PSYCHOLOGY

Learning Rediscovered

Elizabeth Bates and Jeffrey Elman

In a report in this week's issue, Saffran, Aslin, and Newport have proven that babies can learn (*1*). Eight-month-old infants exposed for only 2 min to unbroken strings of nonsense syllables (for example, "*bidakupado*....") are able to detect the difference between three-syllable sequences that appeared as a unit and sequences that also appeared in their learning set but in random order. This result means that infants can use simple statistics to discover word boundaries in connected speech, right at the age when systematic evidence of word recognition starts to appear in real life (*2*). It is obvious that this is important; it may be less obvious to those outside the field why it flies in the face of received wisdom.

First, the nature of this learning is surprising: a purely inductive, statistically driven process, based on only 2 min of incidental input, with no reward or punishment other than the pleasure of listening to a disembodied human voice. Second, it contradicts the widespread belief that humans cannot and do not use generalized statistical procedures to acquire language (*3–7*). Noam Chomsky, the founder of generative linguistics, has argued for 40 years that language is unlearnable; he and his followers have generalized this belief to other cognitive domains, denying the existence of learning as a meaningful scientific construct:

"We may usefully think of the language faculty, the number faculty, and others, as 'mental organs' [that] develop in specific ways, each in accordance with the genetic program...multipurpose learning strategies are no more likely to exist than general principles of 'growth of organs' that account for the shape, structure and growth of the kidney" (*3*, pp. 138–139).

"I, for one, see no advantage in the preservation of the term 'learning'...we would gain in clarity if the scientific use of the term were simply discontinued" (*7*, p. 2).

"It is possible that the notion 'learning' may go the way of the rising and setting of the sun" (*3*, p. 245).

This belief is based on the famous "poverty-of-the-stimulus" argument: linguistic knowledge is "perfect," and it is impossible in principle to extract perfect knowledge from the imperfect data of everyday language use. A formal proof by Gold (*5*) appeared to support this assumption, although Gold's theorem is relevant only if we make assumptions about the nature of the learning device that are wildly unlike the conditions that hold in any known nervous system (*8*). There are, in fact, a number of ways to get around the poverty-of-the-stimulus argument.

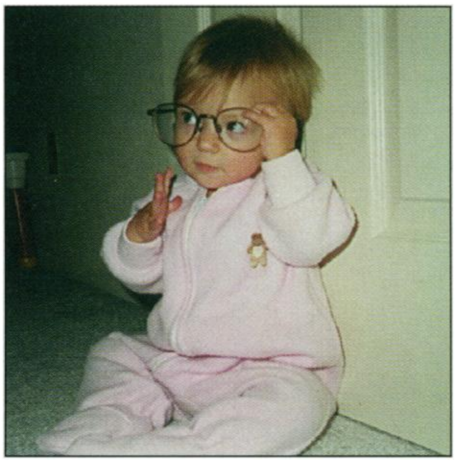

PHOTO: MICHAEL AND JEANETTE TWA

Learning language.

First, we could relax our definition of knowledge, defining successful learning to include behavior that is asymptotically correct but somewhere short of perfect ("close enough for government work"). Although there is plenty of evidence that humans use language creatively (saying and understanding things that have never been said before) and well (with very low error rates), there is very little evidence for the claim that "perfect" knowledge underlies our (occasionally) imperfect behavior.

Second, we could base our estimates of learnability on a more robust learning device than the one assumed by Chomsky. There is now a large body of evidence showing that artificial neural networks can induce regular patterns from imperfect, but quasi-regular input, and generalize those patterns to novel instances (*8–10*). Within the language domain, examples include the extraction of phonetic and phonological structures from raw speech (*9*), the discovery of word boundaries from connected speech (*8*), and the extraction of grammatical regularities from unlabeled strings of words generated by an artificial grammar with many of the properties of natural language (*10*).

Third, we now know that real speech contains a host of statistical regularities that are sufficient to support the kind of robust learning observed in neural networks (*11*). This knowledge has emerged from the analysis of huge computerized corpora of written and spoken language, revealing regularities that are not visible to the naked eye (or audible to the naked ear). Chomsky's belief in the impoverished nature of linguistic input holds only if we look "locally" at relatively short segments of speech. Such imperfections wash out with a large enough sample.

This brings us to the central contribution of the Saffran *et al.* report. Although we now know that linguistic regularities are learnable by neural networks with an imperfect but very large database, it has been argued that human infants do not learn in this way, and even if they did, their memory and attention span are insufficient to support the kind of statistical learning required to get language off the ground. This conclusion was premature: The new work (*1*) has shown that infants are capable of extracting statistical regularities from only 2 min of spoken input with little effort. To be sure, this experiment is not the first demonstration of early learning. For example, studies show that newborns prefer to listen to passages of speech from their native language, which means that some unspecified form of auditory learning has taken place in utero (*12*). Saffran *et al.* take us several steps further, with careful controls that make it absolutely clear what was learned, when, and how. Learning is much more powerful than previously believed, and arguments about the innateness of language and other forms of cognition need to take that undeniable fact into account.

The authors of the new work are quick to point out that their discovery does not justify a return to the *tabula rasa*. Learning is powerful, but it is not everything. In fact, relatively small variations in the initial architecture of a neural network can make the difference between "learnability" and "unlearnability" in the language domain (*8*, *10*). Even if we assume that a brain (real or

An enhanced version of this Perspective, with live links, can be seen in *Science* Online on the Web at http://www.sciencemag.org/

The authors are in the Departments of Cognitive Science and Psychology, University of California, San Diego, CA 92093, USA. E-mail: bates@cr1.ucsd.edu

artificial) contains no innate knowledge at all, we have to make crucial assumptions about the structure of the learning device, its rate and style of learning, and the kinds of input that it "prefers" to receive. The emergence of language in the hominid line must have involved a certain amount of tinkering with the primate brain, leading ultimately to a brain that was capable of learning language.

References

1. J. R. Saffran, R. N. Aslin, E. L. Newport, *Science* **274**, 1926 (1996).
2. L. Fenson *et al.*, *Monogr. Soc. Res. Child Dev.* **59** (1994).
3. N. Chomsky, *Rules and Representations* (Columbia Univ. Press, New York, 1980).
4. S. Crain, *Behav. Brain Sci.* **14**, 597 (1991).
5. E. Gold, *Inf. Control* **16**, 447 (1967).
6. D. Lightfoot, *The Language Lottery: Towards a Biology of Grammars* (MIT Press, Cambridge, MA, 1980); F. Newmeyer, *Linguistic Theory in America* (Academic Press, New York, 1980).
7. M. Piatelli-Palmarini, *Cognition* **31**, 1 (1989).
8. J. Elman *et al.*, *Rethinking Innateness: A Connectionist Perspective on Development* (MIT Press and Bradford Book, Cambridge, MA, 1996).
9. J. Elman and D. Ziupser, *J. Acoust. Soc. Am.* **83**, 1615 (1988).
10. J. Elman, *Cognition* **48**, 71 (1993).
11. K. Lund and C. Burgess, *Behav. Res. Methods Instrum. Comput.* **28**, 203 (1996).
12. P. Jusczyk, A. Friederici, J. Wessels, V. Svenkerud, A. Jusczyk, *J. Mem. Lang.* **32**, 402 (1993); J. Mehler, G. Lambertz, P. Jusczyk, C. Amiel-Tison, *C. R. Acad. Sci. Paris* **303**, 637 (1986).

BIOCHEMISTRY

Crossing the Hydrophobic Barrier: Insertion of Membrane Proteins

Donald M. Engelman

Lipid bilayers are thin, flexible self-sealing boundaries that are used by cells to create regions of different composition and electrochemical potential. To accomplish transmembrane functions, proteins inserted within and across the hydrophobic barrier must cope with hydrophilic interactions with the solutions inside and outside a cell or compartment and hydrophobic interactions with the membrane. Usually, proteins are assisted in their insertion by proteinaceous machinery. But can they insert spontaneously? The structure of α hemolysin, reported by Song *et al.* in this issue of *Science*, reveals how a protein, in this case a toxin produced by a pathogenic bacterium, can penetrate a lipid bilayer (*1*)—by the spontaneous formation of an oligomeric β barrel (see figure).

It is much easier to understand the stability of observed transmembrane structures than to fathom the process by which they are positioned within a membrane. The membrane proteins whose high-resolution structures have been solved contain either bundles of α helices or β barrels in the regions presumed to span the lipid bilayer (*2–5*). One can rationalize each of these structures by recognizing that the main-chain hydrogen bonds need to be satisfied in an environment that lacks hydrogen bond donors or acceptors, and that the hydrophobic effect will stabilize the association of a transmembrane structure with the hydrophobic region of a lipid bilayer if the amino acid side chains contacting this region are predominantly apolar (*6*). Detailed knowledge of the final, folded state of a protein, however, does not inform us directly about the process of insertion, particularly if the protein in question is inserted into the membrane in a process catalyzed by cellular machinery. On the other hand, proteins whose functions require them to be stable in an aqueous environment and also capable of inserting themselves into membranes provide an opportunity to examine, biochemically and structurally, the determinants of an insertion event.

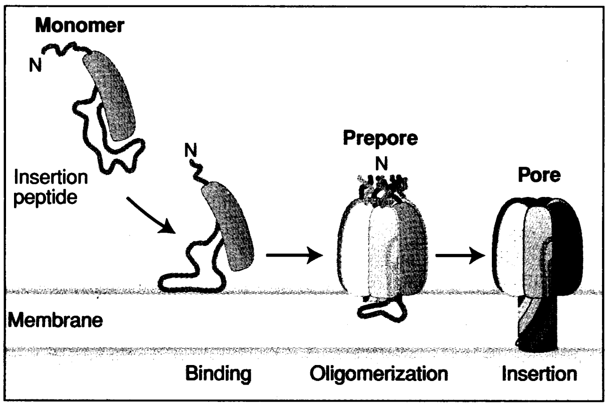

ILLUSTRATION: K. SUTLIFF

Binding, oligomerization, and transmembrane insertion of α hemolysin. The hemolysin is soluble as a monomer, binds through interactions of loops to the surface of a lipid bilayer, oligomerizes to form a heptamer stabilized by amino-terminal "latches," and inserts two strands of a β barrel from each subunit to form a 14-stranded β barrel across the bilayer. The structure that is known definitively is the final pore structure. There is less direct evidence to support the details of other steps. The molecular features of the bilayer at different stages are as yet completely undefined.

Spontaneous transmembrane insertions of both α helices and β barrels are found in the world of toxins, where the capacity to insert is packaged in a soluble molecule. Colicin A has a membrane-insertion domain that sequesters a hydrophobic helical hairpin whose insertion into the bilayer is postulated to be the primary step in colicin action (*7*). The α-hemolysin toxin studied by Song *et al.* is now shown to act by inserting β-barrel structures into bilayers, as previously surmised for aerolysin (*1*, *8*). Other cases of spontaneous insertion have been documented, including the insertion of porins from denatured states in solution into lipid bilayers (*9*). Despite the fact that proteinaceous machinery is used for the insertion of many membrane proteins, a number of cases exist in which the insertion event does not require the participation of structures other than the inserting polypeptide and the lipid bilayer. How might this occur? Which intermediate states might one imagine?

The answer to the issue of intermediates will undoubtedly vary for specific cases, but three themes emerge in the examples we have thus far. The first is the role of oligomerization in the process. Both the α hemolysin and proaerolysin change oligomeric state in the process of insertion. α Hemolysin binds to the membrane as a monomer, subsequently forms a heptamer, and then inserts. This sequence is based on studies of mutants that block steps in the process (*10*, *11*). Thus, one wonders whether the energy of oligomerization may drive the process, producing an intermediate state that relaxes to the transmembrane form. The structure shows that a large surface area is buried in the oligomerization event; hence, a large amount of energy could be available. This is undoubtedly one of the directions that will be explored by Song *et al.*

A second theme is the exposure of regions of the protein that are kept sequestered in the soluble form of the molecule. The colicin A structure shows a hydrophobic helical hairpin surrounded by another structure, sequestering it from the aqueous environment (*7*). A body of work indicates that confor-

The author is in the Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520, USA. E-mail: don@paradigm.csb.yale.edu