

EMERGENTISM

William O'Grady

A significant body of linguistic research can be situated in the philosophical and scientific tradition known as emergentism. This essay offers a brief overview of this work, with a focus on its guiding principles and on the proposals it makes concerning the nature of human language.

The emergentist tradition

The roots of emergentism can be traced to the work of John Stuart Mill (1930 [1843]), who proposed that a system can have properties that amount to more than the sum of its parts. The physical world offers many examples of this, as Mill observes.

The chemical combination of two substances produces, as is well known, a third substance with properties different from those of either of the two substances separately, or both of them taken together. Not a trace of the properties of hydrogen or oxygen is observable in those of their compound, water. (p. 243)

Mill's insight is relevant to the study of so-called "Complex Systems"—ranging from atoms to the weather—whose dynamic non-linear behavior involves many interacting and interconnected parts. (A system is dynamic if it is constantly in flux; it is non-linear if effects are out of proportion to causes, as when a neglected candle causes a fire that destroys an entire city. See *DYNAMICAL SYSTEMS* and *SELF-ORGANIZING SYSTEMS*.) However, the question of whether and to what extent language is an emergent phenomenon remains controversial.

Linguistic emergentism

Although it is widely agreed that emergentist approaches to language necessarily stand in opposition to theories of the language faculty that posit an innate *UNIVERSAL GRAMMAR*, other tenets of linguistic emergentism are less

well defined and there is no consensus within the field as to how precisely the standard problems of linguistic analysis should be confronted. Nonetheless, the starting point for a substantial portion of emergentist work seems to involve a commitment to the following thesis.

(1) *The emergentist thesis for language*

The phenomena of language are best explained by reference to more basic non-linguistic (i.e., “non-grammatical”) factors and their interaction.

An appealing tag line for linguistic emergentism comes from Bates and MacWhinney (1988, 147): language, they say, is a “new machine built out of old parts.” While there is no general agreement concerning just what those parts might be, the list is relatively short, ranging from features of physiology and perception, to processing and working memory, to pragmatics and social interaction, to properties of the input and of the learning mechanisms.

A significant amount of emergentist work within linguistics adopts the techniques of *CONNECTIONISM*, an approach to the study of the mind that seeks to model learning and cognition in terms of networks of (assumedly) neuron-like units. In its more extreme forms, connectionism rejects the existence of the sorts of symbolic representations (including syntactic structure) that have played a central role in explanatory work on human language. Marcus (1998, 2001) and Gregg (2003) offer a critique of this sort of “eliminativist” program, while Smolensky (1999) and Steedman (1999) discuss ways to reconcile it with traditional symbolic approaches to language, including the possibility that representations might be abstract, higher-level descriptions that approximate the patterns of neuronal activation that connectionist approaches seek to model.

Although connectionist modeling provides a useful way to test various predictions about language acquisition, processing, change, and evolution, the eliminativist position is far from universally accepted within emergentism. Symbolic representations of one form or another are evident in the work of many emergentists (e.g., Goldberg 1999, Tomasello 2003, O’Grady 2001, 2005), who nonetheless reject the view that the properties of those representations should be attributed to innate grammatical principles.

Language acquisition

To date, emergentist work within linguistics has focused most strongly on the question of how language is acquired (see, e.g., the many papers in MacWhinney

1999). The impetus for this focus stems from opposition to the central claim of grammatical nativism, which is that the principles underlying a good deal of linguistic knowledge are underdetermined by experience and must therefore be innate.

Emergentism is not opposed to nativism *per se*—the fact that the brain is innately structured in various ways is not a matter of dispute. However, there is opposition to “representational nativism,” the view that there is direct innate structuring of particular grammatical principles and constraints (Elman et al. 1996:369ff, Bates et al. 1998), as implied by many of the proposals associated with Universal Grammar.

Contemporary emergentism often includes a commitment to explaining linguistic development by reference to the operation of simple learning mechanisms (essentially, inductive generalization) that extract statistical regularities from experience. Interestingly, there is as of yet no consensus as to what form the resulting knowledge might take—local associations and memorized chunks (Ellis 2002), constructions (Goldberg 1999, Tomasello 2003), or computational routines (O’Grady 2001, 2005). In addition, there is variation with respect to the exact relationship that is assumed to hold between learning and relative frequency in the input. Some work implies a quite direct relationship (e.g., Ellis 2002), but other work suggests something less direct (e.g., Elman 2002).

Emergentist work on language acquisition often makes use of computer modeling to test hypotheses about development. Jeffrey Elman and his colleagues (e.g., Elman 2002) have been able to show that a Simple Recurrent Network (SRN) can achieve at least some of the milestones associated with language acquisition in children, including the identification of category-like classes of words, the formation of patterns not observed in the input, retreat from overgeneralizations, and the mastery of subject-verb agreement. (An SRN learns to produce output of its own by processing sentences in its input; it is specifically designed to take note of local co-occurrence relationships, or “transitional probabilities” — given the word X, what’s the likelihood that the next word will be Y?)

Emergentist modeling has yielded impressive results, but it raises the question of why the particular statistical regularities exploited by the SRN are in the input in the first place. In other words, why does language have the particular properties that it does? Why, for example, are there languages (such as English) in which verbs agree only with subjects, but no language in which verbs agree only with direct objects?

Networks provide no answer to this sort of question. In fact, if presented with data in which verbs agree with direct objects rather than subjects, an SRN would no doubt “learn” just this sort of pattern, even though it is not found in any known human language.

There is clearly something missing here. Humans don’t just learn language; they shape it. Moreover, these two facts are surely related in some fundamental way, which is why hypotheses about how linguistic systems are acquired need to be embedded within a more comprehensive theory of why those systems (and therefore the input) have the particular properties that they do. There is, simply put, a need for an emergentist theory of *grammar*.

Emergentist approaches to grammatical theory

A substantial amount of analytic work has addressed the traditional concerns of linguistic analysis, including “core” phenomena in the major areas of traditional grammatical theory.

Syntax. It is possible to identify several strands of emergentist work on *SYNTAX*, each devoted to explaining the structural properties of sentences without reference to inborn grammatical principles. Differing views have been put forward by MacWhinney (2005) and O’Grady (2001, 2005), both of whom address a series of issues that lie at the heart of contemporary syntactic analysis—the design of phrase structure, co-reference, agreement, the phonology-syntax interface, and constraints on long-distance dependencies. MacWhinney seeks to explain these phenomena in terms of pragmatics, arguing that grammar emerges from conversation as a way to facilitate accurate tracking and switching of perspective. In contrast, O’Grady holds that syntactic phenomena are best understood in terms of the operation of a linear, efficiency-driven processor that seeks to reduce the burden on working memory in the course of sentence formation and interpretation.

Still other work, such as that done within *CONSTRUCTION GRAMMAR*, seeks to reduce syntax to stored pairings of form and function (“constructions”). Some of this work has a strong emergentist orientation (e.g., Goldberg 1999, Tomasello 2003), but some retains a commitment to Universal Grammar (Goldberg and Jackendoff 2004, 563).

Morphology. Very early connectionist work on *MORPHOLOGY* called into question the existence of morphological rules and representations, even for phenomena such as regular past tense inflection. Instead, it was suggested, a pattern-associator network learns the relationship between the phonological form of stems and that of past tense forms (*run~ran*, *walk~walked*, etc.), gradually establishing associations (“connections”) of different strengths and levels of

generality between the two sets of elements—the most general and strongest involving the *-ed* past tense form. McClelland and Patterson (2002) offer a succinct overview of this perspective.

More recent work has raised important questions about the nature of *MORPHEMES* in general. A key claim of this research is that morphological structure emerges from statistical regularities in the form-meaning relationships between words. (Hay and Baayen 2005 offer an excellent review of this research.)

Intriguing experimental work by Hay (2003) suggests that the internal structure of an affixed word is gradient rather than categorical, reflecting its relative frequency compared to that of its base. The words *inadequate* and *inaudible* are a case in point. Because *adequate* is more frequent than the affixed form *inadequate*, its presence in the derived word is relatively salient, leading to a high native speaker rating for structural complexity. In contrast, *inaudible*, which is more frequent (and therefore more salient) than *audible*, receives a low rating for structural complexity.

If this is right, then morphological structure exists, but not in the categorical form commonly assumed. Rather, what we think of as morpheme boundaries emerge to varying degrees of strength from the interaction of more basic factors such as frequency, semantic transparency, and even phonotactics. (The low-probability sequence in *inhumane* creates a sharper morpheme boundary than the high-probability sequence in *insincere*.)

The lexicon. There have been various attempts to develop an emergentist approach to the *LEXICON*, which is traditionally seen as the repository of information about morphemes and words. One possibility, suggested by Bybee (1998), among others, is that the lexicon emerges from the way in which (by hypothesis) the brain responds to and stores linguistic experience—by creating units whose strength and productivity is determined largely by frequency of occurrence. Some of these units correspond to words, as in a traditional lexicon, but many are phrases and other larger units of organization, including possibly abstract constructions (see *USAGE-BASED THEORY*).

Elman (2004) also argues against a pre-structured lexicon, proposing instead that lexical knowledge is implicit in the effects that words have on the mind's internal states, as represented in the activation patterns created by an SRN. Because an SRN focuses on co-occurrence relationships (see above), these effects are modulated by context—a word's meaning, like its syntactic category, emerges from the contexts in which it is used rather than from an a priori vocabulary of linguistic primitives.

Phonology. Pioneering work on emergentist *PHONOLOGY* was carried out by Donegan (1985), who noted the unhelpfulness to language learners of classic

distributional analysis. As she observed, it is implausible to suppose that children record sets of phonetic representations in memory and then compare them in the hope of determining which phonetic contrasts are distinctive and which are predictable from context (see *SPEECH PERCEPTION IN INFANTS* and *SPEECH PRODUCTION*). Instead, Donegan suggests, children begin with a set of processes (nasalization, devoicing, and so forth) that emerge as responses to the physical limitations of the human vocal tract and the auditory apparatus. (These limitations are inborn of course, but are not inherently linguistic in character, despite their linguistic consequences.) A language's phonemic inventory and allophonic patterns then emerge as specific processes are suppressed in response to experience.

A simple example involves the process that palatalizes /s/ in front of a high front vowel, giving the pronunciation [ʃi] for /si/ in many languages (e.g., Japanese). A child learning English is forced to suppress this process upon exposure to words such as *see*, which is pronounced [si], without palatalization. This in turn results in the admission of /ʃ/ to the phonemic inventory of English: because the palatalization process has been suppressed, the [ʃ] in words such as [ʃi] 'she' must be interpreted as a "sound in its own right" rather than as a process-induced variant of /s/.

Crucially, this conclusion is drawn without the need for comparison of minimal pairs or similar distributional analysis; the phonemic inventory emerges in response to a much simpler and more basic phenomenon—the suppression of processes based on exposure to particular individual words. Boersma (1998) and Hayes, Kirchner and Steriade (2004) discuss a broad range of other phonological phenomena from an emergentist perspective.

Concluding remarks

There is currently no comprehensive emergentist theory of language or its acquisition, but there are various emergentist-inspired research programs devoted to the construction of such a theory. For the most part, this work is based on the simple thesis that the core properties of language are best understood by reference to the properties of quite general cognitive mechanisms and their interaction with each other and with experience. The viability of this idea can and must be measured against its success in confronting the classic empirical challenges of linguistic analysis—figuring out how language works and how it is acquired.

Acknowledgments: Thanks to Kevin Gregg, Patrick Hogan, and an anonymous referee for helpful comments.

Works Cited and Suggestions for Further Reading

- Bates, Elizabeth and Brian MacWhinney. 1988. "What is Functionalism?" *Papers and Reports on Child Language Development* 27: 137-52.
- Bates, Elizabeth and Judith Goodman. 1999. "On the Emergence of Grammar from the Lexicon." *The Emergence of Language*, ed. B. MacWhinney, 29-79. Mahwah, NJ: Erlbaum.
- Bates, Elizabeth, Jeffrey Elman, Mark Johnson, Annette Karmiloff-Smith, Domenico Parisi and Kim Plunkett. 1998. "Innateness and Emergentism." *A Companion to Cognitive Science*, ed. W. Bechtel and G. Graham, 590-601. Oxford, UK: Blackwell.
- Boersma, Paul. 1998. *Functional phonology: Formalizing the Interactions between Articulatory and Perceptual Drives*. The Hague: Holland Academic Graphics.
- Bybee, Joan. 1998. "The Emergent Lexicon." *Proceedings of the 34th Regional Meeting of the Chicago Linguistic Society: The Panels*, 421-435. An excellent example of emergentist thinking about the lexicon.
- Donegan, Patricia. 1985. "How Learnable is Phonology?" *Papers on Natural Phonology from Eisenstadt*, ed. W. Dressler and L. Tonelli, 19-31. Padova, Italy: Cooperativa Libreria Editoriale Studentesca Patavina.
- Ellis, Nick. 2002. "Frequency Effects in Language Processing." *Studies in Second Language Acquisition* 24: 143-88.
- Elman, Jeffrey. 1993. "Learning and Development in Neural Networks: The Importance of Starting Small." *Cognition* 48: 71-99. A much cited and widely admired illustration of the value of computational modeling in the study of language acquisition.
- Elman, Jeffrey. 2002. "Generalization from Sparse Input." *Proceedings of the 38th Regional Meeting of the Chicago Linguistic Society*, 175-200. A highly readable summary of several important SRN-based studies of language acquisition.
- Elman, Jeffrey. 2004. "An Alternative View of the Mental Lexicon." *Trends in Cognitive Science* 8: 301-06.
- Elman, Jeffrey, Elizabeth Bates, Mark Johnson, Annette Karmiloff-Smith, Domenico Parisi, and Kim Plunkett. 1996. *Rethinking Innateness. A Connectionist Perspective on Development*. Cambridge, MA: MIT Press.
- Goldberg, Adele. 1999. "The Emergence of the Semantics of Argument Structure Constructions." *The Emergence of Language*, ed. B. MacWhinney, 197-212. Mahwah, NJ: Erlbaum.
- Goldberg, Adele and Ray Jackendoff. 2004. "The English Resultative as a Family of Constructions." *Language* 80: 532-68.
- Gregg, Kevin R. 2003. "The State of Emergentism in Second Language Acquisition." *Second Language Research* 19: 42-75.
- Hay, Jennifer. 2003. *Causes and Consequences of Word Structure*. New York: Routledge.
- Hay, Jennifer and R. Harald Baayen. 2005. "Shifting Paradigms: Gradient Structure in Morphology." *Trends in Cognitive Science* 9: 342-48. An excellent survey of work on emergentist morphology.
- Hayes, Bruce, Robert Kirchner and Donca Steriade (eds.). 2004. *Phonetically Based Phonology*. Cambridge, UK: Cambridge University Press.
- MacWhinney, Brian. 1998. "Models of the Emergence of Language." *Annual Review of Psychology* 49: 199-227.
- MacWhinney, Brian. 2002. "Language Emergence." *An Integrated View of Language Development: Papers in Honor of Henning Wode*, ed. P. Burmeister, T. Piske, and A. Rohde, 17-42. Trier: Wissenschaftliche Verlag.
- MacWhinney, Brian. 2004. "A Multiple Process Solution to the Logical Problem of Language Acquisition." *Journal of Child Language* 31: 883-914.

- MacWhinney, Brian. 2005. "The Emergence of Grammar from Perspective." *Grounding Cognition: The Role of Perception and Action in Memory, Language and Thinking*, ed D. Pecher and R. Zwaan, 198-223. Cambridge, UK: Cambridge University Press.
- Marcus, Gary. 1998. "Rethinking Eliminative Connectionism." *Cognitive Psychology* 37: 243-82.
- Marcus, Gary. 2001. *The Algebraic Mind*. Cambridge, MA: MIT Press.
- McClelland, James and Karalyn Patterson. 2002. "Rules or Connections in Past-Tense Inflection: What does the evidence rule out?" *Trends in Cognitive Science* 6, 465-72. An update and survey of connectionist work on inflection.
- Mill, John Stuart. 1930 [1843]. *A System of Logic Ratiocinative and Inductive*. London: Longmans, Green and Co.
- O'Grady, William. 2001. "An Emergentist Approach to Syntax." Available at <http://www.ling.hawaii.edu/faculty/ogrady/>. A summary of the detailed arguments for an emergentist theory of syntax found in O'Grady (2005).
- O'Grady, William. 2005. *Syntactic Carpentry: An Emergentist Approach to Syntax*. Mahwah, NJ: Erlbaum. An emergentist approach to syntax that seeks to understand many of the classic problems of syntactic theory in terms of processing.
- Palmer-Brown, Dominic, Jonathan Tepper, and Heather Powell. 2002. "Connectionist Natural Language Parsing." *Trends in Cognitive Science* 6: 437-42.
- Smolensky, Paul. 1999. "Grammar-Based Connectionist Approaches to Language." *Cognitive Science* 23: 589-613.
- Steedman, Mark. 1999. "Connectionist Sentence Processing in Perspective." *Cognitive Science* 23: 615-34.
- Tomasello, Michael. 2003. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press. A widely cited example of a usage-based approach to language acquisition.