2023-7

# Pattern-based target selection for treatment of irregular past tense: A single-case experimental design for children with DLD

Kirsten M. Hannig Russell
*Brigham Young University - Provo*, kirsten.russell@byu.edu

Amy Wilder
*University of Utah*

Julie L. Wambaugh
*University of Utah*



## Original Publication Citation

Hannig Russell, K. M., Wilder, A., & Wambaugh, J. L. (2023). Pattern-based target selection for treatment of irregular past-tense: A single-subject experimental design for children with DLD. American Journal of Speech-Language Pathology, 32, 2057-2074. https://doi.org/10.1044/2023_AJSLP-22-00038

# Pattern-based target selection for treatment of irregular past tense: A single-case experimental design for children with DLD

Kirsten M. Hannig Russell, Amy Wilder, Julie L. Wambaugh

University of Utah

## Author Note

Kirsten Hannig Russell and Amy Wilder have no conflicts of interest to disclose. Julie Wambaugh receives a salary for conducting research.

Correspondence concerning this article should be addressed to Kirsten M. Hannig Russell, University of Utah, 390 South 1530 East, Suite 1201, Salt Lake City, UT 84112. Email: kirsten.h.russell@utah.edu

**Abstract**

Purpose: Many children with developmental language disorder (DLD) demonstrate difficulty inflecting irregular past tense verbs. We evaluated using phonological-based patterns (i.e., schemas) to select targets for treatment of irregular verb inflection for children with DLD. We hypothesized this approach would result in acquisition of treated verbs, generalization to untreated verbs within the same phonological schema, and increased use of correct irregular verbs in naturalistic contexts.

Method: Treatment was provided to three 7-year-old participants with DLD in the context of multiple baseline design across behaviors and participants. Phonological schemas included vowel change, final alveolar, and dual change irregular verbs. Treatment was provided using established therapeutic approaches, including narrative generation, sentence imitation, and naturalistic play activities. Acquisition of treated and untreated targets was assessed at the beginning of sessions, and use of verbs in naturalistic contexts was observed throughout treatment sessions.

Results: Positive acquisition effects were noted for two of three participants. Generalization to untreated items occurred within and across treatment sets for two participants, whereas increases in accurate irregular verb production in naturalistic contexts were not observed.

Conclusions: Outcomes demonstrated support for a pattern-based approach to target selection for treatment of irregular past tense verbs. One participant, who demonstrated general difficulty with the probe and treatment tasks, showed no treatment effect, suggesting the treatment may be effective for some children with DLD. Further research is warranted to determine what factors might have influenced these variable outcomes across participants.

**Keywords:** *target selection, developmental language disorder, children, tense marking, intervention, efficacy.*

# Pattern-based target selection for treatment of irregular past tense: A single-case experimental design for children with DLD

Difficulty with the acquisition of past tense verb forms is a hallmark characteristic of English-speaking children who have developmental language disorder (DLD; Bishop, 2013; Leonard et al., 1992; Oetting & Horohov, 1997; Rice et al., 1995; Rice et al., 1998; Rice et al., 2000). Children with DLD may demonstrate difficulty with grammatical aspects of language, including past tense verb inflection, despite the absence of hearing problems, primary biomedical conditions (e.g., autism, Down syndrome), and neurological deficits. Research examining the acquisition of past tense verb forms has typically focused on a group of children who have specific language impairment (SLI), a stricter designation found within the broader designation of DLD. In addition to the requirements for DLD, children with SLI also demonstrate typical nonverbal IQs. Past tense verbs include regular (i.e., *-ed*) and irregular forms (e.g., *ran*, *slept*). Marking irregular past tense begins earlier in development than regular past tense for most children, but mastery of irregular past tense takes longer than mastery of regular forms (Kuczaj II, 1977; Shipley et al., 1991). For regular past tense verbs, selection of treatment targets includes a pattern- or complexity-based approach (Owen van Horne & Fager, 2015; Owen van Horne et al., 2017, 2018). No current recommendations exist for selecting targets for irregular verb treatment. The current study aims to address this gap by evaluating the efficacy of using a pattern-based approach to select targets for treatment of irregular past tense verb forms.

## Treatment for Regular Past Tense

While there are many studies examining possible factors that promote accurate production of past tense, there are only a few studies that have tested these factors during treatment of children with SLI, a more specific form of DLD that has been studied extensively. The primary difference between DLD and SLI is that children with DLD may demonstrate a range of low-to-typical nonverbal abilities (in the absence of intellectual disability) while children with SLI must demonstrate nonverbal abilities in the

typical range (Bishop et al., 2017). As most of the literature examining past tense marking has focused on children with SLI, we summarize this literature and apply it to the current study.

Most grammatical treatment research has focused on therapeutic methods (Law et al., 2005; Rinaldi et al., 2021), with the consensus that both explicit and implicit methods during treatment result in positive acquisition and generalization of past tense forms, with a possible advantage for explicit methods (Calder et al., 2021; Ebbels, 2014; Kulkarni et al., 2014; Rinaldi et al., 2021). There is, however, little guidance in the literature for selecting optimal treatment targets when treating past tense. Based on studies that examined factors associated with accurate production of regular past tense in typically developing children and those with SLI, one might conclude that selecting targets with high lexical frequency or phonological similarity to known words (i.e., neighborhood structure; Marchman, 1997; Marchman et al., 1999; Oetting & Horohov, 1997; Owen Van Horne & Fager, 2015; Tomas et al., 2017), low phonological complexity (e.g., *cried*, instead of *cuddled*; Marshall & van der Lely, 2007; Tomas et al., 2015), and high telicity (i.e., a clear endpoint; *dropped*) would be optimal (Johnson & Fey, 2006; Leonard, Deevy et al., 2007; Owen & Fager, 2015). Current recommendations support this (Crystal, 1985; Weiler, 2013). Yet, in the few treatment studies that examined these factors, the outcomes did not necessarily support these recommendations.

For example, Smith-Lock (2015) provided treatment for regular past tense marking to five children with SLI and examined the patterns of generalization to untreated regular past tense verbs. She found that children generalized to other regular past tense verbs irrespective of the phonological similarities and lexical frequency of the words. In contrast, a single-case study (Eyer & Leonard, 1994) involving one five-year-old participant demonstrated generalization to untreated past tense verbs, which seemed to be influenced by phonological neighbor features (e.g., showed/goed; cried/flied/dried). One caution in interpreting these outcomes is that both of these studies had small samples and only used pre-treatment and post-treatment measures to examine the generalization

patterns. Yet together, the studies suggest the possibility that the effect of phonological similarities of regular past tense marking may exist on a smaller scale than system-wide effects on a child's language system.

The effects of the recommended factors of phonological complexity, telicity, and frequency of inflection for regular past tense were directly tested by Owen Van Horne et al. (2017, 2018). They designed a treatment that selected treatment targets for regular past tense that were rated as "easy" for less complex verbs (low phonological complexity, high telicity, high frequency), or "hard" for more complex verbs (high phonological complexity, low telicity, low frequency). One group of children received treatment beginning with the "easy" verbs and progressing to the "hard" verbs, while a second group began treatment with the "hard" verbs and then progressed to the "easy" verbs. Results indicated that participants who began treatment targeting the "hard" verbs demonstrated higher accuracy on post-treatment measures and greater generalization to untreated verbs compared to the participants who began treatment with the "easy" verbs. These outcomes challenge the current recommendations for selecting targets for regular past tense treatment (Crystal, 1985; Weiler, 2013).

Currently, no recommendations exist for selecting irregular past tense treatment targets. Nor, to our knowledge, have any treatment studies sought to evaluate the influence of the factors of phonological similarities, complexity, frequency, or telicity on irregular past tense verbs. However, based upon the few treatment studies evaluating regular past tense verb targets, the question remains open as to whether phonological similarity influences the acquisition of verbs during treatment. Thus, this study may shed light on the issue for irregular past tense verb targets.

**Historical Basis for a Pattern-Based Approach to Target Selection**

Current irregular past tense verbs were originally rule-based constructions in earlier English language versions (Pinker, 1999). Although relatively few in number (~160 to 200), irregular verbs are among the most frequently used verbs in the language (Bybee & Slobin, 1982; Pinker, 1999; Shipley et

al., 1991). Typically developing children acquire some irregular past tense verbs as early as 3 years of age (e.g., *went*) and have acquired most irregular verbs by 8 to 9 years of age (Shipley et al., 1991). Most irregular verbs can be consolidated into one of three distinct phonological patterns. Internal vowel change, the most common pattern, maintains the consonant sounds while the vowel changes (e.g., *fall-fell*). The second common pattern is an internal vowel change with a final alveolar consonant (e.g., *ride-rode*). Children are more likely to omit past tense inflections on verbs that end with alveolar consonants, which provides a false signal that the verb has already been inflected for regular past tense (i.e., *-ed*; Marchman et al., 1999). The third pattern consists of changes to the internal vowel and the final phoneme(s). In addition to the vowel change, (a) the final phoneme or consonant cluster changes to an alveolar (e.g., *think-thought*), (b) an alveolar phoneme is added to the end of the word (e.g., *sleep-slept*), or (c) the final phoneme has a voicing change and an alveolar phoneme added (e.g., *leave-left*). Phonological patterns found among irregular past tense verbs suggest the possibility that selecting treatment targets according to a pattern-based approach may be efficacious.

## Theoretical Explanations of Learning Past Tense Forms

### *The Dual-Mechanism Theory*

The Dual-Mechanism Theory (Ullman, 2001) suggests irregular past tense verbs are held in declarative (i.e., episodic) memory as a "list" whereas regular past tense *-ed* verbs are formed via procedural (i.e., pattern-based) memory processes. This theory posits that when a verb must be inflected for past tense, a list of learned irregular forms held in declarative memory is reviewed and, if the verb form is not found, the verb is inflected according to the procedural rule of adding the *-ed* form. Treatment for regular past tense aligns with this theory by taking a rule-based approach (e.g., Owen van Horne et al., 2017, 2018; Smith-Lock et al., 2015). Treating irregular past tense would require memorizing all irregular verbs, as learning one verb would not result in generalization to other verbs.

### *Phonological Schemas Theory*

In contrast, the Phonological Schemas Theory (Bybee & Slobin, 1982) suggests that individuals identify phonological patterns (i.e., schemas) across irregular past tense verbs based on their inflected form (e.g., *drove*). These forms are then categorized according to their respective phonological patterns and used to inflect unfamiliar irregular past tense verbs.

The Dual-Mechanism Theory and the Phonological Schemas Theory suggest that, initially, irregular past tense verbs are acquired through rote learning. However, the Dual-Mechanism Theory suggests that rote learning persists as children mature, whereas the Phonological Schemas Theory suggests that pattern identification allows children to acquire irregular past tense verbs indirectly. Following the Phonological Schemas Theory, selecting treatment targets based on their phonological pattern would facilitate generalization to untreated verbs within each phonological schema.

### The Current Study

Based on the Phonological Schemas Theory, it is proposed that target selection utilizing a pattern-based approach may enhance treatment efficiency for irregular past tense verb inflection. This study evaluated whether (1) selecting irregular past tense verb targets using a pattern-based approach would result in acquisition of those forms (i.e., treatment effects) by three children with DLD, (2) generalization to untreated irregular past tense verbs would occur, and (3) treatment would lead to increased use of accurately inflected irregular past tense verbs in discourse contexts. Specific phonological patterns utilized were internal vowel change (hereafter *vowel change*), internal vowel change with a final alveolar consonant (hereafter *final alveolar*), and internal vowel change plus a change to an alveolar phoneme (hereafter *dual change*).

Pilot data from one child revealed positive acquisition and generalization of irregular verbs. Based on these data, it was hypothesized that (1) the use of a pattern-based approach when selecting irregular past tense verb targets would result in acquisition of the forms during treatment, (2)

generalization to untreated verb forms within patterns would occur, and (3) an increased proportion of accurate irregular past tense verbs would be used in discourse contexts.

## Methods

### Participants

The institutional review board at the University of Utah approved all aspects of this study. Three participants who met our criteria for DLD were recruited (Table 1). To be identified as having DLD for the study, participants were required to score within normal limits (i.e., standard score between 85 and 115) on the Naglieri Nonverbal Ability Test (NNAT; Naglieri, 2003), pass a binaural audiometric hearing screening and, per parental report, have no history of neurological damage or primary biomedical condition (Bishop et al., 2017) causal to their language impairment; although they could have co-occurring conditions (e.g., apraxia of speech, ADHD). Further, participants were identified as having DLD by scoring below one standard deviation (<85; M=100, SD=15) or below criterion on at least two of the following clinical indices: (a) Clinical Evaluation of Language Fundamentals-4$^{th}$ Edition Core Language Score (CELF-4; Semel et al., 2003); (b) Test of Early Grammatical Impairment (TEGI) Elicited Grammar Composite; (c) Redmond Sentence Recall task (RSR; Redmond, 2005); (d) Dollaghan and Campbell's (1998) Nonword Repetition task (NWR).

**[Table 1 near here]**

#### *Language Measures*

The CELF-4 (Semel et al., 2003) is a measure of general language ability. We utilized the Core Language Score and used a cutoff score of 85 due to its 100% sensitivity and 82% specificity for identifying DLD (Nitido & Plante, 2020). The CELF-4 is commonly used by researchers and clinicians (Betz et al., 2013; Caesar & Kohler, 2009; Finestack & Satterlund, 2018).

The TEGI (Rice & Wexler, 2001), a robust clinical marker for identifying DLD (Nitido & Plante, 2020), assesses children's accuracy with marking tense. The Elicited Grammar Composite (EGC) score

reflects children's ability to mark tense on a variety of verbs. All participants were required to score below the criterion score for their age (93%); associated with acceptable sensitivity and specificity levels (Rice & Wexler, 2001).

Nonword repetition tasks are another robust clinical marker for identifying DLD (Graf Estes et al., 2007). Dollaghan and Campbell's (1998) NWR requires examinees to repeat 16 nonwords aligning with the phonotactic rules of the English language. The production of correct phonemes was scored as a percentage and converted to a standard score based on community norms reported by Redmond et al. (2019). Standard scores <85 were associated with the optimal cut-off (raw score of 85%) for the identification of language impairment (Redmond et al., 2011).

Redmond's sentence recall task (2005), another robust clinical marker for identifying DLD (Archibald & Joanisse, 2009), consists of eight active and eight passive sentences presented verbally to children. The children's repetitions were scored according to the number of errors produced (Archibald & Joanisse, 2009). Standard scores <85 were associated with adequate sensitivity and specificity for language impairment (Redmond et al., 2019).

***Descriptive Measure***

As portions of our intervention were narrative-based, it seemed prudent to note participants' narrative abilities. We used the Test of Narrative Language-Second Edition (TNL-2; Gillam & Pearson, 2017), a norm-referenced assessment that measures examinees' comprehension and production of narratives through (a) responses to scripted questions following narratives presented orally and (b) the number of critical elements included in narrative productions.

***Additional Eligibility Requirements***

All participants were required to be monolingual English speakers between the ages of 5 and 9 years of age. Additionally, a portion of the TEGI was used to identify potential participants who demonstrated sufficiently low performance for marking irregular past tense. To participate in the study,

all participants were required to demonstrate <60% accuracy on the TEGI when marking irregular past tense verbs on the past tense probe. This level was chosen because it indicates approximately chance levels of marking irregular past tense correctly, which suggests the treatment may be appropriate for the participants. The irregular past tense score was calculated as the percent correct of all irregular past tense verbs produced on the TEGI past tense probe.

***Participant 1***

Participant 1 (P1) was a male aged 7;2 at entrance into the study. He received early intervention services for speech and language delays beginning at 2 years of age, which continued into preschool and elementary school. According to parental report, P1 was diagnosed with childhood apraxia of speech (CAS) and was receiving private treatment weekly for CAS. He was also receiving services at school targeting speech and language deficits.

***Participant 2***

Participant 2 (P2) was a male aged 7;10 at entrance into the study. He received early intervention services beginning at age 2 for speech and language delays. Like P1, he was currently receiving services at school targeting speech and language deficits. P2 received medication on school days (methylphenidate) to treat suspected ADHD symptoms but had not yet received a diagnosis of ADHD. Upon shutdown of the school system due to COVID-19 (see “Unexpected Events” section), his parents halted his medication, and P2 remained off medication for the remainder of the study.

***Participant 3***

Participant 3 (P3) was a male aged 7;10 upon his admittance to the study. P3 and P2 were twins and lived in the same household. P3 received early intervention services to treat speech and language delays beginning at age 2. Like P1 and P2, he was currently receiving services at school targeting speech and language deficits.

All participants were working with speech-language pathologists on various goals and continued to do so throughout the study. The parents of participants were unsure whether participants were receiving direct treatment at school for past tense marking, but they contacted their school-based clinicians and requested that they not provide treatment targeting past tense marking for the duration of this study.

## Experimental Design

### *General Description*

Each participant received treatment in the context of a multiple baseline design (MBD) across behaviors and participants. The MBD provides experimental control when systematic changes to the dependent variable (i.e., participants' accurate production of irregular past tense verbs) align with the onset of treatment (Kratochwill et al., 2010). Causality of treatment to behavior change (i.e., treatment effects) is indicated when systematic replication of the targeted behavior across treatment sets occurs at three different points in time for a single case (i.e., MBD across behaviors for a single participant) or across three different cases (i.e., MBD across participants; Kratochwill et al., 2010). Three participants were included to allow for the possibility of demonstrating treatment effects across participants and treatment sets. The MBD across participants was nonconcurrent for P1, but concurrent for P2 and P3.

Production of the dependent variable was measured repeatedly in probes during the baseline phase. The number of baseline probes was determined *a priori* and extended across participants (i.e., P1 = 5 baseline probes; P2 = 6 probes; P3 = 7 probes) to demonstrate treatment effects were not aligned with familiarity with the baseline probes. However, each participant needed to demonstrate stable performance for the predetermined number of baseline probes to begin treatment with any experimental set. Initiation of treatment for a given participant was not linked to the other participants' performance. For P2 and P3, extended baseline probing was required to establish stability of responses prior to initiating treatment with subsequent experimental sets, but not with the initial treatment set.

Treatment was applied sequentially to sets of experimental items following the baseline phase while probing continued with treated and untreated sets of stimuli. Three experimental treatment sets were incorporated into the MBD, with the potential to apply treatment with each set depending upon treatment response. Prior to the start of the study, criteria for determining cessation of treatment were established. Treatment was discontinued (a) if a participant demonstrated 80% accuracy or higher on treated items for 3 of 4 consecutive probes, (b) if a maximum of 20 treatment sessions was provided, or (c) if no changes in behavior were evident following five consecutive probes.

Treatment was applied to all three sets as planned *a priori* for Participant 2. Participants 1 and 3 received treatment for two of the three sets for different reasons. Participant 1 had limited treatment response with the first two sets; therefore, treatment was discontinued. Participant 3 demonstrated generalization effects to the third set of stimuli without treatment. Follow-up probes, which measured participants' maintenance of treatment effects across several weeks with no treatment, were completed 2-, 4-, and 8-weeks following completion of all treatment.

***Experimental Stimuli***

Sets of eight treated and eight untreated stimuli for each of the three irregular past tense patterns (i.e., vowel change, final alveolar, and dual change) were developed for each participant. Stimuli for each treatment set (i.e., pattern) were initially selected using 41 verbs provided by Shipley et al. (1991). These verbs represented a range of expected acquisition ages (e.g., 3-9+ years) and fit the parameters of the phonological patterns targeted (e.g., *drive-drove*; *find-found*; *catch-caught*). The verbs in the Shipley et al. (1991) report did not provide enough stimuli for treated and untreated verbs for each treatment set; therefore, a search was conducted online to identify additional lists of irregular verbs that fit the criteria for each treatment set. Those verbs that were concrete enough to address in treatment were added to complete the stimulus sets. Tentative sets of stimuli were assessed with participants prior to beginning the baseline phase. Stimuli that showed prior acquisition through >50%

accurate performance on repeated trial probes were replaced with new stimuli from the same pattern-based category, resulting in individualized sets of stimuli (Supplemental Material S1).

To elicit production of the experimental verbs, sentence completion items that provided a subject were devised for each verb (e.g., Prompt: *"Every day they sit. Yesterday, they?"*). Participants were required to provide the verb at the conclusion of the sentence (e.g., *"sat"*). Probes were constructed with randomized pronouns (i.e., he, she, or they), sentence types (i.e., simple present tense or present tense auxiliary with present progressive aspect), and order of presentation to ensure participants could not memorize a response sequence. Items within sets were randomly assigned as treated or untreated to reflect acquisition or generalization effects of treatment, respectively. In a few instances, baseline performance demonstrated the need to balance items that were partially acquired across treated/untreated subsets prior to treatment to ensure approximate equivalence of item difficulty across subsets.

***Baseline Phase***

Five baseline probes were determined the minimum based upon an *a priori* decision to use the Conservative Dual Criterion method in analysis (CDC; Fisher et al., 2003). The three experimental sets were presented in randomized order for each probe session. For each experimental set, the 16 sentence completion items (i.e., eight acquisition and eight generalization) were randomized along with four control (i.e., regular past tense *–ed*) stimuli for a total of 20 items per set (Supplemental Material S2). Control items were included to prevent potential learning effects. Once baseline performance stability was established, as described previously, treatment for participants' first set of treated stimuli began.

***Treatment Phase***

During each treatment phase, probes for the set under treatment were administered every other treatment session, beginning with Session 2. Probes for sets not currently receiving treatment alternated every third session. Probes were administered before the treatment session began to avoid

the immediate influence of treatment and represented short-term maintenance effects. During the intervention phase, only the treated and untreated stimuli, with no control items, were included in the probes (Supplemental Material S3). The lack of inclusion of control items in probes was a design error and addressed in the Discussion. The order in which treatment sets were provided to participants was determined randomly to reduce the likelihood of conditional learning circumstances (i.e., order effects).

### *Maintenance and Follow-Up Phases*

Maintenance probes for previously treated set(s) occurred on a predetermined, rotating schedule to track progress maintenance during treatment of other experimental sets. Follow-up probes were administered for all experimental sets once the final treatment phase was completed. Follow-up probes were administered at 2-, 4-, and 8-weeks following the conclusion of the final treatment session. The 16-item sets were used during maintenance testing and the 20-item sets (used in the baseline phase) were used for follow-up probes.

## Procedures

### *Dependent Variable*

The dependent variable—percent correct production of irregular past tense verbs measured in sentence completion probes—was calculated separately for treated and untreated stimuli.

**Sentence Completion Probes.** As described previously, each experimental set included eight items designated for treatment and eight items designated as untreated to measure response generalization. During probes, participants completed each sentence with a single verb (e.g., *"Every day she bites something. Yesterday she?"*). No feedback was provided regarding accuracy. The prompt was repeated if the participant used a different verb than the prompted one, did not hear the prompt due to inattentive behavior, or requested a repetition. Each response that used the targeted verb (e.g., *ran*, *ranned*, or *run*; but not *walk*) was transcribed and scored. Standard orthographic spelling was used for transcription, except for the word "read" which could have been interpreted as marked for present or

past tense. For this word, /rid/ or "red" was transcribed for differentiation. Incorrect responses included unmarked (e.g., *run*), overregularized (e.g., *runned*), and double-marked (e.g., *ranned*) verbs.

**Discourse Measurement Contexts.** Proportions of accurate irregular past tense verbs produced across narrative activities (Supplemental Material S4) during the treatment sessions were tracked. No expected number of verb productions was determined *a priori.* No feedback was provided during these activities to ensure the best opportunity to observe verb use in a discourse setting. As in the sentence completion probes, irregular past tense verbs produced as unmarked, overgeneralized, or double-marked were considered incorrect productions. Because the narratives were incorporated into the treatment procedures, these data were not considered part of the experimental design *per se,* and stability of production was not measured during the baseline phase.

## Intervention

### *General Procedures*

Participants received two individual sessions per week in their homes. Treatment was provided by the first author, a graduate student in speech-language pathology, and the second author, a speech-language pathologist completing a clinical fellowship. All sessions were audio/video recorded for fidelity and reliability purposes. Intervention sessions were designed to be 50 minutes, aligning with standards across treatment studies (Law et al., 2004). Actual session duration (excluding probes) ranged between 35 and 84 minutes, with a mean duration of 52 minutes. All eight treatment stimuli in an experimental set were presented each week, randomly presented in groups of four across the two sessions. Because P2 and P3 were twins and assigned different treatment sets, they received simultaneous in-person sessions in separate rooms at their home. When treatment moved to remote delivery (see “Unexpected Events” section), sessions were delivered without their twin present. Parents were instructed to read a homework story (described later) to each participant individually five times per week.

Because past tense is obligatory only when speaking about events that have already occurred (Spencer and Peterson, 2020), a narrative-based protocol was used as the context for intervention. Past tense irregular verbs were contrasted throughout the session with regular past tense *-ed* verbs (Connell, 1982) through feedback and during the presented syntax stories (Supplemental Material S5). Therapeutic techniques were applied as needed, including verbal models, evoked imitation, and elicited production via open-ended question prompts. Session protocols were identical for each treatment set, with changes only to the treated stimuli and associated activities.

***Specific Intervention Procedures***

Intervention procedures followed established protocols using a script and a response chart (Supplemental Material S4 and S6). Each session began with an elicited independent retell (e.g., "What did you do today?") followed by a probe (if scheduled), a report on homework fidelity, and a delayed retell of the homework story. Next, the interventionist began a series of evoked imitations of sentences containing the treatment stimuli followed by a syntax story read to the participant while they looked at a picture prompt (Supplemental Material S5). Each story incorporated all eight treatment stimuli five times each, for 40 cumulative target presentations. Afterward, an immediate retell of the syntax story was elicited from the participant. Next, naturalistic activities (i.e., games) commenced. Following each turn in the activity, the participant described what happened (e.g., "I *sank* the boat!"), for a total of ten productions before moving to the next activity. Throughout the session, the participant wrote keywords on a whiteboard to assist them in recalling the events of the session. At the conclusion of the session, the interventionist and participant used the whiteboard to review the events of the session (e.g., *"We sank the boat"*). Following this review, the participant retold the session events to a parent without assistance from the interventionist while the interventionist recorded the past tense verbs produced by the participant. Participants and parents were reminded each session to read the homework story (the

syntax story presented during the session) at least five days each week. Incentives for homework completion included stickers and positive feedback from the interventionist at the next session.

**Unexpected Events**

Midway through the study (March 2020), the emergence of COVID-19 resulted in a mandatory shutdown of intervention studies. Treatment had just concluded for Participant 1 and he was in the 2-week period prior to his first follow-up probe. Participants 2 and 3 were between their first and second experimental treatment phases. The shutdown required moving treatment delivery from in-person to remote via video conferencing tools. Upon approval by the institutional review board, re-consent by parents of all participants was obtained before continuing treatment and data collection. The transition resulted in P1 missing his 2-week follow-up probe. P2 and P3 had an 11-week break due to parental request, as the parents felt overwhelmed with the sudden change to at-home schooling. Once summer break began, probes for P2 and P3 began again to ensure performance stability before initiating the second treatment phases. Visual inspection of the data indicated that the transition to remote treatment administration did not result in any noticeable changes in probe performance relative to preceding face-to-face probes.

**Treatment Fidelity Procedures**

Both interventionists completed specific training in the study intervention procedures. The first author developed a treatment protocol that was studied and utilized by both interventionists during treatment administration. To facilitate fidelity of administration of treatment, a response and fidelity chart was developed and used for every treatment session (Supplemental Material, S6). This chart was designed to ensure that all components of treatment were completed in the correct order and the desired number of responses were elicited during specific activities. All treatment sessions were audio/video recorded for the interventionists to watch each other's sessions to ensure consistency of

treatment application. The response and fidelity charts were used to review recorded sessions periodically to evaluate interventionist performance.

**Reliability: Dependent and Independent Variables**

Point-to-point inter-rater scoring reliability was completed for 20% of the probes across each phase and participant. The first author used a random number generator to randomly select probes from each design phase for scoring by a second observer. The second observer rescored the probes using the session recordings. Differences between the first and second scorer's recorded responses regarding their accuracy (i.e., correct/incorrect) were considered disagreements. The total number of agreements across all comparisons was divided by the sum of agreements and disagreements to obtain a reliability rate. The mean agreement for all participant probes was 98% (range 81-100%). Because of the satisfactory inter-rater reliability, the original rater's scores were then used for all further analyses.

Ten percent of all recorded sessions, balanced across interventionists, were used to determine reliability of the independent variable. Provision of each session component was verified according to items on the response chart (Supplemental Material S6) to determine accuracy of treatment administration. Total accuracy of treatment delivery was 99.3%, indicating that elements of the intervention that were considered essential were administered according to the treatment protocol.

## Results

**Analyses**

Percent correct productions of irregular past tense forms on sentence completion probes is shown in Figures 1, 2, and 3 for Participants 1, 2, and 3, respectively. Within each figure, each of the three graphs displays responses to a categorical set of irregular forms (i.e., vowel change, final alveolar, dual change), with treated and untreated items portrayed separately. The order of the graphs from top to bottom in each figure indicates the order of treatment application.

**[Figure 1 near here]**

**[Figure 2 near here]**

**[Figure 3 near here]**

Identifying a systematic treatment effect requires evaluating the level, trend, and variability of probe data within and across design phases (Kratochwill et al., 2010). The Conservative Dual Criterion method (CDC; Fisher et al., 2003) was used to assist in the interpretation of the data for treatment effects. The CDC method used the baseline performance on treated items to create level (i.e., mean) and trend (i.e., slope) lines, adjusted upward by 0.25 standard deviations and extended through the treatment phase data. These criterion lines indicated participants' expected performance with irregular verbs across repeated probing if they did not receive treatment. With the CDC method, changes in behavior can be attributed to treatment when a predetermined number of data points fall above both the level and trend lines. For example, a treatment phase containing 8 data points must have 7 points above the level and trend lines to meet the CDC criterion for determining causality of treatment effects. The CDC method has been shown to convey rigorous control for type-1 error even in the presence of autocorrelation (Fisher et al., 2003). Due to improved performance on untreated stimuli sets by P2 and P3, data from an extended baseline period were used to create level and trend lines for the second and third applied treatment sets. Probe data from treated items reflected possible acquisition effects of treatment (as opposed to response generalization effects of treatment) and thus, were considered appropriate for determining if a functional relationship existed between treatment and behavior.

The Nonoverlap of All Pairs (NAP; Parker & Vannest, 2009) was used to statistically estimate the effect of treatment (Table 2). The NAP summarizes the amount of overlap between each data point in the baseline phase compared to each data point in the treatment phase. The NAP equals the number of comparison pairs with no overlap across phases divided by the total number of comparisons made and represents the probability that a random score from the treatment phase would exceed a random score selected from the baseline phase. NAP scores correspond to the following effect sizes: weak 0-.65;

medium .66-.92; strong .93-1.0. A score of .50 denotes chance effect, and scores below .50 indicate deteriorating performance. Extended baselines, when present, were used to calculate the NAP to estimate the effect of treatment application beyond observed improvements in the initial baseline during application of prior treatment sets.

**[Table 2 near here]**

Proportion of accurate irregular past tense verbs produced by participants in the four narrative contexts that occurred during treatment sessions are reported as Supplemental Material S7, S8, and S9. These data are provided for descriptive purposes only.

***Participant 1***

As shown in Figure 1, P1 received treatment for vowel change verbs first, followed by final alveolar verbs. Dual change verbs were left untreated due to nonresponse to the other treatment sets. P1 received 30 treatment sessions across two experimental sets with a mean session duration of 52 minutes. Five baseline probes ranged between 0 to 13% accuracy for Set 1 and 0% accuracy for the remaining sets. Following treatment application for Set 1, no change in behavior was observed in the treated or untreated sets. Likewise, no change in behavior was observed for treatment Set 2.

As seen in the top graph of Figure 1, two of 11 data points during the treatment phase fell above both CDC lines. According to CDC criteria, nine data points were required to exceed the dual criterion lines to attribute a treatment effect. Thus, there was a lack of treatment effect for this set. Similarly, treatment Set 2 had five data points. P1 did not exceed the mean or trend lines for any data point. NAP estimates of effect size (Table 2) indicated weak treatment effects for Set 1 (.39) and Set 2 (.50). Together with the visual analyses and CDC interpretation, the data reflect a non-effect of treatment for P1. However, P1 demonstrated possible emerging growth in his verb production during the personal retell at the conclusion of each session (Supplemental Material S7).

There were a few deviations from the planned experimental protocol with P1. Although gains were not evident after 10 sentence completion probes, it was decided to continue treatment to the maximum of 20 sessions during the first treatment phase to see if the participant needed more time to acquire these forms. The second treatment set was halted following 10 sessions due to non-effects of treatment (Figure 1) as planned. The two-week follow-up probe was canceled because of the mandatory shutdown of research due to the COVID-19 pandemic. The 4- and 8-week follow-up probes were administered remotely, but probe data for the 8-week follow-up was lost due to equipment malfunction.

In summary, planned probes for P1 included five initial baseline probes, 10 probes of treated items during Treatment Phase 1, five probes of treated items during Treatment Phase 2, 10 probes of treated items during Treatment Phase 3, and three follow-up probes of all items. However, the realization of the plan differed in that there were only five probes of treated items during Treatment Phase 2 due to lack of treatment response, and only one of the three follow-up probes was reported, as described above. A third treatment phase was not attempted due to lack of treatment effect in the first two treatment phases. There was no functional relationship demonstrated between treatment and behavioral responding for P1.

***Participant 2***

As displayed in Figure 2, P2 first received treatment for the final alveolar verbs (Set 2), followed by dual change verbs (Set 3), and then the vowel change verbs (Set 1). He received 55 treatment sessions across three experimental sets with a mean session duration of 51 minutes. Across the six baseline sessions, accuracy on the Set 2 items designated for treatment was consistently 0%. Following application of treatment for Set 2, increases in accuracy were observed for the trained verbs following three treatment sessions. P2 reached 100% accuracy on the trained verbs after 17 treatment sessions. Increases in accuracy for the untrained Set 2 verbs were also observed, reaching 50% accuracy.

During treatment application for Set 2, accuracy for Set 3 and Set 1 verbs increased up to 38%. Because of this, extended probing for Set 3 was conducted to establish behavioral stability prior to beginning treatment of this set. An additional six probes indicated stability had been established, with accuracy ranging between 25% and 38% for Set 3 treatment items. Performance remained variable when treatment was applied to Set 3, but gains were noted, with accuracy up to 75%. Untreated items also rose in accuracy. Treatment for Set 3 was halted following eight probes due to a lack of change noted on the final five probes administered during that treatment set.

During treatment of Set 3, variable performance on the untrained Set 1 (range 38-63%) and trained Set 2 (range 50-75%) was noted. As Set 1 was slated for the final treatment set, an additional six probes were administered before applying treatment. The extended probing of Set 1 showed a downward trend from 75% to 50-63% accuracy; thus, treatment was initiated. Upon treatment application, production accuracy was initially variable, but an upward trend was evident, with 100% accuracy reached for trained items and 75% accuracy for untrained items after 19 treatment sessions.

Maintenance probes for Set 2 and Set 3 indicated variability in performance during treatment of Set 1, but treatment gains were generally maintained. Follow-up probes indicated some loss in treatment gains, but performance remained above initial baseline levels for all sets. At 8-weeks post-treatment, accuracy of production of treatment items was as follows: Set 2 = 63%, Set 3 = 38%, and Set 1 = 63%. Decreases in gains for untreated exemplars were observed with all sets at follow-up, but accuracy levels remained above baseline for untreated items from Set 3 and Set 1.

The top graph of Figure 2 shows that nine of 11 data points for treated items during the treatment phase fell above both CDC lines, indicating systematic changes in behavior were associated with treatment for set 2. For Set 3, CDC lines were generated using the extended baseline data. The middle graph revealed that only five of eight data points fell above both CDC lines during the treatment phase, indicating a probable lack of treatment effect beyond gains already achieved, likely due to

generalization from Set 2 treatment. However, extension of the CDC lines through the maintenance phase for Set 3 suggests that additional behavior change was associated with Set 1 treatment application. That is, 12 of 15 combined treatment and maintenance phase probe points fell above both CDC lines. As seen in the final graph of Figure 2, only seven of 10 data points fell above CDC lines, indicating a probable lack of treatment effect beyond initial gains acquired during the initial treatment phase for Set 1 items. NAP estimates (Table 2) revealed a strong treatment effect for Set 1 (.91) and medium treatment effects for Sets 1 and 3 (.67 and .84, respectively). No consistent changes in production of irregular verbs in discourse were observed (Supplemental Material S8).

In summary, planned probes for P2 included six initial baseline probes, 10 probes of treated items during Treatment Phase 1, 10 probes of treated items during Treatment Phase 2, 10 probes of treated items during Treatment Phase 3, and three follow-up probes of all items. As noted, the pandemic necessitated suspension of the study at the end of Treatment Phase 1. When remote conduct of the study was allowed, an additional probe was conducted at the end of Treatment Phase 1 using remote measures to ensure that performance was stable (i.e., 11 probes for Treatment Phase 1). Prior to Treatment Phase 2, six additional probes of the to-be-treated set (Set 3) were required. During Treatment Phase 2, eight probes instead of 10 were completed due to plateauing of behavior. Six additional probes were also required for Set 1 prior to beginning Treatment Phase 3. Treatment Phase 3 and follow-up phase probes were completed as planned. A functional relationship between treatment and behavior was apparent for the first treated set. A functional relationship between treatment and behavior was likely for the second treated set but was not as clear due to variability in probe performance. Although gains were seen with the third treated set, a functional relationship with treatment and behavior was not established.

***Participant 3***

As depicted in Figure 3, P3 received treatment for Set 3 (dual change) first, followed by Set 2 (final alveolar). P3 received 29 treatment sessions across two experimental sets with a mean session duration of 54 minutes. After seven baseline probes that indicated stability at 13% accuracy, treatment was applied to Set 3. Immediate increases in accuracy for treated items were observed, with performance reaching up to 100% accuracy. The criterion for treatment termination was reached immediately prior to treatment session 17 and P3 moved into the extended baseline phase for Set 2. Performance with untreated items in Set 3 was variable but showed improvement across the treatment phase. Midway through treatment for Set 3, increases in accuracy on the untrained Sets 2 and 1 were also observed. Accuracy on untrained Set 2 and Set 1 items increased up to 75% by the completion of Set 3 treatment.

At the completion of treatment for Set 3 an 11-week break ensued (see "Unexpected Events" section). Extended probing was administered to ensure performance stability before applying treatment for Set 2. Stability was established (i.e., accuracy ranged from 50% to 38%), and treatment for Set 2 began. Accuracy in performance on treated behaviors increased up to 100%, and treatment was halted just before treatment session 14 because P3 met the performance criterion on the probe. Untreated items in Set 2 did not improve beyond the extended baseline data, suggesting limited generalization. Treatment was not applied to Set 1, due to high levels of accuracy achieved during treatment of Sets 3 and 2.

Accuracy of production of trained items from Set 3 was maintained throughout treatment of Set 2. Follow-up probes indicated treatment gains were maintained at levels well above baseline. At 8-weeks post-treatment, accuracy levels for treated items were as follows: Set 3 = 63%, Set 2 = 83%, and Set 1 = 100%. Accuracy of untreated items was also maintained at levels above baseline for Sets 3 and 1. As indicated in the top graph of Figure 3, all nine data points during the treatment phase fell above both CDC lines, indicating that systematic change in behavior can be attributed to treatment. The middle

graph shows 6 of 7 data points fell above both CDC lines, again meeting CDC criteria, indicating changes in behavior were associated with treatment. NAP estimates (Table 2) revealed a strong treatment effect for Set 3 (1.0) and Set 2 (.94). Production of irregular past tense verbs in discourse was variable (Supplemental Material S9) and did not indicate changes in production for any of the discourse conditions.

In summary, planned probes for P3 included seven initial baseline probes, 10 probes of treated items during Treatment Phase 1, 10 probes of treated items during Treatment Phase 2, 10 probes of treated items during Treatment Phase 3, and three follow-up probes of all items. P3 required fewer probes during Treatment Phases 1 and 2 due to performance criteria being met (i.e., Phase 1 = 9 probes, Phase 2 = 7 probes). An additional probe was conducted with the start of remote conduct of the study to ensure stability of performance. Five additional probes were required with Set 2 prior to initiating treatment with that set. Additionally, the third treatment phase was not required due to apparent generalization effects with Set 1. A functional relationship was established between treatment and behavior for both sets that received treatment. There was no opportunity to demonstrate another replication in a third treatment phase due to generalized responding with Set 1.

## Discussion

This study evaluated the efficacy of utilizing phonological schemas for target selection in treating irregular past tense verbs for children with DLD. This study is a first step in identifying optimal approaches for treating irregular past tense verbs. Patterns used for verb target sets were based on phonological schemas and included vowel change, final alveolar, and dual change verbs. Experimental control was demonstrated for two of three participants. Baselines, and extended baselines where appropriate, demonstrated stability before initiating treatment. Treatment effects were noted for P2 and P3 and were supported by CDC analyses. A functional relationship between treatment and behavioral change was clearly observed for both treatment phases for P3 and the first treatment phase

for P2. A likely association between treatment and behavioral change was also observed for the second treatment phase for P2. Probable generalization effects from the first two treatment phases obfuscated effects for the third treatment phase for P2 and rendered a third treatment phase unnecessary for P3. P1 demonstrated no direct treatment effects.

***Participant 1***

Data from P1 do not support using a pattern-based approach for selecting targets. During sessions and on probe tasks, P1 tended to produce a non-related word or state that he forgot the word, requiring more therapeutic support (e.g., repetitions, models) compared to P2 and P3. This difficulty was not observed during his eligibility assessments, including his responses during the TEGI and the comprehension portions of the TNL-2, nor did this difficulty seem related to his diagnosis of CAS. Specifically, during his eligibility assessment, P1 produced phonological patterns associated with repeated nouns and verbs (but without past tense marking) with regularity. Based on these observations during his eligibility assessments, his speech intelligibility is not considered to be the reason for his difficulty. However, P1 did demonstrate significant difficulty with his overall language expression, as evidenced by his scores on the CELF-4 and TNL-2.

The difference between P1's scores on eligibility measures compared to P2 and P3 may indicate key characteristics impacting their responses to the intervention. P1's score on the nonverbal assessment fell within normal limits but was considerably lower than P2 and P3's scores. Recall that the nonverbal ability task required examinees to identify visuospatial patterns. It may be that there is a threshold for identifying verbal or visuospatial patterns as a requisite skill for applying the Phonological Schemas theory to treatment. P1's relatively lower score reflected greater difficulty identifying patterns within sequences compared to P2 and P3 and could represent one reason for differences in response to treatment. Further, P1's consistent difficulty with recall after treatment began suggested the possibility that an unidentified cognitive issue may have also impacted his performance. This study did not assess

specific cognitive processes, such as memory, that could have contributed to P1's outcomes. Recognizing factors that support a positive treatment response or predict non-effects of treatment is essential for identifying which children might benefit from this approach to target selection.

Although the data from P1 do not provide support for the Phonological Schemas Theory, it also does not provide clear support for the Dual Mechanism Theory. Recall that the Dual Mechanism Theory posits that the irregular past tense forms would need to be taught directly, one at a time, for acquisition. P1 did not demonstrate acquisition of the treated irregular past tense verbs. However, P1's data do not rule out the possibility of the Dual Mechanism Theory altogether. It could be that to acquire the irregular past tense verb forms P1 needed increased intensity of treatment (i.e., dose frequency or total duration of treatment) to be successful. It is also the case that this treatment used implicit instruction and practice. It could be that P1 required explicit instruction to support his acquisition of the irregular past tense verbs. If increased intensity of treatment and/or explicit rather than implicit instruction had been provided to P1, his results may have been different and he may have shown acquisition of the treated irregular past tense verbs, supporting the Dual Mechanism Theory.

***Participants 2 and 3***

Data from P2 and P3 indicated a pattern-based approach holds promise as an effective method for selecting targets for some children with DLD. Acquisition and generalization to untreated stimuli were demonstrated within and across treatment sets. This may be particularly meaningful, as there is an apparent disadvantage for linguistic development for twins compared to singleton children (Hay et al.,1987; Thorpe et al., 2003) thought to result from less conversational interaction with parents due to increased competition for adult attention. While this may be accurate, P2 and P3 lived in a highly supportive language environment. It was observed by the researchers that the parents were very responsive to the questions and other linguistic interactions of P2 and P3. Therefore, their outcomes may reflect the combined effects of a positive linguistic environment upon treatment.

Our data indicated possible problems with our application of the Phonological Schemas Theory. Treatment effects appeared to partially generalize across sets, suggesting that the patterns we selected may not be categorically different from each other. Behavior changes across all three categories occurred within the first five treatment sessions for P2 and P3. However, the Dual-Mechanism Theory cannot provide a satisfactory explanation for our data either. A key tenet of this theory is that irregular forms are stored in the lexicon and retrieved individually, predicting no generalization would occur to untreated items. However, both P2 and P3 demonstrated improvement on untreated forms within and across phonological categories, indicating generalization and providing evidence against this hypothesis.

A hypothesis not previously considered in our thesis might better explain our data. The Multiple-Rules Model (Albright & Hayes, 2003) suggests inflection of verbs is based on probabilistic rules. That is, when inflectional marking is needed for a verb, the probability of an irregular or regular inflection is calculated and the most probabilistic inflection is applied. This model could explain why adults and children sometimes mark a single verb with various inflections (e.g., *spit, spitted, spat*). It could also explain the generalization patterns noted by P2 and P3. Their exposure to the irregular verb patterns may have provided morphophonological data used to inflect unknown verbs on probes.

The phonological complexity of the dual change verb category could also have influenced the generalization patterns observed. Dual change verbs require multiple steps for accurate inflection. Research examining the effects of complexity on treatment targets indicates an increased likelihood of generalization to simpler yet related forms (Gierut, 2001; Owen van Horne et al., 2017, 2018). Future studies should examine whether our findings may have resulted from probability calculations or complexity factors rather than phonological schemas.

**Limitations and Future Directions**

While this study provides some support for the use of a phonological schemas approach for treating irregular past tense with children who have DLD, it should be noted that the children who were

included in this study all demonstrated nonverbal IQ scores in the average range (i.e., standard scores of 80-120). The outcomes of this study may have been different if some or all of the participants demonstrated language impairment with lower nonverbal cognitive ability (i.e., standard scores 70<85). Children with low nonverbal ability and language impairment may slow in their acquisition of irregular past tense verbs at roughly 7 years of age (Rice et al., 2004). Future studies might compare the impact of cognition when treating irregular past tense.

Another limitation centers on the homework component of treatment. Although the parents of P2 and P3 were instructed to read the homework stories to each participant separately, their fidelity to this procedure was not tracked. Thus, mixed treatment effects via homework is possible. However, P2 and P3's different acquisition patterns suggest it unlikely that this oversight impacted the outcomes significantly.

***Treatment Design***

This study demonstrated that selecting irregular past tense verbs according to their phonological patterns was efficacious for acquisition and generalization. Evaluating which type of pattern is most efficacious or comparing these effects to other approaches (e.g., selecting targets based on frequency, telicity, or neighborhood structure) was outside the scope of this study.

A significant consideration for any treatment study is how well the approach would translate to clinical practice. One way this can be examined is in terms of effort, or the time required to produce the results. Our study aimed for approximately 50 minutes of treatment twice weekly. While this amount of effort is standard across treatment studies (Law et al., 2004), the majority of speech-language pathologists (SLPs) work within a school setting (ASHA, 2021) and are unlikely to provide sessions as long as those used in this study. Thus, future studies should consider whether this pattern-based approach may be efficacious in a condensed format or with increased treatment intensity (Warren et al., 2007).

***Study Design***

Weaknesses in the current study design should be addressed in future research. Specifically, there was a flaw in the implementation of probes. As noted, the baseline and follow-up probes included four regular past tense *–ed* verbs per set (i.e., 20 items per set) while the probes used throughout the treatment and maintenance phases of the study did not include these verbs (i.e., 16 items per set). A crucial premise of single-case experimental designs is that measurement conditions remain the same throughout all phases of the study. The inclusion of the regular past tense *-ed* verbs in only the baseline and follow-up phases was due to miscommunication among investigators and weakened the design controls. Fortunately, probe performance speaks against any substantial influence of this design error. If removal of the probe items after baseline had influenced probe performance, behavioral changes would have been expected soon after removal and should have similarly impacted treated and untreated items. For P2, improvement in performance with the first trained set was not noted until after three treatment sessions for treated items and after 11 treatment sessions for untreated items. For P3, gains in accuracy were seen in the first probe after baseline, but untreated items were at baseline levels after three and five treatment sessions. Further, treatment effects were evident with the second treated set for P3 and were likely for the second treated set for P2. These effects followed extended probing that did not include the regular past tense *-ed* verbs.

The MBD across participants design was not concurrent for P1 but was concurrent for P2 and P3. A totally concurrent design would have provided more experimental control. That is, if all participants had been enrolled concurrently, we could have coordinated initial baselines across participants, which would have provided additional controls to historical threats to internal validity. However, concurrent MBDs are difficult to conduct due to the amount of interventionist time required. In particular, when studies are unfunded (such as the current investigation) and/or are conducted by clinician-researchers, the personnel demands of a concurrent MBD may be prohibitive.

As noted, the number of planned probes differed from the number of completed probes for all participants. The loss of two follow-up probes for P1 is certainly a weakness. Back-up recordings (i.e., using more than one recording device) could have mitigated the loss of one of those follow-up probes. The other differences in number of probes conducted represent the flexibility of single-subject experimental designs. The use of extended probing prior to initiation of treatment for subsequently trained sets allowed the demonstration of behavioral stability. The coordination of number of probes with performance criteria permitted us to tailor the design to the individual participants. Such flexibility is considered a strength of single-case experimental designs (Connell & Thompson, 1986).

**Concluding Remarks**

The findings from this multiple-baseline single-subject experimental design investigation provide preliminary support for a pattern-based approach to target selection when treating irregular past tense verbs. Although the data indicated the possibility that the selected phonological schemas were not as distinct as anticipated, observed generalization to untreated forms on probes suggests this approach to target selection may lead to more efficient acquisition of irregular past tense forms during treatment.

**Acknowledgments**

We gratefully acknowledge the contribution of our participants and their families to this investigation.

**Data Availability Statement**

Data collected for the current study are available on request from the authors.

**References**

Albright, A., & Hayes, B. (2003). Rules vs. analogy in English past tenses: A computational/experimental study. Cognition, 90, 119-161. https://doi.org/10.1016/S0010-0277(03)00146-X

Archibald, L. M. D., & Joanisse, M. F. (2009). On the sensitivity and specificity of nonword repetition and sentence recall to language and memory impairments in children. *Journal of Speech, Language, and Hearing Research, 52*, 899–914. https://doi.org/10.1044/1092-4388(2009/08-0099)

ASHA (2021). *Member and Affiliate Profile*. https://www.asha.org/siteassets/surveys/2021-member-affiliate-profile.pdf

Betz, S. K., Eickhoff, J. R., & Sullivan, S. F. (2013). Factors influencing the selection of standardized tests for the diagnosis of specific language impairment. *Language, Speech, and Hearing Services in Schools, 44*(2), 133-146. https://doi.org/10.1044/0161-1461(2012/12-0093)

Bishop. (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), 20120401–20120401. https://doi.org/10.1098/rstb.2012.0401

Bishop, D. V., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & the Catalise-2 Consortium (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. https://doi.org/10.1111/jcpp.12721

Bybee, J. L., & Slobin, D. I. (1982). Rules and schemas in the development and use of the English past tense. *Language, 58*(2), 265-289. https://doi.org/10.1353/lan.1982.0021

Caesar, L. G., & Kohler, P. D. (2009). Tools clinicians use: A survey of language assessment procedures used by school-based speech-language pathologists. *Communication Disorders Quarterly*, *30*(4), 226-236. https://doi.org/10.1177/1525740108326334

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

Connell, P. J. (1982). On training language rules. *Language, Speech, and Hearing Services in Schools*, *13*(4), 231-240. https://pubs.asha.org/doi/pdf/10.1044/0161-1461.1304.231

Connell, P.J., & Thompson, C.K. (1986). Flexibility of single-subject experimental designs. Part III. Using flexibility to design or modify experiments. *Journal of Speech and Hearing Disorders, 51*(3), 214-225. https://doi.org/10.1044/jshd.5103.214

Crystal, D. (1985). Some early problems with verbs. *Child Language Teaching and Therapy, 1*(1), 46-53. https://doi.org/10.1177/026565908500100106

Dollaghan, C., & Campbell, T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research, 41,* 1136-2246. https://doi.org/10.1044/jslhr.4105.1136

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. https://doi.org/10.1177/0265659013512321

Eyer, J. A., & Leonard, L. B. (1994). Learning past tense morphology with specific language impairment: A case study. Child Language Teaching and Therapy, 10(2), 127-138. https://doi.org/10.1177/026565909401000201

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

Fisher, W. W., Kelley, M. E., & Lomas, J. E. (2003). Visual aids and structured criteria for improving visual inspection and interpretation of single-case designs. *Journal of Applied Behavior Analysis*, *36*(3), 387-406. https://doi.org/10.1901/jaba.2003.36-387

Gierut. (2001). Complexity in Phonological Treatment: Clinical Factors. *Language, Speech & Hearing Services in Schools, 32*(4), 229–241. https://doi.org/10.1044/0161-1461(2001/021)

Gillam, R., & Pearson, N. (2017). *TNL-2 : Test of Narrative Language* (Second edition). Pro-Ed.

Graf Estes, K. G., Evans, J. L., & Else-Quest, N. M. (2007). Differences in the nonword repetition performance of children with and without specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 50*, 177-195. https://doi.org/10.1044/1092-4388(2007/015)

Hay, D. A., Prior, M., Collett, S., & Williams, M. (1987). Speech and language development in preschool twins. *Acta Geneticae Medicae Et Gemellologiae: Twin Research, 36*(2), 213-223. https://doi.org/10.1017/S000156600000444X

Johnson, B. W., & Fey, M. E. (2006). Interaction of lexical and grammatical aspect in toddlers' language. *Journal of Child Language*, *33*(2), 419-435. https://doi.org/10.1017/S0305000906007410

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). Single-case designs technical documentation. *What works clearinghouse*. https://eric.ed.gov/?id=ED510743

Kuczaj II, S. A. (1977). The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior, 16*(5), 589-600. https://doi.org/10.1901/jaba.2003.36-387

Kulkarni, A., Pring, T., & Ebbels, S. (2014). Evaluating the effectiveness of therapy based around Shape Coding to develop the use of regular past tense morphemes in two children with language impairments. Child Language Teaching and Therapy, 30(3), 245-254. https://doi.org/10.1177/0265659013514982

Law, J., Garrett, Z., & Nye, C. (2004). The efficacy of treatment for children with developmental speech and language delay/disorder. *Journal of Speech, Language, and Hearing Research, 47,* 924–943. https://doi.org/10.1044/1092-4388(2004/069)

Law, J., Garrett, Z., & Nye, C. (2005). Speech and language therapy interventions for children with primary speech and language delay or disorder. *Campbell Systematic Reviews, 1*(1), 1-85. https://doi.org/10.4073/csr.2005.5

Leonard, L. B., Davis, J., & Deevy, P. (2007). Phonotactic probability and past tense use by children with specific language impairment and their typically developing peers. *Clinical Linguistics & Phonetics*, *21*(10), 747-758. https://doi.org/10.1080/02699200701495473

Leonard, L. B., Deevy, P., Kurtz, R., Chorev, L. K., Owen, A., Polite, E., ... & Finneran, D. (2007). Lexical aspect and the use of verb morphology by children with specific language impairment. https://doi.org/10.1044/1092-4388(2007/053)

Marchman, V. A. (1997). Children's productivity in the English past tense: The role of frequency, phonology, and neighborhood structure. *Cognitive science*, *21*(3), 283-303. https://doi.org/10.1016/S0364-0213(99)80025-1

Marchman, V. A., Wulfeck, B., & Weismer, S. E. (1999). Morphological productivity in children with normal language and SLI: A study of the English past tense. *Journal of Speech, Language, and Hearing Research, 42*(1), 206-219. https://doi.org/10.1044/jslhr.4201.206

Marshall, C. R., & Van Der Lely, H. K. (2007). The impact of phonological complexity on past tense inflection in children with Grammatical-SLI. *Advances in Speech Language Pathology*, *9*(3), 191-203. https://doi.org/10.1080/14417040701261509

Naglieri, J. A. (2003). *Naglieri Nonverbal Ability Test: Individual administration.* San Antonio, Tx: Harcourt Assessment.

Nitido, H., & Plante, E. (2020). Diagnosis of developmental language disorder in research studies. *Journal of Speech, Language, and Hearing Research*, *63*(8), 2777-2788. https://doi.org/10.1044/2020_JSLHR-20-00091

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. https://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. https://doi.org/10.1044/2018_LSHSS-STLT1-17-0142

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

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 Speech-Language Pathology*, *17*(6), 605-616. https://dx.doi.org/10.3109%2F17549507.2015.1034174

Parker, R. I., Vannest, K. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy, 40*, 357-367. https://doi.org/10.1016/j.beth.2008.10.006

Pinker, S. (1999). *Words and rules: The ingredients of language.* Basic Books.

Redmond, S. M. (2005). Differentiating SLI from ADHD using children's sentence recall and production of past tense morphology. *Clinical Linguistics and Phonetics, 19(2),* 109-127. https://doi.org/10.1080/02699200410001669870

Redmond, S. M., Thompson, H. L., & Goldstein, S. (2011). Psycholinguistic profiling differentiates specific language impairment from typical development and from attention-deficit/hyperactivity disorder. *Journal of Speech, Language, and Hearing Research, 54*(1), 99-117. https://doi.org/10.1044/1092-4388(2010/10-0010)

Redmond, S. M., Ash, A. C., Christopolus, T. T., & Pfaff, T. (2019). Diagnostic accuracy of sentence recall and past tense measures for identifying children's language impairments. *Journal of Speech, Language, and Hearing Research, 62*(7) 2438-2454. https://doi.org/10.1044/2019_JSLHR-L-18-0388

Rice, M. L., Tomblin, J. B., Hoffman, L., Richman, W. A., & Marquis, J. (2004). Grammatical tense deficits in children with SLI and nonspecific language impairment. *Journal of Speech, Language, and Hearing Research, 47*(4), 816-834. https://doi.org/10.1044/1092-4388(2004/061)

Rice, M. L., Wexler, K., & Cleave, P. L. (1995). Specific language impairment as a period of extended optional infinitive. *Journal of Speech, Language, and Hearing Research, 38(4),* 850-863. https://doi.org/10.1044/jshr.3804.850

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. https://doi.org/10.1044/jslhr.4106.1412

Rice, M.L., & Wexler, K. (2001). *Test of Early Grammatical Impairment*. San Antonio, TX: The Psychological Corporation.

Rice, M. L., Wexler, K., Marquis, J., & Hershberger, S. (2000). Acquisition of irregular past tense by children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 43(5), 1126-1144. https://doi.org/10.1044/jslhr.4305.1126*

Rinaldi, S., Caselli, M. C., Cofelice, V., D'Amico, S., De Cagno, A. G., Della Corte, G., ... & Zoccolotti, P. (2021). Efficacy of the treatment of developmental language disorder: A systematic review. *Brain Sciences, 11*(3), 407. https://doi.org/10.3390/brainsci11030407

Semel, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical evaluation of language fundamentals-4$^{th}$ edition*. Pearson.

Shipley, K. G., Maddox, M. A., & Driver, J. E. (1991). Children's development of irregular past tense verb forms. *Language, Speech, and Hearing Services in Schools, 22*(3), 115-122. https://doi.org/10.1044/0161-1461.2203.115

Smith-Lock, K. M. (2015). Rule-based learning of regular past tense in children with specific language impairment. *Cognitive neuropsychology*, *32*(3-4), 221-242. https://doi.org/10.1080/02643294.2014.951610

Spencer, T. D., & Petersen, D. B. (2020). Narrative intervention: Principles to practice. *Language, Speech, and Hearing Services in Schools, 51*(4), 1081-1096. https://doi.org/10.1044/2020_LSHSS-20-00015

Thorpe, K., Rutter, M., & Greenwood, R. (2003). Twins as a natural experiment to study the causes of mild language delay: II: Family interaction risk factors. *Journal of Child Psychology and Psychiatry*, *44*(3), 342-355. https://doi.org/10.1111/1469-7610.00126

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. https://doi.org/10.1044/2016_JSLHR-L-16-0138

Ullman, M. T. (2001) The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research, 30*(1)*,* 37-69. https://doi.org/10.1017/S1366728901000220

Warren, 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

Weiler, B. (2013). Verb selection and past-tense morphology: Crystal's criteria revisited. *Topics in Language Disorders*, *33*(2), 152-164. https://doi.org/10.1097/TLD.0b013e31828f50b8

**Table 1**

*Participant Characteristics*

| | Participant 1 | Participant 2 | Participant 3 |
|---|---|---|---|
| Age | 7;2 | 7;10 | 7;10 |
| NNAT[a] | 88 | 119 | 119 |
| RSR[ab] | 56 | 89 | 83 |
| NWR[ab] | 40 | 62 | 69 |
| TEGI EGC[c] | 21 | 75 | 61 |
| Phonological Probe | Pass | Pass | Pass |
| Irregular past tense probe | 33% | 57% | 13% |
| CELF-4 CLS[a] | 50 | 75 | 88 |
| TNL-2[a] | 75 | 100 | 105 |

*Note.* Age reported as years;months. NNAT = Naglieri Nonverbal Ability Test. RSR = Redmond Sentence Recall task. NWR = Nonword Repetition task. TEGI EGC = Test of Early Grammatical Impairment Elicited Grammar Composite. CELF-4 CLS = Clinical Evaluation of Language Fundamentals-Fourth Edition Core Language Score. TNL-2 = Test of Narrative Language-Second Edition.

[a]Standard score reported.

[b]Standard score conversion for RSR and NWR obtained from Redmond et al., 2019.

[c]Criterion score used.

**Table 2**

*NAP Calculations for each treatment set per participant*

| Participant | Treatment Set | No Overlap | Total Comparisons | NAP |
|---|---|---|---|---|
| Participant 1 | Set 1 | 21.5 | 55 | 0.39 |
| | Set 2 | 27.5 | 55 | 0.50 |
| Participant 2 | Set 2 | 60.0 | 66 | 0.91 |
| | Set 3 | 40.5 | 48 | 0.84 |
| | Set 1 | 40.0 | 60 | 0.67 |
| Participant 3 | Set 3 | 63.0 | 63 | 1.00 |
| | Set 2 | 33.0 | 35 | 0.94 |

*Note.* No Overlap = the number of comparison pairs with no overlap. Total comparisons = the total number of comparisons made across the baseline and treatment phases. Effect sizes: weak = 0-0.65; medium = 0.66-0.92; strong = 0.93-1.0. Calculations made according to Parker and Vannest (2009).

**Figure 1**

*Percent correct irregular past tense verbs produced on probes for Participant 1*

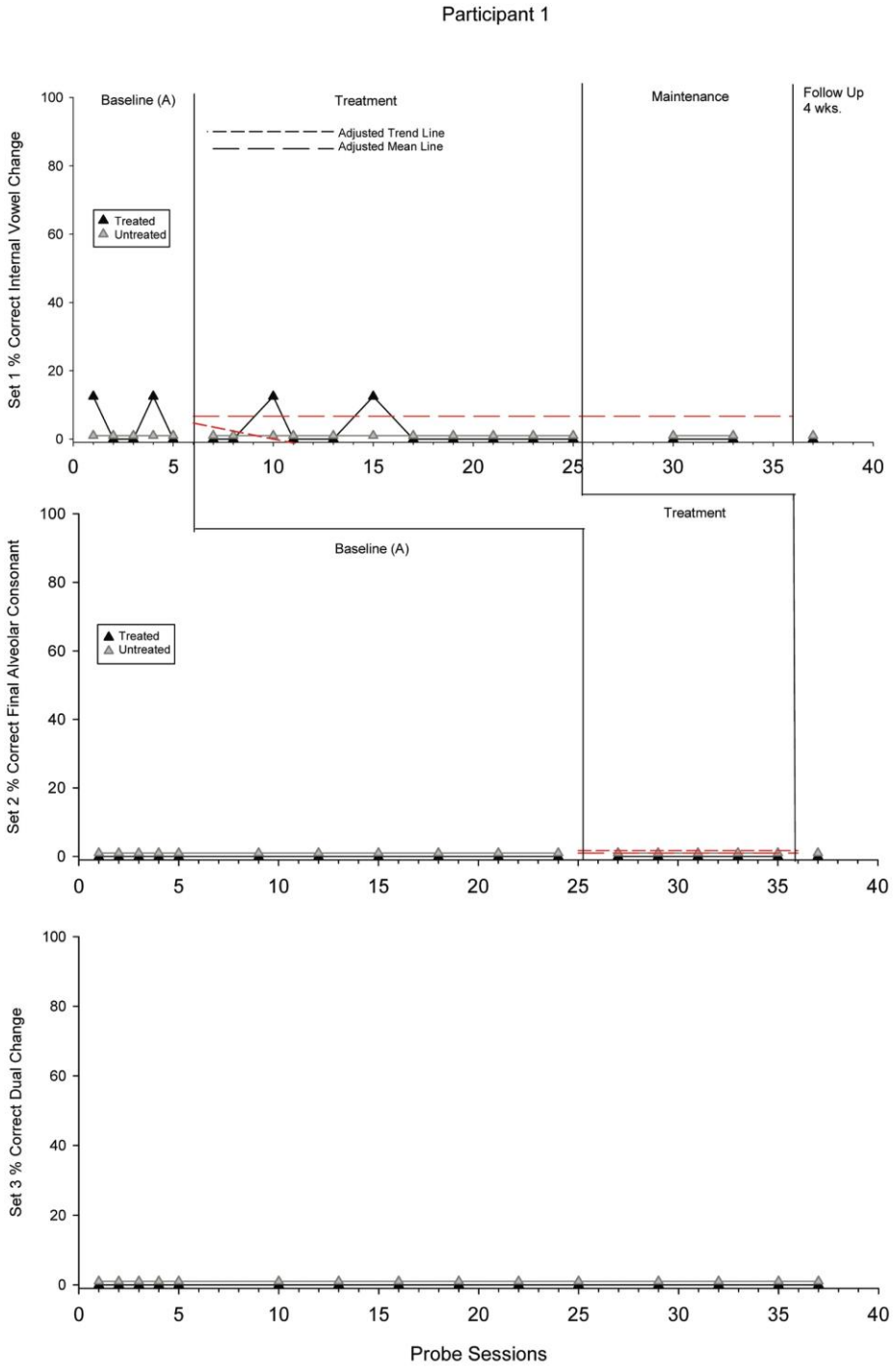


**Figure 2**

*Percent correct irregular past tense verbs produced on probes for Participant 2*

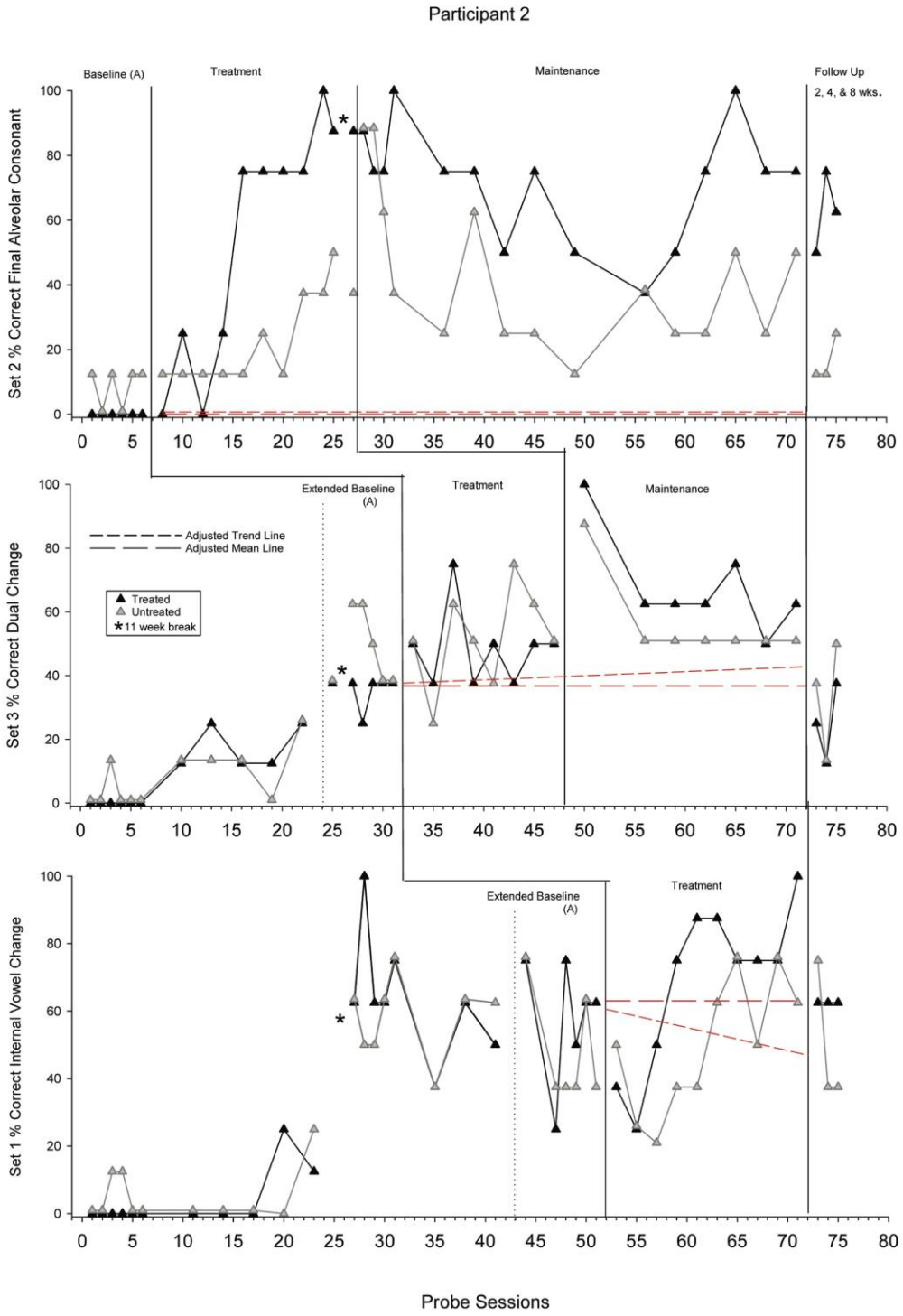
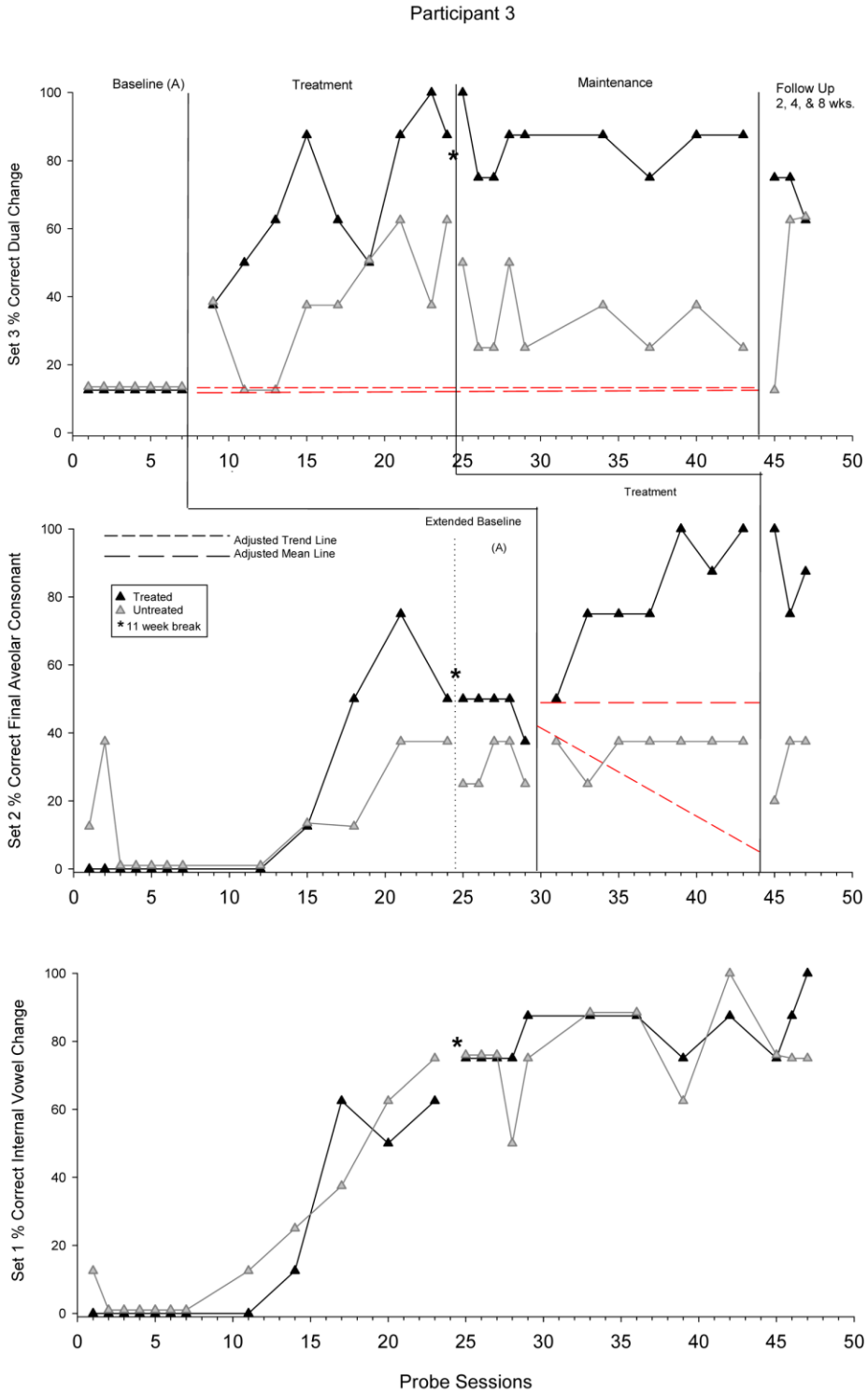


**Figure 3**

*Percent correct irregular past tense verbs produced on probes for Participant 3*



**Supplemental Material S1**

Target and generalization verbs for participants.

**Supplemental Material S2**

Sample probe for baseline and post-treatment phases.

**Supplemental Material S3**

Sample probe for treatment phase.

**Supplemental Material S4**

Sample treatment script.

**Supplemental Material S5**

Sample syntax story provided as part of the treatment script.

**Supplemental Material S6**

Sample recording and protocol fidelity chart.

**Supplemental Material S7**

Proportion of accurate irregular past tense verbs produced in narratives for Participant 1.

**Supplemental Material S8**

Proportion of accurate irregular past tense verbs produced in narratives for Participant 2.

**Supplemental Material S9**

Proportion of accurate irregular past tense verbs produced in narratives for Participant 3.

**Supplemental Material S1.** Target and Generalization Verbs for Participants.

*Treated and Untreated Stimuli for Participant 1*

| | Set 1<br>Vowel Change | Set 2<br>Final Alveolar Consonant | Set 3<br>Dual Change |
|---|---|---|---|
| Treated | Sink-Sank | Slide-Slid | Sell-Sold |
| | Give-Gave | Write-Wrote | Kneel-Knelt |
| | Take-Took | Find-Found | Leave-Left |
| | Run-Ran | Shoot-Shot | Keep-Kept |
| | Drive-Drove | Read-Read | Tell-Told |
| | Hang-Hung | Sit-Sat | Bring-Brought |
| | Win-Won | Bite-Bit | Feel-Felt |
| | Fall-Fell | Hold-Held | Teach-Taught |
| Untreated | Dig-Dug | Meet-Met | Buy-Bought |
| | Break-Broke | Fight-Fought | Say-Said |
| | Speak-Spoke | Hide-Hid | Catch-Caught |
| | Come-Came | Light-Lit | Sleep-Slept |
| | Wake-Woke | Lead-Led | Creep-Crept |
| | Swim-Swam | Eat-Ate | Hear-Heard |
| | Shake-Shook | Ride-Rode | Sweep-Swept |
| | Drink-Drank | Feed-Fed | Lose-Lost |

**Supplemental Material S2.** Sample Probe for Baseline and Post-treatment Phases.

*Examiner:* **I am going to say some words to make a sentence. But I am going to leave the last word for you to say.** (Start tapping the table while speaking…) **So, if I said: "I am tapping the…." you would say "table".**
**Let's try another one. "Your name is …."** (let the child insert their name).

- If they respond correctly, say **That's right. Let's do some more.**
- If they respond incorrectly to the trial, say: **That's not quite right. Your name is *(child's name),* so if I say "Your name is" you would finish my sentence by saying *(child's name)".* Let's try that again.** Then repeat the stimulus.

If the child does not know the meaning of a word that is presented, provide the definition or use an example, using the infinitive form (e.g., creeping means to crawl), then re-administer the prompt. DO NOT provide a past tense-marked example.

**Set 2:**

1. Every day they sit. Yesterday they ____.
2. The girl is finding things. Now she's all done. Tell me what she did. She___.
3. Every day he bites. Yesterday he ____.
4. The kids are feeding things. Now they're all done. Tell me what they did. They___.
5. Every day they camp. Yesterday they ____.
6. Every day she leads. Yesterday she ____.
7. Every day he writes. Yesterday he____ .
8. The kids are holding things. Now they're all done. Tell me what they did. They____.
9. Every day he eats. Yesterday he ____.
10. Every day he reads. Yesterday he ____.
11. The boy is hiding . Now he's all done. Tell me what he did. He____.
12. Every day they slide. Yesterday they ____.
13. Every day she bakes. Yesterday she ____.
14. The boy is painting. Now he's all done. Tell me what he did. He____.
15. Every day she lights things. Yesterday she ____.
16. Every day she works. Yesterday she ____.
17. The girl is riding. Now she's all done. Tell me what she did. She ____.
18. The girl is meeting. Now she's all done. Tell me what she did. She____.
19. The kids are fighting. Now they're all done. Tell me what they did. They____.
20. Every day they shoot things. Yesterday they ____.

**Supplemental Material S3.** Sample Probe for Treatment Phase.

*Examiner*: **I am going to say some words to make a sentence. But I am going to leave the last word for you to say.** (Start tapping the table while speaking…) **So, if I said: "I am tapping the…." you would say "table".**
**Let's try another one. "Your name is …."** (let the child insert their name).

- If they respond correctly, say **That's right. Let's do some more.**
- If they respond incorrectly to the trial, say: **That's not quite right. Your name is *(child's name),* so if I say "Your name is" you would finish my sentence by saying *(child's name)".* Let's try that again.** Then repeat the stimulus.

If the child does not know the meaning of a word that is presented, provide the definition or use an example, using the infinitive form (e.g., creeping means to crawl), then re-administer the prompt. DO NOT provide a past tense-marked example.

**Set 2:**

1. The boy is writing. Now he's all done. Tell me what he did. He _____.
2. Every day she rides something. Yesterday she _____.
3. The boy is leading someone. Now he's all done. Tell me what he did. He _____.
4. Every day they meet someone. Yesterday they _____.
5. Every day he eats. Yesterday he _____.
6. The girl is reading. Now she's all done. Tell me what she did. She _____.
7. Every day she slides. Yesterday she _____.
8. Every day he finds something . Yesterday he _____.
9. The kids are sitting. Now they're all done. Tell me what they did. They _____.
10. Every day they light something. Yesterday they _____.
11. The girl is feeding something. Now she's all done. Tell me what she did. She _____.
12. Every day he fights. Yesterday he _____.
13. Every day she bites something. Yesterday she _____.
14. Every day he shoots something. Yesterday he _____.
15. Every day they hide. Yesterday they _____.
16. The girl is holding something. Now she's all done. Tell me what she did. She _____.

941

942

**Supplemental Material S4.** Sample Treatment Script.

1) Establish rapport using *personal narrative retell*: "Tell me about something you did today"
   - Prompting may include: at school, at home, for fun, but no prompting with verbs.

2) Administer Probe if scheduled.

3) Review homework chart
   - Praise/encourage compliance.

4) *Delayed narrative retell* of homework syntax story
   - "Tell me all the parts you remember of your homework story"
   - No other prompting/correction/help with words

5) Write on cueing board
   Should be determined (and written) by the participant. Spelling is not the focus here. If they ask for help spelling something, you may provide it. If the participant selects a verb to describe a task that may not be expected (e.g., "played a game" for the probe instead of "told you a story" or used a verb unmarked, etc.) allow it to be used. Any verb is acceptable for this, if it is marked with past tense (e.g. regular past tense). If the word is marked incorrectly or not marked with past tense, recast the word with appropriate tense-marking and encourage them to write the correctly-marked word they selected. If the participant suggests an identical word for each item throughout the session, ask "what's another word we could use?" or "how else could we say that?". If they do not provide an appropriate response (new word), respond with **two** suggestions ("We could say we told a story or we could say that we talked about a story. What do you think should say?"). The goal is to have each verb repeated no more than once across the 7 verbs in the personal narrative at the end of the session.

6) Direct Imitation of Sentences

Write on cueing board

7) Presentation of syntax story

8) Child *immediate narrative retell* of syntax story

Write on cueing board

9) Naturalistic activities: (4 per session; all 8 targets per week): sank, drove, took, ran
   ⇒ Correction and requests for repetition of the correctly-produced response is necessary for each of the ten trials for each target word.

943

**Supplemental Material S5**

**Week 8 Syntax Story**
Tx Targets: Dug, threw, stuck, shook, chose, flew, rang, blew

Daniel looked around him. Everyone was ready for Halloween. Daniel chose to be a pirate this year. He chose to be a pirate because he loved ships. He loved how the wind blew against the ship, and how the waves flew up onto the deck during a storm. He could imagine himself a real pirate captain of his own ship.

His best friend Henry chose a completely different costume. He chose to be a superhero because he loved comics. Henry looked at Daniel. “You ready?” he asked. “Yep” Daniel responded. “Which street should we choose to go trick-or-treating on first?” Henry asked. “Hmm,” Daniel stuck his hand under his chin. He tapped his chin for a minute, then threw Henry a smile. “Let’s start on this street, then go around the corner to the next street, and let’s try to get one more street before we come home.” Henry grinned, and the boys started out.

Henry and Daniel walked to the first house. They rang the doorbell. “Trick or treat!” they yelled. The lady at the door threw some candy into each of their bags. The boys thanked her and moved on to the next house.

They rang the next doorbell. This time a man answered the door. He threw some candy into the boys’ bags. They thanked him. Daniel and Henry rang the next doorbell and the next. Each time they rang the doorbell, people threw candy into their bags. Halloween was amazing!

After the first street, Henry shook his bag. “Haha!” he smiled as he shook it. “I have so much candy!” Daniel shook his bag too. “Me too!” Daniel stuck his hand into his bag and dug around. He pulled out a laffy taffy. “I love laffy taffy!” he exclaimed.

Suddenly, the wind started blowing. It nearly blew Daniel’s hat off his head. “Oh no!” Daniel grimaced. He stuck his hat more firmly on his head and looked up at the sky. The wind blew and the leaves flew across the ground. “We’d better hurry. I think there is a storm coming. We want to finish our trick-or-treating before the storm comes!”
Henry agreed with Daniel. They chose their next house and hurried to the door. They rang the doorbell and the man at the door threw more candy into their bags.

The wind blew and blew as the boys hurried to each house on the street. Leaves flew across the sidewalk. They flew into the street. Once, the boys noticed an owl in the sky. He flew right over their heads.

Eventually, the boys noticed their bags were getting heavy. They shook them. “Wow,” Daniel smiled. He dug his hand into the bag. “I think my bag is full. We should probably go home now.” Henry dug his hand into his bag as he shook it. “Yep, my bag is full too. Let’s go home.”

When the boys arrived at Daniel’s house, they walked in. They stuck their bags on the kitchen counter. Daniel’s mom smiled. “You have a lot of candy,” she observed. Daniel stuck his hand in his bag and dug out a handful. “Yep. Look at all my candy.” Henry dug some candy out of his bag too.

**Supplemental Material S6.** Response Recording and Protocol Fidelity Chart.

Participant: _____ Date: _____ Session #: _____ Length of Session: _____ Clinician: _____

**1) Independent retell verbs:**

**3) Homework completed?** **Days Completed:**

**4) Syntax Story delayed retell verbs:**

| Cueing Board Verbs | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| | | | | | | |

| **6) Direct Imitation** | Target 1 | Target 2 | Target 3 | Target 4 | Target 5 | Target 6 | Target 7 | Target 8 |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | | |

**8) Syntax Story immediate retell verbs:**

| **9) Naturalistic Activities Target produced** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Target 1: | | | | | | | | | | |
| Target 2: | | | | | | | | | | |
| Target 3: | | | | | | | | | | |
| Target 4: | | | | | | | | | | |

**12) Personal Narrative retell verbs:**

945

**Supplemental Material S7**

*Proportion of accurate irregular past tense verbs produced in narratives during treatment sessions for Participant 1*

| | Session Number | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Narrative Task | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| Set 1 | | | | | | | | | | | | | | | | | | | | |
| Independent | 0/5 | 0/7 | 0/3 | 0/3 | 0/2 | 0/7 | 0/4 | 0/4 | 0/4 | 0/3 | 0/6 | 1/10 | 0/4 | 0/4 | 0/4 | 0/9 | 0/3 | 1/3 | 0/6 | 0/4 |
| Delayed | n/a | 0/5 | 0/2 | 0/2 | 0/2 | 0/3 | 0/1 | 0/3 | 0/2 | 0/3 | 0/4 | 0/2 | 0/3 | 0/3 | 0/0 | 0/4 | 0/2 | 0/3 | 0/2 | 1/2 |
| Immediate | 0/5 | 0/2 | 0/4 | 0/4 | 0/7 | 0/4 | 0/3 | 1/4 | 0/4 | 1/4 | 0/5 | 0/8 | 1/8 | 1/6 | 1/4 | 1/3 | 0/3 | 0/5 | 1/3 | 1/5 |
| Personal | 1/8 | 2/8 | 3/5 | 3/5 | 1/7 | 1/5 | 1/7 | 3/6 | 0/5 | 3/6 | 2/6 | 1/6 | 0/5 | 1/5 | 1/8 | 3/7 | 0/7 | 2/6 | 4/8 | 1/6 |
| Set 2 | | | | | | | | | | | | | | | | | | | | |
| Independent | 0/3 | 1/7 | 0/5 | 0/5 | 1/4 | 0/7 | 0/2 | 0/3 | 1/7 | 0/4 | | | | | | | | | | |
| Delayed | n/a | 0/2 | 0/2 | 1/3 | 1/2 | 1/1 | 2/4 | 1/5 | 0/1 | 0/1 | | | | | | | | | | |
| Immediate | 0/3 | 0/5 | 0/4 | 1/6 | 1/4 | 1/3 | 2/6 | 1/5 | 0/5 | 0/5 | | | | | | | | | | |
| Personal | 1/7 | 1/6 | 3/5 | 2/6 | 2/7 | 1/6 | 5/7 | 1/6 | 3/6 | 3/6 | | | | | | | | | | |

*Note.* Data from treatment sessions delivered are reported. Proportion of irregular past tense verbs compared to total verbs produced in the retell. Participant 1 received 20 treatment sessions for Treatment Set 1 and 10 treatment sessions for Treatment Set 2. Independent: a narrative in which Participant 1 shared something that happened that day or the day before. Delayed: a retell of the homework story from the previous week. Immediate: an immediate retelling of the story provided during the session. Personal: a retelling of the events that occurred during the session, using keywords chosen by the participant as visual cues for support.

946

**Supplemental Material S8**

*Proportion of accurate irregular past tense verbs produced in narratives during treatment sessions for Participant 2*

| | Session Number | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Narrative Task | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
| Set 2 | | | | | | | | | | | | | | | | | | | |
| Independent | 1/3 | 0/3 | 1/8 | 1/5 | 4/5 | 1/3 | 2/5 | 1/7 | 2/5 | 0/3 | 1/2 | 1/6 | 1/9 | 1/2 | 1/7 | 3/3 | 3/4 | 1/5 | 1/4 |
| Delayed | n/a | 0/3 | 4/9 | 2/5 | 1/2 | 2/4 | 2/6 | 2/4 | 2/6 | 2/3 | 2/2 | 3/9 | 2/7 | 4/5 | 3/3 | 6/13 | 5/13 | 5/11 | 2/9 |
| Immediate | 4/7 | 5/14 | 1/3 | 1/4 | 2/4 | 5/8 | 2/4 | 3/9 | 2/3 | 1/7 | 4/8 | 2/4 | 7/15 | 4/7 | 8/18 | 8/28 | 4/12 | 4/11 | 4/14 |
| Personal | 4/9 | 5/6 | 6/7 | 3/5 | 5/8 | 3/6 | 4/5 | 5/6 | 5/6 | 6/6 | 5/6 | 6/6 | 5/5 | 6/6 | 7/10 | 6/7 | 5/6 | 4/6 | 5/6 |
| Set 3 | | | | | | | | | | | | | | | | | | | |
| Independent | 0/3 | 3/3 | 1/10 | 4/4 | 0/4 | 1/2 | 2/3 | 0/1 | 2/7 | 1/3 | 0/7 | 0/2 | 0/2 | 1/4 | 0/5 | 1/3 | | | |
| Delayed | n/a | 4/10 | 3/12 | 2/11 | 3/8 | 2/3 | 3/4 | 2/4 | 1/3 | 3/4 | 2/3 | 4/7 | 2/5 | 3/9 | 6/8 | 1/5 | | | |
| Immediate | 1/10 | 4/8 | 4/7 | 2/8 | 3/4 | 1/2 | 1/3 | 1/2 | 0/1 | 2/5 | 3/4 | 4/8 | 5/29 | 4/8 | 5/9 | 0/2 | | | |
| Personal | 4/6 | 4/6 | 4/8 | 6/6 | 4/10 | 4/6 | 5/6 | 5/6 | 4/6 | 4/6 | 5/7 | 5/7 | 3/6 | 4/6 | 4/6 | 6/6 | | | |
| Set 1 | | | | | | | | | | | | | | | | | | | |
| Independent | 0/4 | 0/3 | 0/4 | 3/6 | 0/5 | 3/8 | 1/2 | 0/2 | 0/3 | 2/3 | 1/3 | 1/1 | 0/4 | 0/1 | 0/5 | 1/4 | 1/4 | 0/5 | 0/4 |
| Delayed | n/a | 2/6 | 2/6 | 2/6 | 4/8 | 5/10 | 2/5 | 6/13 | 12/22 | 5/11 | 2/9 | 4/11 | 4/12 | 0/1 | 0/7 | 2/4 | 2/2 | 1/5 | 1/3 |
| Immediate | 1/6 | 1/9 | 7/13 | 3/5 | 3/6 | 4/5 | 6/17 | 3/9 | 5/9 | 3/3 | 3/6 | 6/20 | 1/8 | 1/4 | 5/10 | 2/4 | 1/5 | 4/6 | 1/4 |
| Personal | 5/6 | 6/6 | 5/6 | 1/6 | 4/6 | 5/6 | 4/6 | 5/6 | 6/6 | 6/6 | 6/6 | 6/6 | 4/6 | 4/6 | 5/6 | 6/7 | 5/6 | 5/6 | 6/6 |

*Note.* Data from treatment sessions delivered are reported. Proportion of irregular past tense verbs compared to total verbs produced in the retell. Participant 2 received 19 treatment sessions for Treatment Set 2, 16 treatment sessions for Treatment Set 3, and 20 treatment sessions for Treatment Set 1. Independent: a narrative in which Participant 2 shared something that happened that day or the day before. Delayed: a retell of the homework story from the previous week. Immediate: an immediate retelling of the story provided during the session. Personal: a retelling of the events that occurred during the session, using keywords chosen by the participant as visual cues for support.

947

**Supplemental Material S9**

*Proportion of accurate irregular past tense verbs produced in narratives during treatment sessions for Participant 3*

| | Session Number | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Narrative Task | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| Set 3 | | | | | | | | | | | | | | | | | | | | |
| Independent | 0/5 | 0/3 | 8/17 | 1/4 | 1/6 | 2/11 | 2/5 | 1/6 | 1/3 | 0/4 | 2/8 | 1/3 | 3/6 | 1/6 | 4/8 | 4/6 | | | | |
| Delayed | n/a | 2/12 | 2/7 | 3/4 | 1/1 | 4/11 | 1/1 | 5/14 | 0/0 | 0/1 | 1/2 | 1/1 | 5/12 | 2/7 | 2/5 | 4/8 | | | | |
| Immediate | 6/19 | 8/22 | 2/3 | 4/6 | 4/7 | 11/38 | 2/6 | 0/3 | 0/2 | 1/2 | 1/9 | 3/10 | 6/11 | 1/7 | 5/6 | 3/11 | | | | |
| Personal | 5/6 | 4/6 | 5/6 | 6/7 | 4/6 | 4/6 | 6/6 | 6/6 | 4/6 | 4/6 | 4/6 | 3/6 | 3/6 | 6/6 | 6/7 | 3/6 | | | | |
| Set 2 | | | | | | | | | | | | | | | | | | | | |
| Independent | 3/9 | 3/12 | 0/7 | 0/9 | 2/5 | 2/6 | 1/11 | 4/12 | 3/6 | 0/2 | 2/4 | 1/5 | 0/6 | | | | | | | |
| Delayed | n/a | 6/12 | 1/4 | 3/7 | 2/3 | 0/0 | 2/2 | 0/2 | 7/11 | 1/3 | 2/4 | 3/3 | 4/8 | | | | | | | |
| Immediate | 4/15 | 8/21 | 3/5 | 0/5 | 0/2 | 1/10 | 2/4 | 2/3 | 6/14 | 4/5 | 3/3 | 4/12 | 2/4 | | | | | | | |
| Personal | 6/6 | 5/7 | 5/7 | 6/8 | 6/6 | 6/8 | 5/6 | 5/6 | 5/6 | 6/6 | 5/6 | 5/10 | 5/6 | | | | | | | |

*Note.* Data from treatment sessions delivered are reported. Proportion of irregular past tense verbs compared to total verbs produced in the retell. Participant 3 received 16 treatment sessions for Treatment Set 3 and 13 treatment sessions for Treatment Set 2. Independent: a narrative in which Participant 3 shared something that happened that day or the day before. Delayed: a retell of the homework story from the previous week. Immediate: an immediate retelling of the story provided during the session. Personal: a retelling of the events that occurred during the session, using keywords chosen by the participant as visual cues for support.

948