Department of Psychology   Dietrich College of Humanities and Social Sciences

12-1-2006

Emergentism—Use Often and With Care

Brian MacWhinney
Carnegie Mellon University, macw@cmu.edu

Follow this and additional works at: http://repository.cmu.edu/psychology

Emergentism—Use Often and With Care

BRIAN MACWHINNEY

Carnegie Mellon University

Applied Linguistics has been receptive to input from a wide range of disciplines and perspectives. Researchers have explored theoretical and methodological inputs from linguistics, sociology, and psychology. Language planners have accepted ideas from sociology, political science, history, and demography. First and second language teachers have been open to all of these fields, as well as to approaches from education, rhetoric, and literature. This openness to new research and new ideas promotes innovation and dynamism in Applied Linguistics. At the same time, researchers and practitioners have often found themselves awash in a sea of conflicting claims and recommendations from these various theoretical sources. Should reading be taught by phonics or by the whole word method? Should we encourage second language learners to speak in the L2 from the first day or should we allow them to first consolidate their ability to learn from comprehensible input? Should we provide focused vocabulary practice or should vocabulary learning be subsumed in culturally relevant experiences? Should we provide explicit teaching of grammatical rules or rely instead on the student's ability to learn implicitly?

In the great Cervantes classic, Sancho Panza advises Don Quixote that, 'the proof of the pudding is in the eating.' Can we apply this adage to the task of evaluating the utility of theories of second language acquisition? I believe we can. Consider the situation in particle physics. Quantum dynamics is able to yield detailed predictions regarding atomic forces and collider events that are accurate to at least the eighth decimal place. For data collected by the high-energy collider, physicists rely on both string theory and quantum theory to provide accurate predications regarding the products emerging from the collision between two atoms of gold at 3 trillion degrees centigrade. In pools of heavy water buried deep in mountain caverns, physicists wait patiently for the appearance of rare varieties of quarks to provide evidence regarding the origins of the universe. Physicists are happy to pursue a variety of competing cosmological theories. But, in the end, they are committed to testing the worth of these theories in terms of quantitatively accurate models and validated experimental predictions. Although theories of language functioning will never attain this level of precision, we can still expect them to yield predictions that are accurate within a few percentages from observed data.

What is the current status of scientific prediction for second language learning? Consider first the status of predictions deriving from generative

grammar. This theory holds that first language learning is heavily determined by a species-specific ability that centers on the recursive nature of syntactic structure (Hauser *et al*. 2002). The current theory of grammatical minimalism (Chomsky 1995) has relegated large areas of language structure to the 'periphery' and reserved the 'core' for the essential elements of phrase structure, including X-bar theory, recursion, thematic role assignment, and the processes of unification. Elements assigned to the linguistic periphery include lexicon, phonology, morphophonology, dialect, idioms, constructions, discourse processes, associations, and processing load. In other words, virtually all of the substance of L2 learning lies outside of the purview of linguistic theory.

This means that we cannot consider generative grammar as a comprehensive theory of second language learning, although it could still have value as an account for some small subsection of second language learning. Thinking of generative grammar as a pudding that we can only consume and test on special holidays, it might still be great at those times.

However, the consumability of this restricted pudding is hampered by two additional theoretical commitments. The first is a commitment to the theory of parameter setting. This theory holds that first language learning is accomplished by the setting of values on a small set of parameters on the basis of evidence from 'triggers.' If learners were actually using triggers to set parameters, we would expect language development to move through a series of discontinuous and invariant steps. There is no evidence to support this prediction for either first or second language learning. The application of generativist ideas for second language acquisition is also blocked by a second theoretical commitment. This is the idea, articulated first by Lenneberg (1967), that L2 learning is not based on the natural processes underlying L1 learning. According to Lenneberg, there is a critical period for the learning of a first language that terminates at puberty with the attainment of full lateralization of language to the left hemisiphere. Many SLA researchers have accepted and extended this analysis. However, MacWhinney (2005a, 2005b) has shown how age-of-acquisition effects can be attributed not to the termination of a critical period, but to competition between an entrenched L1 and a weak and parasitic L2.

These various assumptions regarding parameter setting, critical periods, and the core–periphery distinction may be useful for work in generative theory. However, there seems to be no clear commitment in generative theory to providing quantitative empirical predictions regarding specific processes in L2 learning. For example, generative theory might characterize two languages as varying in their parameter settings, but this analysis provides no quantitative predictions regarding the course of L2 learning or for between-language interactions in learners. In effect, the work of producing quantitative accounts of L2 learning is turned over to the psycholinguist.

Against this backdrop, it is not surprising to see that researchers and teachers have turned to emergentist accounts of second language learning offers for empirical prediction. Emergentism is the most promising new trend in language studies. It offers new methods of prediction, new understandings of complex dialogic interactions, and new views of the learning process. It is having its impact in a wide range of subject areas, including child language, computational linguistics, phonology, conversation analysis, neurolinguistics, and the study of language–thought relations. Despite the enormous promise of this new trend, there are also dangers and limitations that can arise from misapplication of emergentist thinking. As Ellis and Larsen-Freeman explain so clearly, the forces determining second language attainment are immensely complex and their interactions are nonlinear and dynamic. Given these complexities, can we expect emergentist theory to attain a satisfactory level of empirical prediction regarding second language learning? I believe that the answer here is 'yes.' However, to attain this accurate level of prediction, emergentist explanations must be applied consistently, thoroughly, and carefully. In other words, one has to use emergentism often, but with great care. Let me explain what I mean by this.

First, I believe that emergentist thinking should be applied early, often, and consistently to all areas of thinking and practice regarding second language learning. In the end, emergentist thinking is basic to scientific explanation. The role of emergentist thinking in science can be illustrated through the examination of emergent patterns in the physical, geological, climatological, and biological worlds. Consider the shape of the bills of the 13 species of finches found by Darwin on the Galapagos Islands. Each finch has a bill shape that is adapted to the task of cracking and crushing the foods in their individual mini ecological niche. The insects and seeds eaten by the finches vary in distribution, size, and hardness. Each bill shape arose as an adaptation to these variations in local diet. Darwin's profound discoveries regarding the evolution of the species are grounded on emergentist thinking about these processes of mutation, selection, and adaptation. The study of the formation of shapes in crystals, proteins, schools of fish, cortical maps, termite mounds, snowflakes, and even leaves is grounded on emergentist thinking about packing constraints, molecular attractions, competitive interactions, and fractals. These forms arise not from supervision by some controlling architect, but through the local interaction between simple units. In Chemistry, local properties involve atomic size, valence, packing, and bonding forces. The higher-level properties of crystals, metals, fluids, and colloids all emerge from these local properties. Human physiology also represents an emergent response to specific pressures derived from exercise, aging, growth, and diet, as they impinge upon local units in a body that is continually regenerating itself on the basis of a remarkably stable set of instructions in the genetic code.

These are not isolated examples. In fact, scientific explanation in these areas cannot be separated from reference to emergent properties and

processes. Rather than viewing emergentism as opposed to some alternative viewpoint, it is best to view it as equivalent to basic scientific methodology. It is a mistake to view emergentism as residing at one pole of the traditional nature–nurture debate. Instead, emergentism shows how the long-term evolutionary timescale becomes meshed with the short-term developmental and interactional timescales by activities that occur in the present moment (MacWhinney 2005d). Some analysts have assumed that emergentist accounts derive all of language structure from general properties of cognition (Sabbagh and Gelman 2000). Although it is true that connectionists (Elman *et al.* 1996) have often derived emergence from general cognitive principles, connectionism is just one of the various sub-components of the broader emergentist framework. More fully embodied and grounded emergentist accounts (Pecher and Zwaan 2005; Thelen and Smith 1994) hold that linguistic form arises from interactions between the shape of the body, cognitive processing, and the nature of social interaction across a wide variety of timescales.

In new fields, such as the study of second language acquisition, there is always an initial formative period during which the emphasis is on the collection and systematization of observational data. However, as this systematization progresses and the database builds, the pressure for mechanistic explanation grows. Initial mechanistic explanations are typically grounded on simple cause–effect relations (Bechtel and Richardson 1993). However, as the field advances, explanations turn increasingly to feedback loops, catalytic processes, and complex dynamics. In each case, the operation of the complex system must be articulated mechanistically. In the study of second language learning, emergentist explanations must explain where a linguistic behavior comes from. It is not enough to point to the complexity of some linguistic behavior and to declare that it must be emergent. We may be able to agree that all patterns are emergent in some way. But, from what do they emerge and how?

This concern brings me to the second part of my title, which is that we must use emergentism with care. It is easy to come up with emergentist accounts that seem plausible, but which are simply wrong. Consider the attempt by Lieberman (1975) to link the emergence of language to the descent of the larynx as evidenced in endocast reconstructions from fossils in Africa and the Levant at about 100,000 years ago. According to Lieberman's application of the source-filter model of speech production, the descent of the larynx contributed to the bending of the vocal tract thereby facilitating a clearer articulation of the vowel space. Although much of Lieberman's analysis makes sense, the idea that the bent vocal tract emerges from a descent of the larynx is problematic. As Ohala (2000) has noted, females do not display a bent vocal tract and males do not show the descent of the larynx until puberty. However, women's articulatory abilities are certainly equal to those of men. Moreover, in other species, such as red deer and whooping cranes, laryngeal descent has its onset with puberty and is used to

mark male dominance and threat displays. It appears that the descent of the larynx has more to do with male courtship display than it does with achieving precision of vowel articulation. Both accounts of laryngeal descent rely on evolutionary emergentist reasoning. However, one of these accounts seems to match the observed distribution of adaptational patterns better than the other.

The articulation of emergentist accounts depends on strong methodological support. Because emergentist accounts emphasize complex interactions between multiple factors across multiple time scales, they rely heavily on the powerful computational methods introduced by the digital revolution. Without the enormous recent explosion in computational power and usability, the recent flowering of emergentism would not have been possible. The paper in this volume by Meara illustrates how the rise of neural network simulation has opened up avenues for emergentist thinking about second language learning and language evolution. Ke and Holland go further and link neural network modeling to agent-based modeling (ABM). Their method examines ways in which embodied computer agents (such as robots) can acquire language from complex situated input (Marocco *et al.* 2003). The paper by Larsen-Freeman shows how we can use computerized databases, graphing, and statistics to track complex patterns of variation in learners. The papers by Mellow and Cameron and by Deighan illustrate the impact of large new computerized corpora such as CHILDES and TalkBank (MacWhinney *et al.* 2004) on theories of language learning and functional linguistic analysis.

We are now at the beginning of a technological revolution that will illuminate the study of emergent processes in L2 even more powerfully. This technology involves the linking of desktop databases, statistics, and multimedia to interactive communal computing over the web. The TalkBank Project (http://talkbank.org) is now making available several large databases on second language learning. These databases include transcripts of learners' written productions, as well as spoken productions linked to audio and/or video. In addition, the CHILDES database (http://childes.psy.cmu.edu) now includes 12 in-depth longitudinal studies of childhood bilingualism, along with additional cross-sectional data. As these databases grow, we will develop increasingly powerful data-mining methods that will allow us to observe variation and consistency across time in classroom and naturalistic learning. Complementing this rich naturalistic data on learning processes, we are now also creating new databases that study in complete detail students' interactions with online computer tutors. Within the context of online courses in French and Chinese, developed by the Pittsburgh Science of Learning Center (http://learnlab.org), we are now able to track whole terms of online student responses to vocabulary training, pinyin dictation for Chinese, tone discrimination, sentence repetition, and other methods for training second language skills. The goal here is to apply detailed computational models (MacWhinney *et al.* 1989; Pavlik and Anderson 2005)

to the exploration of the emergence of second language skills in online exercises and to track the generalization of these skills to face-to-face interaction.

With these various prospects and caveats in mind, let take a look at the six contributions to this special issue. How well do these analyses succeed in generating precise predictions for patterns in language learning? Can we use these predictions to improve language learning?

Consider first Paul Meara's use of neural networks to explore bilingual lexical processing. Meara focuses on patterns of activation in simple Boolean networks built up from units that are either on or off. Each unit receives two inputs. If the unit is an AND unit, both units must be on for it to fire. If the unit is an OR unit, it will fire when either unit is on. Meara then programs the network with items in both L1 and L2. Most of the L1 items are connected only to other L1 items and most of the L2 items are connected only to other L2 items. However, a few items are entangled by connections with the other language. These connections are set up so that entangled words are activated if only one of their inputs is activated, but become deactivated if both their inputs are activated. When the network is given an initial wave of random activation, it settles into a steady state in which two-thirds of the dominant L1 forms are active and only 10 per cent of the weaker L2 forms are active. If the L2 words are forced on across several iterations of activation, eventually the activations of the L1 words are decreased.

This pattern of competitive activation provides a fascinating emergentist account of the switching of attention from a dominant L1 to a weaker L2 in bilinguals. It appears that this effect arises primarily from the way in which the entanglement was designed. Input from two languages will lead to deactivation of entangled items and removal of activation from the items to which they then connect. This will serve, over time, as a damper on coactivation of the two languages. Simple local mechanisms of this type can lead to global effects that have otherwise been described in terms of strategic language switching.

Meara also shows how the interference of L2 words with L3 forms can arise from entangled networks. For this simulation, L1 is entangled with L2 and L3, but L2 and L3 are not mutually entangled. This form of entanglement seems to correctly represent the claims of models such as the Revised Hierarchical Model (Kroll and Stewart 1994). Meara shows that forcing of activation in L3 can lead eventually to destabilization of the pattern of activation in the unconnected items of L2. What happens here is that the instability caused by forcing L3 activation spreads first to L1 and then to L2. Since L2 is inherently less stable than the dominant L1, it is the system that shows the greatest intrusions. Meara suggests, reasonably enough, that these instabilities are reflected in the L2 intrusions to L3 reported by bilingual speakers. Again, we see evidence for the emergence of patterns based on the operation of local connections without conscious control.

In his survey of the literature on the development of the bilingual lexicon, Meara notes that the DEVLEX model of Li, Farkas, and MacWhinney (2004) relies on an architecture that is far more complex that the simple Boolean networks he has been exploring. In fact, the basic units in the Kohonen self-organizing feature map operate in a localist way that is not radically different from the units in the Boolean network. A greater difference between the two systems is that, in order to link up with actual input data from the CHILDES database, the DEVLEX system is supplemented with a system that compresses co-occurrence data derived from maternal input. Because the input to DEVLEX is a faithful representation of input to the child, we can treat interactions in the network for specific words as representing plausible real effects. It would be interesting to extend this approach to Meara's problem. In particular, Meara is currently forced to assign entanglement to individual lexical items on a random basis. However, with real bilingual input of the type available in the bilingual corpora in CHILDES, we could distinguish forms that are truly entangled in the input from forms that are not. This would then allow us to understand in better detail exactly what entanglement entails.

Jinyun Ke and John Holland show how emergentist forces can operate in a similar fashion in a very different domain. Using agent-based modeling (ABM), they show how word order patterns can emerge within a communicative system based on initially unstructured, holistic signs. In order to get this system off the ground, Ke and Holland need to make some basic assumptions about the representation of meaning. Fortunately, none of these assumptions are controversial. One assumption is particularly central to the operation of the model. This is the idea that early signals can be structured either holistically or by rule. This contrast between rote and combination was at the core of the model of language acquisition proposed in MacWhinney (1978) and repeated in Pinker (1984). This earlier account (MacWhinney, 1975, 1982, 1987) provided detailed mechanistic proposals about how word order patterns can be extracted from item-based patterns. The Ke and Holland model allows for random generation of both types of communication on each trial. Despite this questionable assumption, the model ends up with the standard developmental pattern of rote preceding combination (MacWhinney 1978). In this case, the preference for early rote arises from the needs for communication between agents. In the case of the child, it may also depend on the relation between the processes that segment out items from the input and those that store forms for later production. In both cases, the developmental sequences emerge from properties in the data stream.

Dean Mellow brings emergentist thinking to bear on the issue of the 'poverty of the stimulus,' which has also been called 'the logical problem of language acquisition.' Other recent emergentist approaches to this problem (MacWhinney 2005c; Pullum and Scholz 2002) have focused their attention on first language acquisition, but Mellow is the first emergentist to tackle this

problem in the context of second language acquisition. At the core of his analysis is the emergentist theory of item-based learning. This view of learning emphasizes the extent to which higher-order constructions emerge from the comparisons and combinations of individual lexical items (MacWhinney 1982). Following the analysis of O'Grady (2005), Mellow shows how a combination of item-based processing with syntactic incrementalism can allow us to understand the observed patterns of learning difficulties for embedded clauses in second language acquisition.

The theory of item-based learning was first offered by Braine (1963) as a component of the theory of pivot grammar. Later, both Braine (1976) and MacWhinney (1975) revised the theory to include a semantic characteriza-tion of argument–head relations. MacWhinney (1982) then showed how generalization could extract constructions from groups of item-based patterns. However, the transition from items to constructions was not explored in empirical detail until recent work by Tomasello (2000) and colleagues. Several details in this theory are still being articulated (MacWhinney 2004a), and the application of item-based learning to embedded structures is one such area. Here, Mellow carefully distinguishes the various forms of items involved in DVCs, such as 'that' and 'which.' Just as shown earlier for child language (Maratsos *et al*. 1979), Mellow shows that these items and their attendant constructions are acquired first with particular verbs and then applied more generally across verbs. In addition, Mellow shows that simpler DVC constructions appear before similar, but more complex, DVC constructions. The important point here is that Mellow is able to provide a detailed account of complexity levels in terms of O'Grady (2005). In O'Grady's account, complexity emerges from basic principles of lexical combination. As a result, Mellow can rely on this emergent property to account for the order of acquisition of DVCs in his second language learner.

Responding to a reviewer's critique, Mellow notes that some UG approaches have begun to take a similar approach. An interesting recent example of a hybrid UG/learning model is the APT theory of Truscott and Sharwood Smith (2004). However, as several reviewers noted in their review of this model (Carroll 2004; MacWhinney 2004b; Pienemann 2004), once a UG account adopts core emergentist assumptions such as statistical tracking, working memory constraints, item-based grammar, and competition, it becomes largely equivalent to the more fully elaborated Competition Model account.

Diane Larsen-Freeman presents an emergentist approach to two central issues in SLA research—the nature of continuity and individual differences in L2 learning. She calls into question the idea that the learner progresses through a series of well-defined stages that incrementally approximate the adult target language. Her arguments against this view are convincing. It is clear that language is not a fixed target. At each time slice during learning, learners are fully competent social participants who can configure limited

linguistic resources in many creative and non-standard ways. Faced with this complexity, her analysis predicts wide individual variation among learners and complex and discontinuous developmental trajectories. Examining written data from five Chinese learners of English, this is exactly what she finds. Her quantitative analysis shows how gradual learning on the group level masks radical discontinuity and variation on the individual level. Her qualitative analysis of the writing samples further underscores the spiraling nature of learning. In these spirals, learners advance new constructions at one point, withdraw from using them, but then apply them again later with increased clarity and precision.

Larsen-Freeman's dynamic view of second language acquisition is fundamentally emergentistic. She correctly notes the fundamental similarity between individual variation in L2 learning and the better-documented processes of individual variation in the early stages of L1 learning (Bates *et al*. 1979; Nelson 1981). Understanding the non-linear nature of this variable process is absolutely critical to understanding SLA and shaping effective SLA pedagogy. However, it seems to me that Larsen-Freeman's analysis could be strengthened through additional attention to two of the other core features of emergentist analysis. The first of these is an attention to the mechanisms of emergence. As I have argued elsewhere, emergentist theories must tell us where a given language behavior comes from. In some cases, the component pieces may be relatively obvious. We can note how Larsen-Freeman's Participant Y progressively changes the way in which she explains how she got her marriage certificate. At first, she says they had been dating for five years. But on retelling, she realizes that they had only been dating for four years. So, this dimension changes on retelling. But what causes this change? My guess is that the very process of retelling is what promotes the reorganization. By October, when she tries to combine the notion of dating for four years with getting the marriage certificate, she ends up confusing telicity and saying 'my husband and I got our marriage certificate for about four years.' Does she realize the error here? Perhaps not, but by November, she is able to break the story down into clear sequential segments that no longer trigger this telicity confusion. This type of learning seems to most clearly resemble Siegler and Shipley's (1995) wave model of cognitive development which was formulated specifically as an alternative to the standard Piagetian stage theory. I very much agree with Larsen-Freeman's emphasis on discontinuities and other waves, but it seems to me that we need to explore the mechanisms that generate each of these waves.

Larsen-Freeman emphasizes the extent to which group data can mask variation on the individual level. Denser samplings from her individual learners would remove some of this variation, but much will remain. We would always expect individual learner data to be more variable than group data. But does this mean that emergentist models should ignore group data? It seems to me that the problem here is that we need to distinguish clearly between the alternative functions of individual and group data for

emergentist theory. As Larsen-Freeman notes, individual data provide us
with a window into the complex world of interacting constraints in the
learner. Group data, on the other hand, provide us with a window toward
society and the distribution of the linguistic input. A great deal of emergentist
thinking derives learning from statistical regularities and co-occurrences in
corpora. This external input is semi-stable across the lifetime of the learner
(Labov 2001; MacWhinney 2005d). Therefore, measurements across large
samples and groups will tend to pick up the effects of these inputs as
approximations toward a target grammar. Thus, it seems to me that a full
emergentist account can and should coordinate Larsen-Freeman's emphasis
on individual variation with models that predict gradual approximation
toward a target grammar.

   Lynne Cameron and Alice Deighan present a corpus-based analysis of
the emergence of metaphoremes. These are patterns of semantic extension
involved in phrases such as 'emotional baggage,' 'bear the weight of guilt,'
'walk away from responsibility,' or 'pay a high price for a crime.' Like others
who have spent long hours examining expository narrative corpora,
Cameron and Deighan have come away impressed by the immense
amount of metaphorical language in this genre. In line with recent work
in embodied cognition (Narayanan 1997), Cameron and Deighan see these
metaphors as grounded on both social and perceptual experience. Many of
the issues that Cameron and Deighan raise remain unanswered at the end of
their provocative paper. Despite this, I am tempted to raise still further issues.
I am particularly interested in understanding how we can relate the
perceptual-motoric grounding of terms such as 'baggage' or 'walk away from'
to their social grounding. If all peoples inhabit a basically similar perceptual
world, then why do languages map the core physical processes so differently?
Why do some languages describe anger as a pressure that builds up and
others as a poison? Why do some languages treat responsibility as a trap and
others as a mere location? At the same time, the idea of stumbling as a
failure in communication may be remarkably common across languages.
The emergentist account of metaphoremes will need to explore these cross-
linguistic patterns in great detail. Given these similarities and differences,
should we expect L2 learners to have problems acquiring metaphoremes in a
new language? What about L1 learners? Would we expect to see individual
differences of the type predicted by Larsen-Freeman? Could we use the
theory of item-based learning proposed by MacWhinney (1989) and others
to account for regularities and irregularities in this process?

   I hope that these brief comments on the five papers in this issue have
served to illustrate the basic points with which I began this commentary.
First, emergentism is not just an alternative to UG. It is a method that is basic
to all of science, because nature is full of emergent processes. This means that
we must apply emergentist thinking widely and often. However, and this is
my second point, it is easy enough to come up with emergentist accounts
that are appealing, but wrong. By itself, emergentism is no magic bullet.

We must apply it with caution and discipline. Emergentist thinking provides general guidelines for studying the mechanisms generating complex phenomena. It is the responsibility of the individual researcher to apply these guidelines to specific cases. Thirdly, and this may be the most important point, the elaboration of emergentist thinking and experimentation will depend increasingly on powerful new advances in computation and complexity theory. In order to tease apart the generators of complexity, we will need to continue studying corpora, multimedia databases, neural networks, individual subject growth curves, online language learning, and the details of cognitive neuroscience. We have our work cut out for us. The proof of the pudding is in the eating.

Final version received June 2006

REFERENCES

Bates, E., L. Benigni, I. Bretherton, L. Camaioni, and **V. Volterra.** 1979. *The Emergence of Symbols: Cognition and Communication in Infancy*. New York: Academic Press.

Bechtel, W. and **R. C. Richardson.** 1993. *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. Princeton, NJ: Princeton University Press.

Braine, M. D. S. 1963. 'The ontogeny of English structure: The first phase,' *Language* 39: 1–13.

Braine, M. D. S. 1976. 'Children's first word combinations,' *Monographs of the Society for Research in Child Development*, 41, (Whole No. 1).

Carroll, S. 2004. 'Acquisition by processing theory: A theory of everything?' *Bilingualism: Language and Cognition* 7: 23–25.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Elman, J. L., E. Bates, K. Plunkett, M. Johnson, and **A. Karmiloff-Smith.** 1996. *Rethinking Innateness*. Cambridge, MA: MIT Press.

Hauser, M., N. Chomsky, and **T. Fitch.** 2002. 'The faculty of language: What is it, who has it, and how did it evolve?' *Science* 298: 1569–79.

Kroll, J. F. and **E. Stewart.** 1994. 'Category interference in translation and picture naming: Evidence for symmetric connections between bilingual memory representations,' *Journal of Memory and Language* 33: 149–74.

Labov, W. 2001. *Language Change. Vol. 2: Social Considerations*. London: Blackwells.

Lenneberg, E. H. 1967. *Biological Foundations of Language*. New York: Wiley.

Li, P., I. Farkas, and **B. MacWhinney.** 2004. 'Early lexical development in a self-organizing neural network,' *Neural Networks* 17: 1345–62.

Lieberman, P. 1975. *On the Origins of Language: An Introduction to the Evolution of Human Speech*. New York: Macmillan.

MacWhinney, B. 1975. 'Pragmatic patterns in child syntax,' *Stanford Papers and Reports on Child Language Development* 10: 153–65.

MacWhinney, B. 1978. 'The acquisition of morphophonology,' *Monographs of the Society for Research in Child Development* 43: Whole no. 1, pp. 1–123.

MacWhinney, B. 1982. 'Basic syntactic processes' In S. Kuczaj (ed.): *Language Acquisition: Vol. 1. Syntax and Semantics*. Hillsdale, NJ: Lawrence Erlbaum, pp. 73–136.

MacWhinney, B. 1987. 'The Competition Model' In B. MacWhinney (ed.): *Mechanisms of Language Acquisition*. Hillsdale, NJ: Lawrence Erlbaum, pp. 249–308.

MacWhinney, B. 1989. 'Competition and lexical categorization' in R. Corrigan, F. Eckman, and M. Noonan (eds): *Linguistic Categorization*. Philadelphia: Benjamins, pp. 195–242.

MacWhinney, B. 2004a. 'A multiple process solution to the logical problem of language acquisition,' *Journal of Child Language* 31: 883–914.

MacWhinney, B. 2004b. 'Parameters or cues?' *Bilingualism: Language and Cognition* 7: 35–36.

MacWhinney, B. 2005a. 'A unified model of language acquisition' in J. F. Kroll and

A. M. B. de Groot (eds): *Handbook of Bilingualism: Psycholinguistic Approaches*. New York: Oxford University Press, pp. 49–67.

MacWhinney, B. 2005b. 'Emergent fossilization' in Z. Han and T. Odlin (eds): *Studies Fossilization in Second Language Acquisition*. Clevedon, UK: Multilingual Matters, pp. 134–56.

MacWhinney, B. 2005c. 'Item-based constructions and the logical problem,' *ACL* 2005: 46–54.

MacWhinney, B. 2005d. 'The emergence of linguistic form in time,' *Connection Science*: 191–211.

MacWhinney, B., S. Bird, C. Cieri, and **C. Martell.** 2004. 'TalkBank: Building an open unified multimodal database of communicative interaction' in *LREC 2004*. Lisbon: LREC, pp. 525–8.

MacWhinney, B., J. Leinbach, R. Taraban, and **J. McDonald.** 1989. 'Language learning: Cues or rules?' *Journal of Memory and La*nguage 28: 255–77.

Maratsos, M., S. A. Kuczaj, D. E. Fox, and **M. A. Chalkley.** 1979. 'Some empirical studies in the acquisition of transformational relations: Passives, negatives, and the past tense' in W. A. Collins (ed.): *Children's Language and Communication*. Hillsdale, N.J.: Lawrence Erlbaum.

Marocco, D., A. Cangelosi, and **S. Nolfi.** 2003. 'The emergence of communication in evolutionary robots,' *Philosophical Transactions of the Royal Society of London A*, 361: 2397–421.

Narayanan, S. 1997. 'Talking the talk is like walking the walk,' *Proceedings of the 19th Meeting of the Cognitive Science Society*: 55–59.

Nelson, K. 1981. 'Individual differences in language development: Implications for development and language,' *Developmental Psychology* 17: 170–87.

O'Grady, W. 2005. *Syntactic Carpentry*. Mahwah, NJ: Lawrence Erlbaum Associates.

Ohala, J. 2000. The irrelevance of the lowered larynx in modern Man for the development of speech. Paper presented at The Evolution of Language—ENST, Paris.

Pavlik, P. and **J. Anderson.** 2005. 'Practice and forgetting effects on vocabulary memory: An activation-based model of the spacing effect,' *Cognitive Science* 29: 559–86.

Pecher, D. and **R. Zwaan** (eds). 2005. *Grounding Cognition*. Cambridge: Cambridge University Press.

Pienemann, M. 2004. 'Processing perspectives in SLA research and their compatibility,' *Bilingualism: Language and Cognition* 7: 37–9.

Pinker, S. 1984. *Language Learnability and Language Development*. Cambridge, MA: Harvard University Press.

Pullum, G., and **B. Scholz.** 2002. 'Empirical assessment of stimulus poverty arguments,' *Linguistic Review* 19: 9–50.

Sabbagh, M. and **S. Gelman.** 2000. 'Buzzsaws and blueprints: What children need (or don't need) to learn language,' *Journal of Child Language* 27: 715–26.

Siegler, R. S. and **C. Shipley.** 1995. 'Variation, selection, and cognitive change' in G. Halford and T. Simon (eds): *Developing Cognitive Competence: New approaches to process modeling*. Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 31–76.

Thelen, E. and **L. Smith.** 1994. *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge, MA: MIT Press.

Tomasello, M. 2000. 'The item-based nature of children's early syntactic development,' *Trends in Cognitive Sciences* 4: 156–63.

Truscott, J. and **M. Sharwood Smith.** 2004. 'Acquisition by processing: A modular perspective on language development,' *Bilingualism: Language and Cognition* 7: 1–20.