

- 20 Bottari, P., Chipriani, P. and Chilosi, A. (1996) Root infinitives in Italian SLI children, in *Proceedings of the 20th Boston University Conference on Language Development* (Stringfellow, A. et al., eds), pp. 76–86, Cascadilla Press, Boston
- 21 Dromi, E., Leonard, L. and Shteyman, M. (1993) The grammatical morphology of Hebrew-speaking children with specific language impairment *J. Speech Hear. Res.* 36, 760–771
- 22 Hakansson, G. (1998) Language impairment and the realization of finiteness, in *Proceedings of the 22nd Boston University Conference on Language Development* (Greenhill, A. et al., eds), pp. 314–324, Cascadilla Press, Boston
- 23 Hamann, C., Penner, Z. and Lindner, K. (1998) German impaired grammar: the clause structure revisited *Lang. Acquis.* 7, 193–245
- 24 Jakubowicz, C., Nash, L. and van der Velde, M. (1999) Inflection and past tense morphology in French SLI, in *Proceedings of the 22nd Boston University Conference on Language Development* (Greenhill, A. et al., eds), pp. 289–300, Cascadilla Press, Boston
- 25 van der Lely, H.K.J. (1998) SLI in children: movement, economy and deficits in the computational-syntactic system *Lang. Acquis.* 7, 161–192
- 26 Neville, H. et al. (1993) The neurobiology of sensory and language processing in language-impaired children *J. Cogn. Neurosci.* 5, 235–253

Correspondence

Networks are not 'hidden rules'

Does learning a language involve formulating rules or gathering statistics? Marcus and colleagues offered two pieces of evidence bearing on this debate¹. Their behavioral studies were taken as evidence that babies form 'algebraic rules' and their attempt to model this behavior suggested that it is incompatible with the properties of a 'popular' class of connectionist networks. Both of these claims have been the subject of considerable discussion elsewhere^{2–9}. In this letter we would like to raise a more general issue about the relationship between connectionist models and algebraic rules. More specifically, we would like to examine critically the statement by Marcus that 'Seidenberg and Elman have not gotten rid of the rule; they have simply hidden it', in reference to simulation data that we recently reported^{2,5,7}.

The purpose of our simulation was to demonstrate that the sequential regularities implicit in Marcus et al.'s stimuli provided a sufficient basis for differentiating sequences that conformed to the 'algebraic rule' from ones that did not, and that having acquired this information, a network could generalize appropriately to novel stimuli. The model was not an account of exactly how babies acquire this information; rather, it demonstrated that if they encoded this information, by whatever means, it would provide a basis for the observed behavior. That babies are able to detect such regularities is consistent with an extensive empirical literature¹⁰.

Marcus objected to two aspects of this simulation: instead of the prediction task used in some other models (including Marcus et al.'s failed simulation), our model was trained to categorize stimuli as fitting a pattern or not doing so. We also used a supervised-learning procedure in which the network was provided with explicit feedback. In a sense, then, feedback to the model was structured on the basis of a rule; hence, Marcus concluded that the model must have had the rule 'hidden' in it. However, his statement is a *non sequitur*. Merely training a network to categorize stimuli into two groups using explicit feedback does not cause it to formulate a rule.

Categorization, like prediction, is a task. The theoretical issue is how such tasks are performed. For a while now people have been debating two competing accounts. One is that such tasks involve formulating rules that have specific properties: for example, they operate over variables and are insensitive to incidental characteristics of the stimuli to which they apply. The other theory is different insofar as it states that the task is performed by using information that the rule mechanism explicitly ignores – stimulus properties, easily perceived by infants and learned by neural networks, that create different degrees of similarity between stimuli. Learning rules involves one type of procedure; acquiring statistical information involves another. Training a net to categorize stimuli using direct feedback does not cause one type of knowledge, a constraint-satisfaction network, to metamorphose into another type of knowledge, a rule, or create a rule-learning device out of one that performs gradient descent in weight space.

Marcus' concern was apparently prompted by the fact that the model was explicitly trained to turn a bit on for one type of sequence and off for another. But, the bit is not a variable and the network is not a rule. We choose this method as a simple and direct way to cause the model to encode statistical properties of the stimuli that we hypothesize underlie babies' categorization performance. As the Altmann⁸ and Christiansen and Curtin^{11,12} models show, there are other ways to achieve the same outcome, including ones more closely tied to the infant's experience (although see responses from Marcus^{9,13}).

The network that we described, like other connectionist networks, provides a different account of behavior previously attributed to rules. The properties of such networks cannot be deduced from the properties of rules, at least as they are standardly construed. The only way to maintain the idea that the behavior of such systems is rule-governed is by changing the definition of 'rule' to mean 'having the properties of a neural network'. Stripped of its perjorative as-

pect, that is what the statement that the rule was 'hidden' in the network amounts to.

Pinker and Prince observed that in order to account for linguistic phenomena, connectionist networks would have to conform to ('implement') rule-based theories¹⁴. What has actually happened, as Marcus' comments illustrate, is that the concept of 'rule' is being altered to conform to the properties of connectionist networks. We think that this move involves abandoning what made rules an interesting alternative to connectionism. In any case, what is important is understanding the behaviour of the network and the principles that govern it, not Marcus' attempt to label it.

Mark S. Seidenberg
Neuroscience Program,
University of Southern California,
Los Angeles, CA 90089-2520, USA.
tel: +1 213 740 9174
fax: +1 213 740 5687
e-mail: marks@gizmo.usc.edu

Jeffrey L. Elman
Department of Cognitive Science,
University of California,
San Diego, La Jolla,
CA 92093-0515, USA.
tel: +1 619 534 1147
fax: +1 619 534 678
e-mail: elman@crl.ucsd.edu

References

- Marcus, G.F. et al. (1999) Rule learning by seven-month-old infants *Science* 283, 77–80
- Seidenberg, M. and Elman, J.L. (1999) Do infants learn grammar with algebra or statistics? *Letter Science* 284, 434–435
- Negishi, M. (1999) Do infants learn grammar with algebra or statistics? *Letter Science* 284, 435
- Eimas, P.D. (1999) Do infants learn grammar with algebra or statistics? *Letter Science* 284, 435–436
- Marcus, G.F. (1999) Do infants learn grammar with algebra or statistics? Response to Seidenberg & Elman, Negishi, and Eimas *Science* 284, 436–437
- McClelland, J.L. and Plaut, D.C. (1999) Does generalization in infant learning implicate

abstract algebra-like rules? *Trends Cognit. Sci.* 3, 166–168

- 7 Marcus, G.F. (1999) Connectionism: with or without rules? *Trends Cognit. Sci.* 3, 168–170
- 8 Altmann, G.T.M. and Dienes, Z. (1999) Rule learning by seven-month-old infants and neural networks *Science* 284, 875
- 9 Marcus, G.F. (1999) Rule learning by seven-month-old infants and neural networks. Response to Altmann and Dienes *Science* 284, 875
- 10 Juczyk, P. (1997) *The Discovery of Spoken Language*, MIT Press
- 11 Christiansen, M.H. and Curtin, S.L. The power of statistical learning: no need for algebraic rules, in *Proc. 21st Annu. Conf. Cognit. Sci. Soc.* (Mahwah, NJ), Erlbaum (in press) (a Web-based version can be found at <http://www.siu.edu/~psycho/faculty/morten/statlearn.html>)
- 12 Christiansen, M.H. and Curtin, S.L. (1999) Transfer of learning: rule acquisition or statistical learning? *Trends Cognit. Sci.* 3, 289–290
- 13 Marcus, G.F. (1999) Reply to Christiansen and Curtin *Trends Cognit. Sci.* 3, 290–291
- 14 Pinker, S. and Prince, A. (1988) On language and connectionism: analysis of a parallel distributed processing model of language acquisition *Cognition* 28, 73–194

Reply to Seidenberg and Elman

We completely agree with Seidenberg and Elman¹ that what is important is 'understanding the behaviour of the network and the principles that govern it', but are mystified by their discussion of rules. Seidenberg and Elman's claim that 'the concept of "rule" has been altered to conform to the behaviour of connectionist networks' is a clever, ironic twist, but it actually has no substance, and reflects only Seidenberg and Elman's misunderstanding of our earlier point. As I emphasized in a series of recent articles, only a subset of neural networks implement rules^{2–4}; not all of them do. As we stated before², the ones that depend on 'open-ended abstract relationships for which we can substitute arbitrary items' are the ones that implement rules. As it turns out, the *neural network* part of Seidenberg and Elman's system does not implement a rule. But, crucially, another part of the system that Seidenberg and Elman propose *does* implement a rule: the external teacher (this is what we meant by a 'hidden rule'). In particular, the external teacher incorporates a universally open-ended rule of the sort, *for all syllables x, y, if x = y then output 1 else output 0*. This corresponds exactly to our characteriz-

ation of rules. Since part of the system contains a rule, it follows that the system as a *whole* contains a rule.

Seidenberg and Elman also point to models of Altmann and Dienes⁵ and Christiansen and Curtin⁶. We have discussed the Altmann and Dienes model elsewhere⁷. The pointer to the Christiansen and Curtin model is at best premature: the difference underlying their crucial result was small, it has not been shown to be statistically significant, and it has not yet been replicated (for a more detailed response see my reply to Christiansen and Curtin, in this issue⁸).

We do not expect this discussion to die down any time soon; clearly, advocates of certain types of connectionism believe that a lot is at stake. But the proliferation of alternative models should not distract us from noticing limits on particular classes of models where such limits do exist. For it is only by taking limits seriously that we can hope to build better models.

Gary F. Marcus
Department of Psychology,
New York University, 6 Washington Place,
New York, NY 10003, USA.
tel: +1 212 998 3551

fax: +1 212 995 4292
e-mail: gary.marcus@nyu.edu
<http://www.psych.nyu.edu/~gary>

References

- 1 Seidenberg, M.S. and Elman, J.L. (1999) Networks are not 'hidden rules' *Trends Cognit. Sci.* 3, 288–289
- 2 Marcus, G.F., Vijayan, S., Rao, S.B. and Vishton, P.M. (1999) Rule learning in 7-month-old infants *Science* 283, 77–80
- 3 Marcus, G.F. (1998) Rethinking eliminative connectionism *Cognit. Psychol.* 37, 243–282
- 4 Marcus, G.F. (1999) Do infants learn grammar with algebra or statistics? Response to Seidenberg & Elman, Negishi, and Eimas *Science* 284, 436–437
- 5 Altmann, G.T.M. and Dienes, Z. (1999) Rule learning by seven-month-old infants and neural networks *Science* 284, 875
- 6 Christiansen, M.H. and Curtin, S.L. (1999) Transfer of learning: rule acquisition or statistical learning? *Trends Cognit. Sci.* 3, 289–290
- 7 Marcus, G.F. (1999) Rule learning by seven-month-old infants and neural networks. Response to Altmann and Dienes *Science* 284, 875
- 8 Marcus, G.F. (1999) Reply to Christiansen and Curtin *Trends Cognit. Sci.* 3, 290–291

Transfer of learning: rule acquisition or statistical learning?

Thirty years ago, Arthur Reber demonstrated that adults show evidence of transfer of learning in artificial-language experiments, in which the surface vocabulary is changed between training and test items¹. In a series of experiments, Marcus and colleagues² demonstrated that infants as young as seven months old also show evidence of transfer of learning, but incorrectly conclude that the infants were extracting abstract algebraic rules rather than encoding statistical regularities. In contrast, a recent comprehensive review of the artificial-language-learning literature has demonstrated that transfer does not entail the involvement of abstract rules³.

In Marcus *et al.*'s most persuasive demonstration of transfer of learning (Experiment 3 in Ref. 2), the infants were first trained on syllable sequences that

followed either an AAB or ABB pattern (e.g. 'le-le-je' versus 'le-je-je'). The infants were then presented with sequences of novel syllables, either consistent or inconsistent with the training pattern. The infants showed a preference for the inconsistent items, thus demonstrating transfer between the different syllable vocabularies used in habituation and testing. Because there was no phonological overlap between training items and test items, Marcus *et al.* concluded that a statistical learning device could not account for these transfer results without implementing algebraic-like rules (see also the responses by Marcus^{4,5} to commentaries by Seidenberg and Elman⁶ and by McClelland and Plaut⁷). However, we suggest that statistical knowledge acquired in the service of learning to segment fluent speech into words might

provide the basis for these transfer effects in much the same way as knowledge acquired in the process of learning to read can be used to perform experimental tasks such as lexical decision.

Using an existing simple recurrent-network⁸ model of early infant speech segmentation⁹ (Fig. 1), we tested this suggestion and successfully modeled the Marcus *et al.* results¹⁰. Importantly, no modifications were made to the original model, which learned to segment speech by integrating different kinds of probabilistic information derived from the speech stream (phonology, lexical stress and utterance-boundary information). Moreover, the simulation closely replicated the experimental conditions during both habituation and testing. The internal representations of the model were recorded at the end of each test item,