FRIES, CHARLES CARPENTER. 1925. The periphrastic future with *shall* and *will* in Modern English. PMLA 40.963–1024.

IHALAINEN, OSSI. 1994. The dialects of English since 1776. The Cambridge history of the English Language: English in Britain and Overseas, ed. by Robert W. Burchfield, 197–274. Cambridge: Cambridge University Press.

ROMAINE, SUZANNE (ed.) 1998. The Cambridge history of the English language: 1776–1997. Cambridge: Cambridge University Press.

Department of English Language and Literature
3187 Angell Hall
University of Michigan
Ann Arbor, MI 48109-1003
[rwbailey@umich.edu]

The algebraic mind: Integrating connectionism and cognitive science. By GARY F. MARCUS. Cambridge, MA: MIT Press, 2001. Pp. 208. $27.95.

Reviewed by IRIS BERENT, *Florida Atlantic University*

How do speakers produce new linguistic forms? Generative accounts attribute the productivity of language to the grammar. Although these theories greatly differ on their specific accounts of the grammar, they all share the assumption that the grammar, the home of linguistic productivity, is a distinct aspect of linguistic knowledge that is separate from and irreducible to the lexicon. With the rise of connectionism, however, this basic assumption has been the subject of fierce debate in the psychological literature. Connectionism is a computational framework that captures knowledge in terms of the activation patterns in a network of interconnected nodes. There are now hundreds of connectionist models that exhibit some degree of linguistic productivity despite having no separate grammatical component. The challenge that these networks present to linguistic research cannot be underestimated: If linguistic knowledge could be captured in terms of the statistical properties of lexical entries, then the notion of grammar would become all but obsolete.

Does a theory of language need a grammar? Marcus's book is a unique, remarkable achievement that is bound to reshape the discussion on this question. M's argument begins not with grammar but with symbol manipulation. As M explains, symbolic accounts of cognition share three assumptions. First, the mind operates on variables—abstract placeholders such as nouns and verbs, akin to algebraic variables (e.g. X). Second, the mind represents the constituents structure of variables (e.g. recursion, X → XY). Third, the mind distinguishes between types (e.g. 'mouse') and individuals ('Mickey'). Linguistic rules, principles, and constraints are special cases of symbolic processes—algebraic operations whose semantic output is determined by the constituent structure of mental variables. For grammars to exist, the mind must have the capacity to perform symbolic operations.

In his book, M identifies numerous areas of cognition, including language, that implicate a symbolic architecture. He demonstrates that the success of connectionist models in these domains critically depends on their maintenance of the symbolic tenets. The most powerful of those analyses is the discussion of relations between variables in Ch. 3. It is this discussion that offers a principled computational explanation as to why grammatical generalizations are irreducible to the properties of lexical instances.

Key to the ability of symbolic architectures to generalize is their operation over variables—abstract placeholders such as noun and verb. The incorporation of variables crucially determines the scope of generalizations. Because symbolic operations appeal to variables (e.g. *noun*), not instances (e.g. *house*, *dog*, etc.), they can be extended to any new instance, regardless of its properties, its familiarity, or similarity to known instances. For instance, because the English past tense rule concatenates the variables *verbstem* and *suffix* (Pinker 1999), it can apply to both familiar (e.g. *flip*) and novel (e.g. *plip*) stems. The representation of variables and the operation

over variables, however, is not necessary for generalization. Numerous connectionist models can generalize without representing variables or incorporating operations over variables. Proponents of such networks claim that generalizations 'emerge' in connectionist models that are not equipped with mechanisms for operating over variables prior to training.

The implications of such findings crucially depend on what, precisely, such 'emergent' generalizations really are—a pivotal question that has not been fully articulated in the connectionist literature. In one view, such systems learn the ability to operate over variables, an ability with which symbolic architectures are innately (i.e. in advance of any training) equipped. Such emergence would not challenge the observational or descriptive adequacy of symbolic operations in general or grammars in particular as the scope of generalizations in these two architectures is indistinguishable. Such emergent generalizations, if they existed, however, would question the explanatory adequacy of symbolic systems, namely, the hypothesis that the ability to operate over variables must be innately specified. The innateness of operations on variables should not be equated with the innateness of specific grammatical principles or constraints: The innateness of the symbolic machinery is not sufficient argument for the innateness of any specific linguistic principle or constraint. But if operations over variables, the core of many linguistic principles and constraints, were learnable from experience with lexical instances, then the very notion of the grammar as an autonomous computational system would be obsolete.

Most discussions of connectionism (e.g. Rumelhart & McClelland 1986), however, appear to have adopted a rather different view of 'emergence'. In that view, training a connectionist network does not lead to the acquisition of operations over variables. What the system acquires, instead, is knowledge concerning the statistical properties of training instances—the items on which the network is trained. Although such statistical knowledge supports generalizations, they are fundamentally distinct from symbolic generalizations. Rather than operating over variables (e.g. *verbstem*), generalizations are formed by analogizing to similar familiar instances (e.g. *plipped* is formed by analogy to *flipped*). This subtle distinction is significant: If human behavior may be captured by generalizations that do not manipulate mental variables, then symbolic models could offer only a gross characterization of knowledge, an account that, upon closer scrutiny, is bound to reveal inconsistencies with human behavior. Grammars, then, are not mental organs—they are merely gross descriptions devoid of both observational and explanatory adequacy as accounts of linguistic generalizations.

The adequacy of grammars thus critically depends on a precise assessment of the scope of linguistic generalizations and their learnability, two crucial questions that have not been clearly articulated in existing research. M's analysis is an important leap ahead on both fronts. M proposes to define the scope of generalizations relative to a model's training space—the representational space of instances (or instance features) on which a model is trained. For example, for a model that encodes place of articulation and is trained on the labial *b*, a new labial (e.g. *m*) may fall within its training space, but a coronal (e.g. *t*) falls outside it, as the coronal node received no relevant training. M unveils a principled limitation in the ability of a large class of popular connectionist networks to extend such generalizations. While these systems can interpolate generalizations to novel items within their training space, they are incapable of extrapolating generalizations to novel items that fall outside the training space. This limitation, a problem M traces to the learning algorithms used in standard connectionist networks, prevents these systems from capturing human behavior. Unlike these networks, people can extend generalizations beyond their training space. M supports this conclusion by a detailed analysis of two case studies that have been each subject to an ample behavioral and computational investigation: linguistic inflection and reduplication in an artificial language (for a demonstration in a natural language, see Berent et al. 2002). M's review of 30 connectionist networks in these areas suggests that the ability of such models to capture human behavior crucially depends on the implementation of variables: Models that are innately equipped with operations over variables can generalize beyond the training space, those that are not—can't.

The implications of these conclusions are far reaching. Generalization to novel items has long been considered critical for a model's observational adequacy. M's analysis, however, suggests that this may be too coarse of a test of a model's behavior as the ability of many popular

connectionist networks to generalize varies markedly depending on the position of test items relative to their representational space. The adequacy of the grammar, then, cannot be simply evaluated by generalizations to novel instances. Instead, one must consider their type: within or outside the training space. The notion of the training space thus offers a methodological yardstick that allows one to test more precisely the observational adequacy of linguistic theories and connectionist models against human behavior. M's analysis also promises to bring some significant progress to the debate regarding the nature of mental architecture. His discovery of a class of generalizations that falls exclusively within the scope of symbolic models renders symbolic accounts, including grammars, far more readily falsifiable. Finally, M's work advances our understanding of the explanatory adequacy of generative grammars. Generative grammars assume operations over variables at their core. It is only if operations over variables are not learnable from the statistical properties of linguistic instances that grammars could correspond to autonomous mental organs, independent from the lexicon. M's argument that the ability to perform operations over variables is not itself learnable from experience thus greatly reduces the class of possible explanations for the acquisition of grammars.

So, does a theory of language need a grammar? M certainly argues that it needs an innate symbolic computational system that is distinct from the lexicon. Such a system, he argues, is also adaptive from an evolutionary standpoint and is biologically plausible. Whether this innate computational system is language-specific is a crucial question that falls beyond the scope of M's analysis. But, regardless of whether grammars are partly innate, M gives good reason to believe that the machinery of symbol manipulation at their core may well be. Is he right to argue so? As is evident from the dozens of connectionist networks designed to challenge his claims (challenges he reviews and refutes), M's conclusions are not universally accepted. This is hardly surprising given the profound implications of his claims. As with any other major scientific theory, however, the significance of his contribution rests on the proposal of a more precise, falsifiable hypothesis regarding a major scientific question, not the truthfulness of his claims. I can think of few other questions that are more fundamental to cognitive science and just as few analyses whose significance approaches that of M's delightful essay.

REFERENCES

BERENT, IRIS; GARY F. MARCUS; JOSEPH SHIMRON; and ADAMANTIOS GAFOS. 2002. The scope of linguistic generalizations: Evidence from Hebrew word formation. Cognition 83.113–39.
PINKER, STEVEN. 1999. Words and rules: The ingredients of language. New York: Basic Books.
RUMELHART, DAVID. E. and JAMES L. MCCLELLAND. 1986. On learning past tense of English verbs: Implicit rules or parallel distributed processing? Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 2: Psychological and biological models, ed. by J. L. McClelland, D. E. Rumelhart, and T. P. R. Group, 216–71. Cambridge, MA: MIT Press.

Department of Psychology
Florida Atlantic University
777 Glades Rd.
Boca Raton, FL 33431-0991
[IBerent@fau.edu]

Political discourse in transition in Europe 1989–1991. Ed. by PAUL A. CHILTON, MIKHAIL V. ILYIN, and JACOB L. MEY. Amsterdam: John Benjamins, 1998. Pp. x, 272.

Reviewed by ALI EMINOV, *Wayne State College*

This volume was inspired by a conference on political discourse held in Moscow in 1989, two weeks before the Berlin wall was torn down. Over the next two years the communist regimes in Eastern Europe and the former Soviet Union, along with the institutions associated with them,