

Official Title: Retrieval-Based Word Learning in Developmental Language Disorder: Adaptive Retrieval Schedule

Grant #: R01DC014708

NCT#: NCT06995014

Date: August 1, 2024

Definition: Type of uploaded study document. Select one.

- Study Protocol: The written description of the clinical study, including objective(s), design, and methods. It may also include relevant scientific background and statistical considerations (if the protocol document includes the statistical analysis plan, use "Study Protocol with SAP and/or ICF" option). Note: All amendments approved by a human subjects protection review board (if applicable), before the time of submission and that apply to all clinical trial Facility Locations must be included.
- Statistical Analysis Plan (SAP): The written description of the statistical considerations for analyzing the data collected in the study. Includes how data are analyzed, what specific statistical methods are used for each analysis, and how adjustments are made for testing multiple variables. If some analysis methods require critical assumptions, the written description should allow data users to understand how those assumptions were verified.

Purpose

Word learning is one of the principal weaknesses in children with developmental language disorder (DLD). Our previous work has shown that special benefits accrue when these children must frequently recall newly introduced words during the course of learning, however their absolute levels of learning are still low compared to typically developing peers. In this study, we seek to increase the children's absolute levels of learning by adjusting retrieval schedules such that they might improve the initial encoding of new word forms. We tested typically developing peers as a comparison group.

Participant selection

Preschool children with DLD are recruited through local speech-language pathologists and preschool teachers. Children with typical language development are recruited through advertisements in parent newsletters, preschools, and community events.

For interested families, a phone interview is used to screen for the following eligibility requirements:

- 1) current age between 4;0 and 5;11 (years; months)
- 2) monolingual English speaker; or bilingual speaker with significant English exposure from birth.
- 3) no significant neurological condition or neurodevelopmental disorder (e.g. ASD).

Informed consent is obtained at the first session and eligibility testing is then completed. To be eligible, the child must:

- 1) pass a pure tone hearing screening in both ears at 20 dB at 500, 1000, 2000, and 4000 Hz

- 2) score above 75 on a test of non-verbal intelligence (*Kaufman Assessment Battery for Children – Second Edition*).
- 3) complete the *Structured Photographic Expressive Language Test – Preschool 2* and receive a standard score in the following range:
 - a. To be included in the DLD group, children had a standard score below 87. This cutoff score has been found to show good sensitivity and specificity (Greenslade et al., 2009).
 - b. To be included in the TD group, children had a standard score above 87.

All children also complete the Peabody Picture Vocabulary Test – Fifth Edition. Parents complete a questionnaire that includes mother's total years of education. The PPVT standard scores and years of maternal education are both used as covariates in all statistical analyses.

For eligible children, 6 sessions are scheduled to complete the study.

Word learning procedure

The principal means by which we hoped to improve encoding and, as a result, subsequent recall, is to provide additional immediate retrieval trials after incorrect spaced retrieval trials in the adaptive condition in order to increase the likelihood of the child's success on the next spaced retrieval trial.

In the current study, we compare our “standard” procedure to an “adaptive” retrieval schedule in which additional immediate retrieval trials are provided when the child is unable to recall the word. So in both conditions, children first hear the word and what it likes and then are immediately asked to recall this information. For subsequent recall trials, three other words intervene between the last time the child heard the word and the time when the child is asked to recall it again. At this point, the two conditions diverge. In the standard condition, if the child is unable to recall the word, they hear the word presented again as feedback and continue on. In contrast, in the “adaptive” condition, if the child is unable to recall the word, they hear the word presented again and have another immediate opportunity to retrieve it. In both conditions, if the child is able to recall the word correctly, they hear the feedback and then continue to the next item in the spaced retrieval schedule.

The words in the two learning conditions - standard retrieval and adaptive retrieval - are presented sequentially in separate sets. The sets are counterbalanced in order. Each set involved two learning sessions held on consecutive days with a recall test administered five minutes after the second learning session, followed by another recall test one week later. The second set begins the next week, following the same scheme as the first set.

Four novel word “forms” are learned in each set. Colored photographs appearing on a laptop screen are the visual referents for these words. We also associate each word with something the referent “likes” (although ‘what it likes’ was not shown on the photograph). This serves as the “meaning” of the word. During each study trial, the novel word is mentioned three times and its meaning is mentioned once, as in: “This is a /nɛp/. It's a /nɛp/. A /nɛp/ likes rain”. For conditions that included retrieval trials, the child sees the picture and is asked, “What's this called?” and then “And what does this one like?” Audio stimuli is pre-recorded and presented along with the visual stimuli (scanned photographs) on the laptop.

The novel words are monosyllabic (consonant-vowel-consonant) and are matched across conditions according to phonotactic probability and neighborhood density (based on Storkel & Hoover, 2010).

The two learning sessions within each set are identical. For the set constituting the standard condition, each novel word is presented initially in a study trial, followed by an immediate retrieval trial, followed by another study trial. We refer to such retrieval trials as "0" retrieval trials, reflecting the fact that there are zero words intervening between a word's study trial and its retrieval trial. After each novel word has had a 0 trial, all subsequent retrieval trials occur after three other words have intervened since the last time the word appeared in a study trial. We refer to these retrieval trials as "3" retrieval trials, again, reflecting the number of intervening words used. Such trials are followed by a study trial for the same word. The retrieval schedule for each day for the words in the standard condition is 0-3-3-3-3-3.

For the set constituting the adaptive retrieval condition, each word first occurs in a study trial, followed by an immediate ("0") retrieval trial, and then another study trial. The next retrieval trial for a word occurs after three other novel words have intervened since the word's most recent study trial (a "3" trial). As in the standard condition, whether or not the child was able to retrieve the correct word, this retrieval trial is followed by a study/feedback trial. However in the adaptive condition, if the child is unsuccessful in retrieving the word form, this study trial will be followed by an additional immediate retrieval trial for that word. If, on the other hand the child retrieves the word form successfully, the next retrieval - study sequence will appear after three intervening words. In the standard condition, an unsuccessful recall attempt will be followed by a study trial but there will be no additional immediate retrieval opportunity. Note that while children will also be asked to retrieve the word meaning (what it likes), unsuccessful recall of this information will not trigger an additional immediate retrieval opportunity. This is based on previous findings showing that the real difficulty for children with DLD is recalling the phonetic form of words, rather than meaning information.

With this design, the two learning conditions provide the same number of exposures to the word forms and meanings, but differ in the number of retrieval opportunities in accordance with the individual child's pattern of retrieval success.

Outcome measures

Primary outcome measures:

Form Recall and Meaning Recall test at 5 minutes. For each set, after the child has completed the learning session of the second day, a five-minute break is given. Then a recall test of each word form and meaning is administered. Two items for each word form and meaning are used, with the second item for each word presented only after all four words had been tested once. Prompts used for the recall items are identical to those used in the retrieval trials ("What's this called?"; "What does this one like?").

Form Recall and Meaning Recall at one week. The Form and Meaning recall test is re-administered one week after the first day of learning.

Form Recognition at one week. Immediately following the one-week Form and Meaning Recall test, a picture-pointing recognition test is administered. The photo of each novel

word referent is shown along with the referents of the other three referents (in an array of four photos) and the child was asked, "Show me the e.g., /nɛp/."

Secondary outcome measures:

Peabody Picture Vocabulary Test – Fifth Edition is completed. Standard scores are obtained and used as a covariate in all statistical analyses.

Parents complete a questionnaire that includes mother's total years of education. Years of maternal education is used as a covariate in all statistical analyses.

Scoring

Word Form Recall measure: Eight items were used for the word form recall test for each set. Several steps are used in the scoring of children's word productions:

- 1) Productions of real words alternatives to the novel word (e.g., "cactus") are scored as incorrect.
- 2) Productions that appear to be attempts at the novel word but are not accurate are submitted to the scoring system of Edwards et al. (2004). Each consonant is awarded one point each for correct place, manner, and voicing. For vowels, one point is given for each of length, height, and backness. An additional point is credited for correct syllable shape (CVC). Given that all novel words had the syllable shape CVC, all fully adult-like pronunciations earned 10 points. For any non-adult-like production to be scored as correct, the production is required to have a higher point total than the total that would be given if the child is instead trying to produce one of the other novel words.

Meaning Recall measure: Eight items were used for the meaning recall test.

Scoring of the children's responses on the meaning recall test as correct or incorrect is straightforward, as pronunciation, even when non-adult-like, is not a factor in distinguishing among the possible meanings (e.g., grass, butterflies, rain, sun).

Form Recognition measure: Eight items were used for the form recognition test.

The child's picture pointing response is recorded at the time of test. The score for the recognition test was the number of items in which the child pointed to the correct photograph in the four-photograph array.

Data analysis plan

The first set of analyses examined the children's longer-term recall and recognition. Specifically, children's responses on the word form recall test, the meaning recall test, and the recognition test were evaluated using a series of mixed-effects models, with and without the covariates of PPVT-5 standard score and maternal education in years. The outcome was the number of correctly recalled items in a set of 8 items (as 2 items were used for each of 4 novel words). Diagnostic group (DLD, TD) was a between-participant variable; within-participant variables were learning condition (standard, adaptive), and time (5-minute, 1-week, for the word form and meaning tests only). Random slopes for learning condition and time were included in the models when they were not close to zero. As a result, the random slope for the learning condition variable was included in the word form recall and meaning recall models and the random slope for time was included in the meaning recall model.

Main effects models and full factorial models that included all possible two-way and three-way interactions were tested hierarchically.

Effect sizes are reported as partially standardized beta coefficients (b_{std}), which are comparable to a Cohen's d except they represent conditional standardized mean differences, conditioned on other variables in the model. Restricted maximum likelihood (REML) estimation was used. Bootstrapped standard errors with 1,000 replicates were used to account for non-normal error terms in the meaning recall and recognition models. Stata version 17.0 was used for mixed effects model analyses (StataCorp, 2019).