

The Role of Prior Experience in Language Acquisition

Jill Lany, Rebecca L. Gómez, LouAnn Gerken

Department of Psychology, The University of Arizona, Tucson

Received 23 August 2005; received in revised form 19 August 2006; accepted 5 September 2006

Abstract

Learners exposed to an artificial language recognize its abstract structural regularities when instantiated in a novel vocabulary (e.g., Gómez, Gerken, & Schvaneveldt, 2000; Tunney & Altmann, 2001). We asked whether such sensitivity accelerates subsequent learning, and enables acquisition of more complex structure. In Experiment 1, pre-exposure to a category-induction language of the form $aX bY$ sped subsequent learning when the language is instantiated in a different vocabulary. In Experiment 2, while naïve learners did not acquire an $acX bcY$ language, in which aX and bY co-occurrence regularities were separated by a c -element, prior experience with an $aX bY$ language provided some benefit. In Experiment 3 we replicated this finding with a 24-hour delay between learning phases, and controlled for prior experience with the $aX bY$ language's prosodic and phonological characteristics. These findings suggest that learners, and the structure they can acquire, change as a function of experience.

Keywords: Language acquisition; Learning mechanisms; Prior experience; Transfer

1. Introduction

A promising paradigm for studying language acquisition involves using artificial grammars to investigate learning abilities (see Gómez & Gerken, 2000 for a review). This approach allows for precise control over the cues presented to learners, which can be difficult to achieve when using natural language materials. Another advantage is that these languages are novel to learners and, therefore, any learning can be attributed to controlled laboratory experience rather than prior experience with natural language or other relevant stimuli outside of the lab. In other words, using artificial languages increases the probability that observed effects are *not* due to prior experience with experiment materials. While this is an important control in such studies, learners acquiring their native language(s) are affected by prior experience in important ways. For example, more complex forms of learning can be scaffolded on simpler forms (cf. Newport, 1990; Elman, 1993; Goldowsky & Newport, 1993; Conway, Ellefson,

& Christiansen, 2003). This implies that while naïve learners may require large amounts of exposure, or highly salient and reliable cues, they may subsequently need less time or fewer cues to learn similar patterns, or may be able to learn patterns too difficult for naïve learners to acquire.

Recent studies of infant learning have begun to shed light on the role of prior experience in language acquisition, and on learning phonological patterns in particular. For example, Gerken (2004) and Thiessen and Saffran (2003) uncovered confounding effects of prior experience in 9-month-old infants. In both studies, English-learning infants were exposed to an artificial language in which words were characterized by stress patterns, and were then tested for discrimination between grammatical and nongrammatical strings using the Head-Turn Preference Procedure. Infants preferred test strings exhibiting a strong-weak stress pattern (the predominant stress pattern of English words) regardless of the pattern present in their familiarization language. Thus, infants' prior experience with their native language overrode the effects of the brief familiarization phase in the laboratory. Saffran and Thiessen (2003; Thiessen & Saffran, 2004) also tested the effects of prior experience more directly. They exposed infants to lists of words conforming to a phonological generalization (i.e., words might take the form CVCVC, or they might have a strong weak stress pattern), and then exposed them to a continuous stream of words also conforming to that generalization. Infants' prior experience with the phonological generalization allowed them to subsequently segment the fluent speech, indicating that prior experience can facilitate phonological learning.

These infant studies, however, do not require prior knowledge that is very abstract. For instance, strong-weak syllable sequences may share particular acoustic-phonetic manifestations that remain constant over experience. However, if prior learning is to be useful in learning syntax, a greater degree of abstraction is required. For example, in natural language phrases are composed of heads (the main element of a phrase, such as the verb in a verb phrase or the proposition in a prepositional phrase) and complements (any other material in the phrase). Within a language there tends to be a consistent ordering of heads relative to complements across phrase types. Most natural languages are either "head-initial" or "head-final." English is a "head-initial" language, such that the head of a phrase precedes its complement. In verb phrases, such as "walk to the corner," verbs (i.e., "walk") precede complements such as prepositional phrases (i.e., "to the corner"), and similarly in prepositional phrases such as "on my desk," prepositions (i.e., "on") precede complements such as noun phrases (i.e., "my desk").

Given that phrases of different types have common underlying structure, despite the fact that they may not overlap in their vocabulary or prosodic qualities, we might ask whether learners who have acquired this structure in one phrase type can subsequently generalize it to another. Many studies of adult learning and abstraction suggest that once learners are familiarized with an artificial language in one vocabulary, they can transfer to a language with the same abstract structure but novel vocabulary at test (e.g., Reber, 1969; Altmann, Dienes, & Goode, 1995; Gómez, 1997; Gómez, Gerken, & Schvaneveldt, 2000; Tunney & Altmann, 2001). While these studies indicate that learners are sensitive to abstract structure, they do not address how increased sensitivity changes the nature of subsequent learning. Rather than giving learners a test in a new vocabulary, it would be informative to train them on a new language and investigate how their learning differs from that of naïve learners. Therefore, we tested whether learners are faster to acquire a pattern after having already learned a similar pattern, and

	X₁	X₂	X₃	X₄	X₅	X₆
a₁	a ₁ X ₁	a ₁ X ₂	a ₁ X ₃	a ₁ X ₄	a ₁ X ₅	a ₁ X ₆
a₂	a ₂ X ₁	a ₂ X ₂	a ₂ X ₃	withheld	a ₂ X ₅	a ₂ X ₆

	Y₁	Y₂	Y₃	Y₄	Y₅	Y₆
b₁	b ₁ Y ₁	b ₁ Y ₂	b ₁ Y ₃	withheld	b ₁ Y ₅	b ₁ Y ₆
b₂	b ₂ Y ₁	b ₂ Y ₂	b ₂ Y ₃	b ₂ Y ₄	b ₂ Y ₅	b ₂ Y ₆

Fig. 1. A typical *aX bY* pattern. Learners are exposed to a subset of the grammatical pairings of markers and content-words. Learners are tested to see if they will generalize correctly to the withheld pairings.

whether they are better able to acquire a very difficult pattern after exposure to a structurally similar but simpler pattern. Note that this design differs from most previous designs in studies of adult learning, in which participants are trained on a language in one vocabulary, and their ability to transfer is assessed by testing them on grammatical and nongrammatical strings instantiated in another vocabulary.

The artificial language we used requires participants to learn word categories and their co-occurrence relationships, and is based on the languages used by Braine (1987), Frigo and MacDonald (1998), Gerken, Wilson, and Lewis (2005), and Gómez and LaKusta (2004). The language is composed of words belonging to the categories *a*, *b*, *X*, and *Y*. The language structure incorporates restrictions on how words from categories of different types can be combined within a string, such that *a* elements are paired with *X*-elements and *b*-elements with *Y*-elements, but not vice versa. The *aX bY* relation is analogous to the co-occurrence relationship between determiners and nouns, and also between auxiliaries and verbs, in English, in that it involves a functional element preceding a lexical one, and restrictions on the co-occurrences of categories of functional and lexical elements. Thus, acquiring this artificial language recruits sensitivities relevant for learning natural language syntax. Figure 1 presents a schematic of the *aX bY* pattern.

Learners are most successful with *aX bY* languages that have relatively few *a*- and *b*-elements relative to *X*s and *Y*s (Valian & Coulson, 1988). Also, learning is facilitated when the language contains cues signaling words' category membership (Braine, 1987; Frigo & MacDonald, 1998; Gerken et al., 2005; Gómez & LaKusta, 2004; Wilson, 2002). Thus, in all our experiments, *X*- and *Y*-words each had a distinctive ending (e.g., “-ee” and “-oo”).

The structure of the *aX bY* language is two-dimensional (Braine, 1987; Frigo and MacDonald, 1998). One dimension concerns the predictive relationships between *a*- and *b*-elements and the features cueing the category membership of *X*- and *Y*-elements. Specifically, each of the two *a*-elements predicts the distinctive ending present on *X*-elements, and the two *b*-elements predict the distinctive ending on *Y*-elements. Learners sensitive to these predictive relations can generalize to novel strings in which *a*- and *b*-elements are paired with *X*- and *Y*-elements containing the distinctive feature. These strings could contain a novel *aX* pairing (for instance in Fig. 1, *a₂X₄*, if *a₁X₄* but not *a₂X₄* was presented during learners' familiarization), or could involve a novel *X*-element with the distinctive feature (i.e., *a₁X₇*, if *X₇* was never presented during familiarization). The other dimension concerns the distribution

of elements within strings. Because each of the two a -elements always predicts X -elements, the a -elements share a distribution. Similarly, both b -elements always predict Y -elements and thus they share a distribution. Moreover, the distribution of elements in the $aX bY$ language is non-overlapping, such that the two a -elements predict one set of elements, while the two b -elements predict an entirely different set. Learners sensitive to the distributional properties of elements can generalize based on the knowledge that a -elements predict X -elements whereas b -elements predict Y -elements. Upon hearing the string $a_1 X_4$, they can generalize to the $a_2 X_4$ pairing, even if X_4 is not marked by the presence of the distinctive ending. Thus, these learners can generalize on the basis of the abstract $aX bY$ relation, even in the absence of the distinctive features. Interestingly, previous research done with $aX bY$ language indicates that sensitivity to this level of abstraction is heavily dependent on the presence of distinctive features cueing category membership (Braine, 1987; Frigo & MacDonald, 1998; Gerken, Wilson, & Lewis, 2005; Gerken, Wilson, Gómez, & Nurmsoo, in prep). For example, noticing that a - and b -elements predict nonoverlapping sets of elements requires that a majority of elements comprising those sets contain some other cue to their category membership (e.g., distinctive endings). Thus, the two dimensions of structure intersect, such that sensitivity to the fact that a - and b -elements have nonoverlapping distributions typically rests on noticing the co-occurrence relationships between as and bs and the highly frequent endings on X - and Y -elements

In sum, there are two kinds of abstraction possible in learning an $aX bY$ language; the fact that as predict words with a particular feature and bs predict words with a different feature, and the fact that as and bs predict nonoverlapping sets of elements. Generalization to strings containing features that cue category membership indicates sensitivity to the association between each of the a - and b -elements and the distinctive endings, but does not require sensitivity to the fact that as and bs predict nonoverlapping sets. It does indicate abstraction beyond sensitivity to heard strings, and thus represents an important aspect of learning co-occurrence relations between categories. We tested this form of learning in the present experiments.

Three experiments examined the effects of prior experience on learning. In Experiment 1 we exposed participants to either 18 or 6 blocks of an $aX bY$ language. A third group of participants was exposed to 18 blocks of an $aX bY$ language, and then transferred to 6 blocks of a second $aX bY$ language with the same underlying structure as the first but novel vocabulary. The transfer group learned the language better than naïve learners exposed to 6 blocks of the language, and performed as well as naïve learners exposed to 18 blocks, showing that prior experience can speed learning of parallel structure. In Experiment 2, we exposed learners to an $aX bY$ language, and transferred them to an $acX bcY$ language in a new vocabulary. Like the $aX bY$ language, the transfer language required learning co-occurrence restrictions on words from different categories. However, it differed in that the categories were separated by a c -element. Naïve learners of the $acX bcY$ language failed to discriminate between grammatical and ungrammatical strings, but learners previously exposed to an $aX bY$ language showed some evidence of discrimination, indicating that prior experience may facilitate learning a more difficult pattern. In Experiment 3 we found that learners benefit from prior experience even if there is a 24-hour delay between exposure to the $aX bY$ and $acX bcY$ languages. Additionally, we ruled out the possibility that benefits conferred by prior experience resulted

from experience with other dimensions of the language (i.e., the speaker's voice and speaking rate, or the language's vocabulary or phonological characteristics, etc.).

2. Experiment 1

In Experiment 1 we tested whether participants would learn a category induction language when given a relatively long familiarization (18 blocks), but not when familiarization was shorter by two-thirds (6 blocks), and whether learners given prior experience with a pattern would subsequently learn a similar pattern under these abbreviated familiarization conditions.

2.1. Method

2.1.1. Participants

One hundred forty-five University of Arizona undergraduates participated in this experiment for course credit. Data from 7 participants were discarded for failure to pay attention (e.g., answering cell-phone calls, or playing with the keyboard while they were supposed to be listening quietly). Data from another 18 participants were discarded for failure to meet test-performance criteria. As part of the instructions at the start of each test block, participants were specifically instructed that half of their 16 responses should be "yes" and half should be "no." We did not expect that participants would give equal numbers of "yes" and "no" responses, however participants who responded with only one answer type clearly failed to listen to or understand the test instructions. Thus, we did not include data from either participants who gave all "yes" or all "no" responses. Of the remaining 120 participants, 40 each were in the 6 Block, 18 Block, and Transfer groups.

2.1.2. Materials

We constructed two versions of an $aX bY$ language, each with its own vocabulary (See Appendix A). We used two versions so that we could train and test participants on one version of the language, and then transfer them to a second version in a subsequent train-test phase. Each version had two grammars, $aX bY$ and $aY bX$, such that grammatical category pairings in one grammar were nongrammatical in the other. Nongrammatical strings thus contained the same words and word endings as grammatical ones, but differed in their combinations of word categories. For example, strings from Version A, G1 were *alt tamoo* and *erd sufficee*, and strings from G2 were *alt sufficee* and *erd tamoo*. In each version there were two a -elements and two b -elements, six X -elements and six Y -elements, for a total of 24 legal strings in each of the familiarization languages. In each grammar, 4 strings were withheld from familiarization to test generalization, and thus familiarization materials were 20 unique strings (10 aX and 10 bY in G1, and 10 aY and 10 bX in G2.).

The test materials consisted of 16 strings, half grammatical and half nongrammatical. There were four grammatical strings that had been withheld during familiarization (grammatical-unheard strings, or GUH), and four strings that had been presented during training (grammatical-heard strings, or GH). The nongrammatical strings were four each of

two types of strings from the other grammar. There were NGH strings (the GH strings in the other grammar), which were “matched” to GH strings in that they consisted of the same X - and Y -elements paired with an ungrammatical a - or b -element. Thus, if a_1X_1 , or “alt feenoo,” was a GH test string in G1, then b_1X_1 , or “erd feenoo,” was a NGH test string. The NGUH test strings were nongrammatical strings matched to GUH strings, such that if a_2X_2 , or “pel wifoo,” was a GUH test strings in G1, then b_2X_2 , or “erd juhnoo” was a NGUH test string. Because our measures of learning assessed discrimination between grammatical strings and their nongrammatical counterparts, we eliminated potential effects due to the possibility that some strings were endorsed at high rates because the X s and Y s were simply more memorable, easier to learn, etc. Additionally, because the X s and Y s occurring in strings withheld from familiarization occurred less frequently than the X s and Y s occurring only in heard strings, we were able to match the frequency of occurrence of the X s and Y s in the grammatical and nongrammatical test strings.

To create the language materials, we recorded a trained female speaking the strings in an animated voice. Because each version’s two grammars differed only in the ways words were combined, the same token of each word was used in strings from both grammars. This ensured that participants could not distinguish between grammatical and ungrammatical strings based on idiosyncrasies in the pronunciation of words in one grammar or the other. The strings were digitized and edited with SoundEdit software. Strings were approximately 1.7 s in duration, and were separated by 1 s of silence when presented during familiarization. Words within a string were separated by 10 ms of silence.

2.1.3. Procedure

There were 12 conditions resulting from the between-participant manipulations of familiarization condition (18 Blocks, 6 Blocks, Transfer), version (Version A, Version B), and grammar (Grammar 1, Grammar 2).

Aside from instructions at the start of the familiarization phase, which were delivered by the experimenter, the entire experiment was conducted on a PC running SuperLab Pro 2.01 software. Participants were given verbal and written instructions stating that they would listen to an artificial language in which nonsense words were arranged in strings or groups. They were told to listen carefully, as they would be asked questions about the language later. In all conditions, each familiarization block consisted of 20 strings, the order of which was randomized for each participant and each familiarization block, and lasted about one minute. The 18 Block and 6 Block participants listened over headphones to 18 blocks (approximately 18 minutes) and 6 blocks (approximately 6 minutes) of their training language, respectively. Transfer participants were familiarized in turn with both Versions A and B of the category-induction language. They heard 18 blocks of one version of the language, and then 6 blocks of the second version. Thus, half of the Transfer participants were exposed to 18 blocks of Version A and then 6 blocks of Version B, and the other half were exposed to Version B followed by Version A. The Transfer group’s first familiarization phase was identical to that of the 18 Block group, and their second was identical to that of the 6 Block group. The only difference was a modification to the initial instructions informing Transfer participants that the experiment took place in two phases. The learning phases were separated from each other by a brief pause in which participants were told that they were done with the first phase and

would begin the second next. They were then given instructions for the second phase, which consisted of telling them that they would listen to more strings from an artificial language, and that they should pay attention, as they would again be tested.

At test, all participants answered two blocks of 16 test questions, each in a different random order and separated from each other by a brief pause. Transfer participants were tested at the end of each familiarization phase. Before each block of test questions, participants were given both verbal and written instructions that the strings they had heard adhered to a set of rules involving word order. They were also told that they would hear a set of 16 strings, half of which followed the rules, and half of which did not. Participants were given the opportunity to ask the experimenter for clarification of the instructions before beginning the test. Participants heard the test strings one at a time, and made yes/no judgments on the grammaticality of each string using the “Y” and “N” keys on their keyboard. Participants who were sensitive to the predictive relationships between *as* and the endings on *X*-elements, and between *bs* and the endings on *Y*-elements, should endorse grammatical test strings, including those withheld during training, more often than nongrammatical ones.

2.2. Results and discussion

Preliminary analyses indicated that the two versions and grammars of the language with which participants were familiarized were learned equally well, and we therefore collapsed across these conditions in all subsequent analyses. Test performance for the three familiarization conditions, reported as endorsement rates to test items, can be found in Table 1.

2.2.1. Naïve learners (18 and 6 Block participants)

We performed a three-way mixed ANOVA to determine whether the groups differed in their ability to distinguish grammatical from nongrammatical strings. There was a between-participant factor of familiarization condition (6 Blocks versus 18 Blocks). The within-participant factors were test block (test-block 1 versus test-block 2), and discrimination type (grammatical-heard, or GH, versus grammatical-unheard, or GUH).

We used two complementary dependent measures of discrimination.¹ One measure was a difference score obtained by subtracting participants’ endorsement rates to grammatical test strings from their endorsement rates to nongrammatical ones. The other measure was *d'*. In all

Table 1
Experiment 1 Mean Endorsement Rates with Standard Error of the Mean in Parentheses

	Block 1				Block 2			
	GH	GUH	NGH	NGUH	GH	GUH	NGH	NGUH
6 Block	.71 (.030)	.69 (.039)	.57 (.044)	.69 (.037)	.70 (.034)	.55 (.040)	.61 (.037)	.53 (.034)
18 Block	.85 (.026)	.72 (.034)	.61 (.044)	.63 (.047)	.83 (.030)	.64 (.035)	.54 (.047)	.47 (.040)
Transfer	.78 (.040)	.69 (.045)	.43 (.048)	.48 (.056)	.74 (.040)	.71 (.044)	.51 (.056)	.37 (.049)

Note: GH refers to grammatical-heard test strings, and NGH to their matched nongrammatical counterparts. Similarly, GUH refers to grammatical-unheard test strings, and NGUH to their nongrammatical counterparts.

Table 2

Experiment 1 Mean Discrimination for GH and GUH strings in terms of Difference Scores and d' , with Standard Error of the Mean in Parentheses

	Block 1				Block 2			
	GH		GUH		GH-NGH		GUH-NGUH	
	GH - NGH	d'	GUH-NGUH	d'	GH-NGH	d'	GUH-NGUH	d'
6 Block	.14 (.050)	.45 (.289)	-.01 (.043)	.02 (.268)	.09 (.053)	.41 (.316)	.02 (.052)	.05 (.292)
18 Block	.24 (.057)	1.49 (.428)	.09 (.059)	.54 (.414)	.30 (.056)	1.92 (.429)	.17 (.060)	.94 (.348)
Transfer	.34 (.074)	2.46 (.492)	.21 (.085)	1.58 (.594)	.23 (.079)	1.63 (.563)	.34 (.070)	2.17 (.519)

Note: GH refers to grammatical-heard test strings, and NGH to their matched nongrammatical counterparts. Similarly, GUH refers to grammatical-unheard test strings, and NGUH to their nongrammatical counterparts.

analyses, GH discrimination reflects the ability to distinguish grammatical-heard test strings from their nongrammatical mates (i.e., NGH strings), and is therefore an index of sensitivity to familiar phrases. Similarly, GUH discrimination reflects the ability to distinguish grammatical-unheard test strings from NGUH ones, and measures generalization to novel phrases. Table 2 contains mean discrimination scores for the three familiarization conditions.

Analyses using difference scores revealed the predicted effect of familiarization condition, $F(1, 78) = 7.35$, $p = 0.008$, with 18 Block participants showing greater overall discrimination ($M = 0.20$, $SE = 0.037$) than 6 Block participants ($M = 0.06$, $SE = 0.037$). There was also a main effect of discrimination type, $F(1, 78) = 17.15$, $p < 0.001$, with discrimination for grammatical-heard (GH) strings ($M = 0.19$, $SE = 0.029$) better than that for grammatical-unheard (GUH) strings ($M = 0.07$, $SE = 0.031$). No other main effects or interactions reached significance. Analyses using d' as a dependent measure confirmed these findings.

We next asked whether 18 Block and 6 Block participants were able to discriminate GH and GUH strings from their nongrammatical counterparts. Participants could respond to GH strings appropriately on the basis of either grammaticality or familiarity, and thus a correct response to these strings does not necessarily reflect sensitivity to the underlying pattern. In contrast, higher endorsement rates to GUH than to NGUH strings indicate that participants have learned the language's co-occurrence restrictions (i.e., the relationship between the a - and b -words and the distinctive endings on the X - and Y -words). Given that a difference score of zero represents a lack of discrimination, we compared each groups' difference scores for GH and GUH items in each test block to zero using one-sample t tests.² For participants in the 18 Block group, GH discrimination in both test-blocks differed from zero, $t(39) \geq 4.28$, $ps < 0.001$. GUH discrimination was not significant in test-block 1, $t(39) = 1.54$, $p = 0.113$, but was significant in test-block 2, $t(39) = 2.84$, $p = 0.007$. This increase in discrimination for GUH strings likely reflects learning during testing, thus cannot be taken as an index of learning during familiarization. The 6 Block groups' GH discrimination was greater than chance in test-block 1, $t(39) = 2.76$, $p = 0.009$, but not in test-block 2, $t(39) = 1.66$, $p = 0.104$. Their GUH discrimination did not differ from chance in either test block, $ts(39) \leq 0.36$, $ps \geq 0.719$. The same analyses of the two groups' discrimination using d' as a dependent

measure largely confirmed these findings, the only exception being that discrimination was not observed in the 6 Block participants, $t_s(39) \leq 1.57$, $p_s \geq 0.124$.

In sum, participants exposed to 18 blocks of familiarization learned more than participants exposed to 6 blocks. They showed robust discrimination for GH strings in both test blocks, while the 6 Block learners' GH discrimination was significant only in the first block of testing in endorsement-rate analyses, and never reached significance in the d' analyses. Neither group showed robust generalization to grammatical-unheard test strings.

2.2.2. Transfer learners

While brief exposure (i.e., 6 blocks) to the $aX bY$ language may be enough to induce weak sensitivity to familiar strings from the language, learners with more extensive exposure (i.e., 18 blocks) showed better discrimination. Learners in both these conditions were naïve, in that their experience with the $aX bY$ language constituted their first exposure to this type of pattern within the experimental session. We thus asked whether extensive exposure is always necessary for learning, or whether learners with prior experience would learn a new instantiation of the pattern more rapidly. We tested this by comparing the Transfer group's performance to that of the 6 Block group. Both groups had 6 blocks of exposure to the $aX bY$ language in question, however, the Transfer group first had 18 Blocks of exposure to another version of the $aX bY$ language. We performed a three-way mixed ANOVA with a between-participant factor of familiarization condition (Transfer versus 6 Blocks), and within-participant factors of test block (1 versus 2), and discrimination type (GH versus GUH). As before, we conducted the analyses using both difference scores and d' as dependent measures. See Table 2 for the Transfer group's mean GH and GUH discrimination.

Analyses using difference scores revealed that the Transfer group discriminated between grammatical and nongrammatical strings at a significantly higher rate ($M = 0.28$, $SE = 0.052$) than the 6 Block group ($M = 0.06$, $SE = 0.052$), $F(1, 78) = 9.30$, $p = 0.003$. One-sample t tests using difference scores indicated significant discrimination for both old (GH-UG) and novel (GUH-UG) strings in both blocks of testing, $t_s(39) \geq 2.48$, $p_s \leq 0.018$. D' analyses confirmed these findings. Thus, in addition to demonstrating enhanced learning relative to naïve 6 Block learners, Transfer participants showed robust generalization to grammatical strings that had been withheld during familiarization.

We next compared the Transfer and 18 Block groups to determine whether they learned as much as participants given extensive experience with one version of the $aX bY$ language. Analyses using difference scores revealed no main effect of familiarization type: the Transfer group's overall performance on the test ($M = 0.28$, $SE = 0.043$) did not differ from the 18 Block group's ($M = 0.21$, $SE = 0.043$), $F(1, 78) = 1.06$, $p = 0.306$. There was a main effect of discrimination type, $F(1, 78) = 4.55$, $p = 0.036$, with greater GH discrimination ($M = 0.28$, $SE = 0.043$) than GUH discrimination ($M = 0.21$, $SE = 0.043$). This effect was qualified by a marginally significant interaction between familiarization condition and discrimination type, $F(1, 78) = 3.53$, $p = 0.064$.³ Paired-sample t tests indicated that the 18 Block groups' GH discrimination ($M = 0.27$, $SE = 0.046$) was significantly greater than their GUH discrimination ($M = 0.13$, $SE = 0.048$), $t(39) = 3.55$, $p = 0.001$, whereas Transfer participants' discrimination was equally high for GH strings ($M = 0.29$, $SE = 0.074$) and for GUH strings ($M = 0.28$, $SE = 0.072$), $t(39) = .15$, $p = 0.879$. This finding was further

qualified by a marginally significant three-way interaction between familiarization condition, discrimination type, and test block, $F(1, 78) = 3.60$, $p = 0.062$. This interaction reflects the fact that the 18 Block group's GH discrimination was significantly better than their GUH discrimination in test-blocks 1 and 2, $t(39) \geq 2.21$, $ps \leq 0.033$, whereas, the Transfer group's GH discrimination was only marginally better than GUH discrimination in test-block 1, $t(39) = 1.93$, $p = 0.061$, and did not differ in test-block 2, $t(39) = -1.64$, $p = 0.11$.

An ANOVA using d' as the measure of discrimination revealed the same pattern of findings, with only two minor differences. First, the two-way interaction between discrimination type and familiarization condition was significant, $F(1, 78) = 3.95$, $p = 0.05$, while it had been marginal in the endorsement rate analysis. Second, the Transfer group's discrimination for test-block 1 GH strings was significantly greater than that for GUH strings, $t(39) = 2.15$, $p = 0.038$, while the difference was marginal in the endorsement rate analyses.

Both the d' and difference score analyses resulted in a significant or near significant interaction between familiarization condition and discrimination type, and a marginal three-way interaction including test block in addition to these factors. How can we interpret these findings? Together they suggest that performance in the Transfer group was more stable. The 18 Block group's discrimination for GH strings was better than that for GUH strings in both blocks of testing, and they failed to discriminate GUH strings from their nongrammatical counterparts in the first test block. In contrast, the Transfer learners' discrimination for GH strings was better than their discrimination for GUH strings in test-block 1, but did not differ in the second test block, and they discriminated both GH and GUH strings from nongrammatical strings at above-chance rates in both test blocks. Thus, while the 18 Block learners showed little evidence of generalization to unheard strings, Transfer learners, while showing an initial advantage for heard strings, generalized to unheard strings in both test blocks.

Finally, there was an interaction between test-string type and test block for the endorsement rate analyses, $F(1, 78) = 5.59$, $p = 0.021$, as well as the d' analyses, $F(1, 78) = 3.25$, $p = 0.075$. Because this interaction does not involve our key variable of interest (familiarization condition), it is of minimal importance.

In sum, the Transfer learners performed significantly better than the 6 Block learners, and in an important way surpassed the 18 Block learners by being the only group to show robust generalization to unheard test strings. Their performance is impressive given that the 18 block group had three times more exposure to that version of the language, and suggests that exposure to a pattern within the experimental session influenced the course of subsequent learning.

3. Experiment 2

In Experiment 1, Transfer learners were exposed to two versions of an $aX bY$ language that had different vocabulary, but shared several other characteristics. For example, strings from both versions were composed of two words, the first monosyllabic and the second bisyllabic, and the distinctive features of X s and Y s took the form of word-endings (the sounds “-ee” and “-oo” in one version, and “-it” and “-ul” in the other). Moreover, both language versions

involved adjacent co-occurrence restrictions on word categories. An important question is whether learners benefit from prior experience *only* in cases where two patterns are highly perceptually similar, or in cases with directly analogous structure. If so, learners familiar with an adjacent relationship will only benefit from experience if the transfer language consists of adjacent relationships, strings of same length, and similar prosodic characteristics, and thus facilitation due to prior learning would be limited to highly similar structure. We therefore tested whether prior experience facilitates learning a pattern with the same abstract structural characteristics, but more dissimilar surface characteristics than the transfer language used in Experiment 1.

This question has clear relevance to natural language acquisition. Revisiting the analogy between our *aX bY* language and the co-occurrence relationship between determiners and nouns, and auxiliaries and verbs, note that determiners and nouns are sometimes separated by an adjective (e.g., “the rubber ducky,” or “a yellow flower”). Similarly, auxiliaries and verbs can be separated by an adverb (e.g., “is quickly running”), or a negative element (e.g., “is not coming”). Thus, while these co-occurrence relationships are often adjacent, they can occur at a distance in more complex sentences. As a result, instantiations of this pattern will be of variable length, will have different prosodic characteristics, and will involve both adjacent and nonadjacent relationships. While nonadjacent relationships occur in natural languages, they can be more difficult for learners to acquire than adjacent ones (Gómez, 2002; Newport & Aslin, 2004). This raises the question of how learners become sensitive to the co-occurrence relationships in these complex constructions. Perhaps prior experience with co-occurrence relationships in simple constructions, in which relevant elements are adjacent, increases sensitivity to similar nonadjacent relationships. Speech addressed to infants and toddlers tends to consist of shorter, simpler utterances than speech to adults (Newport, Gleitman, & Gleitman, 1977; Pine, 1994), and thus presentation of simpler instances of a pattern before more complex ones has ecological validity. We, therefore, tested whether learners given experience with an *aX bY* language would have an advantage over naïve participants in learning a language in which these elements were separated by an intervening element.

3.1. Method

3.1.1. Participants

Seventy-nine University of Arizona undergraduates participated for course credit. Data from 14 participants were excluded for answering either “yes” or “no” to all questions in either test block. One additional participant reported hearing loss, and his data were excluded. Of the remaining 64 participants, 32 were in the Transfer condition.

3.1.2. Materials

The language used in Experiment 2 took the form *acX bcY*, and was identical to the *aX bY* language from Experiment 1, with the exception that a *c*-element separated *a*- and *X*-elements, and *b*- and *Y*-elements (see Appendix B). There were two versions of the language with distinct vocabulary created from Versions A and B of the *aX bY* language used in Experiment 1. Each version of the *acX bcY* language had 3 monosyllabic *c*-elements. In Version A, they were “hes,” “kaf,” and “sij,” and in Version B they were “tash,” “fis,” and nep.” As in Experiment

1, each version of the language had 2 grammars, such that the pairings were $acX bcY$ in one grammar, and acY and bcX in the other. Thus, the nonadjacent combinations in one grammar were illegal in the other. Examples of strings from Version A, G1 are *alt hes tamoo* and *erd hes suffee*, and examples from G2 are *alt hes suffee* and *erd hes tamoo*. Each grammar generated 72 strings, 12 of which were withheld from familiarization to test generalization.

There was no systematic relationship between c -elements and the elements adjacent to them; all c -elements were preceded equally often by a - and b -elements, and were followed by X -elements as often as by Y -elements. Moreover, the ac and bc combinations occurred with relatively high frequency, thus the adjacent relationships are likely to be learned at the expense of the critical nonadjacent ones (see Gómez, 2002). However, because adjacent relationships could not be used to discriminate between strings from the two grammars, learners focusing on them would be unable to distinguish grammatical from ungrammatical strings.

Test materials were composed of 12 strings that were withheld from training (GUH strings), 12 familiar strings (GH strings), and 12 each of the NGH and NGUH strings taken from the unheard grammar. The strings were divided into two sets, and were presented in two separate test blocks, each consisting of 24 strings. There were 6 each of the GH, GUH, NGH and NGUH strings in each block (see Appendix B).

The same person who generated the stimuli used in Experiment 1 recorded strings, and they were generated in the same manner. Words within a string were separated by 10 ms of silence, and strings were separated by 1 s of silence. Strings were an average of 2 s in duration. Each block of familiarization consisted of 60 unique strings, and was approximately 3 minutes in duration.

3.1.3. Procedure

There were eight between-participant conditions, resulting from the manipulation of familiarization (Control versus Transfer), version (Version A versus Version B), and grammar (Grammar 1 versus Grammar 2).

Both Control and Transfer participants were trained on 4 blocks (about 12 minutes) of an $acX bcY$ language; however, the Transfer group was first trained and tested on 18 minutes of one of the $aX bY$ languages from Experiment 1. Version was counterbalanced in the Transfer condition and equal numbers of control participants were trained on Version A and Version B. All other aspects of the familiarization and test procedure were identical to Experiment 1.

3.2. Results and discussion

There were no differences in learning as a function of the version or grammar to which participants were exposed. Thus, we collapsed across these factors in all subsequent analyses. Endorsement rates are displayed in Table 3.

We tested whether the Transfer group learned the $acX bcY$ language better than the Control group using a three-way mixed ANOVA with a between-participant factor of familiarization condition (Control versus Transfer) and within-participant factors of test block (1 versus 2), and discrimination type (GH versus GUH). The analyses again used difference scores and d' as dependent measures, and Table 4 contains means for both measures of discrimination.

Table 3
Experiment 2 Mean Endorsement Rates with Standard Error of the Mean in Parentheses

	Block 1				Block 2			
	GH	GUH	NGH	NGUH	GH	GUH	NGH	NGUH
Control	.70 (.039)	.69 (.037)	.72 (.037)	.68 (.040)	.58 (.037)	.61 (.033)	.59 (.031)	.51 (.033)
Transfer	.71 (.031)	.68 (.035)	.62 (.040)	.55 (.042)	.63 (.040)	.60 (.040)	.58 (.040)	.53 (.037)

Note: GH refers to grammatical-heard test strings, and NGH to their matched nongrammatical counterparts. Similarly, GUH refers to grammatical-unheard test strings, and NGUH to their nongrammatical counterparts.

Analyses using difference scores indicated that discrimination did not differ between the Transfer and Control groups ($M = 0.08$, $SE = 0.026$) in the Transfer group, and $M = 0.03$, $SE = 0.025$ in the Control), $F(1, 62) = 2.27$, $p = 0.137$.

One-sample t tests using difference scores indicated that the Control group did not discriminate between GH and NGH strings in either test-block, $t_s(32) \leq 0.234$, $p_s \geq 0.630$. They failed to discriminate between GUH and NGUH test strings in test-block 1 ($M = 0.01$, $SE = 0.049$, $t(31) = 0.23$, $p = 0.817$). Their GUH discrimination in test-block 2 ($M = 0.10$, $SE = 0.037$) was significant, however, $t(31) = 2.61$, $p = 0.014$, likely reflecting learning during test.⁴ The Transfer group’s GH discrimination was marginally significant in test-block 1, $t(31) = 1.82$, $p = 0.079$, and their GUH discrimination was significant, $t(31) = 2.48$, $p = 0.019$. Neither GH nor GUH discrimination reached significant levels in the second block of testing, $t_s(31) \leq 1.42$, $p_s \geq 0.163$.

Because Transfer participants showed some ability to discriminate between grammatical and nongrammatical strings in the first test-block, while Control participants showed no such ability, we compared the two group’s on their test-block 1 performance. A two-way ANOVA comparing the two groups’ test-block 1 performance (with the between-participant factor of familiarization condition and the within-participant factor of discrimination type) indicated that the Transfer group’s discrimination ($M = 0.11$, $SE = 0.041$) was better than the Control group’s ($M = -0.010$, $SE = 0.041$), $F(1, 62) = 4.021$, $p = 0.049$.) This finding,

Table 4
Experiment 2 Mean Discrimination for GH and GUH strings in terms of Difference Scores and d' , with Standard Error of the Mean in Parentheses

	Block 1				Block 2			
	GH		GUH		GH		GUH	
	GH-NGH	d'	GUH-NGUH	d'	GH-NGH	d'	GUH-NGUH	d'
Control	-.02 (.051)	-.36 (.312)	.01 (.049)	.03 (.342)	-.02 (.045)	.03 (.312)	.10 (.037)	.34 (.238)
Transfer	.09 (.049)	.32 (.312)	.13 (.051)	.71 (.3210)	.04 (.060)	.34 (.275)	.08 (.052)	.65 (.238)

Note: GH refers to grammatical-heard test strings, and NGH to their matched nongrammatical counterparts. Similarly, GUH refers to grammatical-unheard test strings, and NGUH to their nongrammatical counterparts.

while exploratory, provides some indication that learners benefit from prior experience when learning nonadjacent relationships.

The d' analyses painted a slightly different picture. The three-way ANOVA revealed a marginal effect of familiarization condition, $F(1, 62) = 3.3$, $p = 0.074$, with the Transfer group showing better discrimination ($M = .51$, $SE = 0.193$) than the Control group ($M = 0.01$, $SE = 0.193$). There was also an effect of discrimination type, $F(1, 62) = 4.45$, $p = 0.039$, with discrimination for GUH strings ($M = 0.43$, $SE = 0.161$) better than that for GH strings ($M = 0.08$, $SE = 0.158$). Paired sample t tests indicated that Control participants did not discriminate GH strings from nongrammatical ones in either test block, $t_s(31) \leq 1.27$, $ps \geq 0.212$. They did not discriminate GUH strings from nongrammatical ones in test-block 1, $t(31) = 0.08$, $p = 0.931$, although they did in test-block 2, $t(31) = 2.44$, $p = 0.021$. Again, this finding reflects, at most, learning during test. In contrast, Transfer participants showed discrimination for GUH strings in both test blocks, $t_s(31) \geq 2.13$, $ps \leq 0.041$, while discrimination for GH strings was not significant in either test block, $t_s(31) \leq 0.95$, $ps \geq 0.348$.

In sum, Experiment 2 provides preliminary evidence that Transfer learners' prior experience with an $aX bY$ language facilitated their learning of the more difficult $acX bcY$ language. Endorsement rate analyses indicated that this language was extremely difficult for naïve learners, who demonstrated no ability to distinguish even heard grammatical strings from nongrammatical ones. Transfer learners discriminated grammatical from nongrammatical GH and GUH strings during the first test block, although the difference for GH strings was marginal, and their test-block 1 performance was significantly better than the Control's. The d' analyses indicated that the Transfer learners' overall discrimination was marginally better than the Control learners. Transfer learners also showed significant discrimination for GUH but not GH strings across both blocks of testing, while Control learners showed no reliable evidence of learning. The two analyses, while each suggesting a benefit for transfer learners, failed to converge. Thus, we attempted to provide stronger evidence for the effects of prior experience in Experiment 3.

4. Experiment 3

In Experiment 2, Transfer learners' two train-test phases lasted approximately an hour, and thus they may have been fatigued by the end of the session. We might see clearer evidence of learning if they were given a break between learning phases. We addressed this possibility in Experiment 3 by conducting the two train-test phases on consecutive days rather than in immediate succession. While we hypothesized that the 24-hour delay might benefit learners, it is possible that a greater delay between learning experiences could also diminish the effects of prior experience. Thus, this design permitted us to assess whether benefits from prior experience are limited to conditions of very short delay, or whether more distant prior experience also benefits learning.

Additionally, in Experiments 1 and 2, we compared transfer learners to naïve controls who lacked prior exposure to any aspect of the $aX bY$ language. Reber and Perruchet (2003) demonstrated that use of naïve controls can be problematic, as they sometimes demonstrate inherent

biases for some language strings over others at test. The presence of such biases makes it difficult to interpret differences between such control groups and groups familiarized with the language. Importantly, Reber and Perruchet found that familiarizing control group learners with randomized materials could ameliorate these problematic effects. Another methodological drawback to using a naïve control is that experimental and control learners are not equally experienced with nonspecific aspects of the language and experimental procedure. For example, it is possible that Transfer learners benefited not from exposure to the $aX bY$ structure, but instead from experience with more general characteristics of the materials and experimental procedure. To remedy the problems with the control groups used in Experiments 1 and 2, Experiment 3 included a control group exposed to a language generated by the speaker of the $aX bY$ language, with identical vocabulary and prosody, but which did not conform to the $aX bY$ generalization in which a -elements predicted a set of words with a shared feature, and b -elements predict another set of words with a different shared feature. If these learners fail to acquire the $acX bcY$ language, we will have evidence that prior exposure to the surface characteristics alone of the $aX bY$ language is not sufficient to facilitate subsequent learning. Thus, our third experiment builds on the findings of Experiment 2 by testing the long-term effects of prior experience on learning, and by ruling out alternative explanations for observed benefits to learning.

4.1. Method

4.1.1. Participants

Seventy University of Arizona undergraduate students participated in exchange for course credit. The data from 10 were excluded for failure to comply with instructions ($N = 4$), experimenter error ($N = 5$), and hearing loss ($N = 1$). Thirty were randomly assigned to each group.

4.1.2. Materials

The language materials consisted of the $aX bY$ and $acX bcY$ languages used in Experiments 1 and 2. We also recombined the $aX bY$ materials to create an Uncued language, which was presented to Control participants on day 1. In Grammar 1 of the original $aX bY$ language, a -elements were paired with X_{1-6} and b -elements were paired with Y_{1-6} , and the opposite pairings held in Grammar 2. In the Uncued language, Grammar 1 a -elements were paired with X_{1-3} and Y_{4-6} , while b -elements were paired with Y_{1-3} and X_{4-6} . In Grammar 2, the pairings were reversed, such that a -elements occurred with Y_{1-3} and X_{4-6} , and b -elements occurred with X_{1-3} and Y_{4-6} .

A total of 24 strings were generated by the grammar of the Uncued language, though 4 (one each of the types aX , aY , bX , and bY) were withheld from familiarization to measure generalization to unheard strings. The test strings also contained 4 grammatical-heard strings (one of each of the four types), and 8 nongrammatical strings (the matched grammatical heard and unheard strings from the other grammar). The test materials therefore consisted of 16 unique strings. Appendix C depicts the training and test strings from each of the versions and grammars of the Uncued language.

Importantly, all extraneous properties of the $aX bY$ language were retained in the Uncued language. The Uncued and $aX bY$ languages contained the same vocabulary items, presented with the same frequencies, thus the phonological characteristics of the tokens and strings were equivalent. Moreover, both the Uncued and $aX bY$ languages consisted of two-element strings in which a - and b -elements always occurred in first position, and X s and Y s always occurred in second position. Thus, the only difference between the $aX bY$ language and the Uncued language was that in the $aX bY$ language, there were unique relationships between a -elements and X s and b -elements and Y s, while in the Uncued language a -elements predicted both X s and Y s, as did b -elements.

It is important to note that both the $aX bY$ language and the Uncued language contained co-occurrence restrictions on how elements could be combined into strings, such that as and bs predicted *nonoverlapping sets* of elements. However, in the $aX bY$ language, the cooccurrence restrictions were reliably cued by the distinctive endings on X s and Y s whereas they were not in the Uncued language, (i.e., as and bs predicted words ending in both “ee” and “oo”). Sensitivity to the fact that a - and b -elements have nonoverlapping distributions typically relies on the presence of cues signaling the category membership of words (Braine, 1987; Frigo and MacDonald, 1998; Gerken, Wilson, & Lewis, 2005). Because Control learners are not exposed to the higher-order generalization (that a -elements predict a set of words with a shared feature, and b -elements predict words with a different feature), it is unlikely that they will become sensitive to the structural co-occurrence restrictions of their language. If sensitivity to such structural characteristics, rather than experience with the vocabulary, etc., is responsible for subsequent gains in learning, experience with the Uncued language should not facilitate learning the $acX bcY$ language.

4.1.3. Procedure

There were eight conditions, resulting from the between-participant manipulations of familiarization condition (Control versus Transfer), version (A versus B), and grammar (1 versus 2).

Participants came to the lab at the same time on two consecutive days. On the first day, participants in both groups were exposed to 18 blocks of their respective training language followed by a test. The Transfer group was exposed Version A or B of the $aX bY$ language, whereas the Control was exposed to the Version A or B of the Uncued language. On the second day, participants in both groups were exposed to an $acX bcY$ language, with language version counterbalanced across days. The familiarization and test procedure was equivalent to that of the Transfer group in Experiment 2 in all other respects.

4.2. Results and discussion

We first tested whether the Transfer and Control groups differed in their ability to learn the $acX bcY$ language on Day 2 using a three-way ANOVA with the between-participant factor of familiarization condition (Control versus Transfer) and within-participant factors of test block (1 versus 2) and discrimination type (GH versus GUH). Again, both d' and difference scores were used as dependent measures. Table 5 displays the mean endorsement rates to the three

Table 5
Experiment 3 Mean Endorsement Rates with Standard Error of the Mean in Parentheses

	Block 1				Block 2			
	GH	GUH	NGH	NGUH	GH	GUH	NGH	NGUH
Day 1								
Control	.78 (.044)	.68 (.038)	.68 (.048)	.64 (.050)	.67 (.040)	.61 (.043)	.65 (.039)	.57 (.045)
Transfer	.86 (.031)	.76 (.044)	.64 (.057)	.51 (.064)	.78 (.035)	.62 (.046)	.58 (.052)	.53 (.049)
Day 2								
Control	.64 (.030)	.56 (.042)	.63 (.038)	.60 (.036)	.63 (.042)	.62 (.036)	.62 (.037)	.55 (.043)
Transfer	.73 (.031)	.68 (.043)	.57 (.048)	.53 (.046)	.62 (.044)	.60 (.035)	.53 (.045)	.53 (.046)

Note: GH refers to grammatical-heard test strings, and NGH to their matched nongrammatical counterparts. Similarly, GUH refers to grammatical-unheard test strings, and NGUH to their nongrammatical counterparts.

kinds of test strings for the Control and Transfer groups on days 1 and 2. Table 6 contains measures of their discrimination in terms of difference scores and d' .

The ANOVA using difference scores as a dependent measure yielded a marginal effect of familiarization condition, $F(1, 58) = 2.89, p = 0.095$, with the Transfer group showing better overall discrimination ($M = 0.12, SE = 0.044$) than the Control group ($M = 0.01, SE = 0.04$). There was also a marginal interaction between familiarization condition and test block, $F(1, 58) = 3.42, p = 0.069$. We investigated this interaction by separately comparing the groups' test performance during the first and second block of testing. The Transfer group's discrimination during the first test block ($M = 0.16, SE = 0.06$) was significantly better than the Control's ($M = -0.02, SE = 0.037$), $t(58) = 2.48, p = 0.016$. The two groups' discrimination did not differ during the second test block, $t(58) = 0.53, p = 0.597, M = 0.08 (SE = 0.065)$ in the Transfer group and $M = 0.04 (SE = 0.036)$ in the control. This pattern of findings held when d' was used as a dependent measure, the only exception being that the

Table 6
Experiment 3 Mean Discrimination for GH and GUH strings in terms of Difference Scores and d' , with Standard Error of the Mean in Parentheses

	Block 1				Block 2			
	GH		GUH		GH		GUH	
	GH-NGH	d'	GUH-NGUH	d'	GH-NGH	d'	GUH-NGUH	d'
Day 1								
Control	.11 (.060)	.69 (.396)	.04 (.064)	.03 (.383)	.02 (.054)	-.03 (.361)	.04 (.070)	.27 (.373)
Transfer	.22 (.062)	1.45 (.481)	.25 (.081)	1.86 (.545)	.20 (.065)	1.25 (.459)	.08 (.076)	.46 (.490)
Day 2								
Control	.01 (.047)	-.13 (.306)	-.04 (.055)	-.21 (.344)	.01 (.051)	.18 (.360)	.07 (.045)	.51 (.307)
Transfer	.16 (.064)	.91 (.306)	.15 (.066)	.89 (.344)	.09 (.075)	.73 (.306)	.07 (.066)	.41 (.307)

Note: GH refers to grammatical-heard test strings, and NGH to their matched nongrammatical counterparts. Similarly, GUH refers to grammatical-unheard test strings, and NGUH to their nongrammatical counterparts.

interaction between familiarization condition and test-block was significant, $F(1, 58) = 6.75$, $p = 0.012$, rather than marginal. Thus, the Transfer group showed better discrimination than the Control group during the first block of testing, but this advantage diminished in the second test block.

We next tested whether either group discriminated grammatical from nongrammatical strings at above-chance rates. One-sample t tests using difference scores as a dependent measure indicated that Transfer participants discriminated GH and GUH strings from ungrammatical ones in the first test block, $t_s(29) \geq 2.27$, $p \leq 0.031$, but not the second, $t_s \leq 1.24$, $p_s \geq 0.22$. The Control group showed no discrimination between grammatical and nongrammatical strings, $t_s \leq 1.17$, $p_s \geq 0.254$. T tests using d' confirmed these findings, the only difference being that the Control group showed significant discrimination for GUH test strings in test-block 2, $t(29) = 2.16$, $p = 0.039$, perhaps reflecting learning during the test.

What can we conclude about the Transfer groups' advantage over the Control group? Our claim is that it is due to differences in the two groups' prior experiences. The Transfer group was first exposed to the $aX bY$ language, while the Control group was exposed to the Unued language. Recall that there are two intersecting layers of structure in the $aX bY$ language. First, a - and b -elements predict nonoverlapping sets of elements, and second, the set of words predicted by a -elements shared a distinctive ending, while b -elements predict a set of elements with a different distinctive ending. Learners in the Transfer group, therefore, were exposed to a higher-order generalization in which the co-occurrence restrictions were cued by the highly frequent endings on X - and Y -elements. Any sensitivity to such structure should benefit their learning of the $acX bcY$ language. In contrast, the Control participants had prior exposure to the Uncued language, which contained co-occurrence restrictions (i.e., a - and b -elements predicted nonoverlapping sets), but the restrictions were not cued by the endings on X - and Y -elements. Given that such cues are typically critical for learning (Braine, 1987; Frigo & McDonald, 1998; Gerken, Wilson, & Lewis, 2005), we predicted that Control learners were unlikely to become sensitive to these restrictions, and as a result would not benefit from their prior experience upon exposure to the $acX bcY$ language.

We can test these predictions by assessing the two groups' performance on day 1. Measures of participants' day 1 discriminations can be found in Table 6. Using difference scores as a dependent measure, Control participants' discrimination for GH strings in test block 1 was marginally significant, $t(29) = 1.82$, $p = 0.079$, and no other discriminations differed from chance, $t_s(29) \leq 0.66$, $p_s \geq 0.517$. These findings were confirmed by analyses using d' as a dependent measure. We can conclude on this basis that they were not sensitive to the co-occurrence restrictions of their language, and even their ability to recognize familiar strings was weak.

The Transfer group's day 1 performance was much more suggestive of learning. Paired-sample t tests using difference scores as a dependent measure indicated that their Block 1 and 2 GH discriminations differed from chance, $t_s \geq 3.08$, $p_s \leq 0.005$, and they were also able to discriminate GUH strings from NG ones in test-block 1, $t(29) = 3.08$, $p = 0.005$. However, they did not show discrimination for GUH strings in test-block 2, $t(29) = 1.10$, $p = 0.28$. Again, the same pattern of findings was held using d' as the measure of discrimination. This pattern of results suggests that the Transfer group showed robust discrimination of familiar

strings, and may have developed some sensitivity to the cooccurrence restrictions of the $aX bY$ language.

To summarize, the findings from Experiment 3 provide additional evidence that prior experience with an $aX bY$ structure enables learning of the more difficult $acX bcY$ structure. We found an interaction between familiarization condition and test-block, which was marginal in endorsement-rate analyses and significant in the d' analyses, indicating that such prior experience resulted in enhanced discrimination between grammatical and nongrammatical strings in the first test block but not the second. These findings also rule out the possibility that exposure to the vocabulary, prosodic characteristics, and positional regularities of a -, b -, X -, and Y -elements, rather than the $aX bY$ structure itself, account for the Transfer group's ability to learn the $acX bcY$ language during Phase 2. They also suggest that prior experience need not immediately precede exposure to a parallel structure to assist learning, as the Transfer group benefited from an experience that took place 24 hours before their exposure to the $acX bcY$ language.

There are also several interesting relations between the performance of the Transfer group and that of learners from Experiments 1 and 2. First, the Transfer group from Experiment 3 benefited from prior experience with $aX bY$ structure, but that learning was only expressed during the first block of testing. This finding also held for the difference score analyses of Experiment 2 Transfer learners, and for 6 Block learners discrimination for GH strings in Experiment 1. However, the 18 Block and Transfer learners in Experiment 1 did not show this effect; their discrimination was strong in the first test block and was either maintained or improved in the second test block. Why do we see diminishing sensitivity over testing for some learners, and increasing sensitivity for others? One explanation is that learners who have acquired a relatively strong sensitivity to the language structure (i.e., the 18 Block and Transfer groups from Experiment 1) improve over the course of testing, because they have a good foundation, and more free resources to build upon it, during test. Conversely, the 6 Block learners, having a weaker sensitivity, failed to improve over testing. While the Experiment 2 and 3 Transfer learners, unlike the 6 Block learners, showed sensitivity to the co-occurrence restrictions of the language in the first test block, the $acX bcY$ language was quite complex, containing 48 unique adjacent transitions and 24 unique nonadjacent transitions, as compared to the 24 unique adjacent transitions of the $aX bY$ language. Thus, the $acX bcY$ language potentially placed greater demands on memory. As a result, sensitivity to the language structure may have been more susceptible to interference from nongrammatical strings heard during testing. It may be that familiarization conditions could be manipulated to produce learning more resistant to the influence of nongrammatical instances, but importantly, in real-life language learning, grammatical instances are far more common than nongrammatical ones.

Our second point concerns the Transfer learners in Experiment 2, who showed significant GUH discrimination both in the difference score and d' analyses, while they showed marginal GH discrimination in difference score analyses and none at all on d' analyses. Given that GH strings have the extra benefit of having been heard, GH discrimination is typically better than GUH discrimination. Therefore, it is puzzling that GUH learning tended to be better in Experiment 2. However, the fact that GH discrimination was significant in Experiment 3 suggests that the lack of GH discrimination in Experiment 2 may have been a spurious finding.

Finally, in these experiments we report data from two groups of participants exposed to 18 blocks of an $aX bY$ language: the 18 Block learners from Experiment 1, and the Transfer participants from Experiment 3, who were exposed to the $aX bY$ language on day 1. Both groups of learners showed robust discrimination for GH strings. The 18 Block group from Experiment 1 showed discrimination for GUH strings only in the second block of testing, while the Transfer learners' discrimination for GUH strings was significant only in test-block 1. While these data suggest that neither group showed robust discrimination for GUH strings, they do indicate that 18 blocks of exposure to the $aX bY$ language results in some sensitivity (perhaps weak or fragile) to the cooccurrence relationships. In either case, whatever sensitivity they did acquire was sufficient to both speed learning of a novel $aX bY$ language or enable learning of the $acX bcY$ language.

5. General discussion

In this set of experiments, we demonstrated that what learners acquire from their input changes as they gain experience with a particular type of structure. In Experiment 1, we found that naïve learners start to become sensitive to an $aX bY$ pattern with 18 blocks of exposure, but not with 6 blocks. However, Transfer learners exposed to 18 blocks of an $aX bY$ pattern acquired another $aX bY$ pattern within 6 blocks. They learned the second version of the language at least as well as 18 Block participants, even though they had a third of the exposure to the particular strings in that version. They also showed robust generalization to grammatical strings, while participants given extensive exposure to only one version of the language (the 18 Block group) did not. In Experiment 2, learners with prior exposure to an $aX bY$ language went on to learn an $acX bcY$ category induction language (in which the cooccurrence restrictions pertained to nonadjacent word categories) better than naïve participants given equivalent exposure to the $acX bcY$ language. Because these findings were exploratory, and analyses with discrimination scores and d' differed slightly, we searched for stronger evidence in Experiment 3. Here we found that experience with adjacent relationships facilitates learning nonadjacent ones, and that such sensitivity declines over the course of testing. Although the interaction between familiarization condition and test block was marginal in the endorsement-rate analyses, it was significant in the d' analyses. Additionally, we demonstrated that such experience could occur on the preceding day and still confer a benefit, and eliminated the possibility that the benefit obtained from prior experience was due to nonstructural dimensions of prior experience (i.e., vocabulary, positional information) rather than experience with the $aX bY$ structure itself. The gain in sensitivity afforded by prior experience is noteworthy given that the transfer language required learning of nonadjacent co-occurrence restrictions, as opposed to adjacent ones.

It is important to note that while other studies investigating transfer have focused on determining what aspects of artificial grammar participants are sensitive to when it is reinstated in a novel vocabulary (e.g., Gómez, Gerken, & Schvaneveldt, 2000; Tunney & Altmann, 2001), we assessed how subsequent learning is facilitated by prior experience with similar underlying structure. We found that prior experience facilitates learning in two important ways; first, it accelerates learning of analogous structure, and second, it enables learning of more difficult structure. Designs used in previous transfer studies only gave an indication of the

ability to transfer to identical underlying structure. Our design allowed us to assess changes in the learning process arising from prior experience, and thus provided previously unavailable information.

Our data also suggests the intriguing possibility that learners need not attain full knowledge of abstract structure during initial learning to show enhanced sensitivity to that structure in subsequent learning. Because 18 blocks of exposure is not enough to induce robust generalization to GUH strings (18 Block learners in Experiment 1 did not show GUH generalization, and Transfer learners in Experiment 3 only showed GUH discrimination of the $aX bY$ language in test-block 1 on day 1), Transfer participants in both Experiments 1 and 3 began their second learning phase without robust knowledge of this aspect of structure. However, they became sensitive to it upon exposure to the second version of the language showing evidence of generalization to grammatical-unheard strings. This suggests that either sensitivity to surface characteristics or partial learning of abstract structure can, under some circumstances, enhance subsequent learning. By this view, exposure to two instantiations of the $aX bY$ language may actually promote abstraction, perhaps more so than extensive exposure to a single version of the language. In other words, and in contrast to traditional notions of experience-driven learning (Gold, 1967), exposure to multiple instantiations of a pattern may facilitate abstraction rather than presenting learners with an intractable induction problem.

Our data has shed light on how some potentially difficult natural language structures might be acquired. Recall that one example of parallel structure in language, the cooccurrence relationship in English between determiners and nouns, and between auxiliaries and verbs, contains structure similar to that of the category induction languages used in this experiment. Specifically, dependencies between syntactic categories can be adjacent (e.g., in a “determiner-noun” string) or nonadjacent (e.g., in a “determiner-adjective-noun” string). We found that while such nonadjacent relationships are difficult for naïve learners, prior experience with similar adjacent dependencies enables such learning. Our findings are in line with some connectionist models demonstrating benefits of prior experience in learning nonadjacent syntactic dependencies. For example, Elman (1993) trained a simple recurrent network (SRN) on an artificial language with several characteristics of natural language, such that subjects and verbs agreed in number, verbs differed in their argument structure, and sentences could contain multiple relative-clause embeddings. The network succeeded in learning the language structure when it was trained first with simple sentences, and was gradually exposed to complex strings containing more embeddings, but learning was poor without such staged input.

In related work, Conway, Ellefson, and Christiansen (2003) investigated the conditions under which staged input facilitates adults’ learning of recursive structure. They noted that many of the studies demonstrating an advantage of staged input used visually presented stimuli. They gave adult learners visual or auditory exposure to a language incorporating recursion. Learners given visual exposure were presented with all elements comprising a single string simultaneously, while aural presentation of string-elements was sequential. In the visual exposure condition, learners given staged input learned better than those exposed to the entirety of the language at once. Interestingly, when learners were aurally exposed, they did not benefit from staged input.

The current findings, as well as those of Elman (1993) and Conway et al. (2003), demonstrate that learning complex structure can be facilitated by prior experience with simpler

structure under some circumstances. Thus, our work adds to the previous findings on the effects of staging input. However, the languages in Elman and Conway et al. used the same vocabulary in the simple and complex structures. Thus, they found that prior experience with particular exemplars facilitates learning when those same exemplars are involved in more complex constructions. It may be that learners must have exposure to particular predictive relationships between elements when they are adjacent to become sensitive to nonadjacent relationships between them. Our findings differ in that prior experience facilitated learning when the transfer language did not involve the same vocabulary, instead sharing abstract structural characteristics.

Not all studies have replicated the finding that staged input benefits learning. Rhode and Plaut (1999) failed to replicate Elman's findings using similar network architecture and training language. Moreover, they point out that in natural languages, semantic factors often constrain the kinds of material that can be embedded with different phrase types, resulting in distributional biases. Thus, the embedded material may participate in local predictive relationships or dependencies that bridge the more distant dependencies. However, our data suggest that prior experience with adjacent dependencies can also play an important role in such learning.

A second parallel to natural-language-learning conditions concerns the presence of cues indicating category membership. Because the presence of such cues greatly facilitates learning of $aX bY$ structure (Brain, 1987; Frigo & McDonald, 1998; Gerken, Wilson, & Lewis, 2005), we examined whether nouns and verbs in infant-directed speech have such cues, and whether the incidence of cues differs by category (Lany & Gómez, 2004). We selected four mother-infant dyads from the Bernstein-Ratner corpus (Bernstein-Ratner, 1984) in the CHILDES Database (MacWhinney, 2001) and coded mothers' speech for the presence of these cues. The infants, all girls, ranged in age from 1;5 to 1;8, and had not yet begun combining words in their own productions. We chose to code for morpho-syntactic cues, so for nouns we calculated how often they were preceded by a determiner, contained a plural or diminutive ending ("s" and "ey"), or both. For verbs, we calculated how often they were preceded by an auxiliary or modal (e.g., "is" and "has"), contained an inflectional ending (e.g., "s," "ing," and "en"), or both.

Averaged across the dyads, we found that nouns are more strongly cued than verbs. Table 7 contains these marking percentages broken down by individual dyad. Nouns have some morpho-syntactic cue (either a determiner, an ending, or both) 82% of the time, and two markers (a determiner and an ending) 20% of the time. In contrast, verbs have some morpho-syntactic cue only 21% of the time, and double marking just 1% of the time. Thus, in infant-directed speech nouns are marked more frequently with morpho-syntactic cues than verbs. Learning the verb category and its cooccurrence restrictions should thus be difficult for English-learning infants. However, infants who have already learned the strongly-cued noun category might become sensitive to the verb category by generalization to parallel structure. Our data suggest that prior experience facilitates learning a more difficult pattern, and thus we might expect that learners with prior exposure to a well-cued pattern would have an advantage over naïve learners in acquiring a version of that pattern in which the cues highlighting relevant structure are degraded.

Finally, it is important to consider the possibility that our findings may not apply to infants. While many studies of artificial-language learning have not found differences in adults' and infants' sensitivity to statistical information (Saffran, Aslin, & Newport, 1996; Saffran, Newport, & Aslin, 1996; Gómez, 2002), a recent study (Hudson Kam & Newport, 2005)

Table 7
Percentage of Nouns and Verbs Marked with Morpho-Syntactic Cues in Infant Directed Speech

	Nouns				Verbs			
	Alone	Determiner	Ending	Double	Alone	Auxiliary	Inflection	Double
Anne	.23	.45	.21	.12	.78	.17	.05	0
Amelia	.09	.62	.10	.18	.78	.13	.08	.01
Cindy	.28	.37	.11	.24	.72	.16	.08	.04
Dale	.13	.37	.20	.26	.87	.11	.02	0
Average	.18	.45	.15	.20	.79	.14	.06	.01

found differences in adults' and 5- to 7-year-olds' tendency to regularize inconsistent input. However, studies in progress (using similar artificial languages to those used in Experiments 2 and 3) show that infants as young as 12-months of age are able to transfer to parallel structure, providing support for the possibility that such learning is indeed instrumental in early language acquisition (Lany & Gómez, 2005).

In conclusion, we found that learners' prior experience could facilitate acquisition of new language structure in two important ways. It results in faster learning of parallel structure, and can enable bootstrapping to more complex structure. These findings suggest that the acquisition of natural language structures that are too difficult or complex for naïve learners to acquire may be enabled by prior language experience. Moreover, they suggest that accounts of acquisition should take into consideration that learners change with experience, and thus the input required by naïve learners differs from that of more experienced learners.

Notes

1. We thank an anonymous reviewer for suggesting that we complement our endorsement-rate analyses with d' analyses of discrimination.
2. These, and all subsequent t tests, were two-tailed.
3. We chose to explore marginal interactions of direct relevance to our research hypotheses. This puts us at risk of increased Type I error, but we also did not want to risk missing potentially interesting effects.
4. As with the 18 Block group's increase in discrimination for GUH test strings in Experiment 1, we do not interpret significant discrimination in test-block 2 as an index of learning if test-block 1 discrimination was not significant.

Acknowledgments

This work was supported by NIH R01 HD42170 to LG and RG.

References

- Altmann, G. T. M., Dienes, Z., & Goode, A. (1995). Modality independence of implicitly learned grammatical knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 899–912.

- Bernstein-Ratner, N. (1984). Patterns of vowel modification in motherese. *Journal of Child Language*, *11*, 557–578.
- Braine, M. D. S. (1987). What is learned in acquiring word classes—A step toward an acquisition theory. In B. McWhinney (Ed.), *Mechanisms of Language Acquisition* (pp. 65–87). Hillsdale, NJ: Erlbaum.
- Conway, C. C., Ellefson, M. R., & Christiansen, M. H. (2003). When less is less and when less is more: Starting small with staged input. In R. Alterman & D. Kirsch (Eds.), *Proceedings of the Twenty-Fifth Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum.
- Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition*, *48*, 71–99.
- Frigo, L., & MacDonald, J. (1998). Properties of phonological markers that affect the acquisition of gender-like subclasses. *Journal of Memory and Language*, *39*, 448–457.
- Gerken, L. A. (2004). Nine-month-olds extract structural principles required for natural language. *Cognition*, *93*, B89–B96.
- Gerken, L. A., Wilson, R., Gómez, R. L., & Nurmsoo, E. (in prep). Linguistic category induction without reference: The importance of correlated cues.
- Gerken, L. A., Wilson, R., & Lewis, W. (2005). Seventeen-month-olds can use distributional cues to form syntactic categories. *Journal of Child Language*, *32*, 249–268.
- Gold, E. M. (1967). Language identification in the limit. *Information and Control*, *10*, 447–474.
- Goldowsky, B. N., & Newport, E. J. (1993). Modeling the effects of processing limitations on the acquisition of morphology: The less is more hypothesis. In J. Mead (Ed.), *The Proceedings of the 11th West Coast Conference on Formal Linguistics*. Stanford, CA: CSLI.
- Gómez, R. L. (1997). Transfer and complexity in artificial grammar learning. *Cognitive Psychology*, *33*, 154–207.
- Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*, 431–436.
- Gómez, R. L., & Gerken, L. A. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Science*, *4*, 178–186.
- Gómez, R. L., Gerken, L. A., & Schvaneveldt, R. (2000). The basis of transfer in artificial grammar learning. *Memory and Cognition*, *28*, 253–263.
- Gómez, R. L., & LaKusta, L. (2004). A first step in form-based category abstraction in 12-month-old infants. *Developmental Science*, *7*, 567–580.
- Hudson Kam, C. L., & Newport, E. L. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development*, *1*, 151–195.
- Lany, J., & Gómez, R. L. (2004). The role of prior experience in biasing generalization in artificial language learning. In K. Forbus, D. Gentner, & T. Reiger (Eds.), *Proceedings of the 26th Annual Meeting of the Cognitive Science Society*.
- Lany, J., & Gómez, R. L. (2005). 12-month-olds benefit from prior experience in statistical learning. 30th Boston University Conference on Language Development, Boston, November.
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk. Third Edition*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Newport, E. L. (1990). Maturation constraints on language learning. *Cognitive Science*, *14*, 11–28.
- Newport, E. L., & Aslin, R. N. (2004). Learning at a distance. I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*, 127–162.
- Newport, E. L., Gleitman, H., & Gleitman, L. R. (1977). Mother, I'd rather do it myself: Some effects and non-effects of maternal speech style. In C. Snow, and C. Ferguson (Eds.), *Talking to Children: Language Input and Acquisition* (pp. 109–149). Cambridge: Cambridge University Press.
- Pine, J. M. (1994). The language of primary caregivers. In C. Gallaway & C. J. Richards (Eds.), *Input and Interaction in Language Acquisition* (pp. 38–55). Cambridge: Cambridge University Press.
- Reber, A. S. (1969). Transfer of synthetic structure in synthetic languages. *Journal of Experimental Psychology: General*, *120*, 112–116.
- Reber, R., & Perruchet, P. (2003). The use of control groups in artificial grammar learning. *The Quarterly Journal of Experimental Psychology*, *56A*, 97–115.
- Rohde, D. L. T., & Plout, D. C. (1999). Language acquisition in the absence of explicit negative evidence: How important is starting small? *Cognition*, *72*, 67–109.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926–1928.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, 35, 606–621.

Saffran, J. R., & Thiessen, E. D. (2003). Pattern induction by infant language learners. *Developmental Psychology*, 39, 484–494.

Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Use of stress and statistical cues to word boundaries. *Developmental Psychology*, 39, 706–716.

Thiessen, E. D., & Saffran, J. R. (2004). Infants’ acquisition of stress-based word segmentation strategies. In A. Brugos, L. Micciulla, & C. E. Smith (Eds.), *Proceedings of the 28th Annual Boston University Conference on Language Development* (pp. 608–619). Boston, MA: Cascadilla Press.

Tunney, R., & Altmann, G. (2001). Two modes of transfer in artificial grammar learning. *Journal of Experimental Psychology*, 27, 614–639.

Valian, V., & Coulson, S. (1988). Anchor points in language learning: The role of marker frequency. *Journal of Memory and Language*, 27, 71–86.

Wilson, R. (2002). *Syntactic category learning in a second language*. Unpublished Ph.D. Dissertation, University of Arizona, Tucson, AZ.

Appendix A

aX bY language materials

Version A G1: aX bY G2: aY bX				Version B G1: aX bY G2 aY bX			
A	b	X	Y	a	b	X	Y
Alt	erd	juhnoo	nusee	ush	ong	keerit	bivul
Pel	vot	wifoo	lemee	dak	rud	lepit	choopul
		tamoo	sufee			feegit	habbul
		feenoo	vaymee			soolit	jerul
		zinoo	raffee			yohvit	pogul
		deechoo	durpee			zamit	vummul
G1 Grammatical-Heard Test Strings				G1 Grammatical-Heard Test Strings			
alt feenoo	pel zinoo	erd vaymee	vot durpee	ush soolit	dak zamit	ong vummul	rud pogul
G1 Grammatical-Unheard Test Strings				G1 Grammatical-Unheard Test Strings			
alt juhnoo	pel wifoo	erd nusee	vot lemee	ush lepit	dak keerit	ong choopul	rud bivul
G2 Grammatical-Heard Test Strings				G2 Grammatical-Heard Test Strings			
alt vaymee	pel durpee	erd feenoo	vot zinoo	ush vummul	dak pogul	ong soolit	rud zamit
G2 Grammatical-Unheard Test Strings				G2 Grammatical-Unheard Test Strings			
alt nusee	pel lemee	erd juhnoo	vot wifoo	ush choopul	dak bivul	ong lepit	rud keerit

Appendix B*acX bcY language materials*

Version A G1: acX bcY G2: acY bcX				
a	b	C	X	Y
alt	erd	hes	juhnoo	nusee
pel	vot	kaf	wifoo	lemee
		sij	tamoo	sufee
			feenoo	vaymee
			zinoo	raffee
			deechoo	durpee
G1 Grammatical-Heard Test Strings				
alt hes feenoo	pel hes zinoo		erd hes vamee	vot hes rafee
alt kaf wifoo	pel kaf deechoo		erd kaf rafee	vot kaf durpee
alt sij tamoo	pel sij juhnoo		erd sij durpee	vot sij nusee
G1 Grammatical Unheard Test Strings				
alt hes juhnoo	pel hes wifoo		erd hes nusee	vot hes lemee
alt kaf wifoo	pel kaf tamoo		erd kaf lemee	vot kaf sufee
alt sij tamoo	pel sij feenoo		erd sij sufee	vot sij vamee
G2 Grammatical-Heard Test Strings				
alt hes vamee	pel hes rafee		erd hes feenoo	vot hes zinoo
alt kaf rafee	pel kaf durpee		erd kaf zinoo	vot kaf deechoo
alt sij durpee	pel sij lemee		erd sij zinoo	vot sij juhnoo
G2 Grammatical-Unheard Test Strings				
alt hes nusee	pel hes lemee		erd hes juhnoo	vot hes wifoo
alt kaf lemee	pel kaf sufee		erd kaf wifoo	vot kaf tamoo
alt sij sufee	pel sij vaymee		erd sij tamoo	vot sij feenoo
Version B G1: acX bcY G2: acY bcX				
a	b	C	X	Y
ush	ong	tash	keerit	bivul
dak	rud	fis	lepit	choopul
		nep	feegit	habbul
			soolit	jerul
			yohvit	pogul
			zamit	vummul
G1 Grammatical-Heard Test Strings				
ush tash soolit	dak tash yohvit		ong tash jerul	rud tash pogul
ush fis yohvit	dak fis zamit		ong fis pogul	rud fis vummul
ush nep zamit	dak nep keerit		ong nep vummul	rud nep bivul
G1 Grammatical Unheard Test Strings				
ush tash keerit	dak tash lepit		ong tash bivul	rud tash choopul
ush fis lepit	dak fis feegit		ong fis choopul	rud fis habbul
ush nep feegit	dak nep soolit		ong nep habbul	rud nep jerul
G2 Grammatical-Heard Test Strings				
ush tash jerul	dak tash pogul		ong tash soolit	rud tash yohvit
ush fis pogul	dak fis vummul		ong fis yohvit	rud fis zamit
ush nep vummul	dak nep bivul		ong nep zamit	rud nep kirit
G2 Grammatical-Unheard Test Strings				
ush tash bivul	dak tash choopul		ong tash keerit	rud tash lepit
ush fis choopul	dak fis habbul		ong fis lepit	rud fis feegit
ush nep habbul	dak nep jerul		ong nep feegit	rud nep soolit

Appendix C

Uncued language materials

Version A G1: aX ₁₋₃ bY ₄₋₆ G2: aY ₁₋₃ bX ₄₋₆				Version B G1: aX ₁₋₃ bY ₄₋₆ G2: aY ₁₋₃ bX ₄₋₆			
a	b	X ₁₋₃ , Y ₄₋₆	X ₄₋₆ , Y ₁₋₃	a	b	X ₁₋₃ , Y ₄₋₆	X ₄₋₆ , Y ₁₋₃
alt	erd	juhnoo	nusee	ush	ong	keerit	feegit
pel	vot	wifoo	lemee	dak	rud	lepít	soolít
		feenoo	durpee			zamít	yohvít
		sufee	tamoo			bívuł	habbuł
		vaymee	zinoo			choopul	jerul
		raffee	deechoo			pogul	vummul
G1 Grammatical-Heard Test Strings				G1 Grammatical-Heard Test Strings			
alt feenoo	pel zinoo	erd vamey	vot durpee	ush pogul	dak lepít	ong vummul	rud soolít
G1 Grammatical-Unheard Test Strings				G1 Grammatical-Unheard Test Strings			
alt juhnoo	pel wifoo	erd nusee	vot lemee	ush kirit	dak bívuł	ong feegit	rud habbuł
G2 Grammatical-Heard Test Strings				G2 Grammatical-Heard Test Strings			
alt vamey	pel durpee	erd feenoo	vot zinoo	ush vummul	dak soolít	ong pogul	rud lepít
G2 Grammatical-Unheard Test Strings				G2 Grammatical-Unheard Test Strings			
alt nusee	pel lemee	erd juhnoo	vot wifoo	ush feegit	dak habbuł	ong kirit	rud bívuł