

The emergence of linguistic complexity

Brian MacWhinney
Carnegie Mellon University

Linguists have often argued that recursion produces linguistic complexity. However, recursion relies on preexisting processes such as lexical insertion, lexical combination, memory stacks, and methods of interpretation. In the brain, recursion is an emergent property of a set of adaptations that involve at least six processing systems. Linguistic complexity arises from the interplay of all six of these systems. The complexity of this neuronal support means that the full complexity of human language could not have arisen fortuitously at some single moment in evolution. However, there is evidence that some pieces of the six systems supporting complexity have developed more recently than others.

Human language is enormously complex. Each of us knows tens of thousands of words, and each of these words can specify a microcosm of meaning. When we put these words together into sentences and discourse, still further complexities arise in the form of collocations, phrases, and grammatical relations. Of the many forces that generate all this complexity, the one that has received the greatest attention is the process of recursion. Beginning with a seminal article by Bar-Hillel (1953), linguists have argued that recursion allows human beings to produce an infinite variety of possible sentences. Extending this vision, Chomsky and others have claimed that what makes human language distinct from the communication systems of all other animals is the fact that only humans have access to recursion.

For decades, the characterization of recursion and its effects remained a task for linguists or computer scientists, with little connection to issues in biology. However, recently, Hauser, Chomsky, & Fitch (2002) (henceforth HC&F) hypothesized a link between recursion and human evolution. They distinguished between the faculty of language broadly defined (FLB) and the faculty of language narrowly defined (FLN). Within the FLB, there are many properties that human language shares with other communication systems. For example, both birds and humans rely on a vocal apparatus to produce a broadcast transmission of sound waves that is then decoded by an auditory apparatus. In both birds and humans, all members of the species possess the mechanisms for both production and perception. More generally, even outside the FLB, humans and other mammals share complex systems of social organization, memory systems, and methods for learning.

For HC&F, none of these shared mechanisms are central to the faculty of human language, narrowly defined. They must be present for language to operate, but what is at the core of human language is recursion.

This analysis places a heavy conceptual weight on the process of recursion. Given its centrality, it is surprising to find that HC&F fail to provide a clear account of how recursion operates linguistically.¹ Instead, they simply note that recursion “takes a finite set of elements and yields a potentially infinite arrangement of discrete expressions.” In other words, recursion is whatever it takes to produce infinite generativity. But, if we are to understand recursion as a neurolinguistic process,² we need to find out not just what it produces, but also how it operates. It is easy enough to provide some clear cases of the syntactic effects of recursion. As Bickerton (this volume), Shibatani (this volume), and others have noted, sentences with relative clauses provide just such clear examples. In the initial version of transformational grammar (Chomsky, 1957), a sentence like (3) was produced by a transformation that combined (1) and (2).

- (1) The boy has a black jacket.
- (2) The boy stole the bike.
- (3) The boy who stole the bike has a black jacket.

However, in order to constrain the global power of transformations, Chomsky (1965) shifted away from a transformational view of recursion to a theory that generated recursion directly inside phrase structure rules, using the notion of a generalized phrase marker. In the case of example sentence (3), the insertion of the relative clause was then handled through rules like (4) through (6).

- (4) $S \rightarrow NP + VP$
- (5) $VP \rightarrow V + NP$
- (6) $NP \rightarrow (D) N' (S')$

It is the presence of the symbol “S” on the right-hand side of rule (6) that produces the insertion of a relative clause as the modifier of “the boy”. This process of

1. Tomalin (2007) shows how recursion can be characterized in terms of computability, primitive recursion, general recursion, lambda-definability, and inductive definition. Following Tomalin and Bar-Hillel (1953), this paper accepts the definition of recursion as based on the process of inductive analysis.

2. There are many examples of the operation of recursion and iteration in biological systems outside of communication. The codes specified in the DNA have an embedded and hierarchical structure. Mechanisms and structures are repeated throughout our physiology. So, the issue at stake here is not the importance of recursion to biology, but whether humans are the only animals that utilize recursion in communication.

recursive insertion could go on forever, with sentences like (4) being expanded to (7) and so on.

- (7) The boy who stole the bike that the teacher had painted has a black jacket.

In this example, the subject “boy” is separated from its predicate “has a black jacket” by intervening material that could, in turn, be expanded indefinitely. Of course, if this expansion goes on too long, we will lose track of the identity of the subject and eventually the sentence will become more and more difficult to interpret, as in (8)

- (8) The boy who stole the bike that the teacher Mary saw at the train station had painted has a black jacket.

Center-embedded structures become even more difficult when they are composed of common nouns with definite articles and confusable verbs, as in

- (9) The dog the cat the boy the girl liked saw chased ran.

In practice, it is often difficult to distinguish recursion from iteration. Consider sentences such as (10) and (11).

- (10) Mary invited Tim’s mother to dinner.
- (11) Mary invited Tim’s mother’s friend to dinner.
- (12) Mary invited Tim’s mother’s friend’s family to dinner.

The process of expanding these possessive phrases can be viewed as either recursion as in (13) or as iteration as in (14)

- (13) Mary invited (((Tim’s mother’s) friend’s) family) to dinner.
- (14) Mary invited Tim’s + mother’s + friend’s + family to dinner.

In logical terms, we might view (13) as a superior representation. However, because interpretation is incremental, the resultant understanding may not be very different in the two cases. This contrast between iteration and repetition also arises for structures such as (15) with repeated complementation or (16) with right-branching relatives.

- (15) Mary said that Bill promised to tell Frank to ask Sarah to suggest to ...
- (16) This is cat that chased the rat that scared the boy who

There is an important difference between right-branching sentences like (13) through (16) and center-embedded cases like (7) through (9). The right-branching structures place little burden on working memory, since there are no discontinuous grammatical relations. Instead, words can be assigned to correct grammatical roles in an incremental fashion, as they occur. With center embedding, on the other hand, the embedded material directly interferes with incremental assignment of grammatical roles, placing

a major burden on working memory. Similar problems occur with left-branching structures, such as those found in languages like Japanese that place the relative clause before the noun. In general, it seems that an iterative account for the generation of complex structures makes most sense when there is no interruption of constituents, suggesting that recursion involves more than just syntactic patterns.

These facts show that neither recursion nor iteration could plausibly be based on any single simple neuronal device. In fact, we will see that the infinite complexity of human language arises from the interplay between at least six major linguistic subsystems.

1. Six subsystems — six theories

We can begin this exploration by recognizing the importance in both linguistic and neural terms of the six basic levels of linguistic analysis: auditory phonology, articulatory phonology, lexicon, syntax, phrasal storage, and discourse. The account provided here attributes the computation of recursion to mechanisms operating within each of these six levels or subsystems. Table 1 presents an overview of this levels-based analysis of the emergence of recursion.

| Subsystem | Area | Processes | Theory |
|---------------|-------------------|-----------------------------|----------------------|
| Audition | Auditory cortex | Extracting units | Statistical learning |
| Articulation | IFG, motor cortex | Targets, timing | Resonance, gating |
| Lexicon | Wernicke's | Phonology to meaning | DevLex |
| Syntax | IFG | Slots, sequences | Item-based patterns |
| Storage | DLPFC | Binding, competition, lists | Attachment, roles |
| Mental Models | Dorsal cortex | Deixis, perspective | Perspective |

The core of the argument being developed here is as follows. The first three levels work together to construct a system for lexical learning and processing. To build this system, the child must work for years to extract units through audition (Jusczyk, 1997), to link articulation to audition (Oller, 2000), and to establish articulatory fluency. Without the consolidation of these peripheral processes, lexical processing would be impossible. Within the lexicon, self-organization links the topographic structure of posterior areas (including Wernicke's area) to processors in the inferior frontal gyrus (IFG) (Grodzinsky & Friederici, 2006). The architecture of these processors can produce generativity through iteration and recursion, but without the further systems of storage and mental models, their operation would be vacuous. Storage mechanisms keep words and phrases in short term memory while the Merge operation continues to mediate between syntax and the lexicon. This processing involves a three-way

coordination. This coordination is then further modulated by the construction of embodied mental models (MacWhinney, 2008b) that can then offload phrases from temporary storage.

This analysis shows that linguistic complexity, including recursion, requires the complete participation of all six of these linguistic levels. The development of this system was not a sudden, nearly accidental evolutionary event. Rather, the system has built upon six million years of development in a unique evolutionary niche occupied by a species that combined upright posture with tight, but competitive, social organization (MacWhinney, 2008a).

2. Systems vs. modules

Before exploring in detail the processes and structures operative on each of these six subsystems, we need to consider some basic issues regarding pseudo-modular organization in the brain. It is tempting to think of these six subsystems as computational modules (Fodor, 1983; Levelt, 1989; Pinker, 1997). Indeed, each of these subsystems utilizes uniquely adapted processors in localized brain regions. However, thinking of these processors as protected modules like those in a Java program is not in accord with what we know about the brain (Bullinaria, 2007). In fact, brain regions are heavily interconnected by asymmetric bidirectional connections. These connections cannot pass symbols, as required by the digital computer. Instead neurons must communicate by depending on isotopic mapping, learned patterns of connectivity, firing synchrony, and modulation through supervisory units. Moreover, except in lower organisms or the brainstems of higher organisms, it is seldom the case that a single cell is responsible for a discrete cognitive function. This is certainly true at the level of the cortex, where cells appear to operate in assemblies of thousands of neurons to achieve single cognitive goals.

Brain development in the fetus involves the migration of neurons from the germinal matrix to the periphery. During this migration, cortical areas maintain their connections to the various subcortical areas from which they differentiated. For example, within both the thalamus and the hippocampus, there are separate nuclei that project to separate cortical areas. Although single axons fire in a directional fashion, these larger sets of connections are bidirectional, thereby providing a system of reentrance or interaction between areas of the brain (Tucker *et al.*, this volume). Within each of these thalamic or hippocampal nuclei there may be additional fine-grained structure that allows the subcortical area to maintain a map of the structure of the cortical area, even after it has migrated to a more distal position. The map-like nature of these connections between cortical and subcortical structures is further supplemented by map-like connections of motor and sensory areas to external sense organs and the body. In motor

cortex, there is a somatotopic organization that matches up well with the actual shape of the human body. In sensory areas, cortex is organized to represent the features of the sense. For example, auditory cortex is organized in terms of frequencies, as detected by the neighboring cells in the cochlea. Similarly, the visual cortex is organized in patterns that maintain the position of receptors in the left or right visual field and other peripheral patterns.

This map-like organization of the brain allows areas to communicate in terms of an embodied neural code that is implicit in the position of a neuron within the map. As processing moves away from the periphery, the blending of these codes increases. However, through reentrance, it is possible for the brain to ground cognition on these original body maps (Jeannerod, 1997; Schütz-Bosbach & Prinz, 2007; Wilson & Knoblich, 2006). This basic principle of map-like organization also applies to learning methods that function to organize local maps. One powerful way of modeling this local organization relies on the self-organizing feature maps (SOFM) of Kohonen (1990). In this model, neuron-like units are organized in two-dimensional sheets with connections to an array of input and output features. When an input feature vector is activated, units in the map also gain some activation. Through lateral inhibition, the most strongly activated unit will inhibit its neighbors, leading to a winner-take-all effect. This pattern of activity has been well documented for cortical structures. After the initial inhibition, there is then a learning phase in which the connections with the winner and its neighbors are strengthened. As a result of this learning, responses to certain patterns in the input tended to become parceled out across areas of the feature map, with this self-organized differentiation increasing over time. This type of map is a sparse, distributed memory, since there are typically many possible features of which only a few are active for a given input.

2.1 Subsystem 1: auditory organization

Research on the emergence of auditory processing abilities has identified three major streams of auditory learning and development. The first stream involves the core perceptual features of hearing. These features are derived from the anatomy and physiology of the ear, the cochlea, the auditory nerve, and the auditory cortex. These peripheral structures are shared in common by humans and other mammals. For example, Kuhl and Miller (1975, 1978) demonstrated that chinchillas are able to process the contrast between voiced stops like /b/ and unvoiced stops like /p/ in a categorical, humanlike way.

The second stream of auditory development is also shared, at least in part, with our primate cousins. This is the ability to process the statistical regularities of spoken language. There is now abundant evidence that both infants (Aslin, Saffran, & Newport, 1999; Marcus, 2000) and monkeys (Hauser, Newport, & Aslin, 2001) are able to use sequential statistics to extract background properties of their

language involving prosodies, phonotactics, and possible segments. This learning allows infants to orient to speech and to develop codes for storing sequences of syllables.

The third stream of auditory development involves the actual storage of sounds as potential candidate words. By six months, children are able to respond to their own name. Using the auditory codes they have developed, they can retain large numbers of traces of words they have heard, although often they have not yet learned what these words might mean.

2.2 Subsystem 2: articulatory organization

Each of the three streams of auditory processing discussed above could, in theory, proceed without reference to the articulatory system. However, the human phonological system emerges through a process of vocal imitation that couples audition to articulation. Before infants can display real imitation of conventional forms, they must first practice and solidify their control over the basic production processes of speech production. Oller (2000) shows that, during the first six months, the infant is busy consolidating control over the use of the glottis to achieve consistent phonation. During this period, the infant works to control volume, pitch, and the overall contour of vocalizations. Once this achievement is consolidated, the next goal is the linkage of phonation to articulation by control of the jaw and the oral articulators. This linkage develops during the period of repetitive babbling, as the infant works to develop methods for consistently producing target syllables and syllable sequences. As Macneilage & Davis (2000) have noted, the predominant template in babbling is the CV syllable that has a structure that derives from the close-open lip-smacking gesture in other primates. Using this basic gesture, infants begin to forge a resonant linkage or coupling between the processes of articulation and audition. Piaget (1954) described this linkage as a circular reaction. Basically, the child is learning how to control articulation in order to reliably produce important target sounds.

During the last months of the first year, babbling continues to deepen the linkage between articulation and audition. While this is happening, the child is paying ongoing attention to the encoding of statistical regularities in the input language. The linkage of articulation to audition then places a further set of constraints onto the ways in which patterns are detected, since the infant is now trying to actually imitate aspects of the target language. For example, by the end of the first year, the child learning a tone language such as Mandarin will be trying to develop ways of producing each of the four Mandarin tones.

There are now several well-articulated models of the development of linkages between articulation and audition. These models (Guenther, 2006; Lacerda, 1998; Westermann & Miranda, 2004) emphasize the ways in which units in auditory space come to be constrained by the possible articulations produced by the child.

Lindblom et al. (in press) argue that, to work effectively across contexts, this learning has to rely on the fact that the brain is able to compute motor equivalence between targets. This allows the learner to link together a targeted sound with an auditory effect, even when the actual details of the sound production vary radically across contexts. During this learning, the ambient language is providing ongoing input to auditory organization and this input provides new targets for the process of articulation. In the end, the structures encoded in auditory space become the strongest forces in this coupled system.

2.3 Subsystem 3: lexical organization

Li, Zhao, & MacWhinney (2007) have developed a model of lexical learning based on a self-organizing features model (SOFM). This model, called DevLex, uses three separate self-organizing feature maps for auditory phonology, articulatory phonology, and lexical structure. In effect, DevLex provides us with a fully implemented, neurologically grounded, empirically supported account of organization for these first three linguistic subsystems.

Featural organization on the DevLex auditory map relies on the PatPho representational system that parcels out segments into an autosegmental grid. For example, the word *stoops* is composed of an /st/ onset, the vowel nucleus /u/ and the coda /ps/. Within each of these components there are three slots for phonemes and each phoneme is coded in terms of a set of distinctive features. This model assumes that initial auditory processing has yielded a set of perceptual features that are associated with specific syllables, and slots (onset, nucleus, coda) within syllables. The PatPho system was introduced in MacWhinney & Leinbach (1991) and most subsequent work in neural network modeling of input phonology has used this framework. The activation of segments or syllables in a self-organizing feature map is further controlled through a sequence detection mechanism that expresses the form of a word as a linear trajectory through points in the feature map. Multiple positional variants of a given segment are represented as multiple neighboring nodes in the input phonology map. Because this map relies on the same SOFM framework used by the DIVA (Guenther, 2006) model, the DevLex account of phonological learning is generally compatible with DIVA. Output phonology is also represented through sequence control units that activate articulatory gestures organized in a second motor feature map. Figure 1 below shows the shape of DevLex.

The three separate maps of the DevLex model represent three of the six core linguistic modules. These modules are each located in separate brain regions, connected by axonal projections. DevLex trains these connections using Hebbian learning. However, we will see later that there is reason to believe that other processes are involved. Input phonology is processed in the auditory cortex of the superior temporal sulcus. Output phonology is controlled by parts of Broca's area, along nearby regions

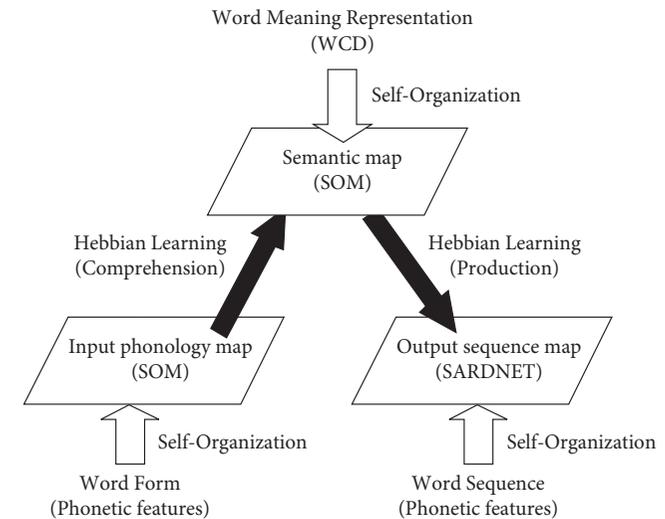


Figure 1. The architecture of the DevLex system.

with motor cortex. The core semantic or lexical map is centered in Wernicke's area, although it is actually far more generally distributed, as we will see later.

Looking first at the control of input phonology, we know that this processing is focused in primary auditory cortex. This area, which spans Brodmann areas 41 and 42, lies in the posterior half of the superior temporal gyrus and the transverse temporal gyri or Heschl's gyri. Within this area, there are in fact multiple tonotopic maps, each of which appears to represent a different view or processing slant on the whole range of the frequency spectrum. Work with rhesus monkeys has shown that the auditory system involves three levels of auditory processing with 15 different tonotopic maps. This pattern of multiple parallel isotopically organized maps is similar to the pattern of multiple parallel maps found in the motor system. Like many other cortical areas, the auditory cortex is also connected to its own specific thalamic nucleus, the medial geniculate nucleus, from which it receives input.

Human auditory processing is fundamentally similar to that of other mammals and even birds. This suggests that the peripheral auditory system and the basic neural engine for auditory feature detection were consolidated prior to the evolution of hominids. Further neural processing can reduce the enormous complexity of the auditory world to a much smaller set of contrasts that can link to output phonology and lexical structure. This reduction of complexity is operative in other mammals. However, it is likely that, under the influence of linkage first to articulation and then to a lexicon,

these categorizational processes extend further and occupy additional neural machinery in humans.

In addition to an overall sharpening of contrasts, human and primate audition must also differ in the extent to which they rely on mechanisms for sequence detection. Although syllables can be perceived as wholes, multisyllabic words need to be encoded in ways that associate sounds with syllable position. Prosodic features, such as syllabic stress or moraic timing, can facilitate and sharpen this encoding, but it is likely that some form of sequence detection is involved in the interfacing of auditory processing with lexical recognition. These sequence detection processes may be present in other mammals, but they are probably elaborated in humans.

3. A distributed lexicon

Lexical representations are far more distributed than the representations of input and output phonology. The broad area of cortex at the intersection of the parietal, occipital, and temporal lobes has further access to wide areas of the whole cortex. Unlike the feature maps for input and output phonology, the core conceptual lexicon must make contact with a very diverse set of connections across the brain. For example, words for tools must make contact with the motor gestures and postures involved in the use of these tools. Words for fruits must make contact with the visual properties of these flowers, including colors, shapes, and smells. Words for actions must make contact with the motor sequences, perceptual changes, and object affordances involved in these actions. The choice between closely related meanings are determined by competitions within the relevant areas. For example, motor and parietal areas involved in tool usage can determine the competition between alternative tools such as *screwdriver* vs. *drill*. The competing cell assemblies within the lower level of this hierarchy then transfer activation back to higher-level units in the central lexical map that connects to phonology. Figure 2 presents a sketch of how this hierarchical organization can operate within a system of self-organizing feature maps (Dittenbach, Rauber, & Merkl, 2002).

To control this hierarchical access, the brain must rely on long-distance connections between the core lexical areas and areas that flesh out the meanings involved in words. Moreover, these hierarchical connections must be structured in a way that allows for a consistent control of competition at both the local areas and the lexical core.

4. Consolidation of lexical patterns

This distributed, hierarchical patterning has important consequences for the consolidation of linguistic complexity. Tucker *et al.* (this volume) argue that ventral stream

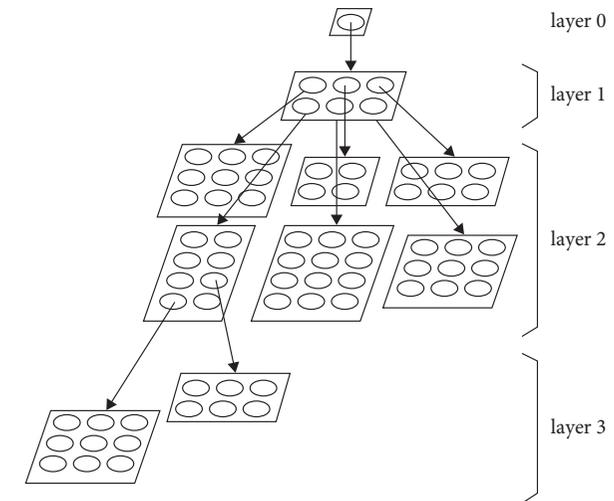


Figure 2. Hierarchical access in self-organizing feature maps.

processing operates upon discrete item-based object representations that are characteristic of processing in temporal cortex. This type of item-based encoding is supported by neostriatal attentional mechanisms and hippocampal reentrant encoding processes. The hippocampus provides a compressed encoding of the distributed patterns related to a word. By maintaining resonant and reentrant reactivation of these patterns, the hippocampus can facilitate the consolidation of these traces into a new cell assembly or lexical pattern. This ventral-hippocampal system provides the basic engine for consolidating and extending linguistic complexity at the lexical level. Here the complexity involves not just the phonological form of the word, but also the diverse connections of the lexical system to many areas of the brain. Because words have become conventionalized mimetic forms, this system then functions to repeatedly consolidate variant meaningful configurations into the same phonological bucket. From this core engine, arise the linguistic complexities of radial semantic structure (Lakoff, 1987), polysemic pathways (MacWhinney, 1989), metonymy, partonomy, and homonymy (Lyons, 1977).

Underneath this linguistic complexity, there is a further level of psycholinguistic complexity that arises from the distributed, hierarchical shape of lexical activation. A simple word like “hammer” is able to trigger both visual images of a hammer in the ventral “what” stream and functional images of wielding a hammer to hit a nail in the dorsal “how” stream. When we come to more complex words such as “grandfather” or “promise”, the meanings involved have to be unpacked in terms of a whole set of

embedded predicates, such as “the father of my father, or the father of my mother” or “tell someone that you will perform an action that you would not otherwise have done with the expectation that, if you fail to complete the action, there would be unpleasant social or interpersonal consequences and that you therefore fully intend to complete the action, even if certain barriers arise.”

The solidification of complexity at the lexical level relies on this system of distributed resonance. When we produce a complex word like “promise”, it is not necessary that all elements of the chain be fully activated in working memory. All that is necessary is that enough of the word be activated to guarantee correct lexicalization of this word as opposed to its competitors. In the case of “promise” it may be that all that is necessary is the notion of saying something seriously. The further pragmatic implications of the word may not be available initially in working memory. However, there are long-term, distributed links available that can call them up if needed. This is equally true for concrete terms such as “hammer”, since we do not always have to think about using the claw of a hammer to pull out a nail when we hear the word “hammer.” In this way, we can think of words as promissory notes or tokens that are issued in the place of the full set of concepts and stances with which they are linked.

5. Sequence analysis within the lexicon

In principle, it would be possible to ground a communication system on sentences or propositions compressed into single words. Polysynthetic languages such as Iroquois or Eskimo push hard in this direction with their inclusion of a wide range of moods, persons, surfaces, and aspects into a single verb-based complex. However, languages achieve this compression by relying on a variant method for controlling the sequencing of lexical items. Consider the contrast between Hungarian and English in the way they form the phrase meaning “my coat.” In English, the possessive appears as a separate word preceding the noun. Variations in the phonological shape of the following noun have minimal effect on the sound of the possessive pronoun. In Hungarian, on the other hand, the suffix *-om* takes on the shape of *-am*, *-om*, *-em*, *-öm*, or *-m*, depending on the shape of the stem. Moreover, the stem changes its shape, depending on the nature of the suffix.

The debate about the cognitive representation of these morphophonological patterns has raged for over three decades in psycholinguistics. The connectionists and analogists view forms such as *kabátom* as produced within the lexicon through interactive activation of analogic patterns. In this model, all lexical forms are produced within the lexicon, without reliance on external routes. The alternative view holds that regular morphological forms are produced by combination between stems and affixes. The third possible formulation is that of MacWhinney (1978, 1982, 1987a, 2005) which

views combinatorial forms as arising through extraction from a core analogic process. Within the framework of self-organizing feature maps, this means that a separate lexical map for affixes emerges from a process that compares similar morphological formations. For example the comparison of *shoe* with *shoes* will lead to the extract of *-s* as the initial productive form for the plural. Similarly, the comparison of *kabát* with *kabátom* leads to the extraction of *-om* as the first person possessive suffix.

This comparison and extraction method is clearly an important additional source of linguistic complexity. This same engine can work within the noun phrase or verb phrase to extract *my* from the combination *my coat*, just as *-om* is extracted from *kabátom*. Moreover, it is an engine that can work in both directions. If the pressures of fast speech work to modify combinations such as *going to* into *gonna*, then the latter can be stored as a single form representing what was earlier a syntactic combination.

The extreme analogist view would hold that the neurological basis of morphology is completely interwoven with the lexical substrate in Wernicke’s area at the juncture of the parietal and temporal lobes. It would view an item such as *-om* as occupying the same lexical map as an item such as *kabát*. One strength of this approach is that the morphological alterations involved in the complex word can be directly controlled through connections within the DevLex maps. However, a weakness in this approach is that it fails to capture the fact that the *-om* suffix occurs positionally after the stem. This problem becomes particularly severe in agglutinative languages where strings of two, three, or even more suffixes appear after the noun. To control these processes within a single local net, sequence detector units would have to be built into the lexical net itself. As an association area, Wernicke’s is not specialized for this type of sequential processing.

5.1 Subsystem 4: positional patterns

To solve this problem and to boost lexical capacity, evolution turned over combinatorial processing to Broca’s area in the inferior frontal gyrus (IFG). Among the various cortical areas specialized for sequence processing, this is the area that was closest to the posterior lexical areas. Grodzinsky & Friederici (2006) argue that the processing of positional patterns depends on connections between the anterior superior temporal gyrus (STG) and the inferior portion of Broca’s area lying in the pars opercularis (BA44). They distinguish this circuit from a parallel, but more dorsal one between posterior STG and the superior portion of Broca’s area in the pars triangularis (B45). Figure 3 below displays these two circuits.

The basic connectivity between these two circuits is already present in monkeys (Deacon, 1989). Therefore, there was no initial need to establish connectivity between the areas. In this sense, the syntactic engine was not built up from scratch. Rather, like all evolutionary advances, these circuits comprise a new machine made

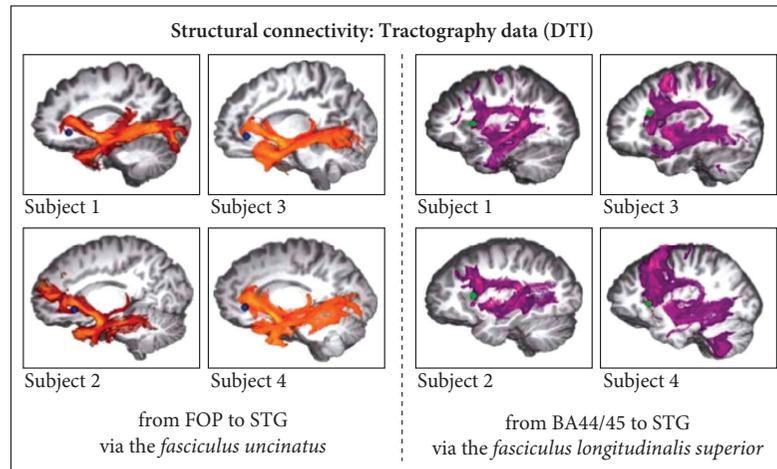


Figure 3. Anterior-posterior tracts measured through DTI.

Source: Friederici, Bahlmann et al., PNAS, 2006.

up of old parts. Within this new machine, there was a need to make sure that this connectivity supported effective control of lexical activation. To do this, it was important for lexical assemblies in posterior cortex to organize themselves in ways that map up with the already existent connections to IFG. Again, this is not some sudden evolutionary invention, but rather the reshaping of an old machine to serve new functions.

The linkage of posterior lexical areas to IFG relies crucially on the establishment of isotopic maps between the two areas. The DevLex model shows how the topological structuring of posterior cortex is achieved through movement of lexical forms on the self-organizing feature map. Figure 4 below illustrates the results of training the DevLex model on parental input derived from the Belfast corpus in CHILDES (MacWhinney, 2000). During this training, words that appear in similar contexts in the parental input self-organize so that they end up being located next to each other in lexical space. In other words, nouns end up next to other nouns and prepositions end up next to other prepositions. This topological self-organization provides support for reliable interactions between syntactic patterns in IFG and the posterior lexical map. In effect, the topological organization of the map is the backbone of a communication protocol between the lexicon and IFG. To understand how this protocol operates to produce complex syntactic structures, we will need to take an excursion into language acquisition theory.

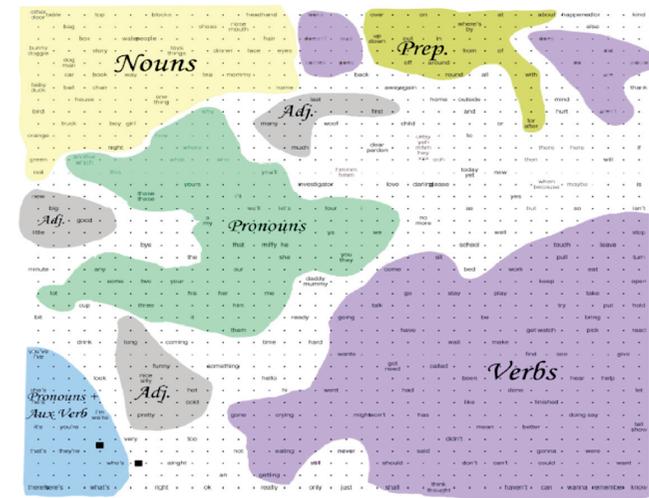


Figure 4. Self-organization in the DevLex model.

6. Item-based patterns

In the early days of acquisitional theory, Braine (1963, 1971) explored ways of applying learning theory to the study of child language. The formulation he devised focused on the idea that function words tend to appear in fixed positions *vis a vis* content words. For example, *the* appears before nouns and the suffix *-ing* appears after verbs. Like Harris (1951), Braine analyzed these constituent structures in terms of slots that could be filled by items of a certain class. Formulating a set of 12 such rules for a small corpus of child utterances, he referred to his account as a “pivot-open” grammar, since it specified the position of pivot words *vis a vis* the open class. Under the influence of Chomsky’s (1957) ideas about deep structure, this model was criticized as failing to pay adequate attention to semantic patterning. Later, Braine (1976) revised his account, emphasizing the role of “groping patterns” that established links based not on lexical class, but semantic relations.

Sticking more closely to Braine’s original formulations, MacWhinney (1975) introduced the notion of the item-based pattern. Applying this construct to a corpus of Hungarian, MacWhinney examined the word order of 11,077 utterances produced by two Hungarian children between the ages of 17 and 29 months. He found that between 85 and 100% of the utterances in these samples could be generated by a set of 42 item-based patterns. Some examples of these patterns in English translation are: *X + too, no + X, where + X, dirty + X, and see + X*. The item-based

pattern model was able to achieve a remarkably close match to the child's output, because it postulates an extremely concrete set of abilities that are directly evidenced in the child's output.

MacWhinney made no general claims about a pivot or open class, focusing instead on the idea that the first syntactic patterns involve links between individual lexical items and other words with which they are prone to combine. An example of an item-based pattern is the structure *the + X*. This pattern states simply that the word *the* occurs before another word with which it is semantically related. In addition to these positional facts, the item-based pattern encodes the shape of the words that can occupy the slot determined by *X* and the nature of the semantic relation between *the* and *X*. This is to say that an item-based pattern is a predicate-argument³ relation which encodes:

- the lexical identity of the predicate,
- the lexical category of the argument(s),
- the sequential position of the predicate vis a vis its argument(s), and
- the semantic relation between the predicate and its argument(s).

During language learning, the child generalizes item-based patterns to produce feature-based patterns or constructions. Feature-based patterns are no longer linked to specific lexical items, but instead apply to classes of items. From item-based patterns like *my + X* the child can extract the feature-based pattern *Possessor + Possession*. In this way, the child slowly pieces together the 23 major grammatical dependency relations of English, as summarized in the work on the GRASP parser (Sagae, Davis, Lavie, MacWhinney, & Wintner, 2007) for the CHILDES database. In this system, predicates can attach to as many as three arguments. Item-based constructions for verbs can also include the verbs of embedded clauses as arguments. And we will see below how item-based constructions for prepositions and auxiliaries include both an endohead and an exohead.

3. This paper uses the predicate-argument relation to describe item-based dependency patterns. This terminology is used to avoid confusions regarding the ways in which clusters inherit head features for X-bar syntax. In the noun phrase, predicates join with their heads to produce new clusters that inherit the features of the head noun. However, in verb phrases and prepositional phrases featural inheritance is driven by the predicate, not the arguments. Because of this, referring to the arguments as the head of a verb phrase would be confusing. The major danger involved in use of predicate-argument terminology for item-based patterns is the possibility that this would be interpreted as applying outside the domain of lexical combinations. Other levels of predicate-argument decomposition and combination exist throughout language and cognition and we are here only focusing on the role of the predicate-argument relation for combinations of words.

There is a third level of argument generalization, above the levels of the item-based pattern and the feature-based pattern. This is the level of the global construction. Just as feature-based constructions emerge from a process of generalization across item-based patterns, so global constructions emerge from generalization across feature-based constructions. For example, in English, there are literally dozens of verb groups that share a common placement of the subject before the verb. Together, these constructions give support for the SV global construction in English. The SV and VO global patterns of English work together to produce prototypical SVO order (MacWhinney, Bates, & Kliegl, 1984). Other languages promote different combinations of global patterns. In Hungarian and Chinese, for example, SV, OV, and VO orders operate to express alternative varieties of object definiteness, producing SVO and SOV orders. Italian combines SV and VO patterns with secondary, but significant use of VS (Dell'Orletta, Lenci, Montemagni, & Pirrelli, 2005) to produce SVO and VSO orders. Other global patterns control the ordering of topic before comment or the tendency to associate animacy with agency.

In addition to this process of generalization, positional patterns can be subjected to a process called composition. Composition takes two positional patterns and hooks them up into a single larger sequence. The important consequence of composition is that it increases the proceduralized nature of syntactic processing. For example, it may be that a single complex network, looking very much like a finite state automaton, processes all variants of noun phrases. In this network, there would be an initial slot for a quantifier, followed by a determiner or possessive, then a series of adjectives, and finally the noun. The compilation of smaller patterns into larger patterns of this type can proceduralize and facilitate both listening and production.

6.1 Subsystem 5: storage

Dependency grammars, such as GRASP, use structures equivalent to the IFG pattern detectors of the type outlined here. However, by themselves, dependency relations are not enough to achieve parsing or generation of longer strings of words. Some additional control mechanism is needed to allow for iteration or recursion. Here, there are two neurologically plausible approaches. One mechanism would compose X-bar structure and trees directly within IFG. However, neurological evidence for such embedded groupings of sequence processors is currently absent. Instead, current evidence suggests that areas outside of IFG are involved in the construction of larger conceptual trees from the sequential fragments detected by IFG. In accounts such as MacWhinney (1987b) or Gibson (1998), smooth processing relies on the incremental construction of interpretable units. Consider a sentence, such as *my coat has a missing button*. As soon as the sequence *my coat* is detected, the predicate is linked to its argument and the whole is then treated as a single cluster in the mental model being constructed.

Mental model construction proceeds in accord with the principle of starting points introduced by MacWhinney (1977) and supported in detail by Gernsbacher (1990). The starting point of *my coat* then becomes the perspective from which the rest of the sentence is interpreted. At this point, resonant activation involves the activation of items in posterior lexical space, activation of the meaning components of these items, and continued processing in IFG. Next, the sequential processor takes this whole active assembly as input to the verb-based frame for *have*. This predicate has argument slots for both a possessor perspective and an object possessed. Even before the second slot is filled, incremental processing activates a mental model expectation for a thing possessed. This mental model structure is structured in terms of the theory of ego-based perspective taking (MacWhinney, 2008b) described below. Next, the phrase *a missing button* is processed by the two relevant sequence processors and the result then fills the second slot of the verb *has*, thereby completing the mental model of a coat that has a missing button. Of course, the model itself may generate additional associated ideas. Perhaps the button is removed in some overt way; perhaps it is seen on the floor; or perhaps there is a focus on the thread left on the coat after the button has fallen off.

The filling of argument slots in feature-based patterns is driven by a series of cues that have been studied in detail in the context of the Competition Model of MacWhinney (1987a, 1987b). The model specifies a series of steps for the ways in which incremental processing triggers competition between constructions:

- Sounds are processed as they are heard in speech.
- Competition during sound processing controls activation of a current word.
- Each new word activates its own item-based patterns along with related feature-based patterns (see below).
- Item-based patterns then initiate tightly specified searches for slot fillers.
- Slots may be filled either by single words or by whole phrases. In the latter case, the attachment is made to the head of the phrase.
- To fill a slot, a word or phrase must receive support from cues for word order, prosody, affixes, or lexical class.
- If several words compete for a slot, the one with the most cue support wins.

Most work on the Competition Model has focused on comprehension, which is easier to control experimentally. However, the model applies equally well as an account for sentence production. The details of the operation of this parser are controlled by the competitions between specific lexical items and the cues that support alternative assignments. Consider the case of prepositional phrase attachment. Prepositions such as *on* take two arguments; the endohead is the object of the preposition, the exohead is the head of the prepositional phrase (i.e., the word or phrase to which the prepositional phrase attaches). Consider the sentence *the man positioned the coat on the rack*. Here, the endohead of *on* is *rack* and its exohead could be either *positioned* or *the coat*.

These two alternative attachment sites for the prepositional phrase are in competition with each other. For detailed examples of the step-by-step operations of this type of processor consult MacWhinney (1987a), MacDonald, Seidenberg, & Perlmutter (1994), or O'Grady (2005).

In this model, syntax involves nothing more than the repetitive clustering of the results of basic linear detectors. Of course, not all sentences are as simple as the one chosen to illustrate the basic process. Often uninterpreted arguments will build up on sentence memory waiting for merger with their predicates. MacWhinney & Pléh (1988) suggested that the capacity of memory for uninterpreted phrases was no greater than three and Gibson (1998) and others argue for a similar limit. But all analysts agree that there must be a mechanism for storing at least two or maybe three such uninterpreted items during processing. Because of its role in the phonological loop and other memory processes, there is reason to believe that dorsolateral prefrontal cortex (DLPFC) provides the necessary store for not-yet-merged items. This frontal mechanism then provides an additional engine for the maintenance and diversification of linguistic complexity.

But can a mechanism like this really control complex syntax? Don't we need the full power of transformational grammar, or at least context-sensitive phrase structure grammars? What about empty categories, traces, indices, interfaces, and so on? It is certainly important to explain how a linear mechanism of this type can compute complex structures. First, the basic generative power in this system comes from the fact that the slots of positional patterns refer to whole classes of items. Because of this, the power of the positional pattern machine is beyond that of finite-state processors that operate only on terminal symbols. As Hausser (1992) has shown, finite state grammars that operate on category symbols are formally equivalent to phrase-structure grammars. Second, the results of individual linear patterns can be combined or clustered through attachment in mental model space. As a result of this, the final model encodes full X-bar structure. Third, many of the linguistic phenomena that have been used to motivate complex syntax are actually better represented through memory processes in mental models, as discussed immediately below. Consider the case of the tangled dependencies caused by Dutch serial verbs or the English "respectively" construction. The fact that *John and Bill ordered steak and fish, respectively* can be interpreted best by a mnemonic device that establishes actual spatial positions in mental model space for John and Bill and then engages in the mental action of parceling out steak and fish to these positions in mental model space. This type of mental model processing is basic for anaphoric processing. There is no reason not to think that it is used to process these constructions too. Of course, the problem here is that, by itself, the syntax would not yield a complete parse tree in such cases. But that is because syntax is not doing this work alone.

6.2 Subsystem 6: mental models

Recent work in neuroscience has benefitted from four fundamental insights, each relating to the construction of mental models. First, in the 1980s, we learned that the visual system separates processing into an image-oriented ventral stream and an action-oriented dorsal stream.⁴ Second, we have learned from imaging work through the last decade that the brain relies on a perception-action cycle to interpret incoming messages. This cycle involves the generation of mental representations for objects in terms of the ways in which we typically act upon them (Knoblich, 2008). Much of this cycle is grounded on interactions that include the action-oriented processing of the dorsal stream. Third, we have learned that the brain provides specific mechanisms for mapping the body images of others onto ours. One consequence of this ability is the fact that “mirror” neurons (Rizzolatti, Fadiga, Gallese, & Fogassi, 1996) controlling actions, facial gestures, and postures can fire equally strongly when the actor is the self or the other. As we are now learning, these mirror systems are just one part of a general system for social cognition. The larger system also includes mechanisms for temporal facial processing (Pelphrey, Morris, & McCarthy, 2005) and eye contact (Pelphrey et al., 2003), as well as amygdala and striatal areas for empathy (Meltzoff & Decety, 2003) and projection (Adolphs & Spezio, 2006). Fourth, we have learned that the basal ganglia and hippocampus play a central role in the consolidation of memories, often driven by rewards and error minimization.

Piecing together these results, and following Tucker *et al.* (this volume), we can see that one of the additional consequences of the dorsal-ventral dichotomy is a shift of discrete processing of individual elements to the ventral stream and a shift of global model construction to the dorsal stream, with particular additional regulatory control from frontal areas. In recent papers (MacWhinney, 2005b, 2008b), I have suggested that this system provides the neurological basis for a system that constructs dynamic mental models from linguistic input. At the core of this system is the notion of the self as actor. During sentence interpretation, this fictive self is then projected onto the role of sentence subject, and the self reenacts the image underlying the sentence. These

4. Following Givon (1995), Hurford (2002) relates the separation of processing into the dorsal and ventral streams to the predicate-argument distinction in language. However, as Bickerton (2002) notes in his commentary to the Hurford’s article, this analysis fails in two important regards. First, predicates and arguments are not “raw sensory feeds” but rather complex lexical items that can themselves involve embedded predications, as we noted earlier in our discussion of words like “promise” or “grandfather”. Second Hurford’s model fails to provide a method by which the brain can integrate predicates and arguments. The mechanism proposed in the current paper is not linked in any clear way to the dorsal-ventral contrast, depending instead on interactions across IFG, distributed lexical processing, and frontal mechanisms for mental model construction.

images place the self into a set of well-understood roles as agent, experiencer, and source. Even locative and temporal relations can be interpreted from the egocentric frame that begins with a projection of the self onto the object located or the event in time.

Because narrative and dialog often involve rapid shifts between agents, this system has to be able to use linguistic devices to control perspective shifting. As a result of this core dynamics, we can refer to this system as the Perspective Shift System. This system constitutes the highest level of support for linguistic complexity. Without the mental model construction supported by this system, complex syntax would be useless. This is because the fundamental purpose of virtually all the devices of complex syntax is the marking of perspective shift. This analysis applies across all the major grammatical constructions, including passivization, relativization, clefting, pronominalization, dislocation, existentials, shift reference, split ergativity, serialization, complementation, conjunction, ellipsis, adverbialization, long-distance anaphora, reflexivization, PP-attachment, and participial ambiguity. Each of these structures allows the speaker to combine, maintain, and shift perspectives in communicatively important ways. And these devices allow the listener to trace these movements of the speaker’s attention across all of these shifts.

7. Building mental models

The traditional view of mental model construction (Budiu & Anderson, 2004; Kintsch, 1998) focuses on the linking of predicates into a coherent propositional graph. This activity is much like the process of clause-combining that we learned in classes in composition. For example, you can combine “the dog chased the bird” and “the bird flew away” to form “the dog chased the bird that flew away.” All one needs here is a grammatical device that serves to mark the fact that *the bird* plays a role in both clauses. Language provides a variety of methods for clause linkage, including conjunction, complementation, relativization, subordination, and adverbialization, as illustrated in (17)–(22):

- (17) You go down to Shattuck, and the bakery is on the corner.
- (18) Jim wanted me to bring him a loaf of bread.
- (19) The bread I bought was unfortunately stale.
- (20) Although it was stale, John wanted to pay me.
- (21) Shaking my head, I accepted the money.
- (22) My refusal of the payment would have made him upset.

These different methods link together clauses in terms of spatial, temporal, causal, and anaphoric relations. Most of these methods rely on both lexical and syntactic processes. Typically, there is a linking lexical element, such as a conjunction or relativizer. In some cases, the lexical element is an affix that forms nominalizations or participials. Use of this linking element then triggers additional syntactic processes, such as extraposition, deletion, constituent reordering, agreement, and so on.

These processes rely on four key aspects of the current model. First, the lexical items involved in these constructions must be stored as phonological forms within the posterior systems. These items include both full lexical items and affixes. Second, these lexical forms must be integrated by IFG into positional patterns that control the positioning of the items in the clauses, as well as movement and deletion. Third, these IFG structures must rely on frontal STM mechanisms that store elements. Fourth, these STM items must be pieced together for final mental model construction.

MacWhinney (2008b) argues that mental model construction is driven by a process of perspective taking. Let us consider an example from relative clause processing. Earlier we noted how clause combining through STM joins “the dog chased the bird” and “the bird flew away” to form “the dog chased the bird that flew away.” In this case, the shift moves smoothly from *bird* as the object of chased to *bird* as the subject of *flew away*. However, if the sentence is “the dog chased the bird that the girl loved” then the perspective shift is far more difficult, since a brand new perspective is introduced and the perspectives of both *the dog* and *the bird* must be dropped. In this case, there is a greater burden on STM for fragment storage and hence a higher overall processing load, as reflected by slower latencies and lesser recall accuracy for object relatives. These shifts of perspective are triggered by syntactic patterns linked to lexical devices. To learn these, the child must figure out how to operate on signals from the lexicon or IFG to control the correct shifting in frontal cortex. As the developmental literature amply demonstrates, the learning of this control takes many years (Franks & Connell, 1996). Later in this paper, we will explore some of these processes in further detail, since this is one of the primary loci of the consolidation of linguistic complexity.

Ambiguous sentences illustrate another face of perspective shifting. Consider sentences such as “John saw the Grand Canyon flying to New York.” Here, the default syntactic mechanism would favor the local attachment of *flying to New York* to *Grand Canyon*. The competing attachment is to *John*. Of course, the latter perspective is far more plausible. Or consider the processing of “Visiting relatives can be a nuisance.” Here, we can either take the perspective of the relatives who become a nuisance to their hosts or the perspective of an unmentioned generalized actor who visits the relatives. In this case, both readings seem plausible. Reflexivity provides another useful example of perspectival processes. Consider these sentences

- (22) Jessie stole a photo of herself/her* out of the archives.
 (23) Jessie stole me a photo of herself/her out of the archives.
 (24) Jessie stole a silly photo of herself/her out of the archives.

In (a) the reflexive is required, because the perspective of Jessie remains active up to the appearance of the anaphor. In (b) on the other hand, the intervening presence of “me” causes a shift of perspective away from Jessie. As a result, when interpretation reaches the anaphor, either the reflexive or the simple pronoun is acceptable. Perspective shift is sensitive not just to other intervening animate perspectives, but also to implicit perspectives triggered by adjectives such as *silly* in (c). This type of phenomena is basic to all levels of mental model construction.

8. Perspective and Gesture

The frontal-parietal system for perspective shifting is not a recent evolutionary adaptation. Chimpanzees (Tomasello, Call, & Gluckman, 1997), dogs, and other mammals make extensive use of symbolic behaviors in social contexts. However, lacking a lexicon and positional patterns, other animals cannot organize these behaviors into recursive structures.⁵ However, Donald (1991) and others have argued that the production of symbolic communication can rely on gestural and vocal devices that may well have been readily accessible to *homo erectus*. Because gestures can be formed in ways that map iconically to their referents, it is relatively easy to build up communal recognition of a gestural system. As Tucker et al. (this volume) argue, such a system would rely primarily on gestures and affordances specific to the action-oriented processes in the dorsal stream. It appears that speakers of sign languages are able to use posterior lexical areas to structure a lexicon of signs, just as they use IFG in the left hemisphere to control the ordering of signs. It is possible that protosign could also have relied on these same neuronal structures for lexical organization. However, looking back two million years, it is likely that the depth of support for lexical storage and positional patterning of gesture was still very incomplete. As a result, it is likely that protosign was incompletely lexical and heavily reliant on dorsal processes for direct perspective taking and shifting.

5. Gentner, Fenn, Margolish, & Nusbaum (2006) claim that starlings demonstrate recursive processing for strings such as AAABBB. However, Corballis (2007) points out that these strings can be detected through a subitization-based counting mechanism that has been demonstrated for birds, along with the ability to match two counts.

Although sign may not have triggered full linguistic structure, it provided a fertile social bed that supported the development of further articulatory, lexical, and sequence systems. As Darwin (1872) notes, vocal and gestural communication coexisted as parallel streams from the beginning of human evolution. Gesture and prosody were able to keep humans engaged in protoconversations, during which the further elaboration of vocal patterns could refine and complement communication in the gestural-prosodic mode. Of course, humans are not the only primates that engage in conversation. However, as argued in MacWhinney (2008a), the shift in *homo habilis* to a full upright posture led to two important consequences. One was the freeing of the hands for additional conversational interaction and the other was the encouragement of full face-to-face interactions linked to full display of the hands and torso. This increasing support for gestural communication brought along with it a supportive social context for the further development of accompanying vocalizations. However, both of these modalities continue to provide important input to conversation in modern humans. Thus, we can best view the transition from a primarily gestural communication to a primarily vocal communication system as gradual, but unbroken, process with no sudden break based on the sudden introduction of an ability to process recursion.

9. Integration of the subsystems

We have now finished our survey of an account of neurolinguistic processing grounded on self-organizing feature maps, sequence processing mechanisms, limbic consolidation, and topological preservation of feature map resonance across six linguistic modules. The core mechanisms of neural connectivity and firing are fundamental to all animals from molluscs to mammals. Mechanisms for sequence detection and control can also be found in both invertebrates such as insects and vertebrates such as amphibians. Systems of topographic organization can be found even in animals with no cortex. Systems controlling memory consolidation and value-based projection are found in bees (Menzel & Giurfa, 2001). What is new in the engines supporting language are not the pieces, but the ways in which the pieces are being combined.

We are now in a position to contrast the sketch of the neurological bases of language processing provided here with the proposal advanced by HC&F. Their claim is that the Merge operation produces recursion and, hence, complex syntax. The problem with this proposal is that it identifies only one of the six subsystems required to support complexity and recursion itself. As we have seen, at least six subsystems must work together to produce complex syntax. Recursion and iteration are not produced within a single subsystem, but through the interaction of subsystems. Of course HC&F could argue that the other five systems are parts of FLB and therefore present in other species. But there is no evidence in other species for a linkage between audition and

articulation of the type seen in human babbling. Nor is the evidence for the presence of a lexicon capable of expressing tens of thousands of concepts, along with detailed shades in their meanings. Moreover, there is no evidence in other species of anything similar to the circuit between IFG and STM areas in DLPFC. Of the various systems supporting recursion in human language, the only one that seems to be least changed by language evolution is the system for mental model construction. However, even here, we have evidence that humans far surpass their primate cousins in tests of perspective-taking and theory of mind (Tomasello, Call, & Gluckman, 1997).

The issue raised by HC&F is a good one, since even an incomplete answer to this puzzle forces us to focus on the core issue of how the brain generates linguistic complexity. Beyond recursion and iteration, complexity arises from processes of anaphora, deixis, derivation, affixation, compounding, prosodic packaging, phonological assimilation, and presupposition. The other chapters in this book discuss in great detail the contributions of these processes to language change. Like recursion and iteration, each of these processes relies on all six neurolinguistic subsystems for its full instantiation. Thus, it seems best to conclude that complexity arises from six permeable subsystems each designed to interact smoothly with the others. Each of the pieces of this integrated engine made important separate, contributions to human evolution. However, the full power of the system could only be realized once all six components were fully integrated.

References

- Adolphs, R. & Spezio, M. 2006. Role of the amygdala in processing visual social stimuli. *Progress in Brain Research* 156: 363–378.
- Aslin, R.N., Saffran, J.R. & Newport, E.L. 1999. Statistical learning in linguistic and nonlinguistic domains. In *The Emergence of Language*, B. MacWhinney (Ed.), 359–380. Mahwah NJ: Lawrence Erlbaum Associates.
- Bar-Hillel, Y. 1953. On recursive definitions in empirical science. In *11th International Congress of Philosophy*, Vol. 5, 160–165.
- Bickerton, D. 2002. Afferent isn't efferent, and language isn't logic, either. *Behavioral and Brain Sciences* 26: 286.
- Braine, M.D.S. 1963. The ontogeny of English structure: The first phase. *Language* 39: 1–13.
- Braine, M.D.S. 1971. On two types of models of the internalization of grammars. In *The Ontogenesis of Grammar: A Theoretical Symposium*, D.I. Slobin (Ed.), New York NY: Academic Press.
- Braine, M.D.S. 1976. Children's first word combinations. *Monographs of the Society for Research in Child Development*, 41, Whole No. 1.
- Budiu, R. & Anderson, J. 2004. Interpretation-based processing: A unified theory of semantic sentence comprehension. *Cognitive Science* 28: 1–44.
- Bullinaria, J.A. 2007. Understanding the emergence of modularity in neural systems. *Cognitive Science* 31: 673–695.

- Chomsky, N. 1957. *Syntactic Structures*. The Hague: Mouton.
- Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge MA: The MIT Press.
- Corballis, M. 2007. Recursion, language, and starlings. *Cognitive Science* 31:697–704.
- Darwin, C. 1872. *The Expression of the Emotions in Man and Animals*. London: John Murray.
- Deacon, T. 1989. The neural circuitry underlying primate calls and human language. *Human Evolution* 4: 367–401.
- Dell'Orletta, F., Lenci, A., Montemagni, S. & Pirrelli, V. 2005. Climbing the path to grammar: A maximum entropy model of subject/object learning. *Association for Computational Linguistics: PsyComp Models*.
- Dittenbach, M., Rauber, A., & Merkl, D. 2002. Uncovering the hierarchical structure in data using the growing hierarchical self-organizing map. *Neurocomputing* 48: 199–216.
- Donald, M. 1991. *Origins of the Modern Mind*. Cambridge MA: Harvard University Press.
- Friederici, A.D., Bahlmann, J., Heim, S., Schubotz, R. I. & Anwander, A. 2006. The brain differentiates human and non-human grammars: Functional localization and structural connectivity. *PNAS* 103(7): 2458–2463.
- Fodor, J. 1983. *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge MA: The MIT Press.
- Franks, S.L. & Connell, P.J. 1996. Knowledge of binding in normal and SLI children. *Journal of Child Language* 23: 431–464.
- Gentner, T.Q., Fenn, K.M., Margoliash, D. & Nusbaum, H.C. 2006. Recursive syntactic pattern learning by songbirds. *Nature* 440: 1204–1207.
- Gernsbacher, M.A. 1990. *Language Comprehension as Structure Building*. Hillsdale NJ: Lawrence Erlbaum.
- Gibson, E. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68: 1–76.
- Givón, T. 1995. *Functionalism and Grammar*. Amsterdam: John Benjamins.
- Grodzinsky, Y. & Friederici, A. 2006. Neuroimaging of syntax and syntactic processing. *Current Opinion in Neurobiology* 16: 240–246.
- Guenther, F. 2006. Cortical interactions underlying the production of speech sounds. *Journal of Communication Disorders* 39: 350–365.
- Harris, Z.S. 1951. *Structural Linguistics*. Chicago IL: The University of Chicago Press.
- Hauser, M., Chomsky, N. & Fitch, T. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298: 1569–1579.
- Hauser, M., Newport, E. & Aslin, R. 2001. Segmentation of the speech stream in a non-human primate: Statistical learning in cotton-top tamarins. *Cognition* 78: B53–B64.
- Hausser, R. 1992. Complexity in left-associative grammar. *Theoretical Computer Science* 106.
- Hurford, J. 2002. The neural basis of predicate-argument structure. *Behavioral and Brain Sciences* 26: 261–316.
- Jeannerod, M. 1997. *The Cognitive Neuroscience of Action*. Cambridge MA: Blackwell.
- Jusczyk, P.W. 1997. *The Discovery of Spoken Language*. Cambridge MA: The MIT Press.
- Kintsch, W. 1998. *Comprehension: A paradigm for Cognition*. Cambridge: CUP.
- Knoblich, G. 2008. Bodily and motor contributions to action perception. In *Embodied Cognition*, R. Klatzky, B. MacWhinney & M. Behrmann (Eds), Mahwah NJ: Lawrence Erlbaum.
- Kohonen, T. 1990. The self-organizing map. In *Proceedings of the IEEE*, Vol. 78, 1464–1480.
- Kuhl, P.K. & Miller, J.D. 1975. Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. *Science* 190: 69–72.
- Kuhl, P.K. & Miller, J.D. 1978. Speech perception by the chinchilla: Identification functions for synthetic VOT stimuli. *Journal of the Acoustical Society of America* 63: 905–917.
- Lacerda, F. 1998. An exemplar-based account of emergent phonetic categories. *Journal of the Acoustical Society of America* 103: 2980–2981.
- Lakoff, G. 1987. *Women, Fire, and Dangerous Things*. Chicago IL: Chicago University Press.
- Levelt, W.J.M. 1989. *Speaking: From Intention to Articulation*. Cambridge MA: The MIT Press.
- Li, P., Zhao, X. & MacWhinney, B. 2007. Dynamic self-organization and early lexical development in children. *Cognitive Science* 31: 581–612.
- Lindblom, B., Diehl, R., Park, S.-H., & Salvi, G. In press. Sound systems are shaped by their users: The recombination of phonetic substance. In *Where do features come from? The nature and sources of pho-nological primitives*.
- Lyons, J. 1977. *Semantics*. Cambridge: CUP.
- MacDonald, M.C., Pearlmuter, N.J. & Seidenberg, M.S. 1994. Lexical nature of syntactic ambiguity resolution. *Psychological Review* 101(4): 676–703.
- MacNeilage, P., Davis, B.L., Kinney, A. & Matyear, C. 2000. The motor core of speech: A comparison of serial organization patterns in infants and languages. *Child Development* 71: 153–163.
- MacWhinney, B. 1975. Pragmatic patterns in child syntax. *Stanford Papers And Reports on Child Language Development* 10: 153–165.
- MacWhinney, B. 1977. Starting points. *Language* 53: 152–168.
- MacWhinney, B. 1978. The acquisition of morphophonology. *Monographs of the Society for Research in Child Development* 43, Whole no. 1, 1–123.
- MacWhinney, B. 1982. Basic syntactic processes. In *Language Acquisition*, Vol. 1. *Syntax and Semantics* S. Kuczaj (Ed.), 73–136. Hillsdale NJ: Lawrence Erlbaum.
- MacWhinney, B. 1987a. The competition model. In *Mechanisms of Language Acquisition*, B. MacWhinney (Ed.), 249–308. Hillsdale NJ: Lawrence Erlbaum.
- MacWhinney, B. 1987b. Toward a psycholinguistically plausible parser. In *Proceedings of the Eastern States Conference on Linguistics*, S. Thomason (Ed.), Columbus OH: Ohio State University.
- MacWhinney, B. 1989. Competition and lexical categorization. In *Linguistic Categorization*, R. Corrigan, F. Eckman & M. Noonan (Eds), 195–242. Amsterdam: John Benjamins.
- MacWhinney, B. 2000. *The CHILDES Project: Tools for Analyzing Talk*, 3rd Edn. Mahwah NJ: Lawrence Erlbaum Associates.
- MacWhinney, B. 2005. Item-based constructions and the logical problem. *ACL 2005*, 46–54.
- MacWhinney, B. 2008a. Cognitive precursors to language. In *The Evolution of Communicative Flexibility*, K. Oller & U. Griebel (Eds), 193–214. Cambridge MA: The MIT Press.
- MacWhinney, B. 2008b. How mental models encode embodied linguistic perspectives. In *Embodiment, Ego-Space, and Action*, R. Klatzky, B. MacWhinney & M. Behrmann (Eds), 369–410. Mahwah NJ: Lawrence Erlbaum.
- MacWhinney, B., Bates, E. & Kliegl, R. 1984. Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior* 23: 127–150.
- MacWhinney, B. & Leinbach, J. 1991. Implementations are not conceptualizations: Revising the verb learning model. *Cognition* 29: 121–157.
- MacWhinney, B. & Pléh, C. 1988. The processing of restrictive relative clauses in Hungarian. *Cognition* 29: 95–141.
- Marcus, G. 2000. Pabiku and Ga Ti Ga: Two mechanisms infants use to learn about the world. *Current Directions in Psychological Science* 9:145–147.
- Meltzoff, A.N. & Decety, J. 2003. What imitation tells us about social cognition: A rapprochement between developmental psychology and cognitive neuroscience. *Philosophical Transactions of the Royal Society of London B*, 358, 491–500.

- Menzel, R. & Giurfa, M. 2001. Cognitive architecture of a mini-brain: The honeybee. *Trends in Cognitive Sciences* 5: 62–71.
- O'Grady, W. 2005. *Syntactic Carpentry*. Mahwah NJ: Lawrence Erlbaum Associates.
- Oller, D.K. 2000. *The Emergence of the Speech Capacity*. Mahwah NJ: Lawrence Erlbaum Associates.
- Pelphrey, K.A., Mitchell, T.V., McKeown, M.J., Goldstein, J., Allison, T. & McCarthy, G. 2003. Brain activity evoked by the perception of human walking: Controlling for meaningful coherent motion. *Journal of Neuroscience* 23: 6819–6825.
- Pelphrey, K.A., Morris, J.P. & McCarthy, G. 2005. Neural basis of eye gaze processing deficits in autism. *Brain* 128: 1038–1048.
- Piaget, J. 1954. *The Construction of Reality in the Child*. New York NY: Basic Books.
- Pinker, S. 1997. *How the Mind Works*. New York NY: W.W. Norton & Company.
- Rizzolatti, G., Fadiga, L., Gallese, V. & Fogassi, L. 1996. Premotor cortex and the recognition of motor actions. *Cognitive Brain Research* 3: 131–141.
- Sagae, K., Davis, E., Lavie, E., MacWhinney, B. & Wintner, S. 2007. High-accuracy annotation and parsing of CHILDES transcripts. In *Proceedings of the 45th Meeting of the Association for Computational Linguistics*. Prague: ACL.
- Schütz-Bosbach, S. & Prinz, W. 2007. Perceptual resonance: Action-induced modulation of perception. *Trends in Cognitive Sciences* 11: 349–355.
- Tomalin, M. 2007. Reconsidering recursion in syntactic theory. *Lingua* 117: 1784–1800.
- Tomasello, M., Call, J. & Gluckman, A. 1997. Comprehension of novel communicative signs by apes and human children. *Child Development* 68: 1067–1080.
- Tucker, et al. (in this volume). Neural mechanisms of recursive processing in cognitive and linguistic complexity.
- Westermann, G. & Miranda, E.R. 2004. A new model of sensorimotor coupling in the development of speech. *Brain and Language* 89: 393–400.
- Wilson, M., & Knoblich, G. 2006. The case for motor involvement in perceiving conspecifics. *Psychological Bulletin* 131: 460–473.