

## The emergence of linguistic form in time

BRIAN MACWHINNEY\*†

†Carnegie Mellon University, Pittsburgh, PA, 15213, USA

Linguistic forms are shaped by forces operating on vastly different time scales. Some of these forces operate directly at the moment of speaking, whereas others accumulate over time in personal and social memory. Our challenge is to understand how forces with very different time scales mesh together in the current moment to determine the emergence of linguistic form.

*Keywords:* Emergence; Word learning; Syntax; Phylogeny; Ontogeny; Resonance

### 1. Introduction

Where do linguistic forms come from? Researchers have proposed accounts based on rules, conventions, constraints, statistical regularities, discourse patterns, language contact, neurology, physical dynamics and genes. What is common to all of these forces is that they must all make their final impact at the moment of speaking. Some of these forces operate within a narrow time frame. For example, when we run up a flight of stairs, there will be an immediate impact on vocal production leading to loss of weak syllables and changes in pitch. This type of exertion has a quick and immediate effect on prosodic and phonetic form. Other processes have a slower time fuse. For example, we may not have made reference to some concept, such as ‘apotheosis’ or ‘ligature’ for years. Yet, when we need to use these words, we can dredge them up out of our memories and apply them in the moment. The effects of Alzheimer’s disease on language can work across decades, leading eventually to an inability to remember even the most common words. Languages themselves undergo changes at a truly glacial pace. For example, the shift in Old English from subject-object-verb (SOV) to subject-verb-object (SVO) word order took centuries to complete.

These various forces constitute dynamical systems operating with very different time frames and histories. What binds all of these systems together is the fact that they must all *mesh* in the current moment. One simple view of the process of meshing is that cues combine in an additive manner (Massaro 1987) and that systems are partially decomposable (Simon 1969). However, psycholinguistic work in the Competition Model framework (MacWhinney in press) has shown that simple cue summation applies only when decisions are not made online. For online decisions, cues interact in far more complex ways, determined by the periodicity linked to their operation. For example, the advantage of first mention will maintain its effect across

---

\*Email: macw@mac.com

the whole sentence, whereas the shift in reference triggered by an implicit causality verb such as criticize will vanish within 200 ms (McDonald and MacWhinney 1995).

The meshing of forces in the moment has two major implications. First, if we want to understand and model specific uses of linguistic forms, we have to understand how diverse forces mesh in the moment. This will allow us, for example, to watch a mother–child interaction and understand the ways in which scaffolding events build on mutual reference, attention, memory and prior interactions. Second, studies of meshing will allow us to understand nonlinearities that emerge from dynamic cue interactions. Third, studies of meshing will allow us to understand how linguistic form emerges from interactions between multiple processes meshing in the moment.

## 2. Varieties of emergence

To begin this exploration, it may be useful to review some examples of emergence across time in the social and natural world. Consider the case of the forces determining the length of checkout lines at a supermarket. Here, you will find that the number of people queuing up in each line stays roughly the same. There are rarely six people in one line and two in the next, unless there is a line with special rules. There is no socially articulated rule governing this pattern. Instead, the uniformity of this simple social structure emerges from other basic facts about the goals and behaviour of shoppers and supermarket managers.

Honey-bees are certainly no smarter than shoppers. However, working together, bees are able to construct an even more complex structure. When a bee returns to the hive after collecting pollen, she deposits a drop of wax-coated honey. Each of these honey balls is round with approximately the same size. As these balls are packed together, they take on the familiar hexagonal shape that we see in the honeycomb. There is no gene in the bee that codes for hexagonality in the honeycomb, nor any overt communication regarding the shaping of the cells of the honeycomb. Rather, this hexagonal form is an emergent consequence of the application of packing rules to a collection of honey balls of roughly the same size.

Nature abounds with examples of emergence. The shapes of crystals emerge from the ways in which atoms can pack into sheets. Alternative crystalline lattice (cubic, hexagonal, monoclinic, orthorhombic) structures emerge from the relative size of the atoms in the ionic compound. The outlines of beaches emerge from interactions between geology and ocean currents. Consider the shape of Cape Cod near Provincetown, where the northeasterly drift of the Gulf Stream works to push the outline of the cape toward the mainland. Weather patterns like the Jet Stream or El Niño emerge from interactions between the rotation of the earth, solar radiation and the shapes of the ocean bodies. Biological patterns emerge in very similar ways. For example, the shapes of the spots on a leopard or the stripes on a tiger emerge from the timing of the expression of a pair of competing genes expressing colour as they set up standing waves governed by B-Z (Belovsov–Zhabotinsky) equilibria across the developing leopard or tiger embryo (Murray 1988). No single gene directly controls these patterns. Rather, the stripes emerge from the interactions of the genes on the physical surface of the embryo. The shape of the brain is very much the same. For example, Miller *et al.* (1989) have shown that the ocular dominance columns described by Hubel and Weisel (1963) can emerge from the competition between projections from the two optic areas during synaptogenesis in striate cortex.

In 1794, Huygens demonstrated that two pendulums moving at different periods would couple together to find a single periodicity if they were mounted on a board with springs. During this coupling, one pendulum serves as the strong attractor that entrains the other pendulum to its periodicity. This form of resonant coupling also occurs within language. For example,

studies of the mechanics of infant babbling have demonstrated that there is an early period when the child moves the jaw with a consistent rhythm (MacNeilage 1998). During babbling, the periodicity of this movement then serves to entrain a similar periodicity in the opening and closing of the glottis. The result of this coupling is the emergence of canonical babbling (Vihman 1996).

The idea of interaction between levels of emergence is a familiar one in sciences like biology and physics. Our biological existence is grounded on the operations of thousands of proteins, each with a subtly different geometry, determined on four emergent levels. The primary structure of a protein is determined by its sequence of amino acids, which is in turn a function of the order of amino acids in a codon of DNA. This is the structure that is most tightly linked to evolution and natural selection. The secondary structure of proteins involves coils, fold and pleats that arise from the formation of hydrogen bonds between CO and NH groups along the polypeptide backbone. Tertiary structure, leading to the folding of single polypeptides, derives from hydrophobic interactions and disulphide bridges that produce bonding between side chains. Quaternary structure emerges from the aggregation of polypeptide subunits, as in the combination of four subunits in haemoglobin. Altogether, 'the specific function of a protein is an emergent property that arises from the architecture of the molecule' (Campbell *et al.* 1999).

Fusion physics (McDonald and MacWhinney 1995) provides a very different example of coupled interactions across divergent temporal frames. In fusion research, the goal is to contain a plasma composed of deuterium and tritium under sufficient pressure and heat to force individual ions to fuse into a helium atom with an extra neutron. If this can be achieved for a sufficient number of atoms in the plasma during a period of time of a fraction of a second, ignition is obtained and the plasma will produce enough energy to continue the reaction. To work within this time frame, atoms must be forcefully contained in a way that can overcome repulsion from Coulomb forces. Once the Coulomb barrier is passed, the resulting fusion occurs in nanoseconds. In the classical model of about 1954, this should be achieved by simple confinement and collision in a toroidal tokamak. However, work with the toroid soon showed that electrons and protons generated banana-shaped poloidal drifts at two separate periodicities, both of which fought against confinement. To battle these emergent forces, physicists added an electric field toroidally. The addition of this field has produced additional emergent forces in super-banana patterns. Fortunately, these additional forces operate at a time frame that will not interfere with containment. The emergence of each of these additional forces in the tokamak illustrates further the ways in which a tightly coupled system such as a tokamak can produce a complex set of forms operating at different temporal frames.

### 3. Seven time frames

We can distinguish seven markedly different time frames for emergent processes and structures (Lorenz 1958).

- (1) *Phylogenetic emergence.* The slowest moving emergent structures are those that are encoded in the genes. Changes across this time frame—which involves millennia rather than minutes—are controlled by natural selection (Darwin 1871). The core engine of emergence is the generation of variation through mutation, followed then by natural selection through both mate choice and differential mortality. Natural selection utilizes the possibilities for reorganization shaped by the DNA and the interactions of polypeptides that it specifies. The unevenness of this underlying landscape makes some mutations more probable and frequent than others, leading to a reliance on the reuse of old forms to serve

new functions. Emergentist accounts in this area have emphasized the ways in which language, society and cognition have undergone co-evolution (MacWhinney 2002) based on the linking of dynamic systems. To trigger this co-evolutionary advantage, changes in linguistic abilities must arise in parallel with advances in cognitive or social abilities. Moreover, both effects must interact at the moment of speaking. When this happens in a way that favours reproductive fitness, the mutation will be preserved.

- (2) *Epigenetic emergence.* The codification of information in the DNA represents a precise meshing between the slow moving process of evolution and the faster-moving process of epigenesis (Waddington 1957). Embryologists have shown that biological structures emerge from processes of induction between developing tissue structures in the embryo. The shape of these interactions is not hard-coded in the DNA. Instead, the DNA encodes information that can push the process of differentiation in particular directions at crucial epigenetic choice points. The precursors of autism in the embryo can be traced to particular epigenetic effects, as can the formation of stripes in the tiger. Epigenetic emergence does not cease at birth. To the degree that the brain maintains a level of plasticity, epigenetic processes allow for recovery of function after stroke through rewiring and reorganization. Before birth, epigenetic interactions with the environment are confined to forces that impinge on the uterus and the embryonic fluid. After birth, the environment can trigger a wide variety of variations in gene expression, from diabetes to brain reorganization for language in the deaf (Bellugi *et al.* 1989).
- (3) *Developmental emergence.* Jean Piaget's (1954) genetic psychology was the first fully articulated emergentist view of development. Impressively complete in its coverage, it failed to specify details regarding mechanisms of development. To provide this missing mechanistic detail, current emergentist accounts of development rely on connectionism, embodiment and dynamic systems theory (Quinlan 2003). Emergentist theory has been used to characterize two different, but interrelated, aspects of development. The first is the basic learning process that involves the continual learning of new facts, forms, relations, names and procedures. Basic models of language learning, such as those that deal with learning of the past tense (MacWhinney and Leinbach 1991), often focus on this type of development. A second type of development involves the learning of new strategies and frameworks that can alter the overall shape of language and cognition, often through cue focusing (Colunga and Smith 2000, Regier 2003). Work linking these two has just begun.
- (4) *Processing emergence.* The most fast-acting pressures on language form are those that derive from online processing constraints (MacWhinney 1999). These pressures are activation of memory mechanisms, attentional focusing, co-ordination of sentence planning, code switching between languages and motor control. Many of these pressures are themselves driven by long-term processes. For example, a child's failure to understand the meaning of the word 'dependability' in a discussion of the reliability of batteries may be the result of problems in understanding previous classroom and computerized lessons on numerical distributions. Similarly, the failure in lexical retrieval that occurs in aphasia is driven by changes to neural tissue subsequent to a stroke. Thus, online processing emergence can reflect the current status of longer-term developmental, neuronal and physiological processes.
- (5) *Social emergence.* Many of the pressures that operate during face-to-face conversations derive from long-term social commitments. Our choice of vocabulary, slang, topics and even language is determined by the nature of the relation we perceive with the people we meet. We can select these options to emphasize solidarity, impose our power, or seek favours. The time course of these social commitments is often measured in terms of years or decades. Some basic social commitments, including those forced by gender and race, can never be fundamentally altered.

- (6) *Interactional emergence*. Apart from our long-term commitments to dialects, languages and subgroup themes, we also make more short-term commitments to ongoing social interactions. For example, we may engage a real estate agent to help us purchase a house. Our linguistic interactions with this agent are then shaped by the current status of the buying process. Even after we terminate one set of transactions with this agent, we will maintain an ongoing relation that will then shape our further interactions, days or weeks later (Keenan *et al.* 1977).
- (7) *Diachronic emergence*. We can also use emergentist thinking to understand the changes that languages have undergone across the centuries (Bybee and Hopper 2001). These changes emerge from a further complex interaction of the previous three levels of emergence (evolutionary, developmental and online).

With these seven time frames in mind, we can construct a revised interpretation of the traditional question ‘Is it innate or learned?’ What this question really means is ‘Across what time frame does this ability emerge?’

Within this general framework, we can further distinguish four ways in which emergent processes impact on linguistic form.

- (1) *Homeorhesis*. Biological systems depend heavily on homeorhetic systems for the preservation of life. These systems maintain balance for hormone levels, ion transport, metabolites, immune functioning and cell growth. In language, homeorhesis operates on physiological, neurological and social levels. Conversation itself can be viewed as a homeorhetic process.
- (2) *Resonance*. The transfer of information between systems can emerge from coupling, entrainment and resonance. In language, multilingual processes can be viewed as emerging from within-language resonance (MacWhinney *in press*). When we code-switch from English to Spanish, the initial moments of speaking in Spanish are still under the influence of resonance operating in English (Grosjean and Miller 1994).
- (3) *Loose coupling*. Systems that operate phylogenetically and epigenetically are typically only loosely coupled with their outcomes. In language, patterns of genetic inheritance often have only a very indirect effect on language functioning. For example, the KE family studied by Gopnik and Crago (1990) has motor dysfunctions that impact language production. However, these effects also lead to a general motor planning problem, including problems with chewing and drooling. Although the genes are clearly involved in specification of our species-specific linguistic ability, the pathway for these effects is extremely indirect and loosely coupled.
- (4) *Memory*. Repeated use of a muscle or bone will lead to its growth and strengthening. Language functions in a similar way. Each use of a sound, word, or construction in a particular context strengthens the memory for that form and increases its ability to compete with alternative forms. These memory processes operate initially within individuals, but then extend over time to whole social and dialect groups.

#### 4. Domain generality

Emergentist accounts often emphasize the extent to which a complex set of behaviours and forms can arise from a few simple mechanisms. Some examples of general cognitive mechanisms include:

- (1) Learning through error propagation, as in the back-propagation algorithm (Rumelhart and McClelland 1986).

- (2) Self-organization in feature maps (Kohonen 2001) and topological sheets (Shrager and Johnson 1995, Elman 1999).
- (3) Item-based learning (MacWhinney 1975a, Tomasello 2000) of grammatical constructions (Goldberg 1999).
- (4) Statistical learning that allows both children and adults to learn patterns of sounds, tones, or visual forms (Saffran *et al.* 1996, Jusczyk 1997), as constituted in sequential groups (Houghton 1990, Gupta and MacWhinney 1997).
- (5) Reorganization of cognitive function for damaged tissue to the contralateral hemisphere (Corina *et al.* 1992, Booth *et al.* 1999, MacWhinney *et al.* 2000).
- (6) Learning to learn, which guides learners' attention to cues that have proven useful in learning previous similar forms (Smith 1999) and which further confirms the importance of these cues for further learning.
- (7) Analogy, which allows problem-solvers and learners to structure domains and relations in parallel alignment (Gentner and Markman 1997).
- (8) Consolidation of short-term memories into long-term memories (McClelland *et al.* 1995).

These domain-general cognitive mechanisms may themselves emerge from the operation of physicalist mechanisms, such as resonance, coupling, competition and selection, that operate on the level of individual neurons.

## 5. Landscapes

Domain-general mechanisms all play a major role in learning and processing. We know that nature loves continually to reuse old forms for new functions, so it would not be surprising to find a process such as statistical learning emerging in many different domains. However, it would be a mistake to imagine, as some critics (Sabbagh and Gelman 2000) have suggested, that emergentist accounts can refer only to domain-general processes. Rather, emergentism holds that domain-general processes operate across a variety of domain-dependent landscapes or substrates, including the brain, the body and society. To attempt to construct an emergentist psycholinguistics that ignores the topology of these landscapes would be like attempting to build an account of honeycomb formation that ignores the honey and the wax.

The importance of the interaction between domain-general processes and specific landscapes has become more and more apparent as emergentist accounts increase their emphasis on the embodied and grounded nature of language and cognition (Feldman and Narayanan 2004). For example, Givón (1998) has argued that language evolution involved a linkage of episodic memory to the auditory system through the support or tutelage of the visual system. The visual system had already established general mechanisms for the episodic encoding of spatial position and form. Primates had already developed a mechanism for recording auditory sequences (Hauser *et al.* 2001). Adapting this mechanism to the task of language learning involved reshaping and relinking previously available cognitive mechanisms. As a result, these general mechanisms underwent a special tuning when they functioned at the local level (Caplan and Waters 1999).

Emergentist accounts make reference to grounding on neuronal, physiological and social landscapes or substrates. Let us take a look at how grounding has this differentiating effect across each of these three landscapes.

### 5.1 *The neuronal landscape*

It is a mistake to think of the brain as a vast, undifferentiated sea of 'connectoplasma'. Each cortical area has many substructures that differ in their connectivity and cytoarchitectonics

(Buck *et al.* 1998). Emergentist accounts hold that much of this differentiation is dependent on general principles of induction (Miller 1994), plasticity (Quartz and Sejnowski 1997, Booth *et al.* 2001) and gene expression (Murray 1988) during epigenesis. Moreover, the brain retains plasticity throughout development.

By the end of infancy, the overall contours of brain differentiation are well established. Because these differentiated structures are themselves coupled to environmentally differentiated functions, homeorhesis within the coupled functional neural circuit maintains stability over time, despite ongoing local instability from cell death and synaptic change. Domain-general processes such as consolidation, analogy and statistical learning are then available to operate across this highly differentiated neuronal landscape, leading to slightly different outcomes and patterns in each local domain.

## 5.2 *The physiological landscape*

On the physiological level, the interaction of domain-general processes with the local substrate is perhaps even more apparent. The relevant mechanisms here are not the cognitive processes of analogy or consolidation, but more basic processes of physical emergence such as resonance, coupling, harmonics, the Bernoulli effect, regulatory feedback, force, pressure and tension. The articulatory system can be characterized as an assemblage of springs and dampers whose functioning is expressed through sets of differential equations (Story 2002, Boersma and Weenink 1996). For example, when singers use vibrato, they set up a resonance at 5–6 Hz between the cricothyroid and thyroarytenoid muscles. This resonance can be modelled by the vibrating string formula, as applied to the underlying dynamic mechanical system (Titze *et al.* 2002). Phonologists have shown that the shape of the vocal mechanism has a wide-ranging impact on phonological processes (Ohala 1974). For example, the preference across languages for unvoiced stops in syllable initial position emerges from the nature of the pressures that build up at the glottis prior to release of a stop closure. These pressures tend to block voicing until well after the stop has been released.

Consider a simple example from phonological development. There is a universal tendency to avoid sequences of nasal consonants followed by voiceless obstruents, as might arise in forms like ‘manpower’. This constraint is grounded on the facts of speech production (Huffman 1993) and figures prominently in recent elaborations of Optimality Theory (Kager 1999). Languages use at least five phonological processes to deal with this problem. These processes include nasal substitution, post-nasal voicing, denasalization, nasal deletion and vowel epenthesis. Initially, children may apply several of these processes (Bernhardt and Stemberger 1998), with only some being in accord with the language they are learning. Which processes are preserved and which are dropped will depend on the shape of the target language, be it Indonesian, Quechua, Toba Batak, English, or Kelantan Malay. In general, we can view the shape of phonological systems as emergent from constraints established by the articulatory system. Rather than stipulating phonological rules or constraints (Bernhardt and Stemberger 1998, Kager 1999), we can view them as emergent responses to underlying pressures.

Discussions of the physiological landscape eventually must also make reference to the external physical environment. For example, our account of the way in which the ear has adapted for hearing must take into account the shape of sound dissipation for low frequencies across distances.

## 5.3 *The social landscape*

The varied topography of the social landscape also plays a major role in determining the emergence of linguistic form. Language is rich in frozen expressions, formulas, exceptions

and irregularities. Many of these expressions survive because they are linked to unique social functions and contexts. For example, the everyday form 'goodbye' has its roots in the phrase 'God be with ye'. We use the phrases 'how about X' and 'what about X' happily enough, but would never dream of using 'when about X' or 'how under X'. We are willing to pronounce the 'ough' in 'plough', 'tough' and 'slough' in three totally different ways despite the similar orthography. When we are in Boston, we expect to hear 'car' pronounced without the final /r/ and when we are in Pittsburgh, we expect to hear 'oil' pronounced as 'earl'. We know that we should refer to the Pope as his Holiness, although we may not know how to address an Anglican bishop. We know not to say 'good night' to someone until we leave, even if it is late at night, whereas we have no problem using 'good morning' and 'good evening' as greetings.

Language is a vast quilt of irregularities, variation and special cases, set against a backdrop of classes with some partial regularity (MacWhinney 1975b, Pinker 1991). Much of this variation arises from specific complexities in the social fabric. These complexities include:

- (1) *Linguistic diversity*. Humans speak about 7000 different languages. Globalization has brought these languages into contact, leading to a myriad of forms of multilingualism, pidgins, creoles and language death. Each of these forms arises because individuals maintain a given level of commitment to a social group, typically on the level of the nation. The tempo and effects of these contact phenomena are increasing as globalization accelerates.
- (2) *Dialect diversity*. Within large language communities, such as English in the UK, one can identify dozens of local dialects. Dialects represent a level of group commitment below that of the nation. These commitments can be allied with political, ethnic and religious divisions. All of these divisions will have their effect on the meshing of linguistic forces in the current moment.
- (3) *Small group solidarity*. Within dialect regions, people make further social commitments to groups organized about family, causes, parties, businesses, sports and religion. These groups then commit themselves to particular ways of talking and thinking. To maintain identification with the group, members must display full command of all these markers.
- (4) *Conversational history*. As we noted earlier, social interactions represent a level of linguistic emergence resting somewhere between group identification and online processing. Crucial features of conversational history include shared knowledge, referents and goals. Recent advances in eye tracking methodology (Allopenna *et al.* 1998, Keysar *et al.* 1998) have given new impetus to the study of language use in realistic situational contexts. Researchers are just beginning to use this methodology to study processes such as the online updating of anaphoric reference ('it', 'that') during real-life activities such as cooking (Hanna and Tanenhaus 2004). Soon it may be possible to link realistic studies of social dynamics of this type to detailed computational modelling, as has already been done for eye movements driven by lexical competition (Tanenhaus *et al.* 2000). Later in this paper, ways in which this new effort can be given a deeper social grounding by tracing the use and comprehension of grammatical devices for shifting perspective will be examined.

Each of these social and conversational forces makes repeated impacts on linguistic form during online sentence processing. When repeated pressures operate consistently in one direction, they produce language change. For example, Du Bois (1987) has argued that ergative marking in languages emerges from the fact that speakers tend to delete the actor in transitive sentences, because it is already given or known. The more often speakers engage in this type of deletion, the stronger the push toward ergativity in a given language. Neural network modellers have begun to use genetic algorithms to model some aspects of these repetitive social processes, as well as language evolution in robots (Steels and Kaplan 2000) and the evolution of robots

that use language (Marocco *et al.* 2003). These models (Hinton and Nowlan 1987) can show how individuals who are better able to learn a task will have greater survival and reproduction, thereby increasing the ability of the overall population to learn the task. Such demonstrations of the 'Baldwin effect' demonstrate a coupling between the ontogenetic and phylogenetic time frames.

Let us now examine two areas of language learning where emergent processes compete and co-operate across time. First, we shall look at the various forces that operate together to determine the learning of new words. Then, we shall look at some of the social and cognitive forces that determine the ways we engage in perspective taking during language use.

## 6. Emergence in word learning

One of the most active areas of current research in child language is the study of early word learning. Researchers have pinpointed a wide variety of cues to word learning that operate on very different time-scales. Before discussing neural network models of this process, it is important to review the overall shape of the word learning challenge.

### 6.1 *The challenge*

Philosophers such as Quine (1960) have emphasized the extent to which word learning needs to be constrained by ideas about what might constitute a possible word. For example, if the child were to allow for the possibility that word meanings might include disjunctive Boolean predicates (Hunt 1962), then it might be the case that the word 'grue' would have the meaning 'green before the year 2000 and blue thereafter'. Similarly, it might be the case that the name for any object would refer not to the object itself, but to its various undetached parts. When one thinks about word learning in this abstract way, it appears to be impossibly hard.

Markman (1989) and Golinkoff *et al.* (1994) have proposed that Quine's problem can be solved by imagining that the child's search for word meanings is guided by lexical principles. For example, children assume that words refer to whole objects, rather than parts of objects. Thus, a child would assume that the word 'rabbit' refers to the whole rabbit, not just some parts of the rabbit. Other putatively innate constraints include the taxonomic constraint (object names refer to classes in a natural taxonomy, such as 'dog'), mutual exclusivity (each referent has only one name), the shape bias (shape is the primary cue for selecting object word meaning), the nominal bias (nouns are easier to learn than other parts of speech) and the prototypical level constraint (names refer to terms like 'bird' that maximize feature conjunction, rather than 'eagle').

In the original formulations from the 1980s, these constraints were proposed as innate components of the species-specific human language facility. However, more recently, researchers have shown how these principles can emerge from the operation of the cognitive system. For example, Merriman (1999) argued that mutual exclusivity emerges naturally from the competition between closely related lexical items (MacWhinney 1989). Similarly, Colunga and Smith (2000) have shown how the shape bias emerges from ongoing word learning; and Gentner (in press) has explained how the shape bias arises from the fact that nouns place fewer demands on the child's memory system for encoding. These revised accounts typically shift the weight of the argument away from attempts to constrain the hypothesis space to attempts to characterize the positive cues toward word learning. In this sense, these new ways of dealing with the logical problem of word learning reflect parallel arguments for dealing

with the logical problem of grammatical acquisition (MacWhinney 2004) where arguments based on learnability are giving way to analyses based on acquirability.

## 6.2 *Initial mappings*

The initial mapping process involves the association of auditory units to conceptual units (Reznick 1990, Naigles and Gelman 1995). This initial mapping is typically fast, sketchy and tentative. In extreme cases, it can rely entirely on short-term cues present only at the moment of initial mapping. For example, the 14-month-old can be brought into the laboratory (Schafer and Plunkett 1998) and shown a picture of an animal called a 'tiv'. The child will then demonstrate understanding of the new word by turning to a picture of the new animal, rather than a picture of a dog, when hearing the word 'tiv'.

In these laboratory experiments, children are learning a new concept in parallel with a new word. However, in the real world, children often have developed a clear idea about a concept well before they have learned the word for that concept. The child comes to the language learning task already possessing a fairly well-structured coding of the basic objects in the immediate environment (Piaget 1954, Sugarman 1982, Stiles-Davis *et al.* 1985). Children treat objects such as dogs, plates, chairs, cars, baby food, water, balls and shoes as fully structured separate categories (Mervis 1984). They also show good awareness of the nature of particular activities such as falling, bathing, eating, kissing and sleeping. This means that, in reality, conceptual organization often precedes initial lexical mapping.

The shift in recent years from reliance on constraints to an emphasis on cues has opened up space for considering the time course of the processes supporting initial mapping. For example, Gentner's view of nouns as benefiting from easier memory encoding involves processes that operate at two different time frames. The first process encodes a current episodic state. At this initial point, nouns would have little advantage over verbs. The second process involves the long-term storage of a series of episodes and the comparison of these stored segments with new exemplars. This longer-term process is certainly easier for nouns, since the representation of invariant properties of nouns requires relatively less attention to changes over time.

Another set of cues has a clearer online effect. These are the cues to word meaning that derive from the syntactic context. Consider these contexts.

- (1) Here is a pum—count noun.
- (2) Here is Pum—proper noun.
- (3) I am pumming—intransitive verb.
- (4) I pummed the duck—transitive (causative) verb.
- (5) I need some pum—mass noun.
- (6) This is the pum one—adjective.

Each of these sentential contexts provides clear evidence that pum is a particular part of speech. Other sentential frames can give an even more precise meaning. If the child hears 'this is not green, it is pum', then it is clear that 'pum' is a colour. If the child hears 'please don't cover it, just pum it lightly', then the child knows that 'pum' is a verb of the same general class as 'cover'. The use of cues of this type leads to a fast, but shallow, mapping of new words to new meanings.

Another set of cues is more clearly linked to the moment of initial learning. These are the cues that involve pointing at an object, shifts in gaze and orientation of the posture toward an object. The ability to follow eye gaze appears to rely on fundamental developments in the visual system that emerge in the first 4 months (Johnson 1992). These developmental changes involve the linkage of basic phylogenetic abilities to ongoing epigenesis. Similar changes arise

in the tracking of postural cues and pointing. By the time the child comes to learn the first words, these cues are generally accessible. Baldwin (1991) and Baldwin and Markman (1989) have shown that children try to acquire names for the objects that adults are attending to. Similarly, Akhtar *et al.* (1996) and Tomasello and Akhtar (1995) have emphasized the crucial role of mutual gaze between mother and child in the support of early word learning.

Going beyond the use of superficial cues, Tomasello (2003) and Bloom (2002) have argued that word learning depends on the child's ability to decode the parent's intentions. For Tomasello and Bloom, this sharing of intentions and trust is not just a matter of superficial cue processing. Tomasello (1999) has argued that human mothers differ significantly from primate mothers in the ways that they encourage mutual attention during language. While not rejecting the overall role of social support in language learning, Samuelson and Smith (1998) have noted that one can also interpret the findings of Akhtar *et al.* in terms of low-level perceptual and attentional matches that help focus the child's attention on novel objects to match up with new words. Whatever the eventual outcome of this debate, it is important to note that the processing of intentionality may involve a longer time-scale than the processing of superficial cues such as gaze and pointing. When processing intentionality, children may refer to facts such as whether the adult had seen hiding objects, or whether the adult and child had previously played with and named an object or activity.

### 6.3 *Beyond initial mappings*

Initial mapping is only the tip of the lexical iceberg. Most lexical learning occurs in a much more extended time frame after the formation of this initial mapping. As the child is exposed repeatedly to new instances of an old word, the semantic range of the referent slowly widens. Barrett (1995), Huttenlocher (1974) and others have viewed this aspect of meaning growth as 'decontextualization'. Harris *et al.* (1988) have shown that the initial representations of words contain components that are linked to the first few contacts with the word in specific episodes or specific contexts. As long as the child sticks closely to attested instances of the category inside the confirmed core, she will tend to under-generalize the word. For example, the word 'car' may be used to refer only to the family car. Anglin (1977) and Dromi (1987) have argued that the frequency of such under-generalizations is typically underestimated, because under-generalizations never lead to errors. If one does a careful analysis of the range of uses of new words, it appears that under-generalization is closer to the rule than the exception.

Gradually, the process of generalization leads to a freeing of the word from irrelevant aspects of the context. Over time, words develop a separation between a 'confirmed core' (MacWhinney 1984, 1989) and a peripheral area of potential generalization. As the confirmed core of the meaning of a word widens and as irrelevant contextual features are pruned out, the word begins to take on a radial or prototype form (Rosch and Mervis 1975, Lakoff 1987). In the centre of the category, we find the best instances that display the maximum category match. At the periphery of the category, we find instances whose category membership is unclear and which compete with neighbouring categories (MacWhinney 1989).

According to the core-periphery model of lexical structure, over-generalizations arise from the pressures that force the child to communicate about objects that are not inside any confirmed core. Frequently enough, children's over-generalizations are corrected when the parent provides the correct name for the object (Brown and Hanlon 1970). The fact that feedback is so consistently available for word learning increases our willingness to believe that the major determinants of word learning are social feedback, rather than innate constraints or even word learning biases.

#### 6.4 *Word learning and schooling*

Current research has tended to focus on word learning in the first 5 years, ignoring the extent to which schooling shapes word meanings. Even for basic concepts such as ‘cat’, schooling can bring about fundamental changes in conceptual structure. For example, Keil (1989) asked children whether a cat that had the fur of a skunk and the shape of a skunk would still be a cat. Preschoolers concluded that the cat would now be a skunk. However, 7-year-olds refused to accept the idea that the change of fur and shape would alter the internal nature of the cat. School children will first support their position by just saying that once you are a cat you are always a cat. However, later, they will make reference to scientific constructs such as genetics and DNA to support their view. There is good reason to believe that schooling supports this shift toward essentialism.

To understand how schooling operates in real time, researchers have begun to rely on detailed analyses of classroom conversations about mathematical and scientific concepts. These analyses are based on transcripts linked to video, many of which are available from <http://talkbank.org>. Let us consider one particular instance of a detailed study of lessons by Paul McCabe and Kay McClain in a seventh-grade classroom studying terminology for describing statistical distributions. The students have been interacting for several weeks with a computer interface that allows them to vary treatments and observe outcome effects. In the current lesson, the teacher is discussing a projected graph of battery life for two types of flashlight batteries. Tests of 30 units of the Always Ready batteries showed that eight lasted longer than 6 h, although 10 lasted less than 4 h. Tests of 20 units of the Tough Cells batteries showed that none of them lasted less than 4 h, although only three lasted longer than 6 h. The teacher begins the discussion by asking which of the two batteries is ‘more dependable’. Caesara responds that Always Ready are more dependable because they yielded more ‘good ones’. This interpretation of the concept of dependability fails to pay attention to the notion of a minimal performance standard. For example, if one were exploring a cave with a flashlight, you might want to make sure that your battery would not give out in the middle of your 4-h exploration. If it did, you would call it ‘undependable’. Understanding this perspective, Blake counters Caesara’s position by noting that ‘All of the Tough Cell is above 80’. At this point, the teacher asks other students to clarify Caesara’s position. Sequoria uses a set of hand gestures to draw a picture of the distribution of batteries in the air and then to point deictically to the segment of the distribution that Caesara is emphasizing. Although Caesara does not overtly withdraw her analysis during this discussion, Blake’s provision of a clear account of the meaning of ‘dependability’ should help the entire class move toward a fuller, statistically grounded understanding of this term.

In this video interaction, we see a wide range of social forces impinging on the collaborative process of discovering meanings. Forces of gender, age, ethnicity and group membership are evident in the discussion. The teacher plays a unique role in terms of her ability to recognize students who want to make contributions, revoice their contributions and question their positions. Apart from the 6-min video segment available at <http://talkbank.org>, the overall study of reasoning from graphs and charts (Sfard and McClain 2002) tracked learning in this domain over several weeks. By examining the growth of interlocked concepts such as ‘distribution’, ‘effectiveness’ and ‘dependability’ over this entire period across various students, we can gain a clearer idea of the process of word learning in the school years.

#### 6.5 *Neural network models of word learning*

Current models of lexical learning are not yet able to integrate emergence across all of these time frames and cue types. Most current models focus on the acquisition of initial mappings and

have little to say about later semantic reorganization. None of the models is able to characterize the interplay of gestural, intentional, attentional, conversational and syntactic cues to meaning. Despite these limitations, these models have made good progress in demonstrating some core aspects of lexical emergence.

One framework that allows us to model several of these forces is the self-organizing feature map (SOFM) architecture (Kohonen 2001), as expressed in the DevLex model of Li *et al.* (2004). SOFM networks model initial word learning as patterns of organization in cortical maps. Three local maps are involved in word learning: an auditory map; a concept map; and an articulatory map. Emergent self-organization on each of these three maps uses the same learning algorithm. Word learning involves the association of elements between these three maps. What makes this mapping process self-organizing is the fact that there is no pre-established pattern for these mappings and no preordained relation between particular nodes and particular feature patterns.

Evidence regarding the importance of syllables in early child language (Bijeljac *et al.* 1993, Jusczyk *et al.* 1995) suggests that the nodes on the auditory map may best be viewed as corresponding to full syllabic units, rather than separate consonant and vowel phonemes. The recent demonstration by Saffran *et al.* (1996) of memory for auditory patterns in 4-month-old infants indicates that children are not only encoding individual syllables, but also are remembering sequences of syllables. In effect, pre-linguistic children are capable of establishing complete representations of the auditory forms of words. Within the self-organizing framework, these capabilities can be represented in two alternative ways. One method uses the slot-and-frame feature notation from MacWhinney *et al.* (1989). An alternative approach views the encoding as a temporal pattern that repeatedly accesses a basic syllable map. A lexical learning model developed by Gupta and MacWhinney (1997) uses serial processes to control word learning. This model couples a serial order mechanism known as an 'avalanche' (Grossberg 1978) with a lexical feature map model. The avalanche controls the order of syllables within the word. Each new word is learned as a new avalanche.

In the self-organizing framework, the learning of a word is viewed as the emergence of an association between a pattern on the auditory map and a pattern on the concept map through Hebbian learning (Hebb 1949, Kandel and Hawkins 1992). When the child hears a given auditory form and sees an object at the same time, the coactivation of the neurons that respond to the sound and the neurons that respond to the visual form produces an association across a third pattern of connections that maps auditory forms to conceptual forms. Initially, the pattern of these interconnections is unknown, because the relation between sounds and meanings is arbitrary (de Saussure 1966). This means that the vast majority of the many potential connections between the auditory and conceptual maps will never be used, making it a very sparse matrix (Kanerva 1993). In fact, it is unlikely that all units in the two maps are fully interconnected (Shrager and Johnson 1995). In order to support the initial mapping, some researchers (Schmajuk and DiCarlo 1992) have suggested that the hippocampus may provide a means of maintaining the association until additional cortical connections have been established. As a result, a single exposure to a new word is enough to lead to one-trial learning. However, if this initial association is not supported by later repeated exposure to the word in relevant social contexts, the child will no longer remember the word.

Neural networks have also been used to model some of the aspects of lexical reorganization that lead to semantic over-generalization and recovery from over-generalization. Schyns (1991) applied a SOFM to the task of learning three competing categories with prototype structures. The individual exemplars of each category were derived from geometric patterns that were blurred by noise to create a prototype structure, although the actual prototypes were never displayed. The simulations showed that the network could acquire human-like use of the categories. When presented with a fourth new word that overlapped with one of the first

three words, the system broke off some of the territory of the old referent to match up with the new name. This competitive behaviour seems to reflect the process of competition between old words and new words discussed for children's word learning by Markman (1989), Clark (1987) and MacWhinney (1989).

A simulation of meaning development by Li and MacWhinney (1996) used a standard back-propagation architecture to model the learning of reversive verbs that used the prefix 'un-', as in 'untie', or 'dis-', as in 'disavow'. The model succeeded in capturing the basic developmental stages for reversives reported by Bowerman (1982) and Clark *et al.* (1995). In particular, the model was able to produce and later correct over-generalization errors such as '\*unbreak' or '\*disbend'. The network's eventual correct performance was based on its internalization of what Whorf (1938, 1941) called the 'cryptotype' for the reversive, which involved a 'covering, enclosing, and surface-attaching meaning' that is present in a word like 'untangle', but absent in a form such as '\*unbreak'. Whorf viewed this category as a prime example of the ways in which language reflects and possibly shapes thought.

## 7. Perspective taking

Our second example of the meshing of forces for linguistic emergence comes from the study of the impact of perspective taking on syntax. Linguistic devices that emerge from the marking of perspective shifts include spatial deixis, extraposition, anaphora, passivization, relativization, and many other structures (MacWhinney 2005).

Successful communication requires three major types of cognitive sharing. First, we must establish mutual reference (Clark and Marshall 1981) to objects and actions in the world. Second, we link this shared reference into more general system of mutual knowledge about the ways in which objects, words and events behave. Third, in order to convey new information based on these mutual understandings, we use the perspective taking system to mark the flow of our ideas across objects, positions, actions, times and beliefs. By giving clear cues to our listeners about which perspectives they should assume and how they should move from one perspective to the next, we maximize the extent to which they can increase our shared understandings and knowledge. When language is rich in cues for perspective taking and perspective shifting, it awakens the imagination of the listener and leads to successful sharing of ideas, impressions, attitudes and narratives. When the process of perspective sharing is disrupted by interruptions, monotony, excessive complexity, or lack of shared knowledge, communication can break down. MacWhinney (2005) formulated this general perspective hypothesis in terms of the following basic claims.

- (1) Perspective taking operates online using images created in five systems: direct experience, space/time deixis, plans, social roles and mental acts.
- (2) Language uses perspective taking to bind these five imagery subsystems.
- (3) Grammar emerges from conversation as a method for supporting accurate tracking and switching of perspective.
- (4) By tracing perspective shifts in language, children are able to learn the cognitive pathways and mental models sanctioned by their culture.

### 7.1 *Integration of perspective across cognitive domains*

Perspective operates on words that are linked to images constructed within each of the five specialized domains. In this way, language provides a general rubric for knitting together understandings across domains. Consider a sentence, such as 'Last night, my sister's friend

reminded me I had dropped my keys under the table behind the garage'. Here, we see how a single utterance integrates information about time (last night), space (under, behind), social relations (sister's friend), mental acts (remind), objects (keys, table, garage) and events (drop). The sentence directs the listener to assume specific perspective on each of these informational levels, including those of the sister, the friend, the speaker and the various locations. The detailed images triggered by these actions, beliefs and referents are elaborated in regions of the brain that are specific to the relevant perceptions. However, language achieves an integration of all this information by setting up a functional neural circuit that integrates across the domains.

## 7.2 *Integration across time*

The computation of perspective operates across three major time frames. The fastest time frame is the one driven by online processing of the sentence. As listeners, we must construct a superficial trace of the perspectival flow of a sentence in order to keep up with the conversation. As speakers, we must quickly channel our thoughts into a set of devices for perspective marking that can be easily understood by our listeners. For example, if we choose to topicalize the object of the verb, we must be ready quickly to select the passive voice to allow the object to be promoted to subject position.

This online manipulation of perspective is built upon a rich system of support for online updating of imagery. Recent research in neurophysiology has shown that, when we imagine objects and actions, we typically activate the same neuronal pathways that are used for direct perception and direct action. For example, when we imagine performing bicep curls, there are discharges to the biceps (Jeannerod 1997). When a trained marksman imagines shooting a gun, the discharges to the muscles mimic those found in real target practice. When we imagine eating, there is an increase in salivation. Neuroimaging studies by Parsons *et al.* (1995), Martin *et al.* (1996) and Cohen *et al.* (1996) have shown that, when subjects are asked to engage in mental imagery, they use modality-specific sensorimotor cortical systems. For example, in the study by Martin *et al.*, the naming of tool words specifically activated the areas of the left premotor cortex that control hand movements. The imagery system relies on the cognitive creation of a body map (Damasio 1999, Kakei *et al.* 1999). This body map then functions as an internal homunculus that can be projected to track the actions of others through the system of motor neurons (Rizzolatti *et al.* 1996).

Linguistic perspective taking uses this underlying support for embodied cognition to shift perspective online. Consider this pair of sentences:

- (1) Visiting relatives can be a nuisance.
- (2) Crying babies can be a nuisance.

Sentence (1) provides us with an ambiguous perspective, since it is not clear whether we are visiting the relatives or the relatives are visiting us. Sentence (2) permits only one perspective, since the verb 'cry' is not transitive so it must be the babies who are crying. But note that the evaluative term 'nuisance' forces a shift to the ego-perspective in both cases. Even though we may begin by evaluating crying from the viewpoint of the babies, we must then shift to ego-perspective to understand how these crying babies can become a nuisance.

## 7.3 *Shifts in perspective over time*

The initial computation of perspective may be quite superficial. We can interpret (2) superficially as if it had the perspectival structure of 'crying babies can be fat'. Here, we must store the fact that the sentence predicates some attribute of crying babies and not worry about

constructing a fully perspectival understanding. This contrast between the fast online interpretation of perspective and the slower, deeper full construction of perspective represents the tension between two time frames. Evidence for this contrast comes from recent studies of eye movements in referential communication experiments. For example, Keysar *et al.* (1998) set up a task in which a director asked a matcher to move objects about on a set of shelves (see figure 1). The shelves were arranged so that some of the objects that were viewable to the matcher were not viewable to the director. If the director asked the matcher to ‘put the bottom block below the apple’ then the matcher initially looked at the block in the bottom left cell. However, this block was not visible to the director. Once matchers realized this, they then shifted their gaze to the block on the next row up, which must have been the block to which the director was referring.

Schober *et al.* (2003) have tracked the gradual imposition of the listener’s perspective over time in the brief conversations involved in telephone interviews and surveys. When we try to explain things to strangers, we often fail to grasp fully the difference between their perspectives and ours. For example, when I called up the Hitachi service centre to describe problems with the operation of my Hitachi universal remote control, the first 5 min of the conversation involved a series of attempts to verify model numbers, clarify terminology regarding input buttons, discuss cable connections, and explain my expectations regarding the functioning of the device. Each of these steps required descriptions of my perspective, adjustment to the standard perspective encoded by the service representative, and then introduction of further aspects of the problem. In the end, once all these terms were shared, it turned out that the actual answer to my problem involved nothing more than the toggling of a single button. However, the complexity of this device with all its buttons and the various connections between the TV, VHS and DVD involved a gradual shared construction of a complex perspective over time.

#### 7.4 Relativization

Restrictive relative clauses provide detailed evidence regarding online perspective shifting. Consider these four types of restrictive relative clauses.

- SS: The dog that chased the cat kicked the horse.    0 switches.  
 OS: The dog chased the cat that kicked the horse.    1– switch.  
 OO: The dog chased the cat the horse kicked.        1+ switch.  
 SO: The dog the cat chased kicked the horse.        2 switches.

In the SS type, the perspective of the main clause is also the perspective of the relative clause. This means that there are no full perspective switches in the SS relative type, although attention shifts partially for a moment to ‘the cat’ in the relative clause. In the OS type, perspective flows

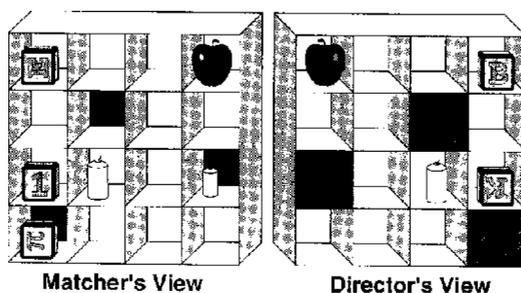


Figure 1.

from the main clause subject ('the dog') to the main clause object ('the cat') in accord with the general principle of partial shift of perspective to the object. At the word 'that', perspective flows further to 'the cat' as the subject of the relative clause. This perspective shift is made less abrupt by the fact that 'cat' had already received secondary focus before the shift was made. In the OO sentence type, perspective also switches once. However, in this case, it switches more abruptly to the subject of the relative clause. In the SO relative clause type, there is a double perspective shift. Perspective begins with the main clause subject ('the dog'). When the next noun ('the cat') is encountered, perspective shifts once. However, at the second verb ('kicked'), perspective has to shift back to the initial perspective ('the dog') to complete the construction of the interpretation.

The perspective account predicts this order of difficulty:  $SS > OO = OS > SO$ . Studies of both acquisition (MacWhinney 1982) and adult processing (MacWhinney and Pléh 1988) have provided support for these predictions. It appears that each shift in perspective exacts a cost during online processing, slowing reading time. Similar results for perspective switches in processing pronouns were reported by McDonald and MacWhinney (1995).

Sentences with multiple centre embeddings have even more switches. For example, 'the dog the cat the boy liked chased snarled' has four difficult perspective switches (dog → cat → boy → cat → dog). Sentences that have as much perspective shifting as this without additional lexical or pragmatic support are incomprehensible, at least at first hearing; but note that the mere stacking of nouns by itself is not enough to trigger perspective shift overload. Consider the sentence, 'My mother's brother's wife's sister's doctor's friend had a heart attack'. Here, we do not really succeed in taking each perspective and switching to the next, but some form of sloppy comprehension is still possible. This is because we just allow ourselves to skip over each perspective and land on the last one mentioned. In the end, we just know that someone's friend had a heart attack.

## 8. Conclusion

Linguistic form emerges from the meshing of disparate forces at the moment of speaking. These forces arise from domain-general processes that operate across domain-specific landscapes. In this paper, we have examined systemic interactions in one aspect of language development and one aspect of language processing. When we looked at word learning, we saw a separation between forces that impact initial mapping and those that impact later enrichment and consolidation of word meaning. Initial mapping depends on systems for processing cues such as gaze, pointing, posture, intentionality, contrast, syntactic context and previous word learning. In many cases, these cues work together to guide the child toward selecting the correct interpretation of a new word. When cues are either missing or contradictory, initial mapping will be less successful. After initial mapping, the forces of conceptual change, conversation and formal schooling work to elaborate detailed word meanings. When we observe real moments of word learning, such as the use of 'dependability' in the Nashville classroom, we can see how weeks and even years of learning are further shaped online by conversational pressures and instructional materials.

When we look at the process of perspective taking, we see a similarly rich meshing of processes with disparate time frames in the moment. Sentences link perspectives across the cognitive domains of space, time, perception, action, belief and social roles. When we listen to sentences, we may manage to pick up only the more superficial perspectives. If we are familiar with the material or if we are given time to process further, we can construct a richer representation with fuller representation of perspective shifts. The time course of this activity

depends heavily on support from linguistic devices, shared social meanings and conversational sharing in the moment.

It is possible that some processes will show partial decomposability. For example, processes of language change may depend on so many millions of repeated cases of usage that detailed online facts about current conversations will not impact their direction. However, historical forces will still have their impact in shaping of the actual uses of forms in the moment. This means that we may be able to model the emergence of long-term structures without repeated attention to detail, but we cannot model the details without attention to long-term structures. In other cases, such as the example we discussed from fusion physics, failure to consider interactions across disparate time-scales may lead to completely erroneous predictions. Capturing the data needed to test these interactions and figuring out how to model this meshing across time frames represents a major challenge to connectionist modelling.

## References

- N. Akhtar, M. Carpenter and M. Tomasello, "The role of discourse novelty in early word learning", *Child Dev.*, 62, pp. 635–645, 1996.
- P.D. Allopenna, J.S. Magnuson and M.K. Tanenhaus, "Tracking the time course of spoken word recognition using eye movements: evidence for continuous mapping models", *J. Mem. Language*, 38, pp. 419–439, 1998.
- J.M. Anglin (Ed.), *Word, Object, and Conceptual Development*, New York: Norton, 1977.
- D.A. Baldwin, "Infants' contribution to the achievement of joint reference", *Child Dev.*, 62, pp. 875–890, 1991.
- D.A. Baldwin and E.M. Markman, "Establishing word-object relations: a first step", *Child Dev.*, 60, pp. 381–398, 1989.
- M. Barrett, "Early lexical development", in *Handbook of Child Language*, P. Fletcher and B. MacWhinney, Eds, Oxford: Basil Blackwell, 1995, pp. 362–392.
- U. Bellugi, H. Poizner and E.S. Klima, "Language, modality and the brain", *Trends Neurosci.*, 12, pp. 380–388, 1989.
- B.H. Bernhardt and J.P. Stemberger, *Handbook of Phonological Development from the Perspective of Constraint-based Nonlinear Phonology*, San Diego: Academic Press, 1998.
- B.R. Bijeljac, J. Bertoncini and J. Mehler, "How do four-day-old infants categorize multisyllabic utterances?", *Devl Psychol.*, 29, pp. 711–721, 1993.
- P. Bloom, *How Children Learn the Meanings of Words (Learning, Development, and Conceptual Change)*, Cambridge, MA: MIT Press, 2002.
- P. Boersma and D. Weenink, *Praat, A System for Doing Phonetics by Computer*, Amsterdam: Institute of Phonetic Sciences of the University of Amsterdam, 1996.
- J.R. Booth, B. MacWhinney, K.R. Thulborn, K. Sacco, J. Voyvodic and H. Feldman, "Functional organization of activation patterns in children: whole brain fMRI imaging during three different cognitive tasks", *Prog. Neuropsychopharmacol. Biol. Psychiatry*, 23, pp. 669–682, 1999.
- J.R. Booth, B. MacWhinney, K.R. Thulborn, K. Sacco, J.T. Voyvodic and H.M. Feldman, "Developmental and lesion effects during brain activation for sentence comprehension and mental rotation", *Devl Neuropsychol.*, 18, pp. 139–169, 2001.
- M. Bowerman, "Reorganizational processes in lexical and syntactic development", in *Language Acquisition: The State of the Art*, E. Wanner and L. Gleitman, Eds, New York: Cambridge University Press, 1982, pp. 319–346.
- R. Brown and C. Hanlon, "Derivational complexity and order of acquisition in child speech", in *Cognition and the Development of Language*, J.R. Hayes, Ed., New York: Wiley, 1970, pp. 11–54.
- G. Buck, T. VanEssen, K. Tatsuoka, I. Kostin, D. Lutz and M. Phelps, *Development, Selection and Validation of a Set of Cognitive and Linguistic Attributes for the SAT I Verbal: Sentence Completion Section*, Princeton, NJ: Educational Testing Service, 1998.
- J. Bybee and P. Hopper, *Frequency and the Emergence of Linguistic Structure*, Amsterdam: John Benjamins, 2001.
- N.A. Campbell, J.B. Reece and L.G. Mitchell, *Biology*, 5th edn, Menlo Park: Addison Wesley, 1999.
- D. Caplan and G.S. Waters, "Verbal working memory and sentence comprehension", *Behav. Brain Sci.*, 22, pp. 77–126, 1999.
- E. Clark, "The principle of contrast: a constraint on language acquisition", in *Mechanisms of Language Acquisition*, B. MacWhinney, Ed., Hillsdale, NJ: Lawrence Erlbaum, 1987, pp. 1–34.
- E. Clark, K. Carpenter and W. Deutsch, "Reference states and reversals: undoing actions with verbs", *J. Child Language*, 22, pp. 633–652, 1995.
- H. Clark and C. Marshall, "Definite reference and mutual knowledge", in *Elements of Discourse Understanding*, B.W.A. Joshi and I. Sag, Eds, Cambridge, MA: Cambridge University Press, 1981.
- M.S. Cohen, S.M. Kosslyn, H.C. Breiter, G.J. DiGirolamo, W.L. Thompson, A.K. Anderson, S.Y. Bookheimer, B.R. Rosen, J.W. Belliveau, "Changes in cortical activity during mental rotation. A mapping study using functional MRI", *Brain*, 119, pp. 89–100, 1996.

- E. Colunga and L.B. Smith, "Committing to an ontology: a connectionist account", in paper presented at the *Cognitive Science Society*, Boston, 2000.
- D.P. Corina, J. Vaid and U. Bellugi, "The linguistic basis of left hemisphere specialization", *Science*, 255, pp. 1258–1260, 1992.
- A. Damasio, *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*, New York: Harcourt Brace, 1999.
- C. Darwin, *The Descent of Man and Selection in Relation to Sex*, London: John Murray, 1871.
- F. de Saussure, *Course in General Linguistics*, New York: McGraw-Hill, 1966.
- E. Dromi, *Early Lexical Development*, New York: Cambridge University Press, 1987.
- J. Du Bois, "The discourse basis of ergativity", *Language*, 63, pp. 805–856, 1987.
- J.L. Elman, "The emergence of language: a conspiracy theory", in *The Emergence of Language*, B. MacWhinney, Ed., Mahwah, NJ: Lawrence Erlbaum, 1999, pp. 1–28.
- J. Feldman and S. Narayanan, "Embodied meaning in a neural theory of language", *Brain Language*, 89, pp. 385–392, 2004.
- D. Gentner, "The development of relational category knowledge", in *Building Object Categories in Developmental Time*, L. Gershkoff-Stowe and D. Rakison, Eds, Mahwah, NJ: Lawrence Erlbaum, 2005, pp. 245–276.
- D. Gentner and A. Markman, "Structure mapping in analogy and similarity", *Am. Psychol.*, 52, pp. 45–56, 1997.
- T. Givón, "On the co-evolution of language, mind and brain", *Evolut. Commun.*, 2, pp. 45–116, 1998.
- A.E. Goldberg, "The emergence of the semantics of argument structure constructions", in *The Emergence of Language*, B. MacWhinney, Ed., Mahwah, NJ: Lawrence Erlbaum, 1999, pp. 197–213.
- R. Golinkoff, C. Mervis and K. Hirsh-Pasek, "Early object labels: the case for a developmental lexical principles framework", *J. Child Language*, 21, pp. 125–155, 1994.
- M. Gopnik and M.B. Crago, "Familial aggregation of a developmental language disorder", *Cognition*, 39, pp. 1–50, 1990.
- F. Grosjean and J. Miller, "Going in and out of languages", *Psychol. Sci.*, 5, pp. 201–206, 1994.
- S. Grossberg, "A theory of human memory: self-organization and performance of sensory-motor codes, maps, and plans", *Prog. Theor. Biol.*, 5, pp. 233–374, 1978.
- P. Gupta and B. MacWhinney, "Vocabulary acquisition and verbal short-term memory: computational and neural bases", *Brain Language*, 59, pp. 267–333, 1997.
- J.E. Hanna and M.K. Tanenhaus, "Pragmatic effects on reference resolution in a collaborative task: evidence from eye movements", *Cognitive Sci.*, 28, pp. 105–115, 2004.
- M. Harris, M.D. Barrett, D. Jones and S. Brookers, "Linguistic input and early word meaning", *J. Child Language*, 15, pp. 77–94, 1988.
- M. Hauser, E. Newport and R. Aslin, "Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins", *Cognition*, 78, pp. B53–B64, 2001.
- D. Hebb, *The Organization of Behavior*, New York: Wiley, 1949.
- G.E. Hinton and S.J. Nowlan, "How learning can guide evolution", *Complex Syst.*, 1, pp. 495–502, 1987.
- G. Houghton, "The problem of serial order: a neural network model of sequence learning and recall", in *Current Research in Natural Language Generation*, R. Dale, C. Mellish and M. Zock, Eds, London: Academic Press, 1990, pp. 287–319.
- D. Hubel and T. Weisel, "Receptive fields of cells in striate cortex of very young, visually inexperienced kittens", *J. Neurophysiol.*, 26, pp. 994–1002, 1963.
- M.K. Huffman, "Phonetic patterns of nasalization and implications for feature specification", in *Phonetics in Phonology V*, M.K. Huffman and P.A. Krakow, Eds, San Diego, CA: Academic Press, 1993, pp. 71–92.
- E. Hunt, *Concept Learning: An Information Processing Approach*, New York: Wiley, 1962.
- J. Huttenlocher, "The origins of language comprehension", in *Theories in Cognitive Psychology: The Loyola Symposium*, R. Solso, Ed., Potomac, MD: Lawrence Erlbaum, 1974, pp. 331–388.
- M. Jeannerod, *The Cognitive Neuroscience of Action*, Cambridge, MA: Blackwell, 1997.
- M. Johnson, "Imprinting and the development of face recognition: from chick to man", *Curr. Directions Psychol. Sci.*, 1, pp. 52–55, 1992.
- P.W. Jusczyk, *The Discovery of Spoken Language*, Cambridge, MA: MIT Press, 1997.
- P.W. Jusczyk, A.M. Jusczyk, L.J. Kennedy, T. Schomberg and N. Koenig, "Young infants' retention of information about bisyllabic utterances", *J. Exp. Psychol.: Hum. Percept. Perform.*, 21, pp. 822–836, 1995.
- R. Kager, *Optimality Theory*, New York: Cambridge University Press, 1999.
- S. Kakei, D.S. Hoffman and P.L. Strick, "Muscle and movement representations in the primary motor cortex", *Science*, 285, pp. 2136–2139, 1999.
- E.R. Kandel and R.D. Hawkins, "The biological basis of learning and individuality", *Scientific American*, 266, pp. 40–53, 1992.
- P. Kanerva, "Sparse distributed memory and related models", in *Associative Neural Memories: Theory and Implementation*, M. Hassoun, Ed., New York: Oxford University Press, 1993, pp. 50–76.
- J. Keenan, B. MacWhinney and D. Mayhew, "Pragmatics in memory: a study in natural conversation", *J. Verbal Learn. Verbal Behav.*, 16, pp. 549–560, 1977.
- F.C. Keil, *Concepts, Kinds, and Cognitive Development*, Cambridge, MA: MIT Press, 1989.
- B. Keysar, D.J. Barr and W.S. Horton, "The egocentric basis of language use: insights from a processing approach", *Curr. Directions Psychol. Sci.*, 7, pp. 46–50, 1998.
- T. Kohonen, *Self-organizing Maps*, 3rd edn, Berlin: Springer, 2001.

- G. Lakoff, *Women, Fire, and Dangerous Things*, Chicago: Chicago University Press, 1987.
- P. Li, I. Farkas and B. MacWhinney, "Early lexical development in a self-organizing neural network", *Neural Net.*, 17, pp. 1345–1362, 2004.
- P. Li and B. MacWhinney, "Cryptotype, overgeneralization, and competition: a connectionist model of the learning of English reverse prefixes", *Connection Sci.*, 8, pp. 3–30, 1996.
- K.Z. Lorenz, "The evolution of behavior", *Scientific American*, 199, pp. 95–104, 1958.
- P. MacNeilage, "The frame/content theory of evolution of speech production", *Behav. Brain Sci.*, 21, pp. 499–546, 1998.
- B. MacWhinney, "Pragmatic patterns in child syntax", *Stanford Papers Rep. Child Language Dev.*, 10, pp. 153–165, 1975a.
- B. MacWhinney, "Rules, rote, and analogy in morphological formations by Hungarian children", *J. Child Language*, 2, pp. 65–77, 1975b.
- B. MacWhinney, "Basic syntactic processes", in *Language Acquisition: Vol. 1. Syntax and Semantics*, S. Kuczaj, Ed., Hillsdale, NJ: Lawrence Erlbaum, 1982, pp. 73–136.
- B. MacWhinney, "Where do categories come from?", in *Child Categorization*, C. Sophian, Ed., Hillsdale, NJ: Lawrence Erlbaum, 1984, pp. 407–418.
- B. MacWhinney, "Competition and lexical categorization", in *Linguistic Categorization*, R. Corrigan, F. Eckman and M. Noonan, Eds, Philadelphia: Benjamins, 1989, pp. 195–242.
- B. MacWhinney (Ed.), *The Emergence of Language*, Mahwah, NJ: Lawrence Erlbaum, 1999.
- B. MacWhinney, "The gradual evolution of language", in *The Evolution of Language*, B. Malle and T. Givón, Eds, Philadelphia: Benjamins, 2002, pp. 233–264.
- B. MacWhinney, "A multiple process solution to the logical problem of language acquisition", *J. Child Language*, 31, pp. 883–914, 2004.
- B. MacWhinney, "The emergence of grammar from perspective", in *The Grounding of Cognition: The Role of Perception and Action in Memory, Language, and Thinking*, D. Pecher and R.A. Zwaan, Eds, Mahwah, NJ: Lawrence Erlbaum, 2005.
- B. MacWhinney, "A unified model of language acquisition", in *Handbook of Bilingualism: Psycholinguistic Approaches*, J.F. Kroll and A.M.B. de Groot, Eds, New York: Oxford University Press, 2005, pp. 49–67.
- B. MacWhinney, H.M. Feldman, K. Sacco and R. Valdes-Perez, "Online measures of basic language skills in children with early focal brain lesions", *Brain Language*, 71, pp. 400–431, 2000.
- B. MacWhinney and J. Leinbach, "Implementations are not conceptualizations: revising the verb learning model", *Cognition*, 29, pp. 121–157, 1991.
- B. MacWhinney, J. Leinbach, R. Taraban and J. McDonald, "Language learning: cues or rules?", *J. Mem. Language*, 28, pp. 255–277, 1989.
- B. MacWhinney and C. Pléh, "The processing of restrictive relative clauses in Hungarian", *Cognition*, 29, pp. 95–141, 1988.
- E. Markman, *Categorization and Naming in Children: Problems of Induction*, Cambridge, MA: MIT Press, 1989.
- D. Marocco, A. Cangelosi and S. Nolfi, "The emergence of communication in evolutionary robots", *Phil. Trans. R. Soc. London, Ser. A*, 361, pp. 2397–2421, 2003.
- A. Martin, C.L. Wiggs, L.G. Ungerleider and J.V. Haxby, "Neural correlates of category-specific knowledge", *Nature*, 379, pp. 649–652, 1996.
- D. Massaro, *Speech Perception by Ear and Eye*, Hillsdale, NJ: Lawrence Erlbaum, 1987.
- J.L. McClelland, B.L. McNaughton and R.C. O'Reilly, "Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory", *Psychol. Rev.*, 102, pp. 419–457, 1995.
- J.L. McDonald and B.J. MacWhinney, "The time course of anaphor resolution: effects of implicit verb causality and gender", *J. Mem. Language*, 34, pp. 543–566, 1995.
- W. Merriman, "Competition, attention, and young children's lexical processing", in *The Emergence of Language*, B. MacWhinney, Ed., Mahwah, NJ: Lawrence Erlbaum, 1999, pp. 331–358.
- C. Mervis, "Early lexical development: the contributions of mother and child", in *Origins of Cognitive Skills*, C. Sophian, Ed., Hillsdale, NJ: Lawrence Erlbaum, 1984, pp. 339–370.
- K. Miller, J. Keller and M. Stryker, "Ocular dominance column development: analysis and simulation", *Science*, 245, pp. 605–615, 1989.
- K.D. Miller, "A model for the development of simple cell receptive fields and the ordered arrangement of orientation columns through activity-dependent competition between ON- and OFF-center inputs", *J. Neurosci.*, 14, pp. 409–441, 1994.
- J.D. Murray, "How the leopard gets its spots", *Scientific American*, 258, pp. 80–87, 1988.
- L.G. Naigles and S.A. Gelman, "Overextensions in comprehension and production revisited: preferential looking in a study of dog, cat, and cow", *J. Child Language*, 22, pp. 19–46, 1995.
- J.J. Ohala, "Phonetic explanation in phonology", in *Papers from the Parasession on Natural Phonology*, A. Bruck, R. Fox and M. La Galy, Eds, Chicago: Chicago Linguistic Society, 1974, pp. 251–274.
- L.M. Parsons, P.T. Fox, J.H. Downs, T. Glass, T.B. Hirsch, C.C. Martin, P.A. Jerabek and J.L. Lancaster, "Use of implicit motor imagery for visual shape discrimination as revealed by PET", *Nature*, 375, pp. 54–58, 1995.
- J. Piaget, *The Construction of Reality in the Child*, New York: Basic Books, 1954.
- S. Pinker, "Rules of language", *Science*, 253, pp. 530–535, 1991.

- S.R. Quartz and T.J. Sejnowski, "The neural basis of cognitive development: a constructivist manifesto", *Behav. Brain Sci.*, 20, pp. 537–596, 1997.
- W.V.O. Quine, *Word and Object*, Cambridge, MA: MIT Press, 1960.
- P.T. Quinlan, *Connectionist Models of Development: Developmental Processes in Real and Artificial Neural Networks*, Hove: Psychology Press, 2003.
- T. Regier, "The emergence of words: from associations to symbols", *Cognitive Science* (in press).
- S. Reznick, "Visual preference as a test of infant word comprehension", *Appl. Psycholinguist.*, 11, pp. 145–166, 1990.
- G. Rizzolatti, L. Fadiga, V. Gallese and L. Fogassi, "Premotor cortex and the recognition of motor actions", *Cognitive Brain Res.*, 3, pp. 131–141, 1996.
- E. Rosch and C.B. Mervis, "Family resemblances: studies in the internal structure of categories", *Cognitive Psychol.*, 7, pp. 573–605, 1975.
- D.E. Rumelhart and J.L. McClelland, *Parallel Distributed Processing*, Cambridge, MA: MIT Press, 1986.
- M. Sabbagh and S. Gelman, "Buzzsaws and blueprints: what children need (or don't need) to learn language", *J. Child Language*, 27, pp. 715–726, 2000.
- J.R. Saffran, R.N. Aslin and E.L. Newport, "Statistical learning by 8-month-old infants", *Science*, 274, pp. 1926–1928, 1996.
- L.K. Samuelson and L.B. Smith, "Memory and attention make smart word learning: an alternative account of Akhtar, Carpenter, and Tomasello", *Child Dev.*, 69, pp. 94–104, 1998.
- G. Schafer and K. Plunkett, "Rapid word learning by 15-month-olds under tightly controlled conditions", *Child Dev.*, 69, pp. 309–320, 1998.
- N. Schmajuk and J. DiCarlo, "Stimulus configuration, classical conditioning, and hippocampal function", *Psychol. Rev.*, 99, pp. 268–305, 1992.
- M.F. Schober, F.G. Conrad and S.S. Fricker, "Misunderstanding standardized language in research interviews", *Applied Cognitive Psychology*, 17, pp. 22–41, 2003.
- P. Schyns, "A modular neural network model of concept acquisition", *Cognitive Sci.*, 15, pp. 461–508, 1991.
- A. Sfard and K. McClain, "Special issue: analyzing tools: perspective on the role of designed artifacts in mathematics learning", *J. Learn. Sci.*, 11, pp. 153–388, 2002.
- J.F. Shrager and M.H. Johnson, "Waves of growth in the development of cortical function: a computational model", in *Maturational Windows and Adult Cortical Plasticity*, B. Julesz and I. Kovacs, Eds, New York: Addison-Wesley, 1995, pp. 31–44.
- H. Simon, *Sciences of the Artificial*, Cambridge, MA: MIT Press, 1969.
- L. Smith, "Children's noun learning: how general processes make specialized learning mechanisms", in *The Emergence of Language*, B. MacWhinney, Ed., Mahwah, NJ: Lawrence Erlbaum, 1999, pp. 277–304.
- L. Steels and F. Kaplan, "AIBO's first words: the social learning of language and meaning", *Evolut. Commun.*, 4, pp. xx, 2000.
- J. Stiles-Davis, S. Sugarman and R. Nass, "The development of spatial and class relations in four young children with right-cerebral-hemisphere damage: evidence for an early spatial constructive deficit", *Brain Cognit.*, 4, pp. 388–412, 1985.
- B. Story, "An overview of the physiology, physics and modeling of the sound source for vowels", *Acoust. Sci. Technol.*, 23, pp. 195–206, 2002.
- S. Sugarman, "Developmental change in early representational intelligence: evidence from spatial classification strategies and related verbal expressions", *Cognitive Psychol.*, 14, pp. 410–449, 1982.
- M.K. Tanenhaus, J.S. Magnuson, D. Dahan and C. Chambers, "Eye movements and lexical access in spoken-language comprehension: evaluating a linking hypothesis between fixations and linguistic processing", *J. Psycholinguist. Res.*, 29, pp. 557–580, 2000.
- I.R. Titze, B. Story, M. Smith and R. Long, "A reflex resonance model of vocal vibrato", *J. Acoust. Soc. Am.*, 111, pp. 2272–2282, 2002.
- M. Tomasello, *The Cultural Origins of Human Communication*, New York: Cambridge University Press, 1999.
- M. Tomasello, "Do young children have adult syntactic competence?", *Cognition*, 74, pp. 209–253, 2000.
- M. Tomasello, *Constructing a First Language: A Usage-based Theory of Language Acquisition*, Cambridge: Harvard University Press, 2003.
- M. Tomasello and N. Akhtar, "Two-year-olds use pragmatic cues to differentiate reference to objects and actions", *Cognitive Dev.*, 10, pp. 201–224, 1995.
- M. Vihman, *Phonological Development: The Origins of Language in the Child*, Cambridge, MA: Blackwell, 1996.
- C.H. Waddington, *The Strategy of the Genes*, New York: MacMillan, 1957.
- B. Whorf, "Some verbal categories of Hopi", *Language*, 14, pp. 275–286, 1938.
- B. Whorf, "The relation of habitual thought and behaviour to language", in *Language, Culture, and Personality: Essays in Memory of Edward Sapir*, L. Spier, Ed., Ogden, UT: University of Utah Press, 1941, pp. 75–93.

