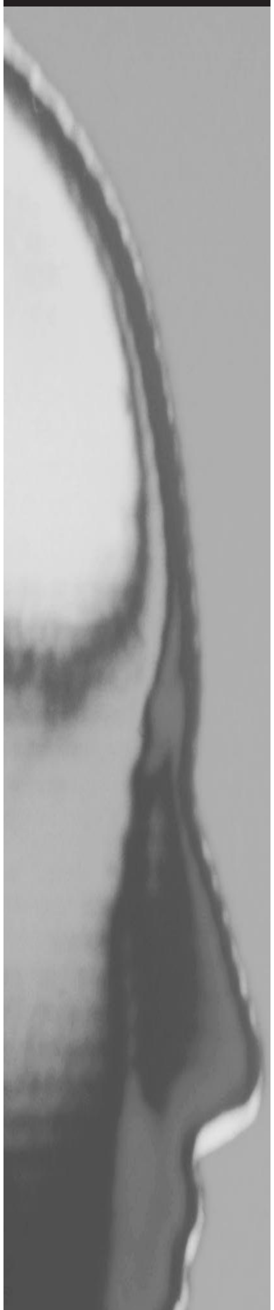
68 Clark, R.E. and Squire, L.R. (1998) Classical conditioning and brain systems: the role of awareness. *Science* 280, 77–81
69 Squire, L.R. (1992) Declarative and non-declarative memory: multiple brain systems supporting learning and memory. *J. Cogn. Neurosci.* 99, 195–231
70 Hebb, D.O. (1949) *The Organization of Behavior*, John Wiley & Sons
71 Murray, E.A. and Bussey, T.J. (1999) Perceptual-mnemonic functions of the perirhinal cortex. *Trends Cognit. Sci.* 3, 142–151
72 Epstein, R. *et al.* (1999) The parahippocampal place area: recognition, navigation, or encoding? *Neuron* 23, 115–125



Infant artificial language learning and language acquisition

Rebecca L. Gómez and LouAnn Gerken

The rapidity with which children acquire language is one of the mysteries of human cognition. A view held widely for the past 30 years is that children master language by means of a language-specific learning device. An earlier proposal, which has generated renewed interest, is that children make use of domain-general, associative learning mechanisms. However, our current lack of knowledge of the actual learning mechanisms involved during infancy makes it difficult to determine the relative contributions of innate and acquired knowledge. A recent approach to studying this problem exposes infants to artificial languages and assesses the resulting learning. In this article, we review studies using this paradigm that have led to a number of exciting discoveries regarding the learning mechanisms available during infancy. These studies raise important issues with respect to whether such mechanisms are general or specific to language, the extent to which they reflect statistical learning versus symbol manipulation, and the extent to which such mechanisms change with development. The fine-grained characterizations of infant learning mechanisms that this approach permits should result in a better understanding of the relative contributions of, and the dynamic between, innate and learned factors in language acquisition.

R.L. Gómez is at the Department of Psychology, Johns Hopkins University, Baltimore, MD 21218-2685, USA.

tel: +1 410 516 4018
fax: +1 410 516 4478
e-mail: gomez@jhu.edu

L. Gerken is at the Department of Speech and Hearing Sciences, University of Arizona, Tucson, AZ 85721, USA.

Language acquisition is one of the most complex learning tasks imaginable. The daunting nature of the undertaking arises from conflicting pressures to generalize beyond the stimuli encountered without generalizing too far. For example, it has been observed that children never erroneously transform a statement like '*The man who is tall is Sam*' into the question '*Is the man who tall is Sam?*' (by moving the subordinate clause verb rather than the main verb to the front of the sentence). The lack of such errors has been taken as evidence that children never consider rules based solely on linear order in sentences, such as 'move the first verb to the front of the sentence'. The computational and logical difficulties raised by these conflicting pressures have caused many researchers to conclude that language is not learnable by an unspecialized learning device[1–3]. Rather, humans must be born with some number of built-in constraints for deciding when and how to generalize from the stimuli they encounter. This view of a constrained language learner has dominated the field for the past 30 years or so. However, recent advances in cognitive science are causing us to reconsider the type and degree of constraints placed on the learner. Of particular interest, and the focus of this article, are recent studies on infants' ability to acquire information about miniature artificial languages after very brief exposure.

The complexity of natural language makes it exceedingly difficult to isolate factors responsible for language learning. For instance in English, when words like *the* and *a* occur at the beginnings of sentences or clauses they tend to be accompanied by intonational patterns involving brief pausing and reduced stress. There has been considerable speculation that such cues might help learners discover the syntax of their native language[4] and, although infants appear to be sensitive to these features of sentences and clauses[5,6], we do not know whether they are responding to pauses, reduced stress, frequently occurring words or some combination of the above. Language researchers have thus turned to artificial languages as a means of obtaining better control over the input to which learners are exposed. Artificial languages can be designed to

1364-6613/00/$ – see front matter PII: S1364-6613(00)01467-4

Box 1. Learning *in utero*

Artificial-language studies with infants demonstrate the presence of remarkably sophisticated learning abilities by seven, eight and 12 months of age (Refs a–d). Such findings inevitably raise questions regarding how early learning might occur. In fact, there is reason to believe that learning begins *in utero*. One of the earliest indications was the finding that newborns prefer their mother's voice to that of another female (Ref. e). Newborns also distinguish sentences from their native language from sentences from another language. Passages read in French produced higher sucking rates (as measured by an operant sucking procedure) in French newborns than passages read in Russian (Ref. f). Presumably, such preferences are shaped by prenatal experience with maternal speech. Intrauterine recordings indicate that the low-frequency components of maternal speech, including its prosodic (or rhythmic) qualities, are audible *in utero* and late-term fetuses consistently respond to sound (Refs g,h), raising the possibility that learning might begin sometime during the last trimester of gestation.

Additional evidence for learning *in utero* comes from experiments showing that newborns discriminate a passage read aloud by their mothers during the last six weeks of pregnancy from an unfamiliar one (Ref. i). Two-day-old newborns were tested, using an operant learning procedure, to see whether the familiar passage would be more reinforcing than an unfamiliar one. The familiar passage was indeed more reinforcing, even when read in another woman's voice, suggesting that infants had learned certain features of their training passage *in utero* (possibly involving the rhythmic qualities of the infant's particular training story). The fact that newborns made the discrimination even when passages were read in another woman's voice, demonstrates that they had acquired information specific to the passages, rather than only to their mother's voice. Further evidence for learning *in utero* comes from a study testing learning in 37-week-old fetuses (Ref. j). Mothers repeatedly recited one of two rhymes out loud, once a day over a four-week period. At the end of this time their fetuses were stimulated with recordings of both the familiar and unfamiliar rhymes. The familiar rhyme consistently elicited a decrease in fetal heart rate whereas the unfamiliar one did not, suggesting that the fetuses discriminated the two.

Although these findings are extraordinary, there are more than likely to be limitations on such learning. For instance, as noted above, the sound transmitted to the fetus is primarily of very low frequency, and hence lacks the type of detail needed for making fine-grained acoustic distinctions. Furthermore, once born, infants do not begin showing preferences for more complex information, having to do with constraints on legal sound patterns in their native language or cues marking phrasal and clausal units, until sometime between six and nine months of age (Refs k–n), suggesting that fetuses could not be acquiring information at this level of detail. However, the fact that learning begins so early (even if in a very rudimentary way) demonstrates that the sensitivities observed at birth are as likely to result from a gradual accumulation of knowledge as from the presence of innate constraints, making it all the more important to obtain a detailed understanding of the developmental trajectories of the mechanisms involved.

References

a Gómez, R.L. and Gerken, L.A. (1999) Artificial grammar learning by one-year-olds leads to specific and abstract knowledge. *Cognition* 70, 109–135

b Marcus, G.F. *et al.* (1999) Rule learning by seven-month-old infants. *Science* 283, 77–80

c Saffran, J.R. *et al.* (1996) Statistical learning by eight-month-old infants. *Science* 274, 1926–1928

d Saffran, J.R. *et al.* (1999) Statistical learning of tonal structure by adults and infants. *Cognition* 70, 27–52

e DeCasper, A.J. and Fifer, W.P. (1980) Of human bonding: newborns prefer their mothers' voices. *Science* 208, 1174–1176

f Mehler, J. *et al.* (1988) A precursor of language acquisition in young infants. *Cognition* 29, 144–178

g Busnel, M-C. *et al.* (1992) Fetal audition. *Ann. New York Acad. Sci.* 662, 118–134

h Lecanuet, J-P. *et al.* (1992) Decelerative cardiac responsiveness to acoustical stimulation in the near term foetus. *Q. J. Exp. Psychol.* 44B, 279–303

i DeCasper, A.J. and Spence, M.J. (1986) Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behav. Dev.* 9, 133–150

j DeCasper, A.J. *et al.* (1994) Fetal reactions to recurrent maternal speech. *Infant Behav. Dev.* 17, 159–164

k Jusczyk, P.W. *et al.* (1992) Perception of acoustic correlates of major phrasal units by young infants. *Cognit. Psychol.* 24, 252–293

l Jusczyk, P.W. *et al.* (1993) Preference for the predominant stress patterns of English words. *Child Dev.* 64, 675–687

m Jusczyk, P.W. *et al.* (1993) Infants' sensitivity to the sound patterns of native language words. *J. Mem. Lang.* 32, 402–420

n Jusczyk, P.W. *et al.* (1994) Infants' sensitivity to phonotactic patterns in the native language. *J. Mem. Lang.* 33, 630–645

test precise characteristics of learning. Knowing what infants can learn should, in turn, lead to more specific hypotheses about the actual mechanisms involved. Training infants on artificial languages also controls for prior learning. This latter feature is important because there is every reason to believe that learning begins even in the womb (Box 1), and such prior learning potentially affects all studies in which infants are tested on properties of their target language (Box 2). Artificial-language research has been conducted for many years with adults[7–9], but only recently with infants[10–14]. These new studies have led to exciting discoveries regarding the learning mechanisms available during infancy.

Infant language researchers have begun by examining four aspects of the language learner's task. The first involves identification of word-like units in speech. The second involves encoding and remembering the order in which words occur in sentences. The third involves generalization of grammatical relations. The last involves learning at the more abstract level of syntactic categories (e.g. determiner, adjective, noun and verb). This fourth sensitivity is at the root of our unique human ability to produce and comprehend novel utterances.

Word segmentation

A problem encountered by all language learners is the identification of words in running speech. This would be easy if words were consistently demarcated by physical cues as they are in written text, but they are not. The difficulty of this task is made all the more salient by recalling what it is like to listen to a completely unfamiliar language. For most of us, the words all seem to run together.

Although there are many plausible candidates for cues to when a series of syllables forms a word[15], a reasonably

Box 2. Nativist and empiricist views of language acquisition

A currently popular view of early language acquisition takes as fundamental the assumption that human infants are born with a language-specific learning device (Ref. a). According to this, the nativist view of language learning, certain aspects of the formal structure of language (thought to be absolute and universal) are genetically specified so that acquiring one's target language is tantamount to fine-tuning language-specific parameters (Ref. b). Central to this view is the argument from the poverty of the stimulus, that linguistic input is too impoverished and learning mechanisms too weak to otherwise explain how young children converge on language (where convergence on a universal grammar is thought to underly linguistic productivity). Convergence on the formal structure of language then, is explained by assuming that certain aspects of linguistic knowledge are given in the form of constraints on the learner. For example, as noted in the introduction, it is assumed that children never consider rules based solely on linear order, such as 'move the first verb to the front of the sentence', because they never erroneously transform statements like '*The man who is tall is Sam*' into ungrammatical questions like '*Is the man who tall is Sam?*'. Given that children hear many simple instances that might lead them to form a rule based on linear order (e.g. '*John is tall. Is John tall?*'), how do we explain the lack of errors in sentences with a subordinate clause (and hence two verbs)? Given the conflicting evidence available in the environment, a classic answer is to assume children are innately constrained to consider the hierarchic organization (or structure dependence) of syntactic phrases as opposed to linear word ordering. Arguments such as these gain considerable momentum from Gold's proof showing that certain classes of languages most like human language are not learnable without some kind of constraint on the hypotheses learners are willing to entertain (Ref. c).

An alternative, empiricist view sees the learner as a blank slate, equipped with general associative learning mechanisms (Refs d,e). According to this view, learning might be constrained by human information-processing abilities, but is not limited to the specific domain of language. Although the nativist view has dominated for many years, recent advances in cognitive science suggest that the assumptions underlying this view might have been overly restrictive (Refs d,e). First, far from being impoverished, the language children hear is rich in statistical regularities (Refs f–h). Such regularities aid learning in humans (Refs i–m) and in neural networks. Second, neural networks have far outstripped early conceptions (Ref. n) of associative learning, especially with respect to their ability for capturing key aspects of linguistic behavior (Refs o–q). Where these models differ from nativist proposals is in their emphasis on learning as a stochastic process over distributed input rather than one involving manipulation of discrete symbols. For example, Rohde and Plaut have recently demonstrated how such an architecture learns without explicit negative feedback (Ref. o). The fact that human infants also capitalize on statistical regularities (Refs m,r) suggests a certain degree of overlap in at least some of the mechanisms involved. Researchers have also begun to argue that although Gold's proof applies under the assumption that all learners converge on one true target, it does not apply under the assumption that the target is stochastically defined (Ref. o). Thus, although the problem of the poverty of the stimulus remains an impervious logical dilemma as framed, increasing evidence suggests that the empirical problem faced by children is not so impenetrable.

It is important to point out that although these views differ in their approach to the language problem, each side acknowledges the contribution of the other. Rather, the difference is one of emphasis in terms of which aspects of language are acquired. For example, we know there is a genetic component to language, as demonstrated by the tendency for certain language impairments to be inherited (Ref. s). There is also evidence suggesting that the timing of certain linguistic milestones could be partially genetically determined (J. Ganger, PhD thesis, MIT, 1998). Although the extent to which associative learning factors into linguistic productivity remains to be seen, certain tasks, such as word learning and acquisition of language-specific sound patterns, are clearly dependent on experience. Thus, the challenge for researchers investigating the learning abilities of young infants will be to sort out the contributions of innate and environmental factors, as well as the dynamic between the two.

The real contribution of infant artificial-grammar learning lies in the potential precision it brings to the investigation of early learning mechanisms. Theorists of the nativist persuasion have traditionally studied acquisition in the context of natural language. A drawback with respect to investigating learning mechanisms in this context is the inability to control for prior learning. Thus, is it not obvious whether sensitivities observed with natural language are the result of constraints or a gradual accumulation of knowledge. By contrast, theorists of the empiricist persuasion traditionally investigate learning by constructing models and observing the degree of match between modeled and human behavior. Until recently, however, a limitation has been the inability to verify such learning with age-appropriate learners. Thus, the ability to acquire more specific data should aid in circumventing limitations in both approaches.

Certain issues, involving structure dependence (and other problems of this ilk), are admitted challenges for learning theorists. However, it is possible that current conceptions of these problems are driven more by theory than by data. For instance, a tenet of the view that children are constrained by Universal Grammar is that once a parameter is set for a particular rule, that rule will be applied across a wide range of category instances. However, fine-grained examination of children's utterances suggests that instead of generalizing widely, children first use a small number of lexically specific constructions (Refs t–v). Additionally, the constructions they use tend to be the ones most frequent in their mothers' speech. Thus, children's utterances reflect knowledge of a much more limited scope than would be predicted by current theory, raising questions regarding certain assumptions of that theory.

The history of science abounds with examples of theories that were internally consistent, but which were either abandoned or radically changed because of evidence to the contrary. The hope is that novel methodologies, such as the one outlined here, will result in more precise hypotheses and ultimately more exacting theories as to the mechanisms involved in early language acquisition.

consistent cue is that syllables within words usually have higher transitional probabilities than syllables spanning words (a 'transitional probability', the conditional probability of Y given X, is calculated by normalizing the co-occurrence frequency of X and Y by the frequency of X)[12]. For example, in the learner's experience, the likelihood that *by* will follow *ba* in the phrase *pretty baby* is much higher than the likelihood that *ba* will follow *ty*. Why? Many words other than *baby* can follow *pretty* (e.g. *pretty doggie*, *pretty mommy*, *pretty girl*, *pretty flower* or *pretty dolly*).

References

a Chomsky, N. (1965) *Aspects of a Theory of Syntax*, MIT Press

b Wexler, K. and Manzini, R. (1987) Parameters and learnability in binding theory. In *Parameter Setting* (Roeper, T. and Williams, E., eds), pp. 41–76, Foris

c Gold, E.M. (1967) Language identification in the limit. *Inf. Control* 16, 447–474

d Bates, E. and Elman, J. (1996) Learning rediscovered. *Science* 274, 1849–1850

e Elman, J.L. *et al.* (1996) *Rethinking Innateness: A Connectionist Perspective on Development*, MIT Press

f Kelly, M. (1996) The role of phonology in grammatical category assignments. In *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition* (Morgan, J. and Demuth, K., eds), pp. 249–262, Erlbaum

g Morgan, J. *et al.* (1996) Perceptual bases of rudimentary grammatical categories: toward a broader conceptualization of bootstrapping. In *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition* (Morgan, J. and Demuth, K., eds), pp. 263–283, Erlbaum

h Sereno, J. and Jongman, A. (1990) Phonological and form class relations in the lexicon. *J. Psycholinguist. Res.* 19, 387–404

i Billman, D. (1989) Systems of correlations in rule and category learning: use of structured input in learning syntactic categories. *Lang. Cognit. Processes* 4, 127–155

j Frigo, L. and McDonald, J.L. (1998) Properties of phonological markers that affect the acquisition of gender-like subclasses. *J. Mem. Lang.* 39, 218–245

k Morgan, J.L. and Newport, E.L. (1981) The role of constituent structure in the induction of an artificial language. *J. Verbal Learn. Verbal Behav.* 20, 67–85

l Morgan, J.L. *et al.* (1987) Structural packaging in the input to language learning: contributions of prosodic and morphological marking of phrases to the acquisition of language. *Cognit. Psychol.* 19, 498–550

m Saffran, J.R. *et al.* (1996) Statistical learning by eight-month-old infants. *Science* 274, 1926–1928

n Skinner, B.F. (1957) *Verbal Behavior*, Macmillan

o Rohde, D.L.T. and Plaut, D.C. (1999) Language acquisition in the absence of explicit negative evidence: how important is starting small? *Cognition* 72, 67–109

p Elman, J.L. (1990) Finding structure in time. *Cognit. Sci.* 14, 179–211

q Rumelhart, D.E. and McClelland, J.L. (1986) On learning the past tenses of English verbs. In *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (Vol. 2) (Rumelhart, D.E. and McClelland, J.L., eds), pp. 216–271, MIT Press

r Gómez, R.L. and Gerken, L.A. (1999) Artificial grammar learning by one-year-olds leads to specific and abstract knowledge. *Cognition* 70, 109–135

s Tomblin, J.B. (1996) Genetic and environmental contributions to the risk for specific language impairment. In *Toward a Genetics of Language* (Rice, M. ed.) pp. 191–210, Erlbaum

t Rowland, C.R. and Pine, J.M. Subject-auxiliary inversion errors and wh–question acquisition: 'what children do know'. *J. Child Lang.* (in press)

u Pine, J.M. and Lieven, E.V.M. (1997) Slot and frame patterns and the development of the determiner category. *Appl. Psycholinguist.* 18, 123–138

v Pine, J. *et al.* (1998) Comparing different models of the development of the English verb category. *Linguistics* 36, 807–830

Saffran, Aslin, and Newport investigated whether infants could use transitional probabilities to identify words in running speech[12]. In their study, eight-month-old infants listened to two minutes of continuous speech consisting of four tri-syllabic nonsense words strung together in random order (e.g. *bidakupadotigolabubidakutupiropadoti*…). Infants were then tested to see whether they would discriminate two of the familiarized words (e.g. *tupiro* and *golabu*) from two non-words (*dapiku* and *tilado*). Infants' listening preferences for different stimuli were measured using the head-turn preference procedure[16]. (Stimuli in this procedure are presented auditorily from the infant's left or right side. The amount of time the infant orients toward the source of sound is taken as the dependent measure.) Words and non-words were drawn from the same syllable set, but differed in terms of the transitional probabilities between syllable pairs (with words having mean transitional probabilities of 1 and non-words having mean transitional probabilities of 0). The only cue to whether or not a stimulus was a word was the difference in mean transitional probabilities, and so discrimination would demonstrate sensitivity to such probabilities. Infants, in fact, showed differential attention to familiar and unfamiliar syllable combinations, suggesting the presence of a fairly sophisticated statistical learning mechanism. Later studies demonstrated that infants were also sensitive to transitional probabilities over tone sequences, suggesting that this learning mechanism was more general than one dedicated solely to processing linguistic stimuli[13]. Whether infants will go on to treat constituents extracted from speech as lexical items is still open, but it is certainly a question that can be investigated empirically.

Words in sequence

In addition to segmenting words in running speech, learners must also acquire the legal ordering of words in sentences. To determine whether infants could learn 'grammatical' word order, Gómez and Gerken[11] exposed 12-month-olds to a subset of strings produced by one of two grammars (see Fig. 1). Note that although word order is constrained by these grammars, there is still considerable variability in terms of the orderings of words in sentences. For example, in Grammar 1, PEL can occur in first position (PEL-TAM-RUD), second position (VOT-PEL-JIC-RUD-TAM), both second and third position (VOT-PEL-PEL-JIC) or not at all (e.g. VOT-JIC-RUD-TAM). Similarly, JIC occurs after either VOT, PEL or TAM, but its position varies as a function of whether the sentence begins with PEL or VOT, whether PEL occurs after VOT or after TAM, and whether PEL repeats in the string.

After brief exposure to a subset of strings in their training grammar (between 50 and 127 seconds), infants were given a short play break, and then were tested to see if they would discriminate new strings from the two grammars. Importantly, both grammars began and ended with the same words and contained the same vocabulary. They differed, however, in terms of the ordering of word pairs. For instance, the transition TAM-JIC found in Grammar 1 never occurred in Grammar 2. Likewise, VOT-RUD found in Grammar 2 never occurred in Grammar 1. Infants listened longer to new strings from their training grammar than to strings from the other grammar, regardless of which grammar they heard during training. Although the constraints placed on word ordering were the same during training and test, infants were never tested on the exact strings encountered during training, demonstrating that learning was not confined to memory for particular strings, but rather generalized to novel strings with familiar co-occurrence patterns. This learning is all the

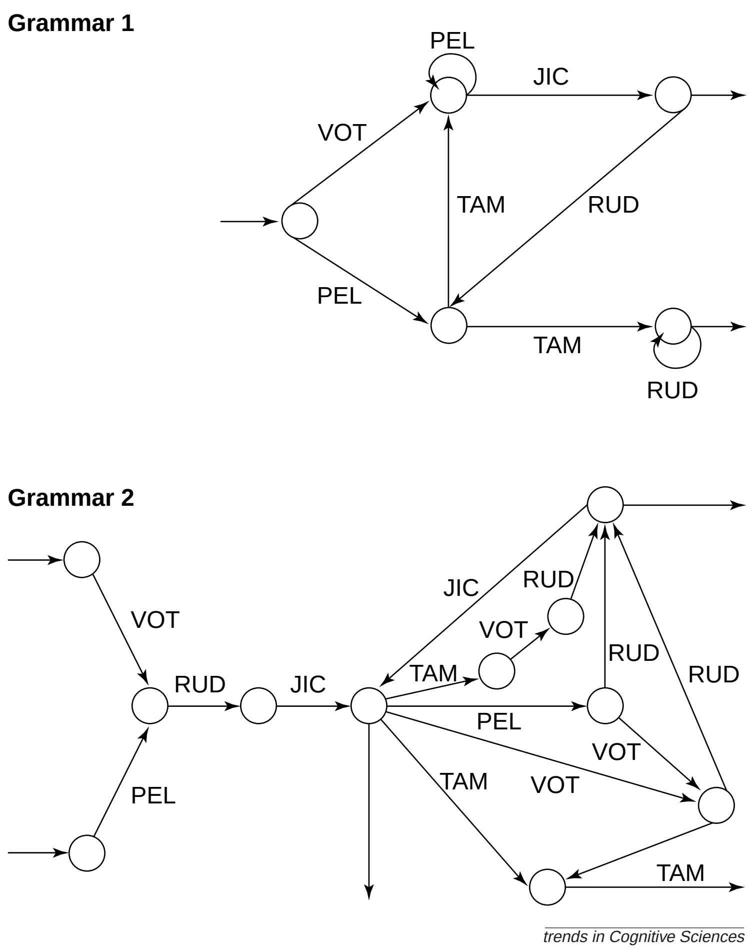


Fig. 1. Artificial Grammars in infant learning. Grammars used in Gómez and Gerken[11]. Grammatical strings are generated by starting at the leftmost position in a grammar and traversing links in the direction of the arrows. Example strings (from Grammars 1 and 2, respectively) are VOT-PEL-JIC-RUD-TAM-RUD and VOT-RUD-JIC-TAM-VOT-RUD. Note that both grammars begin and end with the same words and contain the same vocabulary, but differ in terms of internal word order (e.g. the grammatical transition PEL-JIC in Grammar 1 never occurs in Grammar 2). Learners can be trained on a subset of grammatical strings and then tested to see whether they will generalize to new grammatical strings. Strings can also be instantiated using new lexical items to see whether learners can abstract beyond the ordering of specific word combinations.

more remarkable given that it occurred after less than two minutes exposure and was retained over a short delay.

It is likely that the statistical learning mechanism documented by Saffran and colleagues[12] also explains the learning in these studies. Importantly however, learning is not so static as to prohibit recognition of grammatical word combinations in novel sentences.

Words in abstract patterns

Although sensitivity to word order is necessary for tracking sequential information in sentences, learners must ultimately abstract beyond the ordering of specific words. It is with this aim that researchers have begun investigating early abstraction abilities. For instance, Gómez and Gerken[11] exposed infants to a subset of strings produced by one of the two grammars shown in Fig. 1. Instead of using the vocabulary depicted in the figure, the training set consisted of JED, FIM, TUP, DAK and SOG. The test strings, however, were constructed using the vocabulary VOT, PEL, JIC, RUD and TAM. To give an example, infants trained on Grammar 1 heard strings like FIM-SOG-FIM-FIM-TUP and were tested on new strings like VOT-PEL-PEL-JIC. Thus, although constraints on grammatical word ordering remained constant, vocabulary did not. Critically, because test strings were instantiated in new vocabulary, learners could not distinguish the two grammars based on transitional probabilities between remembered word pairs. This task was all the more difficult because the subset of strings used during training did not overlap with the subset of grammatical strings used at test. That is, none of the *underlying* strings occurred in both training and at test. Infants discriminated grammatical from ungrammatical strings despite the change in vocabulary and despite the fact that none of the underlying test strings were encountered during training, suggesting that they had abstracted some aspect of grammatical structure above and beyond pairs of specific elements. This ability does not appear to be domain specific, at least with respect to adult learners. Adults trained on visually presented consonant and symbol strings generalize to auditorily presented tone and CVC sequences (and vice versa)[17,18]. It remains to be seen whether such learning will prove to be domain general for younger learners.

In a similar series of studies, Marcus and colleagues[14] exposed seven-month-olds to three minute speech samples of strings with ABA (*wi-di-wi* and *de-li-de*) or ABB (*wi-di-di* and *de-li-li*) word patterns. In these studies the underlying pattern was the same for training and for test, however, the vocabulary was different. Infants were subsequently able to discriminate strings with the training pattern from those with a different pattern (e.g. *ba-po-ba* versus *ba-po-po*), despite the change in vocabulary. These results were important for demonstrating that younger infants can also abstract beyond specific word order. Marcus *et al.* further interpreted these findings as evidence that infants are acquiring algebra-like rules (involving substitution of arbitrary elements in abstract variables; an example from language would be the substitution of any plural noun phrase for '*The three daxels*' in the sentence '*The three daxels strolled through the park*')[14]. Marcus has argued that systems sensitive only to statistical regularities (namely connectionist architectures) are, in principle, incapable of such abstraction[19,20]. Arguments against this interpretation (as well as several demonstrations in favor of a statistical learning account of such abstraction) have been mounted by a number of researchers. Thus, although the issue of whether infants are abstracting by means of rules or statistical regularities is still open to debate[21–33], there is no doubt that infants can generalize beyond specific word order. Having demonstrated such abstraction, we must next ask how central it is to acquiring the syntax of one's native language.

Limitations of pattern-based representations

The infant abstraction abilities documented thus far have in common that grammatical and ungrammatical strings were distinguishable by differences in patterns of identical elements (e.g. ABB, ABA, ABCA and ABAAC)[11,14]. No doubt identity is salient for learners. When absent, infants and adults no longer generalize, providing support for the hypothesis that identity underlies this abstraction[34]. Gómez *et al.*[34] exposed learners to a grammar containing strings with one repeating element versus a grammar with no identical elements. Learners acquired robust knowledge of sequential dependencies (as reflected in their ability to discriminate grammatical from ungrammatical strings in their training vocabulary). However, such knowledge did not factor into their ability to generalize to new vocabulary. Abstraction beyond specific word order only occurred for learners trained on the grammar with repeating elements. Such abstraction could be limited, however, with respect to acquiring syntax. The key to understanding this point lies in a contrast between what we will call pattern-based and category-based abstraction.

Pattern-based abstraction can be described in terms of relational operations (e.g. identity, greater-than or less-than) over physical stimuli in sequence. For example, recognizing *ba-po-ba* and *ko-ga-ko* as instances of the pattern ABA entails noting that the first and last syllables in sequence are physically identical. It is perhaps easier to understand this distinction in the context of classic studies from the animal literature. For example, chimpanzees, rats and chickens are able to evaluate relational patterns in training stimuli (e.g. $\text{luminance}_a > \text{luminance}_b$) and generalize to untrained stimuli[35]. Furthermore, starlings trained to respond to ascending tone sequences generalize to new ascending sequences created by various transformations of the training stimuli (where an ascending sequence can be described by a series of relations in which the pitch of sequence element $n+1$ is greater than element n)[36,37]. In each of these examples, a relation is abstracted by comparing the perceptual characteristics of each element in the physical array to those of the other elements. In this way, such relations are perceptually bound.

Category-based generalization, by contrast, involves operations over abstract rather than perceptually bound variables. Compare the pattern-based representation ABA with the category-based representation Noun-Verb-Noun. Although superficially similar, these examples differ along a critical dimension. Recognizing ABA and Noun-Verb-Noun both involve identity, but in the former case, the relation is perceptually bound, whereas in the latter the identity relation holds over abstract categories and thus is at least one step removed from physical identity. That is, abstracting the pattern ABA from *ba-po-ba* involves noting that the first and third elements in a sequence are physically identical. With category-based generalization, however, learners must identify the first and third elements as members of the abstract category 'noun'. These determinations cannot be based on perceptual identity. '*Dogs eat pizza*' and '*John loves books*' share the same category-based structure, despite the obvious physical dissimilarities between category members such as '*John*' and '*books'*. Although abstraction of the ABA pattern could be construed as a relation between the categories 'initial word' and 'final word', successful abstraction still entails identifying the first and last words as identical in form, hence involving pattern-based abstraction.

Although identity plays an important role in other linguistic phenomena (e.g. relations between pronouns and antecedents, and between moved elements and their traces), again, the abstraction required of such phenomena is a step removed from the way in which identity is used in the infant studies. A mature language user presented with '*The bear$_i$ washed himself$_i$*' and shown two pictures, one of a bear washing himself and another of a bear washing another bear, will choose the first picture over the second. Presumably, this ability involves noting an identity relation between the pronoun and its antecedent (indicated in the example by the use of the subscript i). When presented with '*The bear$_i$ washed him$_j$*' the preference is reversed. However, '*himself*' is no more similar to '*bear*' than is '*him*', precluding pattern-based abstraction. Rather, the child must appeal to an abstract referential system in noting the identity relation between the co-indexed pronoun and its antecedent.

	N_1	N_2	N_3
M_1	x	x	?
M_2	?	x	x
M_3	x	?	x

	Q_1	Q_2	Q_3
P_1	x	?	x
P_2	x	x	?
P_3	?	x	x

trends in Cognitive Sciences

Fig. 2. Category-based abstraction. The matrices depict the learning space in studies of category-based abstraction. Learners must acquire dependencies between MN and PQ classes based on a subset of the pairings. For instance, learners might hear M_1 paired with N_1 and N_2, M_2 paired with N_2 and N_3, and M_3 paired with N_1 and N_3. The question is whether they will generalize to the unattested pairings (cells marked with ?). If they have learned the dependencies between categories, then they should accept new grammatical pairings such as M_1N_3, while rejecting ungrammatical ones (e.g. M_1Q_3).

Category-based abstraction

The ability to abstract over categories is fundamental to linguistic productivity. A learner who identifies a novel word as belonging to a particular category has immediate access to all of the rules involving that category[38,39]. Even very young learners are privy to such information. Pre-school children seeing and hearing '*Here is a wug. Now there are two of them*', and asked to complete the sentence '*There are two* ______', respond with the answer '*wugs*'[39].

Category-based abstraction has been of particular interest to researchers investigating language learning mechanisms with older learners[40–44], and has focused on the problem of how learners acquire relations between grammatical classes. For example, English-speaking children need to learn that the determiners *the* and *a* precede nouns and not verbs, whereas auxiliaries like *was* and *is* precede verbs, but not nouns. This problem can be conceptualized in terms of filling in the cells of matrices (such as the ones shown in Fig. 2), where learners must acquire the knowledge that MN and PQ are legal sentences of a language, but MQ and PN are not[44]. In these studies, learners are exposed to most, but not all, grammatical pairings during training to see whether they will generalize to new grammatical pairs at test[44]. If learners acquire the categories M, N, P and Q, and learn the dependencies between them, then they should distinguish a new grammatical pair such as M_1N_3 from the ungrammatical M_1Q_3.

As it happens, these kinds of distributional relations are practically impossible for learners to acquire[40–44]. Although adults learn readily that M- and P-words occur first and N- and Q-words occur second, they do not learn the dependencies between classes. This difficulty is ameliorated, however, when a subset of category members are marked with salient conceptual or perceptual cues[40,43]. For example, Frigo and McDonald augmented a subset of the members from the

Box 3. Systematically related cues and learning

Research on category-based abstraction has been critical for demonstrating the importance of systematically related cues in abstraction of language-like categories (Refs a–c). However, it should come as no surprise that learners rely heavily on correlated cues. Prosodic breaks, function words and concord morphology are more effective at promoting hierarchical packaging of word strings in an artificial grammar when they occur consistently with the hierarchical structure than when they violate that structure (Refs d,e). Learners in an artificial language study were also better at acquiring syntactic rules when they were predicted by systematically related features, than when rules occurred in isolation (rules were indicated by perceptual and conceptual marking of syntactic categories) (Ref. f). This is consistent with findings showing that learners are far more successful at differentiating syntactic categories, and subsequently learning the relationships between them, when some subset of category members are distinguished by perceptual or conceptual cues (Ref. c). Systematicity also factors heavily in concept learning (Ref. g), raising the possibility that certain aspects of language learning (e.g. those dependent on cues to systematic structure), could derive from a general, rather than a domain-specific, acquisition mechanism.

As it turns out, systematically related cues abound in language, providing a potentially rich source of information for learners (Refs h,i). In English, for example, nouns tend to be preceded by frequently occurring determiners sharing a similar vowel sound (e.g. schwa, in *the* and *a*), whereas verbs are preceded by auxiliaries (*was*, *is*) and can also be marked with inflectional endings (*-ed*, *-ing*). Nouns and verbs exhibit different stress patterns, where disyllabic nouns tend to stress the first syllable and disyllabic verbs the second (contrast the noun and verb forms of the word *permit*). Verbs also tend to have fewer syllables than nouns, a cue used even by 4-year-olds (Ref. j). Although such regularities do not predict syntactic structure for every sentence in the English language (Refs k–m), they might occur with enough predictability, or in concert with other types of cues (e.g. semantic), to provide learners with a 'toe-hold' on the acquisition problem, both in terms of breaking the input into smaller, more manageable chunks, and in terms of classifying these chunks into meaningful categories.

A problem for a view emphasizing the role of systematic cues in language acquisition is explaining how learners acquire the cases in direct violation of cues. Another problem (raised by the structure-dependence example used in the introduction) is accounting for performance for which the relevant regularities do not exist. However, an explanation of language acquisition emphasizing the psychological salience of systematically related cues is not meant to supercede all other accounts of how learners acquire language. Furthermore, there is no reason why mechanisms sensitive to statistical regularities cannot act in combination with other sources of information (Ref. n). Rather, the challenge for language researchers is in determining the limits of such learning with respect to acquiring language.

References

a Braine, M.D.S. (1987) What is learned in acquiring word classes: a step toward an acquisition theory. In *Mechanisms of Language Acquisition* (MacWhinney, B., ed.), pp. 65–87, Erlbaum

b Brooks, P.J. *et al.* (1993) Acquisition of gender-like noun subclasses in an artificial language: the contribution of phonological markers to learning. *J. Mem. Lang.* 32, 76–95

c Frigo, L. and McDonald, J.L. (1998) Properties of phonological markers that affect the acquisition of gender-like subclasses. *J. Mem. Lang.* 39, 218–245

d Morgan, J.L. and Newport, E.L. (1981) The role of constituent structure in the induction of an artificial language. *J. Verbal Learn. Verbal Behav.* 20, 67–85

e Morgan, J.L. *et al.* (1987) Structural packaging in the input to language learning: contributions of prosodic and morphological marking of phrases to the acquisition of language. *Cognit. Psychol.* 19, 498–550

f Billman, D. (1989) Systems of correlations in rule and category learning: use of structured input in learning syntactic categories. *Lang. Cognit. Processes* 4, 127–155

g Billman, D. and Knutsen, J. (1996) Unsupervised concept learning and value systematicity: a complex whole aids learning the parts. *J. Exp. Psychol. Learn. Mem. Cognit.* 22, 458–475

h Kelly, M.H. (1992) Using sound to solve syntactic problems: the role of phonology in grammatical category assignments. *Psychol. Rev.* 99, 349–364

i Kelly, M. (1996) The role of phonology in grammatical category assignments. In *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition* (Morgan, J. and Demuth, K., eds), pp. 249–262, Erlbaum

j Cassidy, K.W. and Kelly, M.H. (1991) Phonological information for grammatical category assignments. *J. Mem. Lang.* 30, 348–369

k Fernald, A. and McRoberts, G. (1996) Prosodic bootstrapping: a critical analysis of the argument and the evidence. In *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition* (Morgan, J. and Demuth, K., eds), pp. 365–388, Erlbaum

l Gerken, L.A. *et al.* (1994) When prosody fails to cue syntactic structure: 9-month-olds' sensitivity to phonological versus syntactic phrases. *Cognition* 51, 237–265

m Pinker, S. (1984) *Language Learnability and Language Development,* MIT Press

n Jusczyk, P.W. (1997) *The Discovery of Spoken Language*, MIT Press

N and Q categories with salient word beginnings and endings (e.g. members of N began with *kais* and ended with *rish* as in *kaisemilrish*, whereas Q members began with *wan* and ended with *glot* as in *wanersumglot*)[43]. Once categories are distinguishable, learners can accurately induce category relationships, even for the members not marked with cues to class membership (e.g. *roosa* or *faranu*). Presumably, learners group N- and Q-words into different categories based on their distinguishing features. These features are then used to group M- and P-elements. Once categorized, learners can use the knowledge that M_1 pairs with N_1 to infer that it also pairs with N_2 (Ref. 40).

The results of these studies are instructive in more than one way. First, they demonstrate that humans are not unconstrained learners. People simply do not abstract arbitrary dependencies. Abstraction results only when there is sufficient evidence to distinguish the categories in question. This should not be surprising given previous work on the importance of correlated cues in language learning (Box 3). However, this fact about human learning is also important for circumscribing the nature of the acquisition mechanisms proposed[40]. For example, overly powerful models have assumed that learning involves abstraction of arbitrary structure[3,45], when in fact many of the categories found in natural language (such as gender, declension and conjugation classes) are rich in systematic cues to class membership[46]. For example, in Spanish, feminine nouns often end in -a and masculine nouns in -o. In Hebrew, nouns ending in -a and -t are often feminine.

Now that we have some understanding of the requirements for inducing category-based abstraction in adults, the next step will be to begin investigating how younger learners master such abstraction. Other important issues have to do with whether such learning is rule based or associational in nature and whether learners can induce category structure based on a more limited set of examples (e.g. exhibiting characteristics found in child-directed speech). We are currently investigating these issues in our joint laboratories with infants and adults.

Outstanding questions

- Which aspects of language acquisition are acquired and which are innate? How much knowledge is built into the initial system? What is the dynamic between innate and environmental factors? How does this dynamic change over the course of development?
- Do learning mechanisms develop?
- How do domain-general and domain-specific mechanisms factor into language learning?
- Do these mechanisms operate exclusively by means of rules or associations, or do we make use of both symbolic and associative mechanisms?
- To what extent are the mechanisms identified in artificial-language studies the same as those used in acquiring natural language?

Conclusion and implications

We have reviewed a number of studies investigating the learning abilities exhibited by infants and adults. The results suggest that infants are equipped with remarkable abilities for parsing linguistic input. They are able to identify word-like constituents in fluent speech based on predictive syllable relationships[12]. They learn constraints on grammatical word order[11]. They also exhibit rudimentary abstraction abilities, as reflected in their recognition of familiar patterns in novel vocabulary[11,14]. Finally, they must ultimately discover that the ordering of words in sentences is determined at a more abstract level by dependencies among syntactic categories. We have some understanding of how adults acquire such dependencies[40–44], motivating similar research with younger learners.

How does our growing understanding of infant learning abilities bear on the highly constrained language learner described in the introduction? We can identify at least three ways. First, all of the artificial-language-learning studies discussed have examined infants' sensitivity to linguistic form in the absence of semantic content. In so far as these studies are tapping sensitivities used in real-language acquisition, they challenge many accounts in which language development is driven by a mapping between meaning and form[3,47,48]. This is not to say that learners do not ultimately need to map the syntactic forms they encode during infancy onto meaning. Obviously they do. However, the fact that infants are able to acquire certain aspects of form prior to acquiring the meaning of these forms changes the nature of the language acquisition problem in a fundamental way.

A second implication of the research on infant artificial-language learning concerns the specificity of the constraints on the learner. On many accounts, these constraints have been construed as being language specific, such that for every aspect of language to be acquired, the child is born with a specific constraint or parameter that guides him/her to the correct representation[49,50]. Data showing that infants can use transitional probabilities to segment grammatical tone sequences contrasts with this view, suggesting that they apply statistical learning to linguistic and non-linguistic stimuli alike[13]. The application of statistical sensitivity to the problem of word segmentation is admittedly far from the constraints discussed by linguistic nativists (involving such language-specific notions as whether or not declarative sentences in a particular language must have an overt subject)[51]. Nevertheless, the hypothesis that language (although a specialized human cognitive domain), can be acquired via general-purpose learning mechanisms, is one likely to be investigated with increasing vigor over the next decade.

A third implication of both the infant artificial-language-learning studies reviewed here and the myriad studies of infant language perception preceding them concerns the relevance of children's early utterances as evidence for theories of language acquisition. One of the key observations of linguistic nativists involves errors that children do not make. As noted earlier, children never erroneously transform a statement like '*The man who is tall is Sam*' into a question like '*Is the man who tall is Sam?*' The lack of such errors, along with logical arguments concerning the poverty of the stimulus, have been taken as evidence that children never consider rules based solely on linear order in sentences. Although researchers have begun to address the question of a how a statistical learner might begin to negotiate impoverished input (Box 2), it is equally important to note that if the studies of infants' early linguistic abilities tell us anything, it is that they have become sensitive to many aspects of linguistic form a year or more before they ever begin to produce multiword speech. This is not to say that all of language is acquired by the age of 12 months. However, if infant language-perception studies have one theme, it is in demonstrating the extremely complex (and often contrasting) relationship between aspects of their native language infants and young children have tacitly discerned and those they actually produce. Thus, we must exercise caution in interpreting children's early utterances as evidence for or against the linguistic representations they do and do not entertain.

A final comment is in order. Given the vast differences in artificial grammars and natural language, how do we ensure that the learning observed is representative of language learning in the real world? First, in using this approach it is important to design experiments capturing key linguistic phenomena. If we can isolate a phenomenon of interest experimentally, we can go on to test it using a wide range of manipulations, where, presumably, such manipulations are driven by our knowledge of natural language acquisition. For instance, the finding that 18-month-olds, but not 15-month-olds track grammatical dependencies separated by one to three intervening syllables, suggests that we should see the same pattern with an artificial grammar designed to investigate such learning[52]. Indeed, studies in our joint laboratories show that we do. Another approach, currently being investigated by Saffran and colleagues, is to test whether the output of statistical learning can be used as input to natural language. Ultimately, however, as with any scientific endeavor, the proof of this approach will depend on the extent to which it generates new ways of understanding the mechanisms involved in natural language acquisition. Its real promise lies in the precision it affords with respect to investigating infant learning.

Acknowledgements

The preparation of this manuscript was funded by research grants from NSF to RLG (SES-9910203) and LAG (SBR-9696072). We thank R. Schvaneveldt, D. Plaut and the anonymous reviewers for helpful comments and suggestions.

References

1 Chomsky, N. (1965) *Aspects of a Theory of Syntax*, MIT Press
2 Gleitman, L.R. and Wanner, E. (1982) Language acquisition: the state of the art. In *Language Acquisition: The State of the Art* (Wanner, E. and Gleitman, L.R., eds), pp. 3–50, Cambridge University Press
3 Pinker, S. (1984) *Language Learnability and Language Development*, MIT Press
4 Morgan, J.L. and Demuth, K., eds (1996) *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*, Erlbaum
5 Hirsh-Pasek, K. *et al.* (1987) Clauses are perceptual units for young infants. *Cognition* 26, 269–286
6 Jusczyk, P.W. *et al.* (1992) Perception of acoustic correlates of major phrasal units by young infants. *Cognit. Psychol.* 24, 252–293
7 Esper, E.A. (1925) A technique for the experimental investigation of associative interference in artificial linguistic material. *Lang. Monogr.* 1
8 Braine, M.D.S. (1963) On learning the grammatical order of words. *Psychol. Rev.* 70, 323–348
9 Reber, A.S. (1967) Implicit learning of artificial grammars. *J. Verbal Learn. Verbal Behav.* 77, 317–327
10 Aslin, R.N. *et al.* (1998) Computation of conditional probability statistics by 8-month-olds infants. *Psychol. Sci.* 9, 321–324
11 Gómez, R.L. and Gerken, L.A. (1999) Artificial grammar learning by one-year-olds leads to specific and abstract knowledge. *Cognition* 70, 109–135
12 Saffran, J.R. *et al.* (1996) Statistical learning by eight-month-old infants. *Science* 274, 1926–1928
13 Saffran, J.R. *et al.* (1999) Statistical learning of tonal structure by adults and infants. *Cognition* 70, 27–52
14 Marcus, G. *et al.* (1999) Rule learning by seven-month-old infants. *Science* 283, 77–80
15 Jusczyk, P.W. (1999) How infants begin to extract words from speech. *Trends Cognit. Sci.* 3, 323–328
16 Kemler Nelson, D. *et al.* (1995) The head-turn preference procedure for testing auditory perception. *Infant Behav. Dev.* 18, 111–116
17 Altmann, G.T.M. *et al.* (1995) Modality independence of implicitly learned grammatical knowledge. *J. Exp. Psychol. Learn. Mem. Cognit.* 21, 899–912
18 Tunney, R.J. and Altmann, G.T.M. (1999) The transfer effect in artificial grammar learning: re-appraising the evidence on the transfer of sequential dependencies. *J. Exp. Psychol. Learn. Mem. Cognit.* 25, 1322–1333
19 Marcus, G.G. (1998) Can connectionism save constructivism? *Cognition* 66, 153–182
20 Marcus, G.F. (1998) Rethinking eliminative connectionism. *Cognit. Psychol.* 37, 243–282
21 Altmann, G.T.M. and Dienes, Z. (1999) Rule learning by seven-month-old infants and neural networks. *Science* 284, 875a
22 Christiansen, M.H. and Curtin, S.L. (1999) Transfer of learning: rule acquisition or statistical learning? *Trends Cognit. Sci.* 3, 289–290
23 Eimas, P. (1999) Do infants learn grammar with algebra or statistics? *Science* 284, 435–436
24 Marcus, G.F. (1999) Connectionism: with or without rules? *Trends Cognit. Sci.* 3, 168–170
25 Marcus, G.F. (1999) Do infants learn grammar with algebra or statistics? *Science* 284, 436–437
26 Marcus, G.F. (1999) Reply to Christiansen and Curtin. *Trends Cognit. Sci.* 3, 290–291
27 Marcus, G.F. (1999) Reply to Seidenberg and Elman. *Trends Cognit. Sci.* 3, 289
28 Marcus, G.F. (1999) Rule learning by seven-month-old infants and neural networks: response to Altmann and Dienes. *Science* 284, 875a
29 McClelland, J.L. and Plaut, D.C. (1999) Does generalization in infant learning implicate abstract algebra-like rules? *Trends Cognit. Sci.* 3, 166–168
30 Negishi, M. (1999) Do infants learn grammar with algebra or statistics? *Science* 284, 435
31 Seidenberg, M.S. and Elman, J.L. (1999) Do infants learn grammar with algebra or statistics? *Science* 284, 435–436
32 Seidenberg, M.S. and Elman, J.L. (1999) Networks are not 'hidden rules'. *Trends Cognit. Sci.* 3, 288–289
33 Shastri, L. (1999) Infants learning algebraic rules. *Science* 285, 1673–1674
34 Gómez, R.L. *et al.* (2000) The basis of transfer in artificial grammar learning. *Mem. Cognit.* 28, 253–263
35 Kohler, W. (1925) Simple structural functions in the chimpanzee and the chicken. In *A Sourcebook of Gestalt Psychology* (Ellis, W.D. ed.), pp. 217–227, Harcourt Brace.
36 Hulse, S.H. and Cynx, J. (1986) Interval and contour in serial pitch perception by a passerine bird, the European starling (*Sturnus vulgaris*). *J. Comp. Psychol.* 3, 215–228
37 Hulse, S.H. *et al.* (1984) Discrimination and generalization of rhythmic and arrhythmic sound patterns by European starlings (*Sturnus vulgaris*). *Music Percept.* 1, 442–464
38 Maratsos, M.P. (1982) The child's construction of grammatical categories. In *Language Acquisition: The State of the Art* (Wanner, E. and Gleitman, L.R., eds), pp. 240–266, Cambridge University Press
39 Berko, J. (1958) The child's learning of English morphology. *Word* 14, 150–177
40 Braine, M.D.S. (1987) What is learned in acquiring word classes: a step toward an acquisition theory. In *Mechanisms of Language Acquisition* (MacWhinney, B., ed.), pp. 65–87, Erlbaum
41 Braine, M.D.S. *et al.* (1990) Exploring language acquisition in children with a miniature artificial language: effects of item and pattern frequency, arbitrary subclasses and correction. *J. Mem. Lang.* 29, 591–610
42 Brooks, P. *et al.* (1993) Acquisition of gender-like noun subclasses in an artificial language: the contribution of phonological markers to learning. *J. Mem. Lang.* 32, 76–95
43 Frigo, L. and McDonald, J.L. (1998) Properties of phonological markers that affect the acquisition of gender-like subclasses. *J. Mem. Lang.* 39, 218–245
44 Smith, K.H. (1969) Learning co-occurrence restrictions: rule learning or rote learning? *J. Verbal Learn. Verbal Behav.* 8, 319–321
45 Anderson, J.R. (1983) *The Architecture of Cognition*, Harvard University Press
46 Zubin, D.S. and Kopcke, K.M. (1981) Gender: a less than arbitrary grammatical category. In *Papers from the Seventh Regional Meeting of the Chicago Linguistics Society* (Hendrick, R.A. *et al.*, eds), pp. 439–449, University of Chicago Press
47 Bates, E. (1976) *Language and Context: Studies in the Acquisition of Pragmatics*, Academic Press
48 Bowerman, M. (1973) *Early Syntactic Development: A Cross-linguistic Study with Special Reference to Finnish*, Cambridge University Press
49 Chomsky, N. (1980) *Rules and Representations*, Foris
50 Wexler, K. and Manzini, R. (1987) Parameters and learnability in binding theory. In *Parameter Setting* (Roeper, T. and Williams, E., eds), pp. 41–76, Foris
51 Hyams, N.M. (1986) *Language Acquisition and the Theory of Parameters*, Reidel
52 Santelmann, L.M. and Jusczyk, P.W. (1998) Sensitivity to discontinuous dependencies in language learners: evidence for limitations in processing space. *Cognition* 69, 105–134

Erratum

In the Opinion article by M. Tomasello in the April issue of *Trends in Cognitive Sciences* (Vol. 4, No. 4, pp. 156–163), Table 1 on p. 160 was printed with two errors. In the left-hand column of the table, instead of 'Ref. 43' and 'Ref. 44' it should read 'Lewis and Tomasello (unpublished data)' and 'Childers and Tomasello (unpublished data)', respectively.

We apologize to the author and to readers for this oversight.

PII: S1364-6613(00)01481-9