., Murphy, CA., Saldana, D., Law, J., & Josie
864 Tulip, (2017). Interventions for children with primary language impairment: A COST
865 IS1406 systematic review of the influence of dosage in interventions addressing
866 vocabulary, phonology and morphosyntax outcomes. PROSPERO CRD42017076663
867 Available
868 from: https://www.crd.york.ac.uk/prospéro/display_record.php?ID=CRD42017076663

869 Moher, D. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The
870 PRISMA Statement. *Annals of Internal Medicine*, 151(4), 264.
871 <https://doi.org/10.7326/0003-4819-151-4-200908180-00135>

872 Nicholas, K., Alt, M., & Hauwiller, E. (2018). Variability of input in preposition learning by
873 preschoolers with developmental language disorder and typically-developing
874 language. *Child Language Teaching and Therapy*, 35(1), 55-74.
875 <https://doi.org/10.1177/0265659019830455>

876 OCEBM Levels of Evidence Working Group*. “The Oxford Levels of Evidence 2”.
877 Oxford Centre for Evidence-Based Medicine. <https://www.cebm.net/index.aspx?o=5653>

878 Owen Van Horne, A. J., Curran, M., Larson, C., & Fey, M. E. (2018). Effects of a Complexity-
879 Based Approach on Generalization of Past Tense – ed and Related Morphemes.
880 *Language, Speech and Hearing Services in Schools*, 49(3S), 681–693.
881 https://doi.org/10.1044/2018_LSHSS-STLT1-17-0142

882 Pennington, B., & Bishop, D. (2009). Relations Among Speech, Language, and Reading
883 Disorders. *Annual Review of Psychology*, 60(1), 283-306.
884 <https://doi.org/10.1146/annurev.psych.60.110707.163548>

885 Perry, L. K., Samuelson, L. K., Malloy, L. M., & Schiffer, R. N. (2010). Learn locally, think
886 globally: Exemplar variability supports higher-order generalization and word learning.
887 *Psychological Science*, 21(12), 1894–1902. <https://doi.org/10.1177/0956797610389189>

888 Plante, E., Ogilvie, T., Vance, R., Aguilar, J. M., Dailey, N. S., Meyers, C., Lieser, A.M., &
889 Burton, R. (2014). Variability in the Language Input to Children Enhances Learning in a
890 Treatment Context. *American Journal of Speech-Language Pathology*, 23(4), 530–545.
891 https://doi.org/10.1044/2014_AJSLP-13-0038

892 Plante, E., Tucci, A., Nicholas, K., Arizmendi, G. D., & Vance, R. (2018). Effective Use of
893 Auditory Bombardment as a Therapy Adjunct for Children With Developmental
894 Language Disorders. *Language, Speech, and Hearing Services in Schools*, 49(2), 320–
895 333. https://doi.org/10.1044/2017_lshss-17-0077

896 Proctor-Williams, K. (2009). Dosage and distribution in morphosyntax intervention: Current
897 evidence and future needs. *Topics in Language Disorders*, 29(4), 294–311.
898 <https://doi.org/10.1097/TLD.0b013e3181c29dc0>

899 Proctor-Williams, K., & Fey, M. E. (2007). Recast density and acquisition of novel irregular past
900 tense verbs. *Journal of Speech Language and Hearing Research*, 50(4), 1029–1047.
901 [https://doi.org/10.1044/1092-4388\(2007/072\)](https://doi.org/10.1044/1092-4388(2007/072))

902 Proctor-Williams, K., Fey, M. E., & Loeb, D. F. (2001). Parental Recasts and Production of
903 Copulas and Articles by Children With Specific Language Impairment and Typical
904 Language. *American Journal of Speech-Language Pathology*, 10(2), 155–168.
905 [https://doi.org/10.1044/1058-0360\(2001/015\)](https://doi.org/10.1044/1058-0360(2001/015))

906 Riches, N. G., Faragher, B., & Conti-Ramsden, G. (2006). Verb schema use and input dependence
907 in 5-year-old children with Specific Language Impairment (SLI). *International Journal*
908 *of Language & Communication Disorders*, 41(2), 117–135.
909 <https://doi.org/10.1080/13682820500216501>

910 Smeets, D. J. H., van Dijken, M. J., & Bus, A. G. (2012). Using Electronic Storybooks to Support
911 Word Learning in Children With Severe Language Impairments. *Journal of Learning*
912 *Disabilities*, 47(5), 435–449. <https://doi.org/10.1177/0022219412467069>

913 Smith-Lock, K. M., Leitao, S., Lambert, L., Prior, P., Dunn, A., Cronje, S., Newhouse, S., &
914 Nickels, L. (2015). The effectiveness of two grammar treatment procedures for children
915 with DLD: A randomized clinical trial. *Language, Speech and Hearing Services in*
916 *Schools*, 46(4), 312–324. https://doi.org/10.1044/2015_LSHSS-14-0041

917 Steele, S. (2020). Vocabulary Intervention: A National Survey of School-Based Speech–
918 Language Pathologists. *Communication Disorders Quarterly*, 41(3), 151-161.
919 <https://doi.org/10.1177/1525740119827008>

920 Steele, S. C., Willoughby, L. M., & Mills, M. T. (2013). Learning word meanings during reading:
921 Effects of phonological and semantic cues on children with language impairment.
922 *International Journal of Speech-Language Pathology*, 15(2), 184–197.
923 <https://doi.org/10.3109/17549507.2012.700322>

924 Storkel, H. L., Komesidou, R., Pezold, M. J., Pitt, A. R., Fleming, K. K., & Romine, R. S. (2019).
925 The Impact of Dose and Dose Frequency on Word Learning by Kindergarten Children
926 With Developmental Language Disorder During Interactive Book Reading. *Language,*
927 *Speech and Hearing Services in Schools*, 50(4), 518–539.
928 http://doi.org/10.1044/2019_LSHSS-VOIA-18-0131

929 Storkel, H. L., Voelmle, K., Fierro, V., Flake, K., Fleming, K. K., & Romine, R. S. (2017).
930 Interactive Book Reading to Accelerate Word Learning by Kindergarten Children With
931 Specific Language Impairment: Identifying an Adequate Intensity and Variation in
932 Treatment Response. *Language, Speech and Hearing Services in Schools*, 48(1), 16–30.
933 https://doi.org/10.1044/2016_LSHSS-16-0014

934 Swisher, L., Restrepo, M., Plante, E., & Lowell, S. (1995). Effect of Implicit and Explicit "Rule"
935 Presentation on Bound-Morpheme Generalization in Specific Language
936 Impairment. *Journal Of Speech, Language, And Hearing Research*, 38(1), 168-173.
937 <https://doi.org/10.1044/jshr.3801.168>

938 Twomey, K. E., Ranson, S. L., & Horst, J. S. (2013). That's More Like It: Multiple Exemplars
939 Facilitate Word Learning. *Infant and Child Development*, 23(2), 105–122.
940 <https://doi.org/10.1002/icd.1824>

941 van Berkel-van Hoof, L., Hermans, D., Knoors, H., & Verhoeven, L. (2019). Effects of Signs
942 on Word Learning by Children With Developmental Language Disorder. *Journal of*
943 *Speech, Language, and Hearing Research*, 62(6), 1798–1812.
944 https://doi.org/10.1044/2019_jslhr-1-18-0275

945 Van Horne, A. J. O., Fey, M., & Curran, M. (2017). Do the Hard Things First: A Randomized
946 Controlled Trial Testing the Effects of Exemplar Selection on Generalization Following
947 Therapy for Grammatical Morphology. *Journal of Speech, Language, and Hearing*
948 *Research*, 60(9), 2569–2588. https://doi.org/10.1044/2017_jslhr-1-17-0001

949 Vogt, S., & Kauschke, C. (2017a). Observing iconic gestures enhances word learning in typically
950 developing children and children with specific language impairment. *Journal of Child*
951 *Language*, 44(6), 1458–1484. <https://doi.org/10.1017/S0305000916000647>

952 Vogt, S. S., & Kauschke, C. (2017b). With Some Help From Others' Hands: Iconic Gesture Helps
953 Semantic Learning in Children With Specific Language Impairment. *Journal of Speech*
954 *Language and Hearing Research*, 60(11), 3213–3225.
955 https://doi.org/10.1044/2017_JSLHR-L-17-0004

956 Warren, S., Fey, M., & Yoder, P. (2007). Differential treatment intensity research: A missing link
957 to creating optimally effective communication interventions. *Mental Retardation And*

958 *Developmental Disabilities Research Reviews*, 13(1), 70-77.

959 <https://doi.org/10.1002/mrdd.20139>

960 Wren, Y., Harding, S., Goldbart, J., & Roulstone, S. (2018). A systematic review and
961 classification of interventions for speech-sound disorder in preschool
962 children. *International Journal Of Language & Communication Disorders*, 53(3), 446-
963 467. <https://doi.org/10.1111/1460-6984.12371>

964 Yoder, P. J., Molfese, D., & Gardner, E. (2011). Initial Mean Length of Utterance Predicts the
965 Relative Efficacy of Two Grammatical Treatments in Preschoolers With Specific
966 Language Impairment. *Journal of Speech Language and Hearing Research*, 54(4), 1170-
967 1181. [https://doi.org/10.1044/1092-4388\(2010/09-0246\)](https://doi.org/10.1044/1092-4388(2010/09-0246))

968 Zeng, B., Law, J., & Lindsay, G. (2012). Characterizing optimal intervention intensity: The
969 relationship between dosage and effect size in interventions for children with
970 developmental speech and language difficulties. *International Journal of Speech-
971 Language Pathology*, 14(5), 471-477. <https://doi.org/10.3109/17549507.2012.720281>

972

973 Figures

974 Figure 1. Prisma Flowchart showing literature search process.

975 Supplemental Files

976 Figure 1. Critical Appraisal of each included study, using Cochrane Risk of Bias Tool.

977

978 Table 1. Components of Dose Form and their Definitions

Techniques	The specific actions/ teaching behaviors thought to have benefit
Procedures	The combination and order, of technique delivery

Method of Instruction	The manner in which techniques are delivered i.e. with or without explicit instruction (Explicit versus Implicit)
Intervention contexts	<p>This has 3 sub-components</p> <ul style="list-style-type: none"> • The activity within which the technique/teaching behaviour is being delivered • Where the activity sits in the child-centred, clinician-directed continuum • The degree of variability / uniformity in the linguistic input or materials used.

979

980 Table 3. Dose form comparisons that have been completed across each domain.

981

982

Aspect Manipulated	Domain		
	Vocabulary	Morphosyntax	Phonology
Techniques	Manipulation of Semantic supports	Modelling versus Enhanced conversational recast treatment	Audio-visual speech training versus auditory only
	Manipulation of learning supports – phonological; semantic; combined phonological-semantic; and control.	Dose form in keeping with Competing Sources of Input Hypothesis versus traditional approach not including these features (using story modelling, retell and recasting)	
	Supplementary gesture - signs with speech versus signs alone	Retrospective analysis of differences in types of conversational recasting	
	Supplementary gesture - pseudoword with/without iconic signs	Modelling + recasting versus modelling alone	
	Supplementary gesture - iconic versus attention getting gestures	Recasting versus cueing	
		Grammatical recasting versus prompting followed by a recast or model	

Procedures		Enhanced conversational recast treatment preceded or followed by auditory bombardment	
		Manipulation in the order of verb presentation, (easy to hard, or hard to easy).	
Method of Instruction		Implicit only versus explicit and implicit combined.	
		Explicit and implicit combined versus implicit only- determined by type of auditory prompts given	
Intervention Context	Variability of the physical representation of objects used for each referent in word learning (high versus low variability)	Variability of the linguistic input (high variability versus progressive alignment in exposures to novel construction)	
	Retrieval practice with contextual changes versus immediate retrieval without any intervening linguistic material.	Variability of the linguistic input (high versus low variability in conversational recast treatment).	
	Retrieval practice with contextual changes versus repeated study with no retrieval practice	Variability of the linguistic input (modelling with a noun/pronoun versus noun only frame, in subject /object slots)	
	Using e-books, words illustrated with either static or video images with / without the addition of music and sounds in the video condition		
	Use of story scripts to facilitate word learning, given in sung versus spoken form		

983

984

Figure 1. PRISMA flowchart showing literature search process.

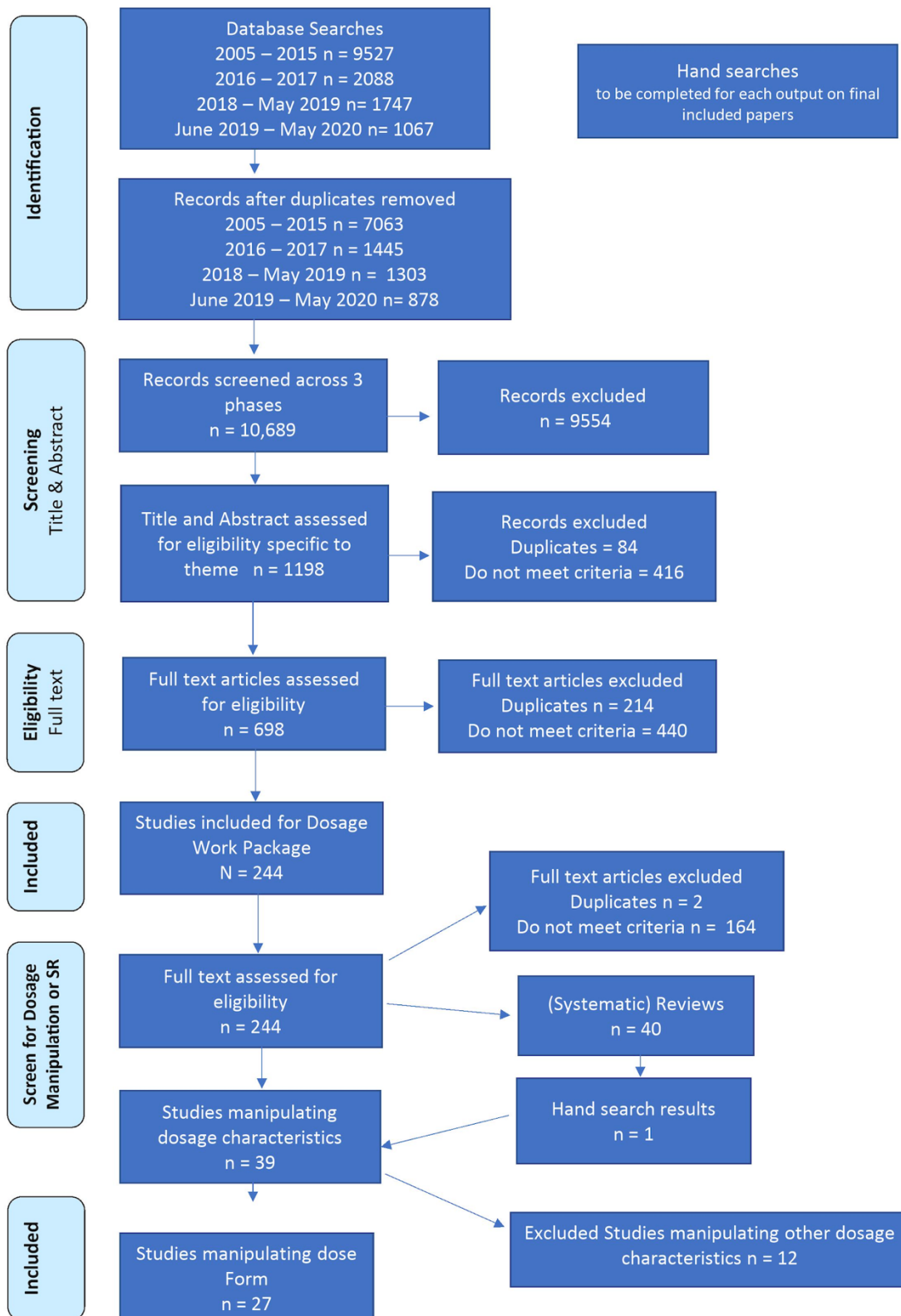


Figure 1. Critical appraisal of each included study.

Aguilar 2018	N/A	N/A	?	+	+	+	-	+
Eidsvag 2019	N/A	N/A	-	+	-	+	?	+
Fey 2017	+	?	-	+	?	+	-	+
Finestack 2018	+	+	+	+	+	+	-	+
Finestack & Fey 2009	+	+	+	+	+	+	-	+
Haebig 2019	N/A	N/A	+	-	+	+	-	+
Hassink 2010	N/A	N/A	-	+	+	+	-	+
Hiekkila 2018	N/A	N/A	?	+	+	+	-	-
Korat 2019	N/A	N/A	+	-	+	+	-	-
Kouri 2006	N/A	N/A	-	-	?	+	-	-
Krzemien 2020	?	?	+	+	+	+	-	+
Leonard 2019a	N/A	N/A	+	-	+	+	-	+
Leonard 2019b	N/A	N/A	+	-	+	+	-	+
Luke 2011	?	?	-	-	?	+	-	-
Owen Van Horne 2018	+	-	+	+	+	+	?	+
Plante 2014	?	?	-	-	+	+	+	-
Plante 2018	-	-	+	-	+	-	?	+
Proctor-Williams 2007	?	?	-	?	+	+	-	+
Riches 2006	N/A	N/A	-	-	?	+	-	-
Smeets 2014 Experiment 1	N/A	N/A	-	-	+	+	?	?
Smith-Lock 2015	N/A	N/A	-	+	?	+	+	-
Steele 2013	N/A	N/A	-	+	-	+	-	+
van Berkel-van Hoof 2019	N/A	N/A	+	-	-	+	-	?
Van Horne 2017	+	-	+	-	+	+	?	+
Vogt & Kauschke 2017a	N/A	N/A	-	+	?	+	-	-
Vogt & Kauschke 2017b	N/A	N/A	-	+	?	+	-	?
Yoder 2011	+	+	-	-	-	+	+	+
	Selection random sequence generation	Selection allocation concealment	Performance bias blinding	Detection bias blinding	Attrition bias incomplete data	Selective reporting	Other- Power	Other- Fidelity Measures

Table 2. Summary of included intervention studies, with vocabulary, morphosyntax or phonology outcomes, in which aspects of *dose form* were manipulated.

Study		Participants (Intervention and comparison)		Treatment					Outcomes		
Aspect Manipulated:Techniques Domain: Vocabulary											
Study	Study design	Number	Age (M, SD)	Dose Form / Intervention context	Treatment /control targets	Dose Form manipulation	Planned/ Received dose	Planned/ received Dose frequency	Intervention duration and session length	Nature and timing of measures	Main findings
1. Korat, O., Graister, T., & Altman, C. (2019).	Quasi Experimental	40 children. 20 with SLI; 20 with typical development All children assigned to each condition	Mean age of 5.58 years (SD = .51)	Word exposure with short dictionary explanation , word exposure with explanation given in the context of the story, or a combination of the previous two. Using animations representing each word meaning – in e-books.	Comprehension and production of 9 verbs	Manipulation of semantic supports for new word learning in e-book activity	Each dose form was given once per session (i.e. short explanation, explanation in story context, combination of the two).	5 sessions at 1-3 day intervals	2 – 3 weeks (sessions 10 minutes in length).	Receptive and expressive probe measures of the target dictionary words (receptive word knowledge test; word definition task; word use task) Administered pre and post (post 1–3 days after the last intervention session)	No significant effect of support between groups. Type of dictionary support did not affect receptive word learning. Dictionary support was the most effective with respect to word use for children with SLI. Explanations in context resulted in the best word definition outcomes. The combined definition approach most

											effective for those with pre-intervention higher levels of language.
2. Lüke, C., Rohlfing, K., & Stenneken, P. (2011)	RCT	12 children with SLI and a semantic lexical disorder (6 children in each group)	Whole group 3;01 to 5;09, M = 57,05 months; SD = 11;08 months) Gesture group median 57,05 months Control group median 63;05	Word exposures in the context of play. Prosodic emphasis and semantic elaboration. In the gesture group, the therapist used signs and speech simultaneously when naming the figures.	9 novel pseudowords (presented as names of comic figures / 'new friends')	New word learning using simultaneous sign and speech or signs alone	SLT named the figures 12 times session 1 16 times session 2, 18 times session 3	Once weekly	3 weeks (45 minutes per session, including testing)	Picture naming procedure and receptive word learning procedure developed to test active word learning during and post treatment. Pretesting (1 week before the intervention). Naming tests after each intervention session; posttesting 1 week after the intervention had concluded.	No difference between groups in active words learned expressively or receptively at post-test. However, there was a statistical trend in favour of the gesture group.
3. Steele, S. C., Willoughby,	Quasi Experimental	12 children with LI from	LI group = 10;03 years	Word exposures through:	20 words—presented in one of four	Semantic word learning	Control Treatment one exposure	One Tx session (the whole study)	One session (length of	Dynamic word learning assessment	Results showed: Children with LI performed

L. M., & Mills, M. T. (2013)		low income backgrounds 11 children with typical language All children assigned to each condition	(9.32 months) TL group = 10;08 years (6.73months)	Phonological segmentation and blending tasks (modelled by the therapist and then completed by the child). Semantic condition - student friendly definitions, and use of word associations / synonyms by therapist and child. Combined approach-phonological and semantic elements as described.	learning conditions, before reading passages in which the words were embedded.	task in 4 conditions – phonological; semantic; combined phonological-semantic; and control.	Phonological Tx = 5 receptive and 2 expressive Semantic Tx = 5 receptive and 2 expressive Combined Tx + 9 receptive and 4 expressive	took place over 4 sessions)	session not specified)	i.e. Ability to define the words (with response dependant contextual clues) immediately post intervention. If contextual clues did not help children were asked multiple choice questions.	significantly better only in the semantic condition relative to the control condition. Despite hearing the words more frequently in the phonological condition performance was similar to the control the Combined condition performance was similar to the semantic even though words were more frequently presented in the combined.
4. van Berkel-van Hoof, L., Hermans, D., Knoors, H., & Verhoeven, L. (2019).	Quasi Experimental	40 children with DLD 26 children with typical development All children assigned to	Children with DLD M= 10;10 (SD = 7.51 months). Age range from 9;3 to 12;2	Pseudo-word and iconic sign exposures using pre-recorded video clips and pictures	Comprehension of 20 pseudo words	Pseudo-word exposures with and without iconic signs One alien from each pair presented in	4 exposures per session (in 4 20 trial blocks)	4 sessions a week 3 were training sessions (1,2 and 3)	1 week, 20 minutes per session	Comprehension probes and response time measures administered pre- during and immediately post-	Children with DLD 1)learned more words with sign than without (mean difference of sign vs no-sign = 1.14.

		each condition	Children with typical development M = 9;11 (SD = 9.80 months), Age range from 8;10 to 12;2	pictures are of paired aliens with one distinguishing feature. Words presented in the context of a carrier phrase 'Look an X' in training and 'where is the X' in testing		the signed condition and one in the no-sign condition		3 testing sessions (2,3 and 4)		intervention to assess children's accuracy and speed of understanding target pseudo-words	2)responded as quickly to words with as without a sign. 3) learned significantly fewer words than TD children (M= 4.81 v's 5.98) No significant effect of sign for TD children
5. Vogt, S. S., & Kauschke, C. (2017a).	Quasi Experimental	60 children – 20 children with SLI 20 TD children matched on chronological age (AM) 20 language matched children (LM) All children exposed to each condition	SLI M= 4;06 years SD = 0;7 AM 4;05 years SD 0.3 LM M= 3;03 years SD;16	12 German words (nouns and verbs) for the AM group. 8 for the LM and SLI groups. Nouns represented rare animal species, verbs were intransitive and represented unusual movement types	Word exposures in the context of a story accompanied by either an Iconic or attention directing gesture.	Manipulation of two gesture conditions (iconic versus attention getting)	Children were exposed to each target word twenty times before the first learning assessment and fifty-seven times before the second (post-test).	3 training sessions over 1-2 week period (2 to 3 days apart)	One to 2 weeks (30 minute sessions, 90 minutes in total)	(2017a) Learning was assessed through target focused naming and comprehension tasks. Measurements taken pre, during (immediately after the first training session) and post intervention (two to three days after	For all three groups, observing the iconic co-speech gestures improved children's comprehension and naming of words to a greater degree than observing attention-directing gestures. There was no particular benefit for

6. Vogt, S., & Kauschke, C. (2017b).										completion of the training) (2017b) As above with the addition of a word definition task	children with SLI. Children defined more words trained with iconic gestures than with attention getting gestures. Children with SLI showed increased semantic knowledge on words taught with the iconic rather than attention getting gestures.
Aspect Manipulated: Techniques											
Domain: Morphosyntax											
7. Eidsvåg, S. S., Plante, E., Oglivie, T., Privette, C., & Mailend, M.L. (2019)	Quasi Experimental	20 children with DLD (10 per group)	Given for group as a whole M = 5;06, (SD = 5.8 months)	Enhanced conversational recast method. Focused recasts targeting one morpheme per child throughout treatment	Production of third-person singular form; past -ed; auxiliary; infinitive; yes/no questions.	Modelling versus Enhanced conversational recast treatment	Children in paired treatment group heard 24 unique recasts per session directed toward themselves and 24 that were directed	5 times per week (25 days). Treatment days received ranged from 22 to 25	5 weeks (30 minutes per session)	use of target and control morphemes on an elicitation probe and correct use of the child's treatment partner's target	Positive treatment effects shown for both conditions. Children in the paired condition showed no significant gains in their ability to produce their

							to the other child. Children treated individually heard 24 unique recasts per session directed solely to themselves.			morpheme where appropriate. Pre, during and 8 weeks post treatment	partner's target morpheme However moderate to large effect sizes emerged, favouring individual treatment re greater spontaneous use of target morphemes and group treatment re children tending to learn their partner's ambient morpheme. Different results may have emerged with a larger sample.
8. Hassink, J. M., & Leonard, L. B. (2010).	Quasi experimental	17 children with SLI (input analysed for all children)	Mean age not reported. Age range 3;03 – 4;04	1) recasts following child utterances that were prompted by clinicians 2) clinicians' recasts of subject-less	Use of 3 rd person singular	Variability in conversational recasting with respect to specific features of the input (retrospective analysis)	12 recasts per session (.8 recasts per minute) 1,152 recasts containing the target form.	four sessions per week	24 weeks (96 treatment sessions). Length of session not stated.	3 rd person singular elicitation probes administered mid (after 48 sessions) and post (after 96 sessions) treatment.	Clinicians' use of non-corrective recasts were the most positive predictor of both short and long term gains in the use of 3 rd person singular. Recasts of

				sentences, 3) clinicians' noncorrective recasts.							subject-less sentences made a small contribution to children's performance but were associated with poorer outcomes.
9. Proctor-Williams, K., & Fey, M. E. (2007).	Cohort Analytical (treatment words randomly assigned)	26 - 13 children with SLI 13 younger TD participants	SLI group - 7;10 years TD group - 5;6 years	Recasts - in the context of a play based activity	Novel verb learning (6 verbs). Syntactically all verbs were transitive, causative and telic. Phonologically all were single-syllable verbs that marked tense with a vowel shift.	Modelling + recasting v's modelling alone	Planned – Total dose of 30 recasts in both density conditions. Low density = .2 per min High density condition = .5 per minute (no recasts in the first three sessions, last 2 sessions included 5 irregular past tense recasts for each of the three high-density verbs Received - low density .19 per min; High density .47.	Distribution of sessions was not tightly controlled - substantial range in the number of days from the first to the fifth experimental sessions (4 to 44, respectively; M = 14 days; SD = 8.95).	Duration in weeks not specified. 2 Training sessions of 31 minutes - 5 experimental sessions - average 31 minutes.	During the intervention - Correct spontaneous productions of irregular past tense novel verbs in obligatory contexts in Sessions 4 and 5. Post intervention - the number of correct irregular past tense verb productions (maximum 12 per condition)	No difference in accuracy of verb production whether recasts were included as part of the dose form or not.

10. Smith-Lock, K. M., Leitao, S., Prior, P., & Nickels, L. (2015)	Quasi Experimental with cluster randomization	31 children with SLI (17 recasting procedure, 14 cueing procedure)	Recasting group = 119.06 months (SD 8.47) cueing group 60.78 months (SD 3.51)	<p>Adult modelling and child production of targets. Correct child responses treated with modelling correct target back and non-specific feedback. Incorrect child responses treated with either recasting or cueing (which also included recasting but which is not explicitly stated in the paper).</p> <p>Structured activities implemented in small groups</p>	Recasting group: 10 children targeted possessive -s, 3 targeted regular past-tense marker -ed, 4 targeted regular third person singular marker -3s. Cueing group: 8 children targeted possessive -s, 1 targeted regular past-tense marker -ed, and 5 targeted regular third person singular marker -3s.	Recasting versus cueing (differing only in the adults response to the child's error)	<p>Received Dose: Total models heard during one 15-min treatment activity were calculated for each group. Mean number of models heard similar in both groups Recasting group = 86.22 (SD 20.05); Cueing group = 87.67 (SD = 52.92). Extrapolating this figure across the whole program results in cumulative intervention intensity of 2,069 models for recasting and 2,104 for cueing group. Mean number of target utterances per child in the recasting group was 10.08 (SD</p>	once weekly	8 weeks (1 hour sessions - whole class 15 minutes, 3 activities 15 minutes each)	Grammar elicitation test (administered 4 times). Gain between Tests 1 and 2 (pre-treatment gain) compared with gains made between Tests 2 and 3. Test 4, 8 weeks post.	Cueing group made significantly more progress than the recasting group – with a medium–large treatment effect in the cueing group and a negligible effect size in the recasting group. No group differences in maintenance of treatment effects 8 weeks post treatment.
--	---	--	---	---	---	--	---	-------------	--	--	---

							= 3.16) and in the cueing group was 14.29 (SD = 7.05). Difference not significant.				
11. Yoder, P. J., Molfese, D., & Gardner, E. (2011)	RCT	57 children with SLI (MLT -Milieu language teaching condition - 27 BTR - Broad target recasts condition - 30)	Mean ages not given. 6 of the children were under 36 months. Range 30 - 60 months	For MLT defined as correct teaching episode: adult prompt to elicit from child, followed by an adult production of targeted structure in either recast or model form BTR: prompts to elicit platform utterances, followed by grammatical recasts	MLT- Researchers selected three targets from a language sample on the basis of the absence of examples of specific structures from a list of potential targets typically present in children with the target child's MLU. Targets were replaced when children used three non-imitative examples of the structure in treatment sessions. BTR; The BTR approach does	Grammatical recasting (Broad target recasts BTR) versus prompting followed by a recast or model (Milieu language teaching MLT)	Received: Rate per minute of grammatical recast for BTR (M = 4.3, SD = 0.74). Rate of correct teaching episode for MLT (M = 3.2, SD = 0.57) Dose (per min) was greater for BTR than for MLT $t(55) = 5.6, p < .001, d = 1.65$. Average proportion of teaching episodes correctly implemented (per min) in the MLT group =	Planned: Three sessions per week (72 sessions) Received: Attendance 96% (BTR M = 68.9, SD = 3.1; MLT M = 68.6, SD = 3.2) of sessions offered.	6 months (24 weeks) 30 minutes per session.	Two 20 minute language samples were collected at 6 time points. Study entry, during the intervention. Immediately post treatment and two follow up maintenance measures (final sample 4 months post). IPSyn score used to reflect syntactic and morphological structures used.	Results revealed a cut point regarding which treatment was most effective. Despite lower MLT dose, for children with an MLU of 1.84 MLT was superior to BTR in facilitating grammatical development. For children with higher MLU, both treatments yielded similar responses.

					not target preselected grammatical structures and uses prompts to elicit children's platform utterances The BTR clinicians recast any 'recastable' child utterance that afforded modelling a developmentally appropriate grammatical structure.		.89 (SD = .17). Average proportion of opportunities (per min) recast in the BTR group = .80 (SD = .07).				
--	--	--	--	--	--	--	--	--	--	--	--

Aspect Manipulated: Techniques

Domain: Phonology

12. Heikkilä, J., Lonka, E., Meronen, A., Tuovinen, S., Eronen, R., Leppänen, P. H., Richardson, U., Ahonen, T.,	Quasi Experimental	20 children with SLI (10 in each group)	Group mean age 8 years 9 months in the audio-visual training group, 9 years 1 month in the auditory training group.	Word exposures without noise (50%), with non-speech noise (25%) and with babble noise (25%). Children were required to extract different pieces of	Development of Phonological skills.	Audio-visual speech training v's auditory only	11 blocks of word-picture tasks, each containing 40 words. (440) 10 blocks of word-letter tasks, each containing 40 words (400). 6 blocks of word-syllable	5 times a week	6 weeks - 10 to 15 minutes per session (between 5 and 7.5 hours)	Neuro-psychological and behavioural assessments completed Pre and post training NEPSY-II Phonological Processing NEPSY Repetition of	Both groups improved in their phonological processing skills. No post treatment differences in how the two groups performed overall.
--	--------------------	---	---	--	-------------------------------------	--	--	----------------	--	--	--

Tiippana, K.(2018).				information from the exposures (word, syllable, phoneme). Intervention delivered through use of Neuro-behavioural systems software			tasks, each containing 40 words. (240). One session consisted of one or two blocks, depending on the response speed of the participant.			Nonsense Words	Repetition of nonsense words the only outcome measure on which there was a significant difference between the two groups, in favour of the audio-visual condition.
Aspect Manipulated: Procedure											
Domain: Morphosyntax											
13. Owen Van Horne, A. J., Curran, M., Larson, C., & Fey, M. E. (2018).	RCT	20 children with DLD (2 groups – 10 easy first, 10 hard first)	easy first 64.7 months (23.76) hard first 70.8 months (22.77)	Sentence imitation, corrective feedback, observational modelling, morpheme exposures within a syntax story and recasts during play based focused stimulation.	Use of past tense morpheme - ed in 30 target verbs	Complexity based approach, manipulation in the order of verb presentation, easy to hard, or hard to easy.	Planned per session: 10 sentence imitation items (two per verb) with corrective feedback. 25 observational modelling examples (five per verb; A minimum of 15 exposures (three per verb); 15 to 25 regular past tense –ed recasts (three	One to three visits per week (scheduling family dependant)	A maximum of 36 visits regardless of progress made. Range 12 – 36. Session length not explicitly stated, approx.- imately 30 minutes based on timing of session components.	Children’s performance on past tense –ed assessed using structured probes and narrative retells, prior to intervention immediately after intervention and 6–8 weeks after intervention.	On structured probes, the hard group first advantage (2017) no longer evident at follow up. In spontaneous language samples, hard group first showed greater gains post treatment and at follow up. No generalisation to untreated morphemes.

							to five per verb)				
14. Plante, E., Tucci, A., Nicholas, K., Arizmendi, G. D., & Vance, R. (2018)	Quasi Experimental	28 children with DLD (14 in each group)	5;03 years (range 4;03 to 6;02 years)	Enhanced conversational recast and auditory bombardment	Expressive morphological targets (chosen on the basis of morpheme use less than 30% of the time in obligatory contexts)	Enhanced conversational recast treatment preceded or followed by auditory bombardment	48 treatment doses per session. 24 presented as recasts, and 24 occurred during auditory bombardment. Cumulative intervention intensity 408-600 across children m=552	Once per day, 5 times per week	5 weeks (total intervention duration of 25 days) 30 minutes per session.	Learning tracked through baseline and generalization probes (given 3 times a week). And spontaneous use of target morphemes during treatment	No significant difference between Bombardment First and Last conditions on: morpheme use in probes; spontaneous morpheme use; unique utterances containing target morphemes. More children showed response to treatment in bombardment last condition.
15. Van Horne, A. J. O., Fey, M., & Curran, M. (2017).	RCT	18 children with DLD (10 easy group first, 8 hard group first)	Easy first 63.1 months (23.02) Hard first 72.75 months (18.48)	Sentence Imitation, structured prompting, observational modelling, focussed stimulation and recasting Therapy was play based	Use of past tense morpheme -ed in 30 target verbs	Complexity based approach, manipulation in the order of verb presentation, easy to hard, or hard to easy.	First Tx visit Structured prompting and sentence imitation resulted in children hearing 20–30 examples of the past-tense -ed (four to six	One to three visits per week (scheduling dependant on family)	A maximum of 36 visits regardless of progress made. Range 12 – 36. Session length not explicitly stated, approx. -imately 30	Sentence completion probes were administered pre- post and during treatment. Probes also used puppet shows and obligated the use of the	Gains in target verb accuracy significantly greater for hard-first group. If analysis included trained verbs only then no group differences.

				and through the use of syntax stories.			examples per verb). Those who didn't meet 80% criterion got 2nd Tx visit where they heard 65- 85 models or recasts at a rate of 0.6 to 1.0 recasts/min during the focused stimulation portion of the treatment sessions		minutes based on timing of session components.	past tense. Both treatment and generalisation sets of verbs were measured.	Gains in generalization verb accuracy significantly greater for the hard-first group Hard group first also made greater gains on all untreated verbs. No differences in time in therapy or progress made on the verbs that were targeted during intervention.
--	--	--	--	--	--	--	--	--	--	--	--

Aspect Manipulated: Method of Instruction
Domain: Morphosyntax

16. Finestack, L. H. (2018).	RCT	25 children with SLI. Implicit only 13 Explicit Implicit combined 12	EI group M = 6.77 years (SD .66) IO group M = 7.35 years (SD 2.46)	Computer presentation Explicit= models, elicitation using cloze procedures, feedback + explicit instruction regarding the rule guiding use of the	acquisition, maintenance, and generalization of the production of 3 different novel grammatical forms (relating to gender, aspect and person)	Implicit and explicit method of instruction/ procedure	morpheme count - 25 for the gender morpheme, 40 for the aspect (habitual action) and 38 for the first person morpheme	5 sessions over a 9 day period (range = 7.08 days to 10.58 days to complete teaching sessions)	Each participant completed up to five teaching sessions (20 minutes long) for each of the three grammatical targets. Max 15 *20 = 300 min or 5	Acquisition, maintenance and generalisation probes. Pre, post and follow up. Follow up (generalisation probe) 1 week post intervention	A significant learning advantage for the E-I group on acquisition, maintenance, and generalization probes when performance was combined across the 3 targets.
------------------------------	-----	---	---	--	---	--	---	--	--	--	---

				novel target form. Implicit = as above but explicit instruction replaced by a filler statement.					hours. Or taking the average 13.5*20 =270 minutes or 4 1/2 hours		Re individual targets significant differences emerged in favour of the E-I group on gender morpheme only Controlling for dose, explicit instruction enhanced morphological learning.
17. Finestack, L.H., & Fey, M.H. (2009)	RCT	32 children with SLI (16 in each group)	Deductive group M = 87.94 months SD 7.71. Inductive group M = 88.31 months SD 8.32	Computer presentation Deductive: Modelling of structure, auditory prompting, production of morpheme (using sentence cloze techniques) followed by recasting if production incorrect. Auditory	Novel marking of subject verb agreement, limited to 3 rd person singular BE in English	Deductive (explicit) v's inductive (implicit) types of auditory prompts given	20 models of the structure 10 auditory prompts to produce the morpheme Number of recasts following incorrect production dependant on child's response (not reported) Generalisation probe 30 items	4 sessions intended to occur within a two week period. Mean number of days between individual treatment sessions for both groups combined was 2.75 (range = 1–11; Deductive: M = 2.73,	The mean number of days required to complete all four treatment sessions for both groups was 9.25 (range = 4–16 days; Deductive: M = 9.19, range = 4–16; Inductive: M = 9.31, range = 4–16).	Teaching probe, generalization probe and maintenance probes were administered during and immediately post treatment	More children in the deductive group successfully used the novel morpheme in the teaching probe (10 v's 3), the generalization probe (10 v's 3), and the maintenance probe (7 v's 2) - despite fewer recasts in deductive group

				prompt was rule based Inductive: as above, but rule based prompt was replaced by a filler prompt				range = 1–11; Inductive: M = 2.77, range = 1–7).	Session length not stated.		
Aspect Manipulated: Intervention Context											
Domain: Vocabulary											
18. Aguilar, J. M., Plante, E., & Sandoval, M. (2018)	Quasi-Experimental	18 children with SLI (9 per group)	High variability group 59.4 months (5.7) No variability group 57.2 months (4.3)	Word exposures – through the presentation of physical objects (with and without variability)	Comprehension of 8 target vocabulary words	The variability of the physical representation of objects used for each referent.	6 exposures per session / procedural fidelity for number of presentations = 95%	3 training sessions once weekly	6 weeks including baseline and outcome measures. 3 weeks intervention. 12.5 minutes average per session	Comprehension probes administered during and after intervention) designed to measure generalisation (ability to identify the object using new object exemplars of the same class)	High Variability group correctly identified more generalization items than the No Variability group at retention test, but not during intervention.
19. Haebig, E., Leonard, L. B., Deevy, P., Karpicke, J., Christ,	Quasi-Experimental	32 children (16 typically developing and 16 with DLD – all children assigned to	TD M = 61.58 months, SD = 5.16; DLD M = 59.60 months, SD = 4.43	Computer presentation Study trial: Word exposure, word	Word learning – 12 novel words (exotic plants and animals)	Retrieval practice with contextual changes (RRCR) versus immediate retrieval	Planned - The total number of exposures of each novel word was 24, and each word meaning was	Intervention carried out over 4 days (2 X 2 day periods)	Each 2 day intervention period was 2 weeks apart. 2 X 10 minute	Word form recall and meaning recall (e.g. What's this called? What does this one like?)	Although RRCR had a reduced expressive dose - Children with DLD achieved an average score of 2.5 points

S.L., Usler, E., Kueser, J.B., Souto, S., Krok, W., & Weberb, C. (2019).	two word learning conditions)	definition i.e. what the object liked, in the presence of a picture referent, Retrieval trial: picture prompts to recall word name and definition. For words in the RRCC condition, retrieval trials for a word occurred only after two other words had been presented. For words in the IR condition, all retrieval trials immediately followed a study trial of the same word.	without any intervening linguistic material (IR)	heard eight times. Each word form and meaning had six retrieval opportunities. Received - Number of exposures constant for each condition Number of words produced were more frequent in the IR schedule (n = 1,033) than in the RRCC schedule (n = 587).	blocks each day. 40 minutes for each set of 6 words. 80 minutes total	Post treatment and follow up. Follow up testing one week post 2 day intervention. Additional form-referent link recognition test at follow up.	higher in the RRCC condition than in the IR condition. Findings indicate that word retrieval exercises in which there are intervening words presented, assist word learning and retention more than repeatedly retrieving and producing a word with no contextual change.				
20. Leonard , L. B.,	Quasi-Experimental	10 children with DLD	Children with DLD M= 63.4	Computer presentation	Retention of 8 novel	Retrieval practice with	Each word in each condition	Treatment completed	4 sessions over a two	Recall (of word form	Both groups showed better

<p>Karpicke, J., Deevy, P., Weber, C., Christ, S., Haebig, E., Souto, S., Keueser, J.B. & Krok, W. (2019).</p>		<p>10 children with typical development</p> <p>All children assigned to both conditions</p>	<p>months, SD = 6.2</p> <p>Children with typical development M = 63.2 months, SD = 4.89</p>	<p>Study trial format – word exposures in the presence of a picture referent and in the context of a 3 sentence sequence in which the novel word name and definition were integrated.</p> <p>Retrieval trial format - Child shown picture and heard a pre-recorded question what's this called (name) and what does this one like? (definition)</p> <p>RRCR condition followed a</p>	<p>consonant-vowel – consonant word forms and their meaning – nouns (targeted as two sets of 4 words</p> <p>Words taken from Storkel and Lee (2011) matched on phonotactic probability and neighbourhood density</p>	<p>contextual changes (RRCR) versus repeated study with no retrieval practice (RS)</p> <p>First retrieval for each word immediately followed the study trial (no other novel words intervening).</p> <p>Subsequent retrieval in the RRCR condition occurred after 3 intervening novel words</p>	<p>was heard and defined 48 times. (3 sentence sequence X 16 trials)</p>	<p>over 2 consecutive days for each set of 4 words – 1 week apart</p>	<p>week period. Session length approximately 25 minutes to include a 5 minute break (2 blocks of 10 minutes on day 1 and 2 blocks of 10 minutes on day 2)</p>	<p>and definition) and multiple choice comprehension task.</p> <p>Both administered 5 minutes post the 4th intervention block and 1 week later.</p>	<p>word form recall following the RRCR condition than in the RS condition – large effect size</p> <p>All but one child with DLD recalled more words in RRCR than in RS condition.</p> <p>Both groups showed better word meaning recall following the RRCR condition than in the RS condition</p> <p>Despite weaker initial encoding, 1 week post intervention there were no differences in retention between the DLD and TD groups</p>
--	--	---	---	--	--	---	--	---	---	--	--

				study-retrieval - study protocol							
				RS condition involved 3 consecutive study trials							
21. Leonard , L. B., Deevy, P., Karpicke, J. D., Christ, S., Weber, C., Kueser, J. B., & Haebig, E (2019)	Quasi-Experimental	14 children with DLD 13 children with typical development All children assigned to both conditions	Children with DLD M= 62.64 months, SD = 5.41, range 53 – 71 months Children with typical development M = 62.54 months, SD = 6.34, range 51 – 71 months	Computer presentation Study trial format – word exposures in the presence of picture referents and in the context of a 3 sentence sequence in which the novel word is integrated. Retrieval trial format - Child shown picture and heard a pre-recorded request to give the	Recall and recognition of 8 novel adjectives (targeted as two sets of 4 words) Generalisation of novel adjectives to new referents	Retrieval practice with contextual changes (RRCR) versus repeated study with no retrieval practice (RS) First retrieval for each word immediately followed the study trial (no other novel words intervening). Subsequent retrieval in the RRCR condition occurred after 3 intervening novel words	44 exposures across two days of learning	Treatment completed over 2 consecutive days for each set of 4 words – 1 week apart	4 sessions over a two week period. Session length approximately 25 minutes to include a 5 minute break (2 blocks of 10 minutes on day 1 and 2 blocks of 10 minutes on day 2)	Recall (of word form, applied to referents used in the treatment and generalisation items). Multiple choice comprehension task. Recall assessed 5 minutes post the second session and 1 week later Recognition / multiple choice task assessed 1	For both groups of children, recall was higher for adjectives learned in the RRCR condition than in the RS condition at both time points. Large effect for DLD group For recognition the DLD group showed greater accuracy for adjectives learned in the RRCR condition than those in the RS condition.

				<p>appropriate adjective to complete the sentence, as in "Tell me about the cat. The cat is very ____.</p> <p>RRCR condition followed a study-retrieval - study protocol (with the exception of the first four items for which there were two study trials)</p> <p>RS condition involved a single study trial (with the exception of the first four items for which there were two study trials)</p>						week post intervention	No effect of condition for recognition in the TD group
--	--	--	--	--	--	--	--	--	--	------------------------	--

22. Smeets, D. J. H., van Dijken, M. J., & Bus, A. G. (2012)	Quasi Experimental	29 children with SLI All children assigned to each condition	(M = 69.34 months SD = 5.92)	Word exposures through electronic stories (static and video with music and sounds).	42 words (half of which were nouns)	Using e-books, words illustrated using either static or video images with the addition of music and sounds in the video condition.	4 exposures per word	Twice weekly	Intervention was carried out over two periods - each lasting 4 weeks (8 weeks in total). Length of session 20 minutes by implication but not specified.	Pre and post intervention, knowledge of target words was assessed using a sentence completion task.	Children learned 23% of words targeted in the intervention. Static books were more effective than those using video.
Experiment 1											
Experiment 2	Quasi Experimental	23 children with SLI	(M = 71.56 months SD = 7.15)	Word exposures through electronic stories (static and video with and without music and sounds).	72 words	Using e-books, words illustrated using either static or video images Both conditions with and without music and sounds.	4 exposures per word	Twice weekly	24 weeks (72 sessions) 30 minutes per sessions	Target vocabulary test as above. (post was 3 to 4 days after the last story exposure) with the addition of the CELF 4 NL, nonword repetition task and digit span from the CELF 4.	Video and static stories were equally effective in children's word-learning. Music and sounds interfered with children's learning in both contexts The effect was moderated by children's language ability.
23. Kouri, T. A., & Winn, J. (2006).	Cohort Analytical	16 children with SLI (12 of which were	M= 4;01 years. Range 3;06 to 5;01	Word exposures through story- telling	Comprehension production and generalisation of 8 novel nouns	Use of story scripts given in sung or spoken form	Words modelled 5 times through the story. All 5	2 Sessions within five days of each other	1 Week (50 – 60 minutes long)	Lexical production, comprehension and	No significant main effects or interactions between

		considered mildly development-ally delayed). These children meet the new criteria for DLD. All children assigned to each condition.		and act out. Character manipulation (by children) during story telling.			exposures were spoken or 2 were spoken and 3 were sung.			generalization probes were administered in each treatment session (8 other objects closely resembling the target items were constructed for use in the generalization probes)	numbers of target or generalization items comprehended in sung and spoken conditions in either session. More spontaneous initiations of novel lexical items produced in the second than first session in the sung condition only.
--	--	--	--	---	--	--	---	--	--	---	---

Aspect Manipulated: Intervention Context

Domain: Morphosyntax

24. Krzemiński, M., Seret, E., & Maillart, C. (2020)	RCT	30 children with DLD 30 children with TD matched on a measure of sentence comprehension Children randomly assigned to	Children with DLD M= 9,11, SD = 1;08 Language matched children M = 7;07, SD = 2;01)	Exposure to the novel construction in two conditions (high or progressive variability) associated with a video in which figurines performed	Comprehension and generalisation of a novel construction of the form NP- NP- V (the first NP is the character who appears, the second NP is the place where it appears and the V specifies how	Variability of the input High variability condition – sentences had no words in common Progressive alignment condition – the second, third, fourth,	Constructions presented in 8 sentences, twice per video - 16 exposures	A single training session	Single session, approximately 15 minutes in duration	Generalisation video probes administered immediately post intervention (6 constructed according to the novel construction, 6 transitive sentences) All sentences had no words	Children in both groups performed better in the progressive alignment condition than in the high variability condition – no significant differences between groups
--	-----	---	--	---	--	---	--	---------------------------	--	---	--

		one of two conditions		various actions	the character appears)	and fifth sentences had words in common, sentences became progressively distinct.				in common with the stimuli used in the training phase.	Both groups performed better on the transitive than novel construction. For children with DLD, performance on the novel construction was at chance in progressive alignment condition and below chance in the high variability condition.
25. Fey, M. E., Leonard, L. B., Bredin-Oja, S. L., & Deevy, P. (2017)	RCT	20 children with SLI (11 TRAD group, 9 CSI group)	45 months	Story modelling, retell and recasting with 1) no competing sources of input 2) ½ to 1/3 presented in competing interrogative form	Production of 3rd person singular and auxiliary is	A dose form in keeping with Competing Sources of Input Hypothesis versus traditional approach not including these features.	CSI intervention: 10 trial <i>is/was</i> comprehension game 12 declarative <i>is</i> models 8 declarative <i>is</i> recasts 10 trial <i>does/did</i> comprehension game (tense contrast focus) 12 declarative 3s models	2 sessions per week	12 weeks – 30 to 40 minutes per session	Morpheme production and comprehension probes which were administered pre-treatment, at midpoint (after 6 weeks) and post intervention (after a	In keeping with the CSI hypothesis the CSI group showed greater gains in their use of <i>is</i> (with a large effect size). Contrary to the CSI hypothesis there were no significant group differences in the production of 3s. This was

							8 declarative 3S recasts TRAD intervent: 10 trial <i>is/was</i> comprehension game 6 declarative <i>is</i> models, 6 interrogative <i>is</i> models 4 declarative <i>is</i> recasts, 4 interrogative <i>is</i> recasts 10 trial <i>does/did</i> comprehension game (semantic focus) 12 declarative 3S models, 6 interrogative <i>does</i> models 8 declarative 3s, 4 interrogative <i>does</i> recasts Actual received very close to what was planned			further 6 weeks).	also the case for the control morpheme <i>-ed</i> .
--	--	--	--	--	--	--	---	--	--	----------------------	---

26. Plante, E., Ogilvie, T., Vance, R., Aguilar, J. M., Dailey, N. S., Meyers, C., Lieser, A.M., & Burton, R. (2014)	RCT	18 children with language impairment 9 per group	low variability group M = 5;02, high variability M = 4;11.	Recasting	morphological deficits	Manipulation of variability in the linguistic input - high versus low variability in conversational recast treatment.	24 for both treatment conditions (high and low variability) Cumulative intervention intensity 408-600 across children m=552	Planned: 5 times per week Received: 20 to 25 sessions. Low variability group mean 23 days, high variability mean 22.9 days	5 weeks and 2 days. 30 minutes per session	Target and control morphological probes, total number of inflected verb tokens and verb types used spontaneously in treatment. Probes administered pre, during and post treatment.	Only those in high variability condition showed significant change in their use of target v's control morphemes. More children showed strong treatment effect in the high variability condition. High variability group spontaneously produced significantly more inflected verb types.
27. Riches, N. G., Faragher, B., & Conti-Ramsden, G. (2006)	Quasi Experimental	24 children with SLI matched to 23 typically developing children on overall language abilities.	SLI - mean age 5;6 years TD - mean age 3;5 years	Modelling act out task, questioning (elicitation question to use verb in any frame, elicitation question to use verb in transitive frame)	Generalising of verbs from a non-transitive to a transitive frame.	Modelling with either a noun/pronoun or noun only frame, in the subject /object slots)	Planned and received - 12 verb exposures during each session. (36 times over the three training sessions). In total, there were 216 verb presentations (36x6).	3 intervention sessions (frequency not specified)	Intervention duration and session length were not specified.	Probe in which the children heard the novel verbs in a non-transitive frame and were required to produce them in a transitive frame.	The likelihood that the children generalized the novel verb to a transitive frame was not dependant on the frame used during the training sessions.

				Play based treatment							<p>As expected more pronoun responses were produced in the mixed condition.</p> <p>The children with SLI showed a greater degree of input dependence than the TD group</p>
--	--	--	--	----------------------	--	--	--	--	--	--	--