Published in final edited form as:

Lang Learn Dev. 2009 April 1; 5(2): 115–145. doi:10.1080/15475440902739962.

More than words: Adults learn probabilities over categories and relationships between them

Carla L. Hudson Kam

Department of Psychology, 3210 Tolman Hall, #1650, University of California, Berkeley, Berkeley, CA 94720

Abstract

This study examines whether human learners can acquire statistics over abstract categories and their relationships to each other. Adult learners were exposed to miniature artificial languages containing variation in the ordering of the Subject, Object, and Verb constituents. Different orders (e.g. SOV, VSO) occurred in the input with different frequencies, but the occurrence of one order versus another was not predictable. Importantly, the language was constructed such that participants could only match the overall input probabilities if they were tracking statistics over abstract categories, not over individual words. At test, participants reproduced the probabilities present in the input with a high degree of accuracy. Closer examination revealed that learner's were matching the probabilities associated with individual verbs rather than the category as a whole. However, individual nouns had no impact on word orders produced. Thus, participants learned the probabilities of a particular ordering of the abstract grammatical categories Subject and Object associated with each verb. Results suggest that statistical learning mechanisms are capable of tracking relationships between abstract linguistic categories in addition to individual items.

Key terms

language acquisition; statistical learning; second language acquisition; artificial languages; probabilistic syntax

Recently, there has been much discussion about the involvement of statistical learning in various aspects of language acquisition (e.g. Marcus, 2000; Marcus & Berent, 2003; Seidenberg, MacDonald, & Saffran, 2002, 2003). Statistical learning, a powerful and relatively modality independent form of learning (Newport & Aslin, 2000), operates over particular forms. On this basis some have suggested that statistical learning is much better suited to tasks such as word segmentation than to acquiring aspects of language that involve categories (Marcus, 2000; Peña, Bonatti, Nespor, & Mehler, 2002). Learning involving categories, it is argued, may be better suited to an algebraic learning mechanism that extracts rules over variables (Marcus, 2000; Peña, et al., 2002). This strict functional distinction between statistical and algebraic learning is not accepted by all, however. Saffran (2001) and Thompson and Newport (2007), for example, have shown that learners can acquire aspects of syntax on the basis of distributional information. Similarly, work by Mintz and his colleagues (among others) has shown that form classes can be extracted on the basis of very specific word co-occurrence information (Mintz, 2002, 2003; Mintz, Newport, & Bever, 2002). While these works do not deny the existence of abstract categories, nor of rules including such categories, their results blur the distinction between statistical and algebraic domains. In so doing they highlight the

need for a greater understanding of the nature of and limits on different types of learning, particularly statistical learning (see Seidenberg, et al., 2002).

The present paper is generally concerned with this issue, exploring the boundaries on statistical learning with respect to syntax. To date, most of the work on this topic has asked whether people can learn phrase structure via statistics (e.g. Saffran, 2001), and further, whether rule-like representations are even really necessary (e.g. Christiansen & Curtin, 1999).¹ As Seidenberg, et al. (2002) point out, these are very difficult questions to answer definitively. We approach the issue from a somewhat different perspective, and one which is (hopefully) more tractable; instead of asking if abstract categories, phrasal units, and rules can (simultaneously) be learned via statistics, we ask if statistics can be learned over higher order categories, and not simply individual items, i.e. when the ‘items’ are categories, rather than particular forms.

To investigate this question, we exposed participants to miniature languages containing probabilistic variation in word order. In particular, the relative ordering of the abstract grammatical categories Subject (S), Object (O), and Verb (V) was variable in participants’ input. Different orders occurred in the input with different frequencies. For instance, in one condition 60% of the sentences participants heard were in VSO order, and 40% were SOV. Importantly, the occurrence of any particular order in a given sentence was unpredictable; there was no difference in the meaning or context associated with one (e.g. VSO) versus the other (e.g. SOV) order. Put differently, the individual orders carried no specific semantic or pragmatic impact. The question was whether participants would learn the probabilities associated with the different orders. If so, it would provide evidence of participants’ abilities to quickly learn probabilities associated with abstract categories, and their relationship to each other. The results from this study, then, have the potential to inform us about whether statistical learning is restricted to lower-levels of language, or is instead a type of learning that can be applied to variables as well as specific entities.

Background

Initial work in statistical learning focused on acquiring individual units such as words (Saffran, Aslin, & Newport, 1996). The computations involved in such learning are computations over specific items, e.g. the individual syllables that make up a word. Learning syntax, however, is different, in that it involves relationships between types or categories of items, and so might be better suited to other kinds of learning mechanisms. While many studies have demonstrated that people can acquire abstract knowledge of finite-state-type-grammars via distributional learning (see Gómez & Gerken, 2000), such knowledge does not involve relationships between categories, a crucial aspect of syntax. Saffran (2001) was the first to demonstrate learning of phrasal groupings involving categories from transitional probabilities. Even so, although participants’ performance was significantly different from chance, learning was not particularly strong, which could indicate that other cues might also be necessary to learn syntax, as suggested by Morgan, Meier, and Newport (1987, 1989). More recently, however, using a modified version of the same language with stronger statistical cues to phrase boundaries, Thompson and Newport (2007) demonstrated very robust learning of the phrasal units. People were able to learn which word types were constitutive of phrases and which were not purely on the basis of distributional cues.

¹Christiansen and Curtin are arguing against the position taken by Marcus, Vijayan, Rao, and Vishton (1999), not Saffran. As our work is not really relevant to this issue, we will not go over that literature in any detail. Interested readers are encouraged to read the references cited in the text.

Although the findings were discussed in terms of statistical learning, one could argue that the learners first used distributional learning to acquire the types or categories (e.g., A, B, C, D, E and F words in the case of Thompson & Newport, 2007), something for which there is independent evidence (e.g., Gerken, Wilson, & Lewis, 2005; Gómez & Lakusta, 2004; Mintz, 2002), but then formed rules regarding the relationships between the types (e.g., their ordering) using an algebraic learning mechanism. Thus, syntax learning would still be non-statistical. If this were the case, however, learners should be insensitive to probabilities associated with the relationships between categories, as the statistics are irrelevant and unrelated to the rules extracted. (Note that the opposite is not true – a lack of sensitivity to the probabilities is not necessarily evidence against statistical learning of syntax – as learners could use statistical distributions to induce rules.) For example, learners exposed to a language where a sentence consists of an A phrase plus a B phrase with A phrases consisting of an A word plus an optional D word and B phrases a B word followed by a word in category E should not be sensitive to the probabilities of a D word versus a B word following an A word, they should only know that both types are possible, well-formed sentences in the language. By implication then, if learners' knowledge of a language includes knowledge of the probabilities associated with relationships between categories, it would constitute very strong evidence for the involvement of statistical learning in the acquisition of syntax.

In previous work we found that adults are able to veridically learn probabilities associated with inconsistently occurring forms in a language (Hudson Kam & Newport, 2005). In particular, when given input in which a determiner occurred probabilistically, adult learners produced determiners with the same probability, i.e. if they heard determiners 75% of the time, they produced them 75% of the time. On the face of it, these data appear to demonstrate that learners can acquire probabilities associated with categories. However, participants did not have to track probabilities associated with categories in order to probability-match. They could have matched the probabilities by computing the probability of the determiner given each individual word that happened to be a noun and then averaging these probabilities, for instance. Although this explanation entails a category over which the averaging occurred, it does not entail that the category was actually an abstract grammatical category such as Noun. The category could simply have been words followed by the word (determiner) *poe*, for instance.² Thus, the statistics need not have been tracked with respect to a category.

Here, we ask if participants can learn probabilities associated with unpredictably variable word orders, in particular, the relative ordering of Subject, Verb, and Object, where the different orders are not meaningful in any way.³ In our language verbs differed in meaning and behaviour from nouns, and so were relatively easy to identify as such. Subject and Object also differed in both meaning and behaviour. The Subject was always the (more) affected or event-defined salient noun, so the lone noun in an intransitive sentence and the noun with the higher semantic role in transitive sentences (where agents are more affected or salient than patients, and themes are more salient than locations, see, e.g., Jackendoff, 1990). In terms of behaviour, when there were two nouns the Subject always preceded the Object, so while absolute position of Subject and Object was variable in the input, relative position was always indicative of a noun's status as Subject or Object. (These sorts of cues to a noun's grammatical status are typical, Keenan,

²Interestingly, participants could not have matched the probabilities without some sort of category. There were two different noun classes, each of which took a different determiner (*poe* and *kaw*). Participants rarely used the incorrect determiner with a noun, suggesting that they knew which determiner went with which nouns.

³Note that the kind of variation we are using here is quite different from the word order variation found in free word order languages which clearly are learned by speakers across the globe every day. Free word order languages do not have truly random or unpredictable word orders: Word order is conditioned on meanings or pragmatic functions such as topic and focus. Thus, although the location of a noun phrase is not determined by its grammatical role in a free word order language (as it is in a language like English), to say that word order is meaningless or that it is truly free is an oversimplification. Therefore, in contrast to the unpredictable variation we are exposing learners to, speakers learning free word order languages are learning meanings or pragmatic functions associated with the different orders (see Newmeyer, 2006, for a related discussion).

1976.) If participants can learn unconditioned word order variation veridically, it would indicate that they can track and compute the probabilities attached to an abstract grammatical category. They cannot match the probabilities by computing the probability of *girl* occurring in the second position, for instance, since *girl* will sometimes occur in second position as a Subject and other times as an Object. Therefore, unconditional probabilities (ignoring grammatical role) will result in different numbers than conditional probabilities (taking into consideration grammatical role). Thus, at the very least, in order to probability-match the word orders they must be tracking the probability of *girl* occurring in second position *when it is the Subject or Object* ($p[\text{'girl' in second position} | \text{'girl' = subject}]$) and averaging over similar computations performed on different nouns. They could also track the probability of the Subject appearing in some particular location ($p[\text{Subject in first position}]$). Alternatively, they could have rules involving the entire sentence, such as $p(\text{SVO})$ or $p(\text{SVO} | \text{'hit'})$.⁴ Importantly, however, each of these possibilities includes abstract grammatical categories. Thus, if we see probability matching, it would indicate that participants are tracking probabilities with respect to abstract categories, not simply individual words.

Despite our previous findings, it is not at all clear that adults will match the probabilities present in their input, precisely because word order patterns are relationships between abstract categories. Peña, et al. (2002), for instance, suggest that statistical learning may be restricted to learning lower-level item-based properties of languages such as words and that an algebraic learning mechanism that extracts rules expressed over variables is involved in learning grammar. While such rules could, in principle, include information about probabilities, the predominant view, at least in linguistics, is that rules and statistics are quite separate (Newmeyer, 2003, 2006; see also Marcus & Berent, 2003). Therefore, the learners might acquire rather strict rules about word order rather than acquire the probabilities. There is some evidence to suggest that such an outcome might be likely. For example, research has shown that children learning Finnish (Bowerman, 1973a) and Russian (Slobin, 1966) initially use strict word order, despite the fact that they are learning free word order languages.⁵ Another relevant piece of evidence comes from Hawaiian Creole English (HCE). The earliest speakers were all adult late learners who spoke the language with unpredictably variable word orders. Importantly, the variable word orders had no functional significance (unlike in Finnish or Russian), they were errors likely due to interference from the speakers' first languages. However, HCE rapidly developed fixed word order; the variation was not sustained as the language was learned by new speakers (Bickerton, 1981; Bickerton & Givón, 1976). Consistency is not the only possible outcome, however. It could be the case that when faced with such unpredictability in sentence structure, the learner simply learns that there is no pattern and so produces word orders at random. Alternatively, the learners might acquire the non-category based statistics in the input, and learn the probabilities associated with particular words and particular locations. Thus, in the present work, word orders learned by participants exposed to variable word orders are examined with respect to each of these possibilities.

⁴One could argue that the first possibility listed here ($p(\text{SVO})$) could result either from statistical learning or an algebraic learning mechanism that gives rise to proposition-like knowledge (e.g. SVO) which is later associated with some frequency information (although the latter blurs the distinction between the two mechanisms and their involvement in language learning). However, this is inconsistent with the distinction made between the two learning mechanisms by those proposing that grammar learning is accomplished by an algebraic mechanisms (e.g. Peña et al., 2002; Marcus, 2000), namely that statistics are associated with items not categories, and thus, are separate from generalizations. [The second possibility - $p(\text{SVO} | \text{'hit'})$ - is less ambiguous, since it mixes item-level (hit) and category-level (SVO) information, and therefore could not result from statistics being associated with rules.] Note that this does not preclude knowing multiple rules, however, only knowing the particular likelihoods associated with each of the possible rules.

⁵Although the children's productions are described in terms of categories like subject and object, it is not entirely clear what categories underlie the word orders – it could be grammatical categories like subject and object, or semantic categories like agent and patient. As Bowerman (1973b) points out, it can be difficult to tell the difference at this early stage.

THE PRESENT STUDY

To investigate whether learners can acquire probabilities over higher order categories, we exposed participants to miniature languages containing unpredictable inconsistency in word order. (The exact specifications of the input are given below). After several days of exposure to the language, participants were tested to see what they had learned about the language.

In our own previous work on language learning from inconsistent probabilistic input we have found both probability matching and overmatching in adult learners (Hudson Kam & Newport, 2005, in press; see also Wonnacott & Newport, 2005).⁶ One factor that affects participants' behaviour is the number of forms that are in variation with each other. When participants are exposed to one form that is randomly present or absent in a given context, they tend to probability match, but when they are exposed to multiple forms randomly alternating in a context they overmatch, producing the more commonly occurring form more often than they hear it (Hudson Kam & Newport, in press). Interestingly, there is a similar trend in the work on non-linguistic probability learning in humans (Gardner, 1957; Weir, 1972). In this work, participants are exposed to stimuli in which different events have different probabilities of occurrence, and their knowledge of the underlying probabilities is assessed (usually via prediction). In the most basic experiments with two possible events, adults' predictions quickly match the actual probabilities of each event occurring. However, when there are more than two possible events, participants often predicting the more frequent event more often than it actually occurs in their experience. That is, as in our language learning studies, they overmatch. It appears then that learners are broadly sensitive to the number of alternatives, with fewer alternatives leading to probability-matching, and more alternatives leading to overmatching. Thus, we exposed participants to different numbers of word orders in transitive sentences, with some hearing two and others three. (There are only two possible orders in intransitive sentences, SV and VS, and so by including variation we necessarily have to expose all participants to both intransitive orders.)

We choose to use higher order categories already known to our participants, thereby avoiding difficulties associated with the formation of categories distributionally (Braine, et al., 1990; Frigo & McDonald, 1998; see discussion in Gerken, et al., 2005). However, this introduces other potential complications; participants' existing linguistic knowledge might affect learning outcomes. Studies of naturalistic second-language (L2) learning have shown that late-learners often transfer word orders from their first language (L1) to their L2 (see Odlin, 1989, for a review), that is, they produce L2 sentences in L1 word orders. If participants are indeed treating our experiment as an exercise in language learning then we might reasonably expect to see some interference from English word order which could affect the degree to which participants match the actual probabilities present in their input. Whether we see interference and how strong it is might be affected by whether or not participants hear sentences with English-like word order. In particular, we reasoned that more English-like sentences might lead to more interference. This is something not often considered in artificial language learning studies, but in other research we have found L1 interference in a statistical word segmentation study (Finn & Hudson Kam, 2008). Therefore, we varied the proportion of transitive sentences participants heard in English-like word order, ranging from 0–40%. (Again, because we are including variation in intransitive sentences, some have to be in English order.)

One might ask why use adult participants given the possibility of L1 interference and the potential for it to obscure participants' true learning abilities. There are three reasons. First, we are using adults specifically because of our earlier findings that they are more likely than

⁶In other work focusing on language change we refer to overmatching as regularization (Hudson Kam & Newport, 2005). Singleton and Newport (2004) use the term frequency boosting for the same phenomenon.

children to probability match given inconsistent variation in linguistic input (Hudson Kam & Newport, 2005), which accords with findings in the probability learning literature more generally (Estes 1964, 1976; Weir, 1964). Given that we are investigating the possibility that statistics can be learned over abstract categories, it is prudent to use an age group that we know can learn probabilities based on other evidence. Second, although interference could make interpretation more difficult, it would also clearly demonstrate that participants are treating this as language learning. In so doing it would stand against the potential criticism that participants could have been using learning mechanisms not typically deployed by language learners, especially should we find evidence for statistical learning over categories. And finally, as mentioned previously, we wanted to introduce variation over grammatical categories known to our learners. Using experienced language users allows us to accomplish this with greater certainty.

Method

Participants

Thirty native-English speaking adults participated in the study (10 males, 20 females, mean age 21.37 years, range 18–24). Participants were students at the University of Rochester and the University of California, Berkeley at the time of the study. They were recruited from a department subject pool (UR) or via flyers posted around campus (UCB). All were paid for their participation.

Stimuli

The basic language to which participants were exposed consisted of 51 words: 36 nouns, 7 intransitive verbs/predicates, 5 transitive verbs/predicates, 1 negator, and 2 determiners, 1 for each of 2 noun classes. (A complete word list can be found in Appendix A of Hudson Kam & Newport, 2005). The only grammatical consequence of noun class is determiner selection: Each class of nouns takes a different determiner. Note that as in natural languages, the distribution of nouns in the two classes is not equal; there were 21 nouns in one class and 15 in the other. The noun classes were included to keep the stimuli as similar as possible to our previous work in order to allow comparisons across different studies.⁷ The intransitive verbs included meanings associated with the grammatical category of adjective in English (e.g. ‘big’, ‘red’) as well as more typical intransitive verbs (e.g. ‘fall’). The transitive verbs likewise were not all verbs in English; some encoded meanings that are expressed using a preposition in English (e.g. ‘inside’). Having these sorts of meanings expressed as verbs may seem unusual, but it does occur in actual languages (see e.g. Dryer, 2005; Stassen, 1997, 2005). The negator is a general negative word that turns a positive statement into a negative, as in (1).

(1)/sɪg fiɪk ɡɛntɔ po/
neg red plastic-donut DET2
‘The donut isn’t red’

Word order varied according to condition, described in more detail below. However, for all participants VS(O) was the most frequent word order. Determiners always followed the noun, as is typical for V-S-O languages (Greenberg, 1963).

The language was created in conjunction with a small world of objects and actions. Even with the semantic restrictions imposed by the referent world, there are over 13,000 possible sentences. There is, therefore, a wide scope for testing participants using novel sentences. There

⁷Nouns were divided into classes on an arbitrary basis. This is atypical for actual languages, but is not unlike the way adult learners treat noun classes (Cain, Weber-Olsen, & Smith. 1987).

were 230 sentences and corresponding visual scenes in the exposure set, 115 intransitive, 115 transitive. Each exposure session contained a different set of approximately 115 sentences drawn from the 230 sentence exposure set. (For the first two sessions participants actually heard half this number.) Each sentence (and scene) was presented three times over the course of the seven exposure sessions.

Each intransitive verb occurred 15–18 times in the exposure sentences, each transitive verb 14–27 times. Each noun occurred 3–4 times in the intransitive sentences, and 3–4 times in each grammatical role (subject and object) in the transitive sentences, with the result that each noun occurred 8–12 times total in the exposure sentences. (Individual nouns did not occur equally often in all three positions, a noun could appear in three intransitive sentences, four times as subject and three times as object.) Since the exposure set was presented three times, actual exposure to each word is three times its occurrence. Negative sentences were included to help the participants learn the meaning of the verbs, especially the intransitives, as well as to expand the number of possible sentences in the language. Individual verbs were presented either in both negative and positive sentences or in only positive sentences; no verb was presented only in negative sentences. This was also true of nouns – no nouns appeared only in negative sentences. Overall, there were relatively few negative sentences in the exposure – 7 transitives and 43 intransitives.

Presentation

Participants were told that they were in a language learning experiment. For the first two days they would start by learning some words in the new language, and then would see some vignettes on a video and hear the corresponding sentences in the new language. On days three through seven, they would only watch videos. They were instructed to pretend that they were on an isolated island, surrounded by people speaking the new language, and that they had to learn it by observing. They were aware that they would be tested at the end of the experiment, and that this testing would involve producing their own new sentences.

Participants were exposed to the language for seven sessions, each lasting 25–29 minutes. The first two sessions began with vocabulary presentation. On a computer monitor the participant saw a picture of an object (for nouns) or a written English gloss (for verbs), and then heard the corresponding word over headphones. For example, the participant saw a picture of a bird and heard ‘fumpoga poe’, or saw **to fall** and heard ‘gern.’ This was the only point in the experiment where participants saw anything written. There were three blocks, each containing all the words in the language in random order. Participants were encouraged to repeat each word as they heard it, but there was no monitoring or correction given.

The second half of the first two sessions, and the whole of sessions three through seven, consisted of exposure to scenes paired with sentences. Participants were seated in front of a video monitor, on which they watched a scene or event. (These were dynamic videos made with real objects.) They then heard a sentence describing the scene. E.g., the participant sees a boat hit a girl and hears:

(2)/flɪm mæuznɐ kə fɛrlukʌ po/

hit boat DET1 girl DET2

‘The boat hits the girl’

Sentences were spoken at a normal rate with English prosody and phonology and sounded very natural and fluent.

Participants were required to learn the grammar of the language solely from the auditory exposure to the sentences; there was no explicit instruction. They were asked to repeat each

sentence after hearing it, although again there was no monitoring of this. They were told that this was pronunciation practice which would be helpful since they would have to produce their own sentences at the end of the experiment.

The entire experiment took eight sessions (seven exposure sessions and one test session). Participants completed the experiment in 10 – 12 days.

Experimental Manipulation

Participants were exposed to languages containing unpredictable variation in the relative ordering of the Subject, Verb, and (for transitive sentences) Object constituents, with the exact orders and proportions of each varying by condition. For all participants, 60% of the intransitive sentences were VS and 40% were SV. We used a 60/40 split because it is slightly off 50/50, and so contains a slight but definite bias toward one order over the other. We reasoned that if this might provide an algebraic- or rule-learner with a reasonable basis for a rule (i.e. VS, although not the only one, as we discuss under Results).⁸ Importantly, we know from several experiments that adults can learn this particular statistical distribution over other aspects of language (i.e., determiner occurrence, Hudson Kam & Newport, 2005, in press), and so if they can learn probabilities over abstract categories, it should be possible for them to learn word orders with these probabilities. Finally, VS was chosen as the more frequent order because it is different from English.

For the transitive sentences, the most frequent sentence type in all conditions was also verb initial (VSO), and again occurred in 60% of the exposure sentences. The experimental manipulation occurred in the remaining 40% of the transitive sentences. One group of participants heard SOV in these sentences. Another group heard SVO. A third group heard 20% SOV and 20% SVO. This is the only thing that differed by input condition. From this point forward, input conditions are identified by the less frequent word orders in the transitive sentences (i.e. *SOV* if they heard 60% VSO and 40% SOV, *SVO* if they heard 60% VSO and 40% SVO, and *SOV & SVO* if they heard 60% VSO, 20% SOV, and 20% SVO).

Recall that we were manipulating both the proportion of English-like sentences participants heard and the number of word orders participants were exposed to in the transitive sentences. With respect to proportions, in the SOV condition participants heard no English-like sentences (aside from those in the intransitive sentences, which were the same for all conditions). In contrast, 20% of the transitive sentences participants heard in the SOV & SVO condition were English-like, as were 40% of the transitive sentences in the SVO condition. As for the number of orders participants heard, participants in the SOV and SVO conditions heard two transitive orders, those in the SOV & SVO condition heard three.

Which sentences occurred in which orders were selected randomly, but were as similar as possible across the three conditions: All participants heard the same sentences in VSO order and sentences that occurred in SOV in condition one were SVO in condition two, and SOV or SVO in condition three.

We did not include any orderings where the Object preceded the Subject, as these are often viewed as unusual, if not impossible, at least as a basic order (Comrie, 1989). This restriction allowed us to include a consistent cue to grammatical role – the subject noun phrase always preceded the object noun phrase. That is, the argument with the highest semantic role ranking (typically the subject cross-linguistically, Keenan, 1976) appeared before the noun with the lower ranking, so agents preceded patients (the hitter preceded the hittee), and themes precede locatives (the located entity preceded its location). There was no agreement or case marking

⁸I.e. VS. Although here we are assuming the extraction or formation of a single rule, we allow for other possibilities (see Results).

in the language. Ordering is often correlated with grammatical role (Keenan, 1976), and is the cue our participants, being native English speakers, are most familiar with (since case and verb agreement are quite limited in English, and were therefore not included in our language).⁹ Related, since they are adult learners who speak a Subject-Object language, we expected that they would automatically categorize things in terms of Subject-Object (Odlin, 1989) rather than Topic (another possible way of organizing sentences, see, e.g., Li & Thompson, 1976).¹⁰

Importantly, the percentages are true only at the level of the abstract categories; they change if we compute the probabilities without respect to the categories. For example, if we think of a transitive sentence as having three slots - first, second, and third - potentially each occupied by some verb or noun phrase, we can compute the unconditional likelihood of any individual noun or verb occurring in each of the three locations. For instance, the noun 'fumpoga' (bird), occurred in the second position of three constituent (or transitive) sentences 38.88% of the time in the SOV condition, 27.28% of the time in the SVO condition, and 33.33% of the time in the SOV & SVO condition. It (fumpoga) occurred in position three more frequently in two of the conditions (SVO and SOV & SVO, at 50% and 44.44%, respectively), and in the other (SOV), fumpoga was equally likely to occur in second and third position. That is, the unconditional probabilities for fumpoga favour position three in two conditions and positions two and three equally in the other. In contrast, if we consider the probability that fumpoga appears in each location conditional on being subject or object, it is always more likely to occur in position two; the likelihood of VSO order for fumpoga is 55.6% in all three conditions. Thus, if participants are learning the unconditional probabilities, participants in the SVO and SOV & SVO conditions should produce more sentences with fumpoga as subject in position three (i.e. X X fumpoga) than the other positions. In contrast, if they are learning the conditional probabilities dependant on grammatical role, they should preferentially produce fumpoga as subject in position two (i.e. X fumpoga X).

Tests

After exposure participants were given three different tests to evaluate what they had learned. Tests were given in the order in which they appear below.

Vocabulary—This test served to evaluate whether participants had learned enough vocabulary to be tested on more complex aspects of the language. Participants were asked to provide a name for 12 objects as each one appeared on a video monitor. Participants were given as much time as they needed to respond. The experimenter kept a tally of correct responses. If the participant knew at least five of the nouns, they progressed on to the next task.¹¹ As a check, the vocabulary tests for a subset of participants were transcribed and coded by a second experimenter. There were no disagreements about the number of vocabulary items correctly produced by participants.

We only tested nouns because of their importance for the following production task (described next). In addition, since scenes can often be described from multiple perspectives, and therefore

⁹In ongoing work we have found that agreement can be quite difficult to learn. In particular, it takes a great deal longer for adults than order of grammatical constituents, even when the two are perfectly correlated.

¹⁰There is also evidence that in the absence of input consistent with topic prominence, the category of subject will naturally emerge (Coppola & Newport, 2005), suggesting that it is more basic than topic/comment structure - another reason to think that our participants will indeed treat the arguments as subjects and objects. If we had found that participants produced orders randomly, one might have been able to argue that they simply failed to use grammatical categories, and so the lack of other cues to subject-/object-hood could have been seen as a problem. However, this was not the case, making this less of an issue.

¹¹Twelve additional participants did not know at least 5 of the words and so did not complete the study. In previous experiments we had participants take an initial vocabulary test after four exposure sessions which served primarily to indicate to them how much they knew/didn't know. In the current experiment we eliminated this first vocab test, and this seemed to have a detrimental effect on participants' learning. In more recent studies we have reinstituted this mid-exposure test, and completion rates appear to have improved.

using multiple verbs (e.g. under and over), it would have been difficult to elicit the intended verb without providing a written English gloss and asking participants to provide the word. This kind of translation task would almost certainly activate their L1 lexicon to a very large degree, something we wished to avoid.¹²

Sentence production task—The test of primary interest was the sentence production task. Participants saw a novel scene on the video monitor and heard a verb. Their task was to use the verb provided in a sentence describing the scene. In order to respond participants had to recall the noun phrases and arrange the sentence in whatever order they felt appropriate. For example, a participant sees a toy bird jump around and hears the word /mɛrt/ 'move'. She should then say something like /mɛrt tɒmpɒgʌ po/ 'move bird det'. Participants were asked to indicate where a noun they could not recall should go in the sentence (by, for instance, saying X instead of the noun). This allowed us to include the data from incomplete responses. Participants were given as much time as they wanted to provide an answer. Responses were videotaped and later transcribed for analysis.

There were 24 test sentences (12 transitive, 12 intransitive). The test set was designed so that 12 nouns (the ones in the vocabulary test) each appeared once in each possible syntactic position (intransitive subject, transitive subject, and transitive object). The first use of any individual noun varied between subject and object in the transitive sentences; some nouns were first used as subjects and others as objects. All sentences used in this and other tests were novel.

Forced-choice grammar test

Next was a forced-choice test assessing what participants had learned about some of the consistent aspects of sentence structure. It served as a check to ensure that participants learned the language. Participants listened to 16 pairs of sentences and were asked to indicate which sounded most like a sentence from the language that they had been learning. The two sentences in each pair were versions of the same base sentence, one grammatical, the other ungrammatical. Test items were pre-recorded and presented via a minidisk deck (UR) or CD player (UCB). Participants heard both versions of the sentence and circled 1 or 2 on an answer sheet, depending on whether they preferred the first or second sentence in the pair. There was a 1-second pause between the sentences in a pair, and a 5-second pause between pairs.

One half of the test items examined participants' knowledge of verb subcategorization, i.e., whether they knew that transitive verbs required two nouns and intransitives only one. For the transitive sentences with a 'missing argument' (transitive verbs occurring with a single noun), either the subject or the object could be missing. That is, the noun phrase missing from the ungrammatical version of the sentence could be the first or second noun phrase in the grammatical version of the sentence. (As there was no visual scene presented alongside the test sentences, there is no actual subject or object.) The remaining items tested whether participants knew that a verb was required in every sentence. In these, the grammatical version was a verb and one or two noun phrases, as appropriate. The ungrammatical version was simply the same noun phrase(s). Whether the grammatical sentence occurred first or second was randomized, as was the ordering of sentence pairs in the test, with the constraint that no more than two pairs could occur in a row that tested the same rule or be of the same valence. Grammatical versions of all sentences in this test were in VS/VSO order.

¹²As noted by one reviewer, we did use written English glosses to teach the verbs' meanings, and so the English lexicons were very likely activated anyway. Moreover, given the evidence that bilinguals always have both lexicons activated to some degree (e.g. Marian & Spivey, 2003; Spivey & Marian, 1999), it would have been almost impossible to completely prevent L1 lexical activation. However, the degree of activation is manipulable (Marian & Spivey, 2003), and we wanted to minimize it as much as possible.

Results

We begin by presenting results from the tests examining participants' general knowledge, and then move on to examine their knowledge of word order patterns.

Forced-choice grammar test

Performance was quite high on this test, indicating that all participants learned the consistent aspects of the language well. Mean scores for the three conditions were: SOV = 14.5 (sd = 2.95), SOV & SVO = 15.9 (sd = .32), SVO = 14.4 (sd = 2.59). (Note that max = 16.) Performance on this task was not affected by input condition ($F_{(2,27)} = 1.358, p = .274$).

Vocabulary test

All participants scored at least 5 out of a possible 12 on the vocabulary test. The overall mean was 9.57 vocabulary items ($SD = 2.25$). Condition means varied slightly (SOV = 9.3, SOV & SVO = 10, SVO = 9.4), but not significantly ($F_{(2,27)} = .268, p = .767$).

Sentence production task

In this test, participants had to produce novel sentences. We noted the order of the constituents and computed the percentage of sentences each participant produced in any particular order. The data set included 355 intransitive and 349 transitive sentences (out of a possible 360 for each). Eight different participants (some in each condition) were missing at least one sentence. One possible intransitive sentence could not be included because the participant failed to produce anything for that sentence. Another was excluded because the participant produced two different orders and refused to choose one as their 'final answer'. In three more sentences, the participant produced only the noun or verb, and so the relative ordering of noun and verb could not be determined. For the transitives, three times a participant produced the same sentence in multiple orders and declined to indicate their preferred response. Four sentences were produced with only one noun, another three with two variables which failed to identify the subject and object nouns as such (e.g. X flimm Y), both of which make it impossible to determine the location of all *three* constituents reliably. The additional missing transitive was lost due to a problem with the video-tape.

Recall that conditions are identified by the less frequent word order(s) in the transitive sentences, thus those who heard 60% VSO and 40% SOV are labelled SOV, those who heard 60% VSO, 20% SOV, and 20% SVO are labelled SOV & SVO, and those who heard 60% VSO and 40% SVO are labelled SVO. Recall also that all participants heard 60% VS and 40% SV orders in the intransitive sentences.

Figure 1a shows the mean percentage of productions of the two word orders for intransitive sentences. Figure 1b shows the mean percentage of productions of the three orders for transitive sentences. Although there are other transitive orders that participants could have used, specifically, ones in which the object preceded the subject, these were quite rare (only 6 in 349 transitive sentences produced by participants), and so they are not shown in the figure.

As is clear from the figure, the productions were variable, and, at least in the intransitive sentences, participants appear to match the input percentages quite closely (SOV: VS = 58.18 sd = 31.05, SV = 41.81 sd = 31.05; SOV & SVO: VS = 58.25 sd = 32.55, SV = 41.74 sd = 32.55; SOV: VS 59.92 = sd = 28.99, SV = 40.08 sd = 28.99). With respect to the transitive sentences, readers may notice that SVO production is consistently high (SOV: VSO = 45.56 sd = 38.78, SOV = 25.99 sd = 30.24, SVO = 28.46 sd = 40.3; SOV & SVO: VSO = 39.17 sd = 30.38, SOV = 12.5 sd = 21.61, SVO = 44.67 sd = 32.71; SOV: VSO = 26.67 sd = 25.4, SOV = 0 sd = 0, SVO = 70.83 sd = 24.61). Indeed, even participants who heard no SVO sentences

produced some. This was not unexpected. This is the transfer or interference we anticipated. The degree and nature of this interference is discussed in more detail later. At this point, however, what is relevant is the variability in word orders produced by participants; they do not appear to be producing a single word order deterministically. We next go over the results for the intransitive and transitive sentences in more detail.

Intransitive Sentences

Possible rules—Although there is no evidence in the overall analyses that participants imposed a deterministic pattern on the language, it is possible that different individuals used different rules, something that would not be evident in the overall data. We therefore examined each individual participant's productions for evidence of a consistent pattern of word order use. Four participants used a single word order in intransitive sentences: three used VS and one used SV 100% of the time. Five participants used a single order in transitive sentences: four used SVO and one used VSO. Only *one* participant produced consistent word orders in both transitive and intransitive productions, that is, was consistent overall. Consistency was not dependant on condition (see Table 1).

There are other linguistically natural rules that participants could have imposed on the language that would lead to apparent variability despite the presence of a rule. These are orders that depend on some aspect of the semantics or other component of the grammar, with one word order being used in some contexts and another in other contexts, and we next examined productions for evidence of such rules.

We first examined the data for rules involving semantic roles. There were three different semantic roles associated with the argument of intransitive verbs: agent, patient, and theme, where agent is a volitional actor, patient is an affected entity, and theme is an entity over which a property is defined (Gasser, 2003; Payne, 1997). Representative verbs in each of the three categories are 'move', 'fall', and 'be blue' respectively. There were four semantic roles in the transitive sentences: agent, patient, theme, and location.¹³ The second analysis involved grammatical categories in the L1. The meanings encoded by intransitive verbs in the language are encoded by both verbs and adjectives in English (e.g. 'fall' and 'be blue' respectively), the transitive verbs have meanings encoded by verbs and prepositions (e.g. 'hit' and 'be inside of'). It is possible that participants treated the verbs differently when they corresponded to different parts of speech in the L1, using one order with meanings they associate with verbs and another with meanings they associate with adjectives, for example. The third possibility we examined was whether the noun class of the argument affected word order choice. Recall that the nouns were divided into two classes. The only consequence of noun class in the language is determiner selection; each class takes a different determiner. It is possible that participants imposed a rule in which different word orders were used with nouns in different classes, using a distinction from one aspect of the grammar to create a distinction in another.

Table 2 presents the mean percent of intransitive sentences produced in VS order by condition for each of these three possible rule types: semantic role, English grammatical category, noun class. Analyses show that participants are not imposing categorical rules based on the semantic role of the argument, and that this does not differ by input group. There are slight differences in the percentages, but these are not significant (Semantic role: $F_{(2,54)} = .1840, p = .169$; Input Group: $F_{(2,27)} = .049, p = .952$; Semantic role by Input Group: $F_{(2,27)} = .086, p = .918$). With respect to English-based grammatical categories, there is a slight trend towards more VS production with English adjectives than verbs that is marginally significant (English

¹³Some of the nouns, e.g. boat, may engage in actions such as falling or hitting without being volitional, and thus may be more correctly referred to as actors (rather than agents).

grammatical category: $F_{(1,27)} = 3.983, p = .056$; Input Group: $F_{(2,27)} = .053, p = .949$; English grammatical category by Input Group: $F_{(2,27)} = .103, p = .903$), however, the distinction was not categorical; they were not always producing one order with English verbs and another with English adjectives. As for noun class, participants did not use different word orders with nouns in different classes, even probabilistically, and this did not differ by input group (Noun class: $F_{(1,27)} = 1.576, p = .220$; Input Group: $F_{(2,27)} = .032, p = .948$; Noun class by Input Group: $F_{(2,27)} = .054, p = .948$).

The immediately preceding analyses were group analyses, but importantly, no individual participant used any of the three rule types categorically either, that is, no individual speaker produced consistently different word orders according to semantic role, English grammatical category, or noun class. Overall then, there is little evidence that participants' productions were governed by systematic rules.

Random variation vs. probability matching—The preceding results need not mean that the participants actually learned the input probabilities, however. It could be the case that when faced with such unpredictability the learner simply learns that there is no pattern with the result that she produces the various word orders at random. That is, she learned that there were multiple possible word orders, but not the particular probabilities associated with each. Thus, we next examined whether participants' variable productions accorded with the statistics present in the input. In particular, asking whether participants were producing the various word orders just as often as they had heard them.

An ANOVA comparing VS production (see Figure 1a) across the 3 conditions confirmed that there were no significant differences between the input groups ($F_{(2,27)} = .01, p = .990$), and so data from the three groups were combined for further analyses of intransitive productions. (Because there are only two possible word orders, VS and SV production percentages are complementary. Therefore, statistical tests using SV production as the dependent variable produce exactly the same results.) A one sample t-test comparing production (overall mean = 58.89) to the input percentage (60%) was not significant (2-tailed $t_{(29)} = -.223, p = .825$). While it is hard to say on this basis that participants were actually matching the input (since it is a null result), it is suggestive. However, there was also substantial individual variation (overall $sd = 29.82$, range = 0 – 100). In order to examine how closely individual participants were matching the input probabilities, we computed their distance from the input percentage and then examined how many people fell within 10, 20, 30, 40, 50 and 60% of the input. These data are shown in Table 3. (Data are shown separately for the three conditions despite the fact that condition did not affect distance from input, $\chi^2 = 13.96, df = 10, p = .174$). Seven of the thirty participants are within 10 percent of the input, and one-third are within 20. Although this is not the majority, clearly some participants are able to match the probabilities present in their input quite well. Note that this distribution is different from chance ($\chi^2 = 12.4, df = 5, p = .03$).

Although it may seem that these numbers do not reflect much probability-matching at the individual level, consider the nature of the task (and therefore the measurement): participants produced 10–12 intransitive sentences and we examined how often they produced the determiner. In this sample just 1–2 sentences can change their production by 8–20 percent. Moreover, they do not know in advance how many sentences they themselves will be producing, and so even if they were consciously tracking their own production probabilities, they would likely not be able to do so with precision. Thus, participants productions are simply a sample of their knowledge, and as such, likely to be an imperfect reflection. (Note also that in other studies showing linguistic probability-matching, participants often undermatch by about 10–15 percent. Thus, being within 10–20 percent is fairly consistent with previous demonstrations of probability-matching.)

There is evidence that multiple actors are more likely to converge on the ‘correct’ solution than any individual (e.g. Black, 1958; Surowiecki, 2004), even when the individuals are queried multiple times (Vul & Pashler, 2008), thus, we might expect probability-matching to be most apparent at the level of the community. Given the 355 total intransitive sentences produced by participants, perfect probability-matching would result in 213 VS and 142 SV sentences. Almost unbelievably, this is exactly what we have – 213 intransitive sentences in VS order, and 142 in SV order. Importantly, this is different from chance, ($\chi^2 = 14.2$, $df = 1$, $p < .001$).

Item-based vs. category-based probability-matching—Although the previous results suggest that participants are learning probabilities associated with abstract categories, it need not be the case. While the overall proportion of VS sentences in the input was .6, there was some variation among the individual verbs and nouns. This variation allows us to examine whether the apparent probability-matching is actually associated with categories, or instead, individual items. Figure 2 shows the relationship between input-based probabilities associated with the nouns/verbs occurring in the intransitive test sentences and participants’ productions. Figure 2a shows the proportion of intransitive sentences participants produced in VS order (X-axis) plotted against the input-based probability of VS order associated with the noun (Y-axis). Each data point represents the proportion of participants in one condition who produced VS order for one of the test sentences (and therefore one of the nouns). Thus, the maximum number of productions involved in computing each VS production-proportion is 10. (Each condition has 10 individuals who may each contribute a single production. Recall that 5 intransitive sentences are missing from the data set and so do not contribute to these computations, thus, some data points are based on fewer than 10 productions.)¹⁴ The input-based probabilities are simply the proportion of VS-ordered intransitive sentences in the input set which contain the noun in the test sentence. This was computed by examining all input sentences containing the particular noun, and computing the proportion of VS versus SV orderings in those sentences. For example, 33% of the time when the noun *flerbit* (cup) occurred in intransitive sentences in the input, the sentence was in VS order (and 67% of the time it was SV). Sometimes only a single noun is represented at a particular probability (e.g. .33), yielding three data-points at that location, one for each condition. Other times several nouns have the same probability of occurring in VS order (e.g. .78), and so more data points occur at that location. Figure 2b is the same, except that it shows the relationship between VS-ordered productions (Y-axis) and verb-based VS probabilities (X-axis). As the figure suggests, participants were not matching the probabilities at the level of individual nouns ($R^2 = .0004$, $B = .02$, $t = .123$, $p = .903$), suggesting the influence of category – based probabilities. However, their productions do reflect the probabilities associated with individual verbs; intransitive verbs that are more likely to occur in initial position in the input are more likely to be produced in initial position ($R^2 = .358$, $B = .988$, $t = 3.456$, $p < .001$). Thus, participants’ intransitive productions reflect some item-based and some category based probabilities.

Transitive Sentences

Participants’ productions of the three transitive word orders (Figure 1b) did not match the input quite as closely. In particular, all participants appear to show some interference from English, over-producing SVO even if there were no SVO sentences in their input and consequently under-producing VSO and SOV (when they heard it).

¹⁴Each of the nouns was used only once, and although some of the verbs were used in multiple sentences (three are used twice), seven of the four of the verbs were used in only one test sentence. Therefore, individual participants’ production-probabilities for any sentence (and therefore any noun or verb input-based probability) are restricted to either 0, .5, or 1. Thus, while a by-participant analysis could show the influence of input-based verb probabilities, it could not show probability matching. Thus, condition proportions were chosen. (Another alternative would have been to use a single proportion based on all three conditions for each sentence. However, that would have reduced the amount of data points (unnecessarily) to one.)

This impression is generally confirmed by the statistical analyses. We first examined whether there were significant differences between the groups' productions of the various word orders. If they are matching the probabilities we might expect differences for the orders that differed between groups (SVO and SOV) but not for VSO, which occurred equally often for all groups. This is exactly the pattern of results we found (SVO: $F_{(2,27)} = 4.155, p = .027$; SOV: $F_{(2,27)} = 3.668, p = .039$; VSO: $F_{(2,27)} = .901, p = .418$). We again conducted t-tests comparing input values with production values for the various word orders to see if the productions differed significantly from the input. Because the productions of SVO and SOV, but not VSO, sentences were significantly different between the three input conditions, comparisons were conducted using the combined data for VSO production, but separately by condition for SVO and SOV production. Participants produced significantly fewer VSO sentences than they heard ($t_{(29)} = -3.928, p < .0001$). All three conditions produced SVO sentences significantly more often than they had heard them (SOV: $t_{(9)} = 2.233, p = .052$; SOV & SVO: $t_{(9)} = 2.384, p = .041$; SVO: $t_{(9)} = 3.962, p = .003$), but underproduction of SOV was not significant for either group that heard SOV sentences (SOV: $t_{(9)} = -1.465, p < .177$; SOV & SVO: $t_{(9)} = -1.098, p = .301$). Thus, in contrast to the intransitive sentences, in general, participants do not appear to be matching the probabilities of the various orders associated with transitive sentences.

Effects of experimental manipulations—We next examined the data to see whether the two variables we manipulated might explain the differences between the input and participants' productions. The first variable we manipulated was the number of alternatives to the main word order (one alternative order in the SOV and SVO conditions versus two in the SOV & SVO condition); we had thought based on work in both linguistic and non-linguistic experiments that increasing the number of alternatives might induce increased overmatching (overproduction of the most frequent form). We thought that the second variable, the amount of English-like sentences, might affect the degree of interference experienced, and so the number of SVO sentences participants would produce, with more English-like sentences in the input leading to increased interference.

Figure 3 shows the data used for these comparisons – the deviation from the input for VSO and SVO sentences, which is simply the input percentage minus the production percentage. Positive deviation indicates over-production, negative indicates underproduction. Positive deviation from VSO would be consistent with overmatching, and positive deviation from SVO would indicate interference.

A comparison of the VSO deviation for participants exposed to one versus two alternative transitive word orders is not significant ($F_{(1,28)} = 1.049, p = .315$) – they were no more likely to overproduce VSO in the SOV & SVO condition than in the SVO and SOV conditions.¹⁵ In contrast, the other variable we manipulated, proportion of SVO sentences in the input (none, 20%, and 40%), did affect the number of SVO sentences participants produced. This is simply the overall effect reported above; more SVO sentences in the input leads to increased production of SVO sentences, an indication that participants are at least somewhat sensitive to their input. However, the more important question for us is whether this factor affects the degree to which participants match the proportion of SVO sentences in their input. In particular, does more SVO in the input induce greater interference and therefore increased over-production of SVO? To examine this, we compared the SVO deviation scores for the three groups. As is clear from the figure, the proportion of SVO sentences present in the input did not affect the severity of interference ($F_{(2,27)} = .113, p = .894$). The over-production of SVO, the order of transitive sentences in our participants' native language, was consistently 25–30% across groups.

¹⁵Despite appearances to the contrary, the SVO group is not underproducing VSO sentences significantly more than either of the other two conditions (SOV: $F_{(1,18)} = 1.66, p = .214$; SOV & SVO: $F_{(1,18)} = .996, p = .331$).

Interestingly, this same degree of overproduction did not occur in intransitive sentences. There, participants matched the probabilities in their input quite well. Why the discrepancy? One possibility is that, due to having three constituents rather than two learners must track, probabilities are more difficult to learn in transitive sentences. This interpretation would be consistent with results from studies of non-linguistic probability learning, where participants overmatch when making predictions over three or four rather than two lights (Gardner, 1957; Weir, 1972). However, those studies would also predict that deviations from the input would consist of overuse of the most frequent word order, VSO, which was not the case.

Another possibility is suggested by the Lexical Competence Hypothesis (LCH, Hudson & Eigsti, 2003). The LCH is based on the idea that production draws on limited cognitive resources (such as working memory), and that the sub-processes of production compete for those resources. Earlier sub-processes, such as lexical access, divert resources away from later processes, like sentence construction (see Bock, 1995, for a more complete description of sentence production).¹⁶ When speakers have difficulty finding all the words they need easily, sentence construction is disrupted. According to the LCH, because of this disruption, when L2 speakers have difficulty constructing a sentence they sometimes fall back on their more automatized L1 knowledge. This will be more of an issue when more words are required, as in transitive sentences. Thus, the LCH predicts that L1 knowledge will be more likely to interfere in transitive than intransitive sentences, something we found to be the case.

The theory makes a second relevant prediction: Speakers who know more words are more likely to find a match for a concept quickly and therefore expend fewer cognitive resources on lexical search. This leaves more resources available for sentence construction, with the result that sentences are more likely to be constructed according to the grammar of the new language. According to the theory then, we might expect to see differences in word order production for participants with larger versus smaller vocabularies, particularly in the transitive sentences; those with smaller vocabularies should show greater interference from English than participants with larger vocabularies. We therefore examined deviation from the input separately for participants with higher (High) and lower (Low) scores on the vocabulary test, which we defined as 9–12 or 5–8 words, respectively. Recall that the test was out of 12 and participants had to know at least 5 words. This division, then, represents a cut-off at the half-way point of possible scores. Nine participants fell into the low-vocab group, 21 into the high. Importantly, there were high- and low-vocab participants in each of the three input groups.¹⁷

Figure 4 shows the mean deviations for VSO and SVO sentences separately for high- and low-vocab participants. The data are collapsed across conditions for two reasons: First, the previous analysis showed that there was no difference in SVO deviation between the conditions. Second, given that there were only nine low-vocab participants, the number in each of the conditions is quite small, too small for any meaningful analyses. As is clear from the figure, although both groups of participants underproduce VSO and overproduce SVO, the difference between the input and productions is significantly larger for the low-vocab group than the high-vocab group for SVO sentences ($F_{(1, 28)} = 11.541, p = .002$), and marginally significant for VSO ($F_{(1, 28)} = 4.021, p = .055$). Thus, participants with lower vocabulary scores fall back on their native language knowledge to a greater degree, and match the input less well, than those with larger

¹⁶The original theory was developed with respect to L2 learners/speakers. (In typical adult L1 usage, production is so well practiced and automatized that competition is not evident.) However, there is some support for this from studies of children. Dapretto and Bjork (2000) found that children with larger productive vocabularies were more likely to name the objects being tested than those with smaller vocabularies, even though all children knew the words being tested.

¹⁷This division was selected because it represents the half-way point in the scale, not to demonstrate the conclusion, and the data need not have turned out as they did. Another way to divide the data would have been to divide the participants in half. However, there was no way to do this evenly, and potentially would have biased the results in our favour even more (at least for the high-vocab group, which would have been restricted to people with even higher scores). For information, the number of participants at each point on the vocab scale was as follows: 5 = 1, 6 = 3, 7 = 4, 8 = 1, 9 = 3, 10 = 5, 11 = 5, 12 = 8.

vocabularies. These data suggest that at least the high-vocab participants actually matched the probabilities present in their input fairly closely, something that was masked in the overall analyses by the participants with the lower vocabulary scores.¹⁸ There was variation in the deviations of the high-vocab participants (deviation from VSO range: $-50 - +40$; deviation from SVO range: $-20 - +75$), however the majority of the high-vocab participants were within $\pm 20\%$ of their input for both word orders.

Item-based vs. category-based probability-matching—Again, however, this apparent matching of overall probabilities could result from a sensitivity to the word-level probabilities only. Given the results from the intransitive sentences, we focused on the influence of verbs. There is a great deal less variation in the input-based probabilities associated with individual verbs in the transitive sentences than intransitive sentences. (Only four transitive verbs occurred in the test sentences, and the range of input-based probabilities they covered was smaller.) Nevertheless, given the intransitive sentence results, we began by examining whether there was a relationship between the input-based verb probabilities of VSO order and the proportion of people producing VSO for each sentence. Figure 5a is a scatterplot of the proportion of VSO ordered sentences produced for each sentence and each condition by the input-based probability of that verb appearing in first position (i.e. VSO). Again, each data point represents the proportion of people in the relevant condition who produced VSO order for one test sentence. As compared to the intransitive sentences, here we see a different pattern; there is no relationship between the input and output proportions ($R^2 = .035$, $B = .652$, $t = 1.11$, $p = 2.74$). However, when we consider only the high-vocab participants (5b), who as we showed previously more closely matched the overall probability of VSO sentences in their input, there is a significant relationship between the input and output ($R^2 = .151$, $B = 1.646$, $t = 2.463$, $p = .019$). The relationship is not as strong as it was for intransitives, but it is significant nevertheless.¹⁹

We next examined the potential effect of joint unconditional probabilities on production, as it is possible that the item-based input probabilities interact in the transitive sentences. These are the products of the input-based probabilities that each particular word in the test sentence will appear in that particular location, independent of its grammatical role in the sentence, and thus are slightly different for the three conditions, due to the differing probabilities of nouns occurring in second and third position. E.g. for the test sentence ‘the bird hits the ladybug’ the input-based probability of ‘hit’ in position 1 is .689 for all three conditions, and the input-based unconditional probability of ‘bird’ in position 2 is .389 in SOV, .353 in SOV & SVO, and .385 in SVO. (The probability of ‘ladybug’ in position 3 is irrelevant for the joint probability computation, as its position is fixed once the other two words have been placed in positions 1 and 2, but for completeness sake, it is .268, .265, and .243, for SOV, SOV & SVO, and SVO respectively. For this same reason, joint probabilities are not relevant for the intransitives.) Thus, by multiplying the probabilities of bird in position 1 and hit in position 2, we get the

¹⁸One might ask if low-vocab participants simply learned the language less well, including vocabulary and grammar. (This, of course, would fit in just as well with the overall conclusion that some participants learned the probabilities associated with higher-order categories.) However, this does not appear to be the case. Vocabulary scores are significantly correlated with how closely participants match the proportion of SVO sentences in their input, whereas scores on the general grammar test are not. This pattern is true in both simple regressions (vocab: $B = -6.535$, $R^2 = .201$, $t = -2.650$, $p = .013$; grammar test: $B = -3.443$, $R^2 = .061$, $t = -1.34$, $p = .186$) and a multiple regression (partial correlations – vocab: $\beta = -.448$, $t = -2.6$, $p = .015$; grammar test: $\beta = -.247$, $t = -1.237$, $p = .196$). This is not due to a lack of variability in the grammar scores. Although the mean grammar test scores are high, they are not perfect, ranging from 8 – 16.

¹⁹A demonstration that noun probabilities are not determining production in the transitives: if categories are ignored (i.e. subject and object) in the SVO and SOV & SVO conditions, for 18 out of the 24 test sentences (12 test sentences $\times$ 2 conditions) the subject noun is most likely to occur in the third position in the sentence in the input. Thus, if noun-level category-independent probabilities were driving production, participants in these two conditions should be more likely to produce these particular test sentences with the subject noun in final position. However, there were only 6 productions out of 236 with the subject at the end. The majority of these (4) were produced by people in the SOV & SVO condition, where the strength of the third-position-bias was actually smaller than in the SVO condition.

joint probabilities of the word order corresponding to VSO for the test sentence ‘the bird hits the ladybug’ for the three conditions: .268 (SOV), .243 (SOV & SVO), and .265 (SVO). We computed the input-based joint unconditional probabilities for each sentence and each condition and compared them to VSO production proportions for each condition that sentence, as shown in Figure 6. Results showed that these joint probabilities were not significantly correlated with production ($R^2 = .05$, $B = .651$, $t = 1.345$, $p = .188$). This was also true when we examined only the high-vocab participants – the input-based joint probabilities were not related to the likelihood of VSO production ($R^2 = .034$, $B = .648$, $t = 1.094$, $p = .282$).²⁰

Thus, for the transitive sentences it does seem that the participants (at least the high-vocab participants) are matching the probabilities of words orders associated with individual verbs only. On the basis of the transitive results, we went back and examined the correlations between the verb-based probabilities and production proportions in the intransitive sentences separately for the high- and low-vocab participants. Interestingly, we only found a significant correlation for the high-vocab participants ($R^2 = .215$, $B = .985$, $t = 3.05$, $p = .004$), not the low-vocab participants ($R^2 = .102$, $B = 1.05$, $t = 1.58$, $p = .128$).

Discussion

This paper examined whether learners can acquire the statistics associated with abstract higher-order linguistic categories and their relationships to each other. Adult learners, who we know can learn probabilistic patterns in languages (e.g., Hudson Kam & Newport, 2005), were exposed to an artificial language containing unpredictable variability in the relative ordering of Subject, Verb, and Object. If grammar learning is accomplished via an algebraic learning mechanism that does not include information about statistics, we would have expected participants to either 1) have ended up with deterministic rule-like knowledge of word order or 2) simply failed to learn any generalization about word order and produce various word orders randomly. In contrast, if learners can acquire statistics over abstract categories as well as items, then participants’ productions should reflect the likelihood of the various word orders in their input. We found little evidence of deterministic rules in participants’ productions, however, productions were variable. Importantly, the variation was not random. Rather, the variation reflected the probabilities associated with the various orders, at least for the high-vocab participants.

They were not the overall category-based probabilities, however, they were the item-based probabilities associated with individual verbs. At first glance, it may seem that this result indicates that participants were not learning probabilities associated with grammatical categories. However, the fact that the same was not true for nouns argues against this. Verbs as a part of speech are important determinants of the structure of the sentence they occur in; they control the number of arguments required or allowed, as well as the ways in which those

²⁰We did another analysis that involved the object-noun probabilities as well and found similar results. We simply added the probability of the verb in Position 1 to the probability of the Subject noun appearing in Position 2 and the probability of the Object noun appearing in Position 3. These numbers were always over 1, and so are not actually probabilities. If the input-based probabilities of all words in the test sentence are contributing to the order produced, however, sentences with higher numbers should be more likely to be produced in VSO order than those with lower numbers. However, these regressions (whether with all participants or high-vocab participants only) were not significant either. Another possible influence is bigrams in the input, in particular, the orders associated with input sentences in which two of the three constituents in the test sentences occurred together. E.g. for the test sentence ‘the bird hits the ladybug’ sentences where a bird hit something or something hit a ladybug would be relevant. There were no input sentences where something hit a ladybug; there was one where a bird hit a tree. Because the input set was repeated three times, participants heard this sentence three times, twice in VSO order. Therefore, the probability of VSO associated with the bigram bird-hit was .667. (Since the test sentences were novel, the three constituents never occurred in the input sentences together, likewise the bigrams in the intransitive test sentences were nonexistent in the input.) Not every test sentence had any input bigrams (8 had SV bigrams, 7 had VO bigrams), and few values were represented for those that did (4 different VSO probabilities for the SV bigrams, same for the VO bigrams). Thus, the data were relatively sparse. However, we did perform regression analyses which showed that the probability associated with VSO order was not significantly correlated with VSO production for either bigram type.

arguments can be expressed. For example, I can ‘eat’, or ‘eat a sandwich’, but can only ‘sleep’ – it is not possible to ‘sleep a nap’. Similarly, ‘donate’ and ‘give’ can both be used to describe the same event, but they place different restrictions on possible sentences. One can ‘give the library a book’ or ‘give a book to the library’, but one can only ‘donate a book to the library’, ‘donating the library a book’ is not possible. Thus, verbs *qua* verbs are important to the actual sentence structures planned and produced. The same is not true for nouns, they do not tend to determine the structural properties of the sentences they occur in. Thus, it seems that verbs are determining the production probabilities precisely *because* of their grammatical category – participants associate the probabilities with individual *verbs*, not simply individual words that happen to be verbs. Moreover, the fact that item-based probability-matching was only found for verbs and not nouns suggests something about the nature of the representations participants are working with, that they are learning something along the lines of $p(\text{VS}|\text{‘fall’})$ or $p(\text{VSO}|\text{‘hit’})$. Thus, although the probabilities are associated with individual verbs, they involve relationships between abstract categories.

This general result was unaffected by the number of alternative word orders participants heard. We had reasoned based on previous work (Hudson Kam & Newport, in press) that participants exposed to more alternatives might be more likely to overuse the more frequent form, and thus show less evidence of having learned the probabilities present in the input. However this was not the case: We found no evidence that the number of alternative word orders affected participants’ learning. This discrepancy is potentially due to having too few alternatives in the current study. In previous work where adult learners regularized over inconsistency there were a large number of low frequency forms alternating with a single high frequency form. (Two and four low frequency forms produced some regularization, however it was only with 8 or 16 low frequency forms that adults regularized to the same degree as children.) However, in order to keep all input as naturalistic as possible in this study, we used only word orders in which the subject preceded the object, restricting us to three possible orders.

As anticipated, there was interference from participants’ native language, at least in the transitive sentences, where participants produced more SVO sentences than they heard. Although the degree of overproduction was the same for all input groups, there were important differences between individuals; participants with lower vocabulary scores showed a high degree of influence from English word order, whereas those who knew more vocabulary (who comprised 2/3 of the sample) deviated substantially less from their input. As mentioned, interference was something we had anticipated, and in fact, were not unhappy to see, as it mitigates the objection that participants were not treating our experiment as an exercise in language learning, and so, could have been using learning mechanisms not typically deployed by language learners. The fact that learners experienced interference from their L1 strengthens our contention that we are studying language learning, at least with respect to adult learners. Note that we are not claiming that children (or L1 learners) will necessarily behave in the same fashion. In fact, given our own previous work we would anticipate that they likely would not. What we are claiming is that we have discovered something about the learning mechanisms involved in language acquisition more generally, in the same way that our other work with adult participants has, and that is that learning probabilities that apply to relationships between abstract categories is possible.

A potential objection to this conclusion comes from the fact that we have no direct evidence that participants have acquired statistics related to the categories Subject and Object. In real languages these are typically categories with consequences in other aspects of the grammar, such as verb agreement or passivization (Keenan, 1976). The miniature language had no such properties, and so there is no way to probe whether participants actually extracted such abstract grammatical categories. It is possible that the participants could have been using semantic roles like agent, and therefore, that the probabilities they learned were attached to semantic and not

grammatical categories. However, semantic roles are also abstract, and so even this sort of knowledge would indicate that we can learn probabilities over categories rather than just particular items, which was the question we set out to address here. In addition, since the verbs used included a variety of semantic roles (e.g. agent, theme, and patient), if participants had learned probabilities over semantic rather than syntactic categories, they would have had to learn separate probabilities for each semantic role. This is a great deal more to have learned, and so, a more complicated explanation. It is much simpler to suppose that they have learned probabilities over grammatical categories.

This is not the first study to show a relationship between statistics and syntactic categories. Numerous studies have shown that learners can acquire categories via distributional learning (e.g., Gerken, et al., 2005; Gómez & Lakusta, 2004; Mintz, 2002), a necessary step in statistical syntax learning. Demonstrations of the acquisition of relationships between categories are relatively more scarce. Saffran (2001) and Thompson and Newport (2007) showed that adults can learn phrase boundaries on the basis of transitional probabilities, presumably computed over categories rather than items. In another examination of statistical syntax learning Kaschak and Saffran (2006) exposed learners to a language that contained occasional exceptions to the major patterns, and found that under certain circumstances people could still acquire the major patterns despite the exceptions. As mentioned previously, these studies were open to the interpretation that the categories were learned distributionally, while the relationships between them were not. Our findings, however, bolster their claims of statistical learning of the syntax itself, by showing quite clearly that statistics can be learned with respect to abstract categories.

There is also a great deal of work in the processing literature showing that speakers/listeners are sensitive to the likelihood of certain types of arguments or constructions associated with individual verbs (e.g., Bock & Irwin, 1980; Gahl & Garnsey, 2004; Garnsey, Pearlmutter, Myers, & Lotocky, 1997; Stallings, MacDonald, & O'Seaghdha, 1998), mirroring our learning results. For example, readers have less difficulty with temporary ambiguities (points at which multiple parses are possible) when the ambiguity is resolved in favour of its bias, or the structure in which it most often appears. *Accept*, for instance, is much more likely to be followed by a direct object (as in 3a) than a sentential complement (3b), while *assume* has the opposite bias; it is more likely to be followed by a sentential complement (4b) than a direct object (4a) (Garnsey, et al., 1997).²¹ (The temporary ambiguity in these examples occurs at the noun phrase after the verb, which can be parsed either as a direct object or the subject of an upcoming sentential complement.)

(3)

- a. The father accepted the money.
- b. The father accepted the story his daughter told him was true.

(4)

- a. The father assumed the blame.
- b. The father assumed the story his daughter told him was true.

Although many people accept that likelihood or bias information is tied to abstract categories such as sentential complement or direct object, and is therefore part of our grammar (see, e.g., Gahl & Garnsey, 2004), others disagree, pointing out that the biases are often related to meanings, and so, not necessarily tied to syntactic categories (Newmeyer, 2006; cf. Gahl &

²¹In almost all previous demonstrations of sensitivity to this sort of frequency information, the speaker/listener has been shown only to distinguish the relatively more and less frequent options from each other (and possibly from unbiased options). However there is at least one other study that, like ours, suggest that this sensitivity is more fine-grained (Jennings, Randall, & Tyler, 1997).

Garnsey, 2006). In the present study this is not the case – the different word orders were not correlated with different sorts of contexts or meanings – and so if participants did know the probabilities, they must have learned them with respect to abstract categories.

Summary

This study was designed to assess whether language learners can acquire probabilities over the relationships between the categories involved in word order (S, V, & O). This type of variation is particularly interesting because it must be computed relative to an abstract category if it is to be learned. The data indicate that at least the high-vocab participants did indeed learn very specific probabilities, demonstrating that, at least in adults, statistical learning is not restricted to tracking specific items and relationships between them (e.g. sounds or syllables to find words). These results therefore provide support for proposals such as those of Thompson and Newport (2007) and Saffran (2001) that statistical learning mechanisms might be involved in the acquisition of syntax.

Acknowledgments

This research was supported by NIH grants HD 048572 to C. L. Hudson Kam and DC 00167 to E. Newport and T. Supalla.

Thank you to Joanne Esse, Xi Sheng, and Nicole Edwards for their help collecting and coding the data. Thanks also to Sarah Creel and Elissa Newport for helpful comments on earlier drafts of this paper.

References

- Bickerton, D. Roots of language. Ann Arbor, MI: Karoma; 1981.
- Bickerton, D.; Givon, T. Pidginization and syntactic change: From SXV and VSX to SVX. In: Steever, SB.; Walker, CA.; Mufwene, S., editors. Papers from the parasession on diachronic syntax. Chicago: Chicago Linguistic Society; 1976. p. 9-39.
- Black, D. The theory of committees and elections. Cambridge: Cambridge University Press; 1958.
- Bock, K. Sentence production: From mind to mouth. In: Eimas, JL.; Eimas, PD., editors. Speech, language, and communication. New York: Academic Press; 1995. p. 181-216.
- Bock JK, Irwin DE. Syntactic effects of information availability in sentence production. *Journal of Verbal Learning and Behavior* 1980;19:467–484.
- Bowerman, M. Early syntactic development: A cross-linguistic study with special reference to Finnish. Oxford: Cambridge University Press; 1973a.
- Bowerman, M. Structural relationships in children's utterances: Syntactic or semantic. In: Moore, TE., editor. Cognitive development and the acquisition of language. New York: Academic Press; 1973b. p. 197-213.
- Braine MDS, Brody RE, Brooks PJ, Sudhalter V, Ross JA, Catalano L, Fisch SM. Exploring language acquisition in children with a miniature artificial language: Effects of item and pattern frequency, arbitrary subclasses, and correction. *Journal of Memory and Language* 1990;29:591–610.
- Cain J, Weber-Olsen M, Smith R. Acquisition strategies in a first and second language: Are they the same? *Journal of Child Language* 1987;14:333–352. [PubMed: 3611246]
- Christiansen, MH.; Curtin, SL. The power of statistical learning: No need for algebraic rules. In: Hahn, M.; Stoness, SC., editors. Proceedings of the Twenty First Annual Conference of the Cognitive Science Society. Mahwah, NJ: Erlbaum; 1999. p. 114-119.
- Comrie, B. Language universals and linguistic typology: Syntax and morphology. 2. Chicago: University of Chicago Press; 1989.
- Coppola M, Newport EL. Grammatical Subjects in home sign: Abstract linguistic structure in adult primary gesture systems without linguistic input. *Proceedings of the National Academy of Sciences* 2005;102:19249–19253.
- Dapretto M, Bjork EL. The development of word retrieval abilities in the second year and its relation to early vocabulary growth. *Child Development* 2000;71:635–648. [PubMed: 10953930]

- Dryer, MS. Order of adposition and noun phrase. In: Haspelmath, M.; Dryer, MS.; Gil, D.; Comrie, B., editors. *The world atlas of language structures*. Oxford University Press; 2005. p. 346-347.
- Estes, WK. Probability Learning. In: Melton, AW., editor. *Categories of Human Learning*. New York: Academic Press; 1964. p. 89-128.
- Estes WK. The cognitive side of probability learning. *Psychological Review* 1976;83:37-64.
- Finn AS, Hudson Kam CL. The curse of knowledge: First language knowledge impairs adult learners' use of novel statistics for word segmentation. *Cognition* 2008;108:477-499. [PubMed: 18533142]
- Frigo L, McDonald JL. Properties of phonological markers that affect the acquisition of gender-like subclasses. *Journal of Memory and Language* 1998;39:218-245.
- Gahl S, Garsney SM. Knowledge of grammar, knowledge of usage: Syntactic probabilities affect pronunciation variation. *Language* 2004;80:748-775.
- Gahl S, Garsney SM. Knowledge of grammar includes knowledge of syntactic probabilities. *Language* 2006;82:405-410.
- Gardner RA. Probability-learning with two and three choices. *American Journal of Psychology* 1957;70:174-185. [PubMed: 13424757]
- Garnsey SM, Pearlmuter NJ, Myers E, Lotocky MA. The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language* 1997;37:58-93.
- Gasser, M. How language works. 2.02003. Retrieved August 6, 2006, from www.indiana.edu/~hlw/book.html
- Gerken L, Wilson R, Lewis W. Infants can use distributional cues to form syntactic categories. *Journal of Child Language* 2005;32:249-268. [PubMed: 16045250]
- Gómez RL, Gerken L. Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences* 2000;4:178-186. [PubMed: 10782103]
- Gómez RL, Lakusta L. A first step in form-based category abstraction by 12-month-old infants. *Developmental Science* 2004;7:567-80. [PubMed: 15603290]
- Greenberg, JH. Some universals of grammar with particular reference to the order of meaningful elements. In: Greenberg, JH., editor. *Universals of language*. Cambridge, MA: MIT Press; 1963. p. 73-113.
- Hudson CL, Eigsti IM. The Lexical Competence Hypothesis: A cognitive account of the relationship between vernacularization and grammatical expansion in creolization. *Journal of Pidgin and Creole Languages* 2003;18:1-79.
- Hudson Kam CL, Newport EL. Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development* 2005;1:151-195.
- Hudson Kam CL, Newport EL. Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology*. in press.
- Jackendoff, RS. *Semantic Structures*. MIT Press; Cambridge, MA: 1990.
- Jennings F, Randall B, Tyler LK. Graded effects of verb subcategory preferences on parsing: Support for constraint-satisfaction models. *Language and Cognitive Processes* 1997;12:485 - 504.
- Kaschak MP, Saffran JR. Idiomatic constructions and language learning. *Cognitive Science: A Multidisciplinary Journal* 2006;30:43-63.
- Keenan, EL. Towards a universal definition of subject. In: Li, C., editor. *Subject and Topic*. New York: Academic Press; 1976. p. 304-333.
- Li, CN.; Thompson, SA. Subject and topic: A new typology of language. In: Li, CN., editor. *Subject and Topic*. New York: Academic Press; 1976. p. 461-466.
- Marcus GF. *Pabiku* and *Ga Ti Ga*: Two mechanisms infants use to learn about the world. *Current Directions in Psychological Science* 2000;9:145-147.
- Marcus GF, Berent I. Are there limits to statistical learning [Letter to the editor]? *Science* 2003;300:53-54. [PubMed: 12677042]
- Marcus GF, Vijayan S, Rao SB, Vishton PM. Rule learning in 7-month-old infants. *Science* 1999;283:77-80. [PubMed: 9872745]
- Marian V, Spivey M. Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition* 2003;6:97-115.

- Mintz TH. Category induction from distributional cues in an artificial language. *Memory & Cognition* 2002;30:678–686.
- Mintz TH. Frequent frames as a cue for grammatical categories in child directed speech. *Cognition* 2003;90:91–117. [PubMed: 14597271]
- Mintz TH, Newport EL, Bever TG. The distributional structure of grammatical categories in speech to young children. *Cognitive Science* 2002;26:393–424.
- Morgan JL, Meier RP, Newport EL. Structural packaging in the input to language learning: Contributions of prosodic and morphological marking of phrases to the acquisition of language. *Cognitive Psychology* 1987;19:498–550. [PubMed: 3677585]
- Morgan JL, Meier RP, Newport EL. Facilitating the acquisition of syntax with cross-sentential cues to phrase structure. *Journal of Memory and Language* 1989;28:67–85.
- Newmeyer FJ. Grammar is grammar and usage is usage. *Language* 2003;79:682–707.
- Newmeyer FJ. Discussion Note: On Gahl and Garnsey on grammar and usage. *Language* 2006;82:399–404.
- Newport, EL.; Aslin, RN. Innately constrained learning: Blending old and new approaches to language acquisition. In: Howell, SC.; Fish, SA.; Keith-Lucas, T., editors. *Proceedings of the 24th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press; 2000. p. 1–21.
- Odlin, T. *Language transfer: Cross-linguistic influence in language learning*. Cambridge: Cambridge University Press; 1989.
- Payne, TE. *Describing morphosyntax: A guide for field linguists*. New York: Cambridge University Press; 1997.
- Peña M, Bonatti LM, Nespor M, Mehler J. Signal driven computations in speech processing. *Science* 2002;298:604–607. [PubMed: 12202684]
- Saffran JR. The use of predictive dependencies in language learning. *Journal of Memory and Language* 2001;44:493–515.
- Saffran JR. Constraints on statistical language learning. *Journal of Memory and Language* 2002;47:172–196.
- Saffran JR, Aslin RN, Newport EL. Statistical learning by 8-month-old infants. *Science* 1996;274:1926–1928. [PubMed: 8943209]
- Seidenberg MS, MacDonald MC, Saffran JR. Does grammar start where statistics stop? *Science* 2002;298:553–554. [PubMed: 12386323]
- Seidenberg MS, MacDonald MC, Saffran JR. Response [Letter to the editor]. *Science* 2003;300:54.
- Singleton JL, Newport EL. When learners surpass their models: The acquisition of American Sign Language from impoverished input. *Cognitive Psychology* 2004;49:370–407. [PubMed: 15342259]
- Slobin, D. The acquisition of Russian as a native language. In: Smith, F.; Miller, GA., editors. *The genesis of language: A psycholinguistic approach*. Cambridge, MA: MIT Press; 1966. p. 129–48.
- Spivey M, Marian V. Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science* 1999;10:281–284.
- Stallings LM, MacDonald MC, O'Seaghdha PG. Phrasal ordering constraints in sentence production: Phrase length and verb disposition in heavy-NP shift. *Journal of Memory and Language* 1998;39:392–417.
- Stassen, L. *Intransitive predication*. Oxford: Clarendon Press; 1997.
- Stassen, L. Predicative Adjectives. In: Haspelmath, M.; Dryer, MS.; Gil, D.; Comrie, B., editors. *The world atlas of language structures*. Oxford: Oxford University Press; 2005. p. 478–81.
- Surowiecki, J. *The wisdom of crowds*. New York: Random House; 2004.
- Thompson SP, Newport EL. Statistical learning of syntax: The role of transitional probability. *Language Learning and Development* 2007;3:1–42.
- Vul E, Pashler H. Measuring the crowd within: Probabilistic representations within individuals. *Psychological Science* 2008;19:645–647. [PubMed: 18727777]
- Weir MW. Developmental changes in problem-solving strategies. *Psychological Review* 1964;71:473–490. [PubMed: 14218072]

Weir MW. Probability performance: Reinforcement procedure and number of alternatives. *American Journal of Psychology* 1972;85:261–270.

Wonnacott, E.; Newport, EL. In: Brugos, A.; Clark-Cotton, MR.; Ha, S., editors. Novelty and regularization: The effect of novel instances on rule formation; BUCLD 29: Proceedings of the 29th Annual Boston University Conference on Language Development; Somerville, MA: Cascadia Press; 2005. p. 663-673.

Figure 1a

Figure 1b

Figure 1. Mean production of sentences in different word orders for intransitive (a) and transitive (b) sentences by input condition.

Figure 2a

Figure 2b

Figure 2. Relationship between noun- (a) and verb-based (b) input and production probabilities (intransitive sentences).

Figure 3.
Mean over- and under-production of SVO and VSO by input group.

Figure 4.
Mean over- and under-production of SVO and VSO for high- and low-vocab participants.

Figure 5a

Figure 5b

Figure 5. Relationship between item-based verb probabilities and production for all participants (a) and high-vocab participants (a).

Figure 6.
Relationship between item-based joint probabilities and production.

Table 1

Consistency of production by Input group

Input Group	Production Consistency Type			
	Intrans	Trans	Both	Other
SOV	1	2	0	0
SOV & SVO	1	0	1	0
SVO	1	2	0	0
				7
				8
				7

Table 2

Mean percent VS production by input condition (SE in parentheses)

'Rule Type'	Input Condition		
	SOV	SOV & SVO	SVO
Semantic Role			
Theme	67.1% (7.7)	64.3% (11.9)	64.3% (9.1)
Patient	60% (16.3)	50% (16.7)	50% (10.5)
Agent	50% (12.4)	50% (11.2)	60% (12.9)
English Category			
Verb	55% (12.4)	49.5% (12.2)	56% (10.7)
Adj	67.1% (7.7)	64.2% (11.9)	64.3% (9.1)
Noun Class			
NC1	64.1% (10.6)	62.5% (10.7)	64.2% (7.6)
NC2	60.9% (9)	56.1% (10.7)	58.8% (10.2)

Table 3

Individual distance from input percentages by input condition

Input Condition	Distance from input (in percentage)						
	0–10	11–20	21–30	31–40	41–50	51–60	
SOV	2	0	3	4	1	0	
SOV & SVO	3	0	1	4	0	2	
SVO	2	3	0	3	0	2	