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

Grant #: R01DC014708

NCT#: NCT05325333

Date: March 1, 2022

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 include 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 create a more gradual spacing of retrieval trials as the learning period proceeds.

In the current study, we compare our "standard" procedure to a gradual or "expanding" retrieval schedule. In both conditions, children first hear each word and what it likes, and then are immediately asked to recall this information. After these trials, the schedule for the two conditions diverge. For subsequent recall in the "standard" condition, three other words intervene between the last time the child heard the word and the time when the child is asked to recall it. In contrast, in the "expanding" condition, the spacing between study and subsequent retrieval of the word is initially narrow -- one intervening word - and expands more gradually to three intervening words.

The words in the two learning conditions - standard retrieval and expanding 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? What do we call this?" and then "And what does this one like? What does it 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. Such trials are followed by a study trial for the same word. We refer to these retrieval trials as "3" retrieval trials, again, reflecting the number of intervening words used. The retrieval schedule for each day for the words in the standard condition is 0-3-3-3-3. However, after the second 3 trial of the four words, we insert one study trial for each word. This study trial is intended as a "refresher" under the assumption that some words will not yet be recalled at that point. The insertion of these study trials does not change the 3 spacing between a word's retrieval trial and its previous study trial. The next day has the 0-3-3-3-3 schedule with a “refresher” study trial after the second 3 trial.

For the set constituting the expanding retrieval condition, each word first occurs in a study trial, followed by an immediate retrieval trial, and then another study trial. The next two retrieval trials occur after only one other novel word has intervened since the word's most recent study trial. These retrieval trials were followed by a study trial. We refer to these retrieval trials as "1" retrieval trials, as there is one intervening word. The two remaining retrieval trials for words in this condition are 3 retrieval trials, with each retrieval trial followed by a study trial for the same word. The schedule for each day, then, is 0-1-1-3-3. After the second 1 retrieval trial for the four words in the set, a study trial is presented for each word as a "refresher," as was done for the words in the standard condition.

With this design, the two learning conditions provide the same number of exposures to the word forms and meanings, and the same number of retrieval opportunities throughout the two-day learning period. The two conditions differ only in the retrieval schedule - 0-3-3-3-3 versus 0-1-1-3-3.

See https://pubs.asha.org/doi/full/10.1044/2024_JSLHR-23-00528 Appendix A and B for an illustration of the details of words and learning schedules.

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 (01133, 03333), 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).

Collapsing across sets, each measure has a range of 0-8 in each of the two conditions.

We hypothesized that the use of “1” retrieval trials prior to proceeding the 3 retrieval trials would increase the children's ability to encode the novel word forms and meanings. Accordingly, we expected that recall and recognition scores would be higher for words in the expanding 0-1-1-3-3 condition than in the standard 0-3-3-3-3 condition. This was a rather stringent test given that our previous work has shown that the standard condition was more successful than comparison conditions that did not involve spacing. Nevertheless, it was necessary to attempt to improve on the standard condition given, using this retrieval schedule, children were unable to learn more than approximately half of the words employed.

We expected that the expanding condition would lead to better results than the standard condition for both the children with DLD and the children with TD. In our earlier work, group differences were often seen favoring the TD children for the standard condition but the difference was smaller than the group differences seen for the comparison conditions. The current study involve two conditions that each involved spaced retrieval, with only the particular retrieval schedule distinguishing the two. We expected that the expanding condition would show narrower differences between the two groups of children than the standard condition.