The effects of variation on learning word order rules by adults with and without language-based learning disabilities

Hope Grunow, Tammie J. Spaulding, Rebecca L. Gómez, Elena Plante *

Department of Speech, Language, & Hearing Sciences, The University of Arizona, 1131 E Second St., Tucson, AZ 85721-0071, USA

Received 19 April 2005; received in revised form 2 September 2005; accepted 8 November 2005

Abstract

Non-adjacent dependencies characterize numerous features of English syntax, including certain verb tense structures and subject–verb agreement. This study utilized an artificial language paradigm to examine the contribution of item variability to the learning of these types of dependencies. Adult subjects with and without language-based learning disabilities listened to strings of three non-words for which the first and third elements had a dependent relationship. In the low variability condition, 12 non-words occurred in the middle position, and in the high variability condition, 24 non-words occurred in this position. Non-disabled adults were able to learn the non-adjacent contingencies and generalize the underlying structure to new strings, but only when variability was high. Adults with language-based learning disabilities did not perform above chance levels under either variability condition. Thus, this group showed poor sensitivity to statistical information in speech input that both infants and non-disabled adults are known to track.

Learning outcomes: As a result of this activity, the reader will: (1) understand the advantages of using an artificial language to investigate language learning; (2) become familiar with a paradigm for studying the rapid learning of syntactic contingencies; (3) comprehend how the ability to map language structure differs for non-disabled adults and adults with a history of language/learning disability as a function of variability in the input the listener receives.

© 2005 Elsevier Inc. All rights reserved.

* Corresponding author. Tel.: +1 520 621 5080; fax: +1 520 621 9901.
E-mail address: eplante@u.arizona.edu (E. Plante).

0021-9924/$ – see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcomdis.2005.11.004
Artificial languages have been widely used in the investigation of language learning in infants, children (e.g., Aslin, Saffran, & Newport, 1998; Gerken, 2004; Gómez, 2002; Gómez & Gerken, 1999; Saffran, Aslin, & Newport, 1996; Saffran, Newport, Aslin, Tunick, & Barrueco, 1997), and adults (e.g., Gómez, 2002; Newport & Aslin, 2004; Plante, Gómez, & Gerken, 2002; Saffran, Newport, & Aslin, 1996). Their use has a number of advantages. First, they control for prior learning. This increases confidence that learning is due to exposure to the artificial language during the experiment as compared to information learned outside the experiment. Second, artificial languages result in more control over the learning environment than natural languages, enabling researchers to isolate and manipulate factors of particular interest. Although there are many cues in the acoustic signal that learners may use to assist in decoding language, little is known about the learning process.

One cue to language structure is found in the ordering of words. This can be examined in terms of adjacent and non-adjacent dependencies. Adjacent dependencies refer to the “rules” by which consecutive words must be ordered. An example in the English language is “the boy ran” (a + b + c), where “the” must precede “boy,” which in turn must precede “ran”. In contrast, “ran boy the” could never occur, as articles must precede nouns in English. A non-adjacent, or remote, dependency occurs when two dependent elements are separated by one or more intervening elements. This is analogous to “is verb-ing,” or third-person agreement in the English language. For these types of structures, the middle element, in this case a verb, is highly variable, but the first and third elements, the auxiliary (is) and inflection (-ing) are invariant.

In comparison to adjacent dependencies, non-adjacent dependencies are extremely difficult to acquire (see Newport & Aslin, 2000, 2004), however, they play an important role in language acquisition. In English, linguistic material intervenes between morphosyntactically dependent auxiliaries and inflectional morphemes (e.g., is rapidly talking) and between nouns and verbs in number and tense agreement (e.g., the birds on the fence are chirping). Additionally, the frequency of co-occurrence of non-adjacent dependencies may be instrumental in leading learners to group the intervening elements as members of a category (Mintz, 2002, 2003; Onnis, Monaghan, Christiansen, & Chater, 2004). In support of this hypothesis, Mintz conducted a corpus analysis of child-directed speech on frequently occurring non-adjacent dependencies with exactly one word position intervening. He found that words occurring in the middle position tended to belong to the same category (e.g., noun, verb, preposition, adjective, adverb), thus suggesting how sensitivity to frequently occurring non-adjacent dependencies might help infants begin to categorize the syntactic roles of words in speech. Thus, on several accounts, learning dependencies between non-adjacent words and morphemes is an important milestone in acquiring syntax.

Gómez (2002) investigated learning of non-adjacent dependencies in infants and adults with normal language. The investigator was interested in the effect of variation on the perception of non-adjacent dependencies. In the infant studies, 18-month-old were given a 3-min exposure to one of two artificial languages producing 3-element strings. Grammar 1 (G1) strings followed the patterns aXc or bXd (e.g., pel wadim jic or vot kicey rud), whereas Grammar 2 (G2) strings followed the patterns aXd or bXc. In G2, the relationship between the first and third elements was reversed, such that pel sentences ended with rud, and vot
sentences ended with jic (e.g., pel wadim rud or vot kicey jic). In both languages, the grammaticality of a sentence depended on the relationship between the first and third elements, while the intervening element varied freely. Since the same set of X-elements occurred in both languages, the two are only distinguishable on the basis of non-adjacent dependencies. As noted above, in natural language the intervening categories are often open-class items comprising much larger sets than the function morphemes associated with non-adjacent structure. Hypothesizing that these set size differences might aid learning, Gómez found that high variability led to better perception of non-adjacent dependencies. Eighteen-month-old were able to acquire the non-adjacent dependencies when the intervening element came from a set of 24 possible words (reflected in differential listening times to strings from their exposure language versus the other language), but showed no discrimination when the intervening set size was smaller (3 or 12).

Gómez replicated this finding with normal language adults, using slightly more complex languages (containing three non-adjacent dependencies instead of two). Adults listened to the exposure language for 18 min, instead of 3, and subsequently made grammaticality judgments on auditory strings from the exposure language versus the other language that violated the non-adjacent dependencies of the training language. This consisted of making yes/no judgments about whether strings were or were not from the exposure language. Adults were exposed to a middle element set size of 2, 6, 12, or 24 words in a between-subjects design. When the set size of possible middle elements was 24, learners performed at 90% accuracy. In contrast, they performed only slightly higher than chance following exposure to set sizes of 2, 6, and 12.

Gómez argued that high variability in the large set size condition acted to increase the salience of the non-adjacent elements compared to the middle element, and in this way, facilitated learning. The radical increase in performance for set size 24 suggests that when variation is high, learners experience a perceptual “pop-out” effect of the non-adjacent dependency.

Despite the fact that reliance on the adjacent dependencies was ineffective for determining word order rules, abandonment of this strategy was not seen with set sizes of 2, 6, or 12, suggesting that variation at those levels was not sufficient for the learners to recognize that alternative strategies were needed (see also Gómez & Maye, 2005, for a replication of the original studies). In particular, learners (both infants and adults) appeared to be focusing on different types of dependencies as a function of their statistical properties. When the conditional probabilities of adjacent elements were relatively high (in the small set size conditions) learners appeared to be tracking the statistical relationships between adjacent, rather than non-adjacent probabilities. However, when adjacent conditional probabilities were sufficiently low (when the set size was 24) adjacent dependencies are not reliable cues to structure, leading learners to focus instead on non-adjacent relationships.

Little is known about sensitivity to sequential word order in individuals with a history of language/learning disability. A study by Plante, Gómez, and Gerken (2002) examined sensitivity to sequential word order in adults with and without language/learning disabilities. Learners were exposed to a subset of strings from a finite-state grammar placing constraints on the orderings of words. Subjects were then asked to make grammaticality judgments on novel strings. The Plante et al. study resulted in a significant difference in performance between the two groups, with language/learning disabled
subjects exhibiting poor sensitivity to the ordering of adjacent dependencies. However, no information is currently available regarding the learning of non-adjacent dependencies in this population or whether variation would assist their learning as was seen in Gómez (2002).

The first purpose of the present investigation is to replicate the findings of Gómez (2002), showing that high variability in an intervening element can assist learning of non-adjacent word order dependencies in non-disabled (ND) adults. We also investigated whether learners would generalize to strings with novel middle elements. Generalization in the presence of a novel middle element provides the strongest test of non-adjacent dependency learning because it ensures that learners are in no way relying on memory of specific strings. Additionally, this study extends the work of Gómez (2002) and of Plante et al. (2002) to examine learning of non-adjacent word order dependencies in adults with a history of language-based learning disabilities (hL/LD). Given previous findings by Plante et al., we predicted that adults with hL/LD would perform less accurately than non-disabled (ND) adults. Given that increased variability appears to aid learning (Gómez, 2002), we predicted that this manipulation would help both hL/LD and ND learners, but that ND learners were likely to derive greater benefit than their hL/LD peers.

1. Method

1.1. Participants

Forty-four students were recruited from the undergraduate population at the University of Arizona. Subjects were tentatively placed into either a group with a history of language/learning disabled (hL/LD) or a non-disabled control group (ND). Subjects in the hL/LD group had self-reported a personal history of therapy and/or services for language impairment, dyslexia, and/or learning disabilities. The majority of these individuals were also receiving academic support and accommodations through the University.

The hL/LD group consisted of 22 adults (10 males, 12 females) who ranged in age from 18 to 26 years (mean age = 19.1 years). Hand preference was self-reported, with 20 participants reporting right-handedness and 2 participants reporting left-handedness. The control group consisted of 22 adults (9 males, 13 females), with 19 participants reporting right-handedness and 3 participants reporting left-handedness. Members of the ND group had reported no personal or family history of speech, language, or learning disorders. They ranged in age from 18 to 19 years (mean age = 18.45 years). All participants were native English speakers and passed an initial hearing screening. All participants reported no history of seizures, head trauma, or diagnosis of attention deficit hyperactivity disorder (ADHD).

The hL/LD and ND status reported by the participants was through a testing method developed by Tomblin, Freese, and Records (1992). They found that four standardized tests yielded a classification sensitivity of 97% and specificity of 100% in discriminating between adults with and without a childhood history of language impairment. We have adopted this set of tests for identifying adults with language-based learning disabilities because of the evidence validating their sensitivity (as a set) to residual impairment in
individuals known to have shown developmental language disorders. We know of no other tests for which equivalent evidence is available. The battery of tests included a modified version of the Token Test (Moric & McNicol, 1985), a written spelling subtest of the Multilingual Aphasia Examination (Benton & des Hamsher, 1978), a measure of speaking rate from a standardized picture description task (Tomblin et al., 1992), and the Peabody Picture Vocabulary Test-Revised (Dunn & Dunn, 1981). Note that the more recent version of the Peabody was not used because the newer version lacks score equivalence with the old and alters classification accuracy (Pankratz, Morrison, & Plante, 2004). Test results were statistically weighted to confirm subject classification through a discriminate analysis. In addition to those included as participants, 11 adults were excluded from this study because the Tomblin battery did not classify them into the group they were initially placed in based on their self-report. Eight of these participants had a history of hL/LD and three did not.

Despite the fact that application of the Tomblin battery results in subject samples that are 100% differentiated on the basis of the language skills, the absolute difference on the individual measures are modest (see Table 1). This is typical for adult samples of individuals with childhood diagnoses of language or learning disorders (e.g., Pankratz et al., 2004; Plante et al., 2002). This may, in part, reflect the relative sensitivity of the tests that comprise the battery. For example, children selected for specific language impairment show only minor discrepancies relative to their peers on measures of single word vocabulary (Gray, Plante, Vance, & Henrichsen, 1999), one of the four measures in the Tomblin battery. The modest absolute differences may also reflect the true profile of residual deficits in adults who have often had previous intervention. Because data on adult subjects are sparse, it is difficult to interpret whether the scores obtained for this sample are representative of a broader population, or if they index levels of language functioning in communicative or academic contexts.

The Test of Non-verbal Intelligence-III (Brown, Sherbenou, & Johnsen, 1997) was administered to ensure all subjects were of normal intelligence. Scores achieved by the

<table>
<thead>
<tr>
<th>Behavioral measures</th>
<th>hL/LD group</th>
<th>ND group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>Peabody Picture Vocabulary Test-Reviseda</td>
<td>92.9*</td>
<td>11.1</td>
</tr>
<tr>
<td>Modified Token Testb</td>
<td>34.3*</td>
<td>3.3</td>
</tr>
<tr>
<td>Speaking Ratec</td>
<td>140.8</td>
<td>35.0</td>
</tr>
<tr>
<td>Multilingual Aphasia Battery Spelling Subtestd</td>
<td>9.5*</td>
<td>1.5</td>
</tr>
<tr>
<td>Test of Non-verbal Intelligence-IIIa</td>
<td>101.1</td>
<td>14.6</td>
</tr>
<tr>
<td>Woodcock–Johnson Letter Word Identification Subtestb</td>
<td>106.9*</td>
<td>18.0</td>
</tr>
<tr>
<td>Woodcock–Johnson Passage Comprehension Subtestb</td>
<td>97.5*</td>
<td>15.1</td>
</tr>
<tr>
<td>Woodcock–Johnson Broad Reading Scorea</td>
<td>103.2*</td>
<td>11.9</td>
</tr>
</tbody>
</table>

a Standard scores with a mean of 100, S.D. of 15.
b Raw score out of 44 possible.
c Words spoken, exclusive of false starts, mazes, fillers, and revisions divided by minutes in the speaking turn.
d Raw score out of 22 possible.
* Significant difference (t-test) from the ND group at p < 0.05.
control group on this test were not significantly higher (mean = 103.2, S.D. = 12.4) than those achieved by the hL/LD group (mean = 101.0, S.D. = 14.6). The Letter Word Identification and Passage Comprehension subtests of the Woodcock–Johnson Psychoeducational Battery (Woodcock & Johnson, 1989) were administered to describe reading skills. Scores achieved by the ND group on these tasks were higher (letter word identification mean = 117.9, S.D. = 18.7; passage comprehension mean = 108.7, S.D. = 12.8) than those achieved by the hL/LD group (letter word identification mean = 106.9, S.D. = 18.0; passage comprehension mean = 97.5, S.D. = 15.1). All tests were administered in random order to control for possible order bias. Means and standard deviations of behavioral measures obtained by the hL/LD and ND groups are reported in Table 1.

1.2. Materials and procedures

Each participant completed a computerized task that presented stimuli auditorily and collected response accuracy data. Participants were told that they would listen to an artificial language that had its own rules, but not told the nature of the testing that would happen after the period of exposure. The participants were exposed to the artificial language for 18 min. This time frame was based on the earlier Gómez (2002) study, which also used this exposure duration. Strings in the language consisted of three elements with non-adjacent word dependencies (e.g., pel kicey rud). Strings were presented with a natural, but standard intonation.

Subjects listened to strings from one of two grammars, which varied the dependent elements to avoid any inadvertent word-dependent effects. Grammar 1 (G1) strings took the form aXd, bXe, and cXf. Grammar 2 (G2) strings took the form aXe, bXf, and cXd. The elements a, b, and c were instantiated as pel, vot, and dak; d, e, and f as rud, jic, and tood. Each participant was exposed to one four conditions: Grammar 1, low variability; Grammar 1, high variability; Grammar 2, high variability; Grammar 2, low variability. Variability was manipulated by selecting the middle element from a set of either 12 or 24 words. The 24 X-elements in the high variability conditions were wadim, kicey, puser, fengle, coomo, loga, gople, tasp, hiftam, deecha, vamey, skiger, benez, gensim, feenam, laeljeen, chila, roosa, plizet, balip, malsig, suleb, nilbo, and wiffle. The set of 12 X-elements in the low variability conditions were drawn from the first 12 words in the list. Note that because there were half as many strings in the set of 12 X-elements relative to the set of 24 X-elements, each string from the 12 element set was heard twice as often during exposure as the strings from the 24 element set. Examples of aXd strings from G1 are pel loga rud and pel taspu rud. Examples of aXe strings from G2 are pel loga jic and pel taspu jic.

Following the exposure phase, participants were told that the strings they heard had followed rules generated according to grammatical word order and that they would now be tested on their knowledge of these rules. Six strings in each language that were common to both sets were used as test stimuli, yielding 12 test strings (6 from G1 and 6 from G2). Thus, all subjects received the same set of test items, but whether an item was grammatical depended on the grammar to which the participant was originally exposed. All test items for each grammar are provided in Table 2. Participants listened to the test strings and were
asked to select “Y” on the computer keyboard if the string followed the rules and “N” if the string violated the rules. Generalization was tested by exposing the participants to another 12 strings with novel middle elements. Again, they were asked to type “Y” if the strings followed the rules and “N” if the strings violated the rules. The generalization items are provided in Table 2. The 12 test and the 12 generalization items were intermixed for presentation. For both the test and generalization items, the participants were given as much time as they needed to respond.

2. Results

The overall performance by the hL/LD and ND groups for items heard during exposure and for generalization items is displayed in Figs. 1 and 2. There were 6 correct responses possible for the items heard during exposure and 6 generalization items, for a total of 12 correct items. There were 12 incorrect test items as well. Therefore, the total number of test items was 24.

The pattern of responses was statistically evaluated using a mixed ANOVA with group (hL/LD versus ND) and set size (12 middle elements versus 24 middle elements) as between group variables and grammaticality (correct accepts versus false positives) and item type (items heard during exposure versus generalization items) as within group variables. The dependent measure was the total number of correct responses. We predicted that the hL/LD group would perform poorly relative to the ND group overall. However, this between-group difference in the ANOVA did not reach statistical significance ($F = 0.47; df = 1, 40, p = 0.4967$). We also predicted that high variability for the middle elements (set size = 24) would be associated with better learning than low variability (set size = 12). Relevant to this prediction was a significant grammaticality effect ($F = 4.14, df = 1, 40$, $p = 0.0517$).
indicating that subjects accepted correct items more frequently than they accepted incorrect items. This effect was qualified, however, by a grammaticality \times set size interaction ($F = 5.52, df = 1, 40, p = 0.0238$). Post hoc testing comparing the number of correct responses (a total of 12 was possible) to chance (6) indicated that the grammaticality effect was significant for the set size of 24 (paired $t = 2.82, df = 1, 21; p = 0.0052$, one-tailed test), but not for the set size of 12 (paired $t = -0.26, df = 1, 21; p = 0.3986$, one-tailed). Finally, participants overall were more accurate for items heard during exposure than for items involving generalization of the underlying pattern ($F = 4.86, df = 1, 40, p = 0.0333$). This is further illustrated by examination of the number of items identified as correct that were correct (correct accepts) versus the items falsely identified as correct (false positives). This calculation reflects the degree of differential performance on test items that were heard previously or required generalization.
The average difference between these two types of items was 1.205 (S.D. = 3.60) for strings heard during exposure, whereas it was 0.841 (S.D. = 3.62) for generalization strings.

Given these significant effects, we asked whether the effect truly reflected performance that was above chance levels in the relevant conditions. To determine this, we evaluated the accuracy of responses collapsed across correct and incorrect items (12 trained and 12 generalization items) compared with a chance rate of 6 items in each condition. Accuracy rates for the combined sample of ND and hL/LD participants who were exposed to a set size of 12 did not exceed chance levels for either trained items or generalization items ($M = 6.14$ for trained items, $M = 5.55$ for generalization items). In contrast, significant differences from chance levels were found for those exposed to the set size of 24 for both the trained (paired $t = 2.89$, $df = 1, 21$, $p = 0.0044$, one-tailed, $M = 8.28$) and generalization items (paired $t = 2.64$, $df = 1, 21$, $p = 0.0075$, one-tailed, $M = 8.14$). Because we predicted that the ND and hL/LD group would benefit more from higher variability in the exposure set, we further examined these effects to determine whether they applied to both each group. Performance by the ND group again exceeded chance levels for both the trained items (paired $t = 2.69$, $df = 1, 10$, $p = 0.0113$, one-tailed, $M = 9.10$) and generalization items (paired $t = 2.17$, $df = 1, 10$, $p = 0.0277$, one-tailed, $M = 8.64$), but only for a set size of 24 middle elements (for a set size of 12, $M = 5.91$ for trained items, $M = 5.55$ for generalization items, paired $t = -0.66$, $df = 1, 10$; $p = 0.4603$). In contrast, performance by the hL/LD group did not reliably exceed chance at either set size or for either item type, paired $t = 1.48$, $df = 1, 10$; $p = 0.085$ (for set size 12, $M = 6.36$ for trained items, $M = 5.54$ for generalization items; for set size 24, $M = 7.45$ for trained items, $M = 7.64$ for generalization items).

3. Discussion

Consistent with the findings of Gómez (2002) and Gómez and Maye (2005), the ND participants demonstrated learning of non-adjacent contingencies (i.e., performance at above chance levels), but only when variability of the intervening element was high (i.e., set size of 24 elements). These subjects not only recognized items they heard during a relatively brief exposure set, but were able to generalize the underlying pattern to three-element strings they had not previously heard. However, we were not able to demonstrate a similar pattern of learning in the hL/LD group. Learning did not statistically exceed chance levels for either set size or item type.

Despite evidence of learning by the ND group and a contrasting lack of learning by the hL/LD group, the group difference was not statistically significant in the overall ANOVA. This is likely due to the relatively low level of learning demonstrated by the ND group. In the present study, the magnitude of the learning effect in the ND group did not provide a sizable interval above chance in order for the chance-level performance of the hL/LD group to show a statistically significant discrepancy. In fact, the magnitude of the group difference we obtained would require very large numbers of participants (e.g., in excess of a thousand) to reach statistical significance (Kraemer & Themann, 1987). Although the performance of our ND adults was consistent with rates seen for a range of artificial language learning studies that have included adult subjects (Gómez, in press), the current
subjects’ level of learning was weaker than that reported previously for the same paradigm (Gómez, 2002). Although the previous study offered little demographic information for comparison to the present sample, these students were enrolled at an Ivy League University. In contrast, our subjects were enrolled at a state university and their test data suggested functioning in areas including vocabulary, non-verbal IQ, and reading comprehension that are comparable to what would be expected for the general population. This raises the intriguing possibility that inter-subject differences in one or more aspect of cognitive functioning might facilitate this type of learning. This possibility remains to be explored in future studies.

The poor learning by the hL/LD group in this study is consistent with that seen by Plante et al. (2002). In that study, participants were asked to learn word order rules which constituted contingent relations among adjacent CVC elements. The results of that study also showed above chance learning by the ND group and chance-level performance by the hL/LD group, with a statistically significant between-group difference in overall performance. When these studies are considered together, the pattern of results suggest that adults with poor language skills are not particularly proficient in mapping statistical dependencies from even highly constrained input. This phenomenon does not appear to be specific to language-like, or even auditory stimuli. Tomblin and Zhang (2004) reported that adolescents with specific language impairment showed poorer implicit learning (as measured by reaction times) of patterns that predicted the appearance of a visual target in a non-verbal serial response task.

Why hL/LD subjects fail to show learning under conditions for which their ND peers do is not fully understood from the data at hand. The collection of available studies suggests that these individuals are less able to take advantage of statistical information in the input that they receive to benefit learning. This would explain why they learn both adjacent and non-adjacent contingencies poorly, and fail to derive benefit from variability in the learning of non-adjacent contingencies. A broad-based deficit such as this could reflect weaknesses in any of more basic cognitive capacities in addition to the mental operations involved in statistical mapping. For example, these tasks have in common significant attentional demands to the ongoing stream of stimuli. Likewise, deficits in the ability to hold stimuli in working memory and compare their order may reduce the capacity of learners to notice regularities.

It is also possible that hL/LD subjects have more constrained limitations in their ability to map statistical information. Indeed, the data available for this population is sparse in comparison to the number of conditions known to result in statistical learning (see Gómez & Gerken, 2000 or Jusczyk, 1997 for reviews). It may be that learning in the language impaired population would improve if they could be induced to adopt effective strategies for processing the input they receive. In the case of the non-adjacent contingencies studied here, Gómez (in press, 2002) has argued that learners may attempt to track relations among adjacent elements as a first-pass strategy, but must abandon this strategy to learn non-adjacent contingencies. When variability of the middle element is low, it appears that the non-adjacent relations lack salience, leaving learners to focus instead on the adjacent relations. High variability makes learning adjacent relations mentally taxing because the probability of any given middle element following the first element is low. When the mental load becomes taxing enough, ND learners may then seek out other forms...
of reliable structure (in this case the contingent relationship between the non-adjacent elements). In contrast hL/LD learners may employ a strategy of tracking the relations among adjacent elements, a strategy that is weak for them (cf. Plante et al., 2002), and continue to persist in this strategy even when their ND peers abandon it for a more productive one. If this is the case, then manipulating their tendency to focus on non-adjacent versus adjacent elements should improve learning under high variability conditions.

The weak abilities of those with language deficits to map statistical relations has interesting theoretical implications for the debate concerning the nature of both normal and impaired language acquisition. Those who have employed artificial language paradigms have argued that statistical learning is a basic skill that infants bring to the task of learning language (Gómez & Gerken, 2000; Saffran, 2003). Weak ability to map this information by individuals whose language skills are poor lends support to this notion. A reduced sensitivity to information that can be used to enhance learning may explain why language acquisition is slow in children with language impairments (but not bizarre in relation to the expected patterns). For example, poor ability to learn contingencies in word ordering may contribute to a delay in these children’s ability to combine words early on, or to use syntactic cues to assist comprehension later in development. The fact that these children eventually produce sentences with correct word orderings (even if the mean length tends to be shorter) suggests that the over-learning that comes with daily language exposure can mitigate this effect to some degree. Likewise, the pattern of common deficits during language acquisition would suggest that all statistical input is not likely to be equally difficult. For instance, some language features that reflect non-adjacent contingencies (e.g., third-person agreement) are substantially more likely to show omission errors than others (e.g., present progressive verb tense). Thus, it may be that language-impaired learners are able to capitalize on other types of cues to support learning in some language contexts that is weak or unavailable in other contexts. As such, discovery of conditions that make grammatical elements and relations salient and thus learnable would be of substantial interest to the field.

Acknowledgement

This work was funded in part by National Institute of Child Health and Development Grant HD42170.

Appendix A. Continuing education questions

1. Which of the following is true regarding non-adjacent contingencies?
   a. They are easier to learn than adjacent contingencies.
   b. They do not occur in natural languages.
   c. They describe structural relations between all morphemes.
   d. They can describe morphological relations that occur in English.
   e. They are not learned until adulthood.
2. How is the characteristic of variability thought to affect learning of non-adjacent contingencies as in the structure aXb (e.g., pel kicey rud)?
   a. It has no effect.
   b. Variability of non-adjacent a and b items aids learning.
   c. Variability of intervening X-items aids learning.
   d. Variability of the first a item aids learning.
   e. Variability for any of the items impedes learning.

3. In this study, *non-disabled* (ND) adults were . . .
   a. Able to learn, but not generalize under the low variability condition.
   b. Able to learn and generalize, but only under the low variability condition.
   c. Able to learn and generalize, but only under the high variability condition.
   d. Able to learn, but not generalize under the high variability condition.
   e. Unable to learn or generalize.

4. In this study, the adults with a history of language/learning disabilities (hL/LD) were:
   a. Able to learn, but not generalize under the low variability condition.
   b. Able to learn and generalize, but only under the low variability condition.
   c. Able to learn and generalize, but only under the high variability condition.
   d. Able to learn, but not generalize under the high variability condition.
   e. Unable to learn or generalize.

5. Poor learning of contingencies by adults with a history of language/learning disabilities (hL/LD):
   a. Appears to be limited to the verbal domain.
   b. Occurs for both verbal and non-verbal learning conditions.
   c. Is significantly affected by variability in the input.
   d. Is identical to the pattern seen in non-disabled adults.
   e. Is limited to non-adjacent contingencies rather than adjacent contingencies.

References


