

and composition of the metal ions of the A-cluster. The presence of copper in the CODH/ACS complex eluded investigators previously, although it was reported in the early characterization of the CODH/ACS complex from a methanogenic archaeobacteria (12).

The presence of copper is surprising given its very low solubility under anaerobic conditions in the presence of sulfide. Although there are a number of examples of copper-containing proteins, the presence of copper in this complex enzyme represents a new role for the metal in biological systems. The occurrence of both copper and nickel at a single active site

has not been reported previously and may reflect the complexity of the biochemical reaction of acetyl-CoA condensation and disassembly.

CODH/ACS is believed to be an ancient enzyme. The ACS reaction has been considered as the primordial initiation reaction for a chemoautotrophic origin of life (13). Recent experiments aimed at mimicking prebiotic conditions indicate that activated acetate can be generated from CO₂ in the presence of a slurry containing NiS and FeS. A role for copper in these prebiotic scenarios for the synthesis of activated acetate will without doubt be the subject of future studies.

References

1. H. G. Wood, *FASEB J.* **5**, 156 (1991).
2. S. W. Ragsdale, M. Kumar, *Chem. Rev.* **96**, 2515 (1996).
3. P. A. Lindahl, *Biochemistry* **41**, 2097 (2002).
4. T. I. Doukov, T. M. Iverson, J. Seravalli, S. W. Ragsdale, C. L. Drennan, *Science* **298**, 567 (2002).
5. H. Dobbek, V. Svetlitchnyi, L. Gremer, R. Huber, O. Meyer, *Science* **293**, 1281 (2001).
6. C. L. Drennan, J. Heo, M. D. Sintchak, E. Schreiber, P. W. Ludden, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 11973 (2001).
7. J. G. Ferry, *Annu. Rev. Microbiol.* **49**, 305 (1995).
8. J. B. Thoden, H. M. Holden, G. Wesenberg, F. M. Raushel, I. Rayment, *Biochemistry* **36**, 6305 (1997).
9. H. Beinert, R. H. Holm, E. Munck, *Science* **277**, 653 (1997).
10. D. C. Rees, *Annu. Rev. Biochem.* **71**, 221 (2002).
11. U. Ermiler, W. Grabarse, S. Shima, M. Goubeaud, R. K. Thauer, *Science* **278**, 1457 (1997).
12. J. A. Krzycki, L. E. Mortenson, R. C. Prince, *J. Biol. Chem.* **264**, 7217 (1989).
13. C. Huber, G. Wächtershäuser, *Science* **276**, 245 (1997).

PERSPECTIVES: NEUROSCIENCE

Does Grammar Start Where Statistics Stop?

Mark S. Seidenberg, Maryellen C. MacDonald, Jenny R. Saffran

Languages exhibit statistical structure—that is, they show inhomogeneities in the distribution of sounds, words, and phrases. The importance of this type of structure in learning a language is a matter of intense debate, and is tackled by Peña *et al.* (1) on page 604 of this issue. The debate originated with Chomsky's 1957 discussion of the sentence "Colorless green ideas sleep furiously" (2). This sentence can be immediately recognized as being well formed (compared to "Ideas colorless sleep furiously green") even though in both cases the probability that these words have previously occurred in this order is close to zero. Chomsky concluded that the statistical properties of language are not central to the characterization of linguistic knowledge, an insight that became part of the foundation of modern linguistic theory. Whether statistical properties are important in language acquisition was largely set aside. Instead, research focused on how the child converges on the rules and other components of grammar using a combination of deductive (nonstatistical) reasoning and innate knowledge (3).

Recently, there has been a resurgence of interest in statistical learning, with evidence showing that infants and young children incorporate statistical cues when learning about the sounds of a language, vocabulary, and the structures in which words occur (4–6). These findings complement evidence

from adults demonstrating the use of statistical information in comprehending and producing utterances, suggesting that similar mechanisms may underlie the learning and use of language (7, 8).

Although this research establishes that statistical information is used in language acquisition, the extent to which acquisition can be explained in these terms is not yet known. Peña *et al.* (1) suggest one possibility: Perhaps there are both statistical processes (based on frequency and distribution of elements in language) and grammatical processes (for example, learning and using rules). Statistical learning may be limited to simpler problems such as learning the sounds of a language and building a lexicon. In contrast, the complexities of grammar may require other nonstatistical procedures. Thus, it seems that learning grammar begins where statistical learning ends.

This reconcilist view is appealing because it preserves the main tenets of the grammar approach while apparently accommodating evidence about statistical learning. In practice, however, it turns out to be difficult to establish a boundary between "grammatical" and "statistical" learning. Any corpus of linguistic stimuli contains a vast array of cues and potential generalizations. Even in carefully designed experiments, conditions intended to isolate grammatical processes may introduce correlated statistical cues that would support performance. For example, in the Peña *et al.* study, adults listened to a continuous stream of nonsense words (see the table). According to the authors, the subjects could extract statistical regularities from the speech

stream, but they could formulate rules only when brief pauses were added at word boundaries. Although the language supplied to subjects by Peña *et al.* consisted of only nine words, the corpus derived by concatenating these words afforded a large number of generalizations about the syllable sequences (some of which are shown in the table). Peña *et al.*'s conclusions about grammatical learning concentrated on some properties of the syllable sequences, but other properties could also have cued subjects' responses.

Two previous attempts to isolate a distinct grammatical form of learning (9, 10) raised similar concerns. In each case subsequent analyses suggested that the behavior could instead have arisen from statistical regularities that occurred simultaneously with the grammatical patterns (11–13). Importantly, these additional findings (like the analysis presented in the table) do not show that grammatical learning does not exist, but rather that statistical learning could also account for the results. Such findings also illustrate the difficulty of working back from observed behavior to the underlying regularities that gave rise to it (14). Knowing how many distinct procedures are involved in learning a language is a critical issue; resolving it will require advances on both the "statistical" and "grammatical" fronts.

Discussions of statistical learning need to consider two questions illustrated by the "colorless" sentence. First, what kinds of statistics are people, particularly infants, capable of computing? As in the "colorless" example, most research has investigated the transition probabilities between words or syllables. The Peña *et al.* study is a welcome step forward insofar as it addresses questions concerning nonadjacent elements (15). Adult learners can track various types of statistics, including some second-order probabilities and long-distance dependencies (14, 16), but the

The authors are in the Department of Psychology, University of Wisconsin, Madison, WI 53706, USA. E-mail: marks@lcnl.wisc.edu

RULES OR STATISTICS IN LANGUAGE LEARNING?

RULES OR STATISTICS IN LANGUAGE LEARNING?		
A Familiarization stream		
... PURAKIBELIGATAFODUPUFOKITALIDUBERAGATARADUPULIKIBEFOGA ...		
B Generalizations		Probability
1.	Words have PU_KI, BE_GA, TA_DU structure (Peña <i>et al.</i> 's rule, AXC)	1.0
2.	Initial syllables begin with a stop consonant	1.0
3.	Final syllables begin with a stop consonant	1.0
4.	Continuant consonants occur word medially	1.0
5.	PU predicts RA	0.33
6.	PU is not followed by BE	1.0
7.	KI predicts BE	0.5
8.	PURA predicts KI	1.0
9.	RAKI predicts BE	0.5
10.	PUBE is not followed by KI	1.0
C Experiment 1 forced choice		
	Choice supported by	Percentage chosen
Word	PURAKI 1 (rule), 2, 3, 4, 5, 8	73.3
Part-word	RAKIBE 3, 7, 9	26.7
D Experiment 2 forced choice		
	Choice supported by	Percentage chosen
Rule-word	PUBEKI 1 (rule), 2, 3, 4, 8	49.8
Part-word	RAKIBE 3, 6, 7, 9, 10	49.2

(A) A section of the speech stream from the Peña *et al.* study (7). The three word families are illustrated in different colors. (B) Ten of the generalizations available from the input in panel A, with their probabilities of occurrence in the familiarization stream. (C) The forced choice alternatives in Peña *et al.*'s experiment 1. The word alternative is favored over the part-word (a sequence that spans a word boundary) by Peña *et al.*'s AXC rule (syllable A, followed by syllable C, with an intervening syllable X) as well as by more of the generalizations from panel B. (D) The forced choice alternatives in experiment 2. The rule-word (a new word generated by the AXC rule) and part-word alternatives were chosen equally often. In subsequent studies, introduction of brief pauses between words in the familiarization phase increased choice of rule-words over part-words, because, according to Peña *et al.*, the pauses switched subjects into a rule-learning mode. The pauses also simplify the word-segmentation task, increase the salience of properties 2-4, and make the part-words less like the familiarization words.

limits on these capacities and whether infants have similar capacities have not been determined. It is also unclear how well such learning mechanisms fit the demands posed by human languages. However, recent results suggest that statistical patterns that occur in natural languages are acquired more readily than patterns not found in natural languages (16–18). Thus, constraints on learning may play a positive role in explaining why language learners acquire only some of the many generalizations afforded by natural language.

The second question asks: Over what types of information are statistics learned? Chomsky's example assumes that people are computing statistics at a single level of linguistic structure—between words in the “colorless” example, and between syllables in the Peña *et al.* study. But language exhibits structure at multiple levels, each of which has its own statistical character. The “colorless” sentence is less puzzling when one looks beyond transition probabilities to

other information that is used in comprehension. For example, words fall into general types—*green* is a property or adjective, and *sleep* is an action or verb—that exhibit characteristic distributions. The “colorless” sentence conforms to these distributions in English, whereas “Ideas colorless sleep furiously green” does not (19).

Our understanding of the contribution of statistical learning is limited by incomplete knowledge of the kinds of statistics infants encode and whether these are the ones relevant to natural language. This view also leaves open the critical question of why only humans acquire language, as many other species are capable of simple forms of statistical learning. Again, there are parallel statistical and grammatical interpretations to explain this fact. For example, the statistics of natural language, which involve correlations over multiple types of information simultaneously, may be too complex for other species to learn; alternatively, other species may lack innate

grammatical capacities that make language learning possible.

Studies centered on understanding rule learning raise two major unresolved issues. First, the distinction between a rule and a statistical generalization remains unclear. Many of the regularities summarized in the table could be called either rules or statistics. Second, how do infants actually find rules in the speech they hear? Is there a procedure that would yield the right rules under realistic learning conditions? The evidence for rule learning is mostly negative: cases where learning occurs but there is no obvious statistical explanation. A theory explaining how rule learners arrive at exactly the correct generalizations given the complexities of their experience would represent substantial progress.

The cascade of potential learning cues and generalizations implicit in the miniature language studied by Peña *et al.* underscores the difficulties in determining how learners acquire the vastly richer structure of natural languages. To some extent this problem may be solved by grammar-specific forms of knowledge or learning. Statistical learning offers another potential explanation insofar as languages may exhibit only those structures that learners are able to track. Thus, the structure of language may have resulted in part from constraints imposed by the limits of human learning.

References and Notes

1. M. Peña *et al.*, *Science* **298**, 604 (2002).
2. N. Chomsky, *Syntactic Structures* (Mouton, the Hague, 1957).
3. S. Crain, R. Thornton, *Investigations in Universal Grammar* (MIT Press, Cambridge, MA, 1998).
4. R. L. Gómez, L. A. Gerken, in *Language Development*, M. Tomasello, E. Bates, Eds. (Blackwell, Oxford, 2001), pp. 42–48.
5. J. Morgan, K. Demuth, Eds., *Signal to Syntax* (Erlbaum, Mahwah, NJ, 1995).
6. R. N. Aslin, J. R. Saffran, E. L. Newport, in *The Emergence of Language*, B. MacWhinney, Ed. (Erlbaum, Mahwah, NJ, 1999), pp. 359–380.
7. M. S. Seidenberg, *Science* **275**, 1599 (1997).
8. ———, M. C. MacDonald, *Cogn. Sci.* **23**, 569 (1999).
9. G. F. Marcus *et al.*, *Science*, **283**, 77 (1999).
10. T. Meulemans, M. Van der Linden, *J. Exp. Psychol. Learn. Mem. Cogn.* **23**, 1007 (1997).
11. G. T. M. Altmann, Z. Dienes, *Science* **284**, 875 (1999).
12. M. S. Seidenberg, J. L. Elman, *Science* **284**, 433 (1999).
13. T. Johnstone, D. R. Shanks, *J. Exp. Psychol. Learn. Mem. Cogn.* **25**, 524 (1997).
14. R. H. Hunt, R. N. Aslin, *J. Exp. Psychol. Gen.* **130**, 658 (2001).
15. E. L. Newport, R. N. Aslin, in *Proc. 24th Ann. Boston Univ. Conf. Lang. Dev.* (Cascadia Press, Somerville, MA, 2000) pp. 1–21.
16. A. Cleeremans, J. L. McClelland, *J. Exp. Psychol. Gen.* **120**, 135 (1991).
17. J. R. Saffran, *J. Mem. Lang.* **47**, 172 (2002).
18. M. H. Christiansen, J. T. Devlin, *Proceedings of the 19th Annual Meeting of the Cognitive Science Society* (Erlbaum, Hillsdale NJ, 1997), pp. 113–118.
19. J. Allen, M. S. Seidenberg, in *The Emergence of Language*, B. MacWhinney, Ed. (Erlbaum, Mahwah, NJ, 1999), pp. 115–151.
20. We thank J. Keidel for phonetic insights and S. Pollak for helpful comments.