



Published in final edited form as:

J Speech Lang Hear Res. 2013 April ; 56(2): 618–629. doi:10.1044/1092-4388(2012/11-0125).

Exemplar variability facilitates rapid learning of an otherwise unlearnable grammar by individuals with language-based learning disability

Janne von Koss Torkildsen¹, Natalie S. Dailey², Jessica M. Aguilar², Rebecca Gómez², and Elena Plante²

¹Formerly at The University of Arizona, now at the University of Bergen

²The University of Arizona

Abstract

Purpose—Even without explicit instruction, learners are able to extract information about the form of a language simply by attending to input that reflects the underlying grammar. Here we explore the role of variability in this learning by asking whether varying the number of unique exemplars heard by the learner affects learning of an artificial syntactic form.

Method—Learners with normal language (n=16) and language-based learning disability (LLD) (n=16) were exposed to strings of nonwords that represented an underlying grammar. Half heard 3 exemplars sixteen times each (low variability group) and half heard 24 exemplars twice each (high variability group). Learners were then tested for recognition of items heard and generalization of the grammar with new nonword strings.

Results—Only those learners with LLD who were in the high variability group were able to demonstrate generalization of the underlying grammar. For learners with normal language, both those in the high and the low variability groups showed generalization of the grammar, but relative effect sizes suggested a larger learning effect in the high variability group.

Conclusion—The results demonstrate that the structure of the learning context can determine the ability to generalize from specific training items to novel cases.

Although language form can certainly be learned through direct instruction, most frequently it is learned without explicit explanation of the rules for combining words into phrases or sentences. This is true for infants acquiring a first language, but is also true for older learners who become immersed in a community in which an unfamiliar language is spoken. This raises the question of what information and abilities learners use to master an unfamiliar syntactic system. This question is of theoretical importance for those interested in learning in general, but also has potentially important clinical implications. If it were known how such unguided learning is accomplished, and what facilitates this type of learning, these principles might be adopted in therapeutic approaches to the remediation of syntactic deficits.

The *Learning Mechanisms* perspective on language acquisition holds promise for addressing this issue. This perspective posits that individuals extract the properties of their language from the structure of the input they receive (Gómez, 2006; Saffran, 2003). It is thought that humans come equipped with a number of general cognitive and perceptual “mechanisms for

learning” that are applied to the task of acquiring language as well as learning in the nonverbal domain (Fiser & Aslin, 2002; Creel, Newport, & Aslin, 2004; Saffran, Johnson, Aslin, & Newport, 1999). Learning can be rapid (over the course of minutes) and is unguided in the sense that no explicit instruction is necessary (Saffran, Newport, Aslin, Tunick, & Barrueco, 1997). *What is learned* is thought to be influenced by the constraints of the cognitive resources the learner brings to the learning task (Gómez, 2006; Newport & Aslin, 2004; Saffran, 2003) as well as by accumulated experience (Dawson & Gerken, 2009; Gerken & Boltt, 2008; Lany, Gómez, & Gerken, 2007; Lany & Gómez, 2008). The methods used to study this type of language learning typically involve the use of miniature artificial languages. Unlike natural language stimuli, which necessarily present multiple and sometimes redundant cues to language structure, artificial languages allow the experimenter to constrain the input to the learner so that learning can be attributed solely to the use of those cues directly under experimental control.

Work conducted within a Learning Mechanisms framework has documented a number of principles concerning the nature of learning, particularly as it applies to language form. Learners are highly sensitive to statistical properties of events in their environment. Specifically, learners track the co-occurrence of events. For example, many syntactic forms are characterized by the predictable co-occurrence of word classes. Articles precede nouns. Verb tense markers occur only after verb root words. In these examples, a component from a closed set (i.e., article, tense marker) is systematically paired with an exemplar from an open class set. There is ample evidence that typical learners can come to recognize features that signal word classes (e.g., nouns, verbs, articles) and discover the underlying “rules” concerning how these different classes of words can be paired solely by hearing input that exemplifies these aspects of the grammar (Frigo & McDonald, 1998; Gerken, Wilson, & Lewis, 2005; Gomez & Lakusta, 2004; Richardson, Harris, Plante, & Gerken, 2006).

It is also the case that the way input to the learner is structured can have a strong influence on whether learning occurs. A now classic example involves learning of strings that reflect an ‘aXb’ form, where ‘a’ elements always co-occurred with ‘b’ elements, and the intervening ‘X’ elements were not dependent on either ‘a’ or ‘b’. Examples of this type of dependency in English include the forms ‘is verbing’ or ‘he verbs’ (as in ‘is running’ or ‘he runs’). When learners hear the a_b structure paired with only a few (1, 3, or 12) X elements, no learning occurs (Gómez, 2002; Newport & Aslin, 2004). However, Gómez (2002) and Gómez and Maye (2005) have demonstrated that presentation of many unique tokens (18–24) representing the ‘X’ element resulted in rapid learning of the same a__b relation. Thus, ‘high variability’, defined in this particular study as the number of unique ‘X’ tokens presented in aXb strings, facilitated learning of the grammatical relation.

It is important to note that token *variability* rather than token *frequency* drove this result. Learners in both the high and low variability conditions heard the aXb strings with equal frequency. That is, the total number of presentations of the a_b grammatical form during the experiment was the same for both high and low variability conditions. The only thing that differed was the number of different X tokens heard for these strings. Gómez (2002) hypothesized that when the variability of the ‘X’ elements was low, learners tended to focus on co-occurrence of only the adjacent elements (e.g., aX₁, aX_x etc.), which prevented learners from recognizing the relation between nonadjacent elements (i.e., a__b) that was key to the grammaticality of the entire string. When variability of the X elements was high, tracking adjacent co-occurrences became untenable. As a result, the stable relation between ‘a’ and ‘b’ elements became salient. This idea has potentially wide-reaching application in the context of language impairment. If poor learning occurs when learners focus on non-informative aspects of the input, strategic application of this variability principle could serve

to redirect attention to those relations that define the underlying grammar and consequently facilitate generalization.

Although there is a large and growing literature from normal language learners that demonstrates the ability to track information in ways that build knowledge of language form, far less is known about how those with impaired language utilize information in the input they receive. The few studies that have examined this type of learning by individuals with impaired language and learning disabilities has shown poor ability to track and use information provided by the input to learn components of a novel language (Evans & Saffran, 2009; Grunow et al., 2006; Plante et al., 2002; Plante, Bahl, & Gerken, 2010; Richardson et al., 2006). Moreover, their poor ability to track information is not limited to the verbal domain (Tomblin, Mainela-Arnold, & Zhang, 2007). This type of evidence has led some to propose that language impairment may result from a general problem in statistical learning (Hsu & Bishop, 2010). However, there is also some evidence that unguided learning is not impossible for this population. Impaired learners do track phonologically-based information (Evans & Saffran, 2009; Richardson et al., 2006) and prosodic information (Plante et al., 2010). Therefore, it is not a foregone conclusion that those with weak language skills will necessarily be poor learners under all conditions. For example, the language-impaired learners in Plante et al. (2010) rapidly acquired and generalized rules for assigning syllable stress only under conditions in which stress and loudness patterns were naturalistic, but not when stress patterns were acoustically enhanced. This type of outcome suggests that impaired learners can show rapid learning that generalizes beyond the input, if the input is structured in ways that are optimal for them.

The present study examined the effect of exemplar variability on learning by adults selected for poor language skills and their typically developing peers. We sought to demonstrate that the learning advantage found previously with high-variability open-class elements (Gómez, 2002; Gómez & Maye, 2005) would also promote learning by adults with impaired language. Specifically, we tested whether high variability of grammatical tokens rather than simple frequency of presentation would result in improved learning by adults with language-learning disability. Furthermore, we aimed to show that the variability principle generalizes beyond the aXb grammatical form used in previous studies (Gómez, 2002; Gómez & Maye, 2005; Newport & Aslin, 2004) by asking adults to learn a different grammatical structure. Learners heard 48 presentations of each of two grammatical forms. These took the forms aX and Yb, in which 'a' and 'b' are each represented by single nonwords and X and Y are represented by either 3 or 24 different nonwords. These phrases were constructed to reflect underlying rules for combining and ordering nonword elements. Test items were constructed to determine whether learners show a preference for information about how elements are paired (i.e., element 'a' must be paired with element 'X', element 'b' must be paired with element 'Y') versus the order in which elements appear in the string (e.g., element 'a' must appear in the initial position and element 'b' in the final position). The X and Y elements differed in terms of the number of syllables each contained, making it possible for learners to distinguish between the two classes of items.

In this study, we distinguish between two levels of learning. One level involves recognition of items that learners actually heard when presented with exemplars of two different grammatical forms (aX and Yb). The other level of learning involves generalization beyond the exemplars actually presented. Generalization of the grammar would indicate that learners have moved beyond recognition of specific strings to an understanding of the underlying patterns those strings represent. We further hypothesized that repeated exposure to a few unique exemplars should promote recognition of the individual exemplars. In contrast, exposure to many unique exemplars should promote generalization of the grammar.

Method

Participants

Thirty-two adults participated in this study. The participants were undergraduate students at the University of Arizona and native speakers of English. Sixteen (10 male, 6 female) belonged to the Language-Based Learning Disability (LLD) group. They ranged in age from 18 to 20 years of age ($M = 18.69$, $SD = 0.70$). Sixteen adults (10 male, 6 female) were members of the Normal Language (NL) group. They ranged in age from 18 to 19 years ($M = 18.50$, $SD = 0.52$). All adults passed a pure-tone hearing screening (500, 1000, 2000, & 4000 Hz at 25 dB HL bilaterally, ANSI, 1996). To rule out mental retardation, all participants were administered the *Test of Nonverbal Intelligence-III* (Brown, Sherbenou, & Johnsen, 1997) and were required to score above 75 ($70+1SEM$) to remain in the study. The scores on this measure (see Table 1) did not differ significantly between groups ($t(1,30)=1.43$, $p=.16$ [two-tailed], $d=.11$).

Adults were classified as having either LLD or NL status based on a combination of self-report and clinical testing. All members of the LLD group self-identified as having either a learning disability ($n=13$) or a history of speech-language services ($n=3$). No adult in the NL group reported a personal or family history for these conditions. The rate of language disorders in individuals identified as having a learning disability is estimated between 79 and 90% (Blalock, 1982; Gibbs & Cooper, 1989). Conversely, the rate of language impairment in the general population is estimated at 13% (Tomblin, Freese, & Records, 1992). Given this, it is clear that an individual's self-identified status is not a perfect predictor of their actual language status. Therefore, we wished to confirm poor language status in the adults in the LLD group and normal language status for adults in the NL group. We used the method of Fidler, Plante, and Vance (2011), which weights scores from a modified version of the Token Test (Morice & McNicol, 1985) and a 15-item written spelling test (Fidler et al., 2011) to identify individuals with scores similar to adults with a childhood history of speech-language impairment. In the present study, adults who self-reported a history of learning disability or speech-language services and whose test scores confirmed poor language skills were retained for study. Conversely, adults whose weighted test scores were consistent with normal language, and who were gender and aged matched (within 1 year) to members of the LLD group, were retained for the NL group.

We tested a pool of 29 potential participants who self-identified as having a learning disability or a history of speech-language services. From this pool, 16 were identified as having poor language skills consistent with an LLD status, and these were retained as subjects. Likewise, we tested 57 individuals who reported no history of speech or language disorder or learning disability. The much larger pool of potential NL participants than were needed occurred because students elected to complete this research study as one way of fulfilling a research requirement for a university general education class. Of these potential participants, 51 met the criteria for typical language on the Fidler et al. (2011) test battery. The final set of participants included in the NL group were randomly selected from this typical language pool until 16 gender- and age-matched participants were identified.

For descriptive purposes, all participants also received the subtests comprising the Broad Reading Cluster of the *Woodcock-Johnson Psycho-Educational Battery-III* (Woodcock, McGrew, & Mather, 2001). The LLD group scored significantly lower than the NL group on this measure ($t(1,30)=3.94$, $p=.0005$, $d=1.19$). These and all other test scores for the 16 study participants in each group are reported in Table 1. Inter-rater reliability was determined by having a second tester score the standardized tests during 8% of the test sessions. Point-to-point reliability for scoring ranged from 91% to 100% agreement with an average of 97.9%.

Stimuli and Procedures

This study was designed to manipulate variability while holding constant the frequency with which participants heard examples of the grammar. The variability manipulation involves the number of unique exemplars used to represent each of two grammatical forms (aX, Yb). The a and b elements each consisted of a single nonword. X and Y in the grammar represents two classes of nonwords, each represented by multiple exemplars. So that listeners could infer the word class, the nonwords associated with each class differed in syllable number. The X elements were one-syllable nonwords and Y elements were two-syllable nonwords. In the high variability condition, 24 unique X- and Y- exemplars for each grammatical form were presented. This will be referred to as Set Size = 24. In the low variability condition, 3 unique X- and Y-exemplars each grammatical form were presented, but each was presented 8 times as often as in Set Size = 24. This constituted the Set Size = 3 condition. Importantly, for both variability conditions, the total number of presentations of each grammatical form was the same (96 presentations). The only difference between conditions was in the variability of X- and Y-exemplars such that there were more unique exemplars in Set-Size = 24. An overview of the study design is presented in Table 2.

The study consisted of a familiarization phase during which exemplars of the artificial grammar were presented, followed by a test phase during which participants were asked to accept test strings that belonged to the grammar and reject those that did not belong. Test items were composed of the 'a' and 'b' elements each paired with an 'X' or 'Y' element. The nonwords comprising the test items were either heard during familiarization or were novel (generalization items). Test items of each type were either consistent or inconsistent with the grammar.

Participants were familiarized with nonword strings that conformed to one of two complementary artificial languages in which two grammatical forms were presented. The structure is represented here as 'aX' and 'Yb' (or 'bX' and 'Ya' in the complementary version of the grammar) in which each letter corresponds to a nonword element. The alternate grammars were used to assure that potentially idiosyncratic factors involved in nonword pairings did not influence the results. The grammars included two CV nonwords ('a' and 'b' elements), one of which always appeared in the initial string position and one of which always appeared in the final position. One of the CV nonwords was paired with CVC nonwords (the 'X' elements) and the other with CVCCV nonwords (the 'Y' elements). For example, in one of the artificial languages, the 'a' element 'poe' was always followed by a CVC-word such as 'jeb', while the 'b' element 'koo' was always preceded by a CVCCV-word such as ritva. Given that there was only one exemplar of each of the 'a' and 'b' elements, and relatively more 'X' and 'Y' elements, the 'a' and 'b' elements are analogous to closed-class elements and the latter are analogous to open class elements in natural languages.

In both experimental conditions (Set Size = 3 and Set Size = 24) and in both versions of the grammar (aX, Yb and bX, Ya), the 'a' and 'b' elements were represented by the nonwords 'poe' and 'koo'. The set of CVC and CVCCV nonwords representing the 'Xs' and 'Ys' are included in the Appendix. These nonwords contained 17 of the English consonants in the initial position. Fifteen different consonants were used in the middle or final position (see the Appendix). Six different vowels used in the CVC nonwords including /a/, /æ/, /ɪ/, /ε/, /ɔ/ and /ʌ/. Vowels used in the CVCCV words included /a/, /æ/, /e/, /i/, /ɪ/, /aɪ/, /u/, /ʌ/ and /o/. This variety of speech sounds used helped to make the individual nonwords perceptually distinct. In addition, the overlapping sound inventory for CVC and CVCCV nonwords prevented phonology from being a strong cue to word class (i.e., words that could be paired with either 'a' or 'b' elements). Instead, the word class of the 'X' and 'Y' elements could only be determined by the number of syllables the nonword contained.

The nonwords used in the experiment were recorded by four speakers, two female and two male. Multiple talkers were used because previous work has shown that recall of real words and encoding of nonwords is better when items are produced by multiple talkers compared to a single talker (Goldinger, Pisoni, & Logan, 1991; Richtsmeier, Gerken, Goffman, & Hogan, 2009). Recordings from both female speakers and one of the two male speakers were used during the familiarization phase, and the test items were presented in the remaining male voice. Participants in both the high and low variability conditions heard each talker the same number of times during the experiment; the only difference was the number of different X- and Y-exemplars heard. The digital recordings were edited using SoundForge 7.0 (Sony, 2003) to ensure that the duration, stress, and loudness for each nonword were comparable across speakers. The nonwords were combined into strings in which the two elements within a string were separated by 150 milliseconds, and the interval between strings was 2000 milliseconds. The experiment was delivered via computer, using Direct RT software (Jarvis, 2004). The software played the audio files and collected data as participants made keyboard responses.

Participants were randomly assigned to one of two set-size conditions. Participants in the LLD and NL groups were likewise equally represented in the two complementary versions of the grammar (i.e., [aX, Yb] or [bX, Ya]) within each set-size condition. Half (8 NL and 8 LLD participants) were familiarized with 3 unique exemplars of each grammatical string (e.g. given the aX, Yb grammar, participants heard 3 aX strings and 3 Yb strings for 6 unique strings total). The other half of the participants (8 NL and 8 LLD) were familiarized with 24 unique exemplars of each of the two grammatical forms. Between-group t-tests indicated no differences for test scores on the Modified Token Test, Written Spelling, or Broad Reading cluster test scores for either the NL or the LLD participants who were assigned to each set size condition.

To control for the total number (i.e. frequency) of exposures, each subject heard exactly 96 strings, with the strings in the Set Size = 3 condition repeated eight times each and strings in Set Size = 24 presented once. This allowed us to contrast *a high repetition of items drawn from a small set size (3)* with *low repetition of items drawn from a large set size (24)*. Frequency of occurrence of each grammatical form was held constant across conditions.

Before beginning the experimental phases, participants were instructed to listen to the stimuli without being provided specific information about the nature of the stimuli or the task they would perform in the subsequent test phase. In order to keep participants engaged during exposure to the training items, they were instructed to press a button whenever they heard a tone. The tone was presented in a quasi-randomized manner so that there were 5–12 familiarization items between each presentation of the tone.

The experiment consisted of four consecutive phases. During the initial familiarization phase of 2 minutes and 50 seconds (exposure to half of the training items) participants in the Set Size = 3 condition heard 8 repetitions of each of the three unique exemplars of the grammatical strings for a total of 48 items. Participants in the Set Size = 24 condition heard 1 presentation of each unique grammatical string for a total of 48 items. Items were presented in a computer-generated randomized order. This was followed immediately by the first of two test sessions. Test items included three exemplars each of the test item types listed in Table 2 for a total of 30 test items. This initial test phase lasted 3 to 4 minutes, depending on the individual participant's response rate. It was followed immediately by a second familiarization session of 2 minutes and 50 seconds. During this phase, the stimuli presented in the first familiarization session (48 items) were replayed in a new computer-randomized order. This was followed by the second test session, in which the original 30 test items were replayed in a new computer-randomized order. This split between familiarization

and test phases was designed to accommodate the large number of test items. Specifically, we were concerned that presenting 60 test items in a row, with three-fifths representing ungrammatical forms (see Table 2), would erode the underlying representation of the grammar before testing was complete. Splitting the familiarization and test phases allowed us to “reset” the representation of the grammar after administering the first 30 test.

Before the onset of the first test phase, subjects were informed that the items they had heard in the training phase were generated using a set of rules, and that their learning of these rules would be tested. Participants were instructed to decide whether the test items followed the rules or not by pressing one of two buttons (one for a “yes” response and the other for a “no” response). For each response, the participant was given feedback about whether their decision was correct or not. If a response was incorrect, a red X appeared on the computer screen. If the response was correct, the computer program moved to the next item.

As summarized in Table 2, there were five different types of test items. These included correct items heard during training, and two types of incorrect items composed of syllables also heard during training. These incorrect test items probed two types of learning strategies concerning the defining role of the closed class ‘a’ and ‘b’ elements within the grammar: (1) attention to linear order and (2) attention to co-occurrence of the nonword pairs. Evidence of *attention to linear order* would consist of high acceptance rates for grammatical strings accompanied by a low acceptance rate of items for which the ‘a’ or ‘b’ item appeared in the incorrect position within the nonword string, but these elements are paired with the correct X- or Y- nonwords (see Table 2). Conversely, evidence of *attention to co-occurrence* would consist of a high acceptance of grammatical strings accompanied by a low acceptance of strings with the a- and b-nonwords in the correct position, but incorrectly paired with X- and Y-nonwords. Sensitivity to both linear order and co-occurrence information would consist of a high acceptance rate of correct items accompanied by a low acceptance rate of the remaining incorrect items. *Learning* is defined as a significant difference between the acceptance of correct items compared with acceptance of either of the two incorrect item types (i.e., Linear Order Violation or Co-occurrence Violation). Therefore, acceptance rates for the three types of test items constituted the dependent variable in this study. In addition to testing items composed of elements heard during familiarization, participants were also tested on two types of generalization items. These included correct generalization items that conformed to the grammar heard and included either a familiar ‘a’ or ‘b’ element (the closed-class elements). The open class ‘X’ and ‘Y’ elements were replaced with new nonwords that conformed to the syllable number and structure heard during familiarization, but were composed of phoneme sets not previously heard during familiarization (referred to as X_{New} Y_{New} elements in Table 2). Incorrect generalization items were ones for which the ‘a’ and ‘b’ elements were in the correct position, but the paired open class nonword in each string violated the number of syllables expected for that pairing (a co-occurrence violation). Note that we did not also include generalization items with Linear Order violations, primarily because the number of test items was already approaching the limits of the participants’ abilities to remain engaged in the task.

Results

Familiar items

The performance on test items that were part of the familiarization set is displayed in Figure 1. Out of a total possible of 12 where a score of 6 represents chance responding, members of the NL group in Set Size = 3 accepted a mean of 10.88 (SD = 1.23) Correct items, 4.13 (SD = 3.18) Linear Order Violation items, and 6.25 (SD = 2.66) Co-occurrence Violation items. The acceptance rates for the Correct items was significantly above chance (single sample t-test, $p < .05$, one tailed) for this group, and the two incorrect item types did not differ from

chance. The NL Set Size = 24 group accepted a mean of 8.88 (SD = 1.73) Correct items, 3.63 (SD = 2.56) Linear Order Violation items, and 6.50 (SD = 1.69) Co-occurrence Violation items. The acceptance rates for both the Correct items and the Linear Order Violation items differed significantly from chance, reflecting the fact that Linear Order Violations were accepted at levels significantly below chance. The LLD Set Size = 3 group accepted a mean of 8.88 (SD = 1.46) Correct items, 7.50 (SD = 1.51) Linear Order Violation items, and 7.38 (SD = 2.33) Co-occurrence Violation items. The acceptance rates for both the Correct items and the Linear Order Violation items were significantly above chance. The LLD Set Size = 24 group accepted a mean of 8.39 (SD = 2.92) Correct items, 4.23 (SD = 2.48) Linear Order Violation items, and 7.38 (SD = 2.33) Co-occurrence Violation items. Again, the acceptance rates for both the Correct items and the Linear Order Violation items differed significantly from chance, with acceptance rates for Correct items significantly above chance and those for Linear Order Violations significantly below chance.

Comparison of responses to individual item types does not consider any underlying response bias (e.g., a tendency to either accept or to reject items overall). Therefore, we were interested in the relative rates at which the different item types were accepted. The data were analyzed statistically with a $2 \times 2 \times 3$ mixed ANOVA in which Group (NL vs. LLD) and Set Size (3 vs. 24) were the between-group factors and Item Type (Correct, Linear Order Violation, and Co-occurrence Violation) was the within-group factor. The number of test items accepted for each item type served as the dependent variable. The ANOVA revealed a significant main effect for Set Size ($F(1,28) = 49.59, p = 0.0015; \eta^2 = 0.31$), reflecting a higher overall acceptance rate for items in Set Size = 24 than Set Size = 3, and a main effect for Item Type ($F(2,56) = 41.99, p = 0.001; \eta_p^2 = 0.60$). The main effect for Group was not significant, but critically, there were several significant interaction effects involving Group. A significant three-way Item Type \times Group \times Set Size ($F(2,56) = 5.51, p = 0.01; \eta_p^2 = .16$) interaction superseded significant Group \times Set Size ($F(1,28) = 6.74, p = 0.02; \eta^2 = .19$) and Item Type \times Group interactions ($F(2,56) = 5.91, p = 0.01; \eta_p^2 = .17$).

We were primarily interested in the effect of token variability on learning within each group, and secondarily in what kinds of information in the input learners were tracking. Therefore, we used a series of planned comparisons to analyze the acceptance rate of different item types within each group. The planned comparisons at each set size are discussed below:

Set Size = 3—The NL group in this condition accepted significantly more Correct items compared to Linear Order Violation items ($t(1,2) = 45.02, p = 0.001, d = 3.65$) and Co-occurrence Violation items ($t(1,2) = 24.50, p = 0.001, d = 3.65$). Thus, there was evidence that this group learned items presented during the familiarization phase. Additionally, participants accepted significantly more Co-occurrence Violation items compared to items with Linear Order Violation items ($t(1,2) = 7.97, p = 0.008, d = 1.77$), indicating this group was more sensitive to linear order than to co-occurrence information. In contrast, planned comparisons in the LLD group revealed no significant differences between the three item types: Correct vs. Linear Order Violation items ($t(1,2) = 1.17, p = 0.29, d = .71$), Correct vs. Co-occurrence Violation items ($t(1,2) = 1.75, p = 0.20, d = .57$) and Linear Order Violation vs. Co-occurrence Violation items ($t(1,2) = 0.02, p = 0.89, d = .05$).

Set Size = 24—Learning of items presented during familiarization occurred for both groups under the high variability condition. Planned comparisons revealed that participants in the NL group in this condition showed significant differences between Correct vs. Linear Order Violation items ($t(1,2) = 18.74, p = 0.001, d = 1.92$), Correct vs. Co-occurrence Violation items ($t(1,2) = 4.72, p = 0.038, d = 1.48$) and Linear Order Violation vs. Co-occurrence Violation items ($t(1,2) = 8.74, p = 0.006, d = 1.18$). Similarly, planned comparisons in the LLD group show significant differences between the number of items

accepted between Correct vs. Linear Order Violation items ($t(1,2) = 85.56, p = 0.001, d = 1.56$) and Correct vs. Co-occurrence Violation items ($t(1,2) = 76.56, p = 0.000, d = 1.48$). There was no significant difference between the number of items accepted in the Linear Order Violation vs. Co-occurrence Violation items ($t(1,2) = 0.07, p = 0.79, d = .13$).

Generalization items

Performance on the generalization items is displayed in Figure 2. Out of a total possible of 12 test items (where 6 represents chance responding), members of the NL group assigned to the Set Size = 3 condition accepted a mean of 5.50 (SD = 3.51) Correct items and 2.88 (SD = 2.42) Co-occurrence Violation items. The acceptance rates for the Correct items did not differ significantly from chance (single sample t-test, $p < .05$, one tailed), but Co-occurrence Violation items were accepted at below chance rates. The NL group assigned to the Set Size = 24 accepted a mean of 7.75 (SD = 2.05) Correct items and 5.25 (SD = 2.05) Co-occurrence Violation items. The acceptance rates for the Correct items only differed significantly from chance. The LLD group assigned to the Set Size = 3 accepted a mean of 2.63 (SD = 1.80) Correct items and 3.50 (SD = 1.69) Co-occurrence Violation items. The acceptance rates for both items types were significantly below chance. The LLD group assigned to a set size of 24 accepted a mean of 7.63 (SD = 2.88) Correct items and 4.5 (SD = 1.60) Co-occurrence Violation items. The acceptance rates for Co-occurrence Violation items were significantly below chance.

The data were analyzed with a $2 \times 2 \times 2$ mixed ANOVA in which Group (NL vs. LLD) and Set Size (3 vs. 24) were the between-group factors and Item Type (Correct, Incorrect) was the within group factor. Note that in this condition, all incorrect items involved co-occurrence violations. An ANOVA revealed a significant effect for Set Size ($F(1,28) = 29.62, p = 0.00008; \eta^2 = 0.51$) and a significant effect for Item Type ($F(2,56) = 9.54, p = 0.005; \eta_p^2 = 0.25$). The main effect for Group was not significant, and there were no significant interaction effects.

Set Size = 3—Planned comparisons for Item Type within group indicated that the NL group in this condition showed a significant difference in the rate of acceptance of Correct vs. Co-occurrence Violation items ($t = 2.42, p = .011, d = .75$). However, the LLD group showed no difference in acceptance rate ($t = -0.81, p = .86, d = 0$).

Set Size = 24—The NL group in this condition accepted Correct items at higher rates than Co-occurrence Violation items ($t = 2.30, p = .015, d = 1.22$). The LLD group also showed a significant difference for these two item types ($t = 2.88, p = .004, d = 1.09$). Therefore, there was evidence of generalization for the NL group at both Set Sizes, but generalization for the LLD participants only occurred for the Set Size of 24.

Discussion

This study investigated the ability of adults to learn two contrasting grammatical forms under high or low exemplar variability conditions. Adults were provided with just over 5 minutes exposure to the novel grammatical forms, but were not provided any instruction other than to listen to the input. Adults with typical language were able to recognize previously heard exemplars and generalize the underlying grammatical structure to new exemplars. This was true for adults who heard just three unique exemplars of each grammatical form as well as for those who heard 24 unique exemplars. Adults with language learning disabilities showed a somewhat different pattern of results. Those exposed to only three exemplars of each grammatical form did not distinguish between items they had previously heard and items that deviated from what they had heard. They also failed to show

evidence of generalization under this condition. However, the LLD group exposed to 24 unique exemplars not only showed learning of the items heard during familiarization, but also showed evidence of generalization to new grammatical strings. These findings demonstrate that rapid learning of grammatical forms can be achieved for individuals with language learning disabilities, if the language input is structured in ways that facilitates rapid, unguided learning.

Adults with normal language required few exemplars of a simple grammar to both learn exemplars of the grammar and to generalize its underlying pattern to novel exemplars. We are familiar with only one other study (Gerken & Boltt, 2008) in which learning occurred with as few as three exemplars. In that study, infants were able to learn prosodic structure based on suprasegmental cues. To our knowledge, this is the first demonstration that typical adults are able to learn a language-like structure with as few as three exemplars. However, it may be that three exemplars were sufficient because of the simple nature of the grammatical forms used. Others (Poletiek & van Schijndel, 2009) have suggested that learning can occur with relatively few exemplars, as long as those exemplars robustly represent the nature of the structure to be learned. For each two-item grammatical string used here, two items (combining two open-class nonwords and one closed-class nonword) would be the minimum number of strings needed to uniquely represent the nature of the grammatical form. Therefore, three strings per grammatical form just exceeded the minimum number proposed as necessary for uniquely defining the grammar.

In contrast to the results for the NL group, familiarization with three exemplars was not sufficient to produce learning and generalization for adults with LLD. This is not necessarily surprising, given that poor language learning might be expected from those selected for poor language skills. The less intuitive finding is that rapid learning and generalization was achieved even by those with LLD when learners were presented with a larger variety of unique exemplars of the two grammatical forms. For those with LLD, the unlearnable grammatical form became learnable, after only minutes of exposure, by a simple manipulation of the input. Furthermore, the relative effect sizes associated with the two set size conditions ($d=0.75$ vs. 1.22) suggest that the NL group also benefitted from higher variability in terms of their ability to generalize the grammar to new exemplars.

For the LLD group, frequent repetition of a small number of training items was not effective for learning and generalization of a simple grammar. Learning for the LLD group rested with the relative variability of open-class training items heard. Those in the low variability condition (Set Size = 3) not only were unable to generalize the grammatical pattern, but were also unable to distinguish between strings they had heard multiple times previously and incorrect strings they had never heard. This happened even though the low variability group heard the correct test strings *sixteen times* more frequently than the high variability group. It is clear from this outcome that sheer repetition of items is not sufficient to generate learning for those with LLD.

The failure of the LLD group to track two-item strings when presented with only three exemplars is unlikely to reflect poor verbal memory for items they had previously heard for several reasons. First, adults with language impairment have not shown deficits on nonword repetition tasks when the total number of syllables in a string is small (Barry, Yasin, & Bishop, 2007; Fidler et al., 2011; Poll, Betz, & Miller, 2010), and the high number of repetitions of these items should have assisted their memorization. Second, learning did occur when the set of open class elements was eight times as large in the Set Size = 24 condition. This would represent a higher memory load than in the Set Size = 3 condition, and thus generalization to new exemplars should have been poor under the Set Size = 24 condition if learners were applying a memory-based strategy of remembering the specific

nonwords heard. Furthermore, under a Learning Mechanisms perspective, memorization of specific strings heard during familiarization should prevent generalization all together, given that it is not the specific pairings of nonwords that defines the grammar, but the statistical relation between word classes.

Instead, the failure of learners with LLD to distinguish between two word strings that were previously heard or not heard under the Set Size = 3 condition might be explained by a different type of learning strategy. When overall token variability was low, participants in the LLD group may have tracked the phonological composition of syllables rather than syllable order or co-occurrence. There are examples from artificial language studies in which LLD participants appeared to preferentially track this type of information as opposed to the structure of the language (Bahl, Plante, & Gerken, 2009; Plante, Gómez, & Gerken, 2002; Richardson, Harris, Plante, & Gerken, 2006). For instance, Richardson et al. reported that LLD learners had the highest acceptance rate for generalization items with the highest phonological similarity to training items, even if these items actually violated the grammar. In the present study, tracking individual nonwords may have been a viable strategy for Set Size = 3, in which learners only heard 8 unique nonwords. However, the 50 unique nonwords encountered by the Set Size = 24 group may have exceeded the capacity of LLD learners for tracking input at the phonological level. This manipulation of the input may have encouraged the LLD group to abandon a default strategy of focusing on the phonological forms that define the individual syllables, freeing them to notice the structure of the language. Although tracking the occurrence of syllables may be useful for some aspects of language acquisition, such as recognizing new lexical items, recognition of language structure requires learning of sequential structure beyond the phoneme level. From a Learning Mechanisms perspective, tracking sequential patterns of words is critical to discovering the structure of a language (e.g. a-elements predict Xs and not Ys).

If learners with LLD are indeed tracking occurrence of individual syllables over grammatical patterns, it may be that overwhelming variability in these units makes the less variable grammatical patterns more salient. There were only two unique grammatical patterns to be learned in the present study, compared to the eight unique nonwords in the Set Size = 3 and the 50 nonwords in the Set Size = 24 conditions. Considered in another way, the variability represented by the nonwords is four times the variability of the grammatical patterns at Set Size = 3 and 25 times the variability of the grammatical patterns at a set size of 24. Therefore, if learners truly seek the most stable components within the input, the grammatical pattern was the more stable component of the input at both set sizes. It appears that a 3:1 variability ratio was sufficient for normal learners to detect the grammatical pattern, but the LLD group required a much greater ratio for learning to occur.

This benefit of high variability for learning morpho-syntactic relations is not unique to this study. Gómez (2002) demonstrated the facilitative effect of high variability in the input for a grammar that is otherwise particularly difficult for typical learners to acquire (Newport & Aslin, 2004). Furthermore, the benefit of high variability was demonstrated for infants by 15 months of age (Gómez & Maye, 2005), suggesting that the benefit of variability in the input is largely age-independent. Our results extend this finding to an alternate grammar, indicating that the results are not unique to the aXb grammatical form. Some variability of exemplars is also important for moving learners from an emphasis on “tokens” to an emphasis on “types” (Gerken, 2006; Homa, Sterling, & Trepel, 1981). Likewise, Hamann, Silke, Apoussidou, and Boersma (2009) demonstrated that typical adults learned a phonotactic constraint better when provided with a large number of tokens presented twice than when a smaller number of tokens were presented multiple times.

The results demonstrate that how the learning context is structured can determine whether a particular grammar is learned and generalized. This is a particularly relevant finding for this population, given that failure to generalize learning has been identified as a significant problem for those with impaired language. This outcome further suggests that lack of generalization in the context of intervention may be linked to the number of exemplars used during training. Unfortunately, the specific number of unique exemplars trained is often unreported.

It is important to note that variability in the present study was not applied across the board to all aspects of the input. Instead, the variability manipulation was confined to the set of open-class elements (X & Y), and low variability was maintained for the closed-class 'a' and 'b' elements. Gómez (2002) previously suggested that when variability among all items is low, learners are able to track individual items in the input (e.g., the vocabulary) and are unlikely to appreciate more subtle patterns that occur between item types (e.g., the grammar). High variability for properties of the input that are extraneous to the grammatical structure (e.g., individual lexical items) may overwhelm the learners' ability to track the individual nonwords. When learners are no longer able to track and remember individual items, the grammatical patterns become the least variable aspect of the input, and these patterns then become salient to the learner (see Gibson, 1991 for a discussion of invariance in perceptual learning). Our data provides converging evidence in favor of this interpretation. The order and pairing of 'a' and 'b' with X and Y elements dictated the grammatical structure to be learned. When X and Y elements were highly variable, learners with LLD appeared to track the more stable information concerning the statistical relation between the 'a' or 'b' and X or Y elements. This permitted generalization of these patterns to new exemplars.

In addition to their ability to abstract patterns from fewer exemplars, the NL group also showed a subtle preference for tracking linear order in the input relative to how subsets of open- and closed-class items were paired. This was seen as a significant difference in the acceptance of Linear Order Violation and Co-occurrence Violation items for this group. This suggests a learning bias towards tracking the sequential order of items over the combination of items. This bias may be inherent to the learner, but may also be a bi-product of exposure to English, in which item order is strongly fixed for many types of morphological elements and grammatical word classes. Evidence from infant studies suggests that not only is phonological perception changed by speech input to the infant (e.g., Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Werker & Tees, 2002), but that native language experience can also rapidly bias subsequent learning of nonphonological linguistic forms (Gerken & Bollt, 2008). In contrast to the NL group, the LLD group showed no significant preference for either order or co-occurrence cues. This may reflect a weaker influence of prior English experience, consistent with their relatively weak language skills. Alternatively, it may reflect a difference in learning strategy that leads them to reject any form of deviation from the original input.

A potential limitation of this study is that we did not measure nonverbal cognitive abilities beyond the general measure of nonverbal intelligence. We do know that the subjects in the high and low variability conditions did not differ on the TONI-3, either of the measures key to diagnosing language impairment in adults (the Modified Token Test and a spelling test), or on any measures of reading comprehension, written phonology, or reading fluency (the Broad Reading cluster of the Woodcock-Johnson Psychoeducational Battery). It may be that measures more focused on memory or sustained attention would have revealed differences between participants with LLD who were assigned to Set Size=3 vs. 24. Although the procedure of random assignment of participants to the two research conditions makes this possibility less likely, it cannot be ruled out entirely.

A second potential limitation of this study is that it did not examine generalization items with linear order violations. This limited our ability to contrast two types of learning strategies, attention to linear order versus attention to co-occurrence. Data from the familiarization items showed that the preference for tracking linear order was confined to the NL group, but converging evidence from the generalization items would have strengthened this conclusion. Moreover, while the NL and LLD groups appeared to generalize the grammar equally well in the high variability condition, group differences in generalization ability might have appeared if linear order items had been included. In order to investigate whether the bias for linear order in NL learners reflects an inherent learning strategy or is a bi-product of exposure to English, the same experimental paradigm should be tested on individuals speaking a language with relatively free word order, such as Hungarian or Portuguese. Results from such a study might in turn inform the question of whether the discrepancy between the NL and the LLD group stems from differences in language-independent learning strategies or their use of native language input.

In summary, the results of this study demonstrate that variability in the input can be critical in determining whether a new grammatical pattern will be learned and generalized. However, variability must be applied strategically in ways that highlight the underlying grammatical pattern. The principle of strategic variability has multiple analogs in natural languages, including the fixed-order pairings of open-class nouns or verbs paired with closed-class grammatical elements (e.g., boys, runs, jumped). Thus, this principle is one that experience with natural languages may predispose both normal and impaired learners to take advantage of when faced with a novel learning task.

Acknowledgments

Preparation of this work was supported by National Institute on Deafness and Other Communication Disorders (NIDCD) grant R01DC04726.

References

- American National Standards Institute (ANSI). Specification for audiometers (ANSI S3.6-1996). New York, NY: Author; 1996.
- Aslin RN, Saffran JR, Newport EL. Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*. 1998; 9:321–324.
- Bahl M, Plante E, Gerken LA. Processing prosodic structure by adults with language-based learning disability. *Journal of Communication Disorders*. 2009; 42:313–323. [PubMed: 19324365]
- Barry JG, Yasin I, Bishop DVM. Heritable risk factors associated with language impairments. *Genes, Brain & Behavior*. 2007; 6:66–76.
- Blalock JW. Persistent auditory language deficits in adults with learning disabilities. *Journal of Learning Disabilities*. 1982; 15:604–609. [PubMed: 7153655]
- Brown, L.; Sherbenou, R.; Johnsen, S. *Test of Nonverbal Intelligence-3*. Bloomington, MN: Pearson; 1997.
- Creel SC, Newport EL, Aslin RN. Distant melodies: Statistical learning of non-adjacent dependencies in tone sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2004; 30:1119–1130.
- Dawson C, Gerken LA. Language and music become distinct domains through experience. *Cognition*. 2009; 111:378–382. [PubMed: 19338982]
- Evans JL, Saffran JR, Robe-Torres K. Statistical learning in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*. 2009; 52:321–335.
- 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]
- Fiser J, Aslin RN. Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences*. 2002; 99:15822–15826.

- Frigo L, MacDonald J. Properties of phonological markers that affect the acquisition of gender-like subclasses. *Journal of Memory and Language*. 1998; 39:218–245.
- Gerken LA. Decisions, decisions: Infant language learning when multiple generalizations are possible. *Cognition*. 2006; 98:B67–74. [PubMed: 15992791]
- Gerken LA, Bollt A. Three exemplars allow at least some linguistic generalizations: Implications for generalization mechanisms and constraints. *Language Learning and Development*. 2008; 4(3): 228–248.
- Gerken LA, Wilson R, Lewis W. Infants can use distributional cues to form syntactic categories. *Journal of Child Language*. 2005; 32:249–268. [PubMed: 16045250]
- Gibbs D, Cooper E. Prevalence of communication disorders in students with learning disabilities. *Journal of Learning Disabilities*. 1989; 22(1):60–63. [PubMed: 2703789]
- Gibson, EJ. *An odyssey in learning and perception*. Cambridge: MIT Press; 1991.
- Goldinger SD, Pisoni DB, Logan JS. On the nature of talker variability effects on recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, & Cognition*. 1991; 17:152–162.
- Gómez RL. Variability and detection of invariant structure. *Psychological Science*. 2002; 13:431–436. [PubMed: 12219809]
- Gómez, RL. Dynamically guided learning. In: Munakata, Y.; Johnson, M., editors. *Attention & Performance XXI: Processes of change in brain and cognitive development*. Oxford, UK: Oxford University Press; 2006. p. 87-110.
- Gómez RL, Lakusta L. A first step in form-based category abstraction by 12-month-old infants. *Developmental Science*. 2004; 7 (5):567–580. [PubMed: 15603290]
- Gómez RL, Maye J. The developmental trajectory of nonadjacent dependency learning. *Infancy*. 2005; 7:183–206.
- Grunow H, Spaulding T, et al. The effects of variation on learning word order rules by adults with and without language-based learning disabilities. *Journal of Communication Disorders*. 2006; 39:158–170. [PubMed: 16376369]
- Hamann, Silke; Apoussidou, Diana; Boersma, Paul. Modeling the formation of phonotactic restrictions across the mental lexicon. *Proceedings of the 45th Meeting of the Chicago Linguistics Society*; 2009.
- Homa D, Sterling S, Trepel L. Limitations of Exemplar-Based Generalization and the Abstraction of Categorical Information. *Journal of Experimental Psychology: Human Learning and Memory*. 1981; 7:418–439.
- Hsu HJ, Bishop DVM. Grammatical difficulties in children with specific language impairment: Is learning deficient? *Human Development*. 2010; 53:264–277. [PubMed: 22003258]
- Jarvis, BG. *DirectRT Research Software (Version 2004)* [Computer software]. New York: Empirisoft; 2004.
- Kuhl PK, Williams KA, Lacerda F, Stevens K, Lindblom B. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*. 1992; 255:606–608. [PubMed: 1736364]
- Lany JA, Gómez RL. Twelve-Month-Olds Benefit from Prior Experience in Statistical Learning. *Psychological Science*. 2008; 19:1247–1252. [PubMed: 19121132]
- Lany J, Gómez RL, Gerken LA. The Role of Prior Experience in Language Acquisition. *Cognitive Science*. 2007; 31:481–507. [PubMed: 21635305]
- Morice R, McNicol D. The comprehension and production of complex syntax in schizophrenia. *Cortex*. 1985; 21:567–580. [PubMed: 4092484]
- Newport EL, Aslin RN. Learning at a distance: Statistical learning of non-adjacent dependencies. *Cognitive Psychology*. 2004; 48:127–162. [PubMed: 14732409]
- Plante E, Bahl M, Gerken LA. Children with specific language impairment show rapid, implicit learning of stress assignment rules. *Journal of Communication Disorders*. 2010; 43:397–406. [PubMed: 20542518]
- Plante E, Gómez R, Gerken LA. Sensitivity to word order cues by normal and language/learning disabled adults. *Journal of Communication Disorders*. 2002; 35:453–462. [PubMed: 12194564]

Poletiek FH, van Schijndel TJP. Stimulus set size and statistical coverage of the grammar in artificial grammar learning. *Psychonomic Bulletin & Review*. 2009; 16:1058–1064. [PubMed: 19966255]

Poll G, Betz S, Miller C. Identification of clinical markers of specific language impairment in adults. *Journal of Speech, Language, and Hearing Research*. 2010; 53:414–429.

Richardson J, Harris L, Plante E, Gerken L. Subcategory learning in normal and language learning-disabled adults: How much information do they need? *Journal of Speech and Hearing Research*. 2006; 49:1257–1266.

Richtsmeier PT, Gerken LA, Goffman L, Hogan T. Statistical Frequency in Perception Affects Children’s Lexical Production. *Cognition*. 2009; 111:372–377. [PubMed: 19338981]

Saffran JR. Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science*. 2003; 12:110–114.

Saffran JR, Johnson EK, Aslin RN, Newport EL. Statistical learning of tone sequences by human infants and adults. *Cognition*. 1999; 70:27–52. [PubMed: 10193055]

Saffran JR, Newport EL, Aslin RN. Word segmentation: The role of distributional cues. *Journal of Memory and Language*. 1996; 35:606–621.

Saffran JR, Newport EL, Aslin RN, Tunick RA, Barrueco S. Incidental Language Learning: Listening (And Learning) out of the Corner of Your Ear. *Psychological Science*. 1997; 8:101–105.

Sony. Sound Forge 7.0. Madison, WI: Sony Pictures Digital Inc; 2003.

Thiessen ED, Saffran JR. When cues collide. Statistical cues to word boundaries in 7 to 9 month old infants. *Developmental Psychology*. 2003; 39:706–716. [PubMed: 12859124]

Tomblin JB, Freese PR, Records NL. Diagnosing Specific Language Impairment in adults for the purpose of pedigree analysis. *Journal of Speech & Hearing Research*. 1992; 35(4):832–844. [PubMed: 1405540]

Tomblin, J Bruce; Mainela-Arnold, Elina; Zhang, Xuyang. Procedural learning in adolescents with and without specific language impairment. *Language Learning and Development*. 2007; 3:269–293.

Werker JF, Tees RC. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior & Development*. 2002; 25:121–133.

Woodcock, R.; McGrew, KS.; Mather, N. Woodcock-Johnson Psychoeducational Battery-Third Edition. Itasca, IL: Riverside Publishing; 2001.

Appendix Stimuli corresponding to familiarization and test phases of the study

Closed Class Elements	One-syllable Open Class Elements		Two-syllable Open Class Elements	
Familiarization & Test	Familiarization	Test	Familiarization	Test
Poe	Bim	Hes	Daivba	Chedwee
koo	Div	Pel	Femkoe	Tilma
	Fam	Tiv	Fapjae	Nabkoe
	Gip	Dis	Gepta	Fedka
	Jeb	Mip	Dasra	Nefkee
	Kes	Sul	Aska	Valsa
	Kiv		Kevma	
	Lif		Leflee	
	Mef		Minroo	
	Mot		Muptae	
	Neb		Mitnee	
	Nug		Neglai	

Closed Class Elements	One-syllable Open Class Elements		Two-syllable Open Class Elements	
Familiarization & Test	Familiarization	Test	Familiarization	Test
	Pob		Pabvoo	
	Ril		Ritva	
	Rudge		Ralza	
	Sog		Safwa	
	Teb		Shamkee	
	Tol		shapna	
	Vep		Sigwoo	
	Vug		Tevzee	
	Wib		Tepvoo	
	Zek		Vipcha	
	Shup		Wagso	
	Chud		zikvoo	

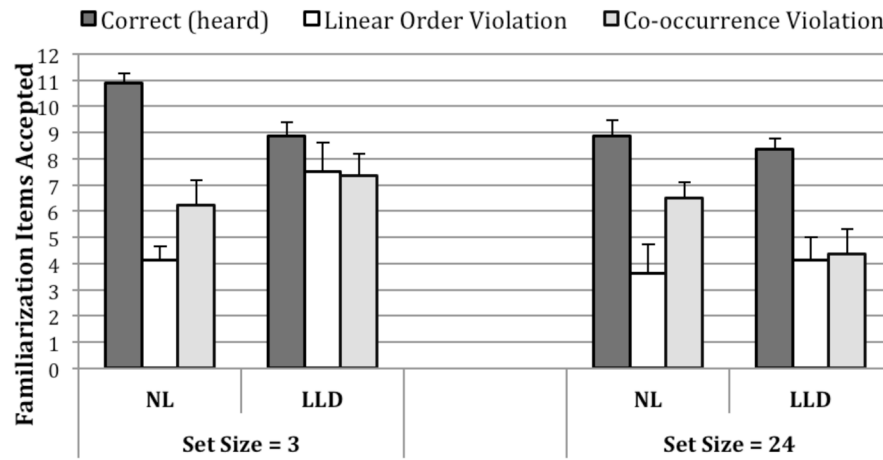


Figure 1. Performance (mean and SE) on test items heard during familiarization. Learning of items heard is defined as significantly greater acceptance of strings heard (Correct items) compared with strings that are composed of the same nonwords, but occur in the wrong order (Linear Order Violation) or involve incorrect item pairings (Co-occurrence Violation).

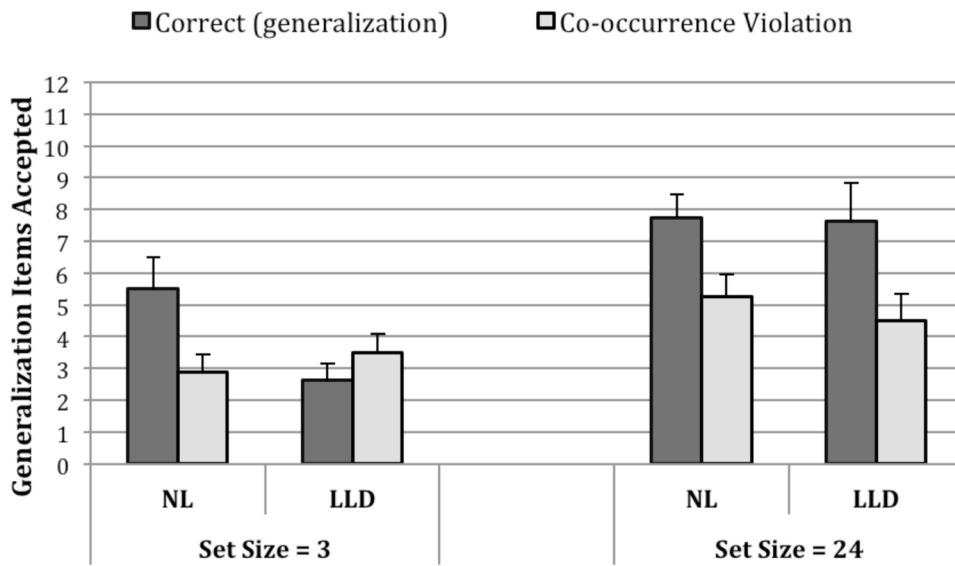


Figure 2. Performance (mean and SE) on generalization test items. Generalization of the grammatical form is defined as significantly greater acceptance of strings that conform to the grammar (Correct items) compared with strings that involve incorrect item pairings (Co-occurrence Violation). In all cases, strings included nonwords not heard during familiarization.

Table 1

Test scores for the participant groups.

	Group			
	Normal Language		Language Learning Disability	
	Mean	SD	Mean	SD
Age	18.50	0.52	18.69	0.70
Modified Token Test ¹	36.56	5.01	31.38*	6.25
Dictated Spelling ²	9.75	2.79	5.19*	3.35
WJ ³				
Letter-Word Identification	99.88	7.50	94.00*	6.65
WJ Reading Fluency	98.44	10.96	87.50*	11.58
WJ Passage Comprehension	102.69	7.38	95.94*	6.20
WJ Word Attack	97.19	9.13	87.06*	6.21
WJ Broad Reading Cluster	100.00	8.28	90.56*	6.55
TONI – 3 ⁴	102.19	16.20	101.56	13.10

* Significant group difference at $p < .05$.

¹ Raw score out of 44 possible items.

² Raw score out of 15 possible items

³ Woodcock-Johnson Psychoeducational Battery-3rd Edition, standard score (mean 100, SD=15)

⁴ Test of Nonverbal Intelligence-3rd Edition, standard score (mean 100, SD=15)

Table 2

Experimental design. For ease of explanation, examples are drawn from the aX, Yb grammar. Half of the participants heard the alternate bX, Ya grammar. Lower case letters are each represented by a single CV nonword. Uppercase letters indicate sets of nonwords, where X elements are CVC words and Y elements are CVCCV words. The X_{New} and Y_{New} indicate items that were not heard previously, but conformed in syllable composition to X and Y items used in the familiarization phase.

Experimental Phase	Item Type	Set Size	String	Number of Strings	Repetition of Each String*	Examples
Familiarization	aX, Yb grammar	3	aX Yb	3 3	16 16	poe jeb ritva koo
		24	aX Yb	24 24	2 2	
Test	Items heard during training	3	aX Yb	3 3	2 2	
		24	aX Yb	3 3	2 2	poe jeb ritva koo
	Items with co-occurrence violations	3	aY Xb	3 3	2 2	
		24	aY Xb	3 3	2 2	poe ritva jeb koo
	Items with linear order violations	3	Xa bY	3 3	2 2	
		24	Xa bY	3 3	2 2	jeb poe koo ritva
Correct Generalization items		3	aX _{New} Y _{New} ,b	3 3	2 2	
		24	aX _{New} Y _{New} ,b	3 3	2 2	poe sul valsa koo

Experimental Phase	Item Type	Set Size	String	Number of Strings	Repetition of Each String*	Examples
	Incorrect generalization items with co-occurrence violations	3	aY _{New}	3	2	poe valsa sul koo
			X _{New} b			
		24	aY _{New}	3	2	
			X _{New} b			

* Number of times each exemplar of the string type was heard during the experiment.