Published in final edited form as:

Int J Speech Lang Pathol. 2017 February ; 19(1): 43–57. doi:10.3109/17549507.2016.1159337.

The encoding of word forms into memory may be challenging for college students with developmental language impairment

KARLA MCGREGOR, TIM ARBISI-KELM, and NICHOLE EDEN

University of Iowa, Iowa City, USA

Abstract

Purpose—To describe the word-learning problems characteristic of developmental language impairment (LI).

Method—College students with LI ($n = 39$) or normal language development (ND, $n = 40$) attempted to learn novel word forms. Training for half of the words was meaning-focused; training for the other half was form-focused. Form recognition and stem completion tasks administered immediately after training tapped encoding of the lexical configuration, and a repetition of the stem completion task one week later tapped consolidation. A visual world paradigm tapped lexical engagement.

Result—At the immediate posttest, the LI group was poorer at recognition and completion of word forms than their ND peers, suggesting a deficit in encoding the lexical configuration. However, the gap between the LI and ND groups in stem completion did not grow over the week, suggesting intact consolidation. Form-focused training yielded better performance than meaning-focused training at immediate- and one-week tests. For both groups, newly trained words slowed the recognition of familiar English words, revealing lexical engagement.

Conclusion—The encoding of word-form configurations is challenging for some, but not all, college students with LI. Training that encourages a focus on the form may be a useful part of vocabulary intervention for those affected.

Developmental language impairment, also known as primary or specific language impairment (LI), can present life-long challenges. Deficits in the learning, comprehending, producing, and defining of words are symptomatic even in the adolescent and adult years (Blalock, 1982; Korkman & Pesonen, 1994; McGregor, Oleson, Bahnsen, & Duff, 2013). These deficits stand to undermine functional ability in daily life. Post-secondary students with LI demonstrate vocabulary deficits that limit understanding of their rights as citizens (Rost & McGregor, 2013) and their learning of new information from classroom lectures (Becker & McGregor, 2015).

To describe the nature of these deficits and, ultimately, to remediate them, it is critical to understand what the problem is and why the problem occurs. A useful framework for

Correspondence: Karla McGregor, University of Iowa, 121C SHC, Iowa City, IA 52242, USA. karla-mcgregor@uiowa.edu.

Declaration of interest

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this paper.

considering the *what* is the distinction between lexical configuration and lexical engagement (Leach & Samuel, 2007). Lexical configuration is the set of information one knows about a given word including its form and meaning whereas the term lexical engagement captures the interactions that occur between words as evident in phenomena such as spreading activation and competition effects. The former is more often studied than the latter.

A useful framework for considering the *why* is the distinction between encoding and consolidation (Walker, 2005). Encoding is the set of processes that establish or update a memory trace whereas consolidation captures the stabilisation or even enhancement of that memory over time. A key difference is that input (e.g., hearing or reading a word) is necessary for encoding but not for consolidation. The passage of time and, in some cases sleep, are necessary for consolidation (Gais & Born, 2004). Behavioural evidence of the encoding of a word is typically gathered by administering tests immediately after the initial experience with the word (McGregor, Licandro, Arenas, Eden, Stiles, Bean, & Walker, 2013; Storkel, 2015). Behavioural evidence of the memory evolution of a word is gathered hours, days, weeks, or more after the initial experience with the word. Relative to an immediate posttest, consolidation will manifest as stable or improved performance.

Configuration

In a recent study, young adults with LI had difficulty learning lexical configurations and the form aspects of the configuration were particularly problematic (McGregor, Licandro, et al., 2013). Specifically, we found that, after equivalent training exposures to 16 novel words and their pictured referents, participants with LI recalled nearly four words, on average, while matched participants with normal language development (ND) recalled six (McGregor, Licandro, et al., 2013). Moreover, as they took tests again 12 hours, one day, and one week after training, the gap between the two groups of participants widened on form, but not meaning. The ND group improved, but the LI group remained steady; they recalled approximately the same number of word forms at each test. This finding corroborates evidence found with younger age groups. Namely, while children with LI may find the learning of word meanings to be challenging (Alt, Plante, & Creusere, 2004; McGregor, Newman, Reilly, & Capone, 2002), the learning of word forms presents a greater obstacle (Gray, 2004; McGregor, Licandro, et al., 2013).

In McGregor, Licandro, et al. (2013), the experience was a training protocol that involved exposures to 16 words and their referents. The gap between the LI and ND groups immediately after training led us to conclude that encoding was a deficient process in learning the word form configuration. Additional evidence supports this conclusion. Young adults with LI have difficulty with verbatim recall of familiar word forms immediately after exposure (Sheng, Byrd, McGregor, Zimmerman, & Bludau, 2015). Also, young children with LI require roughly twice the amount of exposure to learn new word forms as well as their unaffected peers (Gray, 2003).

The widening gap between LD and ND groups over time reported in McGregor, Licandro, et al. (2013) suggests that, in addition to deficits in encoding, individuals with LD also struggle with consolidation. However, that study was not designed to differentiate encoding and

consolidation deficits. The 12-hour and 1-day posttests included a two-alternative forced choice (2AFC) recognition probe and a word association probe, both of which necessitated presentations of the trained words. These additional experiences with the words were, of course, additional opportunities for encoding and retrieval practice (Roediger, & Karpicke, 2006). Therefore, we do not know whether encoding alone was a deficient mechanism, or whether encoding plus consolidation over a retention period were deficient. The extant literature is also equivocal. Alt (Alt, 2011; Alt & Spaulding, 2011) found evidence that children with developmental LI had problems encoding new word forms, especially long words, but the word forms that were encoded were not unusually susceptible to decay. In contrast, Rice, Oetting, Marquis, Bode and Pae (1994) found evidence of both encoding and consolidation problems in this population.

Engagement

Words do not simply reside in the lexicon as isolated configurations of facts; instead, they engage with other words (Leach & Samuel, 2007). For example, Gaskell and Dumay (2003) exposed adults to novel words (e.g. *cathedruke*) that were cohorts of real words (e.g. *cathedral*) over a period of five days. By the third day, reaction times to identify the real words had slowed, presumably because the trained words were competing with the real words for activation.

The visual world paradigm offers a means of observing lexical competition (see Huettig, Rommers, & Meyer, 2011 for a review). For example, when participants viewed the pictured objects *beaker*, *beetle*, *speaker*, and *carriage* on a computer screen and were asked to find the *beaker*, they tended to look at both the *beaker* and the *beetle* as the word *beaker* unfolded. By the second syllable, looks to *beetle* waned as the participants resolved the competition and focused in on the correct target (Allopenna, Magnuson, & Tanenhaus, 1998). In this way, the visual world paradigm reveals lexical competition effects in real time. The competition effects can be measured via eye tracking (e.g. the number of saccades to the cohort vs. the noncohort) (Allopenna et al., 1998) or mouse tracking (e.g. response times to mouse click on the target in the presence of a cohort vs. a noncohort) (Spivey, Grosjean, & Knoblich, 2005).

McMurray, Samelson, Lee, and Tomblin (2010) investigated lexical competition in adolescents with and without LI via a visual world paradigm. Each trial involved four pictured items named by familiar English words, a target (e.g., sandal), a cohort (e.g., sandwich), a rhyme (e.g., candle), and an unrelated foil (e.g., parrot). The two groups demonstrated similar competition effects in that, upon hearing the target name, both looked more to the cohort than the unrelated foil (looks to the rhyme fell between the two) and looks to the cohort declined as the listener found and clicked on the target picture. However, there were some subtle differences in that the listeners with LI looked more to the cohort and rhyme than did the listeners without LI. One hypothesised explanation is that, for those affected by LI, the activation of target words decays quickly, and thus they do not fully inhibit competitors (Rigler, Farris-Trimble, Greiner, Walker, Tomblin, McMurray, 2015).

Whereas McMurray et al. (2010) examined competition between familiar words, the current focus is on whether people with LI demonstrate lexical competition between a newly learned word and a familiar English neighbour, and, if so, whether the competition effects are as robust as those seen in learners with ND. The visual world paradigm does not require metalinguistic judgements, which may be difficult for people with LI, or any overt production, which, in the case of a newly learned word, can be difficult for anyone. Thus, it is an effective way to determine whether newly learned words compete with other, familiar words in the lexicon.

Current Study

To describe the nature of the word learning deficits that characterise LI in young adults, we trained college students with LI and ND on novel words and then compared the outcomes.

Word form configuration

To measure the learning of the word form configuration, we administered probes of word recognition and stem completion immediately after vocabulary training. We also repeated the stem completion probe one week later. Given previous findings, we hypothesised that the word learning challenges associated with LI reflect deficient encoding of word forms into long term memory in response to the input. We predicted that, when compared to peers with ND, college students with LI would demonstrate poorer performance on both probes immediately after training. Given equivocal findings in the extant literature, we tested the hypothesis that the word learning challenges associated with LI also reflect the poor consolidation of word form configurations. If so, college students with LI would demonstrate a widening gap in word form completion over the course of one week in comparison to peers with ND.

Anticipating poorer encoding of word forms on the part of learners with LI, we also asked whether vocabulary training that is form- rather than meaning-focused would serve to close the LI-ND encoding gap. To answer this question, each participant learned half of the words in a training protocol that focused on word form and half in a training protocol that focused on word meaning. This comparison was clinically motivated. If individuals with LI have difficulty encoding word forms, it becomes necessary to know how best to structure learning trials to promote encoding of that particular sort of information. Past work suggests that exposure to word forms alone before pairing them with meanings can be advantageous to encoding (Swingley, 2007) as can training that requires learners to produce the to-be-learned word forms (Leach & Samuel, 2007). On the other hand, there is evidence that word-form encoding is sometimes better when those forms are paired with semantic referents rather than presented alone (Hawkins, Astle, & Rastle, 2015) or when they are linked to richer rather than sparser meanings (Hargreaves, Pexman, Johnson, & Zdrzilova, 2012).

Finally, we asked whether an additional test exposure administered the day after training would serve to close the LI-ND encoding gap. Half of the participants were randomly assigned to receive additional input exposures via testing 24 hours after training. The others had no exposures in the interim between training and 1-week posttest. This comparison was motivated by McGregor, Licandro, et al. (2013). In that study, the gap between LI and ND

groups widened over the week, but we were unsure whether this was due to poor encoding from repeated testing or a genuine problem with memory consolidation. By providing only half of the participants with an opportunity for encoding via testing, we hoped to reach a firmer conclusion.

Engagement

We tested the hypothesis that the word learning problems of students with LI extend to engagement by administering a visual world paradigm with mouse tracking one week after vocabulary training. If the hypothesis held, compared to peers with ND, college students with LI would demonstrate smaller competition effects between English words and newly trained cohorts.

There is currently a debate as to whether lexical competition can emerge upon encoding (Fernandes, Kolinsky, & Ventura, 2009; Kapnoula & McMurray, 2015) or whether a period of consolidation is necessary (Bakker, Takashima, van Hell, Janzen, & McQueen, 2014; Dumay & Gaskell, 2007). We hedged our bets by situating the visual world paradigm one week after training, thereby to capture competition that had arisen earlier during encoding or later after consolidation. However, the additional test exposure given to half of the participants allows us to contribute to this debate. If competition effects emerge upon encoding, then more encoding opportunities should make for more robust competition than fewer encoding opportunities.

Method

Participants

Participants were 39 young adults with LI and 40 with ND recruited from the Midwest of the USA. In the LI group, there were 20 people who identified as women and 19 as men; one of the 39 was Hispanic or Latina; 32 were Caucasian, 1 was Asian, 1 was African American, and five were mixed races. In the ND group, there were 20 people who identified as women and 20 as men; one of the 40 was Hispanic or Latino; 33 were Caucasian, three were Asian, one was Native American, one was African American, one was mixed race, and one did not report. Except a single participant in the ND group who had recently completed a 4-year degree, all were currently enrolled in post-secondary institutions. The LI and ND groups were similar in types of post-secondary institutions represented (university, four-year college, community college).

All of the participants in the LI group had an LI diagnosis before the study. Those in the ND group reported no history of language or learning difficulties. To verify the self-reported diagnoses, we used a procedure and weighted score cut-off known to maximise the sensitivity and specificity of identification of LI in young adults (Fidler, Plante, & Vance, 2011). The procedure involves a modified token test that taps the comprehension of syntax (Morice & McNicol, 1985) and a spelling test that taps the knowledge of irregularly spelled words (Fidler et al., 2011). Eighteen people with previously identified LI did not meet the cutoff for language-based LI on this verification measure and one person with no history of LI did. We did not enroll any of these individuals in the study. One version of the Fidler et

al. (2011) procedure includes a word definition test. We administered this test but did not use it for the purpose of verifying LI status because we did not wish to bias our findings by recruiting people with known vocabulary problems.

We administered the nonverbal matrices of the Kaufman Brief Intelligence Test (KBIT; Kaufman & Kaufman, 2004) and a pure-tone audiometric screening at 0.5, 1, 2, and 4 kHz at 25 dB bilaterally. All participants were required to earn a standard score of 85 or better on the KBIT and to pass the audiometric screening to enroll in the study. English as the primary language was an additional selection criterion.

To further describe the sample, we administered various probes and standardised tests of memory and vocabulary (Table I). Consistently, the LI group scored lower than the ND group and the effect sizes were large; however, as might be expected of students with LI who compensate well enough to enter post-secondary studies, the LI group always scored within one standard deviation of the normative mean on the standardised tests.

Stimuli

A female native speaker of Standard American English recorded all stimulus words and task instructions in a sound-treated booth using a Larson-Davis 2560 0.5-inch random incidence microphone and a digitised sampling rate of 44,110 kHz with 16-bit resolution.

Novel word stimuli were 36 disyllabic words of CVCVC structure (Appendix A). Each novel word was constructed to be a cohort of an English word with overlap on the first syllable (e.g. *bukuz/bucket*) and the English words were selected because they have a CVCVC structure, and they are typically acquired before age 10 (Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012). The 36 novel words comprised three sets of 12. Sets were assigned in counterbalanced fashion across participants to be trained in the meaning-focused condition, trained in the form-focused condition, or untrained. When a set was assigned to the meaning-focused condition, the words in that set were paired with 12 colour photographs of unusual objects that have no commonly known English name. Half of the referents were animate, and half were inanimate. The English base words and pictures of their referents served as familiar stimuli in the visual world paradigm.

Procedure

Procedures followed the ethical standards of the University of Iowa Internal Review Board and the Helsinki Declaration of 1975, as revised in 1983. Participants attended three (- interim test) or four (+ interim test) individual sessions of one to two hours in length (Table II). During the first session, they gave informed consent and took tests to determine whether or not they met enrolment criteria and to describe their language abilities.

Training—During the second visit, participants were trained via computer presentation on the 24 target words, 12 in the meaning-focused condition and 12 in the form-focused condition. The timing parameters of the two training conditions were identical. The order of conditions was counterbalanced across participants.

In both conditions, a complete training cycle involved one presentation of the spoken and written word during each of six blocks. Each cycle repeated so that the total number of exposures to each of the 24 words was 12. The order of presentation was randomised within blocks and the order of blocks was counterbalanced across participants.

In the meaning-focused condition, each word was presented in spoken and written form simultaneously with a pictured referent. During each block, participants were asked to monitor the referents for a given feature (Appendix B). For example, in the first block, they determined whether or not the pictured object had eyes. After each presentation of a word-picture pair, the word and picture would disappear, and the question 'did it have eyes?' appeared with the options 'yes' and 'no' to click. We did not give any feedback on accuracy. The process continued with a new question per block. The purpose of the questions was to maintain attention to the training and to emphasise meaningful aspects of the referents. In the form-focused training, the spoken and written forms of the word were presented simultaneously, but there was no pictured referent. During each block, participants were asked to monitor the words for a given sound. For example, in the first block, they determined whether or not the word included a 'p' sound. The purpose was to maintain attention to training and to emphasise the form of the words to be learned.

Immediate test—Immediately after training, participants were given a 3AFC word recognition probe. Participants heard each trained novel word (e.g. bʌkəz) and two untrained novel words that differed from the trained word on the medial segment (e.g. bʌrəz) or the final segment (e.g. bʌkəf). The participants pressed key 1, 2, or 3 to indicate which of the three was in their study set. The order of trials was randomised to mix words trained in the meaning- and form-based conditions. The order of target and foils varied across trials. The dependent variable was the proportion of items accurately recognised.

Next, the participants completed a letter identification probe. Due to experimenter error, the results were uninterpretable and will not be presented. However, to allow a complete picture of the procedure, we will briefly describe the task. On each of 60 critical trials, participants saw a 5-letter string or a single letter embedded in a string of four asterisks for 20 msec. Each of the 24 words in the training set appeared once. Before each critical trial, participants saw arrows pointing to a blank space where the target letter would appear. Following each critical trial, they saw a pattern mask (asterisks in boxes) in each of the five letter positions. Participants then typed in the letter that had been in the location marked by the arrows. The presentation was so quick that participants were often unsure that they saw anything; nevertheless, performance was high for both groups.

The participants were then given a stem completion probe presented in natural voice via computer. Participants heard the first syllable of each trained novel word and then attempted to say the whole word. The order of trials was randomised to mix words trained in the meaning- and form-based conditions. Responses were transcribed phonetically. To determine reliability, two independent transcribers compared their transcriptions of 20% of the data, randomly selected. The phoneme-to-phoneme agreement was 90.55%. The dependent variable was the proportion of segments correctly produced per final syllable.

Twenty-four-hour test—Participants assigned to the + interim test condition returned 24 hours later to repeat the letter identification and stem completion probes. The procedure was identical to that used on day one.

One-week test—All participants returned one week later to repeat the letter identification and stem completion probes. Immediately after the stem completion probe, the participants were assessed with the visual world paradigm. Because this measure involves picture identification, only words trained in the meaning-focused condition were tested.

On each trial, participants saw two pictures: one in the upper right corner and the other in the upper left of a computer screen. One was the target, and one was the foil. The targets and the foils were referents named by familiar English words or trained words. On half of the trials, the foils were lexical cohorts of the targets (e.g. *bacon* and *baker*) and on half they were not (e.g. *bacon* and *rocket*). The balanced crossing of the language of the target (English or trained), the language of the foil (English or trained), and the cohort status of the foil (cohort or non-cohort) for the twelve words trained in the meaning-focused condition resulted in 96 trials (Appendix A). Participants cycled through the 96 randomly ordered trials five times for a total of 480 responses. On each trial, they moved a computer mouse from the bottom middle of the screen up a 'runway'. Upon reaching the top of the runway, they heard a natural human voice name of one of the pictures, and they continued the mouse trajectory to click on that picture.

The paradigm elicits competition effects based on similarity of form. The critical trials, then, are those that determine whether the trained word competes with or slows the processing of a familiar English word in the lexicon. Therefore, we compared the RTs to familiar English words in the presence of trained foils that were cohorts (e.g. *bukuz* and *bucket*) and in the presence of trained foils that were non-cohorts (e.g. *bukuz* and *kettle*). Trials that involve trained targets with familiar English foils were included solely to prevent the listener from developing an expectation that all correct answers would be items with English names. These were not included in the statistical analysis because one can already assume that familiar English words exist in the long-term lexicon and that they compete with cohorts.

The relationship between memory processes and extant vocabulary—A final step in the procedure was to determine whether the memory processes of encoding and consolidation related to extant vocabulary. Therefore, we selected the standardised vocabulary test that yielded the largest range of scores (the Expressive Vocabulary Test, EVT) and correlated those scores with scores on the probes administered immediately after training and one week later. The goal was to link further any deficient processes to manifestations of vocabulary impairment.

Result

Word Form Configuration

Encoding—We first measured the encoding of the word form configuration via a word recognition probe. An application of the Shapiro-Wilk's *W* test revealed that the accuracy data from the 3AFC word recognition probe for the ND group violated assumptions of

normality. Therefore, we compared the recognition performance of the two diagnostic groups via two nonparametric Mann-Whitney U tests. The groups differed on the accuracy of items trained via form-focused training, $U = 419.5$, $z = 3.53$, $p = .0004$, and those trained via meaning-focused training, $U = 562.5$, $z = 2.13$, $p = .03$. In both cases, the ND group outperformed the LI group (Figure 1).

The ranges of the scores were nearly complete in their overlap. The difference between the diagnostic groups was driven largely by the higher number of ceiling-level scores ($> .90$) in the ND group. Note one particularly low score of .25 in the form-focused condition. With this outlier removed from the analysis, the results of the Mann-Whitney U test remained the same, $U = 419.5$, $z = 3.40$, $p = .0007$.

Individual differences were evident. The variation in recognition of newly trained word forms (collapsed across training conditions) was positively correlated with extant vocabulary as measured by the EVT in the LI group, Pearson product-moment correlation $r = .33$, $p = .04$, but not in the ND group, Spearman rank order correlation = .13, $p = .13$. The null effect may reflect the number of near ceiling-level scores in the ND group.

Next we measured the encoding of the word form configuration via a stem completion probe. Immediately after training, the participants in the LI group completed 4% of all words correctly ($SE = 2\%$) while the ND group completed 13% of all words correctly ($SE = 2\%$). Clearly, this was a difficult probe for both groups. Therefore, we applied a finer level of measurement to determine whether partially correct productions distinguished the groups. The dependent variable was the mean proportion of phoneme-sized segments that were correctly produced (out of three possible) in the final syllable of each target word. Levene's Test revealed unequal variances, so this value was square root transformed. The transformed values met the assumption of equal variances, so these constituted the dependent variable in an ANOVA with diagnostic group (LI, ND) as a between-subjects variable and training focus (form-focused, meaning-focused) as a within-subject variable; however, untransformed values are reported and graphed for ease of understanding.

The result was a main effect of group, $F(1,77) = 15.19$, $p = .0002$, $\eta^2_p = .16$. The ND group produced a mean of $.33 \pm .02$ of all final syllable segments correctly whereas the LI group produced $.21 \pm .02$. Given that three segments comprised all final syllable targets, this means that the ND group remembered, on average, one of the three while the LI group remembered less than one. There was also a main effect of training, $F(1,77) = 21.65$, $p < .0001$, $\eta^2_p = .22$. Both groups performed better following form-focused training. The participants produced a mean of $.31 \pm .02$ of all segments correctly following form-focused training and $.24 \pm .02$ following meaning-focused training.

Figures 2 and 3 (immediate post-test) reveal extensive individual differences within groups. The variation in the production of newly trained word forms (collapsed across training conditions) was positively correlated with extant vocabulary as measured by the EVT in the LI group, Pearson product-moment correlation $r = .33$, $p = .04$, and in the ND group, Pearson product-moment correlation $r = .45$, $p = .003$.

Consolidation—To examine the consolidation of the word form configuration, we compared stem completion performance one week after training to performance immediately after training (Figures 2 and 3 immediate and 1-week post-test). We applied an ANOVA with diagnostic group (LI, ND) and schedule (with interim test exposure, without interim test exposure) as between-subjects variables, and time of the test (immediate, 1-week) and training focus (form-focused, meaning-focused) as within-subject variables. Again the dependent variable was the mean proportion of correctly produced segments in the final syllable of each target word. Although the statistical analysis employed the square root-transformed data; the Figures depict untransformed data for ease of interpretation.

The result was a main effect of diagnostic group, $F(1,75) = 15.74$, $p = .0002$, $\eta^2_p = .17$, with the ND group outperforming the LI group, and a main effect of training, $F(1,75) = 24.10$, $p < .0001$, $\eta^2_p = .24$, with the form-focused training yielding better performance than meaning-focused training. There were no main effects of schedule (the interim test exposure did not affect overall performance at the 1-week test), $F(1,75) < 1$, or time (performance did not vary significantly from immediately after training to 1-week post training), $F(1,75) < 1$. There were no significant interaction effects.

A high degree of overlap between groups is apparent in Figures 2 and 3. Fewer very low performances and more very high performances by some individuals in the ND group drove the difference between diagnostic groups. Although, at the group level, performance was stable over the course of one week, within each group some participants demonstrated stable performance while others demonstrated improvements or declines. To determine whether individual differences in consolidation related to differences in extant vocabulary, we derived a difference score (performance at one-week post-test minus performance at immediate post-test, collapsed across training conditions) and tested for correlations between the difference score and scores on the EVT. There was no correlation for the LI group, Pearson product-moment correlation $r = .01$, $p = .95$, or the ND group, Pearson product-moment correlation $r = -.23$, $p = .15$.

Engagement

To examine engagement, we analysed the results of the visual world paradigm. The dependent variable was the median RT in msec for correct identifications of the referents being named by the English words. Accuracy was high for the LI group ($M = 98.3\%$, $SD = 1$) and the ND group ($M = 98.7\%$, $SD = 2$). There were no violations of homogeneity of variances or normality. Figure 4 illustrates the critical result: the LI group, like the ND group, was longer to respond in the context of a trained cohort than a trained non-cohort, LI $t = 2.6$, $df = 38$, $p = .01$, $d = .19$, ND $t = 5.3$, $df = 39$, $p < .001$, $d = .40$.

To further explore the variables that affected RT, we applied an ANOVA with diagnostic group (LI, ND) and schedule (with the interim test, without the interim test) as between-subjects variables, and language of foil (English or trained), and cohort status of foil (cohort, non-cohort) as within-subject variables. The result was a main effect of diagnostic group, $F(1,75) = 8.02$, $p = .006$, $\eta^2_p = .10$, with the ND group responding more quickly than the LI group (Figure 4). There was a main effect of language status, $F(1,75) = 76.74$, $p < .0001$, $\eta^2_p = .51$, with English foils yielding longer RTs than trained foils. There was a main effect

of cohort status, $F(1,75) = 158.63$, $p < .0001$, $\eta^2_p = .68$, with cohort foils yielding longer RTs than non-cohort foils. There was a language $\times$ cohort interaction, $F(1,75) = 55.38$, $p < .0001$, $\eta^2_p = .43$. A Bonferroni posthoc test revealed that the RTs to English and trained foils that were non-cohorts did not differ, but the RTs to English and trained foils that were cohorts did differ, such that RTs were longer in the presence of English cohorts. In other words, English cohorts elicited more robust competition than trained cohorts.

There was a marginal interaction between diagnostic group and cohort, $F(1,75) = 3.30$, $p = .07$, $\eta^2_p = .04$. Because we were primarily interested in whether the LI group had intact competition, we explored this trend. A posthoc Bonferroni test revealed that both groups took significantly longer to identify targets in the presence of cohort foils than noncohort foils, but the gap between RT in the cohort and non-cohort conditions was larger for the ND group than the LI group. In other words, the ND group trended toward more robust competition effects.

There was no significant main effect of schedule, $F(1,75) = 2.87$, $p = .09$, but there was a schedule $\times$ cohort interaction, $F(1,75) = 6.60$, $p = .01$, $\eta^2_p = .08$. A posthoc Bonferroni test revealed that it took significantly longer to identify targets in the presence of cohort foils than non-cohort foils in both schedules, but the gap in RT was larger for individuals who received the interim test schedule (Figure 5). In other words, the additional test exposure elicited more robust competition effects.

To determine whether individual differences in competition related to differences in extant vocabulary, we derived a difference score (RT to identify English targets in the context of trained cohort foils minus RT to identify English targets in the context of trained non-cohort foils) and tested for correlations between the difference score and scores on the EVT. There was no correlation for the LI group, Pearson product-moment correlation $r = .08$, $p = .65$, but there was for the ND group, Pearson product-moment correlation $r = .33$, $p = .04$. Examination of a scatterplot revealed an outlier in the ND group. The participant with the largest raw score on the EVT had an extreme difference score on the visual world paradigm. With this participant omitted, the correlation for the ND group was similar to that of the LI group, and it was no longer significant, $r = .10$, $p = .56$.

Discussion

Configuration

In the current study, college students with LI recognised and produced fewer word forms than their peers with ND on probes administered immediately after training. That said, they recognised the forms with 70–80% accuracy on a probe where the chance of accurate guessing was 33%. So clearly they did encode some word form information. The gap between the LI and ND groups was especially apparent on the stem completion probe where the ND group correctly completed three times as many words as the LI group. This study confirms, with an independent sample, the word-form encoding deficits that we reported in McGregor, Licandro, et al. (2013).

Producing these newly trained word forms was a challenge for all participants. The extent of the challenge will undoubtedly vary with the number of targets, here 24, and the amount of exposure, here 12 times each. Also, by designing the stimuli to be disyllables, we necessarily created words with low neighbourhood density. Words with few similar sounding neighbours tend to be harder to learn than those with many similar sounding neighbours (Storkel, Armbruster, & Hogan, 2006). Finally, the context for learning was not particularly rich. Even in the meaning-focused condition, semantic information was conveyed via a single photograph per target. Words with fewer semantic features tend to be harder to learn than words with richer meanings (Hargreaves et al., 2012).

Manipulating variables such as the number of targets, the dosage of exposure, and the nature of the training context can make learning easier (or harder), but it will likely do so for any learner, not only those with LI. In the current study, both groups demonstrated better encoding of word form information in response to the form-focused training. The implication is not to strip vocabulary lessons of meaning. Words without meanings are useless. Moreover, aspects of word learning such as engagement would likely suffer in the absence of meaning (Leah & Samuel, 2007). However, asking learners to focus on word-form information during vocabulary lessons enhances the encoding of that information into memory. The implication is that this may be an important component of training for learners who are especially challenged by spoken and written form. The best combination of form-focused and meaning-focused training, the optimal order, and the optimal dosage are unknowns awaiting exploration.

In McGregor, Licandro et al., (2013), the gap between LI and ND groups in word form production widened over the course of one week. There were two possible explanations. Either the ND group benefited from the further encoding opportunities given 12- and 24 hours after training while the LI group did not, or the LI group presented with deficient consolidation. In the current study, we found no support for the latter. With interim test exposures eliminated or limited to one at 24-hours post-training, the gap between the LI and ND groups remained steady. Consolidation was not problematic for this group of college students with LI. Alt and Spaulding (2011) found school children with LI less able than age-mates with ND to detect mispronunciations of newly trained word forms immediately after training but equally able to maintain initial performance levels over a 10-second delay. Although 10 seconds is not enough time for the memory to consolidate, this pattern is consistent with the current finding that individuals with LI have difficulty encoding word forms into memory but not with maintaining those that do get encoded.

We did expect that participants with the interim test exposure would fare better than those without by week's end. On the stem completion probe, this was not the case. Recall that the interim test included stem completion and masked letter identification. Performance on the stem completion probe was low, and there was no feedback. As such, these likely provided minimal learning benefits (see summary and exceptions in Richland, Kornell, & Kao, 2009). Performance on the masked letter identification involved exposure to the target words forms that was so fast as to be not consciously accessible. Thus learning failed, at least learning that was robust enough to support production (Gray, 2003, 2004). In fact, the interim test did

improve learning as measured by a probe that did not require production, the visual world paradigm. We now turn to those results.

Engagement

Using a visual world paradigm, we found both LI and ND groups to demonstrate clear patterns of lexical competition between familiar English words with overlapping onsets, such as *bacon* and *baker*. We also found both groups to demonstrate lexical competition between familiar English words and trained words with overlapping onsets, such as *cigar* and *cigub*. Recall that McMurray et al. (2010) also found clear evidence of competition between targets and cohorts among people with LI.

The level of competition between English and trained cohorts was not as robust as between English and English cohorts. The English words that we selected are acquired before age 10; therefore, the participants had likely been exposed to them many times. Words that are more frequent in the language produce greater competition effects than those that are less frequent (Dahan, Magnuson, & Tanenhaus, 2001) and frequency of occurrence is a reasonable proxy for exposure. The finding that English cohorts were stronger competitors than trained cohorts is further informed by the interaction between cohort type and interim schedule; those who received the interim test exposures demonstrated more robust competition than those who did not.

This finding speaks directly to the current debate over whether competition arises immediately, as an outcome of encoding, or after some time delay, as an outcome of consolidation. Kapnoula et al. (2015) found compelling evidence of competition due to inhibition between trained cohorts and familiar English targets immediately after training. They argue that encoding not only involves bottom-up connections between sublexical (phonemic) and lexical (word form) levels to create lexical configurations but also connections between word forms to create lexical engagement. Both are accomplished by (the same) experience-driven mechanisms. Here, in accord with that hypothesis, we have demonstrated that, when a learner is allowed more opportunities to encode, engagement is more robust.

Although both the LI and ND groups demonstrated competition between English and trained cohorts, there was a marginal trend towards less robust competition on the part of the LI group. If the hypothesis above is correct, this pattern follows logically from the encoding deficits that characterise the LI group. Because they are poorer at encoding, they will develop weaker engagement. Given that the trend was marginal and the effect was small, replication with a larger sample size is required before drawing definitive conclusions.

If replicated, it becomes important to consider carefully the nature of the competition effect. Because the critical trials involve the simultaneous presentation of the pictured referents of the English targets (e.g., bucket) and the newly trained novel cohort (e.g., bukuz), both word forms are activated, and the listener must await the uniqueness point to determine his or her response. However, other investigators have demonstrated that these lexical competition effects are not limited to the displayed set (Kapnoula et al., 2015; Magnuson, Tanenhaus, Aslin, and Dahan; 2003). For example, Magnuson et al. (2003) found that people recognised

newly trained words with many trained neighbours more slowly than those with few trained neighbours even when the referents named by the neighbours were not displayed.

It is also important to consider the best interpretation of the competition effects. Because the newly trained words competed with their English neighbours, one interpretation is that the newly trained novel words have become part of the native lexicon. Such lexicalisation is critical to the definition of engagement as put forth by Leach and Samuel (2007). However, Magnuson et al. (2003) have demonstrated that, whereas newly trained novel words do exhibit competition with cohorts, they do so in a way that is largely isolated from the native lexicon. Specifically, they found no reliable effects of the number of English neighbours on the recognition of the newly learned novel words. Such isolated completion effects accord with the distinction between episodic and lexical memory for word forms as presented by Takashima et al. (2014). An early-established episodic memory could support performance in the tasks used here; a lexicalised memory would not be necessary.

Characteristics of the training are likely crucial. Although the novel words trained here obeyed English phonotactic constraints and half of them had believable pictured referents, the college students who participated undoubtedly realised that they were not learning new English words. To refrain from truly lexicalising them would be adaptive. It would be interesting to examine lexical engagement in paradigms that train real English words, or that present novel words in ways that carefully mimic ecologically valid learning contexts.

In the very least, the competition observed here constitutes evidence that the newly trained word forms exist in long-term memory and that they take on some lexical properties, namely, they compete with simultaneously presented cohorts, during word recognition. For now, we consider engagement a strength relative to configuration for the LI group. The trend towards weaker competition on the part of learners with LI and a more specific understanding of the nature of that competition await replication and the application of more sensitive paradigms (see Kapnoula et al., 2015 for an example).

Conclusion

Diagnostic accuracy

As a group, college students with LI presented with word learning problems when compared with a group of unaffected peers carefully matched for age, gender, and educational level. However, these group-level comparisons masked a great deal of inter-individual variability. Many of the participants with LI performed as well as their peers with ND. Similarly, in our previous study, only 1/4 to 1/3 of college students with LI fell more than one standard deviation below the mean of the matched ND group on any given measure of word learning (McGregor, Licandro, et al., 2013). One might expect word learning to distinguish clinical and nonclinical groups in younger or more severely affected populations. However, Gray (2003) reported individual differences in word learning among preschool children with developmental LI as well. The number of words that children learned to comprehend after a training protocol yielded a sensitivity to LI of 93% but a specificity of only 53%. Fourteen of 30 children who came into the study with diagnosed LI learned words as well as their unaffected age-mates. Word learning performance, then, will not serve as a suitable

diagnostic marker of developmental language learning impairments. That is not to say that it is not an important deficit to address in individuals who are affected.

Intervention

There was a positive correlation between encoding as measured by recognition and production immediately after training and extant vocabulary in the LI group. Although cause cannot be attributed, the relationship leads us to hypothesise that strategies to aid encoding would be worthwhile for individuals with LI who have small vocabularies. There are many variables known to improve encoding including multiple vs. few exposures (Gray, 2003; Nash & Donaldson, 2005), small vs. large training sets (Johnson, Gersten, & Carnine, 1987), active vs. passive training (Roediger & Karpicke, 2006), and a mix of repeated and variable input contexts (Horst, 2013). If the encoding of form is a particular challenge, the current results demonstrate that exposures and questions that lead the learner to focus on form are helpful.

Acknowledgments

We thank the participants for their generosity and Efthymia Kapnoula and Katherine Gordon for feedback on an earlier version of this manuscript. This study was funded by NIH-NIDCD- R01 DC011742.

Appendix A: Stimuli

Stimuli for training and probes presented in orthographic and phonetic notations

Set A	Set B	Set C
bukuz	bʌkəz	bisal
cigub	ɪɡʌb	cedil
gofic	ɡɒfɪk	fezob
ketis	kɛtɪs	garip
lemab	lɛməb	gitug
papig	pɛpɪɡ	hamul
rifas	rɪfəs	jaken
robag	rɒbɪɡ	kabuk
sadeb	sædəb	muzit
tebod	tebɒd	rulot
wilud	wɪləd	sodyn
zipyt	zɪpəɪt	tunam

Note: Two items in sets B and C had identical onsets (see underlined items). For participants assigned sets B and C, these items were not scored on the stem completion probe.

Stimulus pairings for the visual world paradigm

Language of Target	Language of Foil	Competitor status of Foil	Number of items	Example
English	English	Competitor	24	Bacon-Baker
		Non-competitor	24	Bacon-Rocket
English	Trained	Competitor	12	Cigar-Cigub

Language of Target	Language of Foil	Competitor status of Foil	Number of items	Example
Trained	English	Non-competitor	12	Cigar-Zipyty
		Competitor	12	Bukuz-Bucket
		Non-competitor	12	Bukuz-Kettle

Appendix B: Instructions for training and probes

All instructions were presented in spoken and written form.

Meaning-focused training

Following are 12 sets of exercises meant to introduce you to new objects and the words that name them. For each of the 12 sets, there will be a question to answer about each of the objects contained therein. For example, we might ask, 'Is it blue?' To answer yes to the question, press the button on your left; to answer no, press the button on your right. Press any button to begin the experiment.

Training questions:

1. Does it have eyes?
2. Is it edible?
3. Is it soft?
4. Is it man made?
5. Is it alive?
6. Is it heavy?

Form-focused training

Following are 12 sets of exercises meant to introduce you to new words. For each of the 12 sets, there will be a question to answer about each of the words contained therein. For example, we might ask, 'Does it contain a z sound?' To answer yes to the question, press the button on your left; to answer no, press the button on your right. Press any button to begin the experiment.

Training questions:

1. Does it contain an 'r' sound?
2. Does it contain an 's' sound?
3. Does it contain a 't' sound?
4. Does it contain an 'm' sound?
5. Does it contain a 'p' sound?
6. Does it contain a 'd' sound?

3AFC Word recognition task

You will hear a series of word triplets. The words in each triplet will sound similar. One of the three is a word from your study session. If you think the first word is from your study session, press '1'. If you think the second word is from your study session, press '2'. And if you think the third word is from your study session, press '3'. Press any button to begin the experiment.

Stem completion task

This is a word-completion test. Each word is a two-syllable word that you learned in your training sessions. You will hear the first syllable of the word, and your job is to respond by saying the whole word. For example, if you learned the word 'puppet', you would hear 'pu' and you would respond by saying 'puppet'. Each trial starts with the phrase, 'Please say the word that begins with...'. You will then hear the first syllable of one of the new words that you have learned. After you have heard the phrase and the first syllable, please make your response. If you don't remember the word exactly, please try to say what you think might be correct. Press any button to begin.

Visual World Paradigm

In this task you will hear a word and click on the picture that the word represents. You will use the mouse to drag the cursor up a runway. Once you launch from the runway you will then hear a word. Steer the cursor to the picture that matches the word you heard and then click on that picture. Please answer as quickly as possible, this is a timed task. Let's practice a few before you begin.

References

- Allopenna PD, Magnuson JS, Tanenhaus MK. Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*. 1998; 38:419–439.
- Alt M. Phonological working memory impairments in children with specific language impairment: Where does the problem lie? *Journal of Communication Disorders*. 2011; 44(2):173–185. [PubMed: 20943232]
- Alt M, Plante E, Creusere M. Semantic features in fast-mapping: Performance of preschoolers with specific language impairment versus preschoolers with normal language. *Journal of Speech, Language, and Hearing Research*. 2004; 47:407–420.
- Alt M, Spaulding T. The effect of time on word learning: An examination of decay of the memory trace and vocal rehearsal in children with and without specific language impairment. *Journal of Communication Disorders*. 2011; 44(6):640–654. [PubMed: 21885056]
- American Speech-Language-Hearing Association. Guidelines for audiologic screening. Rockville, MD: ASHA; 1997.
- Bakker I, Takashima A, van Hell JG, Janzen G, McQueen JM. Competition from unseen or unheard novel words: Lexical consolidation across modalities. *Journal of Memory and Language*. 2014; 73:116–130.
- Becker TC, McGregor KK. How much do college students with learning disabilities learn from listening to lectures. 2015 Manuscript submitted for publication.
- Blalock JW. Persistent auditory language deficits in adults with learning disabilities. *Journal of Learning Disabilities*. 1982; 15:604–609. [PubMed: 7153655]

- Dahan D, Magnuson JS, Tanenhaus MK. Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive psychology*. 2001; 42(4):317–367. [PubMed: 11368527]
- Delis, DC., Kramer, JH., Kaplan, E., Ober, BA. *The California Verbal Learning Test-Second Edition*. San Antonio, TX: Pearson; 2000.
- Dollaghan C, Campbell TF. Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research*. 1998; 41:1136–1146.
- Dumay N, Gaskell MG. Sleep-associated changes in the mental representation of spoken words. *Psychological Science*. 2007; 18(1):35–39. [PubMed: 17362375]
- Dunn, LM., Dunn, DM. *Peabody Picture Vocabulary Test, Fourth Edition*. San Antonio, TX: Pearson; 2007.
- Fernandes T, Kolinsky R, Ventura P. The metamorphosis of the statistical segmentation output: Lexicalization during artificial language learning. *Cognition*. 2009; 112(3):349–366. [PubMed: 19545864]
- Fidler LJ, Plante E, Vance R. Identification of adults with developmental language impairments. *American Journal of Speech-Language Pathology*. 2011; 20:2–13. [PubMed: 20739630]
- Gais S, Born J. Declarative memory consolidation: Mechanisms acting during human sleep. *Learning and Memory*. 2004; 11:679–685. [PubMed: 15576885]
- Gaskell MG, Dumay N. Lexical competition and the acquisition of novel words. *Cognition*. 2003; 89:105–132. [PubMed: 12915296]
- Gray S. Word-learning by preschoolers with specific language impairment: What predicts success? *Journal of Speech, Language, and Hearing Research*. 2003; 46:56–67.
- Gray S. Word learning by preschoolers with specific language impairment: Predictors and poor learners. *Journal of Speech, Language, and Hearing Research*. 2004; 47:1117–1132.
- Hargreaves IS, Pexman PM, Johnson JC, Zdravilova L. Richer concepts are better remembered: number of features effects in free recall. *Frontiers in Human Neuroscience*. 2012; 6:73. [PubMed: 22514526]
- Hawkins E, Astle DE, Rastle K. Semantic advantage for learning new phonological form representations. *Journal of Cognitive Neuroscience*. 2015; 27:775–786. [PubMed: 25269110]
- Horst JS. Context and repetition in word learning. *Frontiers in Psychology*. 2013; 4:149. [PubMed: 23580347]
- Huetting F, Rommers J, Meyer AS. Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta psychologica*. 2011; 137(2):151–171. [PubMed: 21288498]
- Johnson G, Gersten R, Carnine D. Effects of instructional design variables on vocabulary acquisition of LI students: A study of computer-assisted instruction. *Journal of Learning Disabilities*. 1987; 20:206–213. [PubMed: 3553400]
- Kapnoula EC, McMurray B. Newly learned word forms are abstract and integrated immediately after acquisition. *Psychonomic Bulletin & Review*. 2015:1–9. [PubMed: 24847901]
- Kapnoula EC, Packard S, Gupta P, McMurray B. Immediate lexical integration of novel word forms. *Cognition*. 2015; 134:85–99. [PubMed: 25460382]
- Kaufman, AS., Kaufman, NL. *The Kaufman Brief Intelligence Test-2*. Circle Pines, MN: AGS Publishing; 2004.
- Korkman M, Pesonen AE. A comparison of neuropsychological test profiles of children with attention deficit hyperactivity disorder and/or learning disorder. *Journal of Learning Disabilities*. 1994; 27:383–392. [PubMed: 8051511]
- Kuperman V, Stadthagen-Gonzalez H, Brysbaert M. Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*. 2012; 44:978–990. [PubMed: 22581493]
- Leach L, Samuel AG. Lexical configuration and lexical engagement: when adults learn new words. *Cognitive Psychology*. 2007; 55:306–353. [PubMed: 17367775]
- Magnuson JS, Tanenhaus MK, Aslin RN, Dahan D. The time course of spoken word learning and recognition: studies with artificial lexicons. *Journal of Experimental Psychology: General*. 2003; 132(2):202. [PubMed: 12825637]

- McGregor KK, Licandro U, Arenas R, Eden N, Stiles D, Bean A, Walker E. Why words are hard for adults with developmental language impairments. *Journal of Speech, Language, and Hearing Research*. 2013; 56:1845–1856.
- McGregor KK, Newman RM, Reilly RM, Capone N. Semantic representation and naming in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*. 2002; 45:998–1014.
- McGregor KK, Oleson J, Bahnsen A, Duff D. Children with developmental language impairment have vocabulary deficits characterized by limited breadth and depth. *International Journal of Language and Communication Disorders*. 2013; 48:307–319. [PubMed: 23650887]
- McMurray B, Samelson VM, Lee SH, Tomblin JB. Individual differences in online spoken word recognition: Implications for SLI. *Cognitive psychology*. 2010; 60(1):1–39. [PubMed: 19836014]
- Nash M, Donaldson ML. Word learning in children with vocabulary deficits. *Journal of Speech, Language, and Hearing Research*. 2005; 48:439–458.
- Rice ML, Oetting JB, Marquis J, Bode J, Pae S. Frequency of input effects on word comprehension of children with specific language impairment. *Journal of Speech and Hearing Research*. 1994; 37:106–122. [PubMed: 8170118]
- Richland LE, Kornell N, Kao LS. The pretesting effect: Do unsuccessful retrieval attempts enhance learning? *Journal of Experimental Psychology: Applied*. 2009; 15(3):243. [PubMed: 19751074]
- Rigler H, Farris-Trimble A, Greiner L, Walker J, Tomblin JB, McMurray B. The slow developmental time course of real-time spoken word recognition. *Developmental psychology*. 2015; 51(12):1690. [PubMed: 26479544]
- Roediger HL III, Karpicke JD. The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*. 2006; 1:181–210. [PubMed: 26151629]
- Rost GC, McGregor KK. Miranda rights comprehension in young adults with Specific Language Impairment. *American Journal of Speech-Language Pathology*. 2012; 21:101–108. [PubMed: 22230180]
- Secord, W., Semel, EM., Wiig, EH. *Clinical Evaluation of Language Fundamentals-4*. San Antonio, TX: Psychological Corporation; 2003.
- Semel, EM., Wiig, EH., Secord, W. *Clinical Evaluation of Language Fundamentals-3*. San Antonio, TX: Psychological Corporation; 1995.
- Sheng L, Byrd CT, McGregor KK, Zimmerman H, Bludau K. List memory in young adults with language learning disability. *Journal of Speech, Language, and Hearing Research*. 2015; 58(2): 336–344.
- Spivey M, Grosjean M, Knoblich G. Continuous attraction toward phonological competitors: Thinking with your hands. *Proceedings of the National Academy of Sciences*. 2005; 102:10393–10398.
- Storkel HL. Learning from input and memory evolution: Points of vulnerability on a pathway to mastery in word learning. *International Journal of Speech-Language Pathology*. 2015; 17:1–12. [PubMed: 25539474]
- Storkel HL, Armbruster J, Hogan TP. Differentiating phonotactic probability and neighborhood density in adult word learning. *Journal of Speech, Language, and Hearing Research*. 2006; 49(6):1175–1192.
- Swingle D. Lexical exposure and word-form encoding in 1.5-year-olds. *Developmental Psychology*. 2007; 43(2):454–464. [PubMed: 17352552]
- Takashima A, Bakker I, Van Hell JG, Janzen G, McQueen JM. Richness of information about novel words influences how episodic and semantic memory networks interact during lexicalization. *NeuroImage*. 2014; 84:265–278. [PubMed: 23962957]
- Walker, Matthew P. A refined model of sleep and the time course of memory formation. *Behavioral and brain sciences*. 2005; 28(01):51–64. [PubMed: 16047457]
- Williams, KT. *Expressive Vocabulary Test*. Second. Circle Pines, MN: AGS Publishing; 2007.

Figure 1.

Accuracy of recognizing trained words in a 3-alternative-forced-choice task administered immediately after form-focused training (a) and meaning-focused training (b). The white dots represent individual participants. The black bars represent the group median performance. The group with developmental language impairment (LI) differed from the group with normal development (ND) on the accuracy of items trained via form-focused training, $U = 419.5$, $z = 3.53$, $p = .0004$, and those trained via meaning-focused training, $U = 562.5$, $z = 2.13$, $p = .03$.

Figure 2.

Stem completion performance of individuals with developmental language impairment (LI) and normal development (ND) immediately after and one week after form-focused training. The black bars represent group mean performance.

Figure 3.

Stem completion performance of individuals with developmental language impairment (LI) and normal development (ND) immediately after and one week after meaning-focused training. The black bars represent group mean performance.

Figure 4.
Performance on the visual world paradigm by diagnostic group, cohort status, and foil type.
LI = developmental language impairment; ND = normal development

Figure 5.
The interaction between test exposure and cohort effects.

Table 1

Demographic characteristics and cognitive/linguistic scores of the groups with developmental language impairment (LI) and normal development (ND)

Domain	Measurement tool	LI		ND		Unit of measure	Normative mean (SD)	Effect size d
Age	Self-report	249 (16)		251 (19)		Months		n.s.
Education	Self-report	13.7 (1.3)		14.2 (1.4)		Years		n.s.
Nonverbal IQ	KBIT	103 (12.2)		112 (12.4)		Standard score	100 (15)	-.75
Spelling	* nonstandardised probe	5.1 (2.9)		11.6 (2.8)		Raw score out of 15		-2.46
Sentence comprehension	Token Test	35 (4.4)		41 (1.8)		Raw score out of 44		-1.78
Phonological memory	** nonstandardised probe	87 (4.3)		90 (5.0)		Raw score out of 96		-.66
Sentence memory	Sentence recall from CELF-3	8.21 (2.6)		11.53 (2.0)		Standard score	10 (3)	-1.43
List memory	CVLT-2	44 (8.6)		56 (9.0)		t-score	50 (10)	-1.37
Receptive vocabulary	PPVT-4	102 (11.2)		115 (12.0)		Standard score	100 (15)	-1.12
Expressive vocabulary	EVT-2	104 (18.0)		116 (11.2)		Standard score	100 (15)	-.80
Defining vocabulary	Word definition from CELF-4	12.2 (2.2)		14.8 (1.2)		Scaled score	10 (3)	-1.47

Note: KBIT = Kaufman Brief Intelligence Test; CELF = Clinical Evaluation of Language Fundamentals; CVLT = California Verbal Learning Test; PPVT = Peabody Picture Vocabulary Test; EVT = Expressive Vocabulary Test.

* Fidler et al. (2011);

** Dollaghan & Campbell (1998). Unless marked as nonsignificant (n.s.), all comparisons via t-tests for independent samples were significant at $p < .004$.

Table II

Ordered protocol

Enrolment day		Training and immediate test day		24-hour test day (for + interim test group only)		1-week test day	
1	Informed consent	1	Vocabulary training	1	Letter ID	1	Letter ID
2	Qualifying questionnaires and tests	2	3AFC word recognition probe	2	Stem completion probe	2	Stem Completion probe
3	Descriptive tests	3	Letter ID			3	Visual World probe
		4	Stem completion probe			4	Completion of any remaining descriptive tests

Note: 3AFC = 3-alternative-forced-choice task