Evaluating Two Different Dose Frequencies and Cumulative Intervention Intensities to Improve Past Tense Production for Early School-Aged Children with Developmental Language Disorder

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

Author Affiliations: [1]Curtin School of Allied Health, 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., Leitão, S., & Ebbels, S. (2021). Evaluating Two Different Dose Frequencies and Cumulative Intervention Intensities to Improve Past Tense Production for Early School-Aged Children with Developmental Language Disorder. *International Journal of Language and Communication Disorders.* hiips://doi.org/10.1111/1460-6984.12667

Address correspondence to:

Samuel D Calder
Curtin School of Allied Health, Curtin University, Perth, Western Australia,
GPO Box U1987, Perth
Western Australia, 6845.
Email: sam.calder@curtin.edu.au

Conflict of Interest Statement

The intervention approach described in this article was 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.

What this paper adds

What is already known on the subject

- Understanding the parameters of dosage and intensity are important for clinical practice.
- Research evaluating the efficacy and/or effectiveness of interventions delivered in different dose/intensity conditions is scarce.
- There appears to be different interpretations of what constitutes dosage and intensity in published research.

What this paper adds to existing knowledge

- This study retrospectively compared dosage and intensity conditions of intervention provided twice per week to intervention provided once per week. Both dose frequencies could be delivered in clinical settings.
- Results from this study were analysed by grouping data from multiple testing points, rather than comparing pre-post results. This approach demonstrated the variability of individual performance that would otherwise be lost with conventional methods of analysis.
- This study demonstrated that all past tense allomorphs improve to a similar degree when treated with this intervention.

What are the potential or actual clinical implications of this work?

- Parameters of dosage and intensity are still not clearly defined well enough for translation to clinical practice. In consideration of current research, this intervention may be more effective if delivered twice per week.
- If clinicians are treating past tense, all allomorphs should be considered as priorities for intervention targets.

ABSTRACT

Purpose: This study compared two dose frequency conditions of an explicit intervention with 50 trials per session designed to improve past tense marking in early school-aged children with developmental language disorder (DLD). The influence of allomorphs on intervention effects was also examined.

Methods: Data from previously conducted intervention studies were combined and analysed. Participants included $n = 9$ (mean age = 6;5 years) who received 20-30 minute intervention sessions provided twice per week for 10 weeks (1000 trials; 400-600 minutes) and $n = 20$ (mean age = 6;6) who received 20-30 minute intervention sessions provided once per week for 10 weeks (500 trials; 200-300 minutes). Repeated measures included criterion-referenced probes for production of untrained past tense verbs collected throughout baseline, intervention, and maintenance phases. The rate of progress in each phase was analysed using logistic regression. The proportion of participants who produced past tense allomorphs correctly at pre- intervention, post-intervention, and maintenance testing points was analysed.

Results: Logistic regression showed a stable baseline, highly significant progress during the intervention phase, and a marginally significant shallow decline during the maintenance phase. Those in the twice per week group showed a greater rate of progress during the intervention phase leading to significantly higher scores in the maintenance period when compared with the once per week group. The allomorphic category of past tense verbs did not appear to influence outcomes.

Conclusions: Participants receiving intervention twice per week appeared to demonstrate a greater rate of progress with intervention than those receiving it once per week, although once per week was also effective. However, these results should be interpreted with caution. Limitations to study design indicate a larger randomised controlled trial is required. All past tense allomorphs improve to a similar degree when treated with this intervention.

Developmental language disorder (DLD) is experienced by roughly 7% of the population (Norbury et al., 2016). The condition affects the ability to use and understand language, with language developing at a slower pace compared to age-matched peers (Bishop et al., 2017). Morphosyntax skills, such as tense marking, are particularly affected in this group (Bishop, 2014). Diagnosis of DLD usually involves collecting evidence of language difficulties through standardised assessments, language sampling, functional impact through curriculum-based assessment, and parent and teacher report (Bishop et al., 2016). Children with DLD are likely to face challenges with social communication and academic participation, with the impact of oral language deficits on interaction and literacy acquisition well established (Windsor et al., 2000).

Acquisition of certain finiteness markers is further complicated for young English language users, both communicatively and in their literate form, by the morphophonological constraints of the language. For example, there are morphophonological variations (allomorphs) of the regular past tense *-ed* morpheme, which are distributed based on the phonological properties of the verb stem. If the final sound in a verb is a voiceless posterior plosive, affricate or fricative, regular past tense is marked with [t] (e.g., *walk**ed**, watch**ed**, kiss**ed***). However, if a verb ends in a voiced posterior plosive, affricate or fricative, or a vowel, regular past tense is marked with [d] (e.g., *jogg**ed**, ag**ed**, buzz**ed***). Finally, if a verb ends in an alveolar plosive such as [t] or [d], regular past tense is marked with the unstressed syllable [əd] (e.g., *tast**ed**, wad**ed**, need**ed***). Therefore, without robust grounding in oral language skills to encompass the allomorphic variations in tense marking, children with DLD are likely to face greater barriers in terms of their literacy success.

There is a need to evaluate interventions aiming to improve those skills necessary to access school-based curriculums. Research should not only measure change on those items targeted during intervention, but also any impact on measures of generalisation (Owen Van

Horne et al., 2018). Further, the components of interventions should be clearly operationalised in order to evaluate effectiveness. For example, Plante et al. (2019) evaluated whether the same number of doses delivered in half the time could be effective when treating verb morphology for 20 four to five year old children with DLD. Results from between group post-intervention testing suggests dose density can be manipulated as such, and offers evidence that dosage is an important variable to explore regarding intervention. This finding supports the study of intervention dosage with the goal of providing sufficient intervention at a critical time to support foundation skills for functional participation, such as literacy acquisition at early school-age.

Dose Frequency, Intervention Duration, and Cumulative Intervention Intensity

Warren et al. (2007) outlined a framework for defining intervention components to encourage researchers and clinicians to identify elements contributing to optimal intervention effectiveness. This includes clearly defining *dose*, which refers to "the number of properly administered teaching episodes during a single intervention session" (p. 71); *dose form*, which refers to "the typical task or activity within which the teaching episodes are delivered" (p. 71); *dose frequency*, which refers to "the number of times a dose of intervention is provided per day and per week" (p. 72); *total intervention duration*, which refers to "the time period over which a specified intervention is presented" (p. 72), and; *cumulative intervention intensity*, which refers to "the product of *dose* x *dose frequency* x *total intervention duration*" (p. 72). For example, Calder et al. (2020) reported on an explicit intervention that included 50 properly administered teaching episodes (i.e., trials) to produce regular past tense in 20-30 minute sessions provided twice per week over 10 weeks, resulting in cumulative intervention intensity of 1000 trials over 400-600 minutes of intervention. By defining the multiple parameters that comprise intervention intensity, researchers and clinicians are able to determine if children differ in their response to receiving different intervention intensities.

Warren et al. (2007) also stressed the need to look beyond the notion that more practice is always better. For instance, the Procedural Deficit Hypothesis (Ullman & Pierpoint, 2005) suggests children with DLD experience grammatical difficulties as a result of an implicit learning deficit through impaired procedural memory. Importantly, the hypothesis predicts children with DLD have relatively spared declarative memory, which underlies learning of explicit information. Therefore, grammar, which is learned implicitly by typically developing children, may potentially be learned explicitly by children with DLD if intervention is presented as such. Notably, information learned through declarative memory is strengthened by repeated exposures (Ullman, 2016), so the theory would indicate more practise is better.

A current survey of speech language pathology practices in the US evaluated how often clinicians were able to implement intervention parameters, such as dose frequency, as well as probing speech language pathologists' desired dose frequency if resources were unlimited (Finestack & Satterlund, 2018). Results indicated that most clinicians provided sessions once weekly but would prefer to provide twice the dose frequency to essentially double the cumulative intensity. Therefore, evaluating potential differences in the responsiveness to intervention through contrasting dose frequencies is of clinical interest.

Dosage and Intensity, and Intervention Effectiveness

Warren et al. (2007) claimed that there had been virtually no systematic analyses of the effects of varying intervention intensities. Since this time, there have been advancements in the area, but still very few studies compare the same dose form across other intervention parameters (e.g., dose frequency, total intervention duration, cumulative intervention intensity) to demonstrate efficacy and efficiency. A recent systematic review and narrative synthesis concluded that when dose, total intervention duration, and cumulative intervention intensity (in hours) is controlled, there is no advantage to increasing dose frequency to

improve morphosyntax for young children with DLD (Frizelle et al., 2021). For example, an efficacy study of 16 four-to-five year old children revealed that children receiving enhanced conversational recasting intervention in spaced (three 10 minute sessions daily) versus massed (one 30 minute session daily) conditions performed comparably at post-intervention testing points (Meyers-Denman & Plante, 2016). Results suggested daily dose frequency may be a flexible component of this particular intervention. However, it is difficult to comment on the transferability of this finding to clinical practice, as the daily provision of intervention over a period of five weeks is unlikely to suit real-life clinical contexts (see Finestack & Satterlund, 2018).

In another study of 36 five-year-old children with DLD, Smith-Lock et al. (2013b) compared a daily intervention condition (i.e., once a day for eight days) to a weekly intervention condition (i.e., once a week for eight weeks). Pre- to post-intervention between group differences revealed the once weekly condition was more effective than the once daily condition in improving individualised morphosyntax targets for children with DLD. However, single case analyses revealed individual differences between children within both groups, which stresses the need to individualise dose frequencies depending on the specific needs of each child in clinical practice. Notably, the provision of intervention in the study was at the classroom and small group level. Effects may therefore have been washed out through the provision of intervention in small group contexts. In addition, dose (i.e., trials) was not reportedly pre-determined, thus there is no way to verify if one participant received a higher dose than another, which may have influenced outcomes. Frizelle et al.'s (2021) recent review also identified a lack of advantage for interventions with longer total intervention durations defined as total number of intervention sessions. Collectively, results from studies evaluating dose frequency and total intervention sessions as parameters of intervention suggest that there may in fact be a point of diminishing returns for morphosyntax. If this is

the case, then perhaps conceptualising dose in terms of cumulative intervention intensity in terms of the total number of doses (rather than in hours or minutes) is the most important parameter to consider when evaluating intervention efficacy.

Allomorphy of Regular Past Tense and Implications for Intervention

Another element of interest when evaluating interventions to improve regular past tense marking is that of the allomorphic categories associated with inflection. Several explanations have been presented to account for potential patterns of difficulty with regular past tense. For example, Leonard and colleagues have suggested processing deficits underpin errors with past tense morphology due to its low perceptual salience relative to the bare stem to which it is affixed (Leonard, 1989; Leonard et al., 1997). A processing account would predict an advantage to learning morphemes that are more perceptually salient, for example [ɘd] in *lift**ed*** (Leonard et al., 1997). Effects of phonological complexity are reported to influence past tense production, where inflected verbs without monomorphemic cognates (e.g., voiced obstruent +/d/: *ru**bbed**, jogg**ed**, ju**dged***), are produced with lower accuracy in children with DLD compared to typically developing children (Marshall & van der Lely, 2006), and segmental allomorphs inflected as singleton + [d] codas (e.g. *cri**ed**, stirr**ed**, play**ed***) are produced with higher accuracy than segmental allomorphs produced as consonant + [d]/[t] codas (Oetting & Horohov, 1997; Tomas et al., 2015). Van der Lely and Ullman (2001) suggested lexical frequency effects may explain difficulties in past tense production, where children with DLD produce high-frequency verbs with higher accuracy than low-frequency verbs while no such difference in production was observed in typically developing children. Owen Van Horne and Green-Fager (2015) have suggested that acquisition of regular past tense is influenced by lexical frequency phonological complexity, and lexical aspect for both children with and without DLD. In particular, verbs that are infrequently marked for past tense (e.g., *fish**ed***), phonologically complex (e.g., *jump**ed***), and atelic (e.g., *walk**ed***) (harder verbs) are produced

with lower accuracy than verbs that are frequently marked for past tense (e.g., *play**ed***), phonologically simple (e.g., *cri**ed***), and telic (e.g., *clos**ed***) (easier verbs). The implication for intervention, however, appears to be that targeting harder verbs through intervention results in more rapid improvement and generalisation to connected speech (Owen Van Horne et al., 2018).

Recent evidence suggests an effect of syllabicity in the acquisition of regular past tense, where the syllabic [əd] allomorph appears to be the most difficult for children with DLD (Calder et al., 2021b; Tomas et al., 2015, 2017). In fact, Tomas et al. (2017) suggested that even in typically developing children, verbs inflected for past tense with the [əd] allomorph are still developing at five years old, while verbs inflected with [t] or [d] appear to already be mastered by this age. Counter to processing accounts, which suggest an advantage to learning syllabic allomorphs (Leonard et al., 1997), a broad effect of syllabicity in regular past tense acquisition may reflect a failure to correctly detect the difference between grammatical and ungrammatical rules through a statistical learning deficit (Leonard & Deevy, 2017). For example, although segmental allomorphs may be correctly learned by some children with DLD, uninflected forms with [d] / [t] codas (e.g., *land*, *taste*) may be incorrectly interpreted as grammatical in past tense contexts, suggesting more general deficit in the use of rule-governed computations of morphological sequencing (Plante & Gomez, 2018; Ullman & Pierpoint, 2005). Intervention could plausibly act to increase statistical regularity through clinician-child interactions (e.g., Plante et al., 2014), or alternatively aim to reduce demands on implicit learning through environmental modifications such as visual support and explicit instruction to direct the child to draw upon meta-awareness to facilitate application of a learned grammatical rule (e.g., regular past tense verbs ending in alveolar obstruents are marked with [əd]).

The Current Program of Research

The Procedural Deficit Hypothesis (Ullman & Pierpoint, 2005) suggests that explicit interventions may be beneficial for children with DLD to improve grammar due to an implicit learning deficit. Recent findings from Calder et al. (2018) and Calder et al. (2020, 2021a) suggest explicit grammar intervention combining metalinguistic training and a systematic cueing hierarchy is efficacious for treating production of regular past tense for children aged around six years with DLD. The metalinguistic training aspect is based on the SHAPE CODING™ system (Ebbels, 2007), which uses specific visual cues such as colours, shapes and arrows to explicitly teach children rules of syntax and morphology. The systematic cueing hierarchy was developed by Smith-Lock et al. (2015) and scaffolds children from least to most support to produce morphosyntax targets correctly following an error.

The primary components of the intervention are first to explicitly outline the goal of the session, which is to mark verbs for regular past tense when describing something that has already happened. Subsequent to the initial session, this also includes reminding the child that s/he had practised this in the last session. The vocabulary necessary to use the past tense verbs appropriately is checked by having the child label intervention materials. Relevant shapes from the SHAPE CODING™ system are introduced (Subject: Oval, Verb Phrase: Hexagon, Object: Rectangle; see Balthazar et al., 2020 or Ebbels, 2007 for a detailed description of the system and shapes) and linked to the intervention materials. Subject + Verb/+ Object (SV/O) clauses are modelled including target verbs marked with one allomorph from each of the allomorphic categories, e.g., "*I tap the drum. What DID I just DO? I tapp**ed** the drum. The [t] at the end of tapp**ed** lets us know it's already happened*."; "*I twirl the drumstick. What DID I just DO? I twirl**ed** the drumstick. The [d] at the end of twirl**ed** lets us know it's already happened.*"; "*I lift the drum. What DID I just DO? I lift**ed** the drum. The [əd] at the end of lift**ed** lets us know it's already happened."* Additional visual support in the form of left "down arrows" from the SHAPE CODING™ system are

highlighted at this point. For the practise component, the child is then provided 25 opportunities to produce past tense in response to an interrogative to elicit the target verbs, e.g., "*You tap the drum. What DID you just DO?*" If the child produces an error on the target verbs, s/he is cued systematically in the following sequence: 1) *request for clarification*; 2) *emphatic recasting with elicited response*; 3) *forced choice with elicited response*; 4) *elicited imitation*. A consolidation component is then implemented, and the child responds to comprehension questions relating to the targeted syntactic structure, e.g. "*WHO tapped the drum?*"; "*What DID you DO?*"; "*WHAT did you tap?*" The consolidation component is repeated for one exemplar from each of the allomorphic categories providing the child with three opportunities to produce the target SV/O structure including a verb marked for regular past tense. The shapes and arrows are then removed, and the child is prompted to produce the same SV/O structures elicited at the beginning of the consolidation component. Finally, SV/O sentences are modelled for the child either grammatically or ungrammatically (by omission of past tense *–ed*), and the child decides if it sounds right. The process is repeated with a selection of different verbs for a second activity until a total of 50 trials (i.e., doses) are achieved within the session. Each session includes a summary component where the child is reminded that it is important to include the sounds at the end of past tense verbs when describing something that has already happened.

Calder et al. (2020) reported on a single case experimental design study with $n = 9$ children. The intervention was delivered twice weekly in 20-30 minute sessions for 10 weeks where each child received a cumulative intervention intensity of 1000 trials over 400-600 minutes. Results indicated children significantly improved in the production of trained ($p < .001$, $Tau = 0.88$) and untrained ($p < .001$, $Tau = 0.64$) past tense verbs. In a subsequent study, Calder et al. (2021a) conducted a randomised controlled trial with $n = 21$ children allocated to intervention ($n = 10$) and waiting control ($n = 11$) conditions. The intervention

was delivered once weekly in individual 20-30 minute sessions for 10 weeks where each child received a cumulative intervention intensity of 500 trials over 200-300 minutes. The intervention group significantly outperformed the waiting control group on a measure of untrained past tense production ($p < .001$, $d = 3.03$), and once the control group crossed over to the intervention condition, between-group differences disappeared, but significant improvement pre-post intervention was shown for both groups combined ($p < .001$, $d = 1.22$). These results suggest this explicit intervention is efficacious for treating regular past tense for early school-aged children with DLD. Considering the relative lack of evidence evaluating grammar interventions in different dose conditions (see Frizelle et al., 2021), and the reported preference to provide more intervention rather than less (see Finestack & Satterlund, 2018), re-analysis of previously reported data comparing dose frequency may provide valuable information regarding dosage to optimise efficacy.

The Current Study

The current study combines data from previously reported research to systematically analyse whether intervention for past tense provided twice per week or once per week results in greater rate of progress on a generalised measure of untrained regular past tense production and is therefore optimally efficacious. Such research should inform intervention protocols for determining dose to establish effectiveness in future research and may inform clinical practice in the current absence of research with large group comparison studies. Our primary research question was:

1. Do children in the two different dose frequency and cumulative intervention intensity groups differ in their response to intervention?

   We hypothesised that intervention provided twice per week (i.e., cumulative intervention intensity 1000 trials over 400-600 minutes over 10 weeks) would result in a greater rate

of progress on measures of past tense production compared to intervention provided once per week (i.e., cumulative intervention intensity 500 trials over 200-300 minutes over 10 weeks).

In addition, exploratory analyses of the characteristics of verb finiteness marking, such as allomorphic categories, may increase our understanding of patterns within disordered grammar to inform the design of interventions. Our second research question was:

2. Is there any difference in response to intervention between the three allomorphs? We hypothesised, on average, children with DLD will show greater progress in percentage accuracy in production of past tense marking of [d] and [t] allomorphs compared to [əd].

METHOD

Study Design

The current study tested hypotheses by retrospectively comparing results from previously conducted studies. Data from n = 9 children receiving intervention twice per week for 10 weeks as part of a single case experimental design (twice per week = 2PW) (Calder et al., 2020) were compared to data from n = 20 children who received the same intervention once per week for 10 weeks as part of a randomised control trial (once per week = 1PW) (Calder et al., 2021a).

Participants

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)**. All participants were recruited through specialised educational programmes. Children who attended the specialised

education programs had a previously established diagnosis of DLD as determined by assessment through a multidisciplinary team, which included speech-language pathologists, psychologists, teachers, and in some cases, paediatricians. Children met diagnostic criteria outlined by Bishop et al. (2016), including: language skills significantly below age expectancy with onset in early development, and absence of other biomedical conditions that may otherwise account for language difficulties, such as autism spectrum disorder or intellectual disability. Further, evidence to support a diagnosis of DLD to meet eligibility criteria for the specialised educational programme included functional impact demonstrated through various sources, including language samples, student achievement reports, parent report, and teacher observation.

Recruitment involved consent from the principals, and identification of potential study participants by programme staff. Once identified, parents of the children were sent information letters and consent forms, as well as child friendly forms for the children to complete. Forms were returned to the school to confirm consent. Once consent was obtained, participants' case files were examined to confirm they met criteria for DLD.

Initial Assessment

Following recruitment, participants were screened in hearing acuity and passed at 20 dB HL for each ear at 500, 1000, 2000, and 4000 Hz. Additionally, all participants passed the Phonological Probe from the Test of Early Grammatical Impairment (Rice & Wexler, 2001) which confirmed they were able to articulate phonemes necessary for expressive morphosyntax targets.

To determine participant suitability to receive intervention designed to improve past tense marking, the regular past tense subtest of the Grammar Elicitation Test (GET-ed) (Smith-Lock et al., 2013a) was administered. This criterion-referenced subtest includes 30

items evenly distributed for all possible allomorphs (i.e., 10x [d] verbs, 10x [t] verbs, and 10x [əd] verbs). The GET has been used in several studies to identify intervention targets and/or measure intervention progress (Calder et al., 2018, 2020, 2021a; Smith-Lock et al., 2013a, 2013b, 2015). The results from children with DLD on the GET have been further analysed to explore patterns of morphosyntactic difficulty (Tomas et al., 2015) as well as conduct preliminary psychometric assessment of the measure (Calder et al., 2021b). For the purpose of the current study, verbal elicitation procedures for each item were pre-recorded to ensure consistency in administration and embedded within a Microsoft PowerPoint delivered via laptop.

All participants were also assessed on standardised tests of grammar to report descriptively on their general grammar abilities. The Structured Photographic Expressive Language Test 3rd Edition (SPELT-3) (Dawson et al., 2003) measures expressive morphosyntax using 54 items across a range of structures and was normed on children aged 4-to-9 years. A score of 95 or less (-0.33 standard deviation) has been recommended when using the SPELT-3 for the purpose of differentiating preschool and early school-aged children with and without DLD at 90% sensitivity and 100% specificity (Perona et al., 2005). The Test for Reception of Grammar 2nd Edition (TROG-2) (Bishop, 2003) measures the comprehension of a total of 20 different grammatical structures and was normed on children aged 4-to-16 years. Both tests are reported to have strong reliability and appropriate validity. It is important to note that neither the SPELT-3 nor TROG-2 are designed to diagnose DLD, and determining such a diagnosis on standardised assessments alone is not recommended (Bishop et al., 2016). Indeed, children with DLD represent a heterogeneous clinical population with varied profiles of strengths and weaknesses. However, regular past tense appears to be a reliable marker of the disorder at early school-age (Redmond et al., 2019; Rice et al., 1998).

Table 1.

Demographic and initial assessment information for all study participants.

Participant ID	Age at initial assessment	Age at onset of intervention	Sex	SPELT-3	TROG-2	GET -ed (%)
2PW						
P01	6;3	6;4	M	69	74	36.7
P02	6;2	6;5	M	90	97	40.0
P03	5;10	5;11	M	79	86	33.3
P04	6;8	6;11	M	71	81	30.0
P05	6;6	6;7	M	57	81	3.3
P06	6;2	6;5	F	72	65	3.3
P07	6;7	6;9	M	84	62	23.3
P08	6;0	6;2	M	69	79	16.7
P09	6;1	6;3	M	57	65	36.7
M(SD)	6;3 (0;3)	6;5 (0;4)	88.9%/11.1%	69 (17.8)	76.7 (8.9)	24.8 (13.3)
1PW						
P10	6;0	6;2	F	78	79	26.7
P11	5;11	6;1	M	98	72	46.7
P12	6;1	6;3	F	98	74	36.7
P13	6;1	6;3	M	82	88	23.3
P14	6;4	6;6	M	82	111	50.0
P15	6;0	6;1	M	72	69	3.3
P16	6;0	6;2	M	57	74	20.0
P17	6;4	6;5	M	55	79	10.0
P18	6;8	6;9	M	48	55	10.0
P19	6;3	6;5	M	94	74	46.7
P20	6;7	7;0	M	63	72	16.7
P21	5;9	6;1	M	105	90	43.3
P22	5;10	6;2	M	81	81	13.3

P23	6;7	7;0	M	59	76	3.3
P24	6;2	6.6	F	63	65	3.3
P25	6;2	6;7	F	82	83	33.3
P26	6;7	7;0	M	48	72	3.3
P27	6;2	6;6	M	40	65	6.7
P28	6;4	6;9	M	61	83	16.7
P29	6;7	7;1	F	71	90	23.3
M(SD)	6;3 (0;3)	6;6 (0;4)	75%/25%	71.9 (18.0)	77.6 (11.5)	21.8 (15.7)

Notes. 2PW= intervention 2x per week; 1PW= intervention 1x per week; CONTROL= 'treatment-as-usual' waitlist control group; GET -ed= Grammar Elicitation Test-regular past tense; 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; SD= standard deviation. All scores on standardised assessments reported in this table are scaled scores.

Study assignment, and Demographic and Baseline Variables

Table 1 summarises study assignment, and demographic and other baseline information. The participants in the 2PW condition were not randomly assigned as the Calder et al. (2020) study used a single case experimental design. Rather, study enrolment was capped at $n = 9$ to meet standards for single case reporting (Kratochwill et al., 2012). Participants in the 1PW condition were randomly assigned to either the intervention condition ($n = 10$) or a waiting control ($n = 11$) by a researcher blinded to the purpose of the study (Calder et al., 2021a). Of note, one child was exited from the waiting control group prior to receiving intervention, and their data have been excluded from all analyses resulting in $n = 29$. Therefore, group assignment was not randomised for the purpose of evaluating efficacy relative to dosage in the current study. The intervention for the two studies was also provided at different times. These are considered limitations.

There were no significant differences between the 2PW and 1PW groups on age at initial assessment, $t(27) = -.151$, $p = .881$, or onset of intervention, $t(27) = -.480$, $p = .635$, and no differences in distribution of sex, $U = 77.50$, $z = -.839$, $p = .401$. There were no between group differences on the GET-ed at initial assessment, $t(27) = .479$, $p = .635$, or at the onset of intervention, $t(27) = .404$, $p = .689$. There were no significant group differences on the SPELT-3, $t(27) = -.022$, $p = .982$. Assumptions of normality were violated for 1PW on the TROG-2 (kurtosis = 2.53). Therefore, a Mann-Whitney U test was conducted, which was non-significant, $U = 87.50$, $z = -.118$, $p = .906$. Of note, overall three participants (P11, P12, P21) were considered within the average range on the SPELT-3 at initial assessment using the 95 cut score (Perona et al., 2005), and six participants (P02, P03, P13, P14, P21, P29) would be considered within the average range on the TROG-2 using a one standard deviation cut-off. It is important to note that, of 54 items, the SPELT-3 only has one item testing regular past tense production, and the TROG has none. Further, the SPELT tests multiple

exemplars of structures suggested to be relatively spared in children with DLD, such as progressive *–ing*. Thus, Perona et al. (2005) recommends a stricter -0.33 SD to identify (not diagnose) language disorder, as the test is likely to otherwise overestimate children's expressive grammar skills. Nonetheless, when considering the lower bound of the average range on the standardised tests, all included participants demonstrated difficulties on at least one of the standardised grammar measures. Further, measures of past tense production have been shown to be a reliable indicator of DLD (e.g., Redmond et al., 2019; Rice et al., 1998). In the current study, all included participants performed poorly on the GET-ed as a measure of past tense accuracy (≤ 50% accuracy). Bivariate correlations between the pre-intervention scores on the GET-ed and standardised grammar tests were conducted to evaluate the relationship between past tense production and general grammar measures. Correlations between the GET-ed and pre-intervention standardised grammar measures were positive and strong (SPELT-3: $r = 0.72$, $p < .05$) or moderate (TROG-2: $r = 0.47$, $p < .05$). The weaker correlation between the GET-ed and the TROG-2 is unsurprising given the differences in test items, format, and whether the task is receptive or expressive. Therefore, it is possible participants may have shown disproportionate strengths in receptive grammar compared to past tense production.

Overall, considering various evidence, including an existing diagnosis of DLD determined through multidisciplinary assessment and placement at a specialised educational programme, poor performance on a measure of past tense, and performance predominantly below the average range on at least one standardised (although not diagnostic) grammar test, the participants were considered to have a valid DLD diagnosis.

Intervention

The SHAPE CODING™ system (Ebbels, 2007) was used in combination with a systematic cueing hierarchy (Smith-Lock et al., 2015) to explicitly teach children with DLD when to mark regular past tense verbs. Intervention was carried out in a quiet space on site at the participants' educational programme. Each session was videoed for later fidelity rating. Scoring for fidelity involved the first author and two separate blinded researchers for each study rating 20% of total sessions on percentage accuracy for inclusion of intervention components, including explicit instruction with visual scaffolds, number of verbs elicited, errors cued correctly, and an intervention session plenary. Between-observer agreement was calculated using intraclass correlation coefficients (ICC) with absolute agreement and single measures in a two-way random effects model. The average score across raters was 97.95% for percentage accuracy, and ICC for procedures was .976 indicating excellent agreement.

For convenience, the dose, duration, and intervention intensity conditions are summarised briefly here. For the 2PW group, 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 twice weekly; total intervention duration was 10 weeks; resulting in a cumulative intervention intensity of 1000 trials over 400-600 minutes over 10 weeks. The only difference for the 1PW group was that dose frequency was once weekly, halving the cumulative intervention intensity to 500 trials over 200-300 minutes over 10 weeks. All intervention was delivered within age-appropriate and engaging activities, such as snakes and ladders, and playing with animal manipulatives. All intervention was delivered in 1:1 sessions by the first author, who is an experienced speech-language pathologist. The study involving the twice weekly sessions was carried out before the study involving once weekly sessions.

Outcomes

Past Tense Production

The full 30-item GET-ed subtest was administered to all participants at initial assessment, immediately prior to the commencement of intervention, immediately following the cessation of the 10-week intervention phase, and at the end of the five-week maintenance phase. The sequence of presentation was randomised at each testing point. Importantly, verbs included in the GET-ed were not trained as part of the intervention, and therefore served as measures of generalisation of past tense production. For the 2PW group, baseline phases varied in duration from five-, seven- and nine-weeks. All participants in the 1PW group had a five-week baseline phase.

Repeated Measures. Sets of nine past tense verbs were probed weekly during baseline and maintenance phases. During the intervention phase, past tense production was probed at the beginning of the second intervention session, and every even session in both the 2PW and 1PW conditions during the intervention phase. The verbs were equally distributed for allomorphs and randomly selected from the GET-ed at each timepoint.

RESULTS

The primary aim of this study was to systematically evaluate whether the participants in the two dose frequency conditions differed in their response to intervention. We also conducted an exploratory analysis of whether allomorphs differ in response to intervention.

Outcomes and Estimation

Past Tense Production

All data for past tense production, including initial assessment, pre-intervention, post-intervention and maintenance (i.e., GET-ed /30), as well as repeated measures probed

throughout baseline, intervention and maintenance phases (i.e., GET-ed /9) were included for analysis. Figure 1 shows mean scores and standard errors at each week together with fitted regression lines for each phase, split by group. Visual inspection of the graph indicates a very slight increase by Week during the baseline period, a steeper linear increase from Week 1 to Week 10 during the intervention period for both groups. This increase is steeper for the 2PW group, leading to higher scores during the maintenance period for the 2PW than 1PW group. There appears to be a shallow decline across the maintenance period, particularly for the 2PW group. However, the final scores remain higher than the pre-intervention scores.

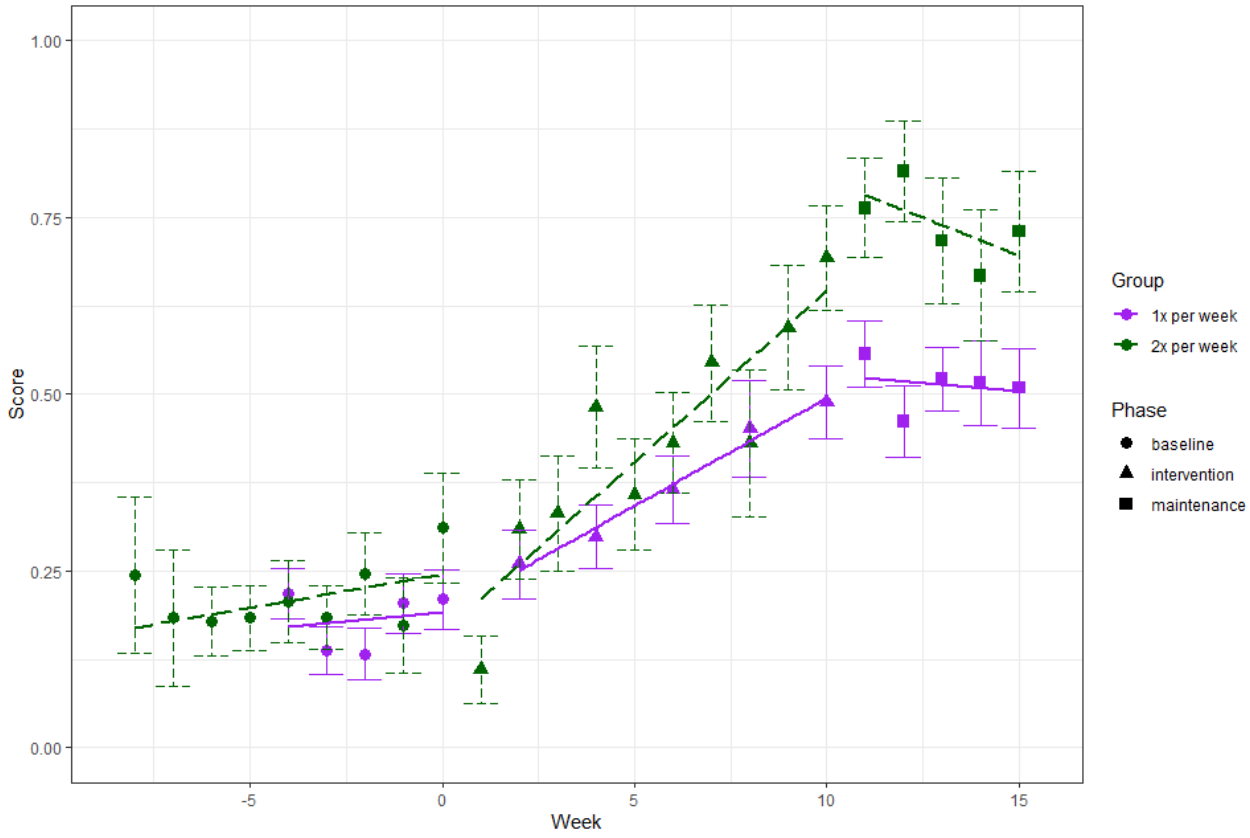


Figure 1. *Mean percent accuracy of past tense production during baseline, intervention, and maintenance phases for the two groups.*

Data in each Phase were analysed separately using logistic regression. This type of general linear model (Howell, 2010) is used for a binomially distributed dependent variable (i.e., correct versus incorrect), and accounts for random effects, such as differences

between individuals. We predicted the proportion of correct responses with the fixed effects of Week and Group (1PW vs 2PW). The interaction between Group and Week was our primary interest, i.e., did receiving one versus two sessions per week affect the rate of progress during the intervention phase, or the retention of any progress during the maintenance phase? Because of this, we ran the full interaction model at each Phase. The models also included a random intercept for participants to model individual differences in initial performance. Analyses were carried out in R version 4.0 using the glmmTMB package (version 1.0.2.1). The results are shown in Table 2.

Table 2.

Results from the logistic regression with week and group as predictors of past tense production in baseline, intervention, and maintenance phases.

	Baseline Phase		**Intervention Phase**		**Maintenance Phase**	
Predictors	*Odds Ratios*	*p*	*Odds Ratios*	*p*	*Odds Ratios*	*p*
Group [2PW]	1.42 (0.61 – 3.28)	0.412	0.88 (0.37 – 2.09)	0.774	16.01 (2.51 – 102.29)	**0.003**
Week	1.01 (0.94 – 1.09)	0.815	1.17 (1.11 – 1.23)	**<0.001**	0.94 (0.88 – 1.00)	**0.040**
Group [2PW] x Week	1.03 (0.93 – 1.13)	0.603	1.09 (1.00 – 1.17)	**0.040**	0.89 (0.79 – 1.01)	0.070
Random Effects						
σ^2	3.29		3.29		3.29	
τ_{00}	0.95 $_{ID}$		0.76 $_{ID}$		1.24 $_{ID}$	
ICC	0.22		0.19		0.27	
N	29 $_{ID}$		29 $_{ID}$		29 $_{ID}$	
Observations	163		190		139	
Marginal R^2 / Conditional R^2	0.004 / 0.228		0.077 / 0.251		0.080 / 0.332	

Week was not a significant predictor during the baseline phase, indicating a stable baseline, but was a significant predictor in the intervention and maintenance phases. In the intervention phase, for the reference group (1PW), the odds of a correct response increased significantly by Week ($p < .001$), whereas during the maintenance phase, this decreased (although this decline was of marginal significance, $p = 0.04$). A marginally significant interaction of Group with Week in the intervention phase ($p = 0.04$), indicates that the rate of progress differed between groups in favour of the 2PW condition. The odds of a correct response were also significantly higher at the start of the maintenance phase for the 2PW group compared with the reference group of 1PW ($p = 0.003$). Although it appears their decline during this phase was steeper than the 1PW group, this did not reach significance ($p = 0.07$).

The results therefore suggest that while progress with intervention was significant in the 1PW group, the rate of progress with intervention was steeper in the 2PW group. In addition, the 2PW group showed significantly higher scores in the maintenance phase, but with indications of a steeper decline. This suggests that while intervention once per week leads to significant progress, there may be an advantage to receiving intervention twice per week. Limitations to this interpretation are presented in the Discussion.

The rate of progress on past tense production for each individual participant through baseline, intervention and maintenance phases is presented in Figure 2. This figure demonstrates that although there was significant improvement in both dose conditions when the participants were considered as a group, there was great variability in individual performance in response to the intervention, regardless of how often the dose was administered.

Predictors of Intervention Progress

Bivariate correlations between pre- to post-intervention progress, and SPELT-3 and TROG-2 scores were run to evaluate whether pre-intervention scores may predict progress across participants. Both standard and raw scores on the SPELT-3 and TROG-2 were included. Correlations for the SPELT-3 standard (r = -0.16, p = 0.41) and raw (r = -0.03, p = 0.87) and TROG-2 standard (r = 0.27, p = 0.16) and raw (r = 0.12, p = 0.55) scores, and pre- and post-progress were weak and non-significant.

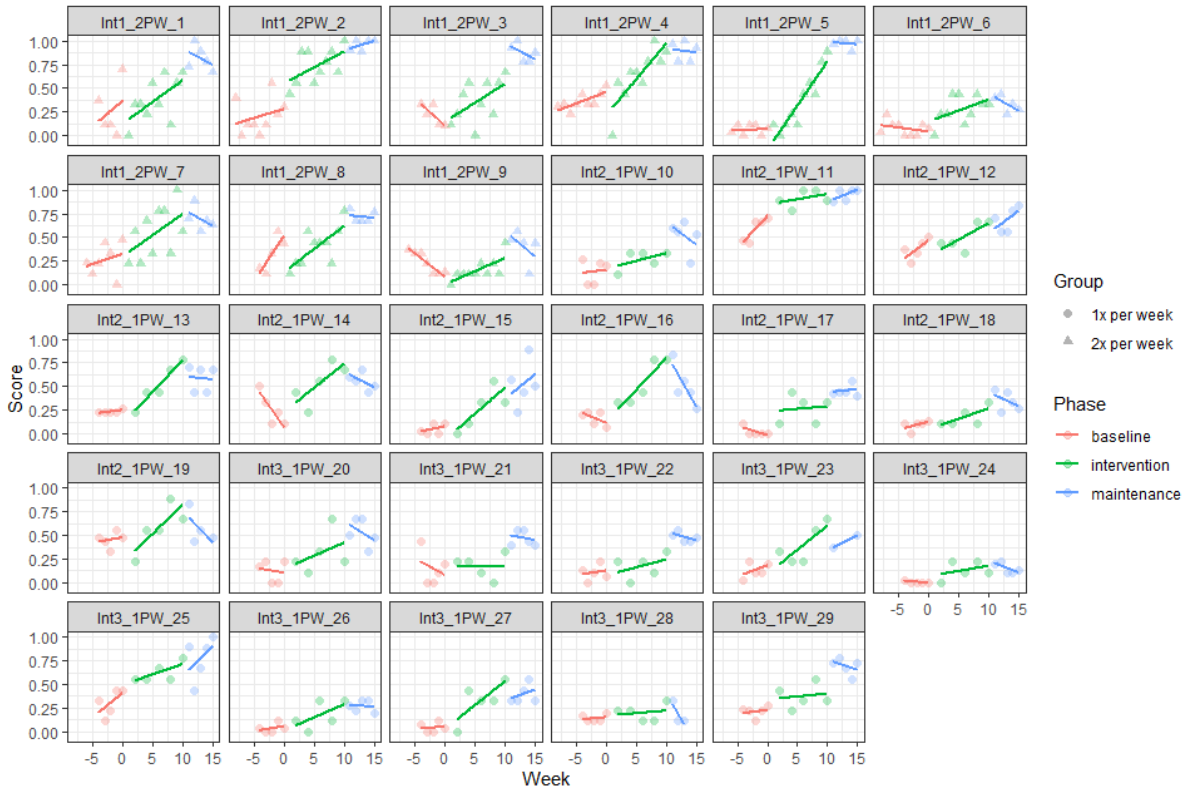
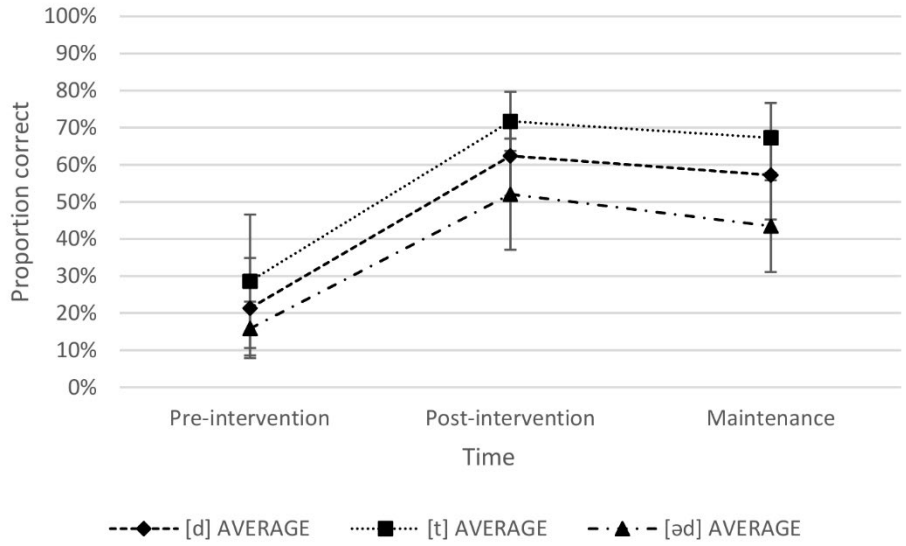


Figure 2. *Rate of progress in past tense production for all individual participants during baseline, intervention and maintenance phases.*

Analysis of Allomorphs

Data for past tense production from immediately pre-intervention, post-intervention and maintenance testing points (i.e., GET-ed /30) were included for analysis. Performance



from all participants who completed intervention were combined. Responses to past tense production probes were organised according to allomorphs ([d], [t], [əd]). Correct/incorrect responses for each item (*Total* = 30) across participants were summed and converted to percentage of participants correct at pre- and post-intervention, and maintenance testing points, i.e. *X*% of participants produced item *Y* correctly pre- intervention, *X*% of participants produced item *Y* correctly post-intervention, *X*% of participants produced item *Y* correctly at the maintenance testing point. To evaluate whether allomorphic categories accounted for variance in intervention effects, a 3x3 mixed design ANOVA was used where time (pre- vs. post-intervention vs. maintenance) was the within-subject variable and allomorphic category ([d] vs. [t] vs. [əd]) was the between-subject variable (see Figure 3). This analysis was carried out using IBM SPSS version 25.

Figure 3. *Average performance across past tense allomorphs.*

There was a significant main effect of Time, $F(2, 54) = 151.71$, $p < .001$, $\eta^2 = .85$. Post hoc pairwise comparisons with Bonferroni adjusted α values indicated significant

differences between the total number of items produced correctly at all testing points, *p*s < .005, with a mean improvement of 40.1% between pre-intervention and post-intervention, and a mean decline of 6.1% between post-intervention and the maintenance testing points. There was also a significant main effect of Allomorphic Category, $F(2, 27) = 8.38$, $p = .001$, $\eta^2 = .38$. Post hoc pairwise comparisons with Bonferroni adjusted α values revealed a significant difference of 18.7% between the total number of [t] allomorphs and [əd] produced correctly, $p = .001$, $d = 0.88$. All other pairwise comparisons were non-significant. There was no significant Time x Allomorphic Category interaction, $F(4, 54) = .91$, $p = .465$, $\eta^2 = .06$. These results suggest that all allomorphic categories improved to a similar extent as a result of intervention, with a shallow decline during the maintenance phase.

DISCUSSION

The primary aim of this study was to evaluate whether children aged between 5;9 – 7;1 years with DLD in two different dose frequency groups, resulting in different cumulative intervention intensities, differ in their response to an explicit intervention to improve past tense marking in children. Data from previously reported intervention studies were analysed. We evaluated rate of progress on criterion-referenced measures of past tense production (i.e., the GET-ed) probed during baseline, intervention, and maintenance phases. Dose conditions included an explicit intervention provided twice weekly (2PW) (1000 trials over 400-600 minutes) and once weekly (1PW) (500 trials over 200-300 minutes) over an intervention duration of 10 weeks. On the assumption that more is better regarding dosage, we hypothesised intervention provided twice per week would result in a higher rate of progress in past tense production during the intervention phase.

Additionally, we conducted exploratory analysis of past tense production and allomorphic categories of regular past tense verbs (i.e., [t], [d], and [əd]) by grouping data pre- and post-intervention from all participants who completed intervention in previously reported studies (Calder et al., 2020, 2021a) to determine if the type of allomorph contributes to the intervention effect, i.e., are some allomorphs easier/harder to learn than others? Based on previous research (e.g., Calder et al., 2021b; Tomas et al., 2015, 2017) and the assumption that children have difficulty detecting the rule-governed computations of morphological sequencing (Plante & Gomez, 2018; Ullman & Pierpoint, 2005), we expected verbs marked with syllabic allomorphs (i.e., marked with [əd] inflection) to be more difficult to learn for children with DLD.

Intervention Effect

The primary outcome in this study was production of untrained verbs. Previous studies in this programme of research indicated a statistically significant difference in mean scores of past tense production of untrained verbs following intervention compared with a baseline period prior to intervention (Calder et al., 2020) and compared to a waiting control group (Calder et al., 2021a). The current analyses indicated both groups showed significant progress with intervention and a shallow decline in performance during the maintenance phase. In line with our first hypothesis, the rate of progress during the intervention phase appeared to show a marginal statistical advantage in providing intervention for past tense production twice per week compared to once per week. The 2PW group also had higher scores during the maintenance phase. Visual inspection of the plotted data in Figure 1 suggests the 2PW group continued progress in the first two weeks of the maintenance phase before performance started to decline to a reasonably stable level of around 70% correct (which was similar to the immediate post-intervention score), whereas the 1PW group's performance showed an increase for one week before decreasing again to near the post-

intervention score of around 50% correct. Thus, it appears that both groups maintained progress made during intervention, but that due to the greater rate of progress during the intervention period, the 2PW group had higher scores.

Notably, individual response to the intervention varied regardless of dose frequency condition (see Figure 2). For example, P04 and P05 in the 2PW condition and P13 in the 1PW condition showed a relatively stable baselines, positive progress in the intervention phase, and stable performance in the maintenance phase. P16 in the 1PW condition showed a similar profile with the exception of a steep decline in the maintenance phase, whereas P15 and P25 continued to improve in the maintenance phase. In contrast, P21, P22 and P29 showed little progress during the intervention phase, but demonstrated higher performance in the maintenance phase compared to the baseline phase, and P24 and P28 demonstrated little improvement throughout any phase. Similar to Smith-Lock et al. (2013b), the existence of individual differences regardless of dose frequency conditions demonstrates the heterogeneity of this clinical population. Therefore, intervention procedures will likely need to be tailored to the individual needs of children within clinical contexts.

It is also worth noting that six participants (P02, P03, P13, P14, P21, and P29) scored within the average range on the TROG-2 as a receptive grammar measure, which may suggest relative strengths in comprehension compared to expressive grammar. Of these six participants, four made steady progress throughout the intervention phase, with two who had 2PW (P02, P03) performing near ceiling during the maintenance phase, and two who had 1PW (P13, P14) declining to around 50% – 60% accuracy at the end of the maintenance phase. P21 and P29 (who had 1PW) demonstrated limited progress during intervention, but increased to higher accuracy during the maintenance phase. It is however or note that of the children who performed below average on the TROG-2, P04, P05 and P11 performed at ceiling during maintenance, with P25 improving to around ceiling by the end of the

maintenance phase. With the exception of P11 (who performed within the average range on the SPELT-3), these participants who scored above 80 on the TROG-2 had higher receptive grammar scores than expressive grammar scores. Therefore, at an individual level, it may appear that overall relative strength in grammatical comprehension provides an advantage to learning through this explicit intervention. This strength may have established the foundation for the metalinguistic aspect of learning to draw upon explicit knowledge of the grammatical rule, which in turn facilitated correct production. However, correlation analyses did not support this observation statistically, so we cannot conclude there is a possible predictive relationship between strength in receptive grammar and pre- to post-intervention progress.

Analysis of Past Tense Allomorphs

A secondary aim of this study was to evaluate whether there is any difference between the three possible past tense allomorphs in response to intervention. Verbs marked with the syllabic [əd] allomorph are suggested to be more difficult for children with DLD than verbs marked with [t] or [d] (Calder et al., 2021b; Tomas et al., 2015, 2017). Consistent with the existing literature, there was a main effect of Allomorphic Category, and pairwise comparisons indicated that fewer children produced past tense of verbs marked with [əd] correctly compared to verbs marked with [t]. However, the significant main effect of Time with no interaction of Allomorph and Time indicates an effect of our explicit intervention on past tense production that, contrary to our hypothesis, was equivalent for all three allomorphs. Therefore, clinicians should not favour one allomorph over another when selecting intervention targets if using this explicit intervention approach. Further, this may provide evidence that this explicit intervention circumvents a potential implicit learning deficit, and that children with DLD can learn through relatively spared declarative learning capacity. This is consistent with the Procedural Deficit Hypothesis account of DLD (Ullman & Pierpoint, 2005). That is, while children may not be able to detect the probabilistic regularities of

applying past tense inflection (particularly the syllabic [əd] allomorph) due to an implicit learning deficit, syllabicity in past tense production had no effect on learning through explicit intervention. Findings from the current study also demonstrate this skill is generalisable to learning untrained verbs.

Dose Frequency, Intervention Duration, Intensity, and Intervention Effectiveness

Prior to implementing effectiveness research with large-scale randomised controlled trials, it is useful to conduct efficacy studies evaluating parameters of intervention intensity. Recent research into the intervention for morphosyntax suggests varying dose in terms of daily frequency (Meyers-Denman, 2016) or dose density within sessions (Plante et al., 2019) does not appear to influence intervention efficacy when other dosage and intensity parameters are held constant (i.e., delivered once a day for five days for five weeks). Although in both studies participants were probed on performance throughout intervention, Meyers-Denman (2016) analysed between group effects based on performance on post-intervention correct use of morphemes only, and Plante et al. (2019) compared effect sizes and post-intervention correct spontaneous production of morphemes. Thus, the richness of the repeated measures data was not utilised.

Smith-Lock et al. (2013b) evaluated the effects of dose frequency (daily versus weekly) while maintaining cumulative intervention intensity (in minutes) for an expressive grammar intervention with five-year-old children with DLD. The weekly dose frequency condition was more effective than the daily condition; however, it is important to note that dose was not specified as part of the dosage and intensity parameters. They analysed between group effects on pre-post intervention progress using analysis of variance, and single case analyses were reported. Single case analyses indicated that responsiveness to intervention within each group was dependent on individual profiles.

Unlike the aforementioned studies, we used repeated measures throughout the baseline, intervention, and maintenance phases to analyse potential between group differences using logistic regression. The repeated measures provided much richer data than just pre- and post-intervention measures and logistic regression could include random participant factors (which is important given the wide individual variation). Indeed, inspection of Figure 1 suggests if we had used only the assessments delivered pre- and post-intervention, this may have led to misleading results, as the pre-intervention score for the 2PW group appears to be an outlier.

Another key aspect to consider is the ways in which dosage and intensity parameters differ across studies, despite operational definitions provided by Warren et al. (2007). Table 3 reports how these parameters have been defined by authors in the aforementioned studies. Differences across studies tend to be by how dose frequency and total intervention duration are defined. For example, Plante et al. (2019) define duration in days, Meyers-Denman and Plante (2016) define duration in weeks, and Smith-Lock et al. (2013b) define duration in days and weeks. As a variable in the equation for which cumulative intervention intensity is a product, this has implications for how intervention procedures have been defined. In fact, the differences between intervention parameters in Plante et al. (2019) cannot be captured with the terminology presented by Warren et al. (2007). It seems the purpose of the Plante et al. (2019) study was therefore to highlight *dose density* by examining how the manipulation of doses within a set number of minutes within individual intervention sessions may influence outcomes. Specifically, comparing dose spacing between sparse (24 dose over 30 minutes) and dense (24 doses over 15 minutes) conditions. This may be a more clinically salient way to conceptualise dosage, as it provides specific information to clinicians in the context of one-to-one contact with the child.

Table 3. *Summary of intervention studies evaluating elements of dosage and intensity to improve morphosyntax in young children with DLD as reported by the authors.*

Study	Sample size	Dose Form	Dose	Dose Frequency	Total Intervention Duration	Cumulative Intervention Intensity	Analysis
Plante et al. (2019)	*n* = 10 per group (total *n* = 20)	Enhanced conversational recasting	Single therapeutic event (p. 1234)	24 doses per session (p. 1237) Session frequency 1x day, 5x days per week (p. 1237)	Up to 25 days (p. 1237)	Mean 566.4 recasts (range = 528 – 600)	Two-tailed *t* test comparing effect sizes and post-intervention spontaneous correct morpheme use: No group differences
Meyers-Denman & Plante (2016)	*n* = 8 per group (total *n* = 16)	Enhanced conversational recasting (p. 340)	24 recasts (p. 340)	3x sessions *versus* 1x session per day (p. 340)	~5 weeks (p. 340)	600 recasts (range = 504 – 624) over an average of 750 min (range = 630 – 780) over 25 days (range = 21 – 26)	Two-tailed *t* test comparing post-intervention correct morpheme use: No group differences
Smith-Lock et al. (2013b)	*n* = 19 daily, *n* = 15 weekly (total *n* = 34)	Expressive grammar programme	Not reported	1x daily *versus* 1x weekly (p. 261)	8 days *versus* 8 weeks (p. 261)	480 minutes (p. 261)	Analysis of variance in pre-post intervention gain between groups: Weekly intervention more effective than daily.
Calder et al. (current)	*n* = 9 2PW, *n* = 20 1PW (total *n* = 29)	Explicit intervention	50 trials	1x 20-30 minute session, 2x per week *versus* 1 x 20-30 minute session, 1x per week	10 weeks	1000 trials over 400-600 minutes over 10 weeks *versus* 500 trials 200-300 minutes over 10 weeks	Logistic regression: Twice weekly more effective than once weekly

Alternative definitions for cumulative intervention intensity have been reported. For example, Smith-Lock et al. (2013b) referred to Gillam's (2012) definition, *session duration* x *dose frequency* x *total intervention duration*, which reportedly omits *dose* from intensity in favour of *session duration*. This is seemingly due to the perceived difficulty in accurately controlling and reporting dose in research and clinical contexts (Smith-Lock et al., 2013b). Schmitt et al. (2017) report cumulative intervention intensity as "the product of dose (i.e., average time spent targeting language skills per each session), frequency (i.e., number of sessions each child received), and total frequency duration (over one academic year)" (p. 159). Given the potential difficulty in controlling for the precise number of doses as 'teaching episodes' (Warren et al., 2007), perhaps defining dose in terms the time spent targeting a specific skill is more clinically salient. However, therein lies the issue of operationally defining 'targeting a specific skill' within an intervention session, or alternatively, a single therapeutic event (Plante et al., 2019).

It appears, overall, more consensus is needed to re-define parameters that contribute to cumulative intervention intensity in ways that inform the design of interventions for empirical evaluation, and which are ultimately clinically applicable. Nonetheless, findings from the previous evidence-base indicate that there is no difference in morphosyntax outcomes when intervention is spaced or distributed sessions while total intervention duration and cumulative intervention intensity is maintained (Meyers-Denman, 2016; Plante et al., 2019). This suggests clinicians can allow for flexibility in their service delivery of intervention. However, although these studies were designed to evaluate efficacy of intervention, providing intervention daily does not align with current speech-language pathologist practices (see Finestack & Satterlund, 2018). Further, there is evidence that spacing dose frequency (in sessions) over weeks rather than days may be more effective than

blocking sessions within weeks (Smith-Lock et al., 2013b). Though, it is unclear whether the dose had any influence on outcomes.

Findings from the current study suggest there may be an advantage to providing twice the cumulative intervention intensity (in doses and minutes) over 10 weeks to improve past tense marking through explicit intervention procedures. Ullman and Pierpoint (2005) have suggested that there is benefit to teaching grammar to children with DLD through explicit intervention in the presence of an implicit learning deficit. Since the declarative memory system is responsible for learning explicit information and functions to store episodic and semantic information (Ullman, 2016), it was hypothesised that repeated exposures to stimuli (i.e., increased trials) would strengthen memories and thus be efficiently retrieved for use in expressive language. Further research is required to determine whether the total number of doses spread over a longer intervention duration (i.e., twice per week for 10 weeks versus once per week for 20 weeks) results in similar intervention effects. If so, this may suggest cumulative intervention intensity is more important to consider than spacing when delivering this intervention. Indeed, Frizelle et al. (2021) concluded that "more is not always better" (p.752), since the evidence-base indicates a point of diminishing returns in terms of dose frequency and the total number of intervention sessions.

Limitations and Future Directions

We cannot draw conclusive inferences for clinical practice as a result of this study due to the following limitations. Firstly, evaluation of rate of progress on past tense production was a post hoc re-analysis of data from previously reported intervention studies conducted at different times. Therefore, drawing causal inferences from findings should be done with caution. Nonetheless, given the relative paucity of evidence related to dose and grammar interventions, findings from this study may provide the foundation for future randomised

controlled trials with larger sample sizes. In addition, evaluating the effects of differing intensity of intervention on response to intervention on performance after intervention has ceased is urgently needed, as this could have important implications for clinical practice.

Secondly, not all participants received the intervention at the same time, so the overall provision of the intervention may have improved by the time the 1PW group received the intervention. Thirdly, the participants may not be entirely representative of the DLD population at large as they were recruited through convenience sampling from specialised educational programmes (Redmond et al., 2019). Some participants also performed within the average range on a measure of receptive grammar, however, correlations between pre- to post-intervention progress on the GET-ed and TROG-2 showed no meaningful relationship. A future study with a larger sample size may be adequately powered to test pre-intervention assessment scores as predictors for progress to identify a profile for responsiveness to intervention. Finally, the primary outcome of this efficacy research was a relatively static measure of morphosyntax production, which may not reflect use in functional communication. Naturalistic measures of expressive grammar, such as those attained through language sample analysis, are required to report on the effectiveness of explicit grammar interventions confidently.

Conclusions

This study retrospectively, yet systematically, evaluated different dose frequencies, resulting in different intervention intensities, of an explicit grammar intervention to improve past tense production in young school-aged children with DLD. Factors relating to cumulative intervention intensity, such as dose frequency, affected the rate of change in past tense production throughout the intervention phase. However, due to the design limitations of this study, further research using large-scale randomised controlled trials is required before

drawing strong conclusions for clinical practice. Notably, the allomorphic category of regular past tense verbs does not appear to influence response to this intervention, with all three possible allomorphs improving equivalently. Our results continue to provide evidence of the efficacy of explicit grammar interventions for children with DLD, provided both once and twice per week, and thus should be considered a viable treatment option for this clinical population.

REFERENCES

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

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

Bishop, D. V. M. (2014). Problems with tense marking in children with specific language impairment: not how but when. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1634). 1-8. hiips://doi.org/10.1098/rstb.2012.0401

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

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & consortium, C. (2016). CATALISE: A Multinational and Multidisciplinary Delphi Consensus Study. Identifying Language Impairments in Children. *PLoS ONE [Electronic Resource], 11*(7), e0158753. hiips://doi.org/10.1371/journal.pone.0158753

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

Calder, S. D., Claessen, M., Leitão, S., & Ebbels, S. (2021b). A profile of expressive inflectional morphology in early school-age children with developmental language

disorder. *Clinical Linguistics & Phonetics*, 1-18. hiips://doi.org/10.1080/02699206.2021.1931454

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 and Hearing Services in Schools, 51*(2), 298-316. hiips://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. hiips://doi.org/10.1177/0265659017735392

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

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. hiips://doi.org/10.1191/0265659007072143

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. hiips://doi.org/10.1044/2018_ajslp-17-0168

Frizelle, P., Tolonen, A. K., Tulip, J., Murphy, C. A., Saldana, D., & McKean, C. (2021). The influence of quantitative intervention dosage on oral language outcomes for children with developmental language disorder: A systematic review and narrative synthesis. *Language, Speech, and Hearing Services in Schools*, *52*(2), 738-754. hiips://doi.org/10.1044/2020_LSHSS-20-00058

Gillam , R . (2012) . *How much is enough? Dosage in child language intervention* . Paper presented at the American Speech-Language-Hearing Association Annual Convention, Atlanta, GA.

Howell, D. C. (2010). *Statistical methods for psychology* (7th ed.). Belmont CA: Wadsworth Cengage Learning

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2012). Single-Case Intervention Research Design Standards. *Remedial and Special Education, 34*(1), 26-38. hiip://dx.doi.org/10.1177/0741932512452794

Leonard, L. B. (1989). Language learnability and specific language impairment in children. *Applied Psycholinguistics, 10*(2), 179-202. hiip://dx.doi.org/10.1017/S0142716400008511

Leonard, L. B., Eyer, J. A., Bedore, L. M., & Grela, B. G. (1997). Three accounts of the grammatical morpheme difficulties of English-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 40*(4), 741-753. hiip://dx.doi.org/10.1044/jslhr.4004.741

Leonard, L. B., & Deevy, P. (2017). The changing view of input in the treatment of children with grammatical deficits. *American Journal of Speech-Language Pathology, 26*(3), 1030-1041. hiips://doi.or/10.1044/2017_AJSLP-16-0095

Marshall, C. R., & van der Lely, H. K. J. (2006). A challenge to current models of past tense inflection: The impact of phonotactics. *Cognition, 100*(2), 302-320. hiips://doi.org/10.1016/j.cognition.2005.06.001

Meyers-Denman, C. N., & Plante, E. (2016). Dose schedule and enhanced conversational recast treatment for children with specific language impairment. *Language, Speech,*

and Hearing Services in Schools, 47(4), 334-346. hiips://doi.org/10.1044/2016_lshss-15-0064

Mirman, D. (2014). *Growth curve analysis and visualization using R*. CRC press.

Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., . . . Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: evidence from a population study. *Journal of Child Psychology and Psychiatry, 57*(11), 1247-1257. hiips://doi.org/10.1111/jcpp.12573

Oetting, J. B., & Horohov, J. E. (1997). Past-tense marking by children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research, 40*(1), 62-74. hiips://doi.org/10.1044/jslhr.4001.62

Owen Van Horne, A. J., Curran, M., Larson, C., & Fey, M. E. (2018). Effects of a complexity-based approach on generalization of past tense -ed and related morphemes. *Language, Speech, and Hearing Services in Schools, 49*(3s), 681-693. hiips://doi.org/10.1044/2018_lshss-stlt1-17-0142

Owen Van Horne, A. J., & Green Fager, M. (2015). Quantifying the relative contributions of lexical and phonological factors to regular past tense accuracy. *International Journal of Speechlanguage Pathology, 17*(6), 605-616. hiips://doi.org/10.3109/17549507.2015.1034174

Perona, K., Plante, E., & Vance, R. (2005). Diagnostic accuracy of the structured photographic expressive language test: third edition (SPELT-3). *Language, Speech & Hearing Services in the Schools, 36*(2), 103-115. hiips://doi.org/10.1044/0161-1461(2005/010)

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

Plante, E., Mettler, H. M., Tucci, A., & Vance, R. (2019). Maximizing treatment efficiency in developmental language disorder: positive effects in half the time. *American Journal of Speech-Language Pathology*, *28*(3), 1233-1247. hiips://doi.org/10.1044/2019_AJSLP-18-0285

Plante, E., Ogilvie, T., Vance, R., Aguilar, J. M., Dailey, N. S., Meyers, C., Lieser, A. M., & Burton, R. (2014). Variability in the language input to children enhances learning in a treatment context. *American Journal of Speech Language Pathology, 23*(4), 530-545. hiips://doi.org/10.1044/2014_ajslp-13-0038

Redmond, S. M., Ash, A. C., Christopulos, T. T., & Pfaff, T. (2019). Diagnostic Accuracy of Sentence Recall and Past Tense Measures for Identifying Children's Language Impairments. *Journal of Speech Language & Hearing Research, 62*(7), 2438-2454. hiips://doi.org/10.1044/2019_jslhr-l-18-0388

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

Rice, M. L., Wexler, K., & Hershberger, S. (1998). Tense over time: The longitudinal course of tense acquisition in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *41*(6), 1412-1431. hiips://doi.org/10.1044/jslhr.4106.1412

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

Smith-Lock, K., Leitao, S., Lambert, L., Prior, P., Dunn, A., Cronje, J., . . . Nickels, L. (2013b). Daily or weekly? The role of treatment frequency in the effectiveness of grammar treatment for children with specific language impairment. *International*

Journal of Speechlanguage Pathology, 15(3), 255-267. hiips://doi.org/10.3109/17549507.2013.777851

Smith-Lock, K. M., Leitão, S., Prior, P., & Nickels, L. (2015). The Effectiveness of Two Grammar Treatment Procedures for Children With SLI: A Randomized Clinical Trial. *Language, Speech, and Hearing Services in Schools, 46*(4), 312-324. hiips://doi.org/10.1044/2015_LSHSS-14-0041

Schmitt, M. B., Justice, L. M., & Logan, J. A. (2017). Intensity of language treatment: Contribution to children's language outcomes. *International Journal of Language & Communication Disorders, 52*(2), https://doi.org/155-167. 10.1111/1460-6984.12254

Tomas, E., Demuth, K., & Petocz, P. (2017). The Role of Frequency in Learning Morphophonological Alternations: Implications for Children With Specific Language Impairment. *Journal of Speech Language, and Hearing Research, 60*(5), 1316-1329. hiips://doi.org/10.1044/2016_jslhr-l-16-0138

Tomas, E., Demuth, K., Smith-Lock, K. M., & Petocz, P. (2015). Phonological and morphophonological effects on grammatical development in children with specific language impairment. *International Journal of Language and Communication Disorders, 50*(4), 516-528. hiips://doi.org/10.1111/1460-6984.12152

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. hiips://doi.org/10.1016/S0010-9452(08)70276-4

van der Lely, H. K.J., & Ullman, M. T. (2001). Past tense morphology in specifically language impaired and normally developing children. *Language and Cognitive Processes, 16*(2-3), 177‑217. hiips://doi.org/10.1080/01690960042000076

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 and 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, and Hearing Research, 43*(6), 1322-1336. hiips://doi.org/10.1044/jslhr.4306.1322