

Word learning by children with developmental language disorder: Identifying gaps in our understanding of spaced retrieval effects

Autism & Developmental Language
Impairments
Volume 9: 1–15
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/23969415241275940
journals.sagepub.com/home/dli



Laurence B. Leonard  and Patricia Deevy

Department of Speech, Language, and Hearing Sciences, Purdue University, West Lafayette, IN, USA

Justin B. Kueser

Center for Childhood Deafness, Language and Learning, Boys Town National Research Hospital, Omaha, NE, USA

Abstract

Background and aims: Current evidence shows that children with developmental language disorder (DLD) benefit from spaced retrieval during word learning activities. Word recall is quite good relative to recall with alternative word learning procedures. However, recall on an absolute basis can be improved further; many studies report that fewer than two-thirds of the words are learned, even with the assistance of spaced retrieval during the learning activities. In this article we identify details of spaced retrieval that are less well understood in an effort to promote more effective learning through retrieval practice.

Main contribution: We discuss the importance of factors such as: (a) integrating immediate retrieval with spaced retrieval trials; (b) determining whether gradual increases in spacing have more than short-term benefits relative to equal spacing; (c) discovering the number of successful retrievals sufficient to ensure later recall; (d) using spaced retrieval to avoid erosion of phonetic details on later recall tests; and (e) whether the well-documented difficulties with learning word forms might be tied to a particular subgroup of children with DLD. We also speculate on some of the possible reasons why spaced retrieval is beneficial in the first place.

Conclusions: Although many children with DLD make gains in word learning through procedures that incorporate spaced retrieval, there are numerous details involved in the process that can alter its success. Until we have a better understanding of the boundaries of spaced retrieval's effectiveness, we will not be taking full advantage of this promising addition to word learning procedures.

Implications: Spaced retrieval activities can be an important addition to the resources that clinicians and educators have available to assist children in their word learning. With a deeper understanding of the issues discussed here, we should be able to put spaced retrieval to even greater use.

Keywords

Developmental language disorder, specific language impairment, retrieval practice, word learning, word recall

Corresponding author:

Laurence B. Leonard, Department of Speech, Language, and Hearing Sciences, Purdue University, 715 Clinic Drive, Lyles-Porter Hall, West Lafayette, IN 47907, USA.

Email: xdxl@purdue.edu



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access page (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

Introduction

Individuals with developmental language disorder (DLD) experience a longstanding difficulty learning and using language (Bishop et al., 2017; Dubois et al., 2020; Leonard, 2014; Norbury et al., 2016; Tomblin, 2014). Although the language symptoms change with age and experience, word learning remains a consistent challenge (e.g., McGregor, Arbisi, et al., 2017; McGregor et al., 2021; Storkel et al., 2017). Both the number of different words acquired and the degree to which these words are truly known distinguish individuals with DLD from their peers (McGregor, Oleson, et al., 2013). Vocabulary expands with age, but each new word is hard won. In fact, children with DLD appear to fall further behind their peers from childhood into adolescence (Rice & Hoffman, 2015). Studies of novel word learning reveal that children with DLD require more encounters with a word to reach the same level of learning as their same-age peers (e.g., Alt, 2011; Gray, 2003; see reviews in Jackson et al., 2019; Kan & Windsor, 2010).

Spaced Retrieval and Word Learning in DLD

Many studies have been devoted to developing and evaluating methods of assisting the word learning of individuals with DLD. In recent years, some of these efforts have included the use of “retrieval practice” during the learning period (e.g., Adlof et al., 2021; Gordon et al., 2024; see review in Gordon, 2020). Retrieval practice involves an individual’s attempts throughout a learning period to recall material that had been studied. The crucial insight behind retrieval practice is that attempts are not merely an assessment of how much was just retained; they actually improve learning. The benefits of retrieval practice have been reported in the scientific literature for more than a century beginning with Abbott (1909a, 1909b) with a resurgence in this line of inquiry in the past 25 years (see reviews in Adesope et al., 2017; Fazio & Marsh, 2019; Karpicke, 2017; Latimier et al., 2021; Rowland, 2014).

Most studies of retrieval practice have involved adult participants. Benefits of retrieval with this population have included more effective learning of English translations of Japanese words (Kang et al., 2014) and Swahili words (Pye & Rawson, 2009), course material in statistics (Lyle & Crawford, 2011), special courses in immunology (Dobson, 2012), procedures for the treatment of emergency neurological complications (Larsen et al., 2009), and written material on mechanical devices such as braking systems (McDaniel et al., 2009).

Studies of retrieval practice with children have been concentrated on the elementary and middle school ages. Benefits have been reported for learning material about U.S. history (Carpenter et al., 2009), learning locations on

a map (Rohrer et al., 2010), and learning new vocabulary words (Goossens et al., 2014). Fewer studies have focused on younger children, though this literature is emerging (see review in Fazio & Marsh, 2019). One such study found that preschoolers recalled the novel names of toys better through retrieval practice (Fritz et al., 2007). The application of retrieval practice to assist the word learning of individuals with DLD, then, has a sound empirical basis.

Two details affect the efficacy of retrieval practice: the frequency of retrieval attempts and the spacing of these attempts. When retrieval attempts are repeated throughout learning, retention is better (Karpicke & Roediger, 2008; Roediger & Karpicke, 2006). These gains are even greater when spacing is used between retrieval attempts, such as through the insertion of additional material between the time the material of interest was last studied and the time it had to be retrieved (Karpicke & Roediger, 2007). Relative to immediate retrieval, retrieval with spacing is less successful during early retrieval attempts. However, retention in the long term is superior, constituting a “desirable difficulty” (e.g., Latimier et al., 2021). The participants’ age and the material to be learned usually dictate the degree of spacing that permits gradual learning while avoiding excessive forgetting (Vlach, 2019).

Much remains to be learned about the conditions under which repeated spaced retrieval (hereafter, “spaced retrieval”) is effective, especially for word learning by individuals with DLD. In this paper, we single out several aspects of spaced retrieval for closer consideration. We then identify gaps in our understanding that need to be filled before this promising addition to word learning procedures can be put to more effective use. We focus on word recall—the recall of word forms (e.g., /sprɪg/), and meanings (definitions or characteristics of the referents, e.g., a small stem with leaves or flowers). Tasks such as recognition (e.g., “Which one is a /sprɪg/?”) are, of course, informative but have been less consistent in identifying differences between groups and between learning conditions (e.g., Gray, 2004; Leonard et al., 2021). Our emphasis is on children with DLD but we refer to relevant studies of adults with DLD when they help in the interpretation of the child data. Before focusing on specific areas of spaced retrieval that warrant closer consideration, we begin with a brief review of what seem to be some of the major findings thus far in the DLD literature.

What (We Think) We Know So Far

The use of retrieval to help children with DLD learn new words has appeared in the literature at least since 2014 (Chen & Liu, 2014) and numerous studies have been published since that date. Over this period, certain findings emerge consistently in the literature.

One major finding is that recall of novel word forms is stronger when learning incorporates spaced retrieval than

when it provides only passive study of the words (“study trials”) with the same degree of exposure (e.g., Leonard, Karpicke, et al., 2019; Leonard et al., 2023; McGregor, Gordon, et al., 2017). This finding applies to novel words representing nouns (Leonard, Karpicke, et al., 2019), adjectives (Leonard, Deevy, et al., 2019), and verbs (Leonard et al., 2023). Effect sizes are large for learning word forms for novel nouns and adjectives; medium effect sizes are seen for novel verb learning.

Spaced retrieval advantages are also seen for meanings, though the relevant studies have been limited to novel words referring to nouns. The effect sizes reflecting the advantage of spaced retrieval over repeated study are not as large for meaning as they are for word form, ranging from small to medium (McGregor, Arbisi, et al., 2017; see aggregated data analysis in Leonard et al., 2021).

Once a minimal level of success is seen for spaced retrieval, providing additional retrieval trials with no additional study will lead to better recall than providing additional study trials with no further opportunity for retrieval (Leonard et al., 2020).

Fewer studies have compared spaced retrieval with learning conditions other than passive study. The available evidence shows greater recall of word form and meaning in spaced retrieval conditions compared to strictly immediate retrieval conditions, with large and small effect sizes, respectively (Haebig et al., 2019). Evidence from real-word learning in a classroom setting suggests that spaced retrieval produces greater gains than learning through rich vocabulary instruction without retrieval opportunities—a finding receiving moderate support based on Bayesian analysis (Levlin et al., 2022).

During the learning period, retrieval trials that involve free recall (e.g., “What is this one called?”) are associated with lower accuracy during learning than trials that involve cued recall (e.g., recalling the whole word after being given the first syllable of the word; Gordon, Storkel, et al., 2021). However, the opposite is true for longer-term retention; free recall trials during the learning period lead to better retention outcomes (Gordon, McGregor, et al., 2021).

Overall, children with DLD recall fewer word forms and meanings than their same-age peers with typical language development (e.g., Gordon, Storkel, et al., 2021). However, the group differences are smaller for spaced retrieval conditions than for comparison conditions (Leonard et al., 2021).

Studies of novel adjectives and novel verbs show that children can apply the novel word to novel referents (Leonard, Deevy, et al., 2019; Leonard et al., 2023). In the case of adjectives, the novel adjective can be successfully applied to previously unseen objects with the same novel attribute. In the case of verbs, the novel verb can be applied to a new agent performing the novel action. The advantage of spaced retrieval over repeated study is seen

for these generalization items as well as for the items that were presented during the learning period.

Advantages of spaced retrieval conditions over comparison conditions are seen for longer-term recall (e.g., 1 week), not just shortly after the end of the learning period (e.g., Gordon, Storkel, et al., 2021; Leonard & Deevy, 2020).

Instead of longer-term retention being a major weakness, it seems to be a relative strength (e.g., Gordon, Storkel, et al., 2021; Pomper et al., 2022). The children’s biggest challenge appears to be encoding, especially the encoding of word forms, but also of meanings (e.g., Bishop & Hsu, 2015; Gordon, Storkel, et al., 2021; Jackson et al., 2021; McGregor, Gordon, et al., 2017).

Our confidence in the above findings should be tempered by the fact that they are based on a relatively narrow database. With a few exceptions (e.g., Chen & Liu, 2014; Levlin et al., 2022), most participants have been English-speaking children. The youngest ages at which spaced retrieval might still be effective have not yet been discovered; most evidence thus far comes from studies of children 4 years of age and older. We also have not identified factors that might explain why spaced retrieval is more effective for some children than others. Apart from such basic questions, we also need a greater understanding of the *process* of spaced retrieval, especially as applied to word learning in children with DLD. We turn to this issue next.

Gaps in Our Understanding of Spaced Retrieval Effects

In the following sections, we examine several issues related to spaced retrieval that warrant a closer look. Previous studies have provided a foundation for further study of some of these issues, and we take advantage of this groundwork to develop a deeper understanding of how spaced retrieval might operate in DLD.

Next-day consolidation effects

Whether teaching real words to children or studying the children’s learning of novel words, investigators often use more than one session scheduled on different days. However, there is one significant difference between spaced retrieval on a subsequent day of learning versus spaced retrieval on the first day of learning. Testing on multiple days allows for consolidation. Consolidation is the process by which recently learned material becomes stabilized and integrated with information in long-term memory. Intervening sleep promotes consolidation of material that has been learned (Dumay & Gaskell, 2007; see meta-analysis in Schimke et al., 2021). According to Davis and Gaskell (2009), the process involves initial encoding of novel words primarily in hippocampal areas.

During sleep there is communication with the neocortex, allowing integration of the novel words with other details in semantic and phonological memory. This is a type of processing of encoded information that is automatic and internal (Walker, 2005). The integration results in a stronger memory trace and greater long-term retention. Children appear to benefit from consolidation to an even greater extent than adults (James et al., 2019). However, the degree to which words have been encoded appears to influence the likelihood of consolidation. Words do not have to be strongly encoded before sleep, but if encoding is very weak, consolidation may not occur (Drosopoulos et al., 2007; Schoch et al., 2017).

When spaced retrieval is incorporated across separate days, it is difficult to distinguish any effects due to spacing from effects due to consolidation. This is complicated further by reports of parallels between spaced retrieval and consolidation because in spaced retrieval, as in sleep, there is hippocampal-to-neocortical communication (Antony et al., 2017).

Perhaps the best research designs to isolate the effect of consolidation are those that compare morning learning with evening testing and evening learning with next-morning testing, with duration in hours between learning and testing kept constant (e.g., 12 h; see Henderson et al., 2012). Studies of this type are sparse in the DLD literature, but McGregor, Licandro, et al. (2013) provided an informative study with such a design. Young adults with DLD and those with typical language skills learned novel words referring to fantasy animals and objects. Half of the participants in each group learned the novel words in a morning session and were then tested 12 h later on the same day. The other half of the participants learned the novel words in the evening and were then tested 12 h later the next morning. The same tests were then administered 12 h, 24 h, and 1 week after the immediate posttests. Across tests and time, the participants with typical language skills outperformed the participants with DLD. Overall, for both groups, gains on the later tests of word meaning were larger for those participants who learned the words in an evening session and received the immediate posttests the next morning, after a night of sleep. There was only a numerical difference favoring the next-day participants for word forms.

During the learning period itself, the participants in the McGregor, Licandro, et al. (2013) study heard the words and saw the referents but were not asked to provide the words or meanings until the immediate posttest. An important study for the future would be one that employed morning subgroups that engaged in spaced retrieval or repeated study and evening subgroups that engaged in spaced retrieval or repeated study. Comparisons across the four subgroups might reveal whether spaced retrieval adds to the recall advantages provided by consolidation through sleep.

Several studies of spaced retrieval employed tests of recall only after the second day of learning. At least one of these studies ended the first day with a spaced retrieval trial and began the second day with a spaced retrieval trial, thus providing a glimpse at possible consolidation effects. Leonard, Deevy, et al. (2019) asked 4- and 5-year-old children with DLD and same-age peers to learn novel adjectives referring to unusual attributes. Four novel adjectives were learned in each of two sets, separated by 1 week. In each set, two novel adjectives were assigned to a spaced retrieval condition and two were assigned to a repeated study condition. For words in the spaced retrieval condition, day 1 for each set began with two immediate retrieval trials. All trials thereafter were spaced retrieval trials with three other novel adjectives intervening between the time the to-be-retrieved word was last heard and when it had to be retrieved. Figure 1 provides an example of each of these two types of trials. This schedule meant that the final trial of day 1 and the first trial of day 2 were both spaced trials. There was a re-familiarization of the novel adjectives and referents at the beginning of day 2, but there were no opportunities for production or retrieval until the first spaced retrieval trial. In Table 1, we show the children's retrieval accuracy on the final spaced trial of day 1 and the first spaced trial on day 2.

As can be seen, for both groups, there was more correspondence than lack of correspondence between the final retrieval trial on day 1 and the first retrieval trial on day 2. There was no evidence of an increase in accuracy due to consolidation. At most, one might speculate that consolidation served as a protective factor against forgetting from one day to the next, given that, for both groups, correct recall on the final trial on day 1 was much more likely to be preserved than lost on day 2 (typical development: 15 preserved versus 4 lost; DLD: 19 preserved versus 3 lost).

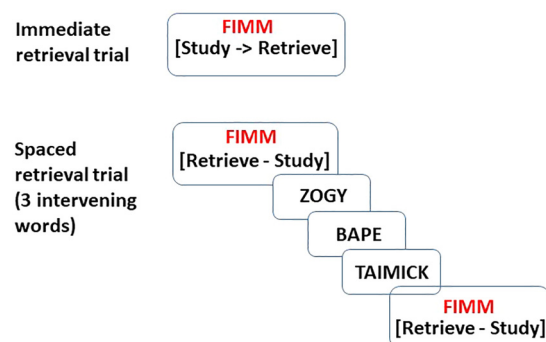


Figure 1. An example of an immediate retrieval trial with no intervening words between the study trial and the retrieval trial, and an example of a spaced retrieval trial with three other words intervening between the last time the to-be-retrieved word was heard (in a study trial) and when it had to be retrieved. Examples from Leonard, Deevy, et al. (2019).

Table 1. Accuracy on the final spaced retrieval trial of day 1 and the first spaced retrieval trial of day 2 for the children with developmental language disorder and their peers with typical language development in the novel adjective learning study of Leonard, Deevy, et al. (2019).

Group	Total number of words	Number of different children
<i>Typical language development</i>		
Correct final day 1, correct first day 2	15	11 of 13
Incorrect final day 1, incorrect first day 2	29	13 of 13
Correct final day 1, incorrect first day 2	4	4 of 13
Incorrect final day 1, correct first day 2	4	4 of 13
<i>Developmental language disorder</i>		
Correct final day 1, correct first day 2	19	11 of 14
Incorrect final day 1, incorrect first day 2	31	11 of 14
Correct final day 1, incorrect first day 2	3	3 of 14
Incorrect final day 1, correct first day 2	3	3 of 14

Kueser et al. (2021) provided an important view of potential consolidation effects through an analysis of data from three previous studies (Haebig et al., 2019; Leonard, Deevy, et al., 2019; Leonard, Karpicke, et al., 2019). Each study involved preschool-age children with DLD and their age mates with typical language development. Kueser et al. fit growth curves to the trial-by-trial learning data from the spaced retrieval conditions in those studies. All three studies covered two consecutive days. One of those studies was the one reviewed earlier (Leonard, Deevy, et al., 2019), in which both the final trial of day 1 and the first trial of day 2 were spaced retrieval trials. For the remaining two studies, the final trial of day 1 was a spaced trial but day 2 began with an immediate retrieval trial for each word. Spaced trials for each word occurred after that point.

Kueser et al. (2021) found that across studies, both groups of children showed similar linear growth during learning, despite the data crossing from day 1 to day 2. Figure 2 provides an illustration of the spaced retrieval data from the three studies. Note especially how similar the changes are from day 1 to day 2 for the two groups in each study. Given the appearance of an immediate retrieval trial before the first spaced trial of day 2 in two of the studies, we don't have a perfect view of possible consolidation. It is possible, for example, that the immediate

retrieval trial that began day 2 served as a type of “refresher.” What does seem clear is that there is no evidence that the retrieval pattern for children with DLD is any different from that of their same-age peers. (We will return at a later point to the role that the insertion of immediate retrieval trials plays in spaced retrieval schedules.)

Longer-term retention as reconsolidation

A frequent finding in the literature is that spaced retrieval leads to better recall of words by the second day than alternative procedures. For words that were successfully recalled at that point, retention over longer periods is quite stable (Leonard & Deevy, 2020). This finding is compatible with earlier studies of both children and adults with DLD (McGregor, Gordon, et al., 2017; McGregor et al., 2020). In some studies, individuals with DLD recall less than their peers, but this is true from the outset; group differences do not become greater over time. As noted earlier, the culprit from the very start is encoding—getting the word form to “stick” in the first place, not retention thereafter (e.g., Bishop & Hsu, 2015; Gordon, Storkel, et al., 2021; Jackson et al., 2021; McGregor, Arbisi, et al., 2017).

Longer-term retention can be viewed in terms of reconsolidation. This term refers to the changes in memory that occur when retrieval takes place after an initial consolidation period. For example, when sleep occurs after a learning period (setting the stage for consolidation), retrieving the learned items on the second day will further change memory through reconsolidation. One explanation of these additional changes is that after an initial consolidation period, subsequent retrieval renders the memory less stable and more amenable to further elaborations (Alberini, 2011; Smith & Scarf, 2017).

In the 2-day learning period in the studies by Leonard and colleagues, such reconsolidation could have been triggered through retrieval trials and testing on the second day and reflected in the testing that occurred 1 week later. However, further improvements in recall across the 1-week period were not observed; instead, there was notable stability in the recall scores across time. For the young ages of the participants in those studies (4- and 5-year-olds), reconsolidation, much like initial consolidation, could have served more as a protection against forgetting than as an occasion for further elaboration. (As we will see in the next section, this protective factor may apply more for word learning under spaced retrieval conditions than for word learning in general, especially for phonetic details of the word.)

Another look at reconsolidation is available from a study by Gordon, Storkel, et al. (2021). These investigators used retrieval practice in teaching novel words across six sessions with a group of preschoolers with DLD and their typical peers. Although not described in terms of reconsolidation, free recall measures were obtained at the beginning

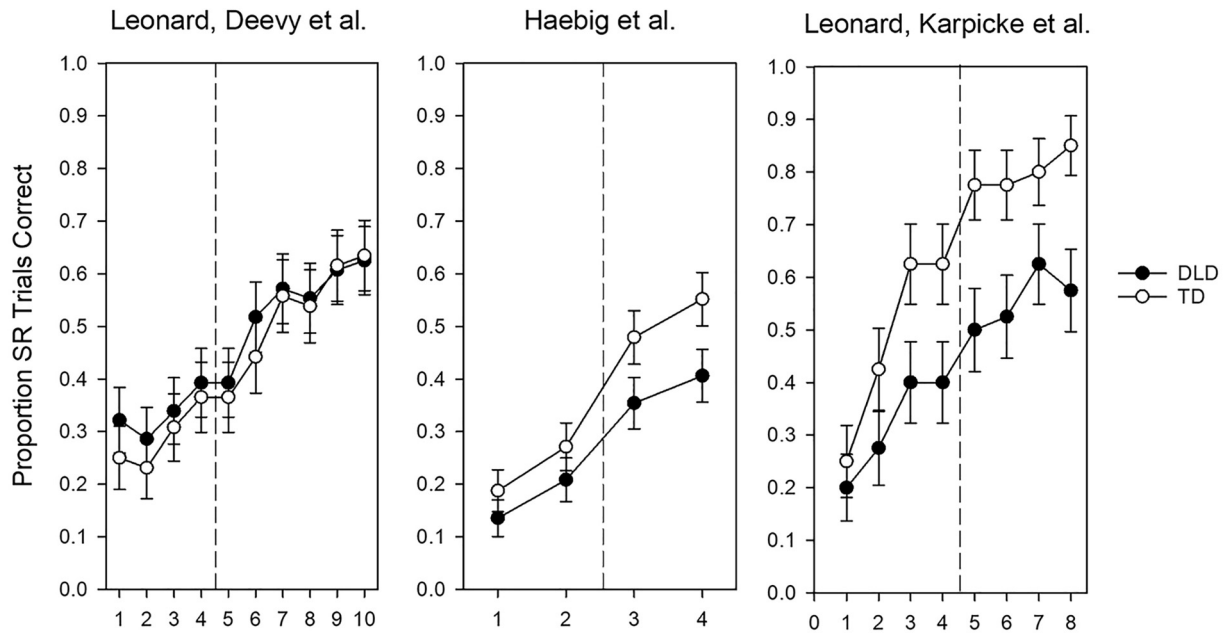


Figure 2. An illustration of data reported by Kueser et al. (2021) in which the trial-by-trial accuracy on spaced retrieval trials is shown for three different studies (Haebig et al., 2019; Leonard, Deevy, et al., 2019; Leonard, Karpicke, et al., 2019). The dotted lines represent the separation of trials for the first and second day of learning. Changes in accuracy from the first to the second day were very similar for the children with developmental language disorder and their same-age peers with typical language development.

of each session (except the first) and at the end of each session. By comparing the children's recall from the end of one session to the beginning of the next over 6 days, Gordon, Storkel et al. acquired important data about successive overnight retention of the words. (There is one qualification: if a child was incorrect on the free recall at the end of a session, a cued recall prompt for the same word was provided [e.g., "It starts with /b/."] and feedback was provided.) Although the children with typical development were more accurate than the children with DLD throughout the learning period, the two groups were quite similar in their stability from the end of one session to the beginning of the next. Both groups showed some decline when tested again 1 month later, but there were no group differences in the degree of decline.

There is one finding regarding reconsolidation that on reflection would not have been expected. If consolidation and reconsolidation enable newly learned material to be integrated with existing semantic and phonological knowledge, the children with typical language development should have benefitted more from these processes than the children with DLD. Based on the standardized language tests serving as selection criteria and on the vocabulary tests serving as covariates, the children with typical development showed significantly greater linguistic knowledge than their counterparts with DLD. This suggests that for the typical children, there was more in long-term semantic and phonological memory for the novel words to integrate with during consolidation and reconsolidation. Yet, the (limited)

changes over time in the two groups were comparable. Similarly, the two groups were similar in their ability to generalize novel adjectives and verbs to new scenarios. Consolidation and reconsolidation seemed to preserve over time what the children had achieved, but there is no clear evidence that these processes included the integration of the prior linguistic knowledge that distinguished the two groups. Future research should include more direct measures of integration (e.g., word association or lexical categorization tasks) to solve this puzzle.

Spaced retrieval preserves longer-term retention of phonetic details

A finding by Leonard et al. (2022) requires a modification of the "stable long-term retention" view for children with DLD. Yet this modification does not alter the view that encoding is the chief problem. Recall that Haebig et al. (2019) found that spaced retrieval led to greater recall of word forms than immediate retrieval at tests 5 min after learning and 1 week later. However, even though fewer words were recalled in the immediate retrieval condition, those that were recalled showed good stability over time. A new look at these data suggests that a qualification of this interpretation is in order.

The scoring procedures used by Haebig et al. (2019) were based on Edwards et al. (2004) and allowed for some phonetic imprecision. For example, a production of

/topik/ instead of /pobik/ was scored as correct if it met other scoring criteria. Leonard et al. (2022) examined the novel words meeting the original criteria as “correct” and assigned them a score based on the full range of the Edwards et al. system. For example, /topik/ would earn a score of 14 out of 16, with one point each deducted for the error of place of articulation in initial position and the error of voicing in medial position.

Leonard et al. (2022) found that, overall, the children with typical language development were more phonetically accurate than the children with DLD, even when articulation accuracy on real words was taken into account. The more informative findings were seen in the results for the learning condition across time. An illustration of the results appears in Figure 3. For both groups, the phonetic accuracy for the final retrieval trial of the learning period was higher for words in the immediate retrieval condition than for words in the spaced retrieval condition. This could have been due to the higher success rates for immediate retrieval during the learning period which gave the children more practice in production and hence an early encoding advantage. However, when tested 5 min later, words in the immediate retrieval condition declined in accuracy whereas accuracy for words in the spaced retrieval

condition remained the same or improved. Most notably, no further decline was seen for the children with typical development 1 week later for words in the immediate retrieval condition but a steep decline was seen for the children with DLD. This was a rather selective decline, however, because the accuracy of the children with DLD on words in the spaced retrieval condition showed no decline 1 week later. Their phonetic stability from the 5-min test to the 1-week test was the same as for the children with typical language development. Furthermore, at the 1-week mark, the children with DLD were more phonetically accurate on words in the spaced retrieval condition than for words in the immediate retrieval condition—a reversal of the pattern seen on the final retrieval trial of the learning period.

The findings suggest additional benefits to the spaced retrieval condition. The previous studies by Leonard and colleagues indicated that more words are learned in the spaced retrieval condition but for words learned in any condition, long-term retention is quite good. However, this interpretation now seems more appropriate only for words learned in the spaced retrieval condition. Even when encoding seems to be assisted by immediate retrieval, the phonetic representations appear to be fragile in children with

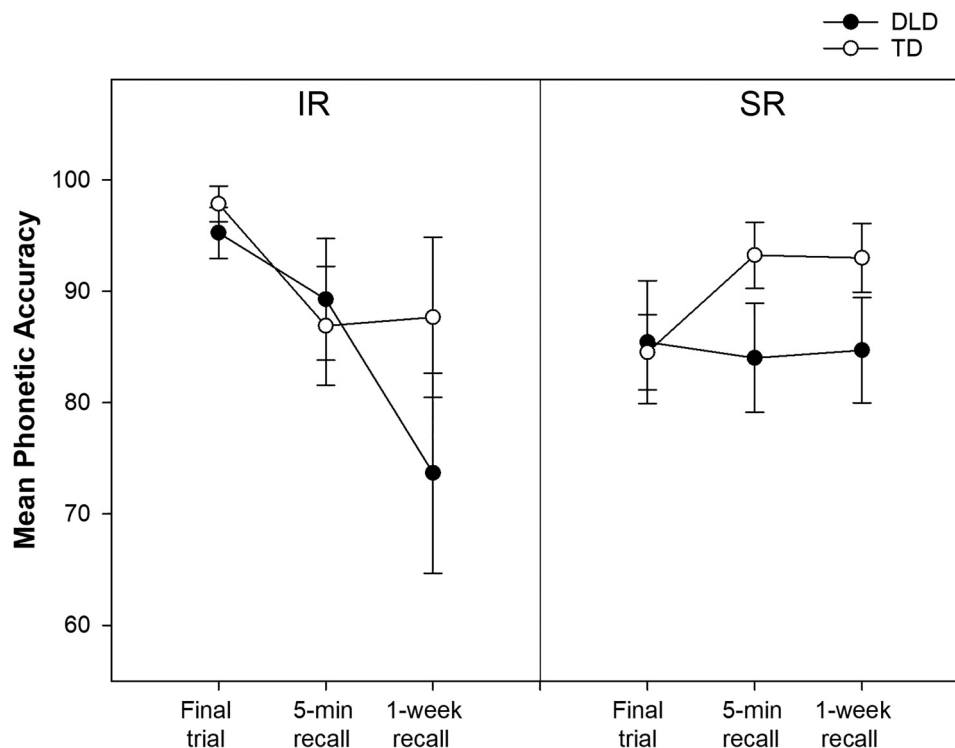


Figure 3. An illustration of data reported by Leonard et al. (2022) based on Haebig et al. (2019) which involved a comparison between the immediate retrieval (IR) condition and the spaced retrieval (SR) condition. Shown is the phonetic accuracy for: the final recall trial during the learning period; the recall test administered 5 min after the learning period; and the recall test administered 1 week after the learning period. Of note is the dramatic drop in phonetic accuracy at 1 week by the children with developmental language disorder, but only for novel words that were in the immediate retrieval condition.

DLD, resulting in a decline in accuracy over time. Although spaced retrieval provided the children with less production practice by virtue of the children's low success rate in early spaced retrieval trials, the phonetic representations that were formed were more robust.

How the course of learning affects subsequent recall

Thus far, we have discussed the relation between the learning period and later recall testing mostly in terms of how the condition to which the words had been initially assigned (e.g., spaced retrieval versus repeated study) influences final recall. Also relevant is what happens *during the learning period itself*.

In the study by Gordon, Storkel, et al. (2021) discussed earlier, preschoolers with DLD and their peers participated in six novel word learning sessions with retrieval practice. Although the children with typical development were more accurate overall, the gains from the beginning of each session to the end of the session were similar for the two groups. Despite their encoding weaknesses that represent a significant challenge from the very start, children with DLD do not seem to exhibit within-session plateaus in their word learning with retrieval practice.

Gordon, McGregor, et al. (2021) analyzed the trial-by-trial data of young adults with DLD and those with typical language ability who were participants in McGregor, Gordon, et al. (2017). Sets of two-syllable novel words were learned on a single day using retrieval practice, and recall was tested 24 h later. Two of the findings reported by Gordon, McGregor et al. reveal the importance of looking at the nature of the retrieval attempts during the learning sessions themselves. Novel words that were more frequently retrieved during learning were more likely to be recalled 24 h later. The second finding was even more illuminating. High phonetic accuracy on the final retrieval trial was predictive of better recall the next day, but especially when the participant's poorest phonetic accuracy of the novel word during the learning period was also relatively high. When the poorest production was relatively low, high accuracy on the final trial was less predictive of next-day recall. It seems that the trajectory of a child's gains during learning is an important factor to consider.

In some of the studies employing spaced retrieval, the first retrieval trial in one or more sessions was an immediate retrieval trial. We learned from the Haebig et al. (2019) study that a condition with a small number of immediate retrieval trials and a larger number of spaced retrieval trials produced better recall than a condition consisting entirely of immediate retrieval trials. In spaced retrieval conditions, immediate retrieval trials were included to provide the children with an encoding opportunity when retrieval demands were minimal. However, whether and

how the inclusion of immediate retrieval trials in these instances actually facilitates learning has not been clear.

The Kueser et al. (2021) study introduced earlier provided an informative look at the possible contribution of immediate retrieval trials in spaced retrieval conditions. Using the trial-by-trial data from three prior studies, Kueser et al. found that successful immediate retrieval trials did not predict final recall scores by themselves. However, successful retrieval on immediate retrieval trials did predict success on subsequent spaced retrieval trials. And greater success with spaced retrieval trials predicted higher final recall scores. Though insufficient by itself, then, successful immediate retrieval appears to play a supportive role in creating greater spaced retrieval stability and, as a consequence, greater final recall.

Should spacing be gradual?

Although spaced retrieval holds an advantage over immediate retrieval, there is always the question of finding the right degree of spacing for the individual and for the type of material to be learned. Unfortunately, even the studies showing advantages of spaced retrieval over repeated study or immediate retrieval have not produced ideal results. For example, in a study of novel nouns with three intervening words, Leonard, Karpicke, et al. (2019) found that the children with DLD showed later recall of only 63%. In a similar study with novel verbs, recall was only 35% for the children with DLD (Leonard et al., 2023).

In an attempt to improve the results of spaced retrieval, Leonard et al. (2024) compared a retrieval schedule with three intervening words with a schedule that expanded the spacing more gradually. Both conditions began with an immediate retrieval trial. In the "expanded" condition, the next two retrieval trials had only one intervening word, followed by two trials with three intervening words. Figure 4 provides an example of each of the three types of retrieval trials in the expanded condition. In the "equally spaced" condition, the immediate retrieval trial was followed directly by four trials with three intervening words. The same two schedules were employed on a second (consecutive) day.

In the end, the two retrieval conditions resulted in very similar recall 5 min after the second learning session and 1 week later. However, an examination of the trial-by-trial data revealed different paths on the way to similar final recall. As expected, the trials with only one intervening word were more likely to be retrieved successfully than the corresponding trials with three intervening words in the equally spaced condition. It was also the case that the trials with three intervening words that directly followed the one-intervening-word trials in the expanded condition were retrieved more successfully than the corresponding trials in the equally spaced condition even though in both instances, three words intervened. This seemed to give

children in the expanded condition a head start by giving them earlier success with three intervening words. However, as trials progressed, retrieval accuracy in the two conditions began to converge, with greater relative gains seen for the equally spaced condition. For the children with typical language development, convergence occurred by the end of the first day. For the children with DLD, this occurred by the end of the second day. Table 2 shows the retrieval points at which the probabilities of successful retrieval begin to converge in the two conditions.

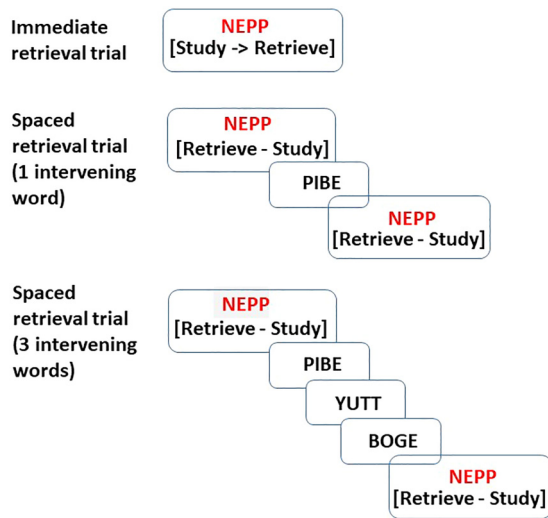


Figure 4. An example of the three degrees of spacing used in the expanded retrieval condition in Leonard et al. (2024). The learning period began with an immediate retrieval, followed by two spaced retrieval trials with one other word intervening between the last time the to-be-retrieved word was heard (in a study trial) and when it had to be retrieved, followed by two spaced retrieval trials with three other words intervening between the last time the to-be-retrieved word was heard (in a study trial) and when it had to be retrieved.

Table 2. The probability of successfully retrieving a novel word at the fourth and fifth retrieval point on each day in the Leonard et al. (2024) study.

		Day 1		Day 2	
		Trial 4	Trial 5	Trial 4	Trial 5
DLD	Expanded	0.279***	0.462*	0.436**	0.487
	Equal	0.123***	0.298*	0.317**	0.472
TD	Expanded	0.400**	0.531	0.774	0.841
	Equal	0.248**	0.509	0.660	0.811

Both retrieval points had spacing with three intervening words for both conditions. However, for the expanded retrieval condition, the fourth retrieval point directly followed a retrieval point with the shorter (one intervening word) spacing. Asterisks indicate significant differences between the expanded and equally spaced conditions within the same retrieval point, within each day and group: * $p < .05$; ** $p < .01$; *** $p < .001$.

The finding of no advantage in final recall with an expanded schedule was disappointing. Furthermore, it is unclear why the benefits of shorter spacing for subsequent trials with greater spacing are only temporary. We speculate on possible reasons in the next section of this article where we discuss how the contribution of context to learning might interact with spacing.

Features of context during spaced retrieval

Retrieval is often treated as a process apart from encoding, given that, in retrieval, the emphasis is on accessing something already in memory. However, retrieval is also an active form of encoding (e.g., Johnsson et al., 2021; Karpicke & Grimaldi, 2012). Furthermore, features of the context may also be part of the information encoded during the retrieval process. As described below, these features can remain associated with the basic representation and thus strengthen the consistency of retrieval which, in turn, provides the learner with more opportunities for the representation to strengthen.

The contribution of context is easy to imagine in circumstances in which one tries to recall a previously studied item in a new location, such as testing one's recall of class material first in a library and then in a coffee shop. The library and coffee shop surroundings do not form part of the defining features of the material studied. However, they can remain associated with the study event. In more structured laboratory studies, the notion of context is much more subtle. A study by Whiffen and Karpicke (2017) provides an example. These investigators asked young adult participants to study two lists of words, separated by a brief unrelated task. The participants were then shown the words from the two lists mixed together, with half of the participants asked during this presentation which of the two lists each of the words had originally appeared in. During subsequent recall, better recall was seen by the participants who made judgments about the specific list that each word had been seen in. Note that during the initial study period, the participants were not informed that the list that a word appeared in was relevant to the task. Yet, this information seemed to be available when participants were asked to decide on list membership, and this process of retrieving contextual (list) information led to better recall.

In studies of this type, the physical context changed very little if at all, but the temporal context underwent change, even if the time span was quite short. Bäuml (2019, p. 177) defined temporal context as “the current pattern of activity in the individual's mind...” when the material is being studied. When participants retrieve an item, there is partial reactivation of the original context. Successful retrieval allows this partial context to join with the context present during retrieval. With further retrieval, more portions of context get added, resulting in a composite of contextual features that is not identical to the temporal

context present in any single study or retrieval event. This composite becomes more unique when spacing is employed, because with more spacing (as in more intervening items), the temporal context changes to a greater extent with each act of retrieval (see Karpicke, 2017; Karpicke et al., 2014).

One of the findings discussed in the previous section seems to support the idea of changing temporal contexts assisting word learning and recall. Leonard et al. (2024) attempted to determine if a gradual increase in the spacing between words (the expanded condition) would better prepare children with DLD and their peers to succeed in retrieving words with greater spacing. It was found that during the learning period, shorter-spaced retrieval trials were easier than the corresponding trials in the equally spaced condition that had greater spacing. Furthermore, the gradual spacing seemed to be helpful in the near term because the first one or two trials with three intervening words that directly followed trials with shorter spacing were more successful than the corresponding trials in the equally spaced condition, even though at that point the trials in both conditions had the same degree of (greater) spacing. However, as noted earlier, words in the condition with greater spacing all along showed steeper acceleration across the learning period and by the end of the learning period, success on the trials with three intervening words was the same in the two conditions.

Why would this be true? One possibility is that the equally spaced condition had twice as many retrieval trials with greater spacing which, based on a context account, provided more opportunities for the temporal context to change with each retrieval attempt. This, in turn, can lead to a building-up of a composite of features that is more distinct than a composite based on retrieval of words with shorter spacing and less change in context.

Effort and the possible contribution of feedback

“Manipulations that speed the rate of acquisition during training can fail to support long-term posttraining performance, while other manipulations that appear to introduce difficulties for the learner during training can enhance post-training performance” (Bjork, 1994, p. 185).

One such manipulation is “effortful” retrieval, and spaced retrieval is perhaps the most prototypical form of effortful retrieval. In some studies, participants provide subjective judgments of effort (e.g., Karpicke & Roediger, 2008; Kornwell & Bjork, 2008), but in most instances, effort is only inferred, based on low success rates on early trials. As is well documented, once some degree of success begins to occur on these challenging trials, long-term retention proves to be better than is seen for less effortful trials (such as those with no spacing). In these instances, it can be said that the particular spacing represents the level of “desirable difficulty.”

Vlach (2019) has pointed out that effortful retrieval—achieved through procedures like spacing, for example—may be a particularly important factor for younger children. As younger children have limitations in memory and quickly forget newly presented words, retrieving words often requires effort. When children have success, though, words can be learned and retained for longer periods in effortful conditions than in situations that do not challenge memory.

One factor that can likely make effortful retrieval more successful is feedback. Feedback refers to the participant hearing or seeing the correct response after trying to retrieve it, even without being explicitly informed whether the retrieval attempt was correct. Although benefits can still accrue from effortful retrieval without feedback, providing feedback appears to increase accuracy on subsequent retrieval trials (e.g., Ma et al., 2020). Feedback appears to be most helpful when participants attempted to engage in effortful retrieval, as for example, when the previous retrieval attempt was unsuccessful (Rowland & DeLosh, 2015), or the retrieval attempt was correct but the participant was unsure of its accuracy (Butler et al., 2008). In these instances, receiving feedback about the correct form when a participant is unsure is more helpful than receiving feedback about a form that the participant was already confident about.

Because participants are often unsuccessful and less confident in the early trials of a spaced retrieval regimen, feedback is particularly important during this period of learning. This may be one of the reasons why in the study by Haebig et al. (2019), a spaced retrieval condition was more effective than a condition in which all retrieval trials involved immediate retrieval. Children were much more successful in their responses in the immediate retrieval condition, even in the early trials. When feedback was provided directly after these accurate responses, it was less informative to the children and therefore more likely to maintain current learning than to create new learning. Greater benefit likely occurred after a failed attempt to retrieve a word—an occurrence much more likely in the spaced retrieval condition, especially early on.

Learning to criterion or a fixed learning period?

Gordon et al. (2024) taught a set of real words to a group of 4- to 7-year-old children with DLD. Nine additional words were used as control words. Only four words from a larger set were taught per session, using retrieval practice procedures. Each word was taught for up to seven sessions. However, if a child retrieved the words correctly at the beginning of two different sessions, the word was no longer included in the training and was replaced by another word from the larger set. This procedure constituted a criterion-based procedure, differing from other studies that had a pre-established number of exposures and retrieval

opportunities independent of the child's success with any given word.

As expected, testing after the intervention period revealed greater recall of the words that were included in teaching than the words serving as control words. The words that met the criterion for early exclusion from the training list showed greater recall than the words for which the children never reached criterion and continued to study and retrieve for all seven sessions. This difference was also seen on testing 8 weeks after the conclusion of intervention. An examination of the individual children's performance indicated that six of the seven children clearly benefitted from the intervention. For the six children showing some degree of success, the average number of sessions before a word was excluded was 4.68, leaving 2.32 additional sessions before recall of the word was tested. This finding speaks to the value of retrieval success; later recall is better even when fewer sessions were devoted to such words and more time transpired before testing.

Leonard et al. (2020) presented data that were compatible with the findings of Gordon et al. (2024). Children with DLD and their same-age peers learned a set of novel words to a minimal criterion, at which point half of the words continued to be presented in study-only trials and the other half continued to be tested in retrieval-only trials, with no additional exposures. These conditions were referred to as more study/less retrieval and more retrieval/less study, respectively. For both groups of children, later recall was superior for the novel words in the more retrieval/less study condition. This was true for testing directly after the learning period and 1 week later. The Leonard et al. study differed from the Gordon et al. study in that words no longer studied continued to appear in retrieval trials. However, both investigations showed that once some clear level of success has occurred for a word, additional study of those words may not be required. As Gordon et al. note, this could render vocabulary instruction more efficient than procedures that continue with all words throughout the intervention period. (As further evidence that this also applies to children with typical language development, Gordon and Lowry [2024] found that 4- to 6-year-old children with typical development could recall novel words with high probability 1 month later if they could successfully retrieve the words at the beginning of four different learning sessions. No further long-term benefit derived from the children recalling the novel word at the beginning of a fifth learning session.)

Could co-morbid dyslexia play a role?

As we noted earlier, several studies have shown that preschool-age children with DLD have greater weaknesses in learning word forms than in learning their meanings. However, the results of a study of school-age children by

Adlof et al. (2021) do not comport with those findings. In the Adlof et al. study, the children learned words using a spaced retrieval procedure. As might have been expected, a group of children with both DLD and dyslexia were found to differ from their same-age peers in both word form and meaning recall. However, a group with DLD only differed from their peers on meaning recall but not on word form recall.

This finding raises the possibility that a large percentage of preschoolers with DLD who were participants in previous studies might have met the criteria for (comorbid) dyslexia when they reached school age, and their pattern of especially weak word form learning as preschoolers might have been more closely related to the same vulnerabilities seen in dyslexia than to their vulnerabilities in the areas of semantics and grammar (see related arguments in studies comparing children with dyslexia and children with both dyslexia and DLD; Alt et al., 2019; Malins et al., 2020). We note as well that word form weaknesses are also seen in adults with DLD (e.g., McGregor, Arbisi, et al., 2017) and many of these individuals have a documented history of comorbid reading deficits (e.g., McGregor et al., 2020).

Conclusions

Spaced retrieval appears to assist the word learning of individuals with DLD. However, the exact conditions under which this occurs and just why it occurs are still not clear. This review has identified several aspects of spaced retrieval in need of closer scrutiny. We need to be more confident about the actual drivers of change when spaced retrieval procedures outperform alternative approaches. It will be important for us to isolate spaced retrieval effects from effects from other sources. The following are some details that should be considered in future work.

Immediate retrieval seems to increase the likelihood of children's success with subsequent spaced retrieval trials with the same word. However, longer-term retention depends on success on prior spaced retrieval trials, not immediate retrieval trials. It also seems that increasing spacing more gradually may have short-term effects on subsequent proximal retrieval trials with greater spacing, but this advantage may not be sustainable. A more uniform schedule of somewhat greater spacing may produce the same results by the end of the learning period.

There is also evidence suggesting that high phonetic accuracy during the final retrieval trial does not tell the whole story. If earlier attempts were quite low in phonetic accuracy, longer-term recall will not be as successful. In addition, it appears that spaced retrieval may help to prevent the erosion of phonetic details of words learned by children with DLD. The notion of stable long-term retention of words by children with DLD may need to be

qualified according to the presence or absence of spaced retrieval opportunities during learning.

Findings from both the laboratory and the classroom suggest that once a word is beginning to be retrieved more consistently by a child, it may not be necessary to retain it on the list of study words. Children appear to remember the words, at least when recall testing occurs several days or a week later. Once a word is dropped from further study, re-testing it in occasional retrieval trials during the learning period should prove helpful. By removing words on which the child has shown success, practitioners can include new words to the study material, and as a result, increase treatment efficiency and, it is hoped, the inventory of words known by the child.

There remains the important question of why spaced retrieval works quite well. One key to its success is, ironically, the *relatively low success rate* of early spaced retrieval attempts. When feedback is provided in the form of hearing the correct word, useful information is transmitted to the learner. Feedback is probably less useful when the learner is confident in the accuracy of the retrieved word. Thus, the confidence of children with DLD during their retrieval attempts may be an important sign that additional (explicit) feedback would be helpful. Another potential key is more related to what might occur when learners *successfully retrieve* a word on a spaced retrieval trial. Portions of the context of the original exposure can join with features of the present context to form a composite. Further success will be built with additional retrieval, as more portions of prior context will be added, which will in effect reduce the memory search space for the learner because the feature composite associated with the word will be unique and distinct from potential competitors. Future work that would be informative could focus on determining what these context features are and whether manipulations of these features promote heightened rates of learning or more durable long-term recall.

Future findings of importance are no doubt in the offing. In the meantime, it is hoped that the details presented here help to move the study of spaced retrieval forward. A deeper understanding of how best to employ spaced retrieval will be an important piece in assisting the word learning of individuals with DLD.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This work was supported by Research Grant R01 DC014708 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health, USA.

ORCID iD

Laurence B. Leonard  <https://orcid.org/0000-0002-9189-4438>

References

- Abbott, E. E. (1909a). *On the analysis of the factor of recall in the learning process* [Master's Thesis]. University of Illinois.
- Abbott, E. E. (1909b). On the analysis of the factor of recall in the learning process. *Psychological Monographs: General Applied*, 11, 159–177. <https://doi.org/10.1037/h0093018>
- Adesope, O., Trevisan, D., & Sundararajan, N. (2017). Rethinking the use of tests: A meta-analysis of practice testing. *Review of Educational Research*, 87, 659–701. <https://doi.org/10.3102/0034654316689306>
- Adlof, S., Baron, L., Bell, B., & Scoggins, J. (2021). Spoken word learning in children with developmental language disorder or dyslexia. *Journal of Speech, Language, and Hearing Research*, 64, 2734–2749. https://doi.org/10.1044/2021_JSLHR-20-00217
- Alberini, C. (2011). The role of reconsolidation and the dynamic process of long-term memory formation and storage. *Frontiers in Behavioral Neuroscience*, 5, 12. <https://doi.org/10.3389/fnbeh.2011.00012>
- Alt, M. (2011). Phonological working memory impairments in children with specific language impairment. *Journal of Communication Disorders*, 44, 173–185. <https://doi.org/10.1016/j.jcomdis.2010.09.003>
- Alt, M., Gray, S., Hogan, T., Schlesinger, N., & Cowan, N. (2019). Spoken word learning differences among children with dyslexia, concomitant dyslexia and developmental language disorder, and typical development. *Language, Speech, and Hearing Services in Schools*, 50, 540–561. https://doi.org/10.1044/2019_LSHSS-VOIA-18-0138
- Antony, U. J., Ferreira, C., Norman, K., & Wimber, M. (2017). Retrieval as a fast route to memory consolidation. *Trends in Cognitive Sciences*, 21, 573–576. <https://doi.org/10.1016/j.tics.2017.05.001>
- Bishop, D. V. M., & Hsu, H. J. (2015). The declarative system in children with specific language impairment: A comparison of meaningful and meaningless auditory-visual paired associate learning. *BMC Psychology*, 3, 3. <https://doi.org/10.1186/s40359-015-0062-7>
- Bishop, D. V. M., Snowling, M., Thompson, P., & Greenhalgh, T., & CATALISE Consortium. (2017). CATALISE: A multi-national and multidisciplinary Delphi consensus study of problems with language development. Phase 2: Terminology. *Journal of Child Psychology and Psychiatry*, 58, 1068–1080. <https://doi.org/10.1111/jcpp.12721>
- Bjork, R. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe, & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). MIT Press.
- Bäuml, K. H. (2019). Context retrieval as a critical component in selective memory retrieval. *Current Directions in Psychological Science*, 28, 177–182. <https://doi.org/10.1177/0963721419827520>

- Butler, A., Karpicke, J., & Roediger, H. (2008). Correcting meta-cognitive error: Feedback increases retention of low-confidence correct responses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 918–928. <https://doi.org/10.1037/0278-7393.34.4.918>
- Carpenter, S., Pashler, H., & Cepeda, N. (2009). Using tests to enhance 8th grade students' retention of U. S., History facts. *Applied Cognitive Psychology*, *23*, 760–771. <https://doi.org/10.1002/acp.1507>
- Chen, Y., & Liu, H.-M. (2014). Novel-word learning deficits in Mandarin-speaking preschool children with specific language impairments. *Research in Developmental Disabilities*, *35*, 10–20. <https://doi.org/10.1016/j.ridd.2013.10.010>
- Davis, M., & Gaskell, M. G. (2009). A complementary systems account of word learning: Neural and behavioural evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1536). <http://rsta.royalsocietypublishing.org/content/364/1536/3773> <https://doi.org/10.1098/rstb.2009.0111>
- Dobson, J. (2012). Effect of uniform versus expanding retrieval practice on the recall of physiology information. *Advances in Physiology Education*, *36*, 6–12. <https://doi.org/10.1152/advan.00090.2011>
- Drosopoulus, S., Schulze, C., Fischer, S., & Born, J. (2007). Sleep's function in the spontaneous recovery and consolidation of memories. *Journal of Experimental Psychology: General*, *136*, 169–183. <https://doi.org/10.1037/0096-3445.136.2.169>
- Dubois, P., St-Pierre, M.-C., Desmarais, C., & Guay, F. (2020). Young adults with developmental language disorder: A systematic review of education, employment, and independent living outcomes. *Journal of Speech, Language, and Hearing Research*, *63*, 3786–3800. https://doi.org/10.1044/2020_JSLHR-20-00127
- Dumay, N., & Gaskell, M. G. (2007). Sleep-associated changes in the mental representation of spoken words. *Psychological Science*, *18*, 35–39. <https://doi.org/10.1111/j.1467-9280.2007.01845.x>
- Edwards, J., Beckham, M., & Munson, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research*, *47*, 421–436. [https://doi.org/10.1044/1092-4388\(2004/034\)](https://doi.org/10.1044/1092-4388(2004/034))
- Fazio, L., & Marsh, E. (2019). Retrieval-based learning in children. *Current Directions in Psychological Science*, *28*, 111–116. <https://doi.org/10.1177/0963721418806673>
- Fritz, C., Morris, P., Nolan, D., & Singleton, J. (2007). Expanding retrieval practice: An effective aid to preschool children's learning. *The Quarterly Journal of Experimental Psychology*, *60*, 991–1004. <https://doi.org/10.1080/17470210600823595>
- Goossens, N., Camp, G., Verkoeijen, P., Tabbers, H., & Zwaan, R. (2014). The benefit of retrieval practice over elaborative restudy in primary school vocabulary learning. *Journal of Applied Research in Memory and Cognition*, *3*, 177–182. <https://doi.org/10.1037/h0101800>
- Gordon, K. (2020). The advantages of retrieval-based and spaced practice: Implications for word learning in clinical and educational contexts. *Language, Speech, and Hearing Services in Schools*, *51*, 955–965. https://doi.org/10.1044/2020_LSHSS-19i-00001
- Gordon, K., & Lowry, S. (2024). Fostering retention of word learning: The number of training sessions children retrieve words positively relates to post-training retention. *Journal of Child Language*, *51*(3), 710–719. <https://doi.org/10.1017/S0305000923000752>
- Gordon, K., McGregor, K., & Arbisi-Kelm, T. (2021). Optimising word learning in post-secondary students with developmental language disorder: The roles of retrieval difficulty and retrieval success during training. *International Journal of Speech-Language Pathology*, *23*, 405–418. <https://doi.org/10.1080/17549507.2020.1812719>
- Gordon, K., Storkel, H., Lowry, S., & Ohlmann, N. (2021). Word learning by preschool-age children with developmental language disorder: Impaired encoding and robust consolidation during slow mapping. *Journal of Speech, Language, and Hearing Research*, *64*, 42504270. https://doi.org/10.1044/2021_JSLHR-21-00046
- Gordon, K., Storkel, H., Lowry, S., & Sultani, M. (2024). A word-learning intervention pilot study utilizing principles of retrieval- and criterion-based learning for children with developmental language disorder. *American Journal of Speech-Language Pathology*. https://doi.org/10.1044/2023_AJLSP-23-00037
- Gray, S. (2003). Word learning by preschoolers with specific language impairment: What predicts success? *Journal of Speech, Language, and Hearing Research*, *46*, 56–67. [https://doi.org/10.1044/1092-4388\(2003/005\)](https://doi.org/10.1044/1092-4388(2003/005))
- Gray, S. (2004). Word learning by preschoolers with specific language impairment: Predictors and poor learners. *Journal of Speech, Language, and Hearing Research*, *47*, 1117–1132. [https://doi.org/10.1044/1092-4388\(2004/083\)](https://doi.org/10.1044/1092-4388(2004/083))
- Haebig, E., Leonard, L., Deevy, P., Karpicke, J., Christ, S., Usler, E., Kueser, J., Souto, S., Krok, W., & Weber, C. (2019). Retrieval-based word learning in young typically developing children and children with developmental language disorder II: A comparison of retrieval schedules. *Journal of Speech, Language, and Hearing Research*, *62*, 944–964. https://doi.org/10.1044/2018_JSLHR-L-18-0071
- Henderson, L., Weighall, A., Brown, H., & Gaskell, M. G. (2012). Consolidation of vocabulary is associated with sleep in children. *Developmental Science*, *15*, 674–687. <https://doi.org/10.1111/j.1467-7687.2012.01172.x>
- Jackson, E., Leitão, S., Claessen, M., & Boyes, M. (2019). The evaluation of word learning abilities in people with developmental language disorder: A scoping review. *International Journal of Language and Communication Disorders*, *54*, 742–755. <https://doi.org/10.1111/1460-6984.12490>
- Jackson, E., Leitão, S., Claessen, M., & Boyes, M. (2021). Word learning and verbal working memory in children with developmental language disorder. *Autism and Developmental Language Impairments*, *6*, 1–20. <https://doi.org/10.1177/23969415211004109>
- James, E., Gaskell, M., & Henderson, L. (2019). Offline consolidation supersedes prior knowledge benefits in children's (but not adults') word learning. *Developmental Science*, *22*, e12776. <https://doi.org/10.1111/desc.12776>

- Johnsson, B., Wiklund-Hörnqvist, C., Stenlund, T., Andersson, M., & Nyburg, L. (2021). A learning method for all: The testing effect is independent of cognitive ability. *Journal of Educational Psychology, 113*, 972–985. <https://doi.org/10.1037/edu0000627>
- Kan, P. F., & Windsor, J. (2010). Word learning in children with primary language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 53*, 739–756. [https://doi.org/10.1044/1092-4388\(2009/08-0248\)](https://doi.org/10.1044/1092-4388(2009/08-0248))
- Kang, S., Lindsey, R., Mozer, M., & Pashler, H. (2014). Retrieval practice over the long term: Should spacing be expanding or equal-interval? *Psychonomic Bulletin & Review, 21*, 1544–1550. <https://doi.org/10.3758/s13423-014-0636-z>
- Karpicke, J. (2017). Retrieval-based learning: A decade of progress. In J. T. Wixted (Ed.), *Cognitive psychology of memory, volume 2. Learning and memory: A comprehensive reference* (pp. 487–514). Academic Press.
- Karpicke, J., & Grimaldi, P. (2012). Retrieval-based learning: A perspective for enhancing meaningful learning. *Educational Psychology Review, 24*, 401–418. <https://doi.org/10.1007/s10648-012-9202-2>
- Karpicke, J., Lehman, M., & Aue, W. (2014). Retrieval-based learning: An episodic context account. In B. Ross (Ed.), *Psychology of learning and motivation* (pp. 238–284). Elsevier.
- Karpicke, J., & Roediger, H. L. III. (2007). Expanding retrieval practice promote short-term retention, but equally spaced retrieval enhances long-term retention. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 33*, 704–719. <https://doi.org/10.1037/0278-7393.33.4.704>
- Karpicke, J., & Roediger, H. L. III. (2008). The critical importance of retrieval for learning. *Science, 319*, 966–968. <https://doi.org/10.1126/science.1152408>
- Kornwell, N., & Bjork, R. (2008). Learning concepts and categories: Is spacing the “enemy of induction”? *Psychological Science, 19*, 585–592. <https://doi.org/10.1111/j.1467-9280.2008.02127.x>
- Kueser, J., Leonard, L., Deevy, P., Haebig, E., & Karpicke, J. (2021). Word-learning trajectories influence long-term recall in children with developmental language disorder and typical development. *Journal of Communication Disorders, 94*, 106160. <https://doi.org/10.1016/j.jcomdis.2021.106160>
- Larsen, D., Butler, A., & Roediger, H. (2009). Repeated testing improves long-term retention relative to repeated study: A randomized controlled trial. *Medical Education, 43*, 1174–1181. <https://doi.org/10.1111/j.1365-2923.2009.03518.x>
- Latimier, A., Peyre, H., & Franck, R. (2021). A meta-analytic review of the benefit of spacing out retrieval practice episodes on retention. *Education Psychology Review, 33*, 959–987. <https://doi.org/10.1007/s10648-020-09572-8>
- Leonard, L. (2014). *Children with specific language impairment* (2nd ed.). MIT Press.
- Leonard, L., Christ, S., Deevy, P., Karpicke, J., & Kueser, J. (2024). Retrieval practice and word learning by children with developmental language disorder: Does expanding retrieval provide additional benefit? *Journal of Speech, Language, and Hearing Research, 67*, 1530–1547. https://doi.org/10.1044/2024_JSLHR-23-00528
- Leonard, L., Christ, S., Deevy, P., Karpicke, J., Weber, C., Haebig, E., Kueser, J., Souto, S., & Krok, W. (2021). A multi-study examination of the role of repeated spaced retrieval in the word learning of children with developmental language disorder. *Journal of Neurodevelopmental Disorders, 13*, 20. <https://doi.org/https://doi.org/10.1186/s11689-021-09368-z>
- Leonard, L., & Deevy, P. (2020). Retrieval practice and word learning in children with specific language impairment and their typically developing peers. *Journal of Speech, Language, and Hearing Research, 63*, 3252–3262. https://doi.org/10.1044/2020_JSLHR-20-00006
- Leonard, L., Deevy, P., Horvath, S., Christ, S. L., Karpicke, J., & Kueser, J. B. (2023). Can retrieval practice facilitate verb learning in children with developmental language disorder and their peers with typical language development? *Journal of Speech, Language, and Hearing Research, 66*, 2022_JSLHR-22-00509. https://doi.org/10.1044/2022_JSLHR-22-00509
- Leonard, L., Deevy, P., Karpicke, J., Christ, S., & Kueser, J. (2020). After initial retrieval practice, more retrieval produces better retention than more study in the word learning of children with developmental language disorder. *Journal of Speech, Language, and Hearing Research, 63*, 2763–2776. https://doi.org/10.1044/2020_JSLHR-20-00105
- Leonard, L., Deevy, P., Karpicke, J., Christ, S., Weber, C., Kueser, J., & Haebig, E. (2019). Adjective learning in young typically developing children and children with developmental language disorder: A retrieval-based approach. *Journal of Speech, Language, and Hearing Research, 62*, 4433–4449. https://doi.org/10.1044/2019_JSLHR-L-19-0221
- Leonard, L., Karpicke, J., Deevy, P., Weber, C., Christ, S., Haebig, E., Souto, S., Kueser, J., & Krok, W. (2019). Retrieval-based word learning in young typically developing children and children with developmental language disorder I: The benefits of repeated retrieval. *Journal of Speech, Language, and Hearing Research, 62*, 944–964. https://doi.org/10.1044/2018_JSLHR-L-18-0071
- Leonard, L., Kueser, J., Deevy, P., Haebig, E., Karpicke, J., & Weber, C. (2022). The contribution of immediate retrieval and spaced retrieval to word learning in preschoolers with developmental language disorder. *Autism and Developmental Language Impairments, 7*, <https://doi.org/10.1177/23969415221077652>
- Levlin, M., Wiklund-Hörnqvist, C., Sandgren, O., Karlsson, S., & Jonsson, B. (2022). Evaluating the effect of rich vocabulary instruction and retrieval practice on the classroom vocabulary skills of children with (developmental) language disorder. *Language, Speech, and Hearing Services in Schools, 53*, 542–560. https://doi.org/10.1044/2021_LSHSS-21-00101
- Lyle, K., & Crawford, N. (2011). Retrieving essential material at the end of lectures improves performance on statistics exams. *Teaching of Psychology, 38*, 94–97. <https://doi.org/10.1177/0098628311401587>
- Ma, X., Li, T., Duzi, K., Li, Z., Ma, X., Li, Y., & Zhou, A. (2020). Retrieval practice promotes pictorial learning in children aged six to seven years. *Psychological Reports, 123*, 2085–2100. <https://doi.org/10.1177/0033294119856553>
- Malins, J., Landi, N., Ryherd, K., Frijters, J., Magnusson, J., Rueckl, J., Pugh, K., Sevcik, R., & Morris, R. (2020). Is that

- a *pibu* or a *pibo*? Children with reading and language deficits show difficulties in learning and overnight consolidation of phonologically similar pseudowords. *Developmental Science*. <https://doi.org/10.1111/desc.13023>
- McDaniel, M., Howard, D., & Einstein, G. (2009). The read-recite-review study strategy: Effective and portable. *Psychological Science*, *20*, 516–522. <https://doi.org/10.1111/j.1467-9280.2009.02325.x>
- McGregor, K., Arbisi-Kelm, T., Eden, N., & Oleson, J. (2020). The word learning profile of adults with developmental language disorder. *Autism and Developmental Language Impairments*, *5*, 1–19. <https://doi.org/10.1177/2396941519899311>
- McGregor, K., Arbisi, T., & Eden, N. (2017). The encoding of word forms into memory may be challenging for college students with developmental language impairment. *International Journal of Speech-Language Pathology*, *19*, 43–57. <https://doi.org/10.3109/17549507.2016.1159337>
- McGregor, K., Gordon, K., Eden, N., Arbisi-Kelm, T., & Oleson, J. (2017). Encoding deficits impede word learning and memory in adults with developmental language disorders. *Journal of Speech, Language, and Hearing Research*, *60*, 2891–2905. https://doi.org/10.1044/2017_JSLHR-L-17-0031
- McGregor, K., Licandro, U., Arenas, R., Eden, N., Stiles, D., Bean, A., & Walker, E. (2013). Why words are hard for adults with developmental language impairments. *Journal of Speech, Language, and Hearing Research*, *56*, 1845–1856. [https://doi.org/10.1044/1092-4388\(2013\)12-0233](https://doi.org/10.1044/1092-4388(2013)12-0233)
- McGregor, K., Oleson, J., Bahnsen, A., & Duff, D. (2013). Children with developmental language impairment have vocabulary deficits characterized by limited breadth and depth. *International Journal of Language and Communication Disorders*, *48*, 307–319. <https://doi.org/10.1111/1460-6984.12008>
- McGregor, K., Owen Van Horne, A., Curran, M., Wagner Cook, S., & Cole, R. (2021). The challenge of rich vocabulary instruction for children with developmental language disorder. *Language, Speech, and Hearing Services in Schools*, *52*, 467–484. https://doi.org/10.1044/2020_LSHSS-20-00110
- Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., Vamvakas, G., & Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: Evidence from a population study. *Journal of Child Psychology and Psychiatry*, *57*, 1247–1257. <https://doi.org/10.1111/jcpp.12573>
- Pomper, R., McGregor, K., Arbisi-Kelm, T., Eden, N., & Ohlmann, N. (2022). Direct instruction improves word learning for children with developmental language disorder. *Journal of Speech, Language, and Hearing Research*, *65*, 4228–4249. https://doi.org/10.1044/2022_JSLHR-22-00300
- Pye, M., & Rawson, K. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? *Journal of Memory and Language*, *60*, 437–447. <https://doi.org/10.1016/j.jml.2009.01.004>
- Rice, M., & Hoffman, L. (2015). Predicting vocabulary growth in children with and without specific language impairment: A longitudinal study from 2; 6 to 21 years of age. *Journal of Speech, Language, and Hearing Research*, *58*, 345–359. https://doi.org/10.1044/2015_JSLHR-L-14-0150
- Roediger, H. L. I. I. I., & Karpicke, J. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*, 249–255. <https://doi.org/10.1111/j.1467-9280.2006.01693.x>
- Rohrer, D., Taylor, K., & Sholar, B. (2010). Tests enhance the transfer of learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 233–239. <https://doi.org/10.1037/a0017678>
- Rowland, C. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin*, *140*, 1432–1463. <https://doi.org/10.1037/a0037559>
- Rowland, C., & DeLosh, E. (2015). Mnemonic benefits of retrieval practice at short retention intervals. *Memory (Hove, England)*, *23*, 403–419. <https://doi.org/10.1080/09658211.2014.889710>
- Schimke, E., Angwin, A., Cheng, B., & Copeland, D. (2021). The effect of sleep on novel word learning in healthy adults: A systematic review and meta-analysis. *Psychonomic Bulletin and Review*, *28*, 1811–1838. <https://doi.org/10.3758/S13423-021-01980-3/Tables/6>
- Schoch, S., Cordi, M., & Rasch, B. (2017). Modulating influences of memory strength and sensitivity of the retrieval test on the detectability of the sleep consolidation effect. *Neurobiology of Learning and Memory*, *145*, 181–189. <https://doi.org/10.1016/j.nlm.2017.10.009>
- Smith, C., & Scarf, D. (2017). Spacing repetitions over long time-scales: A review and reconsolidation explanation. *Frontiers in Psychology*, *8*, 962. <https://doi.org/10.3389/fpsyg.2017.00962>
- Storkel, H., Voelmle, K., Fierro, V., Flake, K., Fleming, K., & Romine, R. S. (2017). Interactive book reading to accelerate word learning by kindergarten children with specific language impairment: Identifying an adequate intensity and variation in treatment response. *Language, Speech, and Hearing Services in Schools*, *48*, 16–30. https://doi.org/10.1044/2016_LSHSS-16-0014
- Tomblin, J. B. (2014). Educational and psychosocial outcomes of language impairment in kindergarten. In J. B. Tomblin, & M. A. Nippold (Eds.), *Understanding individual differences in language development across the school years* (pp. 166–203). Psychology Press.
- Vlach, H. (2019). Learning to remember words: Memory constraints as double-edged sword mechanisms of language development. *Child Development Perspectives*, *13*, 159–165. <https://doi.org/10.1111/cdep.12337>
- Walker, M. (2005). A refined model of sleep and the time course of memory formation. *Behavior and Brain Sciences*, *28*, 51–64. <https://doi.org/10.1017/S0140525X05000026>
- Whiffen, J., & Karpicke, J. (2017). The role of episodic context in retrieval practice effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*, 1036–1046. <https://doi.org/10.1037/xlm0000379>