Verb Learning in Children With SLI: Frequency and Spacing Effects

N. G. Riches
University of Manchester, Manchester, United Kingdom

M. Tomasello
Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

Gina Conti-Ramsden
University of Manchester, Manchester, United Kingdom

Purpose: This study explored the effect of frequency (number of presentations), and spacing (period between presentations) on verb learning in children with specific language impairment (SLI). Children learn words more efficiently when presentations are frequent and appropriately spaced, and this study investigated whether children with SLI likewise benefit. Given that these children demonstrate greater frequency dependence and rapid forgetting of recently acquired words, an investigation of frequency and spacing in this population is especially warranted.

Method: Twenty-four children with SLI (mean age 5;6 [years;months]) and 24 language-matched control children (mean age 3;4) were taught novel verbs during play sessions. In a repeated measures design, 4 experimental conditions combined frequency (12 or 18 presentations) and spacing (all presentations in 1 session, or spread over 4 days). Comprehension and production probes were administered after the final session and 1 week later.

Results: Although the children with SLI benefited significantly from frequent and widely spaced presentations, there were no significant effect in the control group. The language-impaired children showed rapid forgetting.

Conclusions: The frequency and spacing of presentations crucially affect the verb learning of children with SLI. A training regimen characterized by appropriately spaced intervals and moderate repetition will optimally benefit lexical learning.

KEY WORDS: verb learning, specific language impairment, frequency effects, spacing effects

Comparatively few word learning studies have investigated the effect of frequency and spacing (i.e., the number of presentations and the intervals between them). The dearth of studies in this area is surprising given that recent research has demonstrated that spacing and frequency crucially affect word learning performance. Moreover, one particular group of children—those with specific language impairment (SLI)—may be particularly sensitive to the effect of frequency and spacing. This study aimed to investigate the influence of these two factors on the word learning of these children.

Frequency

Many researchers have argued that word learning is an automatic process characterized by extremely rapid learning. The term fast-mapping (Carey & Bartlett, 1978) has been used to describe how young children learn new words after a minimal number of exposures, or perhaps after even one exposure. In addition to being quickly learned, fast-mapped words are also retained for long periods of time (e.g., Carey & Bartlett, 1978; Markson & Bloom, 1997; Woodward, Markman, & Fitzsimmons, 1994).
Although a large number of experiments have demonstrated fast-mapping processes, there is also evidence to suggest that there are limits to children's fast-mapping abilities. Most words are not learned in an all-or-none fashion after only a single presentation. Some studies have indicated that lexical learning is frequency dependent (Childers & Tomasello, 2002; Gray, 2003, 2004; Rice, Oetting, Marquis, Bode, & Pae, 1994; Storkel, 2001). All of these studies have controlled the frequency of presentations, that is, how many presentations the child hears, and have observed that word learning performance follows a trajectory that is characterized by improved performance after increasing numbers of exposures. There is also evidence for frequency dependence during early lexical acquisition. Children's early words are likely to be frequent in the speech of the caregiver (Barrett, Harris, & Chasin, 1991; Hart, 1991; Huttenlocher, Haigh, Bryk, Seltzer, & Lyons, 1991). Although such frequency dependence appears to fade during the 2nd year, the studies listed above have demonstrated that the word-learning performance of older children is sensitive to the number of presentations.

Although frequency effects appear to be inconsistent with the concept of fast-mapping, which proposes that children can learn a word from a single presentation, there is in fact no contradiction. Fast-mapping studies (e.g., Carey & Bartlett, 1978) typically use forced-choice comprehension tasks to demonstrate that children rapidly develop minimal lexical representations. The studies referred to above merely demonstrate that despite this rapid initial learning, children still need to receive a sufficient number of presentations in order to establish sufficiently robust representations. There is a crucial distinction between learning enough about a word to perform successfully in a comprehension task a few days later and developing a representation that is sufficiently strong or detailed for the child to use that word spontaneously in conversation. The results of word learning studies referred to above suggest that input frequency is an important factor that, at least in part, determines the strength of lexical representations.

Although input frequency may subtly affect the word learning performance of typically developing children, it may be a far more crucial factor for children with SLI, whose language difficulties cannot be accounted for in terms of factors normally associated with language impairment such as hearing loss, low IQ, or global cognitive disorders (see Leonard, 1998, for a summary of the diagnostic criteria). There is evidence to suggest that these children's word learning is particularly sensitive to input frequency (Gray, 2003, 2004; Rice et al., 1994; Windfuhr, Faragher, & Conti-Ramsden, 2002).

Rice et al. (1994) used a naturalistic-learning paradigm in which the target words (verbs and nouns) were embedded in a video (known as the Quick Incidental Learning paradigm). The age-matched normal language (NL) group showed significant learning after only 3 presentations, as assessed by the difference in performance on a forced-choice comprehension pretest and an identical posttest. In contrast, the children with SLI (mean age 5:1 [years;months]) needed 10 presentations in order to demonstrate a learning effect. Windfuhr et al. (2002) modeled novel nouns and verbs in a series of play sessions. The number of modeling episodes per session remained constant (40/session), and performance, as measured by the children's spontaneous productions, was plotted over the four sessions. The children with SLI (mean age 5:0) displayed a more concave learning trajectory, with poor performance in the initial sessions and a significant improvement in the final sessions. This profile contrasts with the convex profile of the language-matched NL children (mean age 2:10) who showed rapid gains in the initial sessions followed by stable performance.

Finally, Gray (2003) modeled novel nouns over a 4-day training period and conducted frequent comprehension and production probes. In both comprehension and production, the children with SLI (mean age 4:6) needed about twice as many presentations to learn the novel nouns to criterion—defined as three correct responses from four comprehension/production probes during at least two training sessions—than an age-matched NL group. A follow-up study with a similar experimental design (Gray, 2004) likewise found that children with SLI needed significantly more trials to reach criterion. It is interesting that Gray (2003) found that fast-mapping abilities did not differ across the two groups, suggesting that although the children with SLI may not be slow to establish an initial representation, their difficulties may lie in establishing a robust representation over time. By contrast, Gray (2004) found poorer fast-mapping in the SLI group.

These studies contrast with the findings of a number of earlier studies that found that children with SLI performed as well as, or even better than, language-matched NL peers (Camarata & Schwartz, 1985; Leonard, 1982; Leonard & Schwartz, 1985). However, these early studies used a supportive experimenter-modeling paradigm, in which the experimenter directly labels and object or action, and did not attempt to control for input frequency. Converging evidence from a variety of recent studies indicates that children with SLI perform poorly on word learning tasks that take place in a naturalistic context, such as the Quick Incidental Learning paradigm (Oetting, Rice, & Swank, 1995; Rice, Buhr, & Nemeth, 1990; Rice,
Spacing

In addition to frequency, there is evidence to suggest that the spacing of presentations, that is, the interval between them, is a crucial factor in word learning. Both Schwartz and Terrell (1983) and Childers and Tomasello (2002) have observed that “spaced” training regimes with long intervals between learning episodes were more effective than “massed” training regimes, in which the same number of learning episodes were concentrated in a short period of time. Schwartz and Terrell presented novel nouns and verbs to children between 12 and 15 months old. Some children were given two presentations per training session, whereas others were given only one. Although after 10 training sessions the children who heard two presentations per session were more likely to recall the novel words, the single-presentation regimen was in some ways more efficient. The children spontaneously produced nearly twice as many words after 10 sessions of the single-presentation regimen than after five sessions of the dual-presentation regimen, despite having received the same number of exposures in each condition. Moreover, in the single-presentation regimen the children needed fewer presentations prior to recall, as assessed by their spontaneous productions ($M_s = 7.6$ vs. $12.8$ for the dual-presentation regimen). The difference between the two experimental conditions may be conceptualized in terms of the spacing of the presentations. For example, in the single-presentation condition, all presentations were separated by approximately 24 hr, whereas in the dual-presentation condition half the presentations occurred after very short intervals (i.e., within the same training session). Therefore, the relative benefits of the single-presentation regimen appear to be related to the consistently longer intervals between presentations.

Childers and Tomasello (2002) investigated distributed word learning in 2-year-olds using different conditions that varied both the number of presentations and the intervals between learning sessions. They discovered that, for both novel verbs and novel nouns, four daily presentations resulted in superior word learning (in production, but not comprehension), than four or even eight presentations given on the same day. In this way, the spacing effect was stronger than the effect of number of presentations. Childers and Tomasello also observed that the number of different days on which a word was presented was a key influence on recall probability; for example, one presentation per day over 4 consecutive days is more effective than two presentations per day on 2 consecutive days, which in turn are more effective than four presentations on 1 day.

The two studies described above were motivated by a large body of research that demonstrates that distributed learning, characterized by long intervals between learning episodes, is more effective than massed learning (see Dempster, 1988, and Baddeley, 1997, for reviews). According to Bruce and Bahrick (1992), over the past 100 years, more than 300 studies have investigated the benefits of distributed learning. Spacing effects have been observed in a variety of different cognitive domains, ranging from motor learning (Baddeley & Longman, 1978) to the acquisition of spelling and multiplication tables (Rea & Modigliani, 1985). The existence of a spacing effect across different cognitive domains suggests that it exploits fundamental domain-general learning mechanisms. Another interesting characteristic of distributed learning is that, in addition to boosting initial learning, it leads to longer retention. This effect has been noted in studies of the spacing effect in adult second-language vocabulary learning (Bahrick, Bahrick, Bahrick, & Bahrick, 1993; Bahrick & Phelps, 1987), adult name learning (Landauer & Bjork, 1978), and infant learning of physical movements (Hartshorn, Wilk, Muller, & Rovee-Collier, 1998). It is interesting to note that Bahrick and Phelps (1987) found evidence of improved retention over an 8-year period.

There is reason to believe that distributed training may be particularly beneficial for children with SLI. These children appear to be particularly poor at retaining newly acquired words. Rice et al. (1994) noted that although children with SLI showed significant gains from pretest to posttest for novel verbs, these gains were lost after only 1 to 3 days, and there was a significant rate of forgetting, as assessed via analyses of the differences between the posttest and a retention test. The age-matched group, by comparison, showed good retention, with no significant decline from the posttest to the retention test. Oetting (1999) also observed poor retention of verbs in children with SLI (aged 6;5) who did not score above chance on retention probes conducted on 3 consecutive days. In contrast, the NL children (aged 6;3 and 4;8) showed a significant learning effect. Given that children with SLI present with poor retention of recently acquired words, the potential benefits of distributed training should be fully investigated.

Furthermore, there is evidence to suggest that distributed training is particularly beneficial for poor learners, such as adults with dementia (Camp, Foss, O’Hanlon, & Stevens, 1996) and children with learning
disabilities (Gettinger, Bryant, & Mayne, 1982). Even among nonimpaired adults, it has been observed that poorer learners tend to benefit most from distributed practice, as Cull, Shaughnessey, and Zechmeister (1996) observed in a name–face matching task. If this is the case, then distributed training may be an important method of addressing the retention difficulties of children with SLI.

**Verb Learning**

There is some evidence to suggest that the frequency dependence and poor retention noted in children with SLI may be particularly characteristic of verb learning. Windfuhr et al. (2002) noted that although the verb learning of children with SLI showed a steady rate of increase over four sessions, performance on nouns peaked during the third session. There was a significant interaction among word type (verb or noun), session, and group, indicating that the effect of word type on learning trajectory was significantly different for the children with SLI (mean age: 5;0). It should be noted that Rice et al. (1994) observed a steeper learning trajectory for verbs, but they argued that the children may have found the stories that contained the verbs more interesting than the stories that contained the nouns. In addition to displaying a high degree of frequency dependence, verbs may also be more difficult to retain. The retention difficulties observed by Rice et al. (1994) were evident for verb learning and not noun learning. Oetting (1999) also observed poor retention for recently acquired verbs, but noun retention was not assessed. In the light of the research cited above, this study focused explicitly on verb learning.

**Main Hypotheses**

Given the above findings, this study addressed the following hypotheses related to frequency and spacing.

1. The verb learning performance of the children with SLI will show greater frequency dependence than that of the NL children.

2. Given the powerful domain-general nature of distributed learning, the verb learning performance of both groups will benefit from distributed training. However, a greater distributed training effect will be observed in the SLI group, given that distributed training may be particularly beneficial for poor learners.

3. The children with SLI will show poorer retention of the newly acquired verbs than the NL children.

4. Distributed training will improve retention in both groups, with comparatively greater benefits for the SLI group.

**Method**

**Participants**

Twenty-four children with SLI, with a mean age of 5;6, were recruited from language units attached to mainstream primary schools in northwest England. Language units are specialized classrooms where the educational focus is on the development of language skills. All of the children scored below –1 SD on their expressive and/or receptive language scores, as assessed by the Clinical Evaluation of Language Fundamentals (CELF–Preschool; Wiig, Secord, & Semel, 1992; see Table 1 for details). In contrast, all of the children’s performance IQs were above –1 SD, as measured by the Weschler Preschool and Primary Scale of Intelligence—Revised (Weschler, 1990). According to teacher reports, none of the children with SLI had hearing difficulties or social–emotional difficulties, and all spoke English as their first language.

Twenty-four control children, aged 3;0 through 4;0 (M = 3;5), were also recruited from nursery schools. These children had CELF language scores, and Weschler Preschool and Primary Scale of Intelligence—Revised Performance IQ scores, within the normal range (see Table 1 for details). They were likewise screened for hearing; social and emotional difficulties; and bilingualism, which was assessed via teacher report. Both groups were matched according to mean length of utterance in words, which was calculated from a 10-min transcription of a play session between the child and the experimenter. The two groups were also matched on a talkativeness measure, which was calculated by comparing the number of utterances the children made during the 10-min period. We used this measure because we felt that the children’s talkativeness might be an important confounding factor in an experiment where expressive language was being examined. In addition, two tests were conducted to assess the children’s lexical knowledge: (a) the British Picture Vocabulary Scale (Dunn & Whetton, 1982), a test of receptive vocabulary, and (b) the Expressive Vocabulary Test (EVT; Williams, 1997), a test of expressive vocabulary. The descriptive data are shown in Table 1.

One child with SLI and 2 NL children were excluded from the study because of unforeseen absences, mainly illness related, which occurred randomly across
the participants. The descriptive statistics in Table 1, and analyses in the Results section, are therefore based on the data from 23 children with SLI and 22 NL children.

The results of a series of $t$ tests investigating between-group differences in CELF language age, CELF expressive and receptive raw scores, British Picture Vocabulary Scale and EVT raw scores, the mean length of utterance in words, and the talkativeness measures are shown in Table 1. There were no significant differences between the two groups on the language measures, with the exception of the EVT, on which the SLI group demonstrated significantly better performance ($p = .000$). Therefore, the children with SLI appeared to be at a more advanced expressive vocabulary stage than the NL control children.

With regard to performance IQ, there is a notable difference between the group means, with the control children performing better than the children with SLI. However, it should be emphasized that performance IQs of all the children lay within the normal range.

### Table 1. Results of the psychometric tests.

<table>
<thead>
<tr>
<th>Measure</th>
<th>SLI group ($n = 23$)</th>
<th>NL group ($n = 22$)</th>
<th>Results of $t$ tests comparing the groups on raw scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>$M = 5;6$</td>
<td>$M = 3;5$</td>
<td></td>
</tr>
<tr>
<td>$SD = 6.18$ months</td>
<td>$SD = 3.40$ months</td>
<td></td>
<td></td>
</tr>
<tr>
<td>range = 4;7–6;4</td>
<td>range = 3;1–4;0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CELF Language Age</td>
<td>3;4</td>
<td>3;5</td>
<td></td>
</tr>
<tr>
<td>5.39 months</td>
<td>4.70 months</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2;6–4;4</td>
<td>2;9–4;3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PIQ</td>
<td>100</td>
<td>111</td>
<td>$p = .211$</td>
</tr>
<tr>
<td>12.1</td>
<td>7.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>85–129</td>
<td>94–120</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CELF Expressive raw scores</td>
<td>38.9</td>
<td>40.7</td>
<td></td>
</tr>
<tr>
<td>15.8</td>
<td>15.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9–60</td>
<td>25–71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CELF Receptive raw scores</td>
<td>34.4</td>
<td>37.7</td>
<td>$p = .697$</td>
</tr>
<tr>
<td>11.0</td>
<td>6.38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9–51</td>
<td>26–52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLUw</td>
<td>2.81</td>
<td>3.01</td>
<td>$p = .288$</td>
</tr>
<tr>
<td>0.509</td>
<td>0.560</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.76–3.84</td>
<td>1.85–4.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Talkativeness</td>
<td>97.3</td>
<td>86.3</td>
<td>$p = .190$</td>
</tr>
<tr>
<td>28.2</td>
<td>17.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20–165</td>
<td>40–112</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EVT raw scores</td>
<td>45.7</td>
<td>38.7 **</td>
<td>$p = .000$</td>
</tr>
<tr>
<td>5.68</td>
<td>5.57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>35–63</td>
<td>28–50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BPVS raw scores</td>
<td>41.4</td>
<td>37.5</td>
<td>$p = .190$</td>
</tr>
<tr>
<td>10.7</td>
<td>9.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29–73</td>
<td>21–59</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. SLI = specific language impairment; NL = normal language; CELF = Clinical Evaluation of Language Fundamentals; PIQ = Performance IQ; MLUw = mean length of utterance in words; EVT = Expressive Vocabulary Test; BPVS = British Picture Vocabulary Scale.

**$p < .01$.**

### Materials

The children were taught four novel verbs, shown in Table 2, that were modeled using sets of three objects. One of these objects was used to perform the target verb, while the other two objects were used to perform the foil actions. Foils were necessary as the comprehension task required the children to select the target object from the training set. The verb labels
were all simple CVC constructions, and they were all phonotactically highly probable, as measured by phoneme co-occurrence tables (Kessler & Tremain, 1997).

**Experimental Design**

This study had a $2 \times 2$ experimental design combining 2 training regimes (massed and spaced), with two different numbers of presentations (12 or 18). This created four experimental conditions: (a) Massed 12, with 12 exposures on a single day, (b) Massed 18, with 18 exposures on a single day, (c) Spaced 12, with 12 exposures spread over 4 days (3 per day), and (c) Spaced 18, with 18 exposures spread over 4 days (4,5,4,5).

**Experimental Procedure**

The children were taught the novel verbs in play sessions. At the start of the play sessions, the objects were concealed in bags. The experimenter began the play sessions with the “warm-up” objects. She or he produced the objects in a random order and commented on them. When the cuddly toys were produced, the experimenter merely made comments such as “It’s a giraffe, isn’t it? ... Look, he’s blue.... What’s his name?”, and so on. When the bouncy ball was produced, the experimenter bounced the ball saying “Look, it’s bouncing. See, it bounces.” This was designed to familiarize the child with the target structure that was used to teach the novel verbs—Look, it’s VERBing! See, it VERBs! If the child showed an interest in playing with the object, the experimenter gave the object to the child. When the child had finished playing with/manipulating the object, the experimenter indicated a plastic bucket and asked the child “Can you put it in the bucket?”

A similar game was repeated with the training set. Although the experimenter performed actions with the nontarget objects, he or she did not label the actions but merely commented on some aspect of the objects. The child was then invited to put the object into the bucket. When it came to the turn of the target object, the experimenter acted on the object while producing the novel verb label, for example, “Look, it’s dack! See, it dacks!” This was counted as one presentation of the target verb. The dual morphological frame, alternating between the -ing form and the third-person form, was used to maximize the syntactic cues that would allow the child to identify the novel word as a verb. The intransitive frame was chosen because it is arguably the most simple verb frame, in that it has only one argument and one thematic role, and this would minimize the syntactic demands of the word learning task. However, it is acknowledged that currently little is understood about how such inflectional and argument structure properties may affect verb learning, and it has also been observed that children learn verbs more rapidly when inflections remain constant (Bedore & Leonard, 2000). After a number of presentations had been made, the experimenter invited the child to put the object into the bucket. The experimenter continued playing with the toys until the appropriate number of presentations of the target verbs had been made. The sessions lasted between 2 and 10 min, depending on the number of presentations, which ranged between 3 and 18 depending on the condition (massed vs. spaced and 12 vs. 18 presentations).

The play format provided a degree of flexibility that allowed the experimenter to engage and motivate the child. Consequently, a number of variables, such as the length of the training session, or the number of child turns (i.e., opportunities for the child to

<table>
<thead>
<tr>
<th>Label</th>
<th>Target object, plus action</th>
<th>Foil 1</th>
<th>Foil 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dack</strong></td>
<td>Orange rubber ring</td>
<td>E (experimenter) bounced the object on the floor.</td>
<td>A mat made of wooden slats tied with string</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>E held the mat at either end and rattled it.</td>
</tr>
<tr>
<td><strong>Tam</strong></td>
<td>A rattle on a string</td>
<td>E held onto the string and spun the rattle.</td>
<td>A spinning top</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Meek</strong></td>
<td>A green plastic skittle</td>
<td>E flipped the skittle up into the air.</td>
<td>A plastic box that opens and closes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Gorp</strong></td>
<td>A small wooden rolling pin</td>
<td>E rolled the pin down his or her shins.</td>
<td>A bendy object similar to a large pipe cleaner</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Experimental props.
play with the toys), were not rigorously controlled. The grouping of the target utterances was also open to variation—for example, one target utterance, followed by play with the other toys, followed by four target utterances, and so on. However, despite the flexibility of the play format the number of presentations of the target verb were rigorously counted using a tick-sheet. The experimenter was also careful to avoid providing verb labels for the nontarget objects. This is because if the children associate the nontarget objects with a familiar verb this might help them to respond correctly during the comprehension probe by means of a process of elimination and not because they have learned the novel verb.

**Testing Procedure**

There were two tests: (a) a posttest, which took place immediately after the final presentation of the novel verb, and (b) a retention test, which took place a week after the final training episode.

The warm-up objects were presented one by one in a random order. The child was asked questions about the cuddly toys, for example, “What’s this?” and “What color is it?” When the ball was produced, the child was first asked “What does it do? Can you show me? What does it do?” Then the experimenter bounced the ball and asked “What’s it doing? Can you tell me? What’s it doing?”

A similar procedure was used with the target objects, which were taken from the bags in a random order. First the experimenter asked “What does it do? Can you show me? What does it do?” for each object in turn and handed the object to the child. This was the action probe. Then, using the same order, the experimenter asked “What’s it doing? Can you tell me? What’s it doing?” while performing the appropriate action with the object. This was the production probe. Finally, all three objects were laid out in front of the child, and the experimenter asked “Which one was VERB-ing? Can you tell me? Which one was VERB-ing?” (VERB refers to the particular novel verb being used). This was the comprehension probe.

**Schedules**

Each child was trained on all four novel verbs in one of four different schedules. Although the order of verbs remained constant (dack, tam, meek, gorp), the order of the conditions (Spaced 12, Spaced 18, Massed 12, Massed 18) was counterbalanced to account for potential ordering effects. The schedules were carefully constructed so that children would be trained and tested on no more than one verb on any given day.

**Results**

**Statistical Methods**

The experimental data were coded in a binary fashion: 1 if the child produced/comprehended the action or word and 0 if he or she failed. The logistic regression procedure was considered to be the most appropriate statistical method, as it has been designed to deal with such binary, or dichotomous, data. Logistic regression may be compared to a simple linear regression; however, whereas a line of best fit may be easily calculated for a data set consisting of continuous data, a binary data set does not lend itself to the line-of-best-fit method. The 1s and 0s in the binary data set are therefore transformed into values along a continuous scale by calculating the likelihood of an event—in this case, the likelihood of a child comprehending/producing the target verb, which is coded as 1 in the data set. In logistic regression, likelihood is expressed in terms of odds, that is, the probability of an event happening divided by the probability of the same event not happening (probability/1 – probability). A plot of odds values tends to be S shaped, so a log function is applied to transform the S shape into a straight line. Once a straight line has been plotted, the data may be analyzed using principles identical to those used in linear regressions; that is, parameters may be estimated, confidence intervals may be plotted, and probability values may be calculated.

Although it would have been possible to analyze the data using a more widely known technique, such as analysis of variance, by conducting a frequency count of the 0 and 1 responses, the logistic regression is preferred for two important reasons. First, given the binary/dichotomous nature of the data set, there is an increased chance of ceiling or floor effects. This may result in a non-normal distribution, which violates an important assumption of analysis of variance. Second, frequency counts collapse across means, which may reduce the statistical power of an analysis based on such counts. Another benefit of regression techniques is that the slightly different group sizes used in the current study (23 SLI versus 22 NL) do not greatly affect the analysis.

The logistic regression expresses a relationship between an independent and dependent variable in terms of the odds ratio (OR). This shows the change in the odds of a particular event (i.e., a correct response by the child) with each change in the level of the independent variable. An OR > 1 indicates that the odds of a correct response increase with each increase in the independent variable. Conversely, an OR < 1 shows an inverse relationship, with a decrease in odds for each increase in the independent variable.
The independent variables for the analyses of the present study are as follows: frequency (0 = 12 presentations, 1 = 18 presentations), spacing (0 = massed conditions, 1 = spaced conditions), and time (0 = posttest, 1 = retention test).

Overall Verb Learning

First, analyses were run with experimental conditions collapsed to test whether the two groups showed a similar overall rate of word learning. The mean numbers of correct responses for each group and modality with all experimental conditions collapsed are shown in Table 3. There do not appear to be any substantial differences between the two groups. To test this observation, logistic regressions were conducted using group as the independent variable (coded 0 for the NL group and 1 for the SLI group) and the number of correct responses as the dependent variable. Analyses were conducted in each modality: the expressive modality, which refers to the children’s ability to produce the target verbs, and the receptive modality, which refers to the children’s ability to comprehend the target verbs. These analyses effectively investigate whether the children’s rate of correct responses differs significantly across the two groups. The effect of group was not significant in either modality, OR = 0.734 (95% CI: 0.385, 1.40), p = .348, for comprehension responses, and OR = 1.07 (95% CI: 0.596, 1.93), p = .820, for production.

Within-Group Analyses

Frequency effects. Further analyses were conducted to assess the effect of frequency and spacing within each group. The mean numbers of correct responses for each group in each frequency condition for each modality (time collapsed) are shown in Figure 1. Logistic regressions were conducted to investigate the effect of frequency for each Group x Modality combination. A significant effect of frequency was observed in the SLI group in comprehension (p < .01). All the analyses of main effects are reported in Table 4. An OR > 1 indicates that the probability of a correct response increases as the frequency changes from 12 presentations (coded as 0) to 18 presentations (coded as 1). Results from the other analyses did not approach significance (p > .1).

Spacing effects. Another set of analyses assessed the importance of spacing. The mean numbers of novel verbs produced or comprehended for each Spacing x Modality combination are shown in Figure 2. Logistic regressions were conducted to investigate the effect of spacing on each Group x Modality combination. Significant effects were observed for production in the SLI group (p < .01), and a trend toward a significant effect was observed for comprehension in the SLI group (p < .1). The ORs > 1 indicate that that the children’s performance was better in the spaced conditions.

Analysis of retention. The children’s responses over time are shown in Figure 3. The forgetting gradient appears to be much steeper for the children with SLI, especially in production. To test this observation, we conducted analyses to investigate the effect of time (0 = posttest, 1 = retention test) on the children’s responses in each Group x Modality combination. By using time as the independent variable, these analyses effectively investigate whether the children’s rate of correct responses differs significantly between the posttest and the retention test. A significant effect of time was observed for production in the SLI group (p < .01), with the OR < 1 indicating a decrease in the number of correct responses.
Summary of within-group analyses. The within-group analyses suggest interesting differences between the two groups in terms of frequency dependence, the benefits of distributed training, and retention abilities. Significant effects of frequency and spacing were observed only in the SLI group. Furthermore, analyses of retention revealed a significant rate of forgetting only in the SLI group. All of these findings are consistent with the main experimental hypotheses. However, the finding of significant effects in one group but not another does not necessarily indicate that the effect of a variable differs significantly across the two groups. To test for significant differences across groups, we conducted a series of interaction analyses to investigate the interaction between the grouping variable (group) and the main variables of interest.

Interaction Effects

Analyses investigating the interaction between Group × Frequency, Group × Spacing, and Group × Time were conducted in each modality. There was a significant Group × Time interaction for the children’s production responses, OR = 0.432 (95% CI: 0.229, 0.812), p = .009. Therefore, in production, the rate of forgetting differed significantly across the two groups. None of the other interaction analyses reached significance, although there was a trend toward a significant effect for the Group × Frequency interaction in comprehension, OR = 2.29 (95% CI: 0.940, 5.59), p = .068.

The results of the interaction analyses indicate that only retention differs significantly across the two groups. The interesting findings in relation to frequency and spacing observed in the analysis of the main effects have not been validated by the more stringent interaction analyses.

A Comparison of Frequency and Spacing

This study’s 2 × 2 mixed design conflates both frequency and spacing. Although this design allows one to simultaneously investigate two different variables, it likely complicates the interpretation of the results. For example, the analyses conducted thus far suggest that the spacing effect may be stronger than the effect of frequency (as evidenced by the differences in the means shown in Figures 1 and 2 and the lower significance values for the analyses of training regimen). It is therefore plausible that the spaced training regimen generates a ceiling effect that reduces the variance in the data and therefore masks the effect of frequency.

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Dependent variable</th>
<th>Odds ratio and confidence intervals</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency(a)</td>
<td>Production responses, SLI group</td>
<td>1.26 (0.735, 2.15)</td>
<td>.403</td>
</tr>
<tr>
<td></td>
<td>Comprehension responses, SLI group</td>
<td>1.83 (1.23, 2.73)</td>
<td>.003**</td>
</tr>
<tr>
<td></td>
<td>Production responses, NL group</td>
<td>0.952 (0.597, 1.52)</td>
<td>.840</td>
</tr>
<tr>
<td></td>
<td>Comprehension responses, NL group</td>
<td>0.799 (0.355, 1.80)</td>
<td>.588</td>
</tr>
<tr>
<td>Spacing(b)</td>
<td>Production responses, SLI group</td>
<td>4.02 (1.76, 9.24)</td>
<td>.001**</td>
</tr>
<tr>
<td></td>
<td>Comprehension responses, SLI group</td>
<td>2.36 (0.90, 6.20)</td>
<td>.08</td>
</tr>
<tr>
<td></td>
<td>Production responses, NL group</td>
<td>2.1 (0.78, 5.65)</td>
<td>.14</td>
</tr>
<tr>
<td></td>
<td>Comprehension responses, NL group</td>
<td>1.99 (0.85, 4.63)</td>
<td>.11</td>
</tr>
<tr>
<td>Time(c)</td>
<td>Production responses, SLI group</td>
<td>0.453 (0.293, 0.701)</td>
<td>.000**</td>
</tr>
<tr>
<td></td>
<td>Comprehension responses, SLI group</td>
<td>0.787 (0.470, 1.32)</td>
<td>.08</td>
</tr>
<tr>
<td></td>
<td>Production responses, NL group</td>
<td>1.05 (0.657, 1.68)</td>
<td>.839</td>
</tr>
<tr>
<td></td>
<td>Comprehension responses, NL group</td>
<td>0.429 (0.185, 0.990)</td>
<td>.11</td>
</tr>
</tbody>
</table>

\(a0 = 12\) presentations, \(1 = 18\) presentations. \(b0 =\) massed conditions, \(1 =\) spaced conditions. \(c0 =\) posttest, \(1 =\) retention test.

**\(p < .01\).
To assess the relative benefits of frequency and spacing, the means have been plotted for each Group × Spacing × Frequency × Modality combination (see Figure 4). A visual interpretation of these data suggests that the spacing effect is stronger than the effect of frequency. Furthermore, there appears to be an interesting difference between two groups in terms of the effect of frequency for performance in the massed conditions only, with the SLI group showing a marked effect of frequency, compared with the NL children, who do not appear to benefit from an increased number of presentations. To investigate whether the effect of frequency differs significantly across the two groups, we conducted an analysis to investigate the Group × Frequency interaction (massed conditions only). There was a significant effect in comprehension, OR = 2.44 (95% CI: 1.06, 12.9), p = .04.

**Discussion**

This study identified interesting differences between the word learning performance of the children with SLI compared to the children with NL. The children with SLI showed a significant effect of frequency (for comprehension responses) and a significant rate of forgetting (for production responses) compared with the NL group, whose response rate was not significantly affected by any of the experimental variables. The SLI children also benefited significantly from distributed training (for production responses), whereas there was no significant spacing effect in the NL group. These findings need to be considered within the context of the finding that there were no statistically significant group interaction effects, with the exception of the interaction between group and time. However, it is interesting to note that the results from the analyses of main effects are consistent with previous findings in the literature.

It is interesting to note that with regard to expressive vocabulary (as measured using the EVT), the children with SLI in this study appeared to come to the task with better skills than the younger NL children. This is not surprising given that, on average, the children with SLI were 2 years older. As argued earlier, within this context the difficulties of the children with SLI are emphasized. In other words, despite arguably better expressive vocabulary skills, the children with SLI are more frequency dependent and have poorer retention than younger NL children.

**Frequency**

The children with SLI demonstrated a significant effect of frequency for comprehension responses, compared with the NL group, which did not demonstrate significant frequency dependence. A plot of the mean number of correct responses for each frequency condition (see Figure 1) is consistent with the observation that only the children with SLI benefited from greater numbers of presentations. An analysis of comprehension performance in the massed conditions yielded a significant Group × Frequency interaction, although it should be noted that this analysis is exploratory and not wholly motivated by the main hypotheses.

To our knowledge, only two studies—the current study and Windfuhr et al.’s (2002) study—have indicated that children with SLI may be more frequency dependent than language-matched peers. This suggests a more severe deficit than that which has been noted by Rice et al. (1994) and Gray (2003, 2004), who observed differences between children with SLI and their age peers.

Little is currently known about why children with SLI demonstrate such frequency dependence. One interesting idea was put forward by Leonard (1998), who suggested that SLI may be regarded as a “type of filter such that some but not all of the experiences of a word are registered in semantic memory” (p. 47). According to this account, children process only a subpart of the relevant semantic information on each presentation, and therefore more exposures will be needed to develop a robust representation. Such a proposal is consistent with claims that children with SLI have a
limited processing capacity (Ellis Weismer & Hesketh, 1996; Montgomery, 2000). Word learning may be particularly affected in a limited capacity system because it involves a broad range of cognitive tasks. For example, children must segment the word, analyze its phonological form, designate the appropriate grammatical class, and identify the appropriate meaning/referent.

Conversely, a deficit in specific processing mechanisms may also act as a filter. Research has suggested that children with SLI have poor phonological short-term memories (Bishop, North, & Donlan, 1996; Conti-Ramsden, Botting, & Faragher, 2001; Gathercole & Baddeley, 1990a), and this component of short-term memory may be implicated in lexical learning (Gathercole & Baddeley, 1990b; Gathercole, Willis, Emslie, & Baddeley, 1992). Thus, the filter metaphor could be extended to phonological aspects of word learning. For example, Conti-Ramsden (2003) argued that phonological processing difficulties crucially affect children’s ability to develop fine-grained phonological representations, which in turn will reduce the likelihood that children with SLI will recognize two separate tokens of the same word within different constructions, for example, “Let me play” and “I like to play.” Therefore, for these children, useful exemplars of a word (i.e., exemplars that can be recognized as instances of the same word) may be more widely spaced in time. In the light of these difficulties, frequent exemplars may be particularly beneficial for word learning.

The finding of greater frequency dependence in the massed conditions is open to a number of different interpretations. First, the effect could be artifactual. It may be the case that the relatively large benefits of distributed training mask the effect of frequency in the spaced conditions. Therefore, the greater frequency dependence of the children with SLI is evident only when distributed training effects are minimal. Another possibility is that frequency effects are indeed greater in the massed conditions because the particular combination of frequency and spacing in this condition is particularly beneficial for the children with SLI.

It is interesting that there is evidence to suggest that multiple, closely spaced exemplars may facilitate learning. Giraltommeto, Pearce, and Weitzman (1996) observed that this type of training appears to benefit toddlers (mean age: 2;4) with expressive language delays. They taught caregivers a technique called focused stimulation, which involved providing frequent exemplars of both words and syntactic constructions within a narrow time frame, and observed that children in the experimental group showed significant improvements, compared with a control group, in both lexical and...
learning and syntactic learning, as measured by the parental report on the MacArthur Communicative Development Inventory (Fenson et al., 1993). It has also been observed that repetitive games and routines may play an important role in early language learning (Bruner, 1983). It is interesting to note that both the children with SLI in the current study, and the younger expressive-delayed children observed by Giralometto et al., benefited from multiple, closely spaced exemplars, and this pattern may therefore reflect a particular characteristic of language-impaired children in general. For example, they may have poor attentional resources and may therefore require more learning episodes within a single session. However, although it is interesting to reflect on such parallels, it is also important to note that focused stimulation involves factors besides sheer numbers of presentations. In particular, it attempts to create a communicative context in which the caregiver is responsive to the child’s focus of attention. Although the open-ended play sessions used in this study allowed for a sensitive labeling style, such a labeling style is more rigorously systematized by the technique of focused stimulation. Perhaps this sensitive labeling style is a more important factor in the success of focused stimulation than the frequency of the presentations.

Spacing

This study provides further support for the benefits of distributed learning. Both groups of children, on average, performed better in the distributed condition, and there were significant effects of production in the SLI group, production at the posttest in the SLI group, and comprehension at the retention test in the SLI group. It is interesting that the spacing effect was greater and more significant than the effect of the number of presentations. For example, performance after only 12 presentations in the distributed (i.e., widely spaced) training regimen was better than performance after 18 presentations in the massed training regimen.

On a theoretical level, there are some trends in the data that may contribute to our understanding of why the spacing effect happens. It has been argued that the spacing effect may be related to attention-based factors (Baddeley & Longman, 1978). For example, if the number of presentations remains fixed, then a distributed training regimen will involve shorter sessions than a massed training regimen, and therefore these shorter sessions will be less likely to tax the children’s concentration. However, the improved performance of the children with SLI after 18 presentations in the massed conditions suggests that, at least in this group, the length of the training sessions is not an important factor, and the spacing effect is therefore dependent on factors that are not related to the length of the sessions and the children’s concentration spans. This claim is supported by a number of studies that have found a spacing effect even when there are only one or two presentations per session (Childers & Tomasello, 2002; Schwartz & Terrell, 1983).

Another interesting observation is that a significant spacing effect was observed only in the SLI group, although it should be noted that the Group × Spacing interaction did not reach significance. We argued in the beginning of this article that distributed learning may be particularly beneficial for poor learners, and this claim is consistent with the current data. Bahrick and Phelps (1987) proposed an account of the spacing effect that can explain why this might be the case. They argued that the optimal interval between learning episodes is “likely to be the longest interval that avoids retrieval failures” (p. 349). According to this view, optimal relearning takes place on the cusp of forgetting—in other words, where the retrieval difficulty is greatest. It follows that individuals who form weak representations, or have retention difficulties, will forget sooner than good learners, and therefore for these individuals the optimal interval between learning episodes will be shorter. In the context of this study, it is possible that the interval between training sessions in the spaced conditions (approximately 24 hr) benefited the children with SLI to a greater extent because by this stage they were closer to forgetting than the NL children. Therefore, there may be a direct relationship between the retention difficulties of the children with SLI and the observation that they benefited more from the spaced training regimes. Further research is needed to identify factors that are implicated in the strength of the spacing effect.

Frequency Versus Spacing

The finding of both frequency and spacing effects raises an interesting paradox: Why do the children with SLI benefit from two completely different kinds of training regimens, one in which the exemplars are frequent and closely spaced and another in which the exemplars are widely spaced? It is plausible that training regimens based on intensive repetition (e.g., focused stimulation) and distributed training regimens exploit fundamentally separate learning mechanisms. One possibility is that focused stimulation may facilitate the development of an initial representation, whereas distributed practice enables children to consolidate a pre-existing but weakly elaborated representation. Therefore, a combination of the two kinds of training regimen—for example, an initial frequency-based training session followed by a succession of widely spaced booster sessions—may provide the optimal learning environment. As noted above, there is some evidence to suggest that frequency may be a particularly important factor for children with SLI, and therefore the potential benefits of frequency-based training should be fully explored.
However, it should also be noted that in this study the spacing effect appeared to be stronger than the effect of frequency, in terms of its significance and its effect on the mean response rate. Moreover, although there is evidence to suggest that distributed training is beneficial across different cognitive domains, different populations, and even different species, the weight of experimental evidence in favor of frequency-based learning is less substantial.

**Retention**

The findings for poor retention in the SLI group were the most statistically robust, with a significant interaction for the Group × Time interaction. The data support Rice et al.’s (1994) observation that children with SLI show poor retention of recently acquired verbs. Verb retention in this group did not benefit from distributed training, which appeared to have little impact on the rate of forgetting, as assessed by the gradient of the forgetting slope in Figure 3, which demonstrates a particularly sharp fall for spaced production in the SLI group. However, distributed learning improved retention in the sense that, according to the means, the retention test was better for verbs that had been presented during the spaced training regimen (see Figure 3), and therefore the beneficial effects of distributed learning were evident a week later.

Although there was a steep forgetting gradient for spaced production in the SLI group, it is nonetheless interesting to note that studies investigating the long-term course of forgetting have observed that the rate of forgetting tends to decline after a sharp initial drop, a curve that has been described as “non-monotonic” (Bahrick & Phelps, 1987; Ebbinghaus, 1913). If this is the case, then the initial boost provided by a distributed training regimen may greatly improve the retention period despite a fast initial rate of forgetting. However, it should be noted that little is currently known about the long-term time course of forgetting; more research is needed in this area.

**Implications for the Lexical Learning of Children With SLI**

The findings of this study must be considered in light of some limitations. First, it is evident that the groups were not closely matched according to performance IQ. Nonetheless, it is difficult to account for the main findings of the study, for example, that the children with SLI show poorer retention in terms of differences in performance IQ. The children with SLI all had performance IQs at or above –1 SD from the population mean, and therefore it is not straightforward to argue that their subtle verb learning difficulties may be attributed to cognitive factors. However, it should be acknowledged that overall scores from IQ tests provide only a rough measure of general cognitive abilities, and there may be interesting qualitative differences between the cognitive functioning of children with SLI versus normally developing children that cannot be identified through current standardized tests.

A second difficulty is that, because each verb label was linked with a single object it is plausible that the children merely associated the labels with the objects, as opposed to developing a generalized representation of the verb meaning. Therefore, it is not clear that the training methodology reliably simulated verb learning as opposed to noun learning. Nonetheless, the study identified interesting differences between the two groups in terms of retention and frequency dependence. It may be desirable for future studies to explore the relationship between verb learning and noun learning in a more principled fashion.

It is important to highlight that the results of this study are consistent with previous word learning studies that suggest that children with SLI demonstrate a greater degree of frequency dependence than their typically developing peers. An increased number of experiences with a new word might therefore greatly improve their word learning performance. However, another characteristic of the word learning environment—the spacing of representations—proved to be an even more powerful facilitator of word learning in these children.

Future research will need to investigate further the role of frequency and spacing in the word learning of children with SLI, determine what constitutes appropriately spaced intervals for these children, and explore possible individual differences within this group. In addition, it would be important to ascertain whether current models of therapeutic practice provide a sufficient number of presentations of new lexical items and whether they make systematic use of distributed learning opportunities.

**Acknowledgments**

We thank the Economic and Social Research Council for their 3-year PhD grant (No. R42201343599, awarded to N. G. Riches), and Mike Tomasello at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, for his funding to Gina Conti-Ramsden and his continued support and advice. In addition, we acknowledge the helpful and insightful comments made by the reviewers on earlier versions of this article. We also thank the parents; guardians; speech and language therapists; teachers; careworkers; and, above all, the children who participated in this project. Finally, we thank Kirsten Windfuhr for her help in collecting some of the initial data.
References


Received May 14, 2004
Revision received January 5, 2005
Accepted March 29, 2005
DOI: 10.1044/1092-4388(2005/097)
Contact author: Gina Conti-Ramsden, Human Communication and Deafness, School of Psychological Sciences, University of Manchester, Humanities Building, Oxford Road, Manchester M13 9PL, United Kingdom.
E-mail: gina.conti-ramsden@manchester.ac.uk