

Daily or Weekly? The Role of Treatment Frequency in the Effectiveness of Grammar Treatment for Children with specific language impairment

KAREN SMITH-LOCK¹, SUZE LEITAO², LARA LAMBERT³, POLLY PRIOR³, ANNE DUNN³, JULIA CRONJE³, SARA NEWHOUSE³, LYNDSEY NICKELS¹

1. *ARC Centre of Excellence in Cognition and its Disorders, Macquarie University, Sydney, NSW, Australia*

2. *Curtin University, Perth, WA, Australia*

3. *Department of Education Western Australia*

Correspondence: Dr Karen Smith-Lock, ARC Centre of Excellence in Cognition and its Disorders, Macquarie University, North Ryde NSW Australia 2109. E-mail: karen.smith-lock@mq.edu.au

Key Words: treatment effectiveness, specific language impairment, treatment frequency, grammar intervention, morpho-syntax

Running Head: Treatment Frequency

Abstract

This study compared the effectiveness of a school-based treatment for expressive grammar in 5-year-olds with specific language impairment delivered in two different dose frequencies: eight sessions delivered daily over eight consecutive school days or eight sessions delivered weekly over eight consecutive weeks. Eighteen children received treatment daily and 13 children received treatment weekly. In both groups, treatment consisted of eight one hour sessions of small group activities in a classroom setting. Techniques included explicit instruction, focussed stimulation, recasting and imitation. Results were analysed at the group level and as a case series with each child as their own control in single subject design. The eight weeks group showed significantly greater gain in test scores over the treatment period than in an equal time period prior to treatment, whereas the eight days group did not (Cohen's $d = 1.64$ for eight weeks group). Single subject analyses indicated that 46% of children in the eight-week group and 17% of children in the eight-day group showed a significant treatment effect. It is concluded that expressive grammar treatment was most effective when dose frequency was weekly over eight weeks rather than daily over eight days for 5-year-old children with specific language impairment.

Treatment Frequency

Difficulty with the acquisition of grammar, and morpho-syntax in particular, is a well-established hallmark of Specific Language Impairment (SLI) (Leonard, 1998; Rice, Wexler, & Cleave, 1995). Poor grammar can interfere with the ability to communicate successfully. For example, it is difficult to construct a coherent narrative without competent use of tense to indicate the timing of events and pronominal reference to link the characters in the narrative. Given the importance of grammar for effective communication, it follows that grammar treatment is a high priority. As such, research into the efficacy and effectiveness of grammar treatment is required. There is a growing body of research demonstrating the efficacy, and more recently, the effectiveness, of speech and language intervention and grammar intervention.

Efficacy is established when a treatment is shown to have beneficial results under ideal and tightly controlled conditions (Robey & Schultz, 1998). A number of studies have evaluated the efficacy of therapy using a single technique or contrasting treatment techniques while holding intensity or frequency constant across conditions. For example, Fey, Cleave, and Long (1993) demonstrated a significant treatment effect using focused stimulation and a cyclical goal attack strategy with 30 children aged three to five years with delayed grammatical development. The treatment program involved three hours per week over four-and-a-half months. Similarly, Leonard, Camarata, Pawlowska, Brown, and Camarata (2006) demonstrated treatment effects using focused stimulation and conversational recasts to treat verbs in a group of 25 three and four-year-old children. Children attended four 30-minute treatment sessions per week, condensed within two lab visits per week, for a total of 96 sessions. Connell and Stone (1992) contrasted modelling with imitation plus modelling in a study focusing on the learning of novel morphemes in children with SLI aged 5:0 to 6:11 years, and both age and language matched groups. Children received the same frequency of treatment, four sessions over two weeks. The children with SLI showed significant

Treatment Frequency

improvements in the imitation condition. Weismer and Murray-Branch (1989) also contrasted modelling with modelling plus evoked production to teach syntax and morphology to four children, aged 5 and 6 years of age, using a single subject experimental design. Children received 10-14 treatment sessions spanning 7-10 weeks. Three of the four children demonstrated treatment effects.

While the establishment of treatment efficacy is necessary and important, it is essential to know whether treatment that has been shown to be efficacious in controlled conditions is effective in every day clinical practice. Thus, in contrast to efficacy, treatment effectiveness evidence is drawn from research into the impact of a treatment in the less controlled “real world” of clinical practice (Dollaghan, 2007). To date, evidence of treatment effectiveness has been sparse. In fact, Cirrin, Schooling, Nelson, Diehl, Flynn, Staskowski, Torrey and Adamczyk (2010) note that “direct S/L [speech and language] intervention procedures ... have not been put to adequate experimental tests to determine their effectiveness in facilitating the development of S/L [speech and language] abilities in school age children...” (p. 249). One recent study did examine the effectiveness of expressive grammar treatment in a school setting. Smith-Lock, Leitão, Lambert and Nickels (in press) found that grammar treatment improved the use of grammatical targets in 5-year-old children with SLI. Smith-Lock et al. compared a group of children who received expressive grammar treatment to a group of children who received a control treatment (comprehension of prepositions). Children’s performance on a grammar elicitation task was measured before and after treatment. Treatment consisted of weekly sessions of 1 hour over an eight week duration; the children in the experimental group took part in small group activities designed to teach a grammatical target. Each activity provided opportunities for the teacher to model the grammatical target and for the child to use the correct target and receive feedback through a defined cuing hierarchy. Smith-Lock et al. found a significant group treatment effect with a large effect size

Treatment Frequency

($d = 1.24$). When children with articulation difficulties which prevented them from producing the grammatical targets were excluded from the analysis this effect size increased ($d = 1.66$).

The treatment effect was also significant at the single subject level for the majority of children who had the articulation skills to produce the grammatical target.

Thus, there is a small, but growing body of evidence demonstrating that grammar intervention for children with SLI can be both efficacious and effective. However, the studies have varied substantially in the number of sessions administered (four to 96), the total period of treatment (two weeks to six months) and the amount of treatment per week (30 mins to 2 hours). Given the potential importance of these factors in treatment outcomes it would seem vital to focus investigation on them explicitly.

Warren, Fey and Yoder (2007) outlined a detailed framework for the description of treatment intensity. They defined *dose* as the number of times a teaching episode occurs per treatment session, *session duration* as the length of each treatment session, *dose frequency* as the number of treatment sessions per unit of time (e.g., once per week), and *total intervention duration* as the total period of time over which the treatment was provided (e.g., 6 months). They considered the product of dose x dose frequency x total duration of intervention to be *cumulative intervention intensity*. The focus of this paper is on dose frequency (henceforth, *frequency*).

The frequency of treatment is an important issue for clinicians in the workplace, as they wish to treat as many children as possible as efficiently as possible. Treatment that is delivered at too high a frequency or for too long may lead to wasted resources, while treatment that is delivered too infrequently or for too short a time period may be little better than no treatment (Baker, 2012). Unfortunately, clinicians have little data to guide evidence-based decisions about the frequency of treatment sessions. Law, Garrett and Nye's (2004) meta-analysis of language treatment for preschool children concluded that the role of intensity

Treatment Frequency

(frequency) could not be decisively evaluated, although they did note that larger effect sizes were shown following interventions of longer overall duration than 8 weeks, though the frequency within these studies varied. More recently, Schooling, Venediktov and Leech (2010) conducted a systematic review on the effects of speech and language interventions in children up to 5 years of age. Like Law et al., they concluded that a greater amount of treatment was associated with a positive effect size, but were unable to directly evaluate the role of dose frequency due to a lack of available evidence. The 10 studies included in the review were all for 0- to 5-year-olds and none specifically addressed expressive morpho-syntax.

In the broad field of childhood communication disorders, there is some evidence that increasing dose frequency does not always improve treatment outcomes. For example, in a study of phonological awareness training for kindergarten children at-risk of reading difficulties (some of whom were from a non-English speaking background) Ukrainetz, Ross and Harm (2009) found that over four weeks, just four hours was as effective as 11 hours of treatment. Thus, there was no advantage to treatment that was higher in dose frequency and cumulative intervention intensity. In a study of a print-referencing intervention with pre-schoolers at risk of reading difficulties, McGinty, Breit-Smith, Fan, Justice and Kaderavek (2011) found that increasing frequency from two to four sessions per week with an overall duration of 30 weeks resulted in improved outcomes only when dose per session was low. Thus, there seemed to be a critical dose, below which increasing dose and/or dose frequency benefited learning, but above which there was no impact. Once again however, the effects of manipulating dose frequency and cumulative intervention intensity cannot be distinguished.

While McGinty et al. found that the effect of increasing dose frequency depended on dose, Yoder, Fey and Warren (2012) reported that the effect of increasing dose frequency depended on subject variables. They treated early word learning in two groups of toddlers

Treatment Frequency

with intellectual disability, one group with Down's Syndrome, one group without. They found that children with Down's Syndrome (Yoder, Woynaroski, Fey, & Warren, 2012) and children with initially high object interest (Warren, Fey, & Yoder, 2007) benefitted more from 5 hours per week than one hour per week of intervention. In a study of word learning, Riches, Tomasello and Conti-Ramsden (2005) found that 5-year-olds with SLI learned novel verbs better after 12 presentations spaced across four sessions than after 18 presentations in a single session. Thus, a more distributed frequency was more beneficial than a massed presentation, even with a lower cumulative intervention intensity. In summary, an increase in treatment frequency and cumulative intervention intensity does not appear to lead in a straightforward manner to an increase in treatment outcomes. Potentially, the participant population (at risk readers, children from non-English speaking backgrounds, children with SLI, toddlers with intellectual disability), the type of treatment (phonological awareness, print awareness, word learning) and, importantly, cumulative intervention intensity may all play a role in the effect of treatment frequency.

While the literature described above does not allow the effects of dose frequency and cumulative intervention intensity to be dissociated, there is some research which suggests that dose frequency specifically can influence language intervention outcomes. Barratt, Littlejohns and Thompson (1992) compared the results of 24 sessions of language therapy for children aged 2 to 5 years, provided either weekly for six months or four times per week for three weeks in each three months of a six month period. The more intensive schedule led to a greater improvement in expressive language outcome scores, while comprehension scores improved similarly in both conditions. On the other hand, in a study of 7- to 8-year-old children with SLI, Proctor-Williams and Fey (2007) found no correlation between frequency of sessions (five sessions spread over anything from 4 to 44 days) and the children's ability to learn irregular past tense forms of novel verbs. They did find, however, that the longer the

Treatment Frequency

children took to complete the five sessions, the less accurately they produced the target verbs. This finding suggests that frequent treatment sessions may be more successful than less frequent, more distributed sessions. More recently, Bellon-Harn (2012) compared treatment for semantic and morphology skills in 12 children with language impairment, aged 4-5 years, who were randomly assigned to either a concentrated treatment schedule (4 x 20 minute sessions per week for 6 weeks) or a distributed treatment schedule (2 x 20 minute sessions per week for 12 weeks). There was no untreated control group, or pre-treatment double baseline. The groups did not differ in language scores, phonology scores or nonverbal IQ pre-treatment. Outcome measures were the number of interpretive utterances, Mean Length of Utterance and the use of a variety of morpho-syntactic structures. While both groups improved significantly, there was no difference between the concentrated or distributed treatment groups on any of the outcome measures.

In summary, the evidence regarding optimal treatment frequency is equivocal, with some studies finding more frequent intervention to be more successful, others finding no difference, some finding interactions with frequency and other variables, and some reporting decreased returns after particular cumulative intervention intensity levels have been reached. Only one study has focused on morpho-syntax in children with language impairment, and that study was a pilot with low numbers and a lack of experimental control (Bellon-Harn, 2012).

Given the practical importance of evidence for differing frequencies of treatment and having established the effectiveness of a grammar treatment program when delivered weekly for eight weeks (Smith-Lock et al., in press), we wished to investigate the effect of different dose frequencies on the effectiveness of this same treatment program. In order not to confound the effect of dose frequency with cumulative intervention intensity, we aimed to keep cumulative intervention intensity constant. Thus, we compared the effectiveness of a school-based language intervention program delivered at two different frequencies and

Treatment Frequency

(necessarily) two different durations: daily, (eight sessions over eight days) or weekly, (eight sessions over eight weeks). This addressed the following research questions.

- (1) With cumulative intervention intensity held constant, is treatment more effective when delivered daily over eight consecutive days or weekly over eight consecutive weeks?
- (2) Are differences in treatment effectiveness between the two frequencies maintained when performance is compared after an equal period of time (i.e., after eight weeks for both groups)?
- (3) Are group results reflective of individual children's performance?

Method

Participants

Participants were drawn from two specialized schools for children with language impairment, Language Development Centres (LDCs), in Western Australia. All 52 children in their first year of full time schooling were invited to participate in the study; 51 out of 52 students agreed to participate. The children had a mean age of 5 years, 3 months (64 months). Participant flow is illustrated in figure 1. This figure represents the total number of participants at each stage of the project. Five of the children who agreed to participate were deemed ineligible due to diagnoses other than SLI and ten due to co-occurring speech impairment (see Articulation Screening below for details). This corresponds to figure 1, Exclusions: Outside clinical criteria. All of the 52 children were assigned to treatment conditions, as treatment was part of their regular classroom program. Five children were absent from school for one of the testing sessions and were therefore excluded from the analysis. The timing of their absences can be tracked in figure 1. Absent children were not excluded from subsequent treatment or testing. Of the 31 children included in the analysis, 13 received the treatment program over eight weeks and 18 children received the treatment

Treatment Frequency

program over eight days. This sample size gave us the ability to detect an effect size of .45, with power of 0.8. The eight weeks group contained two girls and the eight days group, five girls.

Insert figure 1 about here

Diagnosis of Language Impairment

Entry to the LDC required diagnosis as specifically language impaired by a speech-language pathologist. Children were assessed either with the Preschool Language Scale (PLS) (Zimmerman, Steiner, & Pond, 2002) or the Clinical Evaluation of Language Function (CELF-P or CELF-P2) (Semel, Wiig, & Secord, 1992; Wiig, Secord, & Semel, 2006) as one part of an extensive assessment process for referral to the school. Referral information also included evidence that children had non-verbal skills in the normal range, as attested by a psychologist or paediatrician. For this purpose, referring clinicians used a variety of tests, including the Wechsler Preschool and Primary Scale of Intelligence (WPPSI-III) (Wechsler, 2002), Cognitive Adaptive Test (Accardo & Capute, 2005), Denver Developmental Screening Test (Frankenburg, Dodds, & Archer, 1992) and Griffiths Mental Development Scales (Griffiths, 1970). We reviewed the school's intake data and included only those children who were unequivocally diagnosed as SLI according the school's intake criteria. We also confirmed the diagnosis of each child with staff speech-language pathologists at time of testing. Thus, diagnosis of SLI was made by three speech-language pathologists independently: the referring clinician, the school clinician and the clinician-researcher. For all children 1) receptive and/or expressive language was impaired, 2) there was documentation of normal non-verbal IQ and 3) clinical opinion was not consistent with an alternate diagnosis. Seventeen out of the 31 children had both receptive and expressive impairment on intake

Treatment Frequency

testing, nine in the eight weeks group and eight in the eight days group. The remainder had expressive language impairment only. Intake test data can be found in table 1.

All testing and intervention was carried out at the children's schools (LDCs). The study was approved by the Macquarie University Human Ethics Committee.

Insert table 1 about here

Group Allocation

Treatment conditions were allocated on the basis of site. One LDC carried out the program over eight weeks and the other over eight days (spread over two weeks). Each site delivered the treatment at the intervention frequency already in use in that LDC. One LDC typically delivered treatment weekly over an eight week period because that suited the duration of the school term. The other LDC preferred a more intensive model which allowed the speech-language pathologists to provide short, intensive periods of treatment to children and training to staff, while freeing them for other assessment, planning and consultative duties for the remainder of the term. For the purposes of this research, the duration of eight days was chosen to match the eight sessions of treatment received by the eight weeks group. The two sites did not differ in class size or access to professional development. There is no reason to believe that the general level of teaching expertise or the language environment differed between the sites. The sites drew on similar socio-economic populations.

Test Materials

Three tests were developed for the project: the Grammar Screening Test, the Articulation Screening Test, and the Grammar Elicitation Test (Smith-Lock et al., in press). In both the Grammar Tests, children were shown a series of pictures, and asked a question for each picture designed to elicit a specific grammatical structure. For example, to elicit the past

Treatment Frequency

tense, the child was shown a picture of a boy kicking a ball. The tester then asked “What did this boy do yesterday?”. Vocabulary items used in tests were not used in treatment activities.

The purpose of the Grammar Screening Test was to identify potential treatment targets for each child. The test consisted of six possessive *s*, six past tense, six present tense and four subject pronouns (two *he*, two *she*). For the possessive, present and past tense, two items were included for each allomorph ([s], [z] and [əz]; [t], [d] and [əd]).

The purpose of the Articulation Screening Test was to establish if the children had the articulation skills to produce the targeted grammatical goals. The children were asked to repeat 26 single syllable non-words which contained the final consonant clusters in the grammatical items tested (e.g. pept, avz).

The purpose of the Grammar Elicitation Test was to obtain multiple instances of a particular grammatical target to act as a pre-and post-test measure of treatment effectiveness. The test contained four sections: possessive, present tense, past tense and subject pronouns *he* and *she*. Each section consisted of 30 items, divided either into ten for each allomorph, or, in the case of the pronoun test, 15 items for each pronoun (he and she). Each child completed only the section of the test that applied to his or her grammatical target (i.e., 30 items). Items were presented in a different random order for each round of testing. Ten percent of the tests were scored by a second scorer for the purposes of inter-rater reliability, yielding a correlation of .99 (Pearson’s Product-Moment Correlation). Details of the Grammar Elicitation Test can be found in Supplemental Appendix A.

Each child first completed the Grammar and Articulation Screening Tests. Only children who made fewer than two errors on the articulation screen for their particular grammatical target were included in the study. Following the screening tests, a potential treatment target was selected for each child, according to the procedure below. The Grammar Elicitation Test was then carried out, in order to collect further data on each child’s selected

Treatment Frequency

grammatical target. The timeline of the project can be seen in figure 2. The Grammar Elicitation Test was administered three times, at eight week intervals for the eight-week group and at two week intervals for the eight-day group. Test 1 and Test 2 provided a pre-treatment baseline. Treatment occurred between Test 2 and Test 3. Test 3 measured post-treatment skills. Test 4 measured post-treatment skills in the eight-day group eight weeks after treatment began, to match the time frame of the eight-week group.

Insert figure 2 about here

A grammatical construction was considered a potential treatment target if the child made at least six errors on the Grammar Elicitation Test. If a child made errors on several target constructions, the construction which occurred first on a developmental continuum was chosen, in the following order: pronouns, possessives, past tense, present tense.

Interventions

The interventions are described in detail in Smith-Lock et al. (in press) and outlined here for convenience. A single treatment session involved both whole class (approximately 12 children) and small group activities. First, the staff speech-language pathologist delivered a whole group lesson focused on one of the treatment goals. All of the targets were addressed an equal number of times over the eight group sessions. Children were then divided into three groups, based on their identified treatment target. Each small group carried out three activities designed to teach them their specific grammatical target. The activities were led by the speech-language pathologists, classroom teachers and education assistants. Groups rotated such that each child completed activities with each treatment provider over the course of the eight sessions. The number of children in each group ranged from 3 to 6. The entire session, including the whole class introduction and the three small group activities, lasted one hour.

Treatment Frequency

The grammatical target was first introduced using explicit teaching. For example, the speech-language pathologist explained that “when we talk about boys, we say *he* and when we talk about girls, we say *she*.” Subsequently, the children were engaged in activities which were designed to demonstrate the meaning of the target and allowed for frequent modelling of the structure (focused stimulation). Modelling included a slight emphasis of the target, typically achieved by slightly increasing the loudness and/or length of the grammatical target. Each child was asked a question which would elicit the grammatical target in question. If the child made an error, the teacher would follow the protocol for the provision of feedback as seen in figure 3.

Insert figure 3 about here

Treatment (both whole class and small group sessions) consisted of play-based activities designed to incorporate the targeted grammatical items. Teachers were provided with detailed activity plans, scripts and vocabulary, and were encouraged to model and cue grammar targets systematically, in the context of naturalistic play. An example of a treatment plan can be found in Supplemental Appendix B.

The key factors that distinguished the treatment program from the general classroom program were 1) the identification of specific targets for each individual 2) the repeated modeling of grammatical targets, followed by 3) multiple opportunities for the children to produce the targets, followed by 4) feedback to the child and 5) opportunities for the child to correct him/herself.

Outside of these treatment sessions, teachers were asked to carry out their classroom program as they normally would. They were provided with a list of the items that would be targeted during the treatment. Teachers were not discouraged from modelling or reinforcing

Treatment Frequency

the targets in whatever fashion they normally would throughout the day, but neither were they specifically asked to do so.

Intensity details of the intervention can be seen in table 2.

Insert table 2 about here

Treatment Fidelity

Treatment fidelity was facilitated by a number of measures: staff training, manualised therapy activities, detailed recording of children's responses and observations of treatment sessions. At the beginning of the project, teachers, speech-language pathologists and education assistants attended an education session which provided hands-on practice of the treatment activities and techniques. Before each treatment session, clinicians and teachers were provided with detailed documentation for each activity, including scripts and materials. Throughout each session, a record was kept of the children's responses to each item.

Blinding

By necessity, the speech-language pathologists and teachers were not blind to the intervention condition they were administering. The children involved in the study saw the treatment as a regular part of their classroom activities and had no contact with children in the other treatment condition. The children were accustomed to language instruction and regular testing and therefore can be considered blind to the research process.

Five testers carried out the pre and post intervention testing. Each tester tested the same children in each testing phase to reduce the likelihood of test score changes being due to different testers. Four of the five testers were blind to the nature of the study. Post-study interviews confirmed that the blind testers had remained blind to the purpose of the testing. Of the four children the non-blinded tester tested, all were from the eight weeks group. A

Treatment Frequency

comparison of the gain scores of the children in the eight weeks group tested by the blind tester versus the children tested by the non-blind tester found no difference. The mean gain from the average of the two pre-tests to time 3 for children tested by the blind testers was 8.89 (SD = 5.42) and for the non-blind tester 10.25 (SD = 8.39); $t(11) = .35$ ($p = .729$). It appears that the scores of the non-blind tester were not affected by her knowledge of the treatment conditions.

Data Analysis

We used both a standard group comparison approach comparing the experimental groups and a case series approach, in which each child acted as his own control in a single-subject design (Logan, Hickman, Harris, & Heriza, 2008; Nickels, 2002). In each analysis, the dependent variable was performance on the Grammar Elicitation Test. For the group analysis, the measure was a child's total score, out of 30. Research question (1) was addressed by comparing the gain made between Test 1 and Test 2 (pre-treatment gain) with the gain made from Test 2 to Test 3 (post-treatment gain), for both groups. Pre-treatment gain was calculated by subtracting Test 1 scores from Test 2 scores. Post-treatment gain was calculated by subtracting Test 2 scores from Test 3 scores. A mixed analysis of variance was conducted, with one between-groups factor (treatment frequency) and one repeated measure (gain score), using SPSS Statistics (version 19) software¹. Effect size was calculated using Cohen's d , with an online Effect Size Calculator (<http://www.cognitiveflexibility.org/effecsize/>).

Research question (2) was addressed by comparing the gain made from Test 1 to Test 2 (pre-treatment) with the gain made from Test 2 to a post-test eight weeks later. For the eight weeks group, this was immediately following treatment. For the eight days group, this was approximately six weeks after treatment (as their treatment took roughly two weeks). The procedure for this analysis was identical to the first.

Treatment Frequency

Question (3) was addressed by comparing each child's gain from Test 1 to Test 2 (pre-treatment baseline) to their gain from Test 2 to Test 3 (pre-treatment to post-treatment). For each item, gain was calculated, yielding 30 data points per child. The gain scores for each child for the two different periods were then compared using Related Samples Wilcoxon Signed Rank Tests, again with SPSS Statistics Software (version 19).

Results

Baseline Equivalence

Age at time of treatment and intake language scores can be found in table 1. The two groups did not differ significantly in age ($t(29) = 1.12, p = .276$) or in total language score on intake (PLS: $t(14) = .44, p = .666$; CELF-P2: $t(9) = .11, p = .917$).² They also did not differ significantly on initial score of the Grammar Elicitation Test (results in table 3 and discussed further below) ($t(29) = 1.11, p = .276$).³ The groups had equal numbers of children targeting possessive (five in each group) and past tense (five in each group). Six children in the eight days group and three children in the eight weeks group targeted pronouns, while two children in the eight days group targeted present tense *s*.

Insert table 3 about here

Group analyses yielded a significant main effect for gain score ($F(1,29) = 10.29, p = .003$), no significant main effect for group ($F(1,29) = 1.73, p = .198$), and a significant interaction between group and gain score ($F(1,29) = 4.29, p = .047$). Post hoc testing, with a Bonferroni correction for repeated t-tests, indicated no significant difference between Test 1-Test 2 gain scores and pooled pre-test to Test 3 gain scores for the eight days group ($t(17) = 0.75, p = .465$), but a significant difference for the eight weeks group ($t(12) = 5.1, p = .000$). Thus, treatment significantly improved grammar test scores when carried out over eight

Treatment Frequency

weeks, but not when carried out over eight days. The eight weeks group showed an effect size of .02 from Test 1 to Test 2 (pre-treatment baseline) and 1.66 from average of the pre-treatment scores to post-treatment, considered a large treatment effect (Cohen, 1988). In contrast, the eight days group showed an effect size of .41 between Test 1 to Test 2 (pre-treatment baseline) and .45 between average of the pre-treatment scores to post-treatment. Thus, there was a large treatment effect in the eight weeks group when compared to the effect size over the pre-treatment baseline ($1.66 - .02 = 1.64$), whereas the eight days group showed a small effect size ($0.45 - 0.41 = .04$).

The second analysis of variance, comparing performance of the two groups after a similar period of time, yielded a significant main effect for gain score ($F(1,29) = 9.37, p = .05$), no significant main effect for group ($F(1,29) = 1.82, p = .188$), and a significant interaction between group and gain score ($F(1,29) = 4.74, p = .038$). Post hoc testing with a Bonferroni correction for repeated t-tests, indicated no significant differences in gain scores for the eight days group ($t(17) = 0.58, p = .569$), but a significant difference in gain scores for the eight weeks group ($t(12) = 5.1, p = .000$). Thus, treatment significantly improved gains on test scores when carried out over eight weeks, but not when carried out over eight days, when groups were tested after equal time intervals. The eight days group showed an effect size (Cohen's d) of .41 between Test 1 to Test 2 (pre-treatment baseline) and .38 between the average of the pre-test scores and eight-week post-test (pre-treatment to post-treatment), indicating no treatment effect ($.41 - .38 = .03$). Effect size for the eight weeks group was unchanged.

Individual Analysis

Results for each individual child can be seen in figures 4 and 5. Seven out of the 13 children (54%) in the eight-weeks group and three out of 18 children (17%) in the eight-days group showed significantly greater gain over the treatment period from Test 2 to Test 3 than

Treatment Frequency

in the pre-treatment baseline period from Test 1 to Test 2. Examination of data indicate that for one child in the eight weeks group (#13 in figure 5), the significant difference in the gain scores was due to a negative gain from Test 1 to Test 2 followed by a return to Test 1 levels in Test 3. This cannot be considered a treatment effect. Therefore, six children (46%) in the eight weeks group showed a significant treatment effect. No trends were identified in demographic or diagnostic variables that might explain the differential responses to treatment amongst the children.

Insert figures 4 and 5 about here

As noted above, five children were excluded from the analysis due to missing data for one of the testing rounds, four from the eight weeks group and one from the eight days group. These children did not discontinue participation in the study, rather they were absent from school for one of the testing periods, typically due to illness. All children completed the treatment. These five children did not differ from the main group on the basis of age (mean age of absent children: 62.6 months (SD = 6 months; mean age of whole group: 62.97 months SD= 3.77), or intake language score. The scores of three children fell within one standard deviation of the group mean on total language score, and two fell more than one standard deviation below the mean.

Discussion

The purpose of this study was to examine the effectiveness of grammar intervention at two different dose frequencies in children with SLI, while holding cumulative intervention intensity constant. We compared the effectiveness of an eight session grammar intervention program administered over eight weeks or eight days. Results indicated that the program was effective over eight weeks, but much less so over eight days. Individual results indicate that

Treatment Frequency

while about half of the children in the eight weeks group improved significantly, the majority of children in the eight days group did not. The children who improved significantly could not be distinguished from those who did not, based on their pre-treatment language scores. Thus it appears that dose frequency was a key factor in the success of the treatment program. Our data suggest that more often is not necessarily better when it comes to the frequency of treatment sessions, at least when the cumulative intervention intensity is kept constant.

Our results are not consistent with several oral language intervention studies reported in the literature. For example, Proctor-Williams and Fey's (2007) results suggested that better outcomes were related to more concentrated treatment sessions. However, their study involved five sessions spread over a range of 4 to 44 days, a much lower treatment frequency than in our study. Bellon-Harn's (2012) findings also differed from ours, as she found no difference in treatment for semantic and morphology skills administered 4 times per week for 6 weeks compared to 2 times per week for 12 weeks. Bellon-Harn administered 24 treatment sessions overall compared to our eight, suggesting that dose frequency differences may be more pronounced when fewer treatment sessions are provided. In addition, the total intervention duration for the more frequent condition in the Bellon-Harn study was substantially greater than ours (6 weeks to our 8 days). Our finding is also in contrast to that of Barratt et al. (1992) who found greater improvement in expressive scores when treatment was delivered in concentrated bursts. The participants in the Barratt et al. study were two years younger than those in our study, and in contrast to our grammar-focused program and tests, their therapy addressed a wide range of speech and language areas and outcome measures were based on a general test of language. Thus, the difference may be explained by differences in dose frequency, participants and outcome measures.

So why might a more distributed schedule have led to better outcomes in our study? Certainly, as Zeng, Law and Lindsay (2012) point out, the issue of distributed learning and

Treatment Frequency

spacing has been considered extensively in the psychological literature. Dempster (1989) explains that “the spacing effect is one of the most dependable and replicable phenomena in the learning literature” (p. 311). Thus, it appears that a general learning principle, often studied with respect to adults’ memory of repeated materials, might be at work with the teaching of grammatical principles to children with SLI.

One account for the benefits of a lower dose frequency in this study may be that the time between sessions allowed the child to practice and consolidate their learning. Even if the child is not receiving structured practice and corrective feedback, natural conversation provides opportunities to practice the grammatical target in question. This practice can be facilitated by teachers or parents informally throughout the day, in between the formal teaching sessions. These opportunities have the effect of increasing dosage in between sessions. In this study, teachers took an active part in the treatment program and learned the language treatment techniques for the sessions. This meant, of course, that they had the skills to apply the language treatment as opportunities arose throughout the day. We do not have a measure of how often teachers applied these skills outside of the sessions. Nevertheless, there is clearly more time available for practice between sessions if they are administered weekly rather than daily.

On the other hand, this opportunity for consolidation and continued practice was insufficient to allow continued improvement for the eight days group post-treatment. The test at time 4 for the eight days group coincided with the test immediately post-treatment for the eight weeks group. A test at this time was included to allow comparison of the two groups after an equal period of time had passed. Thus, this test occurred for both groups eight weeks after the test at time 2, when both groups had received eight treatment sessions. If an intensive boost of eight sessions, followed by a consolidation period facilitated by teachers now up-skilled and attentive to the grammar targets were successful, we would expect no group

Treatment Frequency

differences at this point. However, the eight weeks group showed more gain than the eight days group both immediately after treatment (time 2), and after the same period of lapsed time. This difference cannot be attributed to the fact that the eight weeks group had just completed their treatment, as the same difference was found in the comparison of the two groups immediately post-treatment. Thus, while time for consolidation and informal practice *between* sessions may have contributed to the greater success of the eight weeks program, the data suggest that the opportunity for informal practice alone is insufficient to affect change. If this were the case, the gain for the eight days group would have equaled the gain of the eight weeks group after eight weeks.

The group data for the eight days group showed no significant treatment effect. Consideration of the individual data, however, showed that three children did make significant improvement. These successful children improved substantially, one by nine points (from 1 to 10/30) and two by 17 points (6 to 23/30 and 13 to 30/30). The two most successful children maintained these gains to testing at time 4. These successful children did not differ from their peers on any of the language measures in the study. Teachers did comment, however, that these children were very engaged in the therapy activities and appeared to enjoy them. This suggests that it may not be language scores which contributed most to treatment success in this group, but rather, that social variables were important.

Indeed, while treatment frequency is an important factor to consider in treatment effectiveness (Proctor-Williams & Fey, 2007), as Kamhi (2012) points out, “the notion of fixed, ideal treatment dosage ... seems inconsistent with the flexible, individualized clinical process that characterizes evidence-based practice.” (p.3). As our results reflect, different children may respond differently to treatment frequencies. Clearly, other factors, including treatment techniques, the engagement of the child and their contribution to the outcome, clinician skill and expertise and the clinician-child relationship may also contribute to the

Treatment Frequency

ultimate success of treatment for an individual child (Baker, 2012). The most successful treatment will likely be one that considers treatment frequency as one of several important treatment factors.

While our study did not formally assess the acceptability of the interventions for stakeholders, the design of the study was, in part, dictated by the stakeholders themselves. The two LDCs were interested in addressing the question of frequency, as it related directly to their models of service delivery. Based on the results of this study, the LDC which delivered treatment programs intensively is piloting the delivery of treatment over eight weeks, thus incorporating evidence-based practice into their models of service delivery and directly measuring the outcomes.

Limitations

Due to the “real-world” nature of effectiveness compared to efficacy studies, it is hard to control all experimental variables as tightly as we might wish. One limitation of this study is that random allocation of the subjects to treatment groups was not possible in this setting. Random allocation of participants to groups is used as an attempt to reduce bias due to differing characteristics of participants, sites or clinicians. However, the groups in this study did not differ in age, on standardised test performance on intake to the LDC, or on pre-treatment Grammar Elicitation Test score. Similarly, the potential confounds introduced by the two educational settings in the study were limited, due to the fact that the locations operated under a similar model within the same department of education. Class and group sizes were similar, staff were equally experienced and received similar administrative and professional support. Thus, it is unlikely that results were affected by the lack of randomisation.

It was not possible to obtain clear audio recordings of the treatment in the noisy classroom environment with three treatment groups running concurrently. Because of this, we

Treatment Frequency

were unable to objectively measure treatment fidelity. Nevertheless, we did put in place procedures to facilitate treatment fidelity, and crucially, these measures were the same in both groups. Given that the same procedures were in place for both groups, there is little reason to believe that there were systematic differences in treatment fidelity between the groups that might have affected the results.

A variable that we were unable to control as tightly as we might have wished was dose. Indeed, we do not have a measure of cumulative intervention intensity as defined by Warren et al. (2007). Few efficacy studies, let alone effectiveness studies, report dose information, given the practical difficulties involved in accurately controlling and measuring dose, including the need for highly trained research staff able to consistently deliver a specific dose across children and activities plus clear audio recordings and transcriptions of sessions to accurately measure dose. This practicality led Gillam (2012) to define cumulative intervention intensity, for the purposes of his review, as session duration x dose frequency x total intervention duration. With this definition, we have, in the current study, held cumulative intervention intensity constant at 480 (60 minutes x once weekly x 8 weeks = 480; 60 minutes x once daily x 8 days = 480).

We did attempt to equate the dose in three ways. First, the treatment activities were the same or similar in the two groups. Second, all staff were provided with training which encouraged as high a dose as possible and third, we asked the teachers to keep records of the children's responses during the sessions in order to help them monitor and equate the dose each child received.⁴ With the exception of available time for practice in between sessions (discussed below), there is no reason to believe that dosage differed in any systematic way between the groups.

Similarly, we have no measure of dose in between treatment sessions, which may have been applied informally in the classroom. Having increased the teachers' awareness of

Treatment Frequency

grammar goals and their ability to implement grammar treatment, we could not take that knowledge away from them. This led to the possibility that they may provide teaching episodes to children between sessions. This opportunity was longer in the eight weeks group than the eight days group. However, as discussed above, the total time available for teaching opportunities was equated in the analysis which compared each group 8 weeks after the (pre-) test at time 2. If teachers were providing a sufficient dose of treatment outside of the formal treatment sessions to affect change, we would expect there to be no difference between the groups after an equal period of time had passed. This, however, was not the case. Thus, it does not appear that a difference in dose between sessions can account for the difference in success between the two groups.

The language intervention program brought about a significant change with a large clinical effect over a relatively short intervention period of eight weeks. However, it must be noted that the post-test mean was 17.62 out of 30 for the eight weeks group and 17.06 for the eight days group. This is far from mastery of the grammatical targets. In fact, the similarity of the scores makes it tempting to consider the possibility of 17/30 as ceiling performance for eight treatment sessions. It must be remembered, however, that this is a *mean* post score, representing a group of individuals with a variety of gain scores. In fact, in the eight days group, final scores post-treatment ranged from 0 to 30/30 and, in the eight-weeks group, scores ranged from 6 to 29/30. Not all children reached ceiling, of course, and it remains to be seen how many treatment sessions are required to improve all children's performance to ceiling or close.

Future Research

This study should be replicated with a large scale randomised controlled trial, to confirm that the success achieved with the small numbers in a non-randomised trial can be replicated. Future research should focus on the implications of a variety of distributed and

Treatment Frequency

concentrated frequencies of treatment, for a variety of language targets. For example, would a concentrated start to treatment followed by a more distributed follow-up program be more successful than either a distributed or concentrated program alone? Would treatment focused on areas of language other than morpho-syntax also show greater success with distributed sessions (e.g. semantics, word-finding)? What is the interaction between overall treatment duration and treatment frequency? Research should also focus on the individual's response to differing treatment frequencies in order to determine if it is possible to predict who would benefit from one frequency of treatment over another. In addition, measures of informal application of treatment techniques in the classroom should be examined to allow measurement of the opportunity for children to consolidate skills between formal treatment sessions. Finally, given the recommendations of Baker (2012), Kamhi (2012) and others, future research should also examine the role of treatment procedures themselves in effectiveness of outcomes, such as comparing different treatment techniques, the implementation of recommended techniques and how these factors might interact with treatment frequency (Smith-Lock et al., in preparation).

Conclusion

Our results suggest that a distributed frequency of treatment sessions (eight sessions over eight weeks) was more effective than a concentrated frequency of treatment sessions (eight sessions over eight days) for 5-year-olds with SLI in a school setting. The eight weeks group showed a significant, large effect size, whereas the eight days group did not. 46% of the eight weeks group made significant gains in a single subject analysis, whereas only 17% of the eight days group made significant gains. It appears, that for the majority of children, treatment distributed over eight weeks provides more successful outcomes than treatment concentrated over eight days.

References

- Accardo, P., & Capute, A. (2005). *Cognitive Adaptive Test*. Baltimore, MD: Brookes.
- Baker, E. (2012). Optimal intervention intensity. *International Journal of Speech-Language Pathology, 14*, 401-409.
- Barratt, J., Littlejohns, P., & Thompson, J. (1992). Trial of intensive compared with weekly speech therapy in preschool children. *Archives of Disease in Childhood, 67*, 106-108.
- Bellon- Harn, M. L. (2012). Dose frequency: Comparison of language outcomes in preschool children with language impariment. *Child Language Teaching and Therapy, 28*(2), 225-240.
- Brandel, J., & Loeb, D. F. (2011). Program intensity and service delivery models in the schools: SLP survey results. *Language, Speech, and Hearing Services in Schools, 42*, 461-490.
- Cirrin, F. M., Schooling, T. L., Nelson, N. W., Diehl, S. F., Flynn, P. F., Staskowski, M., Torrey, T.Z. and Adamczyk, D. F. (2010). Evidence-based systematic review: effects of different serivice delivery models on communication outcomes for elementary school-age children. *Language, Speech, and Hearing Services in Schools, 41*, 233-264.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.
- Connell, P., & Stone, C. (1992). Morpheme learning of children with specific language impairment under controlled instructional conditions. *Journal of Speech and Hearing Research, 35*, 844-852.
- Dempster, F.N. (1989). Spacing effects and their implications for theory and practice. *Educational Psychology Review, 1*, 309-330.

Treatment Frequency

- Frankenburg, W., Dodds, J., & Archer, P. (1992). The Denver II: A major revision and restandardization of the Denver Developmental Screening Test. *Pediatrics*, *89*, 91-97.
- Gillam, R. (2012, November). *How much is enough? Dosage in child language intervention*. Paper presented at the American Speech-Language-Hearing Association Annual Convention, Atlanta, GA.
- Griffiths, R. (1970). *The abilities of young children*. Great Britain: Cournswood House.
- Kamhi, A. G. (2012). Pharmacological dosage concepts: How useful are they for educators and speech-language pathologists? *International Journal of Speech-Language Pathology*, *14*, 414-418.
- Law, J., Garrett, Z., & Nye, C. (2004). The efficacy for treatment for children with developmental speech and language delay/ disorder: A meta analysis. *Journal of Speech, Language, and Hearing Research*, *47*, 924-943.
- Leonard, L. B. (1998). Specific language impairment and grammatical morphology: A Discriminant Function Analysis. *Journal of Speech, Language, and Hearing Research*, *41*, 1185-1192.
- Leonard, L. B., Camarata, S., Pawlowska, M., Brown, B., & Camarata, M. N. (2006). Tense and agreement morphemes in the speech of children with specific language impairment during intervention: Phase II. *Journal of Speech, Language, and Hearing Research*, *49*, 749-770.
- Logan, L. R., Hickman, R. R., Harris, S. r., & Heriza, C. B. (2008). Single-subject research design: Recommendations for levels of evidence and quality rating. *Developmental Medicine and Child Neurology*, *50*, 99-103.
- McGinty, A. S., Breit-Smith, A., Fan, X., Justice, L. M., & Kaderavek, J. N. (2011). Does intensity matter? Preschoolers' print knowledge development within a classroom-based intervention. *Early Childhood Research Quarterly*, *26*, 255-267.
- Smith-Lock *et al.* (in press) *International Journal of Speech-Language Pathology*.

Treatment Frequency

- Nickels, L. A. (2002). Theoretical and methodological issues in the cognitive neuropsychology of spoken word production. *Aphasiology, 16*, 3-19.
- Proctor-Williams, K., & Fey, M. E. (2007). Recast density and acquisition of novel irregular past tense verbs. *Journal of Speech, Language, and Hearing Research, 50*, 1029-1047.
- Rice, M. L., Wexler, K., & Cleave, P. L. (1995). Specific language impairment as a period of extended optional infinitive. *Journal of Speech and Hearing Research, 38*, 850-863.
- Riches, N. G., Tomasello, M., & Conti-Ramsden, G. (2005). Verb learning in children with SLI: Frequency and spacing effects. *Journal of Speech, Language, and Hearing Research, 48*, 1397-1411.
- Robey, R. R., & Schultz, M. C. (1998). A model for conducting clinical-outcome research: An adaptation of the standard protocol for use in aphasiology. *Aphasiology, 12*, 789-810.
- Schooling, T. L., Venediktov, R., & Leech, H. (2010). Evidence-based systematic review: Effects of service delivery on the speech and language skills of children from birth to 5 years of age. *National Center for Evidence-Based Practice in Communication Disorders*, American Speech-Language-Hearing Association: Rockville, MD.
Retrieved from: <http://www.asha.org/uploadedFiles/EBSR-Service-Delivery.pdf>
- Semel, E., Wiig, E., & Secord, W. (1992). *Clinical Evaluation of Language Fundamentals-Preschool*. San Antonio, TX: The Psychological Corporation, Harcourt Brace Jovanovich.
- Smith-Lock, K. M., Leitão, S., Lambert, L., & Nickels, L. (in press). Effective intervention for expressive grammar in children with specific language impairment. *International Journal of Language and Communication Disorders*.

Treatment Frequency

- Smith-Lock, K. M., Leitão, S., Prior, P., Lambert, L. & Nickels, L. (in preparation). *A comparison of grammar treatment techniques for children with Specific Language Impairment*.
- Ukrainetz, T. A., Ross, C. L., & Harm, H. M. (2009). An investigation of treatment scheduling for phonemic awareness with kindergartners who are at risk for reading difficulties. *Language, Speech, and Hearing Services in Schools, 40*, 86-100.
- Warren, S. R., 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*, 70-77.
- Wechsler, D. (2002). *Wechsler Preschool and Primary Scale of Intelligence-III (WPPSI-III)*: Pearson Assessments.
- Weismer, S. E., & Murray-Branch, J. (1989). Modeling versus modeling plus evoked production training: A comparison of two language intervention methods. *Journal of Speech and Hearing Disorders, 54*, 269-281.
- Wiig, E., Secord, W., & Semel, E. (2006). *Clinical Evaluation of Language Fundamentals Preschool - Second Edition, Australian and New Zealand Standardised Edition (CELF P-2 Australian and New Zealand)*. Sydney, NSW: Pearson.
- Yoder, P. J., Fey, M. E., & Warren, S. F. (2012). Studying the impact of intensity is important but complicated. *International Journal of Speech-Language Pathology, 14*(5), 410-413.
- Yoder, P. J., Woynaroski, T., Fey, M. E., & Warren, S. (2012). *Is more better? If so, for whom?* Paper presented at the American Speech-Language-Hearing Association Convention, Atlanta.
- Zeng, B., Law, J. and Lindsay, G. (2012). Characterizing optimal intervention intensity: The relationship between dosage and effect size in interventions for children with

Treatment Frequency

developmental speech and language difficulties. *International Journal of Speech-Language Pathology*, 14, 471-477.

Zimmerman, I., Steiner, V., & Pond, R. (2002). *Preschool Language Scale, Fourth Edition (PLS-4) Australian Language Adapted*. Sydney, NSW: Pearson.

Footnotes

¹ Analyses conducted with pre-treatment gain calculated as the difference between Test 3 and an average of Test 1 and Test 2 yielded identical results, as did an analysis using each child's highest pre-treatment test score.

² As seen in table 1, four children in the 8 weeks group were tested using the CELF-P rather than the CELF-P2. As no children in the 8 days group received this test, we have no t-test comparison for these children.

³ Non-parametric tests yielded identical results, as did a comparison of the mean of test 1 and test 2 scores and of test 2 scores.

⁴ While written records of the sessions were kept by the teachers, observation of the treatment sessions suggested that the records were not sufficiently accurate for research purposes. The teachers were focused on the activity, the treatment techniques and behaviour management, as well as completing the online record. While the teachers insist it was a very valuable tool for them to monitor dose and the children's responses, we were not sufficiently confident of their accuracy to use them to calculate dose.

Acknowledgements

This work was funded by the Australian Research Council Centre of Excellence for Cognition and its Disorders (CE110001021) <http://www.ccd.edu.au>. We would like to thank research assistants Sarah Arnott, Joanne Dixon, Caroline Dixon, Shannon Golding, Rachel Reid, Lilian Sklarz, and Catherine Wilkinson. We are very grateful to the staff and students at Fremantle and West Coast Language Development Centres for welcoming us so wholeheartedly into their classrooms. Correspondence concerning this article should be addressed to Dr. Karen Smith-Lock, ARC Centre of Excellence in Cognition and its Disorders, Department of Cognitive Science, Macquarie University, North Ryde NSW 2109, Australia. Email: karen.smith-lock@mq.edu.au; telephone: +61 29850 9599 fax: +61 2 9850 6067.

Treatment Frequency

Supplementary Appendix A. Grammar Elicitation Test

Target: Possessive /s/

Sample Stimulus:

[child is shown a picture of a cat with a ball]

Tester says: “This is a cat. The cat has a ball. Whose ball is it?”

Correct response: “the cat’s ball”

Incorrect responses: “the cat,” “the cat ball”

Prompts: “I think the ball belongs to the cat. Whose ball is it?”

Test items:

Possessive [s]: Hope’s, Jack’s, Pat’s, Brett’s, Blake’s, Kate’s, Pip’s, cat’s, duck’s, sheep’s

Possessive [z]: Doug’s, bee’s, May’s, Carl’s, dog’s, Em’s, Mary’s, Bob’s, boy’s, man’s

Possessive [əz]: Joyce’s, church’s, Josh’s, Grace’s, Blanche’s, Mitch’s, Rich’s, Trish’s,
horse’s, fish’s

Target: Past Tense /t/

Sample Stimulus:

[child is shown a picture of a boy kicking a ball]

Tester says: “This happened yesterday. What did the boy do yesterday?”

Correct response: “the boy kicked the ball”

Incorrect responses: “kick,” “the boy kick ball,” “the boy is kicking the ball”

Prompts: If a verb other than *kick* was used:

Tester says: “Did he kick the ball?”

Child: (if) yes

Tester: “You tell me that. What did the boy do yesterday?”

Child: (if) no

Tester: “I think he did kick the ball. You tell me that. What did the boy do yesterday?”

Treatment Frequency

Test items:

Past [t]: hopped, skipped, touched, danced, walked, shopped, dropped, licked, jumped

Past [d]: squeezed, paddled, smiled, crawled, combed, cried, stirred, buzzed, watered, climbed

Past [əd]: pointed, ended, needed, twisted, added, folded, counted, landed, painted, melted

Target: present tense /s/

Sample Stimulus:

[child is shown a picture of a boy walking]

Tester says: “This boy is walking He does that every day. What does he do every day?”

Correct response: “the boy walks”

Incorrect responses: “walk,” “the boy walk,” “the boy is walking”

Prompts: If a verb other than *walk* was used:

Tester says: “Did he walk every day?”

Child: (if) yes

Tester: “You tell me that. What does the boy do?”

Child: (if) no

Tester: “I think he does walk. You tell me that. What does the boy do every day?”

Test items:

Present [s]: cuts, sits, walks, kicks, counts, coughs, skips, jumps, laughs, picks

Present [z]: wears, climbs, reads, smiles, needs, opens, drives, plays, cries, runs

Present [əz]: squashes, brushes, hisses, washes, squeezes, matches, touches, watches, kisses, freezes

Target: he/she

Sample Stimulus:

Treatment Frequency

[child is shown a picture of a girl dancing]

Tester says: “Tell me about this picture. What’s happening?”

Correct response: “She is dancing.”

Incorrect responses: “the girl is dancing,” “her dance,” “he is dancing”

Prompts: If the child used “the girl”, the tester replied “why?” The child typically then answered with the pronoun, as the referent was previously established.

Treatment Frequency

Supplementary Appendix B. Sample Treatment Plan.

Goal 1: Whole group activity with speech-language pathologist (SLP): To identify items that belong to a pirate and introduce and expose to the possessive ('s sounding [s], [z] and [əz]).

Procedure: To check students can identify the items that belong to a pirate. Place the items in the centre of the circle. SLP points to each item and children take a turn to label it. Model using one item that is put on the pirate, the pirate's beard. "The beard belongs to the pirate. Whose beard is it? It is the pirate's beard". Explicitly teach "We add a possessive 's to the end of the word, pirate, to show that the beard belongs to the pirate". Repeat and emphasise. "Can you hear, the pirate's beard". Children each have an item. Each child gets a turn to name 4 student names using the possessive 's for their item as they put it on the pirate eg. "The hook belongs to Nick". SLP: "Whose hook is it? It is Nick's hook". 3 students will be involved in the items for each dressing of the pirate. After each dressing students will move items to the left so each student has a different item. Dressing of the pirate will be repeated 4X for all students to practise. Point out when the ending sounds /z/ and /əz/. "Sometimes instead of /s/ the possessive sounds like /z/ or /əz/" (For reference: /s/ after k, t, p or /z/ after g, l, m, n, vowel e.g., *bees* or /əz/ after sh, ch, s, x). Errors are recast on-line.

Targets: Model: the pirate's beard. Student targets, one each: _____'s earring, eye patch, shirt, pants, hat, sword/cutlass, hook, map, gold, spade, telescope, binoculars (12)

Assessment by the SLP: Anecdotal notes taken of students who had difficulty using the possessive *s* on a class list and also students who had difficulty listening and following instructions

Task explanation: Sit in a circle. You will each have a turn to name these items that belong to the pirate. Point and say "What is this?" "Yes, it is the pirate's beard". "When something belongs to the pirate, we add a possessive 's to the word pirate and say, *the pirate's beard*."

Treatment Frequency

Now we will dress up _____ as a pirate. Here is an item each that belongs to you.

Each of you will get 4 turns to name a friend and what belongs to them e.g., I want you to say: “The hook belongs to Nick, it’s Nick’s hook”. Then Nick will go up and dress the pirate.

This time _____ will be the pirate. Pass your item to the next person on your left.”

Motivation: A pirate sticker

Materials: Class list for assessment, pirate stickers, 12 targets - earring, eye patch, shirt, pants, hat, sword/cutlass, hook, map, gold, spade, telescope, binoculars and a model - beard

Time: 15 minutes

Treatment Frequency

Table 1.

Age and language standard scores on intake to Language Development Centre (LDC).

Mean scores, standard deviation in brackets.

Group	Age (months)	Total Language Standard Score		
		PLS ¹	CELF-P2 ³	CELF-P ²
eight weeks	62.08 (4.3)	81.67 (9.33) n = 6	81.00 (10.54) n = 3	75.0 (7.12) n = 4
eight days	63.61 (3.55)	84.20 (12.0) n = 10	80.13 (12.47) n = 8	

¹ Preschool Language Scale (Zimmerman et al., 2002).

² Clinical Evaluation of Language Fundamentals-Preschool (Semel et al., 1992)

³ Clinical Evaluation of Language Fundamentals - Preschool 2 (Wiig, Secord, & Semel, 2006).

Treatment Frequency

Table 2. Intensity Details

Group	Session Duration	Dose Frequency	Total Intervention Duration	Cumulative Intervention Intensity
8 Days	60 minutes	Once daily	8 days	480
8 Weeks	60 minutes	Once weekly	8 weeks	480

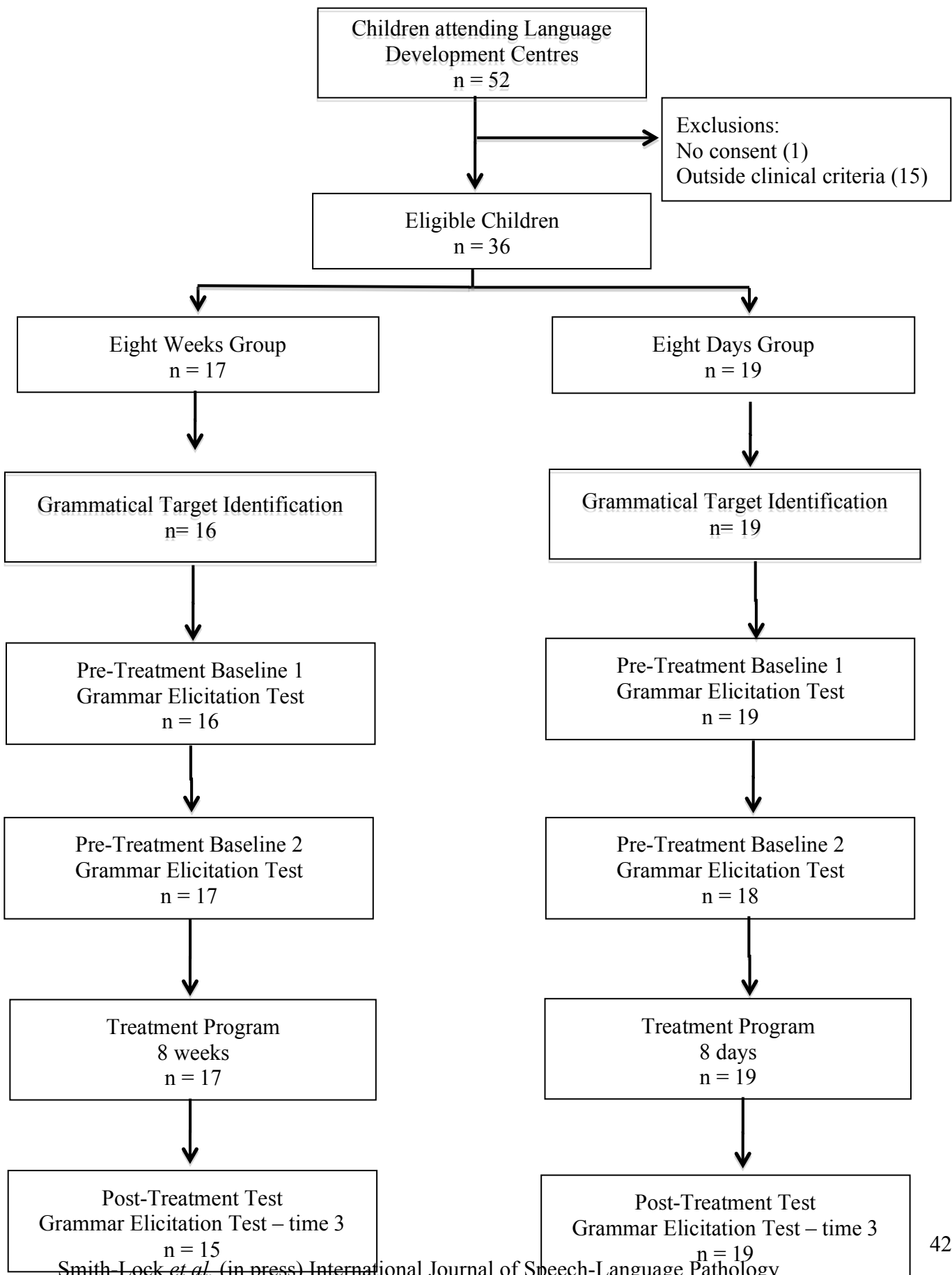
Treatment Frequency

Table 3.

Mean score (standard deviation in brackets) on Grammar Elicitation Test (Smith-Lock et al., in press).

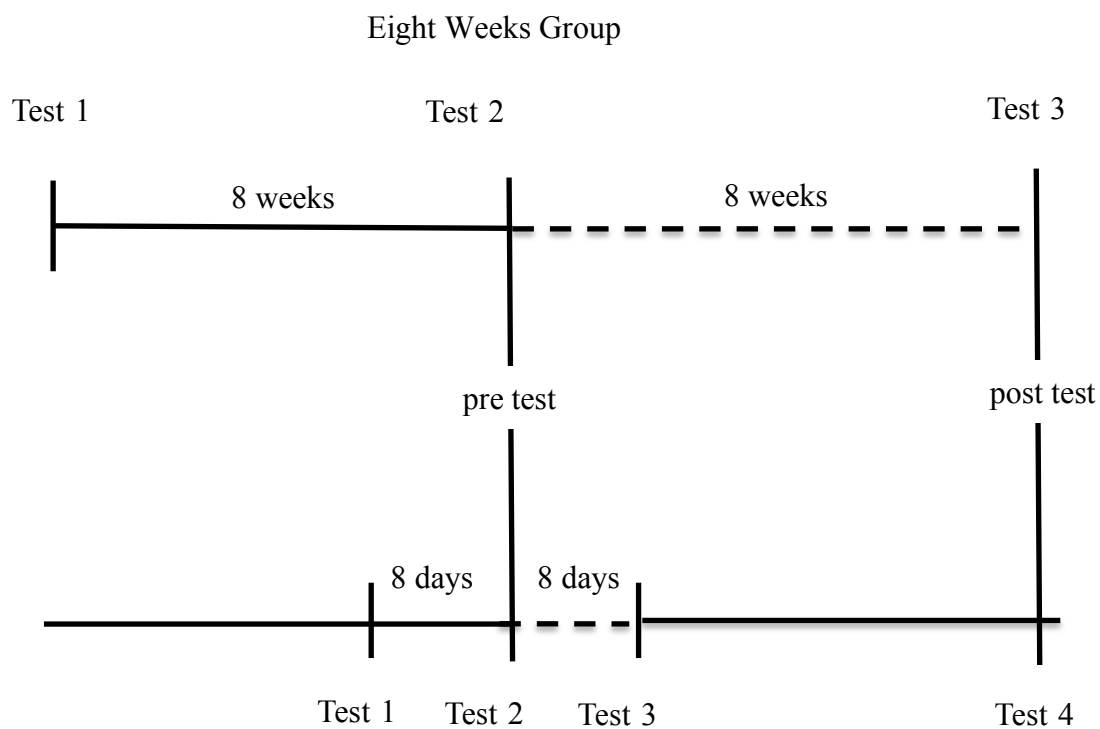
Group	Test 1	Test 2	Test 3 Immediate Post-test	8-weeks Post -Test	Gain Test 1 to Test 2	Gain Mean of Test 1 and 2 to Test 3	Gain Mean of Test 1 and Test 2 to 8-weeks Post-Test
							= Gain Mean of Test 1 and Test 2 to Test 3
8 weeks	8.38 (7.0)	8.23 (6.02)	17.62 (7.16)	= Test 3	-0.15 (3.74)	9.31 (6.14)	
8 days	11.33 (7.52)	13.17 (9.27)	17.06 (9.44)	16.61 (11.08)	1.83 (4.79)	4.81 (7.71)	4.35 (8.62)

Figure 1. Participant Flow.



Treatment Frequency

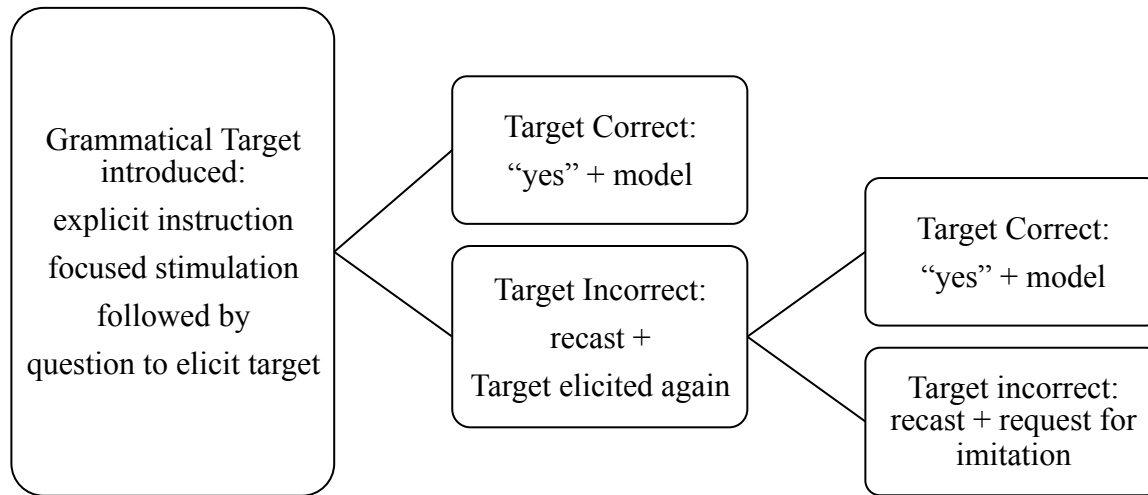
Figure 2. Timeline.



Eight Days Group

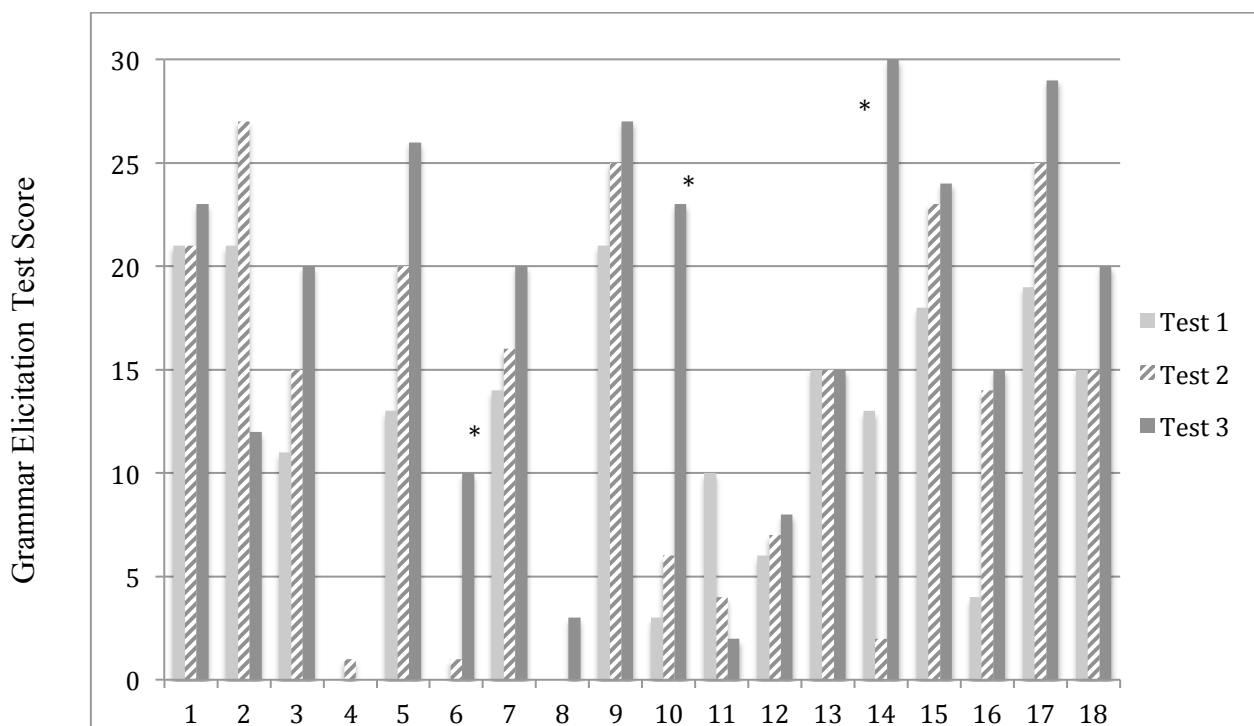
Treatment period - - - -

Figure 3. Treatment Technique.



Treatment Frequency

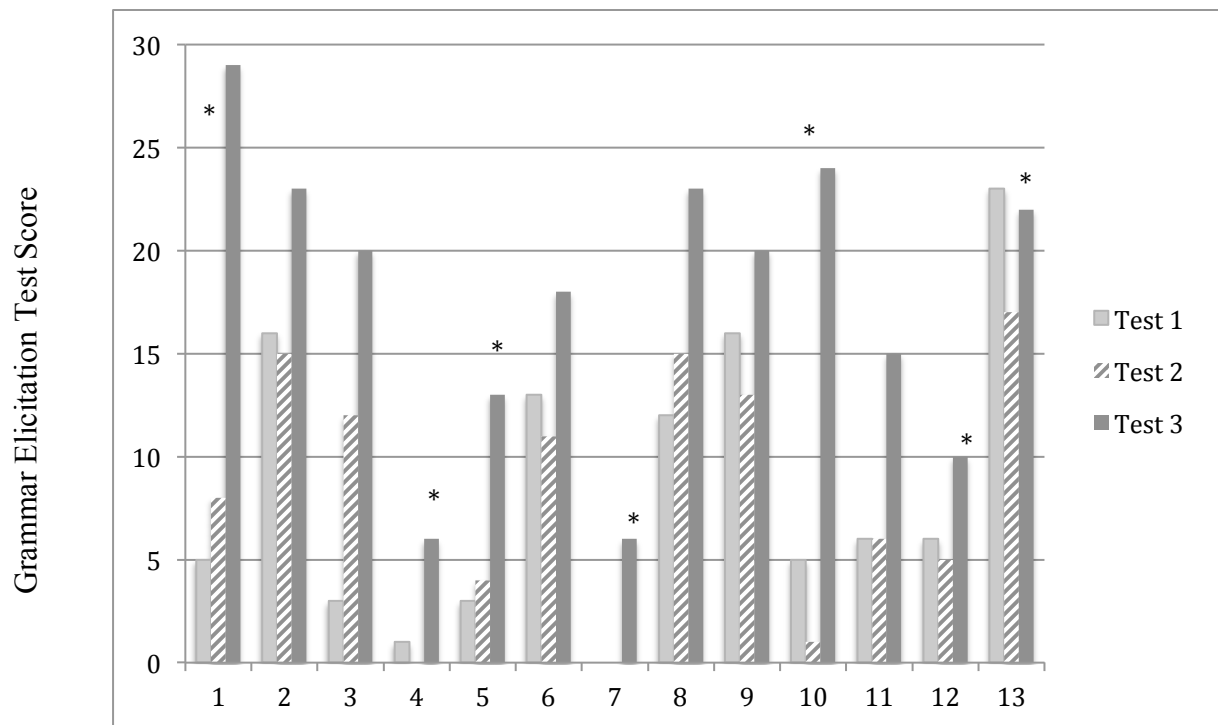
Figure 4. Individual Participant Scores for Grammar Elicitation Test: 8 Days Group.



* significant difference in gain from Test 1 to Test 2 (pre-treatment baseline) versus Test 2 to Test 3 (pre- to post-treatment)

Treatment Frequency

Figure 5. Individual Participant Scores for Grammar Elicitation Test: 8 Weeks Group.



* significant difference in gain from Test 1 to Test 2 (pre-treatment baseline) versus Test 2 to Test 3 (pre- to post-treatment)