our ABA or ABB patterns to novel words (*3*). Seidenberg and Elman appear to abandon (without comment) the usual "predication task" version of the network model in favor of a different kind of model, in which an external teacher decides whether each pair of successive words is identical. Such information is not "directly observable from the environment" (*4*); instead, it is provided by an external teacher (built by Seidenberg and Elman) that itself builds in an algebraic rule. Because, in the human, that external device must be something inside the child rather than something provided by the environment, Seidenberg and Elman have not gotten rid of the rule; they have simply hidden it (*5*).

We find Negishi's model to be more interesting. Negishi points out, quite rightly, that an SRN that uses real numbers rather than binary encoding can capture our results. Why should that be the case? As we noted in our report, "algebraic" rules are "open-ended abstract relationships for which we can substitute arbitrary items." Models that use real-number encoding use their nodes as variables and incorporate operations that treat all instances of a given variable equally. In other words, rather then presenting an alternative to rules, such devices wind up implementing them (*6*).

This is a subtle point, perhaps best understood in a comparison (*3*) between two models, one that represents numbers as sets of discrete binary features, and another that represents numbers as analog values, such as the identity function mentioned by Negishi, $f(x) = x$. Neither architecture is inherently superior: Models that represent inputs as sets of nonarbitrary discrete features can capture transitional probabilities between words such as would be present in the experiments in the 1996 report by Saffran *et al.*, but cannot freely generalize the identity relationships that underlie our studies; models that use nodes as registers can freely generalize identity relationships, but cannot capture the transitional probabilities between words that underlie the experiments in that report. In some broad sense, both architectures might be characterized as "statistical," but the two architectures are suited to different problems.

Our results, in tandem with those of Saffran *et al.*, suggest that infants are capable of discerning both rules and transitional probabilities. As we said in our report (note 24), we aimed "not to deny the importance of neural networks but rather to try to characterize what properties the right sort of neural network architecture must have."

Gary F. Marcus
Department of Psychology, New York University, New York, NY 10003, USA. E-mail: gary.marcus@nyu.edu

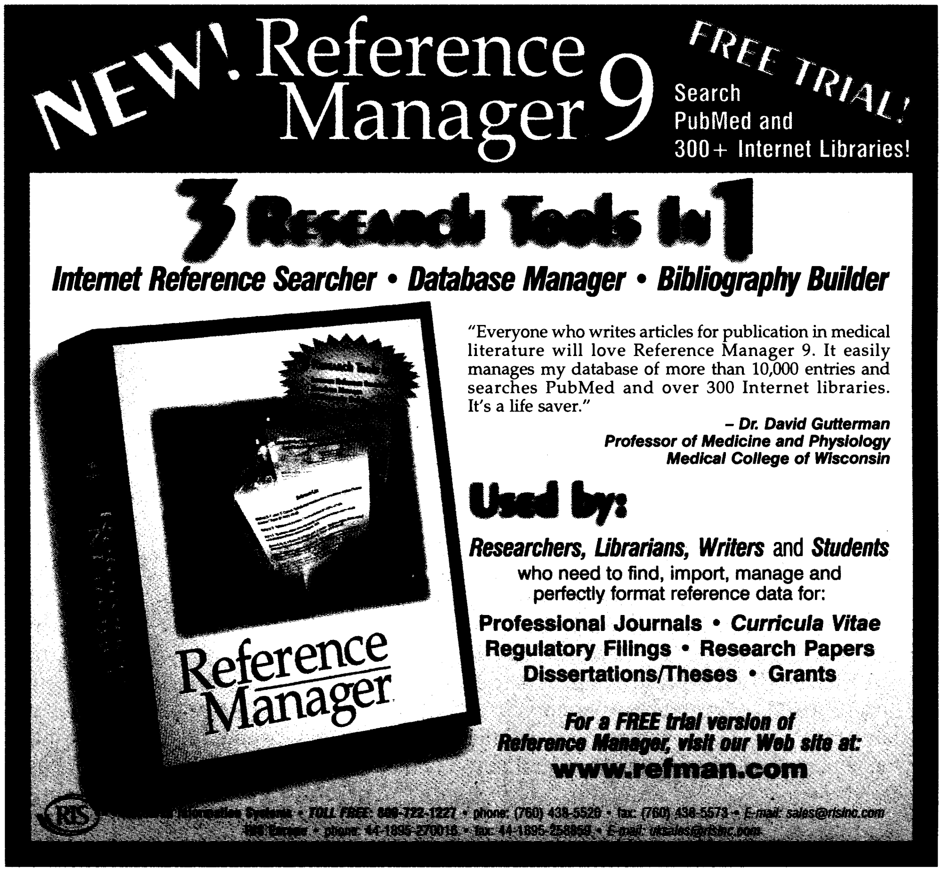
References and Notes

1. In the control experiment, we trained eight 7-month-old infants on sentences from a BAB or an AAB grammar and tested on BAB and AAB sentences made up of novel words. Seven of the eight infants looked longer at the inconsistent sentences than at the consistent sentences.
2. J. L. Elman, in *Mind as Motion: Explorations in the Dynamics of Cognition*, R. F. Port, and T. v. Gelder, Eds. (MIT Press, Cambridge, MA, 1995), pp. 195–223.
3. Discussion, examples, and models at psych.nyu.edu/~gary/science/discussion.html
4. The model was also given "negative evidence"; that is, in the habituation phase, the model was told not only which sentences are ABB sentences (positive evidence), but also which sentences were not (negative evidence). In contrast, the infants in our experiment were given only positive evidence, and not exposed to examples of "ungrammatical patterns." Our experiment, but not the Elman-Seidenberg model, is consistent with the assumption that children are able to learn grammar without negative evidence [R. W. Brown and C. Hanlon, in *Cognition and the Development of Language*, R. Hayes, Ed. (Wiley, New York, 1970); J. L. Morgan and L. L. Travis, *J. Child Lang.* **16**, 531 (1989); G. F. Marcus, *Cognition* **46**, 53 (1993)].
5. Seidenberg and Elman appear to use the term "statistics" to refer to regularity, thus counting rules as statistical regularities. Weakening the terminology in this way does not take away from our point that infants can learn rules.
6. A recent paper of ours, cited in note 22 in our report, made this point explicitly (*7*, p. 275): "While most networks represent inputs by pattern of activation across sets of nodes, in principle one could use a single node to represent all possible inputs, assigning each possible input to some real number...incorporating what is a transparent implementation of a register....The node in question would represent a variable; its value would represent the instantiation of that variable." For two other examples of neural network architectures that explicitly implement relationships between variables and that could capture our findings without a hidden teacher or negative evidence, see K. J. Holyoak and J. E. Hummel, in *Cognitive Dynamics: Conceptual Change in Humans and Machines*, E. Deitrich and A. Markman, Eds. (Erlbaum, Mahwah, NJ, 1999) and L. Shastri and V. Ajjanagadde, *Behav. Brain Sci.* **16**, 417 (1993). For further discussion, see (*7*) and G. F. Marcus, *The Algebraic Mind* (MIT Press, Cambridge, MA, in press).
7. G. F. Marcus, *Cognit. Psychol.* **37**, 243 (1998).

CORRECTIONS AND CLARIFICATIONS

Throughout Constance Holden's article "Dispute over a legendary fish" (News of the Week, 2 Apr., p. 22), the word "*Comptes*" in the title of the French journal *Comptes Rendus de l'Académie des Sciences* was spelled incorrectly.

The photo at the upper right on page 1623 of the issue of 12 March was of climbers at Pinnacles National Monument in California, not Joshua Tree National Monument.



Circle No. 26 on Readers' Service Card