The Efficacy of an Explicit Intervention Approach to Improve Past Tense Marking for Early School-Aged Children with Developmental Language Disorder

Samuel D Calder[1], Mary Claessen[1], Susan Ebbels[2, 3] and Suze Leitão[1]

Author Affiliations: [1]School of Occupational Therapy, Social Work, and Speech Pathology, Curtin University, Perth, Western Australia, [2] Moor House Research and Training Institute, Moor House School & College, Oxted, United Kingdom, [3]Language and Cognition, University College London, UK

Cite as:

Calder, S. D., Claessen, M., Ebbels, S., & Leitão, S. (2021). The efficacy of an explicit intervention approach to improve grammar for early school-aged children with DLD. *Journal of Speech, Language, and Hearing Research*, *64*(1), 91-104. https://doi.org/10.1044/2020_JSLHR-20-00132

Address correspondence to:

Samuel D Calder

School of Occupational Therapy, Social Work, and Speech Pathology, Curtin University, Perth, Western Australia, GPO Box U1987, 6845.

Email: samuel.calder@postgrad.curtin.edu.au

Conflict of Interest Statement

Components of the intervention approach described in this article were developed by co-author, Susan Ebbels, whose employer (Moor House School & College) sells courses and resources relating to the SHAPE CODING™ system.

Funding

The authors acknowledge the contribution of an Australian Government Research Training Program Scholarship in supporting this research.

Abstract

Purpose: The aim of this study was to evaluate the efficacy of a theoretically motivated explicit intervention approach to improve regular past tense marking for early school-aged children with developmental language disorder (DLD).

Method: Twenty one children with DLD (ages 5;9 – 6;9 years) were included in a crossover randomized controlled trial (intervention, $n = 10$; waiting control, $n = 11$). Intervention included once weekly sessions over 10 weeks using the SHAPE CODING^{TM} system in combination with a systematic cueing hierarchy to teach past tense marking. Once the first group completed intervention, the waiting control group crossed over to the intervention condition. The primary outcome was criterion-referenced measures of past tense marking with standardized measures of expressive and receptive grammar as the secondary outcome. Ancillary analyses on extension and behavioural control measures of morphosyntax were also conducted.

Results: There was a significant Time x Group interaction ($p < .001$) with a significant difference in pre-post intervention improvement in favour of the intervention group ($p < .001$, $d = 3.03$). Further analysis once both groups had received the intervention revealed no improvement for either group on past tense production during the five-week pre-intervention period, significant improvement pre-post intervention ($p < .001$, $d = 1.22$), with gains maintained for five weeks post-intervention. No significant differences were found on pre- to post-intervention standardized measures of grammar, or on extension or control measures.

Conclusion: The efficacy of the theoretically motivated explicit grammar intervention was demonstrated. Results contribute to the evidence-base supporting this intervention to improve past tense production in early school-aged children with DLD, suggesting it is a viable option for clinicians to select when treating morphosyntactic difficulties for this population.

Compared with typically developing peers, children with developmental language disorder (DLD) are reported to have a slower pace of language development, and difficulty producing and understanding language in the absence of other biomedical factors (Bishop et al., 2017). This includes particular difficulties with a range of morphosyntactic skills, such as the use and understanding of tense related morphosyntax (Conti-Ramsden et al., 2001; Rice et al., 1999). DLD can have a significant impact on academic (Windsor et al., 2000) and social development (Clegg et al., 2005), and a range of difficulties often persists well into adolescence and adulthood (Law et al., 2009), ultimately affecting employment opportunities (Conti-Ramsden et al., 2018). The precise etiology and contributors to DLD remain unknown. However, exploration of the recommendations derived from theoretical accounts of DLD may inform the development of effective interventions for this at-risk population.

Theory Informing Practice: The Procedural Deficit Hypothesis

The Procedural Deficit Hypothesis (PDH) is based on the assumption of a domain-general deficit in implicit learning for children with DLD in the presence of spared explicit learning (Ullman & Pierpont, 2005). The hypothesis acknowledges the distinction between long term procedural and declarative memory systems. That is, through the procedural memory system, information is learned as a result of repeated exposures to the stimulus. The system underlies the implicit (i.e., non-conscious) learning of skills and habits. Learning through procedural memory is demonstrated through task performance of, for example, early motor development, such as infants learning to walk. In contrast, the declarative memory system is responsible for learning arbitrary items of information and deriving associations between them, underlying explicit (i.e., conscious) retrieval and use of facts and events. Recall or recognition demonstrates the learning of information through explicit memory, for example, events from a birthday party. Information is learned rapidly; however, repeated exposures strengthen memories (Ullman, 2016).

Linguistically, the PDH predicts children with DLD will have impaired morphosyntax learning as this relies on the procedural memory system; whereas vocabulary learning remains relatively intact as it is dependent on the spared declarative memory system (Ullman & Pierpoint, 2005). Certainly, recent evidence points to an implicit learning deficit in children with DLD (Lum et al., 2014) in the presence of relatively spared declarative memory (Lum et al., 2015), particularly in the visual domain (Lum et al., 2012). As such, the way in which linguistic information is presented may assist with language learning. Specifically, explicit instruction delivered with spaced and repeated practice, augmented with visual support, is predicted to improve the learning, storage and use of grammar. Given the competitive nature of the systems, if children with DLD have impaired procedural memory, the PDH suggests using cognitive strategies harnessing spared declarative memory (e.g., explicit intervention) would be more effective than expecting children to learn morphosyntax implicitly.

Recently, Balthazar et al. (2020) summarised the evidence for explicit interventions designed on the principles and perspectives aligned with the PDH to improve grammar for children with DLD. A key component of explicit interventions is metalinguistic training, which refers to the "…verbal description, explanation, and feedback focused on form, the functions of form, and the manipulations of forms" (Balthazar et al., p. 227). The aim of metalinguistic training is to make information conscious and available to recall upon demand. The review included three key approaches to explicit instruction, including Complex Sentence Intervention (Balthazar & Scott, 2018), the SHAPE CODING™ system (Ebbels, 2007), and MetaTaal (Zwitserlood et al., 2015). Of particular interest to the current study is the SHAPE CODING™ system.

The SHAPE CODING™ system provides a systematic way of representing syntax, morphology, and aspects of semantics. The system uses specific visual cues, including colours, shapes and arrows, where: colour coding is used for parts of speech (e.g., nouns, verbs, adjectives); shapes are used to code phrases in accordance with position and role within sentences, and; arrows are used to depict tense. It aligns with an *explicit* intervention approach, and uses metalinguistic training techniques teach the rules of morphosyntax. The system may improve the learning of spoken morphosyntax in the presence of a potential procedural memory deficit by exploiting the functional characteristics of declarative memory. That is, harnessing explicit learning in rich semantic contexts. Further, as an environmental modification, the primarily visual aspect of the system may reduce demands on learning through procedural memory by presenting linguistic information that is less transitory than the spoken modality. The system also allows the presentation of complex sequential information, such as morphosyntax, to be segmented and presented frequently to enhance learning. As such, the SHAPE CODING™ system aligns with recommendations from the PDH in many ways.

Recently, Plante and Gomez (2018) suggested explicit teaching should not be recommended as a therapeutic device, stating, "[b]y intentionally avoiding explicit teaching in favour of implicit learning, clinicians can harness the cognitive resources that support rapid learning" (p. 71). Clearly, this is at odds with the PDH, which suggests explicit teaching is necessary given implicit learning deficits. Rather, Plante and Gomez recommend principles to enhance linguistic input, which are inherently implicit.

The Effectiveness of Interventions Targeting Morphosyntax

Interventions aiming to improve morphosyntax can be considered implicit or explicit (Ebbels, 2014). Implicit interventions are those that aim to enhance the quality and quantity of language input to accelerate language growth without necessarily making the learner consciously aware of the goals of intervention. In contrast, explicit interventions aim to increase the learner's awareness of the goal of intervention, and information learned can be recalled upon demand. The latter aligns with the PDH. Most research has investigated interventions that promote implicit learning, and therefore may be considered best practice; however, there is also an expanding evidence-base supporting the efficacy of explicit interventions (Balthazar et al., 2020), specifically for spoken morphosyntax and usually with older school-aged children and adolescents.

Implicit Interventions

A systematic review and meta-analysis of recasting as an implicit intervention approach has established its effectiveness, with large effect sizes (average $d = 0.96$) on proximal measures of morphosyntax (Cleave et al., 2015). Recently, Eidsvåg et al. (2019) explored whether enhanced recasting was efficacious for improving morphological targets for 20 children aged between 4;8 and 6;7 years with DLD. Children received either individual or paired delivery, once a day, for five days over five weeks. Intervention targets included regular past tense and third person singular. They found significant mean improvement on target morphemes for both delivery conditions with a large effect ($d = 1.16$), but no transference to ambient morphemes. In a randomized controlled trial (RCT) of 18 children aged 4- to 10 years, Owen Van Horne et al. (2017) evaluated a primarily implicit intervention targeting regular past tense production provided for up to 36 intervention sessions. Children in the 'harder' past tense verb condition (i.e., phonologically complex, infrequently marked for inflection and atelic) made greater gains on untrained verbs with a large effect ($g = 1.76$) compared to those in the 'easy' condition (i.e., phonologically simple, frequently marked for inflection and telic). Findings indicated there is an advantage to learning past tense verbs that are more complex.

Explicit Interventions

Ebbels et al. (2007) conducted a randomized controlled trial (RCT) comparing two theoretically motivated interventions (the SHAPE CODING™ system and a semantic intervention) and a no treatment control group with 27 children aged between 10;0 and 16;1

with DLD. Both intervention conditions improved significantly in their use of verb-argument structure with large effects ($d > 1.0$) compared to the control group, with effects generalizing to untrained verbs, suggesting benefit to exploring theoretically grounded interventions. Further, Ebbels et al. (2014) conducted an RCT with 14 children aged between 11;3 and 16;1 comparing progress in comprehension of coordinating conjunctions for children and adolescents with severe receptive language deficits receiving intervention using the SHAPE CODING™ system to a waiting control group. The intervention group showed significant improvement with a large effect ($d = 1.6$), which was maintained for four months. These results suggest benefit in the ongoing evaluation of the SHAPE CODING™ system to target other morphosyntactic deficits, and potentially with younger children.

In a study with two younger children (aged 8;11 and 9;4) with language disorder, Kulkarni et al. (2014) conducted a clinical evaluation of the SHAPE CODING™ system combined with elicited production and recasting to improve the use of regular past tense. Both made statistically significant gains in their use of the target structure after 10 intervention sessions, indicating that further evaluation of the SHAPE CODING™ system to improve past tense marking was warranted.

In an RCT of 31 preschool children with a mean age of five years, Smith-Lock et al. (2015) evaluated the effectiveness two intervention approaches to improve morphosyntax (i.e., regular past tense, third person singular, and possessive *'s*) for children with DLD. One condition combined explicit rule instruction with systematic cueing, and the other condition was recasting alone. In the explicit rule instruction and systematic cueing condition, a cueing hierarchy was used to cue children to correct errors contingent upon their response. If an error occurred when producing a target (e.g., *pulle**d***), the child was first cued with a request for clarification (e.g., *Try that one again.*). Persistent errors were cued until a correct production was achieved, first with emphatic recasting (e.g., *You pulle**d** the cart.*), then forced choice (e.g., *You just pull the cart or you just pulle**d** the cart?*), and finally, elicited imitation (e.g., *You just pulle**d** the cart. Say it like me: pulle**d**.*). It was hypothesized that if an error occurred after being explicitly taught the grammar rule, the cueing hierarchy would provide an opportunity to produce the morphosyntactic target correctly, and therefore would be more effective than using recasts which do not require the child to produce the target. Following eight weeks of intervention, there was a significant medium – large effect ($d = 0.75$) in production for the combined explicit rule instruction and cueing group, but not for the recasting alone group.

In a pilot efficacy study, Calder et al. (2018) combined the SHAPE CODING™ system with a systematic cueing hierarchy (Smith-Lock et al., 2015) for three children aged seven years with DLD. Using a single case experimental design (SCED) the children were seen one-on-one, twice a week for five weeks in 45 minute sessions, resulting in seven and a half hours of intervention. Dose was not held constant, but the children received an average of 49 trials per session. The focus of the intervention for all three children was regular past tense (*–ed*) production. Of the three children, two made significant improvement on *–ed* production. One child demonstrated a decline in performance on *–ed* production during the five week maintenance period, while the other continued to improve significantly. All tree made significant improvement on the Test of Early Grammatical Impairment (Rice & Wexler, 2001), and two of the three children made significant improvement on the Test for Reception of Grammar 2nd Edition (TROG-2) (Bishop, 2003). No child showed improvement on the third person singular (*3S*) extension measure or the possessive *'s* control measure, further increasing confidence that improvements in *–ed* production were due to intervention. This provided early evidence supporting the use of explicit intervention approaches to improve receptive and expressive grammar, particularly production of *–ed* following five weeks of intervention. However, it was acknowledged that a longer period of intervention may be necessary to increase the magnitude of intervention effects. Finally, other measures of grammatical knowledge (e.g., grammaticality judgment) would increase confidence in reporting improvement of *–ed* marking.

Efficacy of the intervention was further evaluated using a study of $n = 9$ through a SCED (Calder et al., 2020). Intervention was provided twice weekly for 10 weeks with 50 trials resulting in 1000 trials for each participant. Repeated measures were probed, including trained and untrained *–ed* verbs, an extension probe (*3S*), and a control probe (*'s*). All probes included expressive and grammaticality judgment contexts. Pre-post measures of standardized expressive grammar and receptive grammar measures were also analyzed. Of the nine children, eight made significant improvement on production of trained verbs with large effects (*Tau* = 0.88), and seven made significant improvement on production of untrained verbs with moderate effects (*Tau* = 0.64). A within-group concurrent analysis was also conducted on production of trained verbs indicating significant progress pre- to post-intervention ($p = .008$). Results suggested a stable pre-intervention phase, significant pre-post-intervention improvement, and maintenance of gains at a group level. For the grammaticality judgment probes, three participants improved significantly on trained verbs with small effects (*Tau* = 0.26), and only one improved on untrained verbs. There was limited to no improvement on extension and control probes for either expressive or grammaticality judgment contexts. For the standardized measures, eight participants exceeded the reliable change index (RCI: > 1.96) indicating significant improvement on the Structured Photographic Expressive Language Test 3rd Edition (Dawson et al., 2003), and one exceeded the RCI on the TROG-2. Results from the study demonstrated that the intervention was efficacious for improving *–ed* marking of trained and untrained verbs, and expressive grammar generally if provided twice a week for 10 weeks. Given the positive results yielded from SCEDs, evaluation of the intervention through an RCT was warranted.

The Current Study

The aim of the current study was to examine the efficacy of an explicit intervention motivated by the PDH to improve *–ed* marking, and a possible generalized effect to standardized grammar scores. The intervention involved a combination of the SHAPE CODING™ system (Ebbels, 2007) as the metalinguistic training and a systematic cueing hierarchy (Smith-Lock et al. 2015). If explicit intervention is efficacious in improving morphosyntax for young children with DLD, this may inform the theoretical motivation for developing intervention procedures well suited to facilitate learning for these children. The research questions were as follows:

1. Do early school-aged children (5;9-6;9 years) with DLD show greater improvement in past tense (*–ed*) marking following 10 weeks of explicit intervention compared to a 'treatment-as-usual' waiting control group?

   It was hypothesised that the intervention group would show significantly greater progress on measures of -ed marking following intervention compared to the control group.

2. Do these children also show greater improvement on standardized expressive and receptive grammar measures following intervention for *–ed* compared to the waiting control group?

   It was hypothesized that the intervention group would show significantly greater progress on generalized measures of grammar compared to the control group.

3. Will results from ancillary analyses determine whether children also improve on a linguistically related grammatical target (third person singular: *3S*), but not on a behavioral control measure (possessive *'s*).

   It was hypothesized that the intervention group may show greater improvement on measures of 3S*, but there would be no between group differences on measures of 's.*

Method

Efficacy of the intervention was tested using a crossover randomized controlled trial (RCT) study design. Pre-post results on measures of past tense (*–ed*) were the primary variables of interest, pre-post results on standardized measures of grammar were the secondary variables of interest, and measures of third person singular (*3S*) and possessive *'s* were part of ancillary analyses. All reporting follows the Consolidated Standards of Reporting Trials (CONSORT) statement (Schulz et al., 2010) (see Supplemental Materials for CONSORT checklist (S1) and flow diagram (S2)). Ethical approval for the study was obtained from the Curtin University Human Research Ethics Committee (Approval number HRE2017-0835) and the Western Australian Department of Education (Approval number D190018955).

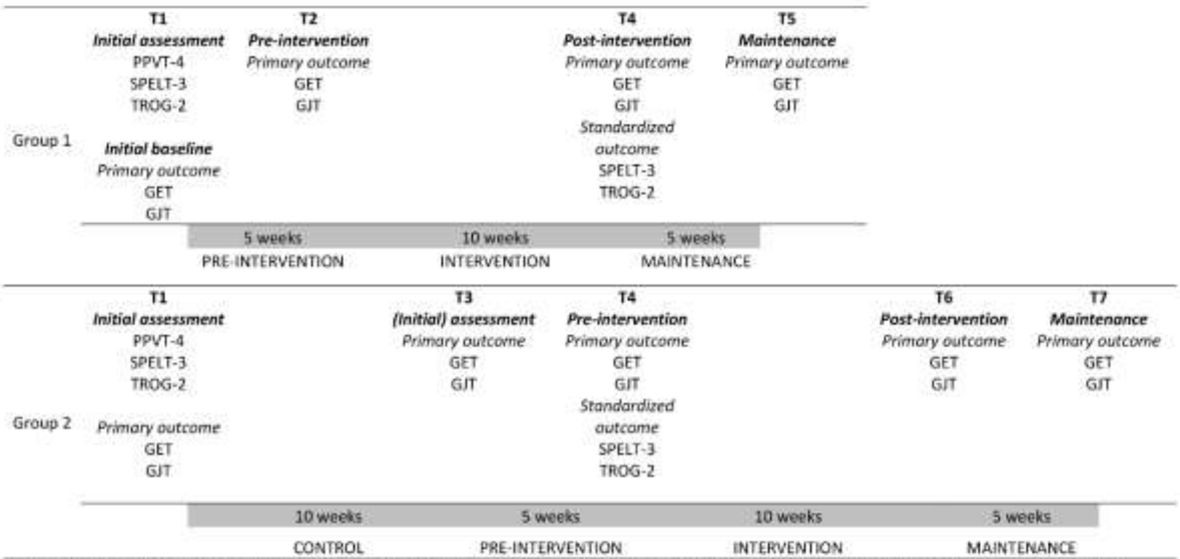
Participants

Participants were recruited from specialized educational programs for students diagnosed with DLD across three sites. The principal consented school participation, and then provided information letters and consent forms to the parents/carers of potential participants identified by speech-language pathologist (SLP) and teachers employed at the educational program. Parents then returned forms consenting to their child's participation. Participant inclusion criteria included children aged between 5;6 and 7;6; English as a primary language, and; grammar difficulties associated with DLD. Exclusionary criteria were based on factors outlined in Bishop et al. (2016) for determining a diagnosis of DLD, including a neurological and/or cognitive impairment. Accessing the participants' school enrolment packages confirmed they met criteria for DLD[1].

[1] There have been recent changes relating to the terminology and classification of DLD, and the use of non-verbal IQ criteria for diagnosis (see Bishop et al., 2016, pp.5-6). At the time of this study, entry requirements to specialized educational programs in Western Australia included non-verbal IQ in the average range.

Figure 1

Summary of the assessment and intervention schedule for Group 1 and Group 2



Notes. Group 1= explicit intervention group; Group 2= 'treatment-as-usual' waitlist control (crossover to explicit intervention at T4); T1= timepoint 1 (Group 1 initial assessment; Group 2 initial assessment); T2= timepoint 2 (Group 1 pre-intervention); T3= timepoint 3 (Group 2 (initial) assessment); T4= timepoint 4 (Group 1 post-intervention; Group 2 pre-intervention); T5= timepoint 5 (Group 1 maintenance); T6= timepoint 6 (Group 2 post-intervention); T7= timepoint 7 (Group 2 maintenance); GET= Grammar Elicitation Test, including –ed, 3S and 's probes; GJT = Grammaticality Judgment Test, including –ed, 3S and 's probes; SPELT-3= Structured Photographic Language Test 3rd Edition (Dawson, Stout, & Eyer, 2003); TROG-2= Test of Reception of Grammar 2nd Edition (Bishop, 2003); PPVT-4= Peabody Picture Vocabulary Test, Fourth Edition (Dunn & Dunn, 2007).

See Figure 1 for a full breakdown of the assessment and intervention schedule. Participants were screened in hearing acuity and passed at 20 dB for each ear at 500, 1000, 2000, and 4000 Hz. All participants passed the Phonological Probe from the Test of Early Grammatical Impairment (Rice & Wexler, 2001) which confirmed that they were able to articulate phonemes necessary for morphosyntax production targets. Participants were assessed for production of morphosyntax using criterion-referenced measures adapted from the Grammar Elicitation Test (GET) (Smith-Lock et al., 2013). This experimental test was designed to elicit multiple instances of specific morphosyntax structures, with subtests for regular past tense (GET-*ed*), third person singular (GET-*3S*), and possessive *'s* (GET-*'s*). Each subtest contains 30 probes each divided into three allomorph groups (i.e., 10x [d], 10x [t] and 10x [əd] for –*ed*; 10x [z], 10x [s], 10x [əz] for *3S* and possessive *'s*). Verbal elicitation for each probe (e.g., *This boy likes to hop. He did it yesterday. What did he do yesterday?*) was pre-recorded to ensure each child received consistent assessment procedures.

Measures of morphosyntax were also developed for a grammaticality judgment task mirroring the above morphosyntactic structures, hereafter referred to as the Grammaticality Judgment Test (GJT: GJT-*ed*, GJT-*3S*, GJT-*'s*). As with the GET, all possible allomorphs were included and distributed equally within each category of probe. Videos of actions depicted the declarative clauses containing *–ed* and *3S*. For possessive *'s*, copyright free still images depicting nouns and ownership were retrieved. As with the GET, accompanying audio for each task item, both grammatical and ungrammatical (e.g., *The boy hopped on one foot* vs *The boy hop* on one foot*), was pre-recorded for administration.

All subtests for both the GET and GJT were embedded into a Microsoft PowerPoint presentation and delivered via laptop. For the GJT, participants wore noise-cancelling headphones during administration and were required to decide if the sentence 'sounded right' by pressing 'yes' or 'no' on a tablet app. Items were counterbalanced for grammaticality so participants neither received the same combination of grammatical/ungrammatical items, nor was there a pattern in presentation of items to counteract a priming effect. For all relevant analyses, scores from the GJT were adjusted to account for 'yes bias' by computing A*'* as $A' = 0.5 + (y - x)(1 + y - x)/4y(1 - x)$, where y = proportion of hits (i.e., the child selected 'yes' for a grammatical item) and x = proportion of false alarms (i.e., the child selected 'yes' for an ungrammatical item) (Rice et al., 1999). Tables report percentage accuracy on the GJT for consistency with the GET.

Initial assessment measures also included the Structured Photographic Expressive Language Test 3rd Edition (SPELT-3) (Dawson et al., 2003). The test measures a range of expressive morphosyntax structures over 54 items. It has strong internal consistency reliability (r = .86) and appropriate construct validity. The Test of Reception for Grammar 2nd

Edition (TROG-2) was used to assess receptive grammar. Test blocks measure a total of 20 different grammatical structures. It has strong internal consistency reliability (*r* = .86) and appropriate construct validity. The Peabody Picture Vocabulary Test, 4th Edition (PPVT-4) (Dunn & Dunn, 2007) was administered as a yardstick for static vocabulary abilities. The test has strong internal consistency reliability (*r* = .94) and appropriate construct validity.

Following distribution of consent forms to the specialized education programs for children with DLD, parents of 23 children aged between 5;9 and 6;9 years provided consent to participate, and the children were assigned a code and assessed for eligibility. These codes were entered into a true random list generator by a researcher blinded to the purpose of the study to ensure concealed allocation sequence. Once the list was randomized from 1-23, the researcher assigned participants to either the intervention group (Group 1) (every odd occurrence on the list) or to the 'treatment-as-usual' waiting control group (Group 2) (every even occurrence on the list). Further to blind random assignment, assessors were blind to group assignment at all testing timepoints. Participants, their caregivers, and teachers were not made aware of the conditions. The purpose of the intervention beyond targeting morphosyntax was not disclosed to caregivers or teachers.

Of the 23 children, two males were excluded from the study because they reached ceiling on initial assessment using the GET and GJT, and were deemed unlikely to benefit from further participation. This resulted in the remaining 21 participants comprising the explicit intervention group (Group 1) (*n* = 10) or the 'treatment-as-usual' waiting control group (Group 2) (*n* =11). Group 1 comprised eight males (80%) and two females (20%) with a mean age of 6;3 years (*SD* = 0;4 years; range: 5;11-6;8 years) at initial assessment. Group 2 comprised seven males (63.6%) and four females (36.9%) with a mean age of 6;6 years (*SD* = 0;3 years; range: 5;9-6;9 years) at initial assessment.

Table 1

Mean and standard deviation values for demographic and initial assessment information for study participants.

	Age at initial assessment	Age at onset of treatment	Sex	GET%	GJT%	SPELT-3 (SS)	TROG-2 (SS)	PPVT-4 (SS)
Group 1	6;3 (0;4)	6;4 (0;2)	M 80%/F 20%	32.0 (23.3)	55.3 (14.9)	76.4 (18.2)	77.5 (14.5)	93.2 (14.9)
Group 2	6;6 (0;3)	7;3 (0;3)	M 64%/F 36%	23.1 (21.6)	53.9 (12.1)	69.2 (18.7)	76.7 (9.3)	90.7 (8.8)

Notes. All scores from standardized assessments are scaled scores; Group 1 = explicit intervention group; Group 2 = 'treatment-as-usual' waiting control (crossover to explicit intervention at T4); GET = Grammar Elicitation Test; GJT= Grammaticality Judgment Test; PPVT-4= Peabody Picture Vocabulary Test, Fourth Edition (Dunn & Dunn, 2007); SPELT-3= Structured Photographic Language Test 3rd Edition (Dawson et al., 2003); TROG-2= Test of Reception of Grammar 2nd Edition (Bishop, 2003); M= male; F= female. SS = Scaled Score. Percentage accuracy for the GJT is presented rather than *A'* to maintain consistency with the GET.

See Table 1 for a summary of initial assessment data including standardized scores from assessments. Potential between group differences on age, sex, grammar scores and the PPVT-4 at initial assessment were evaluated using a one-way between group analysis of variance (ANOVA). Assumptions of normality were violated for Group 1 for age (Shapiro-Wilk statistic: $p = .04$), sex (Shapiro-Wilk statistic: $p < .001$) and the TROG-2 (kurtosis: > 1.96). Assumptions were violated using *A'* on the GJT for both groups (kurtosis: >1.96; Shapiro-Wilk statistic: $p < .001$). Therefore, differences on these variables were analysed with non-parametric Mann Whitney *U* tests. Analyses revealed no group differences on any demographic variables or the GJT *A'* (all *p*'s > .42, two-tailed). *F* tests for the SPELT-3, GET, and PPVT-4 were also non-significant (all *p*'s >.32). Therefore, there were no group differences following allocation to groups, and both groups had mean scores within normal limits on the PPVT-4 using a one *SD* cut-off.

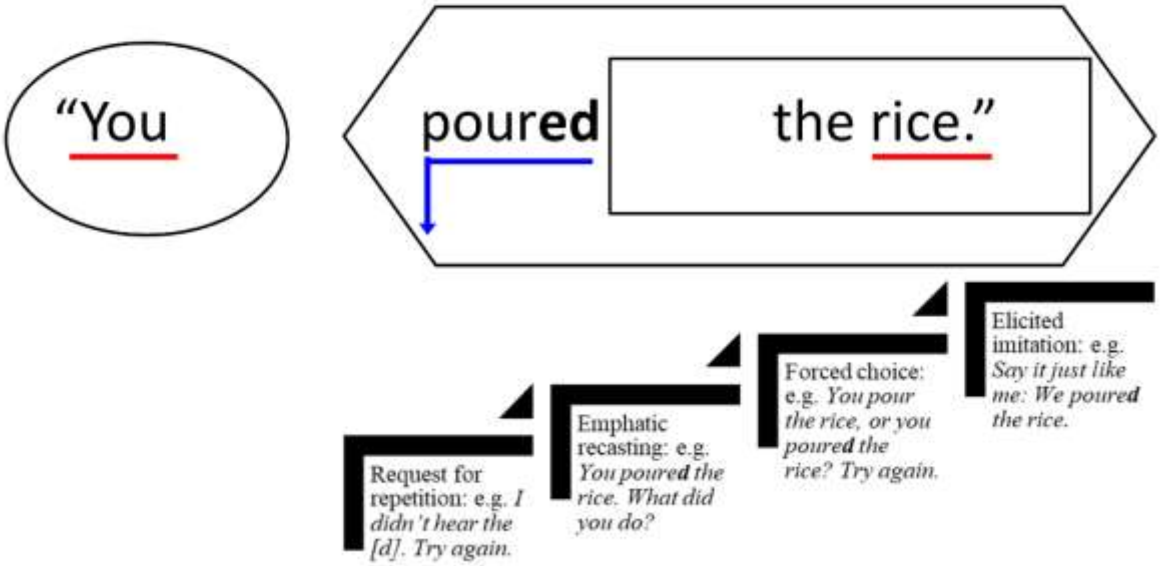
Intervention

The goal of intervention was to improve *–ed* marking. All intervention sessions were videotaped and carried out in a quiet space at the site of the educational programs. Using the framework for explaining intervention suggested by Warren et al. (2007), the dose was 50 trials within 20-30 minute sessions; dose form was explicit intervention combining metalinguistic training using the SHAPE CODING™ system (Ebbels, 2007) with a systematic cueing hierarchy (Smith-Lock et al., 2015); dose frequency was once a week; total intervention duration was 10 weeks, and; cumulative intervention intensity was (50 trials x 1 time per week x 10 weeks), resulting in a total of 500 trials over 10 individual therapy sessions through roughly 3.5-5 hours of intervention. This is half the cumulative frequency reported in Calder et al. (2020), allowing evaluation of the intervention in a more clinically relevant dose frequency (e.g., Finestack & Satterlund, 2018). Training of *–ed* was contextualized within engaging and naturalistic activities suited to early school-aged children, including playdough, board games, puppets, and farm and sea creature manipulatives. Target morphemes were presented in syntactic structures as they occurred felicitously within these activities. The first author, a trained SLP, delivered all intervention.

See Figure 2 for a visual representation of the SHAPE CODING ™ system and systematic cueing hierarchy. Procedures included reminding children of the goal of intervention (i.e., *–ed* marking) through explicit teaching at the beginning of the session. The SLP checked vocabulary knowledge of the participants by asking them to label materials from the session. Three Subject + Verb/ +Object (SV/O) sentences were demonstrated using one exemplar from each of the possible allomorphic categories (i.e., [d], [t] and [əd]). The shapes and arrows from the SHAPE CODING™ system were introduced. The shapes included the oval (subject noun phrase WHO/WHAT?), the hexagon (verb phrase WHAT DOING?), and the rectangle (object noun phrase WHO/WHAT?). The arrows included blue 'left down arrows' to depict orthographic representations for each of the possible *–ed* allomorphs (i.e., [d] = 'd', [t] = 't', and [əd] = 'ed'). Production of *–ed* was targeted through 25 trials in total where the child had the opportunity to respond to an interrogative following priming (e.g., *You pour the rice. What DID you DO?*), while the shapes and arrows were gestured to. Children were cued using the systematic hierarchy in the case of errors in production (e.g., unmarked bare stem *pour* or overgeneralised form *pourded*). A consolidation component included reviewing three exemplars from each allomorphic category, followed by a comprehension task where participants answered *wh-* questions related to the SV/O structures while the SLP gestured to the shapes and arrows (e.g., *WHO poured the rice*; *What DID you DO?; WHAT did you pour?*), and finally the child produced sentences without the visual support of shapes and arrows. 'Silly Sentences' were used whereby three sentences were spoken by the SLP, either grammatically or ungrammatically (i.e., *–ed* morphemes were either included or omitted in the clause), and the child would decide if the sentence 'sounded

Figure 2

Visual cues and systematic cueing hierarchy used during intervention



right'. These procedures were repeated for a second activity until 50 trials in total were achieved. The sessions were concluded by providing a summary of the goals of intervention (i.e., to produce *–ed*). A full session plan fidelity checklist with timing breakdown (S3), a summary of intervention targets and materials (S4), and adherence to the template for intervention and description and replication checklist (Hoffman et al. 2014) (S5) are available in the Supplemental Materials.

Intervention for the 'treatment-as-usual' waiting control involved regular attendance at a specialized educational program, which included intensified oral language instruction with a modified curriculum designed to cater to the academic needs of children with DLD. Typically, oral language instruction was embedded within narrative blocks to enhance exposure to language through strategies similar to those reported in Gillam et al. (2012) and Spencer et al. (2014). Therefore, although grammar instruction was part of the modified curriculum, it was not the primary focus of instruction for the duration of the current study.

Intervention Fidelity

The first author and a researcher blinded to the purpose of the study rated 20% of videotaped sessions on percentage accuracy for intervention procedures (see S3 for a scoring checklist). Between-observer agreement was calculated using intraclass correlation coefficients (ICC) with absolute agreement and single measures in a two-way mixed effects model. The average score across raters was 99.58% for percentage accuracy, and ICC for procedures was .995 indicating excellent agreement (Cicchetti, 1994).

Inter-Rater Reliability for Test Scoring

In addition to the first author scoring all data, a researcher blinded to the purpose of the study scored 20% of experimental measures. For the GET (including *–ed*, *3S* and *'s*), the ICC was .956. For the GJT (including *–ed*, *3S* and *'s*), the ICC was .997. Therefore, excellent between-observer agreement was demonstrated across all experimental measures.

Results

All outcomes were planned to be assessed using intention-to-treat analyses. In the case of participant loss to follow up, last-observation-carried-forward was implemented. Study compliance is reported in detail. Data from all 10 children in the intervention group were analysed for outcomes and estimations. For the waiting controls, data from all 11 children were analysed for between group analyses of an intervention effect on grammar scores, i.e., intervention (Group 1) versus 'treatment-as-usual' waiting control (Group 2). However, one participant was exited from the study following crossover into the intervention condition due to reaching ceiling on pre-intervention measures and was deemed unlikely to benefit further from participation. Therefore, this participant was excluded from further analyses. One additional child in Group 2 was unavailable for testing at the final testing point due to being absent from the specialized educational program because of illness and was unavailable for follow up. The results from this child's final data point during the maintenance phase were included in the analysis using last-observation-carried-forward. All other children participated in all intervention sessions and assessment timepoints.

Outcomes and Estimation

All means and standard deviations on outcome, extension, and control measures are reported in Table 2. For all between group comparisons, mixed ANOVAs were planned, with Time as the within-participant variable, and Group as the between-participant variable. All necessary post hoc pairwise comparisons included Bonferroni adjustments for α values. Following post hoc tests, effect sizes were calculated for between group comparisons on pre-post intervention improvement using Cohen's *d* to interpret small (0.2), medium (0.5) and large (0.8) effects. In the case of violated assumptions, non-parametric Kruskal-Wallis one-way ANOVAs were conducted, and scores from pre- to post-intervention were transformed by subtracting T1 from T4 to account for time as a factor when testing between group improvement. Cohen's *f* were calculated for effect sizes. All statistics were computed using IBM SPSS Version 25.

Research Question 1: Does Past Tense Marking Improve Following Intervention Compared to 'Treatment-As-Usual'?

Refer to Figure 1 for the full testing schedule. The GET-*ed* and GJT-*ed* were administered as primary outcome measures of past tense production and grammaticality judgment, respectively. As verbs from the GET -*ed* and GJT -*ed* were not trained as part of intervention, improvement would unlikely be attributed to practice effects alone. Items were randomized for administration at the initial assessment (T1), one week prior to intervention commencing for Group 1 (T2), one week following intervention for Group 1 (T4) and five weeks following cessation of Group 1's intervention (T5). This differed for the 'treatment-as-usual' waiting control group (Group 2), where there was an initial assessment (T1), beginning

Table 2

Means and standard deviations for all primary and standardized outcomes, and extension and control measures

Group 1		Pre-intervention			Post-intervention			
Outcomes		T1	T2		T4	T5		
Primary	GET-*ed* (%)	27.3 (16.1)	25.3 (21.7)		66 (14.5)* †	54.3 (22.16)		
	GJT-*ed* (%)	52.7 (13.4)	57 (9.1)		58.7 (14.5)	56.3 (16.2)		
Standardized	SPELT-3 (SS)	76.4 (17.2)	-		78.9 (11.9)	-		
	TROG-2 (SS)	77.5 (13.7)	-		84.4 (19.6)	-		
Extension	GET-*3S* (%)	29.3 (24.1)	32.3 (34.06)		45.3 (28.2)	44.7 (31.1)		
	GJT-*3S* (%)	60.3 (15.0)	58 (14.1)		57.0 (14.8)	61.3 (16.1)		
Control	GET-*'s* (%)	39.3 (26.7)	44.3 (30.5)		50.0 (33.4)	40.3 (32.4)		
	GJT-*'s* (%)	56.0 (15.3)	55.0 (14.6)		60.0 (18.0)	56.7 (19.6)		
Group 2		Pre-intervention					Post-intervention	
Outcomes		T1		T3	T4		T6	T7
Primary	GET-*ed* (%)	19.7 (16.3)		17.6 (17.1)	22.4 (21.9)†		45.3 (20.5)*	43.4 (26.0)
	GJT-*ed* (%)	55.2 (11.7)		56.7 (14.1)	60.3 (13.8)		51.3(11.4)	52.2 (8.6)
Standardized	SPELT-3 (SS)	69.0 (17.8)		-	75.7 (13.0)		82.5 (18.4)	-
	TROG-2 (SS)	76.7 (8.9)		-	81.0 (18.3)		84.0 (16.2)	-
Extension	GET-*3S* (%)	23.6 (26.8)		27.3 (22.8)	32.1 (29.6)		32.0 (28.0)	33.0 (29.5)
	GJT-*3S* (%)	52.7 (10.6)		54.8 (18.2)	60.6 (19.0)		61.0 (11.2)	54.8 (17.5)
Control	GET-*'s* (%)	26.1 (19.9)		16.3 (17.6)	31.0 (25.5)		25.3 (18.4)	33.0 (29.5)
	GJT-*'s* (%)	53.9 (13.7)		53.0 (15.5)	54.2 (14.6)		54.3 (11.6)	51.5 (6.1)

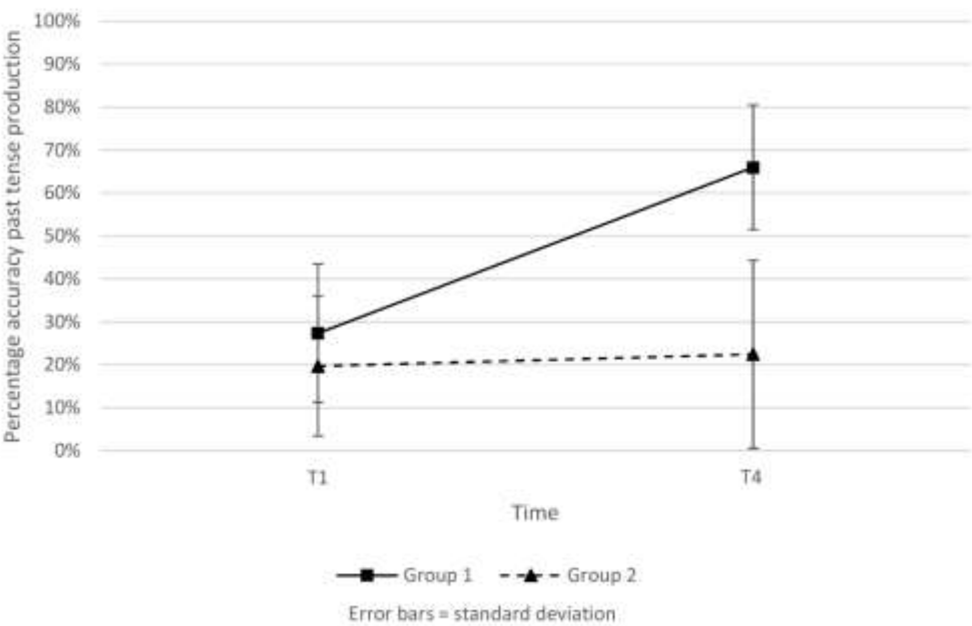
Notes. Group 1= explicit intervention group; Group 2= 'treatment-as-usual' waitlist control (crossover to explicit intervention at T4); T1= timepoint 1 (Group 1 initial assessment); T2= timepoint 2 (Group 1 pre-intervention); T3= timepoint 3 (Group 2 initial assessment); T4= timepoint 4 (Group 1 post-intervention; Group 2 pre-intervention); T5= timepoint 5 (Group 1 maintenance); T6= timepoint 6 (Group 2 post-intervention); T7= timepoint 7 (Group 2 maintenance); GET= Grammar Elicitation Test; GJT= Grammaticality Judgment Test; SPELT-3= Structured Photographic Language Test 3rd Edition (Dawson et al., 2003); TROG-2= Test of Reception of Grammar 2nd Edition. (Bishop, 2003); *-ed* = regular past tense; *3S* = third person singular; *'s* = possessive 's; SS = Scaled Score; *sig. difference to pre-intervention score; †sig. between group difference. Percentage accuracy for the GJT is presented rather than A' to maintain consistency with the GET.

of pre-intervention (T3), one week prior to intervention commencing (T4), one week following intervention (T6) and five weeks following cessation of intervention (T7).

GET –ed: Real Time. Percentage correct of *–ed* production of untrained verbs as the primary outcome is presented in Figure 3. The hypothesis predicted a significant Time x Group interaction on the GET-*ed* as an evaluation of performance in ‘real time’. That is, all participants were assessed on the same measures collected at the same time throughout the study (i.e., T1 and T4). Results showed a significant main effect of Group, $F(1, 19) = 11.71$, $p = .003$, $\eta^2 = .381$, in favour of Group 1. Further, there was a significant main effect of Time, $F(1, 19) = 48.87$, $p < .001$, $\eta^2 = .72$. There was a significant Time x Group interaction, $F(1, 19) = 36.84$, $p < .001$, $\eta^2 = .66$, where post hoc pairwise comparisons revealed mean difference in improvement was significant ($p < .001$) in favour of Group 1 ($M = 38.7\%$) over Group 2 ($M = 2.7\%$) ($d = 3.03$).

Figure 3

Between group comparison of mean percent accuracy past tense production pre- and post-intervention in real time

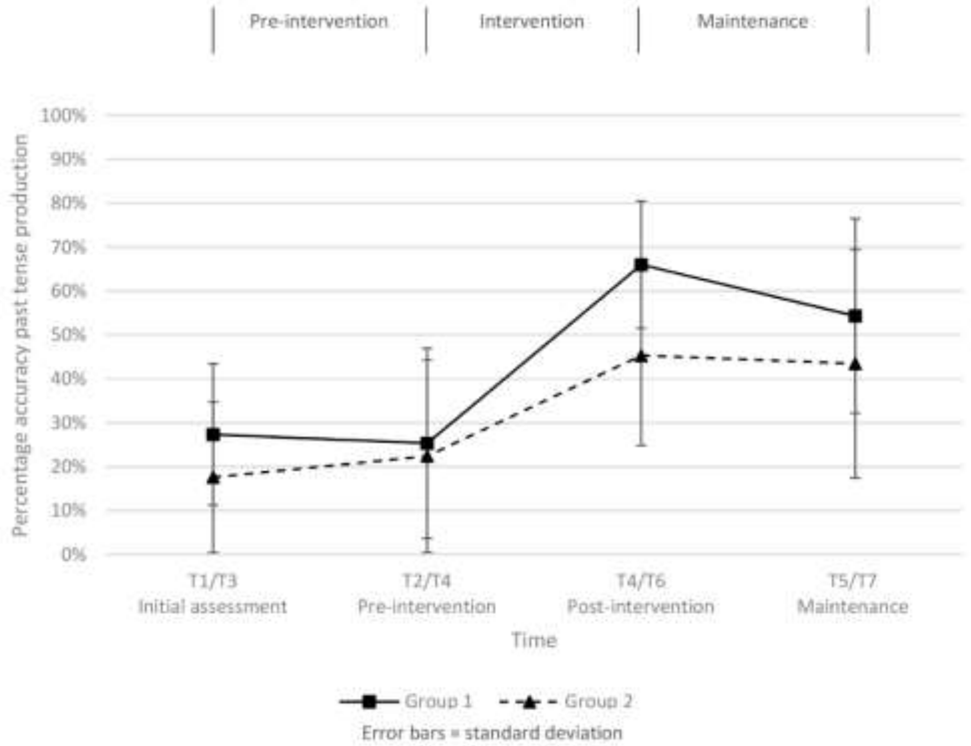


GET –ed: Relative Time. Mean percentage accuracy of *–ed* production at the 'relative' point of testing once Group 2 had crossed over into the intervention condition was compared. That is, although data were collected at different time points, the measures taken were relative to the time of receiving intervention: initial assessment (Group 1: T1, Group 2: T3); pre-intervention (Group 1: T2, Group 2: T4); post-intervention (Group 1: T4; Group 2: T6), and; maintenance (Group 1: T5, Group 2: T7). This counteracted the possibility of environmental factors that may influence analyses of data collected in 'real time'. Percentage accuracy of *–ed* production of untrained verbs is presented in Figure 4. The hypothesis predicted no change in performance during the pre-intervention period, but significant changes in performance following intervention between pre- and post-intervention time points for both groups, i.e., a main effect of Time only.

Results indicated no significant main effect of Group, $F(1, 18) = 2.90$, $p = .11$, $\eta^2 = .36$ or Time x Group interaction, $F(3, 54) = 1.13$, $p = .34$, $\eta^2 = .01$. However, there was a main effect of Time, $F(3, 54) = 56.28$, $p < .001$, $\eta^2 = .68$. Post hoc pairwise comparisons of mean percentage accuracy of *–ed* production from both groups combined revealed no significant difference between the initial assessment and pre-intervention timepoints ($p = 1.00$) or the post-intervention and maintenance timepoints ($p = .52$). However, both the post-intervention and maintenance timepoints were significantly higher than initial assessment and pre-intervention timepoints ($p < .001$ across all comparisons). These results suggest that once the assessment times relative to intervention were compared, there was no improvement during the pre-intervention phase ($d = -.12$), yet *–ed* production improved during the intervention period ($d = 1.22$), and this effect was maintained for five weeks ($d = -.30$).

Figure 4

Between group comparison of mean percent accuracy past tense production across pre- and post-intervention timepoints in relative time



GJT -ed: Real Time. Assumptions were violated on GJT-*ed* A*'* as a measure of grammaticality judgment for both groups. The test of progress between groups was non-significant, $H = 2.20$, $df = 1$, $N = 21$, $p = .138$, $f = .35$, showing both groups made a similar Amount Of Progress, Therefore No 'Relative Time' Analysis Was Conducted.

Research Question 2: Do Standardized Measures Of Grammar Improve Following Intervention for –ed Compared to 'Treatment-As-Usual'?

SPELT-3 and TROG-2. The SPELT-3 and TROG-2 were administered at T1 and T4 as standardized expressive and receptive grammar measures, respectively. For between group comparisons on the SPELT-3, the *F* test was non-significant for a Time x Group interaction, $F(1,19) = .75$, $p = .397$, $\eta^2 = .038$. Normality was violated for Group 1 at initial assessment on the TROG-2. Between group comparisons of progress with intervention were non-significant, $H = .55$, $df = 1$, $N = 21$, $p = .46$, $f = .17$.

Research Question 3: Does Third Person Singular Marking, but not Possessive 's Marking Improve Following Intervention for –ed Compared to 'Treatment-As-Usual'?

The GET-*3S* and GJT-*3S* were administered as extension measures for ancillary analyses. Since *3S* structures are linguistically related to *–ed*, and the *3S* structure was often used to prime children (e.g., *The frog flips. What DID it DO?*), it was hypothesized that there may be an observable intervention effect as a result of the increased linguistic input during intervention. The GET-*'s* and GJT-*'s* were also administered as control measures for ancillary analyses. The inclusion of a linguistically unrelated control measure contributed to the internal validity of the study. The testing schedule was identical to the administration of the GET-*ed* and GJT-*ed* presented in Figure 1 and results are presented in Table 2.

GET 3S and GJT 3S (Extension). For production of *3S*, normality for Group 2 was violated. Results from between group comparison of progress between T1 and T4 was non-significant, $H = .61$, $df = 1$, $N = 21$, $p = .437$, $f = .04$. For grammaticality judgement of *3S,* normality was violated for both groups, and the between group comparison of progress with between T1 and T4 was non-significant, $H = 1.22$, $df = 1$, $N = 21$, $p = .269$, $f = .26$.

GET 's and GJT 's (Control). For the between group performance on production of *'s* as a control measure at T1 and T4, the *F* tests for an interaction were non-significant, $F(1, 19) = 1.048$, $p = .319$, $\eta^2 = .052$. Normality was violated for GJT *'s* for Group 1, and the between group comparison of progress between T1 and T4 was non-significant, $H = .85$, $df = 1$, $N = 21$, $p = .358$, $f = .21$.

Discussion

This study reports on the efficacy of an explicit intervention approach motivated by the Procedural Deficit Hypothesis (PDH) (Ullman & Pierpoint, 2005) which combined metalinguistic training and systematic cueing to improve past tense (*–ed*) marking for early school-aged children with DLD. Recent pilot (Calder et al., 2018) and early efficacy (Calder et al., 2020) studies have demonstrated that explicit intervention using the SHAPE CODING™ system (Ebbels, 2007) in combination with a systematic cueing hierarchy (Smith-Lock et al., 2015) significantly improves *–ed* production of untrained and trained verbs, with no improvement observed on behavioural control measures. The current study used a crossover RCT design to compare intervention provided once per week for 10 weeks (Group 1: *n* =10) compared to a 'treatment-as-usual' waiting control group (Group 2: *n* = 11). Results contribute to the understanding of viable grammar intervention options for children with DLD, particularly interventions motivated by the PDH.

Outcomes and Estimation

Research Question 1: Does –ed Marking Improve Following Intervention Compared to 'Treatment-As-Usual'?

Group comparisons in 'real time' showed a statistically significant improvement pre-post intervention for the explicit intervention group (Group 1) with a large effect (*d* = 3.03), but not for the 'treatment-as-usual' waiting control group (Group 2). Both groups were also compared after Group 2 had crossed over and received the intervention comparing performance in 'relative time'. Analysis showed that for both groups, there was no improvement on *–ed* production during the pre-intervention phase (*d* = -0.12), but improvement with intervention (*d* = 1.22), and although there was a decrease in the effect (*d* = -0.30), this was not significant, so progress was maintained for five weeks, further supporting the efficacy of the intervention. However, there was no significant difference between groups on the measure of *–ed* grammaticality judgment for any analysis. This measure evaluates knowledge of correct finiteness marking use in obligatory contexts. Rice et al. (1999) have shown that this is a clinical marker of DLD, but it has not been shown to be amenable to change through this intervention in an early stage efficacy (Calder et al., 2020) study or the current study. This may indicate DLD persists regardless of highly targeted intervention, and alternative approaches to improve grammatical judgment may be necessary.

Research Question 2: Do Standardized Measures of Grammar Improve Following Intervention for –ed Compared to 'Treatment-As-Usual'?

Between-group comparisons of standardized grammar scores were non-significant. Inconsistent with previous findings (Calder et al., 2018, 2020), mean scores on standardized assessments did not significantly improve following intervention. The current study differs in a critical aspect: the addition of a waiting control group. This allowed the exploration of whether change may have been attributable to maturation or a practice effect instead of intervention, which was not possible within the previous studies. However, the marked improvement by the majority of participants on expressive grammar scores in Calder et al. (2020), supported with analysis using the reliable change index, may suggest another important variable to consider: dosage. Calder et al. (2020) evaluated intervention delivered twice weekly for 10 weeks, whereas the current study evaluated intervention provide once weekly. Perhaps by halving cumulative intervention intensity, significant improvement was only observed on the target structure (i.e., *–ed* production). Other environmental factors cannot be ruled out without direct comparison, so this should be explored in future research.

Research Question 3: Does Third Person Singular Marking, but not Possessive 's Marking Improve Following Intervention for –ed Compared to 'Treatment-As-Usual'?

There were no differences between groups on production or grammaticality judgment of *3S* marking (the extension measure), suggesting no effect of the intervention on these measures. Similar to findings from Calder et al. (2018, 2020), the limited improvement on *3S*, suggests that even linguistically similar structures are unlikely to improve without direct intervention. This finding is consistent with existing studies (e.g., Eidsvåg et al., 2019) suggesting grammar intervention must be targeted to specific targets to yield improvement in children with DLD. Further, there were no between-group differences on production or grammaticality judgment of possessive *'s* as a behavioural control measure. Therefore, results overall support the efficacy of the intervention to improve *–ed* production, which is not attributable to maturation or other general factors such as school environment. This is further supported by a lack of between group differences on a behavioural control measure.

Theoretical Implications

The PDH suggests that in the presence of impaired implicit learning, children with DLD as a clinical population may be suited to learning grammatical information explicitly. The current intervention was designed based on recommendations from the PDH. These include the provision of intervention which modifies the way in which morphosyntax is presented for learning. Specifically, providing explicit instruction, spacing of repeated practice, and visual support to improve the learning, storage, and use of morphosyntax. Results from the current study support the PDH insofar as explicit intervention using metalinguistic training and visual support improves *–ed* production. However, without a direct comparison of explicit and implicit interventions, the explanatory power of the PDH suggesting explicit intervention is best suited to children with DLD cannot yet be confirmed.

Effectiveness of Interventions Targeting Morphosyntax

Given this study evaluated a theoretically motivated explicit intervention approach, future research may serve to compare implicit versus explicit grammar interventions to determine superiority. Findings from the current study are comparable to those of recent studies which have evaluated similar age groups and intervention targets for children with DLD. Eidsvåg et al. (2019) evaluated the efficacy of an implicit intervention using enhanced conversational recasting on children with a mean age of 5.6 years with 45 minute sessions provided five days a week for five weeks. The mean proportion correct on goals pre-intervention was 4.6% and post-intervention was 57% resulting in a mean improvement of 52.4% correct. Owen Van Horne et al. (2017) evaluated a primarily implicit intervention with children with a mean age of six years delivered for up to 36 sessions. The mean proportion correct for the group who showed the greatest advantage was 25% correct on generalisation targets pre-intervention and 60% correct post intervention, providing a mean improvement of 35%. Perhaps most relevant, Smith-Lock et al. (2015) evaluated the effectiveness of an intervention approach that combined explicit rule instruction with a systematic cueing hierarchy. Children with a mean age of 5;1 years received intervention delivered once weekly for one hour over eight weeks in small groups. Results showed that pre-intervention, children performed at 38.57% accuracy and improved to 75.97% accuracy post-intervention, demonstrating a 37.4% improvement. Importantly, the outcomes in this study were measured on the same test as the current study (i.e., the GET). The explicit intervention in the current study resulted in a mean 66% proportion correct post-intervention compared to 25.3% correct pre-intervention with a mean improvement of 40.7% when compared to a waiting control. Therefore, the explicit intervention in the current study yields mean improvement of similar, or even greater proportion to studies using implicit interventions.

Of note, the amount of time participants spent receiving intervention in the current study was markedly shorter than the studies mentioned above. For example, Eidsvåg et al. (2019) reported on an intervention duration of 18.75 hours, which is markedly longer than the 3.5-5 hours reported in the current study. Further, the current study was completed within 10 sessions, compared to 12 to 36 sessions reported in Owen Van Horne et al. (2017). This suggests the explicit intervention under investigation may be more time efficient than implicit approaches. As such, explicit interventions should indeed be considered a viable and perhaps more efficient intervention approach to improve morphosyntax in young children with DLD.

This further warrants future superiority trials comparing explicit and implicit interventions, which may shed further light on theoretical accounts of DLD.

Future Directions

Since recent findings support the use of various approaches to improve morphosyntax, including implicit and explicit, future research could aim to identify the active ingredients within interventions to test how to achieve 'optimal' outcomes. There is a body of literature supporting input-based intervention to treat verb morphology (Cleave et al; 2015; Eidsvåg et al., 2019; Owen Van Horne et al., 2017). Explicit approaches may initially appear converse to the recommendations for input-based implicit interventions. Nonetheless, explicit interventions have been shown to be efficacious in improving *-ed* production (Calder et al., 2018, 2020). Perhaps then, combining recommendations would be of interest. For example, the selection of complex *–ed* verbs (as per Owen Van Horne et al., 2017) for explicit intervention to improve production may lead to greater intervention gains on untrained forms.

Limitations

Firstly, participants were recruited using convenience sampling from a specialized educational program designed for children with DLD. This means participants may not be entirely representative of the clinical (and subclinical) population of children with DLD at large. Secondly, although RCTs are robust for comparing groups, the current design used a 'treatment-as-usual' waiting control instead of a comparison with 'gold-standard' intervention, which may better serve to determine intervention superiority. Nonetheless, the aim of the current trial was to evaluate efficacy of a theoretically motivated explicit intervention, and the implementation of a crossover phase for the control group contributes to the robustness of the current study. Next, the sample size of the current study was relatively small, even though comparable to recently published intervention studies (e.g., Eidsvåg et al., 2019; Owen Van Horne et al., 2017). Increasing sample size would allow the detection of small treatment effects which may be of clinical interest. Finally, the measures used in the current study are relatively static and may not be entirely representative of grammar used in functional communication. Future research should evaluate intervention effects on reliable measures obtained through language sample analysis.

Clinical Implications

The profile of participants from this study conform to predictions from the PDH, in that their static vocabulary skills as measured by the PPVT-4 appear to be within normal limits, whereas expressive and receptive grammar appear to be areas of deficit. Considering the positive response of the intervention group to the theoretically motivated explicit intervention compared to a 'treatment-as-usual' waiting control group examined in the current study, the use of explicit interventions should be considered a viable option for children exhibiting this clinical profile. However, this also highlights another caveat: the PPVT-4 is often used for research (cf., Rice & Watkins, 1996) and clinical (Eickhoff et al., 2010) purposes as a proxy for language skill. If children meeting criteria for DLD are, on average, performing within normal limits on such an assessment, perhaps this particular assessment will not result in the accurate grouping or characterization of children's language skills. Secondly, the limited transference of intervention effects to a linguistically related structure (i.e., *3S*) highlights the need for intervention to be highly targeted to the needs of children with DLD. Lastly, the current study does represent a more clinically relevant intervention frequency of once weekly (Finestack & Satterlund, 2018), and findings do suggest this frequency still results in intervention effects (cf. Calder et al., 2020).

Conclusion

As viewed through the lens of the PDH, the current efficacy study supports recommendations for intervention from this theory. Specifically, explicit grammar intervention using metalinguistic training with visual support and systematic cueing improves

–*ed* production in young school-aged children with DLD. Results from this efficacy study have also established the foundation for future intervention research which would further shed light on theoretical accounts of DLD and methods suited to improve morphosyntax in affected children. Perhaps by combining explicit intervention procedures with empirically supported input-based procedures, both researchers and clinicians can optimise grammar intervention effects for children with DLD.

References

Balthazar, C. H., Ebbels, S., & Zwitserlood, R. (2020). Explicit Grammatical Intervention for Developmental Language Disorder: Three Approaches. *Language, Speech,& Hearing Services in Schools, 51*(2), 226-246. https://doi.org/10.1044/2019_LSHSS-19-00046

Balthazar, C. H., & Scott, C. M. (2018). Targeting complex sentences in older school children with specific language impairment: Results from an early-phase treatment study. *Journal of Speech, Language, & Hearing Research, 61*(3), 713-728. https://doi.org/10.1044/2017_JSLHR-L-17-0105

Bishop, D. (2003). *Test for Reception of Grammar-2* (London: The Psychological Corporation). *CIT0003*.

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & Catalise-consortium, (2017). Phase 2 of CATALISE: a multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *Journal of Child Psychology & Psychiatry, 58*(10), 1068–1080. https://doi.org/10.1111/jcpp.12721

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & Catalise-consortium. (2016). CATALISE: A multinational and multidisciplinary delphi consensus study. Identifying language impairments in children. *PLoS ONE, 11*(7). https://doi.org/10.1371/journal.pone.0158753

Calder, S., D., Claessen, M., Ebbels, S., & Leitão, S. (2020). Explicit grammar intervention in young school-aged children with Developmental Language Disorder: An efficacy study using single-case experimental design. *Language, Speech, & Hearing Services in Schools, 51*(2), 298-316. https://doi.org/10.1044/2019_LSHSS-19-00060

Calder, S. D., Claessen, M., & Leitão, S. (2018). Combining implicit and explicit intervention approaches to target grammar in young children with Developmental Language Disorder. *Child Language Teaching and Therapy, 34*(2), 171-189. https://doi.org/10.1177/0265659017735392

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*(4), 284-290. https://doi.org/10.1037/1040-3590.6.4.284

Cleave, P. L., Becker, S. D., Curran, M. K., Van Horne, A. J. O., & Fey, M. E. (2015). The efficacy of recasts in language intervention: A systematic review and meta-analysis. *American Journal of Speech-Language Pathology*, *24*(2), 237-255. https://doi.org/10.1044/2015_AJSLP-14-0105

Clegg, J., Hollis, C., Mawhood, L., & Rutter, M. (2005). Developmental language disorders--a follow-up in later adult life. Cognitive, language and psychosocial outcomes. *Journal of Child Psychology & Psychiatry & Allied Disciplines*, *46*(2), 128-149. https://doi.org/10.1111/j.1469-7610.2004.00342.x

Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *Journal of Child Psychology & Psychiatry & Allied Disciplines*, *42*(6), 741-748. https://doi.org/10.1111/1469-7610.00770

Conti-Ramsden, G., Durkin, K., Toseeb, U., Botting, N., & Pickles, A. (2018). Education and employment outcomes of young adults with a history of developmental language disorder. *International Journal of Language & Communication Disorders, 53*(2), 237-255. https://doi.org/10.1111/1460-6984.12338

Dawson, J. I., Stout, C. E., & Eyer, J. A. (2003). *Structured Photographic Expressive Language Test—Third Edition (SPELT-3).* DeKalb, IL: Janelle Publications.

Dunn, L. M., & Dunn, D. M. (2007). *PPVT-4: Peabody picture vocabulary test*: Pearson Assessments.

Ebbels, S. (2007). Teaching grammar to school-aged children with specific language impairment using Shape Coding. *Child Language Teaching and Therapy, 23*(1), 67-93. https://doi.org/10.1191/0265659007072143

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

Ebbels, S. H., Maric, N., Murphy, A., & Turner, G. (2014). Improving comprehension in adolescents with severe receptive language impairments: a randomized control trial of intervention for coordinating conjunctions. *International Journal of Language & Communication Disorders, 49*(1), 30-48. https://dx.doi.org/10.1111/1460-6984.12047

Ebbels, S. H., van der Lely, H. K., & Dockrell, J. E. (2007). Intervention for verb argument structure in children with persistent SLI: A randomized control trial. *Journal of Speech, Language, & Hearing Research*, *50*(5), 1330–1349. https://doi.org/10.1044/1092-4388(2007/093)

Eidsvag, S. S., Plante, E., Oglivie, T., Privette, C., & Mailend, M. L. (2019). Individual versus small group treatment of morphological errors for children with evelopmental Language Disorder. *Language, Speech, & Hearing Services in Schools, 50*(2), 237-252. https://doi.org/10.1044/2018_lshss-18-0033

Eickhoff, J., Betz, S. K., & Ristow, J. (2010, June). Clinical procedures used by speech-language pathologists to diagnose SLI. In *Symposium on Research in Child Language Disorders*.

Finestack, L. H., & Satterlund, K. E. (2018). Current practice of child grammar intervention: A survey of speech-language pathologists. *American Journal of Speech-Language Pathology, 27*(4), 1329-1351. https://doi.org/10.1044/2018_ajslp-17-0168

Gillam, S. L., Gillam, R. B., & Reece, K. (2012). Language outcomes of contextualized and decontextualized language intervention: Results of an early efficacy study. *Language, Speech, & Hearing Services in Schools, 43*(3), 276−291. https://doi.org/10.1044/0161-1461%282011/11-0022%29

Hoffmann, T. C., Glasziou, P. P., Boutron, I., Milne, R., Perera, R., Moher, D., . . . Michie, S. (2014). Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ : British Medical Journal*, *348*, g1687. https://doi.org/doi:10.1136/bmj.g1687

Law, J., Rush, R., Schoon, I., & Parsons, S. (2009). Modeling developmental language difficulties from school entry into adulthood: literacy, mental health, and employment outcomes. *Journal of Speech Language, & Hearing Research, 52*(6), 1401-1416. https://doi.org/10.1044/1092-4388(2009/08-0142)

Lum, J. A. G., Conti-Ramsden, G., Morgan, A. T., & Ullman, M. T. (2014). Procedural learning deficits in specific language impairment (SLI): A meta-analysis of serial reaction time task performance. *Cortex, 51*(100), 1-10. https://doi.org/10.1016/j.cortex.2013.10.011

Lum, J. A. G., Ullman, M. T., & Conti-Ramsden, G. (2015). Verbal declarative memory impairments in specific language impairment are related to working memory deficits. *Brain and Language, 142*, 76-85. http://dx.doi.org/10.1016/j.bandl.2015.01.008

Lum, J. A., Conti-Ramsden, G., Page, D., & Ullman, M. T. (2012). Working, declarative and procedural memory in specific language impairment. *Cortex, 48*(9), 1138-1154. http://dx.doi.org/10.1016/j.cortex.2011.06.001

Owen Van Horne, A. J., Fey, M., & Curran, M. (2017). Do the hard things first: A randomized controlled trial testing the effects of exemplar eelection on generalization following therapy for grammatical morphology. *Journal of Speech Language, & Hearing Research, 60*(9), 2569-2588.https://doi.org/10.1044/2017_jslhr-l-17-0001

Plante, E., & Gomez, R. L. (2018). Learning without trying: The clinical relevance of statistical learning. *Language, Speech, & Hearing Services in Schools, 49*(3s), 710-722. https://doi.org/10.1044/2018_lshss-stlt1-17-0131

Rice, M. L., & Watkins, R. V. (1996). Show Me X": New views of an old assessment technique. In: Cole K., Dale, P., & Thal, D. (eds.) *Assessment of Communication and Language.* Baltimore, MD: Paul H. Brookes Publishing Co. 183-206.

Rice, M., & Wexler, K. (2001). Test of early grammatical impairment. *San Antonio, TX: The Psychological Corporation*.

Rice, M. L., Wexler, K., & Redmond, S. M. (1999). Grammaticality judgments of an extended optional infinitive grammar: evidence from English-speaking children with specific language impairment. *Journal of Speech Language, and Hearing Research, 42*(4), 943-961. https://doi.org/10.1044/jslhr.4204.943

Schulz, K. F., Altman, D. G., & Moher, D. (2010). CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ, 340*, c332. https://doi.org/10.1136/bmj.c332

Smith-Lock, K. M., Leitao, S., Lambert, L., & Nickels, L. (2013). Effective intervention for expressive grammar in children with specific language impairment. *International Journal of Language & Communication Disorders, 48*(3), 265-282. https://doi.org/10.1111/1460-6984.12003

Smith-Lock, K. M., Leitão, S., Prior, P., & Nickels, L. (2015). The Effectiveness of two grammar treatment procedures for children with SLI: A randomized clinical trial. *Language, Speech, & Hearing Services in Schools, 46*(4), 312-324. https://doi.org/10.1044/2015_LSHSS-14-0041

Spencer, T. D., Petersen, D. B., Slocum, T. A., & Allen, M. M. (2014). Large group narrative intervention in Head Start preschools: Implications for response to intervention. *Journal of Early Childhood Research, 13*(2), 196–217. https://doi.org/10.1177/1476718X13515419

Ullman, M. T. (2016). The declarative/procedural model: A neurobiological model of language learning, knowledge and use. In G. Hickok & S. A. Small (Eds.), *The Neurobiology of Language*. Elsevier. pp. 953-68.

Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specific to language: the procedural deficit hypothesis. *Cortex, 41*(3), 399-433. https://doi.org/10.1016/S0010-9452(08)70276-4

Warren, S. F., Fey, M. E., & Yoder, P. J. (2007). Differential treatment intensity research: a missing link to creating optimally effective communication interventions. *Mental Retardation & Developmental Disabilities Research Reviews, 13*(1), 70-77. https://doi.org/10.1002/mrdd.20139

Windsor, J., Scott, C. M., & Street, C. K. (2000). Verb and noun morphology in the spoken and written language of children with language learning disabilities. *Journal of*

Speech Language, & Hearing Research, 43(6), 1322-1336. https://doi.org/10.1044/jslhr.4306.1322

Zwitserlood, R., Wijnen, F., van Weerdenburg, M., & Verhoeven, L. (2015). 'MetaTaal': enhancing complex syntax in children with specific language impairment--a metalinguistic and multimodal approach. *International Journal of Language & Communication Disorders*, *50*(3), 273-297. https://doi.org/10.1111/1460-6984

List of Supplemental Materials

S1: CONSORT 2010 checklist of information to include when reporting a randomised trial.

S2: CONSORT 2010 Flow Diagram.

S3: Full intervention session plan with timing breakdown.

S4: Session-by-session summary of intervention targets and materials.

S5: Explicit intervention to improve past tense marking for early school-aged children with DLD template for intervention and description and replication (TIDieR) checklist.